From Single-User to Multi-User

• In Lecture 3 we studied various techniques for multiple access and interference management in cellular systems.

• In Lecture 4 we learned about information theory and investigate the capacity of point-to-point channels.

• In this lecture we extend the information theoretic framework to multi-user channels.

• Present new techniques that emerge from the information theoretic study:
  - Success interference cancellation (SIC)
  - Superposition coding
  - Multi-user diversity
  - Opportunistic communication paradigm
Plot

• Two scenarios:
  - Uplink channel (many-to-one)
  - Downlink channel (one-to-many)

• Start with AWGN (no fading)
  - Uplink channel: successive interference cancellation (SIC)
  - Downlink channel: superposition coding

• Fast Fading
  - CSIR only
  - Full CSI

• Multi-user Diversity
Outline

• Uplink/Downlink AWGN channel

• Uplink/Downlink fading channel

• Multi-user diversity

• Opportunistic beamforming
Uplink/Downlink
AWGN Channel
Uplink and Downlink Channel

- **Channel gains** are fixed over time and known to Tx & Rx
- **Uplink noise**: $\mathcal{CN}(0, \sigma^2)$
- **Downlink noise at User $k$, $k = 1,2$**: $\mathcal{CN}(0, \sigma_k^2)$
**Capacity Region**

- **Point-to-point channel:**
  - Capacity $C$
  - Achievable $\iff P_e^{(N)} \to 0$ as $N \to \infty$

- **Multi-user channel**
  - Each user has its own data
  - Two data rates $R_1$ & $R_2$
  - Capacity region $C$
  - $(R_1, R_2)$ is achievable $\iff$
    - both error probability $\to 0$
    - Not Achievable if $(R_1, R_2) \notin C$
Capacity Region of the UL Channel

\[ C_{\text{Uplink}} = \bigcup \left\{ (R_1, R_2) \geq 0 : \begin{cases} R_1 &\leq \log (1 + \text{SNR}_1) \\ R_2 &\leq \log (1 + \text{SNR}_2) \\ R_1 + R_2 &\leq \log (1 + \text{SNR}_1 + \text{SNR}_2) \end{cases} \right\} \]

\[ \text{SNR}_k := \frac{|h_k|^2 P_k}{\sigma^2}, \ k = 1, 2 \]
Non-Achievability Outside $\mathcal{C}_{\text{Uplink}}$

- $R_k \leq \log(1+\text{SNR}_k)$: obvious, since $\log(1+\text{SNR}_k)$ is the point-to-point capacity as if there is only one Tx.

- $R_1 + R_2 \leq \log(1+\text{SNR}_1+\text{SNR}_2)$: obvious, since the maximum received SNR from the two independent Tx is $\text{SNR}_1+\text{SNR}_2$, and therefore the total rate cannot exceed the capacity of the point-to-point channel with this SNR.
Successive Interference Cancellation

Achieving point **A**

User $k$ encodes its data using a capacity achieving AWGN channel code at rate $R_k$, $k=1,2$

Rx first decodes User 2’s data, treating User 1’s signal $x_1$ as Gaussian noise

$$\implies R_2 = \log \left(1 + \frac{|h_2|^2 P_2}{|h_1|^2 P_1 + \sigma^2} \right) = \log \left(1 + \frac{\text{SNR}_2}{1 + \text{SNR}_1} \right)$$

can be achieved

Rx then subtracts $x_2$ from $y$ and get a point-to-point channel for User 1

$$\implies R_1 = \log (1 + \text{SNR}_1)$$

can be achieved

**Note:** smaller $R_2$ can also be achieved
Equivalent Point-to-Point Channels

- Equivalent channels
  - For User 2, the equivalent noise is $h_1x_1 + w$, with variance $|h_1|^2 P_1 + \sigma^2$
  
  $h_2 \quad h_1x_1[m] + w[m]$

  $x_2[m] \rightarrow \bigcirc \rightarrow \bigcirc \rightarrow y[m]$

  - For User 1, after removing $x_2$, Rx sees a clean point-to-point channel without interference

  $h_1 \quad w[m]$

  $x_1[m] \rightarrow \bigcirc \rightarrow \bigcirc \rightarrow y[m] - h_2x_2[m]$
Similarly point B can be achieved

To achieve a rate point on AB, say, \( qA + (1-q)B \), the system can take the following two strategies with a prescribed portion of time:

**Strategy achieving A**
Decode User 2 first and then decode User 1; \( q \) of time

**Strategy achieving B**
Decode User 1 first and then decode User 2; \((1-q)\) of time
Comparison with Conventional CDMA

• For each user, treat the other user’s signal as noise
  - No successive interference cancellation (SIC)
  - Hence a single-user receiver, not a multi-user receiver
• It is strictly suboptimal (achieving point C)

\[
\begin{align*}
R_2 &= \log (1 + \text{SNR}_2) \\
R_1 &= \log \left(1 + \frac{\text{SNR}_1}{1 + \text{SNR}_2}\right)
\end{align*}
\]
UL Orthogonal Multiple Access

- Consider time-division access
  - User 1 uses the first $\alpha$ of the time
  - User 2 uses the rest $(1-\alpha)$ of the time

- Power constraint:
  - User 1 can now use power $P_1/\alpha$ during its transmission
  - User 2 can now use power $P_2/(1-\alpha)$ during its transmission

- Achievable rates:
  \[
  \begin{align*}
  R_1 &= \alpha \log \left(1 + \frac{\text{SNR}_1}{\alpha}\right) \\
  R_2 &= (1-\alpha) \log \left(1 + \frac{\text{SNR}_2}{1-\alpha}\right)
  \end{align*}
  \]
  $\alpha \in [0, 1]$
  - When $\alpha = \text{SNR}_1/(\text{SNR}_1+\text{SNR}_2)$, the sum capacity is achieved (i.e., $R_1+R_2 = \log(1+\text{SNR}_1+\text{SNR}_2)$ is achieved)
Orthogonal MA is Sum Rate Optimal

\[ \bigcup_{\alpha \in [0,1]} \left\{ (R_1, R_2) : \begin{cases} R_1 = \alpha \log \left( 1 + \frac{\text{SNR}_1}{\alpha} \right) \\ R_2 = (1 - \alpha) \log \left( 1 + \frac{\text{SNR}_2}{1 - \alpha} \right) \end{cases} \right\} \]

Orthogonal multiple access can only achieve the optimal sum rate at a single point, when
\[ \alpha = \frac{\text{SNR}_1}{(\text{SNR}_1 + \text{SNR}_2)} \]

Fairness is an issue

\[ \mathcal{C}_{\text{Uplink}} \]

D: \[ \begin{cases} R_1 = \frac{\text{SNR}_1}{\text{SNR}_1 + \text{SNR}_2} C_{\text{sum}} \\ R_2 = \frac{\text{SNR}_2}{\text{SNR}_1 + \text{SNR}_2} C_{\text{sum}} \end{cases} \]

\[ C_{\text{sum}} = \log (1 + \text{SNR}_1 + \text{SNR}_2) \]
For a general $K$-user uplink channel

- Capacity region:

$$\mathcal{C}_{\text{Uplink}} = \bigcup \left\{ (R_1, \ldots, R_K) \geq 0 : \sum_{k \in S} R_k \leq \log \left( 1 + \sum_{k \in S} \text{SNR}_k \right), \forall S \subseteq [1 : K] \right\}$$

- Sum capacity:

$$C_{\text{Uplink}}^{\text{sum}} = \log \left( 1 + \sum_{k=1}^{K} \text{SNR}_k \right)$$

For example, 3-user uplink channel capacity region:

$$R_k \leq \log (1 + \text{SNR}_k), \ k = 1, 2, 3$$
$$R_1 + R_2 \leq \log (1 + \text{SNR}_1 + \text{SNR}_2)$$
$$R_2 + R_3 \leq \log (1 + \text{SNR}_2 + \text{SNR}_3)$$
$$R_3 + R_1 \leq \log (1 + \text{SNR}_3 + \text{SNR}_1)$$
$$R_1 + R_2 + R_3 \leq \log (1 + \text{SNR}_1 + \text{SNR}_2 + \text{SNR}_3)$$
Capacity Region of the DL Channel

\[ C_{\text{Downlink}} = \bigcup_{\beta \in [0,1]} \left\{ (R_1, R_2) \geq 0 : \begin{cases} R_1 \leq \log (1 + \beta \text{SNR}_1) \\ R_2 \leq \log \left( 1 + \frac{(1-\beta)\text{SNR}_2}{1+\beta} \right) \end{cases} \right\} \]

Note: proof of non-achievability outside this region is beyond the scope of this course

\[ \text{SNR}_k := \frac{|h_k|^2 P}{\sigma_k^2}, \ k = 1, 2 \]

**WLOG assume** \( \text{SNR}_1 \geq \text{SNR}_2 \)

Maximum sum rate is achieved when \( \beta = 1 \)

\[ \implies C_{\text{Downlink}}^{\text{sum}} = \log (1 + \text{SNR}_1) \]
Superposition Coding

• Tx sends $x = x_1 + x_2$, where for $k=1,2$
  - User $k$’s data is encoded onto $x_k$
  - Power of $x_1 = \beta P$; power of $x_2 = (1-\beta)P$

• User 1 has a better received SNR
  - User 1’s channel is better than User 2
  - User 1 can decode whatever User 2 can decode

• Single-user decoding at User 2:
  - Decode $x_2$ by treating $x_1$ as noise
    $\Rightarrow$ can achieve $R_2 = \log \left(1 + \frac{(1-\beta)SNR_2}{1+\beta SNR_2}\right)$

• SIC Decoding at User 1:
  - First decode $x_2$ by treating $x_1$ as noise, and remove it from $y_1$
  - Then decode $x_1$ $\Rightarrow$ can achieve $R_1 = \log \left(1 + \beta SNR_1\right)$
Comparison with Conventional CDMA

- Conventional CDMA: the same as before except that User 1 does not do SIC
- Strictly suboptimal
- Exercise: how to choose $\beta$ such that all DL users have the same received SINR?
Consider time-division access

- User 1 uses the first $\alpha$ of the time with power $P$
- User 2 uses the rest $(1-\alpha)$ of the time with power $P$

Achievable rates:

\[
\begin{aligned}
R_1 &= \alpha \log (1 + \text{SNR}_1) \\
R_2 &= (1 - \alpha) \log (1 + \text{SNR}_2)
\end{aligned}
\]

\[\alpha \in [0, 1]\]

- Strictly suboptimal (except the two corner points when one of the users is shut down)
**K-user Downlink Channel Capacity**

- **WLOG assume** $\text{SNR}_1 \geq \text{SNR}_2 \geq \ldots \geq \text{SNR}_K$

- **Capacity region:**

$$\mathcal{C}_\text{Downlink} = \bigcup_{\beta_1, \ldots, \beta_K \geq 0 \atop \beta_1 + \ldots + \beta_K = 1} \left\{(R_1, \ldots, R_K) \geq 0 : \begin{array}{c} R_k \leq \log \left(1 + \frac{\beta_k \text{SNR}_k}{1 + \sum_{j=1}^{k-1} \beta_j \text{SNR}_k} \right), \\ \forall k \in [1 : K] \end{array} \right\}$$

- $\beta_k$ denotes the portion of power allocated to User $k$’s codeword

- **Sum capacity:** achieved by sending only to the best user

$$C_{\text{Downlink}}^{\text{sum}} = \log (1 + \text{SNR}_1)$$
Uplink/Downlink Fading Channel
Setting

- **Fast fading**: $\forall k, \{h_k[m]\}$ is stationary and ergodic
- **Symmetry**: $\forall k, \{h_k[m]\}$ is identically distributed
- **We shall focus on** ergodic sum capacity

### Uplink

Rx: decodes both users' data

\[
y[m] = h_1[m]x_1[m] + h_2[m]x_2[m] + w[m]
\]

- $x_1$ User 1
- $x_2$ User 2
- $y$

### Downlink

Tx: encodes both users' data

\[
y_1[m] = h_1[m]x[m] + w_1[m] \\
y_2[m] = h_2[m]x[m] + w_2[m]
\]

- $x$ $x_1$ User 1
- $x_2$ User 2
- $y_1$ $y_2$
Uplink Channel Capacity: CSIR Only

• Without CSIT, the ergodic sum capacity is

\[ C_{UL, CSIR}^{\text{sum}} = \mathbb{E} \left[ \log \left( 1 + \sum_{k=1}^{K} |h_k|^2 \text{SNR}_k \right) \right] \]

where \( \text{SNR}_k := \frac{P_k}{\sigma^2}, \ k = 1,2,\ldots,K \)

• Comparison with AWGN capacity:

\[ C_{UL, CSIR}^{\text{sum}} = \mathbb{E} \left[ \log \left( 1 + \sum_{k=1}^{K} |h_k|^2 \text{SNR}_k \right) \right] \leq \log \left( 1 + \sum_{k=1}^{K} \mathbb{E} \left[ |h_k|^2 \right] \text{SNR}_k \right) = \log \left( 1 + \sum_{k=1}^{K} \text{SNR}_k \right) = C_{UL, AWGN}^{\text{sum}} \]

due to Jensen’s inequality

• Similar to the point-to-point case: if there is no CSIT, fading makes things worse
Downlink Channel Capacity: CSIR Only

- Recall channel symmetry:
  - \( \forall k, \{h_k[m]\} \) is identically distributed

- Due to the symmetry assumption:
  - There is a natural ordering of the users based on the noise level \( \sigma_k^2 \)

- WLOG assume \( \text{SNR}_1 \geq \ldots \geq \text{SNR}_K \) where \( \text{SNR}_k := \frac{P}{\sigma_k^2} \)

- Sum capacity is achieved by serving the best user only

\[
C_{\text{DL, CSIR}}^{\text{sum}} = \mathbb{E} \left[ \log \left( 1 + |h_1|^2 \text{SNR}_1 \right) \right]
\]

- By Jensen’s inequality this is strictly worse than the AWGN downlink sum capacity
Impact of Multiple Users

- Under fast fading without CSIT:
  - The ergodic sum capacity of multiuser UL/DL channels is smaller than that without fading
  - Similar to the point-to-point case

- As $K$ (# of users) increases, this capacity loss behaves differently in the uplink and the downlink
  - In uplink, the loss vanishes as $K \to \infty$
  - In downlink, the loss remains as $K \to \infty$

- Explored in Homework 3
Downlink Channel Capacity: Full CSI

- User symmetry assumption: $\sigma_k = \sigma$, $\forall k = 1, \ldots, K$
- Downlink channel sum capacity is achieved by sending only to the instantaneously best user.

**Optimization problem:**

$$
\max_{P(h) \geq 0} \mathbb{E} \left[ \log \left( 1 + \frac{\max_{k \in [1:K]} |h_k|^2 P(h)}{\sigma^2} \right) \right] \\
\text{s.t.} \quad \mathbb{E} [P(h)] = P
$$

**Solution:**

$$
P^*(h) = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+,
$$

$\nu$ satisfies

$$
\mathbb{E} \left[ \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+ \right] = P
$$
Uplink Channel Capacity: Full CSI

- **Full CSI assumption:**
  - At each time $m$, all users know the instantaneous realization of all channel gains $\{h_k[m] \mid k = 1,2,\ldots,K\}$
  - In other words, a user can know not only its own channel but also others’ channels

- **User symmetry assumption:** $P_k = P$, $\sigma_k = \sigma$, $\forall k = 1,\ldots,K$

- **Program in finding uplink sum capacity under full CSI:**
  - Consider an $L$-parallel uplink AWGN channel
  - Relax individual power constraints to a total power constraint
  - Solve the new problem under finite $L$, and take $L \to \infty$
  - By channel and user symmetry, argue that the found solution is also feasible under individual power constraints
Power Allocation Problem

• **Original problem:**

\[
\begin{align*}
\max_{P_k(h) \geq 0, \quad k \in [1:K]} & \quad \mathbb{E} \left[ \log \left( 1 + \frac{\sum_{k=1}^{K} |h_k|^2 P_k(h)}{\sigma^2} \right) \right] \\
\text{s.t.} & \quad \mathbb{E} [P_k(h)] = P, \quad k = 1, 2, \ldots, K
\end{align*}
\]

Individual power constraint

• **Relaxed problem:**

\[
\begin{align*}
\max_{P_k(h) \geq 0, \quad k \in [1:K]} & \quad \mathbb{E} \left[ \log \left( 1 + \frac{\sum_{k=1}^{K} |h_k|^2 P_k(h)}{\sigma^2} \right) \right] \\
\text{s.t.} & \quad \sum_{k=1}^{K} \mathbb{E} [P_k(h)] = KP
\end{align*}
\]

Total power constraint

• **We will solve the relaxed problem first, and verify that the solution found there is also feasible for the original one**
Parallel Uplink Channel

- $L$ parallel $K$-user uplink channel:
  - Channel gains for the $l$-th sub-channel: $\{h_{k,l} \mid k = 1,2,\ldots,K\}$
  - Power allocated to the $l$-th sub-channel: $\{P_{k,l} \mid k = 1,2,\ldots,K\}$

- Optimization problem (total power constraint):

\[
\max_{P_{k,l} \geq 0} \quad \frac{1}{L} \sum_{l=1}^{L} \log \left(1 + \frac{\sum_{k=1}^{K} |h_{k,l}|^2 P_{k,l}}{\sigma^2}\right)
\]

s.t.
\[
\sum_{k=1}^{K} \frac{1}{L} \sum_{l=1}^{L} P_{k,l} = KP
\]

total power constraint
Optimal Allocation in Parallel UL (1)

• Rewrite the total power constraint as

\[
\sum_{k=1}^{K} \frac{1}{L} \sum_{l=1}^{L} P_{k,l} = KP \iff \sum_{l=1}^{L} \left( \sum_{k=1}^{K} P_{k,l} \right) = LKP
\]

• For a fixed partition \( \{P_{l} \mid l \in [1:L]\} \) of the total power \( LKP \), the sum rate of the \( l \)-th sub-channel is maximized if all of \( P_{l} \) is allocated to the best user: (\( P_{l} := \sum_{k=1}^{K} P_{k,l} \))

\[
\max_{P_{k,l} \geq 0 \atop k \in [1:K]} \log \left( 1 + \frac{\sum_{k=1}^{K} |h_{k,l}|^2 P_{k,l}}{\sigma^2} \right) = \log \left( 1 + \frac{(\max_{k \in [1:K]} |h_{k,l}|^2) P_{l}}{\sigma^2} \right)
\]
Optimal Allocation in Parallel UL (2)

- How to determine the best partition \( \{ P_l \mid l \in [1:L] \} \) of the total power \( LKP \)?

- Problem becomes:

\[
\max_{P_l \geq 0 \atop l \in [1:L]} \frac{1}{L} \sum_{l=1}^{L} \log \left( 1 + \frac{\left( \max_{k \in [1:K]} |h_{k,l}|^2 \right) P_l}{\sigma^2} \right)
\]

s.t. \( \sum_{l=1}^{L} P_l = LKP \)

- Water-filling solution:

\[
P^*_l = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_{k,l}|^2} \right)^+ , \quad \nu \text{ satisfies } \sum_{l=1}^{L} P^*_l = LKP
\]
Optimal Allocation in Parallel UL (3)

- Optimal power allocation in $L$ parallel $K$-user uplink channel under the total power constraint:

$$P^*_{k,l} = \begin{cases} P^*_l, & \text{if } k = \arg \max_{j \in [1:K]} |h_{j,l}|^2 \\ 0, & \text{otherwise} \end{cases}$$

where $P^*_l = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_{k,l}|^2} \right)^+$, $\nu$ satisfies $\sum_{l=1}^{L} P^*_l = LKP$

- Take $L \to \infty$, we obtain the solution of the power allocation problem of the uplink fading channel under total power constraint

$$P^*_k(h) = \begin{cases} P^*(h), & \text{if } k = \arg \max_{j \in [1:K]} |h_j|^2 \\ 0, & \text{otherwise} \end{cases}$$

where $P^*(h) = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+$, $\nu$ satisfies $\mathbb{E}[P^*(h)] = KP$
Solution to the Original Problem

- Recall the original vs. the relaxed problem

\[
\max_{P_k(h) \geq 0, k \in [1:K]} \mathbb{E} \left[ \log \left( 1 + \sum_{k=1}^{K} \frac{|h_k|^2 P_k(h)}{\sigma^2} \right) \right]
\]

\[\text{s.t. } \mathbb{E} [P_k(h)] = P, \ k = 1, \ldots, K \quad \text{Original Problem}\]

\[\text{s.t. } \sum_{k=1}^{K} \mathbb{E} [P_k(h)] = KP \quad \text{Relaxed Problem}\]

- Note: solution to the relaxed problem is

\[P^*_k(h) = \begin{cases} P^*(h), & \text{if } k = \arg \max_{j \in [1:K]} |h_j|^2 \\ 0, & \text{otherwise} \end{cases}\]

where \(P^*(h) = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+\), \(\nu\) satisfies \(\mathbb{E} [P^*(h)] = KP\)

- Due to channel symmetry, \(\mathbb{E} [P^*_k(h)]\) are equal \(\forall k \Rightarrow \mathbb{E} [P^*_k(h)] = P \ \forall k \Rightarrow \text{feasible in the original problem!}\)
UL Capacity with Full CSI: Summary

\[
\begin{align*}
    \max_{\substack{P_k(h) \geq 0 \quad k \in [1:K]}} & \mathbb{E} \left[ \log \left( 1 + \frac{\sum_{k=1}^{K} |h_k|^2 P_k(h)}{\sigma^2} \right) \right] \\
    \text{s.t.} & \quad \mathbb{E} [P_k(h)] = P, \quad k = 1, 2, \ldots, K
\end{align*}
\]

- **Solution:**

\[
P_k^*(h) = \begin{cases} 
P^*(h), & \text{if } k = \arg \max_{j \in [1:K]} |h_j|^2 \\ 0, & \text{otherwise} \end{cases}
\]

where

\[
P^*(h) = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+,
\]

\[
\nu \text{ satisfies } \mathbb{E} \left[ \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+ \right] = KP
\]
Remarks

• Sum capacities and optimal power allocation solutions of the DL and the UL channels are essentially the same
  - UL total power constraint: $KP$
  - DL total power constraint: $P$

• Full CSIT requirement in UL:
  - We begin with the assumption that all users know all the channels
  - However, to attain the optimal power allocation, each user only needs to know its own channel and whether it is the best channel
  - Amount of feedback to each user is not increasing with $K$!
Multi-User Diversity
A Key Feature of Wireless Channel

- Time variation!
- Multi-path fading
- Large-scale channel variations (path loss, shadowing)
- Time-varying interference
Multiuser capacity and opportunistic communication

The channel strength scheme that improved in high mobility environments. In Section 6.7.3, we will discuss a

impossible without knowing when the channel is actually good. even though the actual channel fluctuates, opportunistic communication is

has trouble tracking and predicting the channel variations, so that the predicted

increases with the number of users! It turns out that at this speed the receiver

environment? In fact quite the opposite is true. The total throughput hardly

inherent multiuser diversity is more limited in the fixed environment.

are what determines the performance of the scheduling algorithm. Thus, the

fluctuations are likely to be higher in the mobile environment, and the peaks

is more dramatic in the low mobility case. While the channel varies in both

of users in both the fixed and low mobility environments, but the increase

for fairness of comparison. The total throughput increases with the number

Doppler spread of 2 Hz. The Doppler spectrum of this component follows Clarke's model with a

or time-varying component that is assumed to be Rayleigh distributed.

in these examples) the peaks of the channel

refers to the energy in the specular

specular

• Compensates for channel fluctuations
Example: CDMA Systems

• Two main compensating mechanisms:
  - Channel diversity
  - Interference management

• Channel diversity
  - Frequency diversity via Rake receiver
  - Macro-diversity via soft handoff
  - Tx/Rx antenna diversity

• Interference management
  - Intra-cell: power control
  - Inter-cell: interference averaging
What Drives this Approach?

- Main application is **voice**, with tight latency constraints
- Need a **consistent** channel

![Diagram showing channel strength over time and dynamic range comparison between mobile and fixed environments.](image-url)
Opportunistic Communication

• A completely different view!

• Transmit more **when** and **where** the channel is good

• **Exploits fading** to achieve higher long-term throughput, but no guarantee that “the channel is always there”

• **Appropriate for data** with non-real-time latency requirements (file downloads, video streaming)
The capacity of fading channels can be modeled using the law of diminishing marginal return on capacity from increasing the received power. At low SNR, the capacity of the fading channel is:

\[
C = \frac{1}{E_{\text{AMSopen}}} \log_2 \left( 1 + \frac{S/N}{\sigma_r^2} \right) \approx \frac{1}{E_{\text{AMSopen}}} \log_2 \frac{S/N}{\sigma_r^2} \approx C_{\text{awgn}},
\]

where \(C_{\text{awgn}}\) is the capacity of the AWGN channel and is measured in bits per symbol. Hence at low SNR the "Jensen's loss" becomes negligible; this is because the capacity is approximately linear in the received SNR in this regime. At high SNR,

\[
C \approx \frac{1}{E_{\text{AMSopen}}} \log_2 \frac{S/N}{\sigma_r^2} = \log_2 SNR + \frac{1}{E_{\text{AMSopen}}} \log_2 \frac{1}{\sigma_r^2} \approx C_{\text{awgn}} + \frac{1}{E_{\text{AMSopen}}} \log_2 \frac{1}{\sigma_r^2},
\]

The difference is

\[-0.83 \text{ bits/s/Hz} \]

for the Rayleigh fading channel. Equivalently, 2.5 dB more power is needed in the fading case to achieve the same capacity as in the AWGN case. Figure 5.20 compares the capacity of the Rayleigh fading channel with the AWGN capacity as a function of the SNR. The difference is not that large for the entire plotted range of SNR.

5.4.6 Transmitter side information

So far we have assumed that only the receiver can track the channel. But let us now consider the case when the transmitter can track the channel as well. There are several ways in which such channel information can be obtained at the transmitter. In a TDD (time-division duplex) system, the transmitter...
Single-User Channel: Low SNR Regime

The channel gain is $h$. The rate of that code is therefore $\frac{\log_2(1 + \frac{P^*}{h^2/N_0})}{\bar{SC}_h^2}$ bits/s/Hz. No coding across channel states is necessary. This is in contrast to the case without transmitter channel knowledge, where a single fixed-rate code with the coded symbols spanning across different coherence time periods is needed (Figure 5.22). Thus, knowledge of the channel state at the transmitter not only allows dynamic power allocation but simplifies the code design problem as one can now use codes designed for the AWGN channel.

Waterfilling performance

Figure 5.20 compares the waterfilling capacity and the capacity with channel knowledge only at the receiver, under Rayleigh fading. Figure 5.23 focuses on the low SNR regime. In the literature the former is also called the capacity with full channel side information (CSI) and the latter is called the capacity with channel side information at the receiver (CSIR). Several observations can be made:

• At low SNR, the capacity with full CSI is significantly larger than the CSIR capacity.
• At high SNR, the difference between the two goes to zero.
• Over a wide range of SNR, the gain of waterfilling over the CSIR capacity is very small.

The first two observations are in fact generic to a wide class of fading models, and can be explained by the fact that the benefit of dynamic power allocation is a received power gain: by spending more power when the channel is good, the received power gets boosted up. At high SNR, however, the capacity is insensitive to the received power per degree of freedom and varying the amount of transmit power as a function of the channel state yields a minimal gain (Figure 5.24(a)). At low SNR, the capacity is quite sensitive to the received power (linear, in fact) and so the boost in received power from optimal transmit power allocation provides significant gain. Thus, dynamic
Hitting Peaks over Time

Interpretation: at low SNR, one only transmits when the channel is at its peak!
⇒ primarily a power gain at low SNR!
Increase in **spectral efficiency** with number of users $K(\forall K>1)$ at all SNR’s, not just low SNR
Several observations can be made from the plots:

• The sum capacity without transmitter CSI increases with the number of the users, but not significantly. This is due to the multiuser averaging effect explained in the last section. This sum capacity is always bounded by the capacity of the AWGN channel.

• The sum capacity with full CSI increases significantly with the number of users. In fact, with even two users, this sum capacity already exceeds that
Multi-User Gain

- Let us compare the single-user and the multi-user cases:
  - Point-to-point capacity
    
    $$C_{\text{point-to-point}} = \mathbb{E} \left[ \log \left( 1 + |h|^2 \frac{P^*(h)}{\sigma^2} \right) \right]$$
    
    $$P^*(h) = \left( \nu - \frac{\sigma^2}{|h|^2} \right)^+,$$ \(\nu\) satisfies \(\mathbb{E} \left[ \left( \nu - \frac{\sigma^2}{|h|^2} \right)^+ \right] = P\)
  - Multi-user downlink capacity
    
    $$C_{\text{Downlink}} = \mathbb{E} \left[ \log \left( 1 + \max_{k \in [1:K]} |h_k|^2 \frac{P^*(h)}{\sigma^2} \right) \right]$$
    
    $$P^*(h) = \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+,$$ \(\nu\) satisfies \(\mathbb{E} \left[ \left( \nu - \frac{\sigma^2}{\max_{k \in [1:K]} |h_k|^2} \right)^+ \right] = P\)
Multi-User Opportunistic Communication

- Dedicate full power to serve only the best user + the peak value is higher than the mean $\Rightarrow$ multi-user gain!
- Hitting peaks not only over time (at low SNR), but also over users (at all SNR)
Multi-User Diversity

- In a large system with users fading independently:
  - Likely to have a user with very good channel at any time
  - Different users peak at different times
- The more random the channel is, the higher the rate is

The graph illustrates the sum capacity at SNR = 0 dB (bits/s/Hz) versus the number of users for different fading conditions: AWGN, Rayleigh fading, and Rician fading. The sum capacity increases with the number of users, and the Rician fading condition shows the highest sum capacity compared to AWGN and Rayleigh fading, demonstrating the benefit of multi-user diversity.
Multi-User vs. Classical Diversity

• Both due to the existence of independently faded paths
  - Classical diversity: over time, frequency, antennas in a link
  - Multi-user diversity: over multiple users in the network

• Classical diversity is to **compensate** channel fluctuation
• Multi-user diversity aims to **exploit** channel fluctuation

• Classical diversity increases reliability
  - Suitable for application with stringent latency constraints (voice)
• Multi-user diversity increases total throughput (long-term)
  - Suitable for application with long latency constraints (data)
Issues in System Implementation

- **Fairness:**
  - Multi-user diversity offers a system-wide benefit (sum capacity ↑)
  - How to share this benefit among all users in a fair way (in an asymmetric environment)?

- **Slow and limited fluctuations:**
  - Channels are less random ⇒ multi-user diversity gain ↓
  - How to retain the benefit even in a rather static environment?

- **Channel measurement and feedback:**
  - Tracking channel is crucial in getting multi-user diversity
  - Overhead has to be considered
Proportional Fair Scheduler

- At each time slot,
  - Each user will request a data rate from the base station
  - A scheduler decides which user to transmit and at what rate

- To obtain multi-user diversity:
  - Transmit to the best user + stronger user requests higher rates
  - Most likely will select the statistically strongest all the time
  - Highly unfair!

- Solution: schedule the user with the highest ratio $R_k / T_k$, where
  - $R_k :=$ current requested rate of user $k$
  - $T_k :=$ average throughput of user $k$ in the past $t_c$ time slots
The statistically stronger user gets an higher avg. rate
But the statistically weaker user still gets served fairly!
The algorithm serves each user when it is near its peak within the latency time-scale $t_c$. 
Multi-User Diversity in Practice

- Fixed environment has limited fluctuation
- High mobility environment has a lot of fluctuation, but it is difficult to get this gain because the system cannot track the channel!

Fixed environment: 2Hz Rician fading with \( \kappa = 5 \)

Low mobility environment:
- 3 km/hr, Rayleigh fading

High mobility environment:
- 120 km/hr, Rayleigh fading
Inducing Randomness

• Scheduling algorithm exploits the nature-given channel fluctuations by hitting the peaks

• Not enough fluctuations $\Rightarrow$ multi-user diversity gain ↓

• Why not purposely induce fluctuations?
• Multiply the information bearing signal at each Tx antenna by a random complex gain
  - $\alpha(t)$: portion of power allocated to the first antenna
  - $\vartheta(t)$: phase shift
Slow Fading $\rightarrow$ Fast Fading

Before

After

Supportable Rate

User 1

User 2

Time Slots

Supportable Rate

User 1

User 2

Time Slots
Opportunistic Beamforming

• Dumb antennas create a beam in random time-varying directions

• In a large system, there is likely to be a user near the beam at any one time

• By transmitting to that user, close to true beamforming performance is achieved
Performance Improvement

Opportunistic beamforming with dumb antennas increases the performance of the fixed environment significantly.

Fixed environment: 2Hz Rician fading with $\kappa = 5$

Mobile environment: 3 km/hr, Rayleigh fading

- Opportunistic beamforming with dumb antennas increases the performance of the fixed environment significantly.
Dumb, Smart, and Smarter Antennas

• Smart antennas (space-time code in Lecture 3)
  - Improve reliability of point-to-point links
  - Reduce multi-user diversity (less fluctuations)

• Dumb antennas
  - Add fluctuations to point-to-point links
  - Increase multiuser diversity gains

• Smarter antennas
  - With full CSI, antennas can actually form beams pointing to users
  - Coherent beamforming
As the number of users grows, the performance of opportunistic beamforming approaches that of coherent beamforming.
Table 6.1 A comparison between three methods of using transmit antennas.

<table>
<thead>
<tr>
<th></th>
<th>Dumb antennas (Opp. beamform)</th>
<th>Smart antennas (Space-time codes)</th>
<th>Smarter antennas (Transmit beamform)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel knowledge</td>
<td>Overall SNR</td>
<td>Entire CSI at Rx</td>
<td>Entire CSI at Rx, Tx</td>
</tr>
<tr>
<td>Slow fading performance gain</td>
<td>Diversity and power gains</td>
<td>Diversity gain only</td>
<td>Diversity and power gains</td>
</tr>
<tr>
<td>Fast fading performance gain</td>
<td>No impact</td>
<td>Multiuser diversity ↓</td>
<td>Multiuser diversity ↓ power ↑</td>
</tr>
</tbody>
</table>
Thus, for a system designer, the opportunistic beamforming technique provides a compelling case for implementation, particularly in view of the constraints of space and cost of installing multiple antennas on each mobile device. Further, this technique needs neither any extra processing on the part of any user, nor any updates to an existing air-link interface standard. In other words, the mobile receiver can be completely ignorant of the use or non-use of this technique. This means that it does not have to be "designed in" (by appropriate inclusions in the air interface standard and the receiver design) and can be added/removed at any time. This is one of the important benefits of this technique from an overall system design point of view.

In the cellular wireless systems studied in Chapter 4, the cell is sectorized to allow better focusing of the power transmitted from the antennas and also to reduce the interference seen by mobile users from transmissions of the same base-station but intended for users in different sectors. This technique is particularly gainful in scenarios when the base-station is located at a fairly large height and thus there is limited scattering around the base-station. In contrast, in systems with far denser deployment of base-stations (a strategy that can be expected to be a good one for wireless systems aiming to provide mobile, broadband data services), it is unreasonable to stipulate that the base-stations be located high above the ground so that the local scattering (around the base-station) is minimal. In an urban environment, there is substantial local scattering around a base-station and the gains of sectorization are minimal; users in a sector also see interference from the same base-station (due to the local scattering) intended for another sector. The opportunistic beamforming scheme can be thought of as sweeping a random beam and scheduling transmissions to users when they are beamformed. Thus, the gains

<table>
<thead>
<tr>
<th></th>
<th>Conventional multiple access</th>
<th>Opportunistic communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guiding principle</td>
<td>Averaging out fast channel fluctuations</td>
<td>Exploiting channel fluctuations</td>
</tr>
<tr>
<td>Knowledge at Tx</td>
<td>Track slow fluctuations</td>
<td>Track as many fluctuations as possible</td>
</tr>
<tr>
<td></td>
<td>No need to track fast ones</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>Power control the slow fluctuations</td>
<td>Rate control to all fluctuations</td>
</tr>
<tr>
<td>Delay requirement</td>
<td>Can support tight delay</td>
<td>Needs some laxity</td>
</tr>
<tr>
<td>Role of Tx antennas</td>
<td>Point-to-point diversity</td>
<td>Increase fluctuations</td>
</tr>
<tr>
<td>Power gain in downlink</td>
<td>Multiple Rx antennas</td>
<td>Opportunistic beamform via multiple Tx antennas</td>
</tr>
<tr>
<td>Interference management</td>
<td>Averaged</td>
<td>Opportunistically avoided</td>
</tr>
</tbody>
</table>