EE4-65/EE9-SO27 Wireless Communications

Bruno Clerckx

Department of Electrical and Electronic Engineering, Imperial College London

January 2019

Course Ojectives

- Course on wireless communication and communication theory
  - Fundamentals of wireless communications from a 4G and beyond perspective
  - At the cross-road between information theory, coding theory, signal processing and antenna/propagation theory
- Major focus of the course is on 1) fundamentals of wireless, 2) multi-antenna systems and 3) multi-user communications
  - Applications: everywhere in wireless communication networks: 3G, 4G(LTE,LTE-A), 5G, WiMAX(IEEE 802.16e, IEEE 802.16m), WiFi(IEEE 802.11n), satellite,...
- Valuable for those who want to either pursue a PhD in communication or a career in a high-tech telecom company (research centres, R&D branches of telecom manufacturers and operators,...).
- Skills
  - Some mathematical modelling and analysis of wireless communication systems
  - Design (transmitters and receivers) of wireless communication systems
  - Basic system design and performance evaluations

Content

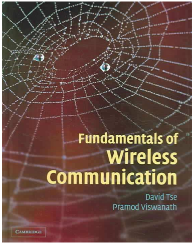
Important Information

- Course material on blackboard
- Prerequisite: For Meng EE3-03 - Communication Systems
- No Exam
- 3 courseworks (using Matlab): $20\%, 30\%, 50\%$. 1) deadline: 3 February 2019, 2) deadline: 24 February 2019, 3) deadline: 20 March 2019. No oral for coursework 1 and 2, but oral for coursework 3 (to be scheduled on 22 March 2019).
- Participate to the lab sessions. See the problem sheets on blackboard.

Two reference books

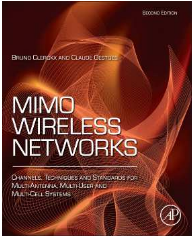
- Introductory

  D. Tse and P. Viswanath, “Fundamentals of Wireless Communication,” Cambridge University Press, May 2005

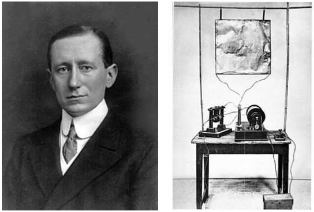


- More advanced and MIMO-focused

  Bruno Clerckx and Claude Oestges, “MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems,” Academic Press (Elsevier), Oxford, UK, Jan 2013.



Wireless Revolution



Marconi 1896

Cellular systems

1G – 1980's

2G (GSM) – 1990's

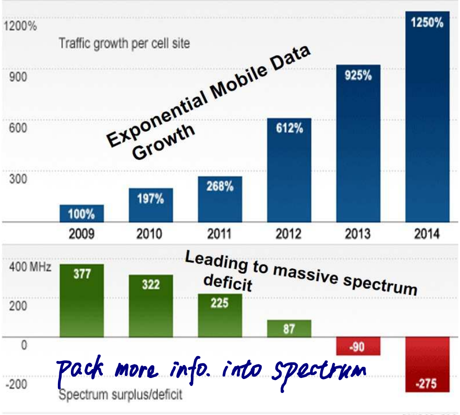
3G (UMTS) – 2000's

4G (LTE) – 2010's

5G – 2020's

Other systems

WiFi, Satellite, Bluetooth, Zigbee, ...



Source: FCC

Challenges

Network/Radio Challenges

- Data rates
- Scarce/limited spectrum
- Reliability and coverage
- Mobility
- Energy efficiency (Tx + Rx)
- Latency
- Explosion of the number of devices

Device/RF/Chip Challenges

- Performance
- Complexity
- Size, Power, Cost
- High frequencies/mmWave
- Multiple Antennas
- Multiradio Integration
- Coexistence

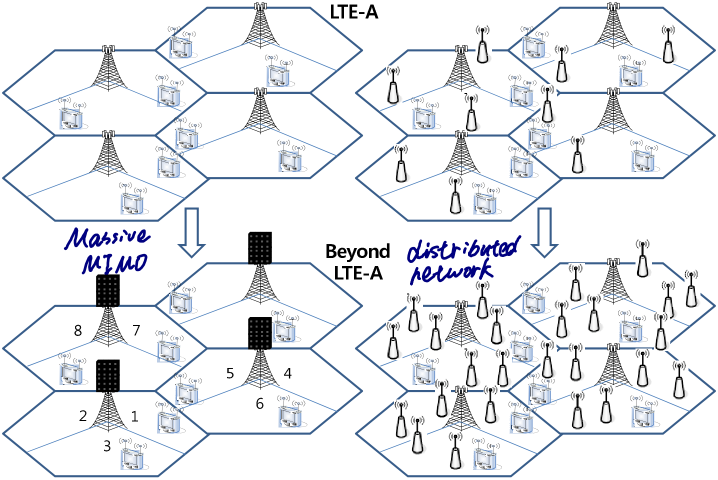
Spectrum Regulation

- Spectrum limited
- Worldwide spectrum controlled by ITU-R
- Regulation/spectrum allocation needed

Standards

- Systems need to interact
- Companies want their systems adopted as standard
- Intellectual property (patents)
- IEEE, 3GPP

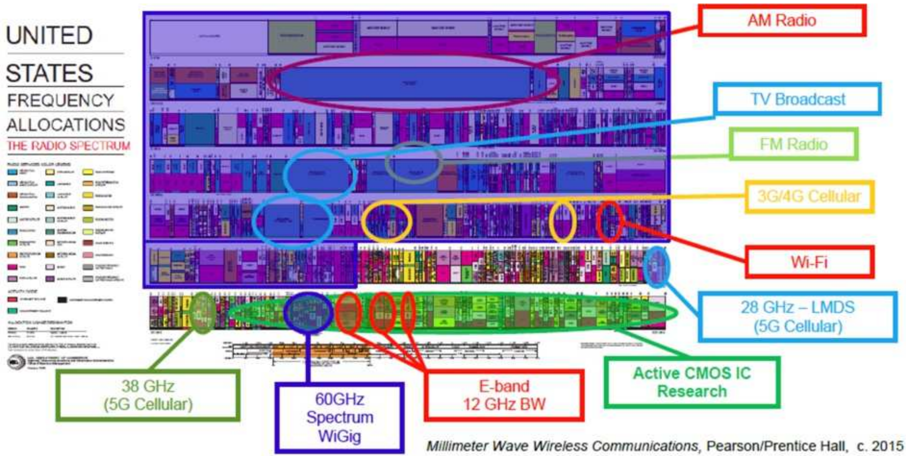
5G: more antennas, more cells



Massive multi-antenna network

Massive multi-cell network

5G: higher frequencies and larger bandwidth



Millimeter Wave Wireless Communications, Pearson/Prentice Hall, c. 2015

Future Wireless Networks

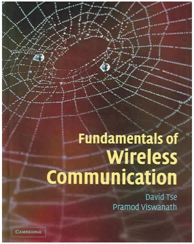
- Next-Gen Cellular/WiFi (5G and beyond)
- Smart Homes/Spaces
- Smart Cities
- Internet of Things
- Autonomous Cars
- Body-Area Networks
- Satellite
- Drones/UAV unmanned aerial vehicle
- Chemical/Molecular Communications
- Applications of Communications in Health, Bio-medicine, and Neuroscience
- Energy harvesting and transfer
- ...

Part 1: Basics of Wireless

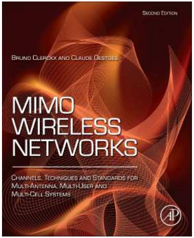
Discrete Time Representation
The Wireless Channel
Fading and Diversity
Capacity of Wireless Channels

Reference Book

Chapter 5 - Sections 5.2.1, 5.3.1, 5.3.2, 5.4.1, 5.4.2, 5.4.3, 5.4.5.

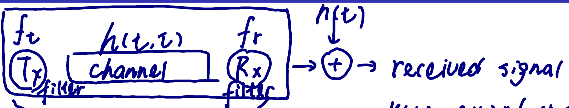


Chapter 1 - Sections 1.2, 1.3, 1.4, 1.5.



$c_3\ c_2\ c_1\ c_0\ c_{-1}\ c_{-2}\ c_{-3}$

$$c(t) = \sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, \delta(t - lT_s)$$



$$h_B(t,\tau) = f_r * h(t,\tau) * f_t$$

$$y(t) = c(t) * h_B(t,\tau) + n(t)$$

Discrete Time Representation

$$
\begin{aligned}
y(t) &= c(t) * h_B(t,\tau) + n(t) \\
&= \int_0^{\tau_{max}} h_B(t,\tau)\, c(t-\tau)\, d\tau + n(t) \\
&= \int_0^{\tau_{max}} h_B(t,\tau) \sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, \delta(t-\tau-lT_s)\, d\tau + n(t) \\
&= \sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l \int_0^{\tau_{max}} h_B(t,\tau)\, \delta(t-\tau-lT_s)\, d\tau + n(t) \\
&= \sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, h_B(t, t-lT_s) + n(t) \\
&= \underbrace{\sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, h_t(\underbrace{t-lT_s}_{\text{delay}}) + n(t)}
\end{aligned}
$$

response at t delay τ

$h_B(t,\tau)$

$\Updownarrow$

$h_B(t-\tau)$ impulse at $t-\tau$

$\boxed{h_t(\tau)}$ delay τ

- at time t, there can be contribution from other symbols.

components: contributions by symbols $c(t)$ at lT_s sample.

Discrete Time Representation

- *channel*: the impulse response of the linear time-varying communication system between one (or more) transmitter(s) and one (or more) receiver(s).
- Assume a SISO transmission where the digital signal is defined in discrete-time by the complex time series $\{c_l\}_{l\in\mathbb{Z}}$ and is transmitted at the symbol rate T_s.

  discrete signal

- The transmitted signal is then represented by

h: impulse response of infinite bandwidth (impossible to measure)
↑
h_B: impulse response of bandwidth B

$$\boxed{c(t) = \sum_{l=-\infty}^{\infty} \sqrt{E_s} c_l \delta(t - lT_s),}$$

impulse response at time t

  where E_s is the transmitted symbol energy, assuming that the average energy constellation is normalized to unity.
- Define a function $\mathsf{h}_B(t,\tau)$ as the time-varying (along variable t) impulse response of the channel (along τ) over the system bandwidth $B = 1/T_s$, i.e. $\mathsf{h}_B(t,\tau)$ is the response at time t to an impulse at time $t-\tau$.

  starting

  impulse time $t-\tau$
  ↓ } duration τ
  response time t

  end

- The received signal $\mathsf{y}(t)$ is given by

$$\boxed{\begin{aligned} \mathsf{y}(t) &= \mathsf{h}_B(t,\tau) \star c(t) + \mathsf{n}(t) \\ &= \int_0^{\tau_{max}} \mathsf{h}_B(t,\tau) c(t-\tau) d\tau + \mathsf{n}(t) \end{aligned}}$$

  where $\star$ denotes the convolution product, $\mathsf{n}(t)$ is the additive noise of the system and τ_{max} is the maximal length of the impulse response.

Discrete Time Representation

- h_B is a scalar quantity, which can be further decomposed into three main terms,

$$\mathsf{h}_B(t,\tau) = f_r \star \mathsf{h}(t,\tau) \star f_t,$$

where

  - f_t is the pulse-shaping filter,
  - $\mathsf{h}(t,\tau)$ is the electromagnetic propagation channel (including the transmit and receive antennas) at time t,
  - f_r is the receive filter.
- Nyquist criterion: the cascade $f = f_r \star f_t$ does not create inter-symbol interference when $\mathsf{y}(t)$ is sampled at rate T_s.
- In practice,
  - difficult to model $\mathsf{h}(t,\tau)$ (infinite bandwidth is required).
  - $\mathsf{h}_B(t,\tau)$ is usually the modeled quantity, but is written as $\mathsf{h}(t,\tau)$ (abuse of notation).
  - Same notational approximation: the channel impulse response writes as $\mathsf{h}(t,\tau)$ or $\mathsf{h}_t[\tau]$.
- The input-output relationship reads thereby as

$$\mathsf{y}(t) = \mathsf{h}(t,\tau) \star c(t) + \mathsf{n}(t) = \sum_{l=-\infty}^{\infty} \sqrt{E_s} c_l \mathsf{h}_t[t - lT_s] + \mathsf{n}(t).$$

Discrete Time Representation

$\frac{1}{50\tau}$

- Sampling the received signal at the symbol rate T_s ($\mathsf{y}_k = \mathsf{y}(t_0 + kT_s)$, using the epoch t_0) yields

$B_{cb} = \frac{1}{T_m}$ coherence bandwidth

$T_{ct} = \frac{1}{B_d}$ coherence time

$\frac{9}{16\pi B_d}$

$W > B_{cb}$: frequency selective fading
$W < B_{cb}$: flat fading

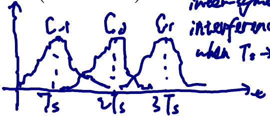
$T > T_{ct}$: fast fading
$T < T_{ct}$: slow fading

$$\mathsf{y}_k = \sum_{l=-\infty}^{\infty} \sqrt{E_s} c_l \mathsf{h}_{t_0+kT_s}[t_0 + (k-l)T_s] + \mathsf{n}(t_0 + kT_s)$$

$$= \sum_{l=-\infty}^{\infty} \sqrt{E_s} c_l \mathsf{h}_k[k-l] + \mathsf{n}_k$$

Sample at nT_s.

inter-symbol interference $\to 0$ when $T_s \to T_{max}$



Example

At time $k = 0$, the channel has two taps: $\mathsf{h}_0[0]$, $\mathsf{h}_0[1]$

$W << B_{cb}$

$t=0$, $\tau=0$; $t=0$, $\tau=1$

$$\mathsf{y}_0 = \sqrt{E_s}\left[c_0 \mathsf{h}_0[0] + c_{-1}\mathsf{h}_0[1]\right] + \mathsf{n}_0$$

- If $T_s >> \tau_{max}$, (symbol period) (max delay) (all symbols arrive at the same time roughly)
  - $\mathsf{h}_B(t, \tau)$ is modeled by a single dependence on t: write simply as $\mathsf{h}_B(t)$ (or $\mathsf{h}(t)$ using the same abuse of notation). In the sampled domain, $\mathsf{h}_k = \mathsf{h}(t_0 + kT_s)$.
  - the channel is then said to be *flat fading* or narrowband

$$\boxed{\mathsf{y}_k = \sqrt{E_s}\mathsf{h}_k c_k + \mathsf{n}_k}$$ (No ISI. flat channel)

- Otherwise the channel is said to be *frequency selective*.

$$y_k = y(t_0 + kT_s) = \sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, h_{t_0+kT_s}[t_0 + (k-l)T_s] + n(t_0 + kT_s)$$

$$= \underbrace{\sum_{l=-\infty}^{+\infty} \sqrt{E_s}\, c_l\, h_k[k-l] + n_k}$$

components: contribution of l-th sample

The Wireless Channel

$h_k[k-l]$ has multiple diracs due to 2 ISI sources:

1. filters do not satisfy Nyquist Criterion (solve by raised cosine filter)
2. delay create former interference between symbols. (tackle by equalisation)

Raised cosine vs. sinc filter:

The tail of raised cosine filter in time domain decays faster than sinc:
- lower ISI
- larger bandwidth

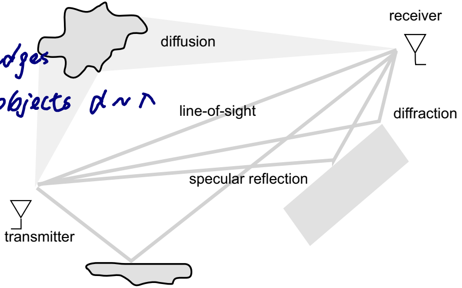
The Wireless Channel

- Multipath

reflection: $d \gg \lambda$

diffraction: sharp edges

diffusion: multiple objects $d \sim \lambda$



$\Delta P(d)$

- Wireless channel varies: → s. min
  - Long time scale: Large Scale Fading (Path-Loss and Shadowing)
  - Short time scale: Small Scale Fading

↳ ms

Large Scale Fading: Path-Loss and Shadowing

- Time constants associated with variations are very long as the mobile moves.
  - many seconds or minutes

Path loss: steady decrease of the long-term average signal power as a function of distance.

- Important for cell site planning and rough estimate of network performance.

- Modeling:
  - Maxwells equations: complex and impractical
  - Free space and 2-path models: (too) simple (LoS + reflection)
  - Ray tracing models: requires site-specific information (modeling)
    find all reflections
  - Simplified power falloff models: good for high-level analysis (based on measurements)
  - Empirical and Standards-based Models: not accurate, used to assess different designs

Free space model

- Path loss for unobstructed LOS path: distance R between Tx and Rx
- Friis equation: received power attenuates like $1/R^2$

$$\boxed{\begin{aligned} P_r &= P_t D_t D_r \left(\frac{\lambda}{4\pi R}\right)^2 \\ P_r|_{\mathsf{dB}} &= P_t|_{\mathsf{dB}} + D_t|_{\mathsf{dB}} + D_r|_{\mathsf{dB}} + 20\log_{10}\left(\frac{\lambda}{4\pi R}\right) \end{aligned}}$$

  - $|_{\mathsf{dB}}$ indicates the conversion to dB
  - D_t, D_r the transmit and receive antenna directivities
  - $R >> \lambda$ the wavelength (far-field)

- $\boxed{\text{Path Loss: } \Lambda_0|_{\mathsf{dB}} = 20\log_{10}\left(\frac{4\pi R}{\lambda}\right)}$

no physical loss
only capture part energy

2-path model

- Path loss for one LOS path and 1 ground (or reflected) bounce

- Ground bounce approximately cancels LOS path above critical distance
- Received power $\frac{1}{R^n}$
  - Inversely proportional to R^2 (small R)
  - Inversely proportional to R^4 ($R > Rc$)

R_c : coherence distance
the distance in which the channel appears to be static/constant.

- Important in rural areas where base stations are placed on roads

Simplified Path Loss Model

- Used extensively in system design
- A real-valued deterministic attenuation term modeled as $\Lambda_0 \propto R^\eta$

measurement distance

$$\boxed{P_r = P_t K \left(\frac{R_0}{R}\right)^\eta}$$

- Important parameter: the path loss exponent η
  - Determined empirically: $2 \le \eta \le 8$.
- Path Loss: $\boxed{\Lambda_0|_{\text{dB}} = L_0|_{\text{dB}} + 10\eta \log_{10}\left(\frac{R}{R_0}\right)}$
  - L_0 is the deterministic path-loss at a reference distance R_0

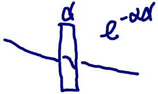
Path loss model: ideal. (deterministic)

Shadowing model: includes attenuation from obstructions/obstacles. (random)

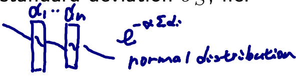
- What about at higher frequencies? mmWave?
  - Much less mature, on-going characterization
  - Path loss large due to frequency, rain, and oxygen
  - Limited to short range or would need massive MIMO

Shadowing

- Models attenuation from obstructions/obstacles
- Random due to random number and type of obstructions



- Modeled as a lognormal random variable S
  - $S|_{\text{dB}} = 10\log_{10}(S)$ is a zero-mean normal variable of given standard deviation σ_S, i.e. $S|_{\text{dB}} \sim \mathcal{N}(0, \sigma_S^2)$.



  - zero mean as mean captured in path loss
  - $4 < \sigma_S < 12$
  - Central Limit Theorem used to explain this model
- Path loss + shadowing

$$\Lambda|_{\text{dB}} = \Lambda_0|_{\text{dB}} + S|_{\text{dB}} = L_0|_{\text{dB}} + 10\eta\log_{10}\left(\frac{R}{R_0}\right) + S|_{\text{dB}},$$

$$P_r|_{\text{dB}} = P_t|_{\text{dB}} + K|_{\text{dB}} - 10\eta\log_{10}\left(\frac{R}{R_0}\right) - S|_{\text{dB}}$$

Path loss + shadowing at the reference point.

  - Λ is sometimes simply known as the path-loss.

Large Scale Fading and Network Performance

only related to path loss + fading.

- For user q in cell i, the *wideband/long-term SINR*

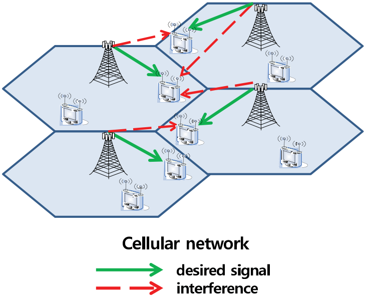
PL + shadowing

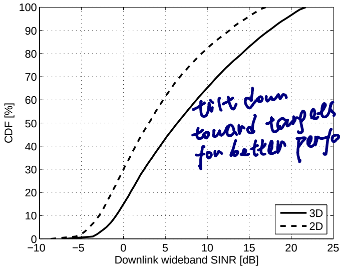
transmitted power (P_t)

$$SINR_{w,q} = \frac{\Lambda_{q,i}^{-1} E_{s,i}}{\sigma_{n,q}^2 + \sum_{j \neq i} \Lambda_{q,j}^{-1} E_{s,j}}.$$

avg. received power

- Function of major propagation mechanisms (path loss, shadowing, antenna radiation patterns,...), base stations deployment and user distribution.
- CDF of $SINR_{w,q}$ in a frequency reuse 1 network (cells share the same frequency band) with 2D and 3D antenna patterns in urban macro deployment.





tilt down toward targets for better performance.

Integrating Large Scale and Small Scale Fading

channel

- Assuming narrowband channels and given specific Tx and Rx locations, h_k is modeled as

$$\boxed{\mathsf{h}_k = \frac{1}{\sqrt{\Lambda_0 \, S}} \, h_k,}$$

fading

pathloss shadowing

where

  - *path-loss* Λ_0: $\Lambda_0 \propto R^\eta$ where η.
  - *shadowing* S: Lognormal random variable, $S|_{\mathsf{dB}} \sim \mathcal{N}(0, \sigma_S^2)$.
  - *fading* h_k: caused by the combination of non coherent multipaths. By definition of Λ_0 and S, $\mathcal{E}\{\, |h|^2 \,\} = 1$.

- Alternatively, $\mathsf{h}_k = \Lambda^{-1/2} \, h_k$ with Λ modeled on a logarithm scale

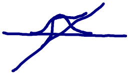
$$\Lambda|_{\mathsf{dB}} = \Lambda_0|_{\mathsf{dB}} + S|_{\mathsf{dB}} = L_0|_{\mathsf{dB}} + 10\eta \log_{10}\left(\frac{R}{R_0}\right) + S|_{\mathsf{dB}},$$

Reminder: Gaussian Random Variable

- *Real Gaussian random variable* x with mean $\mu = \mathcal{E}\{x\}$ and variance σ^2

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

  Standard Gaussian random variable: $\mu = 0$ and $\sigma^2 = 1$
- *Complex Gaussian random variable* $x = x_r + jx_i$: $[x_r, x_i]^T$ is a real Gaussian random vector.
- Important case: $x = x_r + jx_i$ is such that its real and imaginary parts are i.i.d. zero mean Gaussian variables of variance σ^2 (circularly symmetric complex Gaussian random variable).
- $s = |x| = \sqrt{x_r^2 + x_i^2}$ is Rayleigh distributed



$x = x_r + j x_i$ (polar coordinate)
$x = s e^{j\phi}$
Jacobian

$$p(s) = \frac{s}{\sigma^2} \exp\left(-\frac{s^2}{2\sigma^2}\right).$$

- $y = s^2 = |x|^2 = x_r^2 + x_i^2$ is χ_2^2 (i.e. exponentially) distributed (with two degrees of freedom)

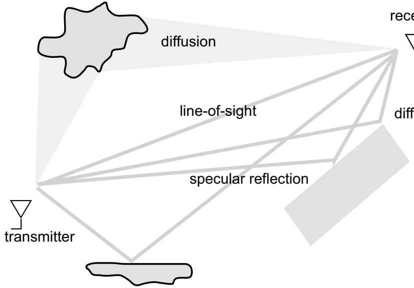
$\mathcal{N}(0, \sigma^2)$

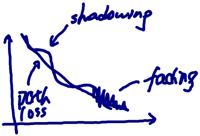
$$p_y(y) = \frac{1}{2\sigma^2} \exp\left(-\frac{y}{2\sigma^2}\right).$$

Hence, $\mu = \mathcal{E}\{y\} = 2\sigma^2$.

Small Scale Fading (or simply Fading)

- Multipaths





- Assuming that the signal reaches the receiver via a large number of paths of similar energy,
  - Central Limit Theorem used to explain this model
  - h is modeled such that its real and imaginary parts are i.i.d. zero mean Gaussian variables of variance σ^2 (circularly symmetric complex Gaussian variable).
  - Recall $\mathcal{E}\{\,|h|^2\,\} = 2\sigma^2 = 1$.

fading brings neither new energy nor loss (captured in path loss)

Small Scale Fading (or simply Fading)

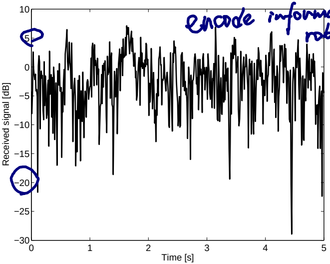
- The channel *amplitude* $s \triangleq |h|$ follows a *Rayleigh* distribution,

$$p_s(s) = \frac{s}{\sigma^2} \exp\left(-\frac{s^2}{2\sigma^2}\right),$$

  whose first two moments are

$$\mathcal{E}\{s\} = \sigma\sqrt{\frac{\pi}{2}}, \qquad \mathcal{E}\{s^2\} = 2\sigma^2 = \mathcal{E}\{\,|h|^2\,\} = 1.$$

- The *phase* of h is uniformly distributed over $[0, 2\pi)$
- Typical received signal strength of a Rayleigh fading channel



encode information robust to channel flactuations

Fading and Diversity

System Model

- Path loss models are identical for both single- and multi-antenna systems.
- For point to point systems, it is common to discard the path loss and shadowing and only investigate the effect due to fading, i.e. the classical model for narrowband channels

path loss ×
shadowing × ⇒ narrowband channels
fading ✓

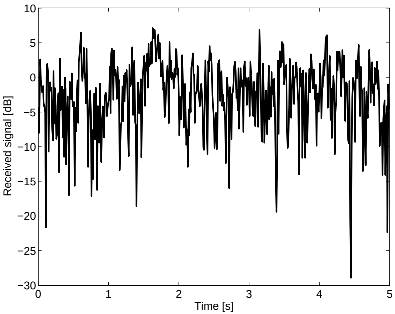
$$y = \sqrt{E_s} hc + n,$$

  where the time index is removed for better legibility and n is usually taken as white Gaussian distributed, $\mathcal{E}\{n_k n_l^*\} = \sigma_n^2 \delta(k-l)$.
- E_s can then be seen as an average received symbol energy.
- The *average SNR* is defined as $\rho \triangleq E_s/\sigma_n^2$.
- The *instantaneous SNR* is $E_s|h|^2/\sigma_n^2 = \rho|h|^2$.

Fading and Diversity

- The signal level randomly fluctuates, with some sharp declines of power and instantaneous received SNR known as *fades*.

Single-user:
multi-path have similar power.
↳ only model fading.



- When the channel is in a deep fade, a reliable decoding of the transmitted signal may not be possible anymore, resulting in an error.
- How to recover the signal? Use of diversity techniques

Maximum Likelihood Detection

- Decision rule: choose the hypothesis that maximizes the conditional density

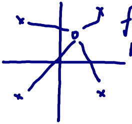
$$\arg\max_x p(y|x) = \arg\max_x \log p(y|x)$$

- If real AWGN $y = x + n$ with $n \sim N(0, \sigma_n^2)$,

maximize.

$$\log \frac{1}{\sqrt{2\pi\sigma_n^2}} \cdot \left(-\frac{(y-x)^2}{2\sigma_n^2}\right)$$

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(y-x)^2}{2\sigma_n^2}\right)$$



find x

minimize $(y-x)^2$

and

$$\underbrace{\arg\max_x p(y|x) = \arg\min_x (y-x)^2}$$

- If $y = \sqrt{E_s}hc + n$, the ML decision rule becomes

need CSIR

$$\underbrace{\arg\min_c \left| y - \sqrt{E_s} hc \right|^2}$$

Impact of Fading

- What is the impact of fading on system performance?
- Consider the simple case of BPSK transmission through an AWGN channel and a SISO Rayleigh fading channel:
  - In the absence of fading ($h = 1$), the symbol-error rate (SER) in an additive white Gaussian noise (AWGN) channel is given by BPSK:

BER: $P_e = P_e(0,1) = Q\left(\sqrt{\frac{d_{0,1}^2}{2N_0}}\right)$

SER: $\bar{P} = \bar{P}(s_i, s_j) = Q\left(\sqrt{\frac{d_{s_i s_j}^2}{2N_0}}\right)$

s_1 s_0

$-\sqrt{E_s}$ $\sqrt{E_s}$

$$\bar{P} = \mathcal{Q}\left(\sqrt{\frac{2E_s}{\sigma_n^2}}\right) = \mathcal{Q}(\sqrt{2\rho}),$$

$d_{s_0 s_1} = 2\sqrt{E_s} \Rightarrow d_{s_0 s_1}^2 = 4E_s$

$\therefore \bar{P} = Q\left(\sqrt{\frac{d_{s_0 s_1}^2}{2N_0}}\right) = Q\left(\sqrt{\frac{2E_s}{\sigma_n^2}}\right) = Q(\sqrt{2\rho})$

where $\mathcal{Q}(x)$ is the Gaussian $\mathcal{Q}$-function defined as CDF of Gaussian distribution

$$\mathcal{Q}(x) \triangleq P(y \geq x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy.$$

  - In the presence of (Rayleigh) fading, the received signal level fluctuates as $s\sqrt{E_s}$, and the SNR varies as ρs^2. As a result, the SER

$$\bar{P} = \int_0^\infty \mathcal{Q}(\sqrt{2\rho}s) p_s(s)\, ds$$

$$= \frac{1}{2}\left(1 - \sqrt{\frac{\rho}{1+\rho}}\right)$$

$$\overset{(\rho\nearrow)}{\cong} \frac{1}{4\rho}$$

although the average SNR $\bar{\rho} = \int_0^\infty \rho s^2\, p_s(s)\, ds$ remains equal to ρ.

$SER_{BPSK\ fading} = \frac{1}{2}\left(1 - \sqrt{\frac{\rho}{1+\rho}}\right)$

high SNR $\frac{1}{4\rho}$ $\boxed{\propto \rho^{-1}}$

fading channel: lower SER ($\propto \rho^{-1}$)

- unit diversity gain
- average SNR not change

$\bar{P} = \frac{1}{2}\left(1 - \sqrt{\frac{\rho}{1+\rho}}\right) = \frac{1}{2}\left(1 - \frac{1}{\sqrt{\frac{1}{\rho}+1}}\right)$

$\because (1+\frac{x}{2})^2 = 1 + x + \frac{x^2}{4} \overset{x\to 0}{\cong} 1+x$

$\therefore \sqrt{1+x} \overset{x\to 0}{\cong} 1 + \frac{x}{2}$

$\Rightarrow \sqrt{1+\frac{1}{\rho}} \overset{\rho\to\infty}{\cong} 1 + \frac{1}{2\rho}$

$\therefore \bar{P} \overset{\rho\to\infty}{\cong} \frac{1}{2}\left(1 - \frac{1}{1+\frac{1}{2\rho}}\right) = \frac{1}{2}\left(1 - \frac{2\rho}{2\rho+1}\right)$

$= \frac{1}{2} \cdot \frac{1}{2\rho+1} = \frac{1}{4\rho+2} \overset{\rho\to\infty}{\cong} \frac{1}{4\rho}$

Diversity in Multiple Antennas Wireless Systems

- How to combat the impact of fading? Use diversity techniques
- The principle of diversity is to provide the receiver with multiple versions (called diversity branch) of the same transmitted signal.
  - Independent fading conditions across branches needed.
  - Diversity stabilizes the link through channel hardening which leads to better error rate.
  - Multiple domains: time (coding and interleaving), frequency (equalization and multi-carrier modulations) and space (multiple antennas/polarizations).

Tx | X Rx diversity in polarization

- *Array Gain*: increase in average output SNR (i.e., at the input of the detector) relative to the single-branch average SNR ρ

far spacing => independent fading

2 Rx. double received signal power
3dB array gain

$$\boxed{g_a \triangleq \frac{\bar{\rho}_{out}}{\bar{\rho}} = \frac{\bar{\rho}_{out}}{\rho}}$$

- *Diversity Gain*: increase in the error rate slope as a function of the SNR. Defined as the negative slope of the log-log plot of the average error probability $\bar{P}$ versus SNR

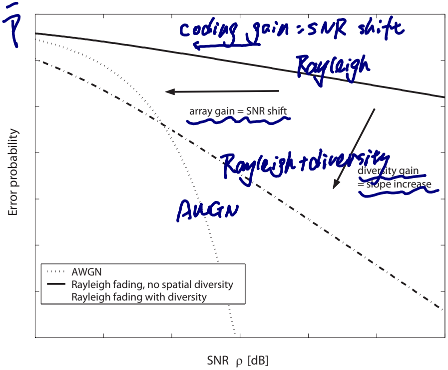
diversity: multiple independent replicas of the transmitted signal

$$g_d^o(\rho) \triangleq -\frac{\log_2\left(\bar{P}\right)}{\log_2\left(\rho\right)}.$$

$\bar{P} \propto \rho^{-g_d}$

The diversity gain is commonly taken as the asymptotic slope, i.e., for $\rho \to \infty$.

- Illustration of diversity and array gains



Careful that the curves have been plotted against the single-branch average SNR $\bar{\rho} = \rho$!
If plotted against the output average SNR $\bar{\rho}_{out}$, the array gain disappears.

- *Coding Gain*: a shift of the error curve (error rate vs. SNR) to the left, similarly to the array gain.
  - If the error rate vs. the average receive SNR $\bar{\rho}_{out}$, any variation of the array gain is invisible but any variation of the coding gain is visible: for a given SNR level $\bar{\rho}_{out}$ at the input of the detector, the error rates will differ.

SIMO Systems

- Receive diversity may be implemented via two rather different combining methods:
  - *selection combining*: the combiner selects the branch with the highest SNR among the n_r receive signals, which is then used for detection,
  - *gain combining*: the signal used for detection is a linear combination of all branches, $z = \mathbf{g}\mathbf{y}$, where $\mathbf{g} = [g_1, \ldots, g_{n_r}]$ is the combining vector.
    1. Equal Gain Combining
    2. Maximal Ratio Combining
    3. Minimum Mean Square Error Combining
- Space antennas sufficiently far apart from each other so as to experience independent fading on each branch.
- We assume that the receiver is able to acquire the perfect knowledge of the channel.

Receive Diversity via Selection Combining

$P(s<S) = \int_0^S \frac{s}{\sigma^2} e^{-\frac{s^2}{2\sigma^2}} ds = -e^{-\frac{s^2}{2\sigma^2}}\Big|_0^S = -e^{-\frac{S^2}{2\sigma^2}} + 1$

- Assume that the n_r channels are independant and identically Rayleigh distributed (i.i.d.) with unit energy and that the noise levels are equal on each antenna.
- Choose the branch with the largest amplitude $s_{max} = \max\{s_1, \ldots, s_{n_r}\}$.
- The probability that s falls below a certain level S is given by its CDF

$s = |h|$: channel amplitude

$$\boxed{P[s < S] = 1 - e^{-S^2/2\sigma^2}.}$$

$\frac{d}{ds}(1-e^{-s^2})^{n_r} = n_r(1-e^{-s^2})^{n_r-1} \cdot 2s\, e^{-s^2}$

- The probability that s_{max} falls below a certain level S is given by

$$P\left[\underline{s_{max} < S}\right] = P[s_1, \ldots, s_{n_r} \le S] = \left[1 - e^{-S^2}\right]^{n_r}.$$

- The PDF of s_{max} is then obtained by derivation of its CDF

$u = \rho s^2$, $du = 2\rho s\, ds$, $dv = n_r 2s e^{-s^2}[1-e^{-s^2}]^{n_r-1} ds$, $v = [1-e^{-s^2}]^{n_r}$

$$p_{s_{max}}(s) = n_r\, 2s\, e^{-s^2}\left[1 - e^{-s^2}\right]^{n_r - 1}.$$

$\bar{P}_{out} = \int_0^\infty \rho s^2 P_{s_{max}}(s) ds = \int_0^\infty \rho s^2 \cdot n_r 2s e^{-s^2}[1-e^{-s^2}]^{n_r-1} ds = uv\Big|_0^\infty - \int_0^\infty 2\rho s[1-e^{-s^2}]^{n_r} ds$

- The average SNR at the output of the combiner $\bar{\rho}_{out}$ is eventually given by

$$\bar{\rho}_{out} = \int_0^\infty \rho s^2 p_{s_{max}}(s)\, ds = \boxed{\rho \sum_{n=1}^{n_r} \frac{1}{n}} \overset{n_r \nearrow}{\approx} \boxed{\rho\left[\gamma + \log(n_r) + \frac{1}{2n_r}\right]}.$$

<u>where $\gamma \approx 0.57721566$ is Euler's constant.</u> We observe that the array gain g_a is of the order of $\log(n_r)$.

CDF → PDF → SNR $= \int_0^\infty \rho s^2 P(s) ds$
→ SER $= \int_0^\infty Q(\sqrt{2\rho}\, s) P(s) ds$

Receive Diversity via Selection Combining

- For BPSK and a two-branch diversity, the SER as a function of the average SNR per channel ρ writes as

$SER_{BPSK, 2branches} = \frac{1}{2} - \sqrt{\frac{\rho}{1+\rho}} + \frac{1}{2}\sqrt{\frac{\rho}{2+\rho}}$

$\xrightarrow{\rho \to \infty} \frac{3}{8\rho^2} \propto \frac{1}{\rho^2}$

$$
\begin{aligned}
\bar{P} &= \int_0^\infty Q(\sqrt{2\rho s})\, p_{s_{max}}(s)\, ds \\
&= \frac{1}{2} - \sqrt{\frac{\rho}{1+\rho}} + \frac{1}{2}\sqrt{\frac{\rho}{2+\rho}} \\
&\overset{\rho\nearrow}{\cong} \frac{3}{8\rho^2}.
\end{aligned}
$$

Selection combining:
array gain $\propto \log(N_r)$
diversity gain $= N_r$

The slope of the bit error rate curve is equal to 2.

- In general, the diversity gain g_d^o of a n_r-branch selection diversity scheme is equal to n_r. Selection diversity extracts all the possible diversity out of the channel.

Receive Diversity via Gain Combining

- In gain combining, the signal z used for detection is a linear combination of all branches,

$$z = \mathbf{gy} = \sum_{n=1}^{n_r} g_n y_n = \sqrt{E_s}\mathbf{ghc} + \mathbf{gn}$$

constructively

Gain combining:
array gain ∝ Nr
diversity gain = Nr

where
  - g_n's are the combining weights and $\mathbf{g} \triangleq [g_1, \ldots, g_{n_r}]$
  - the data symbol c is sent through the channel and received by n_r antennas
  - $\mathbf{h} \triangleq [h_1, \ldots, h_{n_r}]^T$

Tx

$$\text{Rx} \begin{bmatrix} h_1 \\ \vdots \\ h_{n_r} \end{bmatrix}$$

- Assume Rayleigh distributed channels $h_n = |h_n| e^{j\phi_n}$, $n = 1, \ldots, n_r$, with unit energy, all the channels being independent.
- *Equal Gain Combining*: fixes the weights as $g_n = e^{-j\phi_n}$. only need channel phase.
  - Mean value of the output SNR $\bar{\rho}_{out}$ (averaged over the Rayleigh fading):

Rx

problem.
Rx with different channel strength

sum signal power

$$\bar{\rho}_{out} = \frac{\mathcal{E}\left\{\left[\sum_{n=1}^{n_r}\sqrt{E_s}\,|h_n|\right]^2\right\}}{n_r\sigma_n^2} = \ldots = \rho\left[1 + (n_r - 1)\frac{\pi}{4}\right],$$

sum noise power

where the expectation is taken over the channel statistics. The array gain grows linearly with n_r, and is therefore larger than the array gain of selection combining.
  - The diversity gain of equal gain combining is equal to n_r analogous to selection.

Reminder: Chi-Square Distribution and MGF

- χ_n^2 is the sum of the square of n i.i.d. zero-mean Gaussian random variables.

- Assume n i.i.d. zero mean complex Gaussian variables $h_1, \ldots, h_n$ (real and imaginary parts with variance σ^2). Defining $u = \sum_{k=1}^{n} |h_k|^2$, the MGF of u is given by

Moment generating function

$$\mathcal{M}_u(\tau) = \mathcal{E}\{e^{\tau u}\} = \left[\frac{1}{1 - \underbrace{2\sigma^2}_{1}\tau}\right]^{n \to n_r},$$

-ρ

Receive Diversity via Gain Combining

$\boxed{\begin{array}{l} \text{MRC}: g_n = h_n^* \\ \text{array gain} = N_r \\ \text{diversity gain} = N_r \end{array}}$

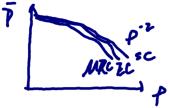
- *Maximal Ratio Combining*: the weights are chosen as $g_n = h_n^*$
  - It maximizes the average output SNR $\bar{\rho}_{out}$

$z = \sqrt{E_s}\,\|h\|^2 c + h^* n$

$$\bar{\rho}_{out} = \frac{E_s}{\sigma_n^2}\mathcal{E}\left\{\frac{\|\mathbf{h}\|^4}{\|\mathbf{h}\|^2}\right\} = \rho\mathcal{E}\left\{\|\mathbf{h}\|^2\right\} = \rho n_r.$$

$\boxed{\begin{array}{l} \bar{\rho}_{out} = \rho N_r \\ \bar{P} = (4\rho)^{-N_r}\binom{2N_r-1}{N_r} \end{array}}$

The array gain g_a is thus always equal to n_r, or equivalently, the output SNR is the sum of the SNR levels of all branches (holds true irrespective of the correlation between the branches).



  - For BPSK transmission, the symbol error rate reads as

$Q(\sqrt{2\rho})$ AWGN
$\downarrow$
$Q(\sqrt{2\rho\|h\|^2})$ fading

$$\bar{P} = \int_0^\infty \mathcal{Q}\left(\sqrt{2\rho u}\right)p_u(u)\,du$$

(re + im) · channel

where $u = \|\mathbf{h}\|^2$ is χ^2 distribution with $2n_r$ degrees of freedom when the different channels are i.i.d. Rayleigh

$\bar{P} = E_h\{Q(\sqrt{2\rho\|h\|^2})\}$

$$p_u(u) = \frac{1}{(n_r-1)!}\,u^{n_r-1}e^{-u}.$$

At high SNR, $\bar{P}$ becomes

$$\boxed{\bar{P} = (4\rho)^{-n_r}\begin{pmatrix} 2n_r - 1 \\ n_r \end{pmatrix}}.$$

The diversity gain is again equal to n_r.

Receive Diversity via Gain Combining

– For alternative constellations, the error probability is given, assuming ML detection, by

SER$_{ML}$ for general constellations.

$$\bar{P} \approx \int_0^\infty \bar{N}_e \mathcal{Q}\left(d_{min}\sqrt{\frac{\rho u}{2}}\right) p_u(u)\, du,$$

$\bar{P} \leq \mathcal{E}\{e^{-\rho\|h\|^2}\}$. MGF

$$\leq \bar{N}_e \mathcal{E}\left\{e^{-\frac{d_{min}^2 \rho u}{4}}\right\} \quad \text{(using Chernoff bound } \mathcal{Q}(x) \leq \exp\left(-\frac{x^2}{2}\right)\text{)}$$

MGF param

where $\bar{N}_e$ and d_{min} are respectively the number of nearest neighbors and minimum distance of separation of the underlying constellation.

Since u is a χ^2 variable with $2n_r$ degrees of freedom, the above average upper-bound is given by

$$\bar{P} \leq \bar{N}_e \left(\frac{1}{1+\rho d_{min}^2/4}\right)^{n_r}$$

$$\overset{\rho\nearrow}{\leq} \bar{N}_e \left(\frac{\rho d_{min}^2}{4}\right)^{-n_r}.$$

The diversity gain g_d^o is equal to the number of receive branches in i.i.d. Rayleigh channels.

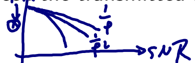
Receive Diversity via Gain Combining

MMSE: $\boxed{g^* = h^H R_{n_i}^{-1} \quad R_{n_i} = \mathcal{E}\{n_i n_i^H\}}$ $\quad SNR = \dfrac{E_s \{h^H R_{n_i}^{-1} h\}^2}{h^H R_{n_i}^{-1} R_{n_i} h}$

- *Minimum Mean Square Error Combining*
  - So far noise was white Gaussian. When the noise (and interference) is colored, MRC is not optimal anymore.
  - Let us denote the combined noise plus interference signal as $\mathbf{n}_i$ such that $\mathbf{y} = \sqrt{E_s}\mathbf{h}c + \mathbf{n}_i$.

    $g = h_{eff}^H R_{n_i}^{-1/2} = h^H R_{n_i}^{-1/2} R_{n_i}^{-1/2} = h^H R_{n_i}^{-1}$
  - An optimal gain combining technique is the minimum mean square error (MMSE) combining, where the weights are chosen in order to minimize the mean square error between the transmitted symbol c and the combiner output z, i.e.,

Tx Rx h_1 h_2 g_1 P P_2



$$\mathbf{g}^\star = \arg\min_{\mathbf{g}} \mathcal{E}\left\{ |\mathbf{g}\mathbf{y} - c|^2 \right\}.$$

$\rho_{out} = E_s h^H R_{n_i}^{-1} h$

$\mathcal{E}\{\tilde{n}_i \tilde{n}_i^H\} = R^{-1/2} \mathcal{E}\{n_i n_i^H\} R^{-H/2} = I$

  - The optimal weight vector $\mathbf{g}^\star$ is given by

$g_1 = h_1^*$ $\quad g_2 = h_2^*$ $\quad \rho_{MRC} = \|h\|^2 \rho$ $\quad h = [h_1\ h_2]$

$$\boxed{\mathbf{g}^\star = \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1},}$$

$R_{n_i}^{-1/2} y = \sqrt{E_s} \underbrace{R_{n_i}^{-1/2} h}_{h_{eff}} c + \underbrace{R_{n_i}^{-1/2} n_i}_{\tilde{n}_i}$

where $\mathbf{R}_{\mathbf{n}_i} = \mathcal{E}\left\{\mathbf{n}_i \mathbf{n}_i^H\right\}$ is the correlation matrix of the combined noise plus interference signal $\mathbf{n}_i$.

$\underbrace{h_{eff}^H R_{n_i}^{-1/2}}_{g} y = \sqrt{E_s}\, h_{eff}^H R_{n_i}^{-1/2} h c + h_{eff}^H \tilde{n}_i$

  - Such combiner can be thought of as first whitening the noise plus interference by multiplying $\mathbf{y}$ by $\mathbf{R}_{\mathbf{n}_i}^{-1/2}$ and then match filter the effective channel $\mathbf{R}_{\mathbf{n}_i}^{-1/2}\mathbf{h}$ using $\mathbf{h}^H\mathbf{R}_{\mathbf{n}_i}^{-H/2}$.

if $R_{n_i} = \sigma^2 I$ (no interf)

  - The Signal to Interference plus Noise Ratio (SINR) at the output of the MMSE combiner simply writes as

$\rho_{out} = \frac{E_s}{\sigma^2} h^H h = \rho \|h\|^2$

$$\boxed{\rho_{out} = E_s \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}.}$$

  - In the absence of interference and the presence of white noise, MMSE combiner reduces to MRC filter up to a scaling factor.

Receive Diversity via Gain Combining

$n_i = n + h_i x \quad R_{n_i} = h_i P_x h_i^H + \sigma_n^2 I_{n_r}$

Example

Question: Assume a transmission of a signal c from a single antenna transmitter to a multi-antenna receiver through a SIMO channel $\mathbf{h}$. The transmission is subject to the interference from another transmitter sending signal x through the interfering SIMO channel $\mathbf{h}_i$. c, x zero mean

$P_x = E[|x|^2]$

The received signal model writes as

$g^* = h^H R_{n_i}^{-1} = h^H (h_i P_x h_i^H + \sigma_n^2 I_{n_r})^{-1}$

$$\mathbf{y} = \mathbf{h}c + (\mathbf{h}_i x + \mathbf{n})$$

$z = g^* y = h^H (h_i P_x h_i^H + \sigma_n^2 I_{n_r})^{-1} (hc + h_i x + n)$

where $\mathbf{n}$ is the zero mean complex additive white Gaussian noise (AWGN) vector with $\mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2 \mathbf{I}_{n_r}$.
We apply a combiner $\mathbf{g}$ at the receiver to obtain the observation $z = \mathbf{g}\mathbf{y}$.
Derive the expression of the MMSE combiner and the SINR at the output of the combiner.

$= h^H R_i^{-1} h c + h^H R_{n_i}^{-1} n_i$ (no interference!)

Receive Diversity via Gain Combining

Example

Answer: The MMSE combiner $\mathbf{g}$ is given by

$$\mathbf{g} = \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1}$$

where $\mathbf{R}_{\mathbf{n}_i} = \mathcal{E}\left\{\mathbf{n}_i \mathbf{n}_i^H\right\}$ with $\mathbf{n}_i = \mathbf{h}_i x + \mathbf{n}$.
Hence $\mathbf{R}_{\mathbf{n}_i} = \mathbf{h}_i P_x \mathbf{h}_i^H + \sigma_n^2 \mathbf{I}_{n_r}$ with $P_x = \mathcal{E}\{\,|x|^2\,\}$, the power of the interfering signal.
Hence,

$$\mathbf{g} = \mathbf{h}^H \left(\mathbf{h}_i P_x \mathbf{h}_i^H + \sigma_n^2 \mathbf{I}_{n_r}\right)^{-1}.$$

At the receiver, we obtain

$$z = \mathbf{g}\mathbf{y} = \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h} c + \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{n}_i.$$

Receive Diversity via Gain Combining

Example

Answer: The output SINR writes

$$
\begin{aligned}
\rho_{out} &= \frac{\left|\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}\right|^2 P_c}{\mathcal{E}\left\{\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{n}_i \left(\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{n}_i\right)^H\right\}} \\
&= \frac{\left|\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}\right|^2 P_c}{\mathcal{E}\left\{\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{n}_i \mathbf{n}_i^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}\right\}} \\
&= \frac{\left|\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}\right|^2 P_c}{\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}} \\
&= \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h} P_c \\
&= P_c \mathbf{h}^H \left(\mathbf{h}_i P_x \mathbf{h}_i^H + \sigma_n^2 \mathbf{I}_{n_r}\right)^{-1} \mathbf{h} \\
&= \mathsf{SNR}\, \mathbf{h}^H \left(\mathsf{INR}\, \mathbf{h}_i \mathbf{h}_i^H + \mathbf{I}_{n_r}\right)^{-1} \mathbf{h}
\end{aligned}
$$

$\mathcal{E}\{n_i n_i^H\} = R_{n_i}$

interference negligible

MMSE ⇒ MRC.

with $P_c = \mathcal{E}\{\,|c|^2\,\}$, SNR $= P_c/\sigma_n^2$ (the average SNR), INR $= P_x/\sigma_n^2$ (the average INR - Interference to Noise Ratio). □

MISO Systems

MISO Tx h_1 Rx   Channel est. need resources.

TDD → same h in uplink and downlink → mobile pilot to BS
reciprocity (not ideal in reality)

- MISO systems exploit diversity at the transmitter through the use of n_t transmit antennas in combination with pre-processing or precoding. BS UE f DL|UL   large b/w

FDD → receive pilot from BS. encode and report to BS.

- A significant difference with receive diversity is that the transmitter might not have the knowledge of the MISO channel. f_1 DL f_2 UL   2× small b/w
  - At the receiver, the channel is easily estimated.
  - At the transmit side, feedback from the receiver is required to inform the transmitter.

- There are basically two different ways of achieving *direct transmit diversity*:
  - when Tx has a *perfect channel knowledge*, beamforming can be performed to achieve both diversity and array gains, low velocity
  - when Tx has a *partial or no channel knowledge of the channel*, space-time coding is used to achieve a diversity gain (but no array gain in the absence of any channel knowledge). high velocity

- *Indirect transmit diversity* techniques convert spatial diversity to time or frequency diversity.

Transmit Diversity via Matched Beamforming

$c \to wc \to hwc\sqrt{E_s} + n = (h_1 w_1 + h_2 w_2) c + n$ $\quad [w_1\ w_2] \quad [h_1\ h_2]$ $\quad \max |wh^H|,\ \|w\|=1 \Rightarrow w = \frac{h^H}{\|h\|}$ paths add constructively

- The actual transmitted signal is a vector $\mathbf{c}'$ that results from the multiplication of the signal c by a weight vector $\mathbf{w}$. $y = h\frac{h^H}{\|h\|}c\sqrt{E_s} + n = \frac{\|h\|^2}{\|h\|}c\sqrt{E_s} + n = \|h\|c\sqrt{E_s} + n$
- At the receiver, the signal reads as $\rho_{out} = \frac{E_s\|h\|^2}{\sigma_n^2} = \|h\|^2\rho \Rightarrow E(\rho_{out}) = \rho E(\|h\|^2) = 2\rho$ $\quad h_1^2 + h_2^2$

$c' = wc$

$$y = \sqrt{E_s}\mathbf{h}\mathbf{c}' + n = \sqrt{E_s}\mathbf{h}\mathbf{w}c + n,$$

where $\mathbf{h} \triangleq [h_1, \ldots, h_{n_t}]$ represents the MISO channel vector, and $\mathbf{w}$ is also known as the precoder.
- The choice that maximizes the receive SNR is given by

$$\mathbf{w} = \frac{\mathbf{h}^H}{\|\mathbf{h}\|}.$$

- Transmit along the direction of the matched channel. Hence also known as *matched beamforming* or *transmit MRC* or *maximum ratio transmission (MRT)*. array gain $= n_t$
- The array gain is equal to the number of transmit antennas, i.e. $\bar{\rho}_{out} = n_t\rho$.
- The diversity gain equal to n_t as the symbol error rate is upper-bounded at high SNR by

diversity gain $= n_t$

MRC $\quad$ Tx $\quad$ Rx $\quad$ MRT $\quad$ Tx $\quad$ LRx

$$\bar{P} \leq \bar{N}_e\left(\frac{\rho d_{min}^2}{4}\right)^{-n_t}.$$

- Matched beamforming presents the same performance as receive MRC, but *requires a perfect transmit channel knowledge*.

Transmit Diversity via Space-Time Coding

(assume h unchanged during $T_1 . T_2$)

- *Alamouti scheme* is an ingenious transmit diversity scheme for two transmit antennas which does not require transmit channel knowledge.
  - Assume that the flat fading channel remains constant over the two successive symbol periods, and is denoted by $\mathbf{h} = [h_1 \ h_2]$.

    scale down for power constraint $\frac{1}{\sqrt{2}}$ $\quad (\frac{1}{\sqrt{2}} c_1)^2 + (\frac{1}{\sqrt{2}} c_2)^2 = 1$
  - Two symbols c_1 and c_2 are transmitted simultaneously from antennas 1 and 2 during the first symbol period, followed by symbols $-c_2^*$ and c_1^*, transmitted from antennas 1 and 2 during the next symbol period:

    slot (channel is unchanged in 2 slots)

    same diversity gain

    no array (beamforming gain)

1	2
c_1	$-c_2^*$
c_2	c_1^*

2 symbols in 2 slots. transmitted twice. cost: beamforming gain

space + time

$$y_1 = \sqrt{E_s} h_1 \frac{c_1}{\sqrt{2}} + \sqrt{E_s} h_2 \frac{c_2}{\sqrt{2}} + n_1, \qquad \text{(first symbol period)}$$

$E(P_{out}) = 2P$ $\quad P_{out} = P$

$$y_2 = -\sqrt{E_s} h_1 \frac{c_2^*}{\sqrt{2}} + \sqrt{E_s} h_2 \frac{c_1^*}{\sqrt{2}} + n_2. \qquad \text{(second symbol period)}$$

Rx h_1 h_2 Tx

  - The two symbols are spread over two antennas and over two symbol periods.
  - Equivalently

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} \qquad H^H = \begin{bmatrix} h_1^* & h_2 \\ h_2^* & -h_1 \end{bmatrix}$$

time 1 $\quad y_1 = \frac{1}{\sqrt{2}}(h_1 c_1 + h_2 c_2) + n_1$

time 2 $\quad y_2 = \frac{1}{\sqrt{2}}(-h_1 c_2^* + h_2 c_1^*) + n_2$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \sqrt{E_s} \underbrace{\begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}}_{\mathbf{H}_{eff}} \underbrace{\begin{bmatrix} c_1/\sqrt{2} \\ c_2/\sqrt{2} \end{bmatrix}}_{\mathbf{c}} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}$$

$\|h\|^2$

$z_1 = \frac{1}{\sqrt{2}}(|h_1|^2 + |h_2|^2) c_1 + n_1'$

$z_2 = \frac{1}{\sqrt{2}}(|h_1|^2 + |h_2|^2) c_2 + n_2'$

$y_2^* = \frac{1}{\sqrt{2}}(-h_1^* c_2 + h_2^* c_1) + n_2^*$

$$H^H H = \begin{bmatrix} |h_1|^2 + |h_2|^2 & h_1 h_2 - h_2 h_1 \\ h_1^* h_2^* - h_2^* h_1^* & |h_1|^2 + |h_2|^2 \end{bmatrix}$$

  - Applying the matched filter $\mathbf{H}_{eff}^H$ to the received vector $\mathbf{y}$ effectively decouples the transmitted symbols as shown below

$z = H^H \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \frac{1}{\sqrt{2}}(|h_1|^2 + |h_2|^2) I \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + n'$

Rx need to consider H

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{H}_{eff}^H \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \sqrt{E_s} \left[|h_1|^2 + |h_2|^2 \right] \mathbf{I}_2 \begin{bmatrix} c_1/\sqrt{2} \\ c_2/\sqrt{2} \end{bmatrix} + \mathbf{H}_{eff}^H \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}$$

decoupled as linear combination.

Transmit Diversity via Space-Time Coding

T.H. Tx/fx

$c_1 \mid -c_2^*$

$c_2 \mid c_1$

– The mean output SNR (averaged over the channel statistics) is thus equal to

if channel is similar for f_1 and f_2

$$\bar{\rho}_{out} = \frac{E_s}{\sigma_n^2}\,\mathcal{E}\left\{\frac{\left[\,\|\mathbf{h}\|^2\,\right]^2}{2\,\|\mathbf{h}\|^2}\right\} = \rho.$$

No array gain owing to the lack of transmit channel knowledge.

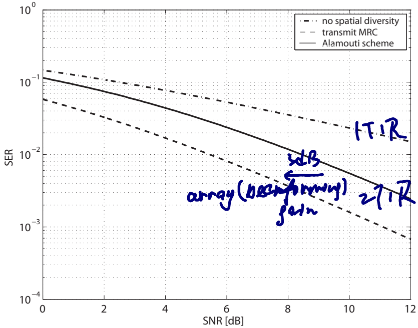
– The average symbol error rate at high SNR can be upper-bounded according to

$$\bar{P} \leq \bar{N}_e \left(\frac{\rho d_{min}^2}{8}\right)^{-2}.$$

same diversity gain

beamforming gain = 1

The diversity gain is equal to $n_t = 2$ despite the lack of transmit channel knowledge.



Transmit MRC vs. Alamouti with 2 transmit antennas in i.i.d. Rayleigh fading channels (BPSK).

Observations:

– At high SNR, any increase in the SNR by 10dB leads to a decrease of SER by 10^{-n} for diversity order n.

- Alamouti, transmit MRC: 2
- No spatial diversity: 1

– Transmit MRC has 3 dB gain over Alamouti

Indirect Transmit Diversity

- It is also possible to convert spatial diversity to time or frequency diversity, which are then exploited using well-known SISO techniques.
- Assume that $n_t = 2$ and that the signal on the second transmit branch is
  - either delayed by one symbol period: the spatial diversity is converted into *frequency diversity* (delay diversity)
  - either phase-rotated: the spatial diversity is converted into *time diversity*
  - The effective SISO channel resulting from the addition of the two branches seen by the receiver now fades over frequency or time. This selective fading can be exploited by conventional diversity techniques, e.g. FEC/interleaving.

orthogonal STBC.

$$CC^H = \alpha I$$

$$\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix} \begin{bmatrix} c_1^* & c_2^* \\ -c_2 & c_1 \end{bmatrix} = \begin{bmatrix} c_1^2 + c_2^2 & c_1 c_2^* - c_2^* c_1 \\ c_2 c_1^* - c_1^* c_2 & c_1^2 + c_2^2 \end{bmatrix} = \|c\|^2 I_2$$

Capacity of Wireless Channels

System Model

- A single-user SISO system with one transmit and one receive antennas over a frequency flat-fading channel.
- The transmit and received signals are related by

$$y_k = \sqrt{E_s} h_k c_k + n_k$$

  where
  - y_k is the received signal,
  - h_k is the complex channel,
  - n_k is a zero mean complex additive white Gaussian noise (AWGN) $\mathcal{CN}(0, \sigma_n^2)$, with $\mathcal{E}\{n_k n_l\} = \sigma_n^2$ if $k = l$ and 0 otherwise.
  - $\rho = E_s/\sigma_n^2$ represents the average SNR.
- Power constraint: $\mathcal{E}\left\{|c_k|^2\right\} \leq 1$.
- Channel time variation: T_{coh} coherence time

  - *slow fading*: T_{coh} is so long that coding is performed over a single channel realization.
  - *fast fading*: T_{coh} is so short that coding over multiple channel realizations is possible.

Capacity of Deterministic SISO Channel

Definition

The capacity of an AWGN channel is

$$C = \log_2 (1 + SNR) \quad [bits/s/H].$$

1. codeword infinity long (block length)
2. X Gaussian distributed (finite constellation in practise)

Definition

The capacity of a deterministic (time-invariant) SISO channel h is

$$C = \log_2 \left(1 + \rho |h|^2\right) \quad [bits/s/H].$$

Instantaneous SNR is $\rho |h|^2$.

Capacity of Deterministic SISO Channel

- Low SNR: $\log_2(1+x) \approx x \log_2(e)$ for x small

$C \propto SNR$

$$C \approx \rho |h|^2 \log_2(e).$$

  - C grows linearly with SNR at low SNR.
  - Double SNR and C is doubled.

- High SNR: $\log_2(1+x) \approx \log_2(x)$ for x large

$C \propto \log(SNR)$

$$C \approx = \log_2\left(\rho |h|^2\right) = \log_2(\rho) + \log_2\left(|h|^2\right).$$

  - C grows logarithmically with SNR (or linearly with SNR in dB) at high SNR.
  - Double SNR (or 3dB increase) and C is increased by 1 bits/s/Hz.

Capacity of Deterministic SIMO Channel

- System model

$$\mathbf{y}_k = \sqrt{E_s}\mathbf{h}_k c_k + n_k$$

  - $\mathbf{y}_k$, $\mathbf{h}_k$ n_r-dimensional vectors

- MRC combining

$$\mathbf{h}_k^H \mathbf{y}_k = \sqrt{E_s}\mathbf{h}_k^H \mathbf{h}_k c_k + \mathbf{h}_k^H n_k$$

Instantaneous SNR $\rho \|\mathbf{h}_k\|^2$.

assume h is constant over time

Definition

The capacity of a deterministic (time-invariant) SIMO channel $\mathbf{h}$ is

$$C = \log_2\left(1 + \rho \|\mathbf{h}\|^2\right) \quad [bits/s/H].$$

Capacity of Deterministic MISO Channel

- System model

$$y_k = \sqrt{E_s}\mathbf{h}_k\mathbf{c}_k + n_k$$

  - $\mathbf{h}_k$, $\mathbf{c}_k$ n_t-dimensional vectors

- MRT $\mathbf{c}_k = \mathbf{w}_k c_k$ with $\mathbf{w}_k = \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}$

$$y_k = \sqrt{E_s}\mathbf{h}_k\mathbf{w}_k c_k + n_k$$

  Instantaneous SNR $\rho\|\mathbf{h}_k\|^2$.

Definition

The capacity of a deterministic (time-invariant) MISO channel $\mathbf{h}$ is

$$C = \log_2\left(1 + \rho\|\mathbf{h}\|^2\right) \quad [bits/s/H]$$

With perfect transmit channel knowledge!

Ergodic Capacity of Fast Fading Channels

block length longer than coherence time.

- Fast fading:
  - Doppler frequency sufficiently high to allow for coding over many channel realizations/coherence time periods
  - The transmission capability is represented by a single quantity known as the ergodic capacity

- Rate of information flow between Tx and Rx at time instant k over channels h_k

$$\log_2\left(1+\rho|h_k|^2\right).$$

Such a rate varies over time according to the channel fluctuations. The average rate of information flow over a time duration $T >> T_{coh}$ is

$T >> T_{coh}$: fast fading

$\downarrow$

ergodic capacity

$$\frac{1}{T}\sum_{k=0}^{T-1}\log_2\left(1+\rho|h_k|^2\right).$$

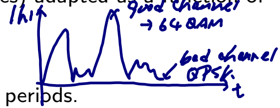
Definition

The ergodic capacity of fast-fading channel is given by

$$\bar{C} = \mathcal{E}\left\{\log_2\left(1+\rho|h|^2\right)\right\}.$$

Impact on Coding Strategy

- Perfect Transmit Channel Knowledge
  - Use a variable-rate code (family of codes of different rates) adapted as a function of the channel state.
  - Code for state h achieves the capacity $\log_2\left(1+\rho|h|^2\right)$.
  - Average throughput $\mathcal{E}\left\{\log_2\left(1+\rho|h|^2\right)\right\}$.
  - No need for the codeword to span many coherence time periods.
  - Code average out the effect of noise



- Partial Transmit Channel Knowledge (only channel distribution known)
  - Same average throughput $\mathcal{E}\left\{\log_2\left(1+\rho|h|^2\right)\right\}$ achievable.
  - Encoding requires a fixed-rate code (whose rate is given by the ergodic capacity) with encoding spanning many channel realizations.
  - Code length large enough $T >> T_c$ to average out both the noise and channel fluctuations.

Ergodic Capacity of Fast Fading Channels

- Low SNR:

Low SNR → no fading

$$\bar{C} \approx \rho \mathcal{E}\{|h|^2\} \log_2(e) = \rho \log_2(e) = C_{AWGN}$$

Same as AWGN

- High SNR:

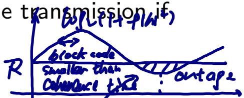
$$\bar{C} \approx \log_2(\rho) + \mathcal{E}\left\{\log_2\left(|h|^2\right)\right\} = \log_2(\rho) - 0.83 = C_{AWGN} - 0.83$$

Fading detrimental: 2.5 dB more power needed in the fading case to achieve the same capacity as in AWGN.

Outage Capacity and Probability in Slow Fading Channels

- In slow fading, the encoding still averages out the randomness of the noise but cannot fully average out the randomness of the channel.
- For a given channel realization h and a target rate R, reliable transmission if

$$\log_2\left(1+\rho|h|^2\right) > R$$



If not met, an outage occurs and the decoding error probability is strictly non-zero.

- Look at the tail probability of $\log_2\left(1+\rho|h|^2\right)$, not its average!

rate vs reliability

Definition

The outage probability $P_{out}\left(R\right)$ of a wireless channel with a target rate R is given by

$$P_{out}\left(R\right) = P\left(\log_2\left(1+\rho|h|^2\right) < R\right).$$

- More meaningful in the absence of CSI knowledge at the transmitter: the transmitter cannot adjust its transmit strategy $\rightarrow$ hopes the channel is good enough

Diversity-Multiplexing Trade-Off in Slow Fading Channels

- For a given R, how does P_{out} behave as a function of the SNR ρ?

Definition

A diversity gain $g_d^\star(g_s, \infty)$ is achieved at multiplexing gain g_s *at infinite SNR* if

multiplexing gain: $R \sim \rho$

$R = g_s \log_2 \rho + c$

$g_s \in [0, 1]$

$\lim_{\rho \to \infty} \frac{R}{\log \rho} = g_s + \frac{c}{\log_2 \rho}$, with $\frac{c}{\log_2 \rho} \to 0$

$$\lim_{\rho \longrightarrow \infty} \frac{R(\rho)}{\log_2(\rho)} = g_s$$

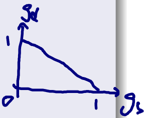
$$\lim_{\rho \longrightarrow \infty} \frac{\log_2\left(P_{out}(R)\right)}{\log_2(\rho)} = -g_d^\star(g_s, \infty)$$

$P_{out}(R) = P(\log_2(1+\rho|h|^2) < R)$

$= P(\log_2(1+\rho|h|^2) < g_s \log_2 \rho)$

$= P(1+\rho|h|^2 < \rho^{g_s})$

$\approx P(|h|^2 < \rho^{g_s - 1})$



The curve $g_d^\star(g_s, \infty)$ as function of g_s is known as the asymptotic diversity-multiplexing trade-off of the channel.

- The multiplexing gain indicates how fast the transmission rate increases with the SNR.
- The diversity gain represents how fast the outage probability decays with the SNR.

Diversity-Multiplexing Trade-Off in Slow Fading Channels

$R = g_s \log_2 \rho$

- Determine for a transmission rate R scaling with ρ as $g_s \log_2(\rho)$, the rate at which the outage probability decreases with ρ as ρ increases.
- Outage probability

$$P_{\text{out}} = P\left(\log_2\left[1+\rho\,|h|^2\right] < g_s \log_2(\rho)\right)$$
$$= P\left(1+\rho\,|h|^2 < \rho^{g_s}\right)$$

- At high SNR,

$$P_{\text{out}} \approx P\left(|h|^2 \le \rho^{-(1-g_s)}\right)$$

- Since $|h|^2$ is exponentially distributed, i.e., $P\left(|h|^2 \le \epsilon\right) \approx \epsilon$ for small ϵ

trade off : $\begin{cases} P_{out} \\ \text{rate} \end{cases}$

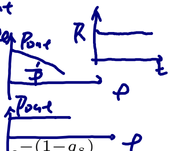
(fix rate) $g_s = 0 \Rightarrow P_{out} \approx \rho^{-1}$

$0 \le g_s \le 1$

$g_s = 1$ (max rate) $\Rightarrow P_{out} = \text{const}$

$$P_{\text{out}} \approx \rho^{-(1-g_s)}$$

An outage occurs at high SNR when $|h|^2 \le \rho^{-(1-g_s)}$ with a probability $\rho^{-(1-g_s)}$.



- DMT for the scalar Rayleigh fading channel $g_d^\star(g_s, \infty) = 1 - g_s$ for $g_s \in [0,1]$.

Impact on Coding Strategy

- $g_s = 0$ (constant rate): $P_{\text{out}} \approx 1/\rho$ $\quad$ R vs. Pout
  - Error probability of uncoded transmission also decays as $1/\rho$.
  - Coding cannot further improve the slope in slow fading.
  - Coding can average out the noise but cannot average out the channel fade.

- $g_s = 1$ (rate scales with SNR): $P_{\text{out}} \approx$ constant
  - P_{out} does not decrease with SNR.

Careful

- AWGN: Send data at rate $R < C$ and error probability as small as desired.
- Slow fading: Cannot be done if the probability of having a deep fade is non-zero.
  - Impossible to code over a large number of independent channel realizations $\rightarrow$ outage as soon as the channel is in deep fade.
  - Capacity of the slow fading channel in the strict sense is zero.
  - ϵ-outage capacity C_ϵ: largest rate of transmission R such that $P_{\text{out}}(R) \leq \epsilon$.

Outage Capacity and Probability in MISO/SIMO Channels

- $P_{out}(R)$ of a SIMO wireless channel with a target rate R

$$P_{out}(R) = P\left(\log_2\left(1+\rho\|\mathbf{h}\|^2\right) < R\right).$$

  - Use receive diversity based on MRC.
  - $g_s = 0$ (constant rate): $P_{\text{out}} \approx 1/\rho^{n_r}$.

- $P_{out}(R)$ of a MISO wireless channel with a target rate R

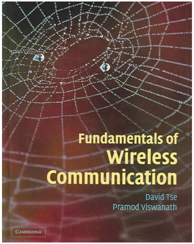
$$P_{out}(R) = P\left(\log_2\left(1+\frac{\rho}{n_t}\|\mathbf{h}\|^2\right) < R\right).$$

  - Use transmit diversity based on Alamouti for 2Tx.
  - $g_s = 0$ (constant rate): $P_{\text{out}} \approx 1/\rho^{n_t}$.

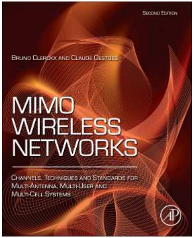
Part 2: MIMO Systems

Basics of MIMO
The MIMO Channel
Capacity of MIMO Channels
Tx and Rx Strategies: Space-Time Coding and Processing
Quantized Precoding

Reference Book



Chapter 1 - Sections 1.2.4, 1.3.2, 1.6
Chapter 2 - Sections 2.2.1, 2.3.1
Chapter 3 - Sections 3.2.1, 3.4.1
Chapter 5 - Sections 5.2, 5.3, 5.4.2, 5.5.1, 5.7, 5.8.1
Chapter 6 - Sections 6.1, 6.2, 6.3.1, 6.5.2, 6.5.4, 6.5.8



Basics of MIMO

Previous Lectures

- Discrete Time Representation
  - SISO: $y = \sqrt{E_s} hc + n$
  - SIMO: $\mathbf{y} = \sqrt{E_s}\mathbf{h}c + \mathbf{n}$
  - MISO (with perfect CSIT): $y = \sqrt{E_s}\mathbf{hw}c + n$
- h is fading
  - amplitude Rayleigh distributed
  - phase uniformly distributed
- Diversity
  - Diversity gain: $g_d^o(\rho) \triangleq -\frac{\log_2(\bar{P})}{\log_2(\rho)}$
  - Array gain: $g_a \triangleq \frac{\bar{\rho}_{out}}{\bar{\rho}} = \frac{\bar{\rho}_{out}}{\rho}$
- SIMO
  - selection combining
  - gain combining
- MISO
  - with perfect channel knowledge at Tx: Matched Beamforming
  - without channel knowledge at Tx: Space-Time Coding (Alamouti Scheme), indirect (time, frequency) transmit diversity

MIMO Systems

- In MIMO systems, the fading channel between each transmit-receive antenna pair can be modeled as a SISO channel.
- For uni-polarized antennas and small inter-element spacings (of the order of the wavelength), path loss and shadowing of all SISO channels are identical.

$$H \begin{array}{c} n_T \\ \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \end{array} \} n_R$$

- Stacking all inputs and outputs in vectors $\mathbf{c}_k = [c_{1,k}, \ldots, c_{n_t,k}]^T$ and $\mathbf{y}_k = [y_{1,k}, \ldots, y_{n_r,k}]^T$, the input-output relationship at any given time instant k reads as

$$\mathbf{y}_k = \sqrt{E_s}\mathbf{H}_k\mathbf{c}'_k + \mathbf{n}_k,$$

Tx		Rx		
1	h	1	SISO	
|||	h^T	1	MISO	$y = h^T c + n$
1	h	|||	SIMO	$y = hc + n$
|||	H	|||	MIMO	< perfect CSIT / no CSIT

  where
  - $\mathbf{c}'_k$ is a precoded version of $\mathbf{c}_k$ that depends on the channel knowledge at the Tx.
  - $\mathbf{H}_k$ is defined as the $n_r \times n_t$ MIMO channel matrix, $\mathbf{H}_k(n, m) = h_{nm,k}$, with h_{nm} denoting the narrowband channel between transmit antenna m ($m = 1, \ldots, n_t$) and receive antenna n ($n = 1, \ldots, n_r$),
  - $\mathbf{n}_k = [n_{1,k}, \ldots, n_{n_r,k}]^T$ is the sampled noise vector, containing the noise contribution at each receive antenna, such that the noise is white in both time and spatial dimensions, $\mathcal{E}\{\mathbf{n}_k\mathbf{n}_l^H\} = \sigma_n^2\mathbf{I}_{n_r}\delta(k-l)$.
- Using the same channels normalization as for SISO channels, $\boxed{\mathcal{E}\{\|\mathbf{H}\|_F^2\} = n_t n_r}$.
- when Tx has a *perfect channel knowledge*: (dominant and multiple) eigenmode transmission
- when Tx has *no knowledge of the channel*: space-time coding (with $\mathbf{c}'_k = \mathbf{c}_k$)

Reminder: Linear Algebra

- *Vector Orthogonality*: $\mathbf{a}^H\mathbf{b} = 0$ (^{H} stands for Hermitian, i.e. conjugate transpose)
- *Hermitian matrix*: $\mathbf{A} = \mathbf{A}^H$
- *Unitary matrix*: $\mathbf{A}^H\mathbf{A} = \mathbf{I}$
- *Singular Value Decomposition* (SVD) of a matrix $\mathbf{H}$ $[n_r \times n_t]$: $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$
  - $\mathbf{U}$ $[n_r \times r(\mathbf{H})]$: unitary matrix of left singular vectors
  - $\mathbf{\Sigma} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}$: diagonal matrix containing the singular values of $\mathbf{H}$
  - $\mathbf{V}$ $[n_t \times r(\mathbf{H})]$: unitary matrix of left singular vectors
  - $r(\mathbf{H})$: the rank of $\mathbf{H}$

  We will often look at Hermitian matrices of the form $\mathbf{A} = \mathbf{H}^H\mathbf{H}$ whose *Eigenvalue Value Decomposition* (EVD) writes as $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ with $\mathbf{\Lambda} = \mathbf{\Sigma}^2$.
- $\mathbf{A} = \mathbf{H}^H\mathbf{H}$ is a positive-semidefinite matrix (≥ 0), i.e. all eigenvalues of $\mathbf{A}$ are nonnegative.
- *Trace* of a matrix $\mathbf{A}$: $\text{Tr}\{\mathbf{A}\} = \sum_i \mathbf{A}(i,i)$.
- *Frobenius norm* of a matrix $\mathbf{A}$: $\|\mathbf{A}\|_F^2 = \sum_i \sum_j |A(i,j)|^2$
- $\|\mathbf{A}\|_F^2 = \text{Tr}\left\{\mathbf{A}\mathbf{A}^H\right\} = \text{Tr}\left\{\mathbf{A}^H\mathbf{A}\right\}$
- $\text{Tr}\{\mathbf{AB}\} = \text{Tr}\{\mathbf{BA}\}$
- $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$

Reminder: Linear Algebra and Matrix Properties

Alamouti 2x2

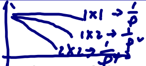
$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix} = \begin{array}{c|cc} & y_1 & y_2 \\ \hline Rx_1 & c_1h_{11}+c_2h_{12} & -c_2^*h_{11}+c_1^*h_{12} \\ Rx_2 & c_1h_{21}+c_2h_{22} & -c_2^*h_{21}+c_1^*h_{22} \end{array}$$

$$y_1 = \begin{bmatrix} c_1h_{11}+c_2h_{12} \\ c_1h_{21}+c_2h_{22} \end{bmatrix} + n_1 .$$

$$y_2 = \begin{bmatrix} -c_2^*h_{11}+c_1^*h_{12} \\ -c_2^*h_{21}+c_1^*h_{22} \end{bmatrix} + n_2$$

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \underbrace{\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{12}^* & -h_{11}^* \\ h_{22}^* & -h_{21}^* \end{bmatrix}}_{H_{eff}} \sqrt{\tfrac{E_s}{2}} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1h_{11}+c_2h_{12} \\ c_1h_{21}+c_2h_{22} \\ c_1h_{12}^*-c_2h_{11}^* \\ c_1h_{22}^*-c_2h_{21}^* \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}$$

- Kronecker product: $\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}(1,1)\mathbf{B} & \dots & \mathbf{A}(1,n)\mathbf{B} \\ \vdots & \dots & \vdots \\ \mathbf{A}(m,1)\mathbf{B} & \dots & \mathbf{A}(m,n)\mathbf{B} \end{bmatrix}$
- $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$
- $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$
- $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD})$
- $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$ if $\mathbf{A}, \mathbf{B}$ square and non singular.
- $\det(\mathbf{A}_{m\times m} \otimes \mathbf{B}_{n\times n}) = \det(\mathbf{A})^n \det(\mathbf{B})^m$
- $\text{Tr}\{\mathbf{A} \otimes \mathbf{B}\} = \text{Tr}\{\mathbf{A}\}\text{Tr}\{\mathbf{B}\}$
- $\text{vec}(\mathbf{A})$ converts $[m \times n]$ matrix into $mn \times 1$ vector by stacking the columns of $\mathbf{A}$ on top of one another.
  - $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\,\text{vec}(\mathbf{B})$
- $\text{Tr}\{\mathbf{ABB}^H\mathbf{A}^H\} = \text{vec}(\mathbf{A}^H)^H (\mathbf{I} \otimes \mathbf{BB}^H)\,\text{vec}(\mathbf{A}^H)$
- $\det(\mathbf{I} + \epsilon\mathbf{A}) = 1 + \epsilon\text{Tr}\{\mathbf{A}\}$ if $\epsilon << 1$



div. gain: 2
4 deg. of freedom

- MIMO without Transmit Channel Knowledge
- Array/diversity/coding gains are exploitable in SIMO, MISO and … MIMO
- *Alamouti scheme* can easily be applied to 2×2 MIMO channels

Tx Rx t_1 t_2 h_1 h_2 $\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$ Rx $\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$

$h_L^H h_L = \begin{bmatrix} |h_1|^2+|h_2|^2 & 0 \\ 0 & |h_1|^2+|h_2|^2 \end{bmatrix}$

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

div. gain = 4
8 deg of freedom (2 Tx . 2 Rx . 2 space)

$y_1 = H\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + n_1$

- Received signal vector (make sure the channel remains constant over two symbol periods!)

$y_2 = H\begin{bmatrix} -c_2^* \\ c_1^* \end{bmatrix} + n_2$

$H_{eff}^H H_{eff} = \begin{bmatrix} |h_{11}|^2+|h_{12}|^2+|h_{21}|^2+|h_{22}|^2 & 0 \\ 0 & |h_{11}|^2+|h_{12}|^2+|h_{21}|^2+|h_{22}|^2 \end{bmatrix}$ $= \|H\|^2$

H remains constant over 2 consecutive slots

$$\mathbf{y}_1 = \sqrt{E_s}\mathbf{H}\begin{bmatrix} c_1/\sqrt{2} \\ c_2/\sqrt{2} \end{bmatrix} + \mathbf{n}_1, \qquad \text{(first symbol period)}$$

$$\mathbf{y}_2 = \sqrt{E_s}\mathbf{H}\begin{bmatrix} -c_2^*/\sqrt{2} \\ c_1^*/\sqrt{2} \end{bmatrix} + \mathbf{n}_2. \qquad \text{(second symbol period)}$$

(if channels change over 2 slots, no orthogonality)

- Equivalently

$\underbrace{H_{eff}^H}_{2\times 4} \cdot \underbrace{y}_{4\times 1} \to 2\times 1$

$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{12}^* & -h_{11}^* \\ h_{22}^* & -h_{21}^* \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2^* \end{bmatrix} = \sqrt{E_s}\underbrace{\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{12}^* & -h_{11}^* \\ h_{22}^* & -h_{21}^* \end{bmatrix}}_{\mathbf{H}_{eff}} \underbrace{\begin{bmatrix} c_1/\sqrt{2} \\ c_2/\sqrt{2} \end{bmatrix}}_{\mathbf{c}} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2^* \end{bmatrix}.$$

div. gain = # independent path
= # Tx X # Rx

Space-Time Coding

decode: H_{eff}^H

- Apply the matched filter $\mathbf{H}_{eff}^H$ to $\mathbf{y}$ ($\mathbf{H}_{eff}^H\mathbf{H}_{eff} = \|\mathbf{H}\|_F^2\,\mathbf{I}_2$)

$V^H W$

$M_T \times M_T$ $\quad M_T \times 1$

$\|H\|_F^2 = |h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2$

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \sqrt{E_s}\,\mathbf{H}_{eff}^H\mathbf{y} = \sqrt{E_s}\,\|\mathbf{H}\|_F^2\,\mathbf{I}_2\,\mathbf{c} + \mathbf{n}'$$

$$\begin{bmatrix} v_1^H \\ v_2^H \\ \vdots \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} v_1^H w_1 \\ v_2^H w_2 \\ \vdots \end{bmatrix}$$

where $\mathbf{n}'$ is such that $\mathcal{E}\{\mathbf{n}'\} = \mathbf{0}_{2\times 1}$ and $\mathcal{E}\{\mathbf{n}'\mathbf{n}'^H\} = \|\mathbf{H}\|_F^2\,\sigma_n^2\mathbf{I}_2$.

$c = \begin{pmatrix} c_1/\sqrt{2} \\ c_2/\sqrt{2} \end{pmatrix}$

w^H $\quad v$

$1 \times M_T$ $\quad M_T \times M_T$

- Average output SNR

$$\bar{\rho}_{out} = \frac{E_s}{\sigma_n^2}\,\mathcal{E}\left\{\frac{\left[\,\|\mathbf{H}\|_F^2\,\right]^2}{2\,\|\mathbf{H}\|_F^2}\right\} = 2\rho.$$

$[w_1^H\ w_2^H \cdots][v_1\ v_2 \cdots]$

$= [w_1^H v_1\ \ w_1^H v_2 \cdots]$

Receive array gain ($g_a = n_r = 2$) but no transmit array gain!

- Average symbol error rate

$$\bar{P} \le \bar{N}_e\left(\frac{\rho d_{min}^2}{8}\right)^{-4}.$$

Full diversity ($g_d^o = n_t n_r = 4$)

Dominant Eigenmode Transmission

- MIMO with Perfect Transmit Channel Knowledge
- Extension of Matched Beamforming to MIMO

$$\mathbf{y} = \sqrt{E_s}\mathbf{H}\mathbf{c}' + \mathbf{n} = \sqrt{E_s}\mathbf{H}\mathbf{w}c + \mathbf{n},$$

$$z = \mathbf{g}\mathbf{y} = \sqrt{E_s}\mathbf{g}\mathbf{H}\mathbf{w}c + \mathbf{g}\mathbf{n}.$$

MRC combining $g = h^H = (Hw)^H = w^H H^H V_1^H$

$SNR_{out} = \rho\|h\|^2 = \rho\|Hw\|^2$

$\max_w \|Hw\|^2$ for any H

dot product

$y = H \cdot w \, c + n$, $[M_R\times1]\ [M_R\times M_T]\ [M_T\times1]$

$M_R = 1$, $w = \frac{h^H}{\|h\|}$

$M_T = 1$, $gy = ghc + gn$, $g = h^H$, $SNR_{out} = \rho\|h\|^2$

$w^H V \Lambda V^H w = \sum_{i=1}^{R(H)} (v_i^H w)^2 \sigma_i^2 \le \sigma_1^2$

$= 1$ if aligned

$\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{R(H)}^2$

- Decompose

$$\mathbf{H} = \mathbf{U_H}\mathbf{\Sigma_H}\mathbf{V_H}^H,$$

$$\mathbf{\Sigma_H} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}.$$

precode w with dominant eigenvalue v

$w = v_{max}$ (beamforming) choose the strongest

- Received SNR is maximized by matched filtering, leading to

$$\boxed{\begin{aligned}\mathbf{w} &= \mathbf{v}_{max}\\ \mathbf{g} &= \mathbf{u}_{max}^H\end{aligned}}$$

transmit c along direction of strongest path

decompose → virtual path

$H = U\Sigma V^H$, $H^H = V\Sigma^H U^H$

$g = w^H H^H = \alpha V_1^H$

where $\mathbf{v}_{max}$ and $\mathbf{u}_{max}$ are respectively the right and left singular vectors corresponding to the maximum singular value of $\mathbf{H}$, $\sigma_{max} = \max\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}$. Note the generalization of matched beamforming (MISO) and MRC (SIMO)!

- Equivalent channel: $z = \sqrt{E_s}\sigma_{max}c + \tilde{n}$ where $\tilde{n} = \mathbf{g}\mathbf{n}$ has a variance equal to σ_n^2.

Dominant Eigenmode Transmission

- Array gain: $\mathcal{E}\{\sigma_{max}^2\} = \mathcal{E}\{\lambda_{max}\}$ where λ_{max} is the largest eigenvalue of $\mathbf{H}\mathbf{H}^H$. Commonly, $\max\{n_t, n_r\} \leq g_a \leq n_t n_r$.

- Diversity gain: the dominant eigenmode transmission extracts a full diversity gain of $n_t n_r$ in i.i.d. Rayleigh channels.

Dominant Eigenmode Transmission

Example

Question: Show that the optimum (in the sense of SNR maximization) transmit precoder and combiner in dominant eigenmode transmission is given by the dominant right and left singular vector of the channel matrix, respectively.
Answer: Let us write

$$\mathbf{y} = \sqrt{E_s}\mathbf{H}\mathbf{c}' + \mathbf{n} = \sqrt{E_s}\mathbf{H}\mathbf{w}c + \mathbf{n},$$
$$z = \mathbf{g}\mathbf{y} = \sqrt{E_s}\mathbf{g}\mathbf{H}\mathbf{w}c + \mathbf{g}\mathbf{n}.$$

where $\|\mathbf{w}\|^2 = 1$ (power constraint). We decompose

$$\mathbf{H} = \mathbf{U_H}\mathbf{\Sigma_H}\mathbf{V_H}^H, \qquad \mathbf{\Sigma_H} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}.$$

In order to maximize the SNR, we choose $\mathbf{g}$ as a matched filter, i.e. $\mathbf{g} = (\mathbf{H}\mathbf{w})^H$ such that

(dot product ≤ 1

$$\mathbf{g}\mathbf{H}\mathbf{w} = \mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w} = \mathbf{w}^H\mathbf{V_H}\mathbf{\Sigma}^2_{\mathbf{H}}\mathbf{V}^H_{\mathbf{H}}\mathbf{w} = \sum_{i=1}^{r(\mathbf{H})}\sigma_i^2\left|\mathbf{v}_i^H\mathbf{w}\right|^2 \le \sigma^2_{max}$$

=1 when w = v_max (aligned)

where $\mathbf{v}_i$ is the i column of $\mathbf{V_H}$ and $\sigma_{max} = \max_{i=1,\ldots,r(\mathbf{H})}\sigma_i$.

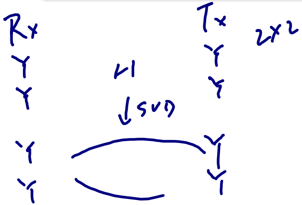
Dominant Eigenmode Transmission

Example

Answer: The inequality is replaced by an equality if $\mathbf{w} = \mathbf{v}_{max}$. By choosing $\mathbf{w} = \mathbf{v}_{max}$,

$$\begin{aligned} \mathbf{g} = \mathbf{w}^H \mathbf{H}^H &= \mathbf{v}_{max}^H \mathbf{V}_\mathbf{H} \mathbf{\Sigma}_\mathbf{H} \mathbf{U}_\mathbf{H}^H \\ &= \sigma_{max} \mathbf{u}_{max}^H \end{aligned}$$

where $\mathbf{u}_{max}$ is the column of $\mathbf{U}_\mathbf{H}$ corresponding to the dominant singular value σ_{max} of $\mathbf{H}$. If we normalize g such that $\|\mathbf{g}\|^2 = 1$, we can write $\mathbf{g} = \mathbf{u}_{max}$. $\square$



Multiple Eigenmode Transmission

$c' = Wc$ DET $\Rightarrow$ $c' = Wc$ MET $Gy = GHWc + Gn$

- Assume $n_r \geq n_t$ an that $r(\mathbf{H}) = n_t$, i.e. n_t singular values in $\mathbf{H}$. Hence, what about spreading symbols over all non-zero eigenmodes of the channel:
  - Tx side: multiply the input vector $\mathbf{c}$ ($n_t \times 1$) using $\mathbf{V_H}$, i.e. $\mathbf{c}' = \mathbf{V_H}\mathbf{c}$
  - Rx side: multiply the received vector $\mathbf{y}$ by $\mathbf{G} = \mathbf{U_H^H}$
  - Overall,

$H = U\Sigma V^H$ $W = V$ $G = U^H$

$Gy = GHW\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + G\begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$

$U^H U \Sigma V^H V \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + V\begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$

$$
\begin{aligned}
\mathbf{z} &= \sqrt{E_s}\mathbf{G}\mathbf{H}\mathbf{c} + \mathbf{G}\mathbf{n} \\
&= \sqrt{E_s}\mathbf{U_H^H}\mathbf{H}\mathbf{V_H}\mathbf{c} + \mathbf{U}^H\mathbf{n} \\
&= \sqrt{E_s}\mathbf{\Sigma_H}\mathbf{c} + \tilde{\mathbf{n}}.
\end{aligned}
$$

$z = \Sigma \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + V\begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$

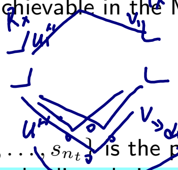
$R(H) = 1$ $R[2\times 2]$ $H = U\Sigma V^H$ $\begin{bmatrix} \sigma & \\ & 0 \end{bmatrix}$ no virtual path

  - The channel has been decomposed into n_t parallel SISO channels given by $\{\sigma_1, \ldots, \sigma_{n_t}\}$.

Rank = # available path = # streams = # orthogonal eigenvector

- The rate achievable in the MIMO channel is the sum of the SISO channel capacities

Tx: V Rx: U^H



$$
R = \sum_{k=1}^{n_t} \log_2(1 + \rho s_k \sigma_k^2),
$$

Choose s_k strength of virture path to max R

fraction of tx power to streams

$R = R_1 + R_2$ $\log(1+SNR_2)$ $\log(1+SNR_1)$ $\frac{P_1}{\sigma_n^2}\sigma_1^2$

where $\{s_1, \ldots, s_{n_t}\}$ is the power allocation on each of the channel eigenmodes.

- The rate scales linearly in n_t.
- In general, the rate scales linearly with the rank of $\mathbf{H}$.

spacing . reflection ...

but not all paths should be used.

Tx close: Rx1 Rx2 n_1 n_2 n_1 n_2

Multiple Eigenmode Transmission

$R_{MET} = \log_2(1+\frac{P/2}{\sigma_n^2}\sigma_1^2) + \log_2(1+\frac{P/2}{\sigma_n^2}\sigma_2^2)$ $R_{DET} = \log_2(1+\frac{P}{\sigma_n^2}\sigma_1^2)$

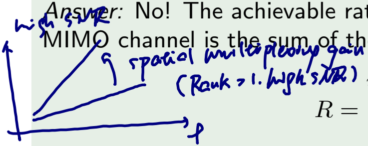
Example

Large SNR: $R_{MET} \sim 2\log(\rho) > R_{DET} \sim \log(\rho)$

Small SNR: $R_{MET} \approx \log_2(1+\frac{P/2}{\sigma_n^2}\sigma_1^2) < R_{DET} = \log_2(1+\frac{P}{\sigma_n^2}\sigma_1^2)$

Question: Is the rate achievable in a MIMO channel with multiple eigenmode transmission and uniform power allocation across modes always larger than that achievable with dominant eigenmode transmission?

Answer: No! The achievable rate with multiple eigenmode transmission in the MIMO channel is the sum of the SISO channel achievable rates



$$R = \sum_{k=1}^{r(\mathbf{H})} \log_2(1 + \rho s_k \sigma_k^2),$$

DET: $w = v_{max}$, $g = u_{max}^H$

MET: $w = V$, $g = U^H$

where $\{s_1, \ldots, s_{r(\mathbf{H})}\}$ is the power allocation on each of the channel eigenmodes.

Large SNR: $R_{MET} > R_{DET}$

Small SNR: $R_{MET} < R_{DET}$

Δ transmit multiple streams is not always optimal.

Two strategies (for a total power constraint $\sum_{k=1}^{r(\mathbf{H})} s_k = 1$):

- Uniform power allocation: $R_u = \sum_{k=1}^{r(\mathbf{H})} \log_2(1 + \rho 1/r(\mathbf{H})\sigma_k^2)$ CSI
- Dominant eigenmode transmission: $R_d = \log_2(1 + \rho\sigma_{max}^2)$

R_u could be either greater or smaller than R_d. For instance, if $\sigma_1 >> 0$ and $\sigma_k \approx \epsilon$ for $k > 1$, $R_u \approx \log_2(1 + \rho\sigma_1^2/r(\mathbf{H})) \le R_d$ for small values of ρ. At very high SNR, despite the little contributions of $\sigma_k \approx \epsilon$, R_u will become higher than R_d. □

Multiplexing gain

- Array/diversity/coding gains are exploitable in SIMO, MISO and MIMO but MIMO can offer much more than MISO and SIMO.

array gain } MISO
diversity gain } SIMO

- MIMO channels offer *multiplexing gain*: measure of the number of independent streams that can be transmitted in parallel in the MIMO channel. Defined as

multiplexing gain MIMO

g_s : # interference-free stream

$$g_s \triangleq \lim_{\rho \to \infty} \frac{R(\rho)}{\log_2(\rho)}$$

$R(\rho)$ → achievable data rate $= \min(n_r, n_t)$

$\log_2(\rho)$ → SISO data rate

where $R(\rho)$ is the transmission rate.

- The multiplexing gain is the pre-log factor of the rate at high SNR, i.e.

$$R \approx g_s \log_2(\rho)$$

- Modeling only the individual SISO channels from one Tx antenna to one Rx antenna not enough:
  - MIMO performance depends on the channel matrix properties
  - characterize all statistical correlations between all matrix elements necessary!

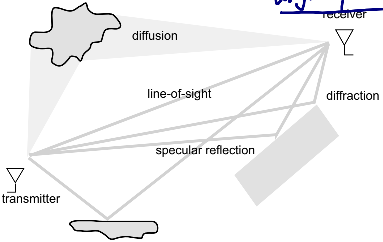
The MIMO Channel

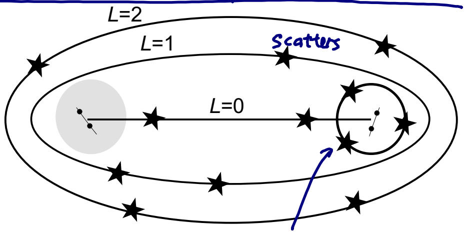
Double-Directional Channel Modeling

- Space comes as an additional dimension
  - *directional*: model the angular distribution of the energy at the antennas
  - *double*: there are multiple antennas at transmit and receive sides
- Spatial distribution of scatterers matters
  - delay-spread $\Longleftrightarrow$ channel frequency selectivity
  - angle-spread $\Longleftrightarrow$ channel spatial selectivity

delay-spread > symbol-spread: f-selecting

angle-spread: a measure of how widespread the signal comes from.





room. many scatters.
energy comes from many directions
→ large angle-spread.

The MIMO Channel Matrix

- $n_r \times n_t$ MIMO channel

$$\mathbf{H}(t,\tau) = \begin{bmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & \dots & h_{1n_t}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & \dots & h_{2n_t}(t,\tau) \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_r1}(t,\tau) & h_{n_r2}(t,\tau) & \dots & h_{n_rn_t}(t,\tau) \end{bmatrix},$$

- Narrowband MIMO channel (the channel is not frequency selective)

\# max streams = rank (H)

$$\mathbf{H}(t) = \begin{bmatrix} h_{11}(t) & h_{12}(t) & \dots & h_{1n_t}(t) \\ h_{21}(t) & h_{22}(t) & \dots & h_{2n_t}(t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_r1}(t) & h_{n_r2}(t) & \dots & h_{n_rn_t}(t) \end{bmatrix},$$

Statistical Properties of the MIMO Channel Matrix

$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$ $\quad H^H = \begin{bmatrix} h_{11}^* & h_{21}^* \\ h_{12}^* & h_{22}^* \end{bmatrix}$ $\quad E(h_i^2) = 1$

how correlation coefs correlated in Tx. Rx.

- Assume narrowband channels, the spatial correlation matrix of the MIMO channel

$$\mathbf{R} = \mathcal{E}\{\text{vec}(\mathbf{H}^H)\text{vec}(\mathbf{H}^H)^H\} = \begin{bmatrix} h_{11}^*h_{11} & h_{11}^*h_{12} & h_{11}^*h_{21} & h_{11}^*h_{22} \\ h_{12}^*h_{11} & h_{12}^*h_{12} & h_{12}^*h_{21} & h_{12}^*h_{22} \\ h_{21}^*h_{11} & h_{21}^*h_{12} & h_{21}^*h_{21} & h_{21}^*h_{22} \\ h_{22}^*h_{11} & h_{22}^*h_{12} & h_{22}^*h_{21} & h_{22}^*h_{22} \end{bmatrix}$$

$\text{vec}(H^H) \cdot \text{vec}(H^H)^H = \begin{bmatrix} h_{11}^* \\ h_{12}^* \\ h_{21}^* \\ h_{22}^* \end{bmatrix} [h_{11}\, h_{12}\, h_{21}\, h_{22}]$

This is a $n_t n_r \times n_t n_r$ positive semi-definite Hermitian matrix.

- It describes the correlation between all pairs of transmit-receive channels:
  - $\mathcal{E}\{\mathbf{H}(n,m)\mathbf{H}^*(n,m)\}$: the average energy of the channel between antenna m and antenna n,
  - $r_m^{(nq)} = \mathcal{E}\{\mathbf{H}(n,m)\mathbf{H}^*(q,m)\}$: the receive correlation between channels originating from transmit antenna m and impinging upon receive antennas n and q,
  - $t_n^{(mp)} = \mathcal{E}\{\mathbf{H}(n,m)\mathbf{H}^*(n,p)\}$: the transmit correlation between channels originating from transmit antennas m and p and arriving at receive antenna n,
  - $\mathcal{E}\{\mathbf{H}(n,m)\mathbf{H}^*(q,p)\}$: the cross-channel correlation between channels (m,n) and (q,p).

h_1 h_2 d. a. $E\{h_1 h_2^*\}$ no d.

Example

2x2 MIMO

$$\mathbf{R} = \begin{bmatrix} 1 & t_1^* & r_1^* & s_1^* \\ t_1 & 1 & s_2^* & r_2^* \\ r_1 & s_2 & 1 & t_2^* \\ s_1 & r_2 & t_2 & 1 \end{bmatrix}$$

$t_1 = h_{12}^* h_{11}$
$r_1 = h_{21}^* h_{11}$
$s_1 = h_{22}^* h_{11}$
$\rightarrow 0$ iff uncorrelated
$\Rightarrow R = I_{n_t n_r}$

$$t_1 = \mathcal{E}\{\mathbf{H}(1,1)\mathbf{H}^*(1,2)\}$$
$$r_1 = \mathcal{E}\{\mathbf{H}(1,1)\mathbf{H}^*(2,1)\}$$

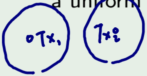
Spatial Correlation

- When the energy spreading is very large at both sides and d_t/d_r are sufficiently large, elements of $\mathbf{H}$ become uncorrelated, and $\mathbf{R}$ becomes diagonal.

Example

Consider two transmit antennas spaced by d_t.

- *isotropic scattering*: (k small) very rich scattering environment around the transmitter with a uniform distribution of the energy.

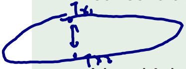


Bessel function

$$\boxed{t = \mathrm{J}_0\left(2\pi \frac{d_t}{\lambda}\right).}$$

The transmit correlation only depends on the spacing between the two antennas.

- *highly directional scattering*: (k large) scatterers around the transmit array are concentrated along a narrow direction $\theta_{t,0}$, i.e., $\mathcal{A}_t(\theta_t) \to \delta(\theta_t - \theta_{t,0})$

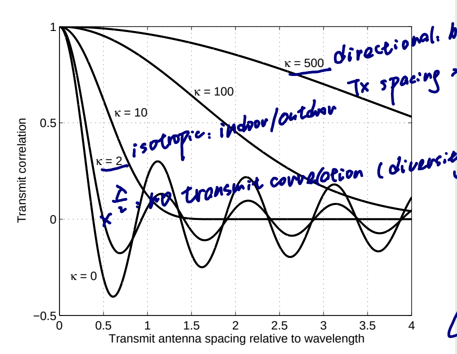


$$\boxed{t \to e^{j\varphi_t(\theta_{t,0})} = e^{j2\pi(d_t/\lambda)\cos\theta_{t,0}}.}$$

Very high transmit correlation approaching one. The scattering direction is directly related to the phase of the transmit correlation.

Spatial Correlation

Example



directional: base station

Tx spacing >> λ

isotropic: indoor/outdoor

$\Sigma\, \kappa^2$: no transmit correlation (diversity)

- $A_t(\theta_t)$ in real-world channels: neither uniform nor a delta.
- isotropic scattering ($\kappa = 0$): first minimum for $d_t = 0.38\lambda$
- directional scattering ($\kappa = \infty$): correlation never reaches 0
- in practice, decorrelation in rich scattering is reached for $d_t \approx 0.5\lambda$
- The more directional the azimuthal dispersion (i.e. for κ increasing), the larger the antenna spacing required to obtain a null correlation.

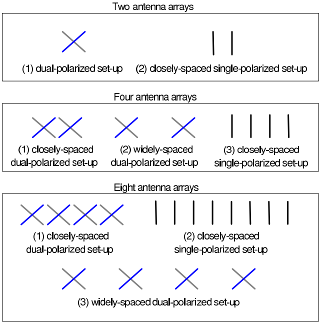
$\kappa \nearrow \Rightarrow d_t \nearrow$

Analytical Representation of Rayleigh MIMO Channels

- Independent and Identically Distributed (I.I.D.) Rayleigh fading
  - $\mathbf{R} = \mathbf{I}_{n_t n_r}$
  - $\mathbf{H} = \mathbf{H}_w$ is a random fading matrix with unit variance and i.i.d. circularly symmetric complex Gaussian entries.
- Realistic in practice only if both conditions are satisfied:
  - the antenna spacings and/or the angle spreads at Tx and Rx are large enough,
  - all individual channels characterized by the same average power (i.e., balanced array).

i.i.d requirements

- What about real-world channels? Sometimes significantly deviate from this ideal channel:
  - *limited angular spread and/or reduced array sizes* cause the channels to becom *correlated* (channels are not independen anymore)
  - a *coherent contribution* may induce the channel statistics to become *Ricean* (channels are not Rayleigh distributed anymore),
  - the use of multiple *polarizations* creates gain imbalances between the various elements of the channel matrix (channel are not identically distributed anymore).



Correlated Rayleigh Fading Channels

$r_1 = r_2 = r$
$t_1 = t_2 = t$
$s_1 = rt$
$s_2 = rt^*$

- For identically distributed Gaussian channels, $\mathbf{R}$ constitutes a sufficient description of the stochastic behavior of the MIMO channel.

Kronecker model: $\boxed{\begin{array}{l} R = R_r \otimes R_t \\ H = R_r^{\frac{1}{2}} H_w R_t^{\frac{1}{2}} \end{array}}$

- Any channel realization is obtained by

$$\text{vec}\big(\mathbf{H}^H\big) = \mathbf{R}^{1/2}\ \text{vec}(\mathbf{H}_w),$$

  where $\mathbf{H}_w$ is one realization of an i.i.d. channel matrix
- Complicated to use because
  - cross-channel correlation not intuitive and not easily tractable
  - Too many parameters: dimensions of $\mathbf{R}$ rapidly become large as the array sizes increase
  - vec operation complicated for performance analysis
- Kronecker model: use a separability assumption

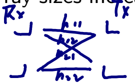
$R_r = \begin{bmatrix} 1 & r^* \\ r & 1 \end{bmatrix}$ $\quad R_t = \begin{bmatrix} 1 & t^* \\ t & 1 \end{bmatrix}$ $\quad R = R_r \otimes R_t = \begin{bmatrix} 1 & t^* & r^* & r^*t^* \\ t & 1 & r^*t & r^* \\ r & rt^* & 1 & t^* \\ rt & r & t & 1 \end{bmatrix}$

$r \cdot t \sim d_r \cdot d_t$

$$\mathbf{R} = \mathbf{R}_r \otimes \mathbf{R}_t,$$

$$\mathbf{H} = \mathbf{R}_r^{1/2}\mathbf{H}_w\mathbf{R}_t^{1/2}$$

$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$



if spacing is small, the correlation is same. $h_{11} \sim h_{22}$. $h_{12} \sim h_{21}$.
$\Rightarrow$ rank collapse.

  where $\mathbf{R}_t$ and $\mathbf{R}_r$ are respectively the transmit and receive correlation matrices.
- Strictly valid only if $r_1 = r_2 = r$ and $t_1 = t_2 = t$ and $s_1 = rt$ and $s_2 = rt^*$ (for 2×2)

$$\mathbf{R} = \begin{bmatrix} 1 & t_1^* & r_1^* & s_1^* \\ t_1 & 1 & s_2^* & r_2^* \\ r_1 & s_2 & 1 & t_2^* \\ s_1 & r_2 & t_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & t^* & r^* & r^*t^* \\ t & 1 & r^*t & r^* \\ r & rt^* & 1 & t^* \\ rt & r & t & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & r^* \\ r & 1 \end{bmatrix}}_{\mathbf{R}_r} \otimes \underbrace{\begin{bmatrix} 1 & t^* \\ t & 1 \end{bmatrix}}_{\mathbf{R}_t}$$

Correlated Rayleigh Fading Channels

Example

Question: Assume a MISO system with two transmit antennas. The channel gains are identically distributed circularly symmetric complex Gaussian but can be correlated and are denoted as h_1 and h_2. Write the expression of the transmit correlation matrix $\mathbf{R}_t$ and derive the eigenvalues and eigenvectors of $\mathbf{R}_t$ as a fonction of the transmit correlation coefficient t.

Answer: We write

$$\mathbf{R}_t = \mathcal{E}\left\{\begin{bmatrix} h_1^* \\ h_2^* \end{bmatrix}\begin{bmatrix} h_1 & h_2 \end{bmatrix}\right\} = \begin{bmatrix} \mathcal{E}\left\{|h_1|^2\right\} & \mathcal{E}\left\{h_1^* h_2\right\} \\ \mathcal{E}\left\{h_1 h_2^*\right\} & \mathcal{E}\left\{|h_2|^2\right\} \end{bmatrix} = \begin{bmatrix} 1 & t^* \\ t & 1 \end{bmatrix}$$

eigenvector $\sim \angle t$

where $t = \mathcal{E}\left\{h_1 h_2^*\right\}$ is the transmit correlation coefficient. The SVD leads to

eigenvalue $\sim |t|$

$$\mathbf{R}_t = \begin{bmatrix} 1 & 1 \\ t/|t| & -t/|t| \end{bmatrix}\begin{bmatrix} 1+|t| & 0 \\ 0 & 1-|t| \end{bmatrix}\begin{bmatrix} 1 & 1 \\ t/|t| & -t/|t| \end{bmatrix}^H.$$

The eigenvalues are only function of the magnitude of t while the eigenvectors are only function of the phase of t.

□

Capacity of MIMO Channels

Previous Lectures

- Transmission strategies
  - Space-Time Coding when no Tx channel knowledge
  - Multiple (including dominant) eigenmode transmission when Tx channel knowledge

$$\begin{aligned}\mathbf{z} &= \sqrt{E_s}\mathbf{G}\mathbf{H}\mathbf{c}' + \mathbf{G}\mathbf{n}\\ &= \sqrt{E_s}\mathbf{U}_{\mathbf{H}}^H\mathbf{H}\mathbf{V}_{\mathbf{H}}\mathbf{c} + \mathbf{U}^H\mathbf{n}\\ &= \sqrt{E_s}\mathbf{\Sigma}_{\mathbf{H}}\mathbf{c} + \tilde{\mathbf{n}}.\end{aligned}$$

  Multiple parallel data pipes $\rightarrow$ Spatial multiplexing gain!

- Performance highly depends on the channel matrix properties
  - Angle spread and inter-element spacing
  - Spatial Correlation: spread antennas far apart to decrease spatial correlation
  - Rayleigh and Ricean distribution

System Model

- A single-user MIMO system with n_t transmit and n_r receive antennas over a frequency flat-fading channel.
- The transmit and received signals in a MIMO channel are related by

$$\mathbf{y}_k = \sqrt{E_s}\mathbf{H}_k\mathbf{c}'_k + \mathbf{n}_k$$

CSCG channel
⇒ Gaussian input
to achieve capacity.

where

  - $\mathbf{y}_k$ is the $n_r \times 1$ received signal vector,
  - $\mathbf{H}_k$ is the $n_r \times n_t$ channel matrix
  - $\mathbf{n}_k$ is a $n_r \times 1$ zero mean complex additive white Gaussian noise (AWGN) vector with $\mathcal{E}\{\mathbf{n}_k\mathbf{n}_l^H\} = \sigma_n^2\mathbf{I}_{n_r}\delta(k-l)$.
  - $\rho = E_s/\sigma_n^2$ represents the SNR.

- The input covariance matrix is defined as the covariance matrix of the transmit signal $\mathbf{c}'$ (we drop the time index) and writes as $\mathbf{Q} = \mathcal{E}\{\mathbf{c}'\mathbf{c}'^H\}$.

$Q = \mathcal{E}\{c'c'^H\}$
$= V_H R_c V_H^H$
$[\ddots]$ power to streams

- Power constraint: $\text{Tr}\{\mathbf{Q}\} \leq 1$. (trace: measure of power)
- Channel time variation: T_{coh} coherence time
  - *slow fading*: T_{coh} is so long that coding is performed over a single channel realization.
  - *fast fading*: T_{coh} is so short that coding over multiple channel realizations is possible.

code over one channel state.

$\text{Tr}(V_H R_c V_H^H) = \text{Tr}(V_H^H V_H R_c)$
$= \text{Tr}(R_c)$

Capacity of Deterministic MIMO Channels

$Q:$ $\begin{cases} \text{trace} \rightarrow \text{power} \\ \text{offdiagonal} \rightarrow \text{correlation} \end{cases}$

Proposition

For a deterministic MIMO channel $\mathbf{H}$, *the mutual information* $\mathcal{I}$ *is written as*

$y = Hc' + n$

$E[yy^H] = HQH^H + R_n$

$$\mathcal{I}(\mathbf{H}, \mathbf{Q}) = \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^H \right]$$

$I(x;y) = H(y) - H(n)$

where $\mathbf{Q}$ *is the input covariance matrix whose trace is normalized to unity.*

Definition

The capacity of a deterministic $n_r \times n_t$ MIMO channel with perfect channel state information at the transmitter is

$$\Delta \qquad \boxed{C(\mathbf{H}) = \max_{\mathbf{Q} \geq 0 : \mathrm{Tr}\{\mathbf{Q}\} = 1} \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^H \right].}$$

Note the difference with SISO capacity.

$C = \log_2(1 + \rho |h|^2)$

Capacity and Water-Filling Algorithm

$Q = E\{c'c'^H\} = V_H R_c V_H^H = [\ddots]\ \text{diag}(s_1 \dots s_n) \quad n = \text{rank}(H_r \cdot H_t)$

- What is the best transmission strategy, i.e. the optimum input covariance matrix $\mathbf{Q}$?
- *First*, create $n = \min\{n_t, n_r\}$ parallel data pipes (Multiple Eigenmode Transmission)
  - Decouple the channel along the individual channel modes (in the directions of the singular vectors of the channel matrix $\mathbf{H}$ at both the transmitter and the receiver)

$H = U_H \Sigma_H V_H^H$

$z = U_H^H y = \underbrace{U_H^H H V_H}_{\Sigma_H} c + n$

$$\mathbf{H} = \mathbf{U_H \Sigma_H V_H^{\mathit{H}}},$$

$$\mathbf{U_H^{\mathit{H}} H V_H} = \mathbf{U_H^{\mathit{H}} U_H \Sigma_H V_H^{\mathit{H}} V_H} = \mathbf{\Sigma_H}$$

$\log_2 \det(I + \rho H Q H^H)$
$= \log_2 \det(I + \rho H V_H \text{diag}(s_1 \dots s_n) V_H^H H^H)$
$= \log_2 \det(I + \rho U_H \Sigma_H \text{diag}(s_1 \dots s_n) V_H^H V_H \Sigma_H U_H^H)$
$= \log_2 \det(I + \rho U_H^H U_H \Sigma_H \text{diag}(s_1 \dots s_n) \Sigma_H)$
$= \log_2 [\Pi(1 + \rho \lambda_k s_k)]$
$= \sum_{k=1}^{n} \log_2(1 + \rho \lambda_k s_k)$

  - Optimum input covariance matrix $\mathbf{Q}^\star$ writes as

$$\mathbf{Q}^\star = \mathbf{V_H}\,\text{diag}\{s_1^\star, \dots, s_n^\star\}\,\mathbf{V_H^{\mathit{H}}}$$

beamforming dir.

- *Second*, allocate power to data pipes
  - $\mathbf{\Sigma_H} = \text{diag}\{\sigma_1, \dots, \sigma_n\}$, and $\sigma_k^2 \triangleq \lambda_k$
  - Capacity: $C(\mathbf{H}) = \max_{\{s_k\}_{k=1}^n} \sum_{k=1}^{n} \log_2[1 + \rho s_k \lambda_k] = \boxed{\sum_{k=1}^{n} \log_2[1 + \rho s_k^\star \lambda_k]}$

Proposition

The power allocation strategy $\{s_1, \dots, s_n\} = \{s_1^\star, \dots, s_n^\star\}$ *that maximizes* $\sum_{k=1}^{n} \log_2(1 + \rho \lambda_k s_k)$ *under the power constraint* $\sum_{k=1}^{n} s_k = 1$, *is given by the water-filling solution,*

$$s_k^\star = \left(\mu - \boxed{\frac{1}{\rho \lambda_k}}\right)^+, \quad k = 1, \dots, n$$

$\log(1+\rho\lambda_1 s_1)$

$\log(1+\rho\lambda_2 s_2)$

$s_1 + s_2 = 1$

where μ *is chosen so as to satisfy the power constraint* $\sum_{k=1}^{n} s_k^\star = 1$.

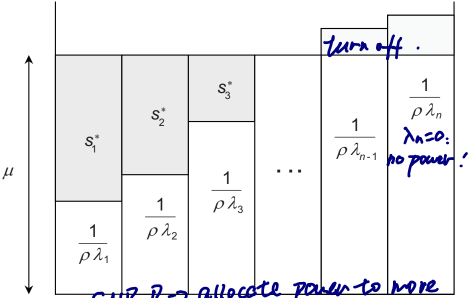
$\lambda_k \uparrow \Rightarrow$ more power.

Water-Filling Algorithm

$s_1 = \mu - \frac{1}{\rho\lambda_1}$. $s_2 = \mu - \frac{1}{\rho\lambda_2}$

$s_1 + s_2 = 1 = 2\mu - \frac{1}{\rho\lambda_1} - \frac{1}{\rho\lambda_2} \Rightarrow \mu = \frac{1}{2}\left(1 + \frac{1}{\rho\lambda_1} + \frac{1}{\rho\lambda_2}\right)$

- Iterative power allocation
  1 – Order eigenvalues λ_k in decreasing order of magnitude
  2 – At iteration i, evaluate the constant μ from the power constraint
  (update μ)
  $$\mu(i) = \frac{1}{n-i+1}\left(1 + \sum_{k=1}^{n-i+1} \frac{1}{\rho\lambda_k}\right)$$
  3 – Calculate power
  (update stream power)
  $$s_k(i) = \mu(i) - \frac{1}{\rho\lambda_k}, \quad k = 1, \ldots, n-i+1.$$
  If $s_{n-i+1} < 0$, set to 0
  4 – Iterate till the power allocated on each mode is non negative.



SNR ↑ ⇒ allocate power to more streams.

SNR → ∞ ⇒ multiplexing gain = # streams.

Water-Filling Algorithm

$$H = \begin{bmatrix} a & 0 & a & 0 \\ 0 & b & 0 & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2}a & 0 \\ 0 & \sqrt{2}b \end{bmatrix} \left[\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^H$$

$$C = \log(1+\rho \cdot 2|a|^2 s_1) + \log(1+\rho \cdot 2|b|^2 s_2)$$

$$P_1 = \left(\mu - \frac{1}{\rho 2|a|^2}\right)^+ \quad P_2 = \left(\mu - \frac{1}{\rho 2|b|^2}\right)^+$$

Example

Question: Consider the transmission $\mathbf{y} = \mathbf{H}\mathbf{c}' + \mathbf{n}$ with perfect CSIT over a deterministic point to point MIMO channel whose matrix is given by

$$\mu = \frac{1}{2} + \frac{1}{2\rho 2|a|^2} + \frac{1}{2\rho 2|b|^2}$$

$$\mathbf{H} = \begin{array}{c} \\ Rx \end{array} \begin{array}{c} Tx \\ \begin{bmatrix} a & 0 & a & 0 \\ 0 & b & 0 & b \end{bmatrix} \end{array}$$

$$s_2 = 0 \text{ if } \frac{1}{2} + \frac{1}{4\rho|a|^2} + \frac{1}{4\rho|b|^2} - \frac{1}{2\rho|b|^2} \le 0$$

where a and b are complex scalars with $|a| \geq |b|$. The input covariance matrix is given by $\mathbf{Q} = \mathcal{E}\left\{\mathbf{c}'\mathbf{c}'^H\right\}$ and is subject to the transmit power constraint $\text{Tr}\left\{\mathbf{Q}\right\} \leq P$.

1. Compute the capacity with perfect CSIT of that deterministic channel. Particularize to the case $a = b$. Explain your reasoning.
2. Explain how to achieve that capacity.
3. In which deployment scenario, could such channel matrix structure be encountered?

Water-Filling Algorithm

Example

Answer:

1. Let us write $\mathbf{Q} = \mathbf{V}\mathbf{P}\mathbf{V}^H$ with the diagonal element of $\mathbf{P}$, denoted as P_k (satisfying $\sum_{k=1}^{n_t} P_k = P$), refers to the power allocated to stream k. The capacity with perfect CSIT over the deterministic channel $\mathbf{H}$ is given by

$$C\left(\mathbf{H}\right) = \max_{P_1,\ldots,P_k} \sum_{k=1}^{\min\{2,4\}} \log_2\left(1 + \frac{P_k}{\sigma_n^2}\lambda_k\right)$$

where λ_k refers the non-zero eigenvalue of $\mathbf{H}^H\mathbf{H}$, respectively equal to $2\left|a\right|^2$ and $2\left|b\right|^2$. Hence,

$$C\left(\mathbf{H}\right) = \max_{P_1,P_2}\left(\log_2\left(1 + \frac{P_1}{\sigma_n^2}2\left|a\right|^2\right) + \log_2\left(1 + \frac{P_2}{\sigma_n^2}2\left|b\right|^2\right)\right).$$

The optimal power allocation is given by the water-filling solution

$$P_1^\star = \left(\mu - \frac{\sigma_n^2}{2\left|a\right|^2}\right)^+, \qquad P_2^\star = \left(\mu - \frac{\sigma_n^2}{2\left|b\right|^2}\right)^+$$

with μ computed such that $P_1^\star + P_2^\star = P$.

Example

Answer:

Assuming $P_1^\star$ and $P_2^\star$ are positive, $\mu = \frac{P}{2} + \frac{\sigma_n^2}{4}\left(\frac{1}{|a|^2} + \frac{1}{|b|^2}\right)$. If $\mu - \frac{\sigma_n^2}{2|b|^2} \leq 0$, i.e. $\frac{P}{2} + \frac{\sigma_n^2}{4|a|^2} - \frac{\sigma_n^2}{4|b|^2} \leq 0$, $P_2^\star = 0$ and $P_1^\star = P$. The capacity writes as

orthogonal transmission in freq. domain:

parallel: no interf. B

f_1 — $\log_2(1+|h_1|^2 s_1)$

f_2 — $\log_2(1+|h_2|^2 s_2)$

f_3 — $\log_2(1+|h_3|^2 s_3)$

all to one:

$$C(\mathbf{H}) = \log_2\left(1 + \frac{P}{\sigma_n^2} 2|a|^2\right).$$

If $\frac{P}{2} + \frac{\sigma_n^2}{4|a|^2} - \frac{\sigma_n^2}{4|b|^2} > 0$, $P_1^\star = \frac{P}{2} - \frac{\sigma_n^2}{4|a|^2} + \frac{\sigma_n^2}{4|b|^2}$ and $P_2^\star = \frac{P}{2} + \frac{\sigma_n^2}{4|a|^2} - \frac{\sigma_n^2}{4|b|^2}$.
The capacity writes as

$\sum_k \log_2(1+|h_k|^2 s_k)$

$\sum_k s_k = P$

$$C(\mathbf{H}) = \log_2\left(1 + \frac{P_1^\star}{\sigma_n^2} 2|a|^2\right) + \log_2\left(1 + \frac{P_2^\star}{\sigma_n^2} 2|b|^2\right).$$

In the particular case where $a = b$, uniform power allocation $P_1^\star = P_2^\star = \frac{P}{2}$ is optimal and

streams in same condition:

$$C(\mathbf{H}) = 2\log_2\left(1 + \frac{P}{\sigma_n^2}|a|^2\right).$$

Water-Filling Algorithm

Example

Answer:

2. Transmit along $\mathbf{V}$, given by the two dominant eigenvector of $\mathbf{H}^H\mathbf{H}$. They are easily computed given the orthogonality of the channel matrix $\mathbf{H}$ as

$\log_2(1+\lambda_1 S_1) + \log_2(1+\lambda_2 S_2)$

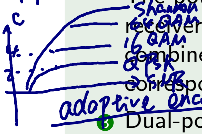
$\lambda_1 > \lambda_2 \Rightarrow$ as SNR $\uparrow$. $S_1 \downarrow . S_2 \uparrow$.

eig. vectors

$$\mathbf{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

digital constellation → Gaussian input

The power allocated to the two streams is given by $P_1^\star$ and $P_2^\star$. At the receiver, the precoded channel is already decoupled and no further combiner is necessary. Each stream can be decoded using the corresponding SISO decoder.



adaptive encoding: approximate Gaussian input when SNR is large. (by denser constellation)

3. Dual-polarized antenna deployment (e.g. VHVH-VH) with LoS and good antenna XPD.

□

Capacity Bounds and Suboptimal Power Allocations

- Low SNR: power allocated to the dominant eigenmode

$$C(\mathbf{H}) \overset{\rho\to 0}{\to} \log_2(1+\rho\lambda_{max}).$$

- High SNR: power is uniformly allocated among the non-zero modes

$$C(\mathbf{H}) \overset{\rho\to\infty}{\to} \sum_{k=1}^{n} \log_2\left(1+\frac{\rho}{n}\lambda_k\right) \cong n\log_2\left(\frac{\rho}{n}\right) + \sum_{k=1}^{n} \log_2(\lambda_k).$$

Observations: $C(\mathbf{H})$ scales linearly with n. The spatial multiplexing gain is $g_s = n$. MISO fading channels do not offer any multiplexing gain. multiplexing gain: $\frac{C_{MIMO}}{C_{SISO}} (\rho\to\infty)$

- At any SNR

$$\boxed{\begin{aligned} C(\mathbf{H}) &\geq \log_2(1+\rho\lambda_{max}), \\ C(\mathbf{H}) &\geq \sum_{k=1}^{n} \log_2\left(1+\frac{\rho}{n}\lambda_k\right). \end{aligned}}$$

Impact on Coding Architecture

- Transmit independent streams in the directions of the eigenvectors of the channel matrix $\mathbf{H}$.

- For a total transmission rate R, each stream k can then be encoded using a capacity-achieving Gaussian code with rate R_k such that $\sum_{k=1}^{n} R_k = R$, ascribed a power λ_k and be decoded independently of the other streams.

- The optimal power allocation based on the water-filling allocation strategy.

Ergodic Capacity of Fast Fading Channels

- Fast fading:
  - Doppler frequency sufficiently high to allow for coding over many channel realizations/coherence time periods
  - The transmission capability is represented by a single quantity known as the ergodic capacity
- With Perfect Transmit Channel Knowledge, similar to deterministic channels
- Focus on Partial Transmit Channel Knowledge

MIMO Capacity with Partial Transmit Channel Knowledge

CDIT

- $\mathbf{H}$ is not known to the transmitter $\rightarrow$ we cannot adapt $\mathbf{Q}$ at all time instants
- Rate of information flow between Tx and Rx at time instant k over channels $\mathbf{H}_k$

$$\log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H}_k \mathbf{Q} \mathbf{H}_k^H \right].$$

$Q = V_H \begin{bmatrix} S_1 & \\ & S_2 \end{bmatrix} V_H^H$

Such a rate varies over time according to the channel fluctuations. The average rate of information flow over a time duration $T >> T_{coh}$ is fix $Q \Rightarrow \log_2(1+\rho|h|^2)$

$$\frac{1}{T} \sum_{k=0}^{T-1} \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H}_k \mathbf{Q} \mathbf{H}_k^H \right].$$

encoding over multiple channel realisations.

i.i.d: a lot scatters

$H = H_w$ (CSCG)

$R = I$

Definition

The ergodic capacity of a $n_r \times n_t$ MIMO channel with channel distribution information at the transmitter (CDIT) is given by

$$\bar{C}_{CDIT} \triangleq \bar{C} = \max_{\mathbf{Q} \geq 0: \text{Tr}\{\mathbf{Q}\}=1} \mathcal{E} \left\{ \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H} \mathbf{Q} \mathbf{H}^H \right] \right\},$$

maximise C for expectation

where $\mathbf{Q}$ is the input covariance matrix optimized as to maximize the ergodic mutual information. $Q : \frac{I}{n_t}$ equal power alloc. (on avg all streams perform similarly) no specific beamforming dir.

- $T >> T_c$ to average out the noise and the channel fluctuations

I.I.D. Rayleigh Fast Fading Channels: Partial Transmit Channel Knowledge

2 channel: $\log_2(1+\rho\lambda_1\frac{1}{2}) + \log_2(1+\rho\lambda_2\frac{1}{2}) = \log_2\det(I + \frac{\rho}{2}HH^H)$

$V_H[\lambda]V_H^H$

- Optimal covariance matrix

$\|H\|_F^2 = \mathrm{Tr}(HH^H) = \mathrm{Tr}(V_H \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} V_H^H) = \lambda_1+\lambda_2$

Proposition

In i.i.d. Rayleigh fading channels, the ergodic capacity with CDIT is achieved under an equal power allocation scheme $\mathbf{Q} = \mathbf{I}_{n_t}/n_t$, *i.e.,*

$C = \log_2(1+\frac{\rho}{a}\lambda_1) + \log_2(1+\frac{\rho}{b}\lambda_2) = \log_2[(1+\frac{\rho}{a}\lambda_1)(1+\frac{\rho}{b}\lambda_2)]$

high SNR $\approx n\log_2\frac{\rho}{n_t} + \mathcal{E}\{\sum_{k=1}^{n}\lambda_k\}$

$$\bar{C}_{CDIT} = \bar{\mathcal{I}}_e = \mathcal{E}\left\{\log_2\det\left[\mathbf{I}_{n_r} + \frac{\rho}{n_t}\mathbf{H}_w\mathbf{H}_w^H\right]\right\} = \mathcal{E}\left\{\sum_{k=1}^{n}\log_2\left[1+\frac{\rho}{n_t}\lambda_k\right]\right\}.$$

$= \log_2(1+\underbrace{\rho(\frac{\lambda_1}{a}+\frac{\lambda_2}{b})}_{\text{low SNR}} + \underbrace{\rho^2\frac{\lambda_1\lambda_2}{ab}}_{\text{high SNR}})$

$H\ (M_r \times M_t)$

rank = multiplexing gain = $\min(M_r, M_t)$ = # singular vectors.

Encoding requires a fixed-rate code (whose rate is given by the ergodic capacity) with encoding spanning many channel realizations.

low SNR: $C \approx \log_2(1+\rho(\frac{\lambda_1}{a}+\frac{\lambda_2}{b}))$ $\Rightarrow C \approx \log_2[\frac{\rho}{2}(\lambda_1+\lambda_2)]$ $\approx \frac{\rho}{2}(\lambda_1+\lambda_2)\log_2 e$

CDIT: $a=b=2$

- Low SNR:

$$\bar{C}_{CDIT} \geq \mathcal{E}\left\{\log_2\left[1+\frac{\rho}{n_t}\|\mathbf{H}_w\|_F^2\right]\right\} \approx \frac{\rho}{n_t}\mathcal{E}\{\|\mathbf{H}_w\|_F^2\}\log_2(e) = n_r\rho\log_2(e)$$

high SNR: $C \approx \log_2(\frac{\rho^2}{ab}\lambda_1\lambda_2)$ $\Rightarrow C \approx 2\log_2(\frac{\rho}{4}\lambda_1\lambda_2)$

CDIT: $a=b=2$ ($Q = \frac{I}{2}$: joint decoding) multiplexing gain = 2 $n_t n_r$

Observations:

- $\bar{C}_{CDIT}$ is only determined by the energy of the channel.
- A MIMO channel only yields a n_r gain over a SISO channel. Increasing the number of transmit antennas is not useful (contrary to perfect CSIT). SIMO and MIMO channels reach the same capacity for a given n_r.

I.I.D. Rayleigh Fast Fading Channels: Partial Transmit Channel Knowledge

- High SNR:

$$\bar{C}_{CDIT} \approx \mathcal{E}\left\{\sum_{k=1}^{n} \log_2\left[\frac{\rho}{n_t}\lambda_k\right]\right\} = n\log_2\left(\frac{\rho}{n_t}\right) + \mathcal{E}\left\{\sum_{k=1}^{n} \log_2(\lambda_k)\right\}$$

Observations:

  - $\bar{C}_{CDIT}$ at high SNR scales linearly with n (by contrast to the low SNR regime).
  - The multiplexing gain g_s is equal to n, similarly to the CSIT case.
  - $\bar{C}_{CDIT}$ and $\bar{C}_{CSIT}$ are not equal: constant gap equal to $n\log_2(n_t/n)$ at high SNR.

- Expressions can be particularized to SISO, SIMO, MISO cases. At high SNR,
  - SISO ($N = n = 1$):

$$\bar{C}_{CDIT} \approx \log_2(\rho) + \mathcal{E}\left\{\log_2\left(|h|^2\right)\right\} = \log_2(\rho) - 0.83 = C_{AWGN} - 0.83$$

  - SIMO ($n_t = n = 1$, $n_r = N$):

$$\bar{C}_{CDIT} \approx \log_2(n_r\rho)$$

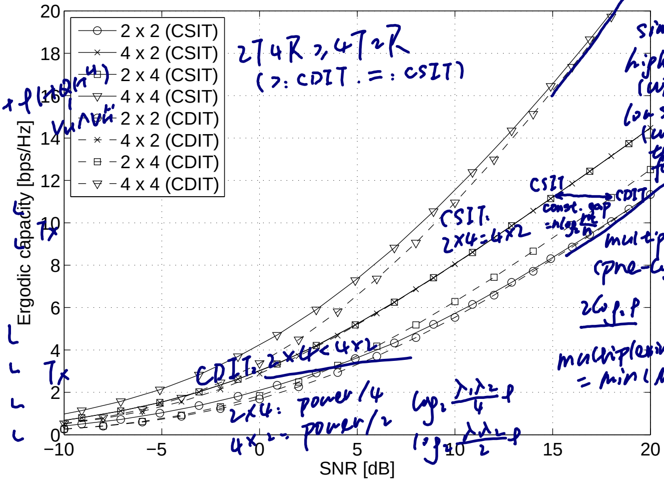
  - MISO ($n_r = n = 1$, $n_t = N$):

$$\bar{C}_{CDIT} \approx \log_2(\rho) + \mathcal{E}\left\{\log_2\left(\|h\|^2/n_t\right)\right\} \overset{n_t\to\infty}{\approx} \log_2(\rho) = C_{AWGN}$$

I.I.D. Rayleigh Fast Fading Channels

- Ergodic capacity of various $n_r \times n_t$ i.i.d. Rayleigh channels with full (CSIT) and partial (CDIT) channel knowledge at the transmitter.

multiplexing gain = 4



$\log_2 \det(I + \frac{\rho}{n_t} H H^H)$

2T4R ≥ 4T2R
(≥: CDIT, =: CSIT)

CSIT: 2x4 = 4x2

CDIT: 2x4 < 4x2

2x4: power/4

4x2: power/2

$\log_2 \frac{\lambda_1 \lambda_2}{4} \rho$

$\log_2 \frac{\lambda_1 \lambda_2}{2} \rho$

Rx 4x2 Tx

Rx 2x4 Tx

similar
high SNR ⇒ coincide.
(WF = uniform)
low SNR ⇒ diff.
(WF ⇒ dominant)
turn off streams
favor 1 dir.

CSIT ⟷ CDIT
const. gap
$= n \log_2 \frac{n_t}{n_r}$

multiplexing gain = 2
(pre-log factor)

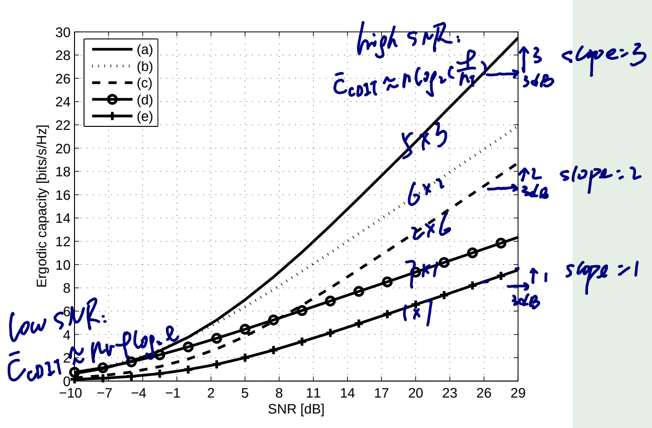
$2\log_2 \rho$

multiplexing gain
$= \min(n_r, n_t)$

I.I.D. Rayleigh Fast Fading Channels

Example

Question: Here is the ergodic capacity of point-to-point i.i.d. Rayleigh fast fading channels with Channel Distribution Information at the Transmitter (CDIT) for five antenna ($n_r \times n_t$) configurations (denoted as (a) to (e)) with $n_t + n_r = 8$.



I.I.D. Rayleigh Fast Fading Channels

Example

Question: What is the achievable (spatial) multiplexing gain (at high SNR) for cases (a), (b), (c), (d) and (e)? Provide your reasoning.

Answer: The multiplexing gain is the pre-log factor of the ergodic capacity at high SNR, i.e. $g_s = \lim_{\rho \to \infty} \frac{\bar{C}_{CDIT}}{\log_2(\rho)}$. Hence by increasing the SNR by 3dB (e.g. from 17dB to 20dB), the ergodic capacity increases by g_s bits/s/Hz.

- (a) $g_s = 3.$
- (b) $g_s = 2.$
- (c) $g_s = 2.$
- (d) $g_s = 1.$
- (e) $g_s = 1.$

Example

Question: For (a), (b), (c), (d) and (e), identify an antenna configuration, i.e. n_t and n_r, satisfying $n_t + n_r = 8$ that achieves such multiplexing gain. Provide your reasoning.

Answer: There are several possible configurations that satisfy to $n_r + n_t = 8$, namely 5×3, 3×5, 6×2, 2×6, 7×1 and 1×7, 4×4. The matching between curves and antenna configurations is easily identified by using the following two arguments: 1) The multiplexing gain with CDIT at high SNR is given by $\min\{n_t, n_r\}$. 2) With CDIT only, the input covariance matrix in i.i.d. channel is $\mathbf{Q} = 1/n_t \mathbf{I}_{n_t}$. This implies that 6×2 and 7×1 outperform 2×6 and 1×7, respectively.

- (a) $n_r \times n_t = 5 \times 3$ or 3×5
- (b) $n_r \times n_t = 6 \times 2$
- (c) $n_r \times n_t = 2 \times 6$
- (d) $n_r \times n_t = 7 \times 1$
- (e) $n_r \times n_t = 1 \times 7$

Correlated Rayleigh Fast Fading Channels: Uniform Power Allocation

$H = R_r^{\frac{1}{2}} H_w R_t^{\frac{1}{2}}$ spatially correlated. (channel not i.i.d.)

- Assume the channel covariance matrix is unknown to the transmitter
- Mutual information with identity input covariance matrix

CSIT: beamform based on correlation.

$\log_2\det(I + \frac{\rho}{2} HH^H)$ $H: 2\times 2$

$$\bar{\mathcal{I}}_e = \mathcal{E}\left\{\log_2 \det\left[\mathbf{I}_{n_r} + \frac{\rho}{n_t}\mathbf{H}\mathbf{H}^H\right]\right\}.$$

high SNR

- Low SNR

$\log_2\det(\frac{\rho}{2}HH^H)$

$\det(R_r^{\frac{1}{2}} H_w R_t^{\frac{1}{2}} R_t^{\frac{H}{2}} H_w^H R_r^{\frac{H}{2}})$

$= \det(H_w H_w^H)\det(R_r)\det(R_t)$

$$\bar{\mathcal{I}}_e \geq \mathcal{E}\left\{\log_2\left[1 + \frac{\rho}{n_t}\|\mathbf{H}\|_F^2\right]\right\}.$$

- High SNR in Kronecker Correlated Rayleigh Channels $\mathbf{H} = \mathbf{R}_r^{1/2}\mathbf{H}_w\mathbf{R}_t^{1/2}$ (with full rank correlation matrices) and $n_t = n_r$

$R_t: \begin{bmatrix} 1 & t^* \\ t & 1 \end{bmatrix}$ $R_r = \begin{bmatrix} 1 & r^* \\ r & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1+|r| & 0 \\ 0 & 1-|r| \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 1+|t| & 0 \\ 0 & 1-|t| \end{bmatrix}$

$$\bar{\mathcal{I}}_e \approx \mathcal{E}\left\{\log_2 \det\left[\frac{\rho}{n_t}\mathbf{H}_w\mathbf{H}_w^H\right]\right\} + \log_2\det(\mathbf{R}_r) + \log_2\det(\mathbf{R}_t).$$

$\det(R_t) = (1+|t|)(1-|t|)$ ($|t| \in [0,1]$)

(max at $t=0$ (no spatial correlation)

min at $|t|=1$ (deficient)

as $|t| \to 1$. correlation ↑ hard to distinguish paths (rank=1.5)

Observations:

- $\det(\mathbf{R}_r) \leq 1$ and $\det(\mathbf{R}_t) \leq 1$: receive and transmit correlations always degrade the mutual information (with power uniform allocation) with respect to the i.i.d. case.
- $\bar{\mathcal{I}}_e$ still scales linearly with $\min\{n_t, n_r\}$

Impact on Coding Architecture

- When the channel is i.i.d. Rayleigh fading, $\mathbf{Q} = (1/n_t)\,\mathbf{I}_{n_t}$.
- Transmission of independent information symbols may be performed in parallel over n virtual spatial channels.
- The transmitter is very similar to the CSIT case except that all eigenmodes now receive the same amount of power.
- Transmit with uniform power allocation over n_t independent streams, each stream using an AWGN capacity-achieving code and perform joint ML decoding (independent decoding of all streams is clearly suboptimal due to interference between streams).

Outage Capacity and Probability in Slow Fading Channels

- In slow fading, the encoding still averages out the randomness of the noise but cannot fully average out the randomness of the channel.
- For a given channel realization $\mathbf{H}$ and a target rate R, reliable transmission if

$$\underbrace{\log_2 \det\left(\mathbf{I}_{n_r} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)}_{\text{upper bound}} > R \text{ rate}$$

If not met with any $\mathbf{Q}$, an outage occurs and the decoding error probability is strictly non-zero.
- Look at the tail probability of $\log_2 \det\left(\mathbf{I}_{n_r} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)$, not its average!

Definition

Δ trade-off : $\{$ rate, Pout

The outage probability $P_{out}(R)$ of a $n_r \times n_t$ MIMO channel with a target rate R is given by

Pout $= \rho^{-r(1-g_s)}$

find Q minimise Pout

$$P_{out}(R) = \min_{\mathbf{Q} \geq 0 : \text{Tr}\{\mathbf{Q}\} \leq 1} P\left(\log_2 \det\left(\mathbf{I}_{n_r} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H\right) < R\right),$$

where $\mathbf{Q}$ is the input covariance matrix optimized as to minimize the outage probability.

- More meaningful in the absence of CSI knowledge at the transmitter: the transmitter cannot adjust its transmit strategy $\rightarrow$ hopes the channel is good enough

Diversity-Multiplexing Trade-Off in Slow Fading Channels

- For a given R, how does P_{out} behave as a function of the SNR ρ?

Definition

A diversity gain $g_d^\star(g_s, \infty)$ is achieved at multiplexing gain g_s *at infinite SNR* if

$R(\rho) = g_s \log_2(\rho)$

$$\lim_{\rho \longrightarrow \infty} \frac{R(\rho)}{\log_2(\rho)} = g_s$$

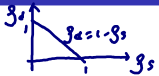
$P_{out} = \rho^{-g_d}$

$$\lim_{\rho \longrightarrow \infty} \frac{\log_2\left(P_{out}(R)\right)}{\log_2(\rho)} = -g_d^\star(g_s, \infty)$$

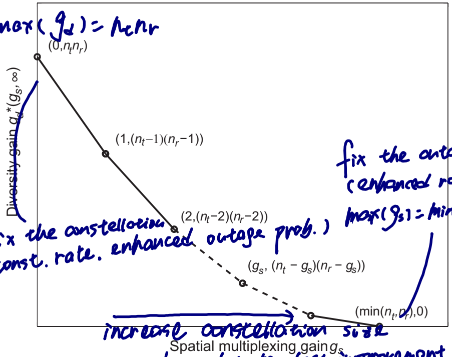
The curve $g_d^\star(g_s, \infty)$ as function of g_s is known as the asymptotic diversity-multiplexing trade-off of the channel.

- The multiplexing gain indicates how fast the transmission rate increases with the SNR.
- The diversity gain represents how fast the outage probability decays with the SNR.

Diversity-Multiplexing Trade-Off in I.I.D. Rayleigh Slow Fading Channels



- Point $(0, n_t n_r)$: for a spatial multiplexing gain of zero (i.e., R is fixed), the maximal diversity gain achievable is $n_t n_r$.
- Point $(\min\{n_t, n_r\}, 0)$: transmitting at diversity gain $g_d^\star =$ (i.e., P_{out} is kept fixed) allows the data rate to increase with SNR as $n = \min\{n_t, n_r\}$.
- Intermediate points: possible to transmit at non-zero diversity and multiplexing gains but that any increase of one of those quantities leads to a decrease of the other quantity.



max(g_d) = $n_t n_r$

fix the constellation (const. rate. enhanced outage prob.)

fix the outage prob. (enhanced rate)

max(g_s) = min(n_t, n_r)

increase constellation size (enhanced rate, less improvement on outage prob.)

Diversity-Multiplexing Trade-Off in I.I.D. Rayleigh Slow Fading Channels

multiplexing gain ⟷ rate ⇒ speed
diversity gain ⟷ outage prob. ⇒ reliability

- For fixed rates $R = 2, 4, ..., 40$ bits/s/Hz,
  - The asymptotic slope of each curve is four and matches the maximum diversity gain $g_d^\star(0, \infty)$.
  - The horizontal separation is 2 bits/s/Hz per 3 dB, which corresponds to the maximum multiplexing gain equal to $n(= 2)$.
- As the rate increases more rapidly with SNR (i.e., as the multiplexing gain g_s increases), the slope of the outage probability curve (given by the diversity gain g_d^*) vanishes.

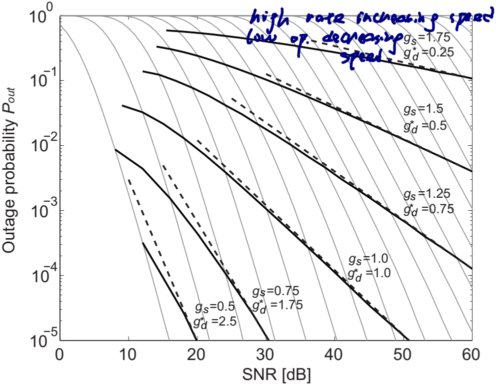


Figure: 2×2 MIMO i.i.d. Rayleigh fading channels

Impact on Coding Architecture

- Impossible to code over a large number of independent channel realizations $\rightarrow$ separate coding leads to an outage as soon as one of the subchannels is in deep fade.

- Joint coding across all subchannels necessary in the absence of transmit channel knowledge!

Tx and Rx Strategies: Space-Time Coding and Processing

Previous Lectures

- Previous lecture
  - Capacity of deterministic MIMO channels

$$C(\mathbf{H}) = \max_{\mathbf{Q}\geq 0:\mathsf{Tr}\{\mathbf{Q}\}=1} \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H\right].$$

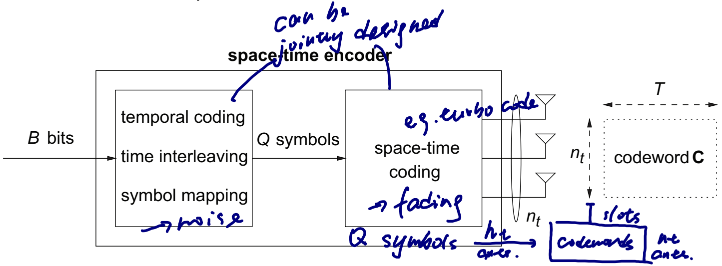
  - Ergodic capacity of fast fading channels
  - Outage capacity and probability of slow fading channels

- MIMO provides huge gains in terms of reliability and transmission rate
  - diversity gain, array gain, coding gain, spatial multiplexing gain, interference management

- What we further need
  - practical methodologies to achieve these gains?
  - how to code across space and time?
  - Some preliminary answers: multimode eigenmode transmission when channel knowledge available at the Tx, Alamouti scheme when no channel knowledge available at the Tx

Overview of a Space-Time Encoder

- Space-time encoder: sequence of two black boxes



$\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$ $Q=2$, $T=2$

- First black box: combat the randomness created by the noise at the receiver.
- Second black box: spatial interleaver which spreads symbols over several antennas in order to mitigate the spatial selective fading.
- The ratio B/T is the signaling rate of the transmission.
- The ratio Q/T is defined as the spatial multiplexing rate (representative of how many symbols are packed within a codeword per unit of time).

symbols per unit time

System Model

- MIMO system with n_t transmit and n_r receive antennas over a frequency flat-fading channel
- Transmit a codeword $\mathbf{C} = [\mathbf{c}_0 \dots \mathbf{c}_{T-1}]$ $[n_t \times T]$ contained in the codebook $\mathcal{C}$
- At the k^{th} time instant, the transmitted and received signals are related by

$$\mathbf{y}_k = \sqrt{E_s}\mathbf{H}_k\mathbf{c}_k + \mathbf{n}_k$$

$H \sim t$

  where
  - $\mathbf{y}_k$ is the $n_r \times 1$ received signal vector,
  - $\mathbf{H}_k$ is the $n_r \times n_t$ channel matrix,
  - $\mathbf{n}_k$ is a $n_r \times 1$ zero mean complex AWGN vector with $\mathcal{E}\{\mathbf{n}_k\mathbf{n}_l^H\} = \sigma_n^2\mathbf{I}_{n_r}\delta(k-l)$,
  - The parameter E_s is the energy normalization factor. SNR $\rho = E_s/\sigma_n^2$.
- No transmit channel knowledge but we know it is i.i.d. Rayleigh fading.
- Codeword average transmit power $\mathcal{E}\left\{\text{Tr}\left\{\mathbf{C}\mathbf{C}^H\right\}\right\} = T$. Assume $\mathcal{E}\left\{\|\mathbf{H}\|_F^2\right\} = n_t n_r$.
- Channel time variation:
  - *slow fading*: $T_{coh} >> T$ and $\{\mathbf{H}_k = \mathbf{H}_w\}_{k=0}^{T-1}$, with $\mathbf{H}_w$ denoting an i.i.d. random fading matrix with unit variance circularly symmetric complex Gaussian entries.
  - *fast fading*: $T \geq T_{coh}$ and $\mathbf{H}_k = \mathbf{H}_{k,w}$, where $\left\{\mathbf{H}_{k,w}\right\}_{k=0}^{T-1}$ are uncorrelated matrices, each $\left\{\mathbf{H}_{k,w}\right\}$ being an i.i.d. random fading matrix with unit variance circularly symmetric complex Gaussian entries.

Error Probability Motivated Design Methodology

- With instantaneous channel realizations perfectly known at the receive side, the ML decoder computes an estimate of the transmitted codeword according to

Alamouti: c_1, c_2 can be decoupled (no interf). $y = \sqrt{E_s} h c + n$
no need for ML.

ML: if c are coupled.
(eg. CSI changes)

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C}} \sum_{k=0}^{T-1} \left\| \mathbf{y}_k - \sqrt{E_s} \mathbf{H}_k \mathbf{c}_k \right\|^2 \quad (z-\mu) R^{-1} (z-\mu)$$

Rx Tx codewords next T

where the minimization is performed over all possible codeword vectors $\mathbf{C}$.

- Pairwise Error Probability (PEP): probability that the ML decoder decodes the codeword $\mathbf{E} = [\mathbf{e}_0 \ldots \mathbf{e}_{T-1}]$ instead of the transmitted codeword $\mathbf{C}$.
- When the PEP is conditioned on the channel realizations $\{\mathbf{H}_k\}_{k=0}^{T-1}$, it is defined as the conditional PEP,

$P = Q(\sqrt{d_{min}} \ldots)$
↳ chernoff bound
$Q(x) \le e^{-\frac{x^2}{2}}$
$\bar{P} = \mathcal{E}(e^{\cdot})$ moment generating func.

correlated in time / space

$$P\left(\mathbf{C} \rightarrow \mathbf{E} \middle| \{\mathbf{H}_k\}_{k=0}^{T-1}\right) = \mathcal{Q}\left(\sqrt{\frac{\rho}{2} \sum_{k=0}^{T-1} \left\| \mathbf{H}_k \underbrace{\left(\mathbf{c}_k - \mathbf{e}_k\right)}_{\text{distance between Tx symbol and error}} \right\|_F^2}\right)$$

where $\mathcal{Q}(x)$ is the Gaussian $\mathcal{Q}$-function.

- The average PEP, $P(\mathbf{C} \rightarrow \mathbf{E})$, obtained by averaging the conditional PEP over the probability distribution of the channel gains.
- System performance dominated at high SNR by the couples of codewords that lead to the worst PEP.

Diversity Gain and Coding Gain

- Assume a fixed rate transmission, i.e., spatial multiplexing gain $g_s = 0$.

Definition

The **diversity gain** $g_d^o(\rho)$ achieved by a pair of codewords $\{\mathbf{C}, \mathbf{E}\} \in \mathcal{C}$ is defined as the slope of $P(\mathbf{C} \to \mathbf{E})$ as a function of the SNR ρ on a log-log scale, usually evaluated at very high SNR, i.e.,

$$g_d^o(\infty) = \lim_{\rho \to \infty} g_d^o(\rho) = -\lim_{\rho \to \infty} \frac{\log_2 (P(\mathbf{C} \to \mathbf{E}))}{\log_2 \rho}.$$

PS: $g_d^o(\infty) \leftrightarrow P(\mathbf{C} \to \mathbf{E})$, $g_d^\star(0, \infty) \leftrightarrow P_{out}$.

Definition

The **coding gain** achieved by a pair of codewords $\{\mathbf{C}, \mathbf{E}\} \in \mathcal{C}$ is defined as the magnitude of the left shift of the $P(\mathbf{C} \to \mathbf{E})$ vs. ρ curve evaluated at very high SNR.

- If $P(\mathbf{C} \to \mathbf{E})$ is well approximated at high SNR by

$$P(\mathbf{C} \to \mathbf{E}) \approx c\,(g_c \rho)^{-g_d^o(\infty)}$$

with c being a constant, g_c is identified as the coding gain.

Derivation of the Average PEP

$C - E = \begin{bmatrix} c_1 - e_1 & -(c_2^* - e_2^*) \\ c_2 - e_2 & c_1^* - e_1^* \end{bmatrix}$ $\quad (C-E)^H = \begin{bmatrix} c_1^* - e_1^* & c_2^* - e_2^* \\ -(c_2 - e_2) & c_1 - e_1 \end{bmatrix}$

- Conditional PEP

$\tilde{E} = (C-E)(C-E)^H = \begin{bmatrix} (c_1-e_1)^2 + (c_2-e_2)^2 & 0 \\ 0 & (c_2-e_2)^2 + (c_1-e_1)^2 \end{bmatrix}$

$$P\left(\mathbf{C} \to \mathbf{E} | \left\{\mathbf{H}_k\right\}_{k=0}^{T-1}\right) = \mathcal{Q}\left(\sqrt{\frac{\rho}{2} \sum_{k=0}^{T-1} \left\|\mathbf{H}_k\left(\mathbf{c}_k - \mathbf{e}_k\right)\right\|_F^2}\right)$$

where $\mathcal{Q}(x)$ is the Gaussian $\mathcal{Q}$-function defined as

$$\mathcal{Q}(x) \triangleq P(y \geq x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy.$$

$P = Q(\sqrt{d^2 |h|^2})$

$\bar{P} = \mathcal{E}_h\{Q\sqrt{d^2 |h|^2}\}$

$\bar{P} \approx \mathcal{E}_h\{e^{-0}\}$

MGF

$\mathrm{Rank}(\tilde{E}) \leq \min(n_t, n_r)$

- Average PEP

$$P(\mathbf{C} \to \mathbf{E}) = \mathcal{E}_{\mathbf{H}_k}\left\{P\left(\mathbf{C} \to \mathbf{E} | \left\{\mathbf{H}_k\right\}_{k=0}^{T-1}\right)\right\}.$$

- This integration is sometimes difficult to calculate. Use Chernoff bound

$$\mathcal{Q}(x) \leq \exp\left(-\frac{x^2}{2}\right).$$

- Average PEP

$$P(\mathbf{C} \to \mathbf{E}) = \mathcal{E}_{\mathbf{H}_k}\left\{P\left(\mathbf{C} \to \mathbf{E} | \left\{\mathbf{H}_k\right\}_{k=0}^{T-1}\right)\right\}$$

  – Computable using moment generating function (MGF), e.g. see in SISO

Slow Fading MIMO Channels

$\det(\mathbf{I}_{n_t} + \frac{\rho}{4}\tilde{\mathbf{E}})^{-n_r} = \det(\mathbf{I}_{n_t} + \frac{\rho}{4} V(\ddots)V^H)^{-n_r}$

$\tilde{\mathbf{E}} = V(\ddots)V^H$ $\quad = \prod_{i=1}^{r(\tilde{\mathbf{E}})} (1+\frac{\rho}{4}\lambda_k)^{-n_r}$

- The conditional PEP writes as

$$P\left(\mathbf{C} \to \mathbf{E} \mid \mathbf{H}\right) = \mathcal{Q}\left(\sqrt{\frac{\rho}{2}\left\|\mathbf{H}\left(\mathbf{C}-\mathbf{E}\right)\right\|_F^2}\right).$$

- The average PEP over Rayleigh slow fading channels is

$$P\left(\mathbf{C} \to \mathbf{E}\right) = \mathcal{E}_{\mathbf{H}}\left\{P\left(\mathbf{C} \to \mathbf{E} | \mathbf{H}\right)\right\}$$

- Using MGF and Chernoff bound

Alamouti

$CC^H = \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}\begin{bmatrix} c_1^* & c_2^* \\ -c_2 & c_1 \end{bmatrix} = \begin{bmatrix} |c_1|^2+|c_2|^2 & c_1c_2^* - c_2^*c_1 \\ c_2c_1^* - c_1^*c_2 & |c_1|^2+|c_2|^2 \end{bmatrix}$

$(C-E)(C-E)^H = (|c_1-e_1|^2+|c_2-e_2|^2)\,I$

error matrix $\tilde{E} = (C-E)(C-E)^H$

rank

$$P\left(\mathbf{C} \to \mathbf{E}\right) \leq \left[\det\left(\mathbf{I}_{n_t} + \frac{\rho}{4}\tilde{\mathbf{E}}\right)\right]^{-n_r} = \prod_{i=1}^{r(\tilde{\mathbf{E}})}\left(1+\frac{\rho}{4}\lambda_i\left(\tilde{\mathbf{E}}\right)\right)^{-n_r}$$

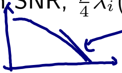
diversity gain $n_r \cdot r(\tilde{E})$

eigenvalues of $\tilde{E}$

with $r(\tilde{\mathbf{E}})$ denotes the rank of the error matrix $\tilde{\mathbf{E}} = \left(\mathbf{C}-\mathbf{E}\right)\left(\mathbf{C}-\mathbf{E}\right)^H$ and $\left\{\lambda_i\left(\tilde{\mathbf{E}}\right)\right\}$ for $i = 1, \ldots, r(\tilde{\mathbf{E}})$ the set of its non-zero eigenvalues.

- At high SNR, $\frac{\rho}{4}\lambda_i\left(\tilde{\mathbf{E}}\right) >> 1$ $\quad P(C\to E) = (g_c\rho)^{-g_d}$ $\quad$ rank ↗ ⇒ P ↘.

$g_d = n_r r(\tilde{E})$



$$P\left(\mathbf{C} \to \mathbf{E}\right) \leq \left(\frac{\rho}{4}\right)^{-n_r r(\tilde{\mathbf{E}})}\underbrace{\prod_{i=1}^{r(\tilde{\mathbf{E}})}\lambda_i^{-n_r}\left(\tilde{\mathbf{E}}\right)}$$

→ SISO: distance between constellations

all Rx · employed Tx

- diversity gain: $n_r r\left(\tilde{\mathbf{E}}\right)$, coding gain: $\prod_{i=1}^{r(\tilde{\mathbf{E}})}\lambda_i\left(\tilde{\mathbf{E}}\right)$. multiplexing gain = min(n_t, n_r)

The Rank-Determinant Criterion

Design Criterion

(**Rank-determinant criterion**) *Over i.i.d. Rayleigh slow fading channels,*

1. **rank criterion**: *maximize the minimum rank* r_{min} *of* $\tilde{\mathbf{E}}$ *over all pairs of codewords* $\{\mathbf{C}, \mathbf{E}\}$ *with* $\mathbf{C} \neq \mathbf{E}$ (error dominated by the weakest pair)

rank $(\tilde{E})$ → the dissimiliarity of C and E

$$r_{min} = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} r(\tilde{\mathbf{E}})$$

$r(\tilde{E}) \leq \min(n_t, T)$

2. **determinant criterion**: *over all pairs of codewords* $\{\mathbf{C}, \mathbf{E}\}$ *with* $\mathbf{C} \neq \mathbf{E}$, *maximize the minimum of the product* d_λ *of the non-zero eigenvalues of* $\tilde{\mathbf{E}}$,

$$d_\lambda = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} \prod_{i=1}^{r(\tilde{\mathbf{E}})} \lambda_i(\tilde{\mathbf{E}}).$$

(full rank)

If $r_{min} = n_t$, *the determinant criterion comes to maximize the minimum determinant of the error matrix over all pairs of codewords* $\{\mathbf{C}, \mathbf{E}\}$ *with* $\mathbf{C} \neq \mathbf{E}$

$$d_\lambda = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} \det(\tilde{\mathbf{E}}).$$

The Rank-Determinant Criterion

n_t $\boxed{C}$ $T=1$, $n_t=2$ $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ only spatial complexity $r(\tilde{E}) \le \min(2,1)$

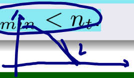
transmit c_1, c_2 independently $\frac{Q}{T} = 2$. rate ↗ P_e ↗ $g_d = 1$

Definition

A full-rank (a.k.a. full-diversity) code is characterized by $r_{min} = n_t$. A rank-deficient code is characterized by $r_{min} < n_t$.

~~spatial multiplexing~~

$T=2$, $n_t=2$ $\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$ $r(\tilde{E}) \le \min(2,2)$



$\frac{Q}{T} = 1$ rate ↘ P_e ↘ $g_d = 2$

Example

Rank-deficient and full-rank codes for $n_t = 2$

- Rank-deficient code

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

- Full-rank code

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}, \mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} e_1 & -e_2^* \\ e_2 & e_1^* \end{bmatrix}$$

$$(\mathbf{C} - \mathbf{E})(\mathbf{C} - \mathbf{E})^H = \frac{1}{2} \begin{bmatrix} |c_1 - e_1|^2 + |c_2 - e_2|^2 & 0 \\ 0 & |c_1 - e_1|^2 + |c_2 - e_2|^2 \end{bmatrix}$$

The Rank-Determinant Criterion

Alamouti: O-STBC

Example

Question: Relying on the rank-determinant criterion, show that delay diversity achieves full diversity. Assume for simplicity two transmit antennas.

Answer: The codeword for delay diversity can be written as

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_1 & c_2 & \ldots & c_{T-1} & 0 \\ 0 & c_1 & c_2 & \ldots & c_{T-1} \end{bmatrix}.$$

Taking another codeword $\mathbf{E}$, different from $\mathbf{C}$,

$$\mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} e_1 & e_2 & \ldots & e_{T-1} & 0 \\ 0 & e_1 & e_2 & \ldots & e_{T-1} \end{bmatrix}.$$

The diversity gain is given by the minimum rank of the error matrix over all possible pairs of (different) codewords, i.e.

$$r_{min} = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} r\big(\tilde{\mathbf{E}}\big) = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} r\big(\mathbf{C}-\mathbf{E}\big).$$

The Rank-Determinant Criterion

Example

With delay diversity, we have

$$\mathbf{C}-\mathbf{E}=\frac{1}{\sqrt{2}}\begin{bmatrix} c_1-e_1 & c_2-e_2 & \dots & c_{T-1}-e_{T-1} & 0 \\ 0 & c_1-e_1 & c_2-e_2 & \dots & c_{T-1}-e_{T-1} \end{bmatrix}.$$

Obviously, $r\left(\mathbf{C}-\mathbf{E}\right)\leq 2$. Actually, $r\left(\mathbf{C}-\mathbf{E}\right)=2$ as long as $\mathbf{C}\neq\mathbf{E}$. Indeed even in the case where all $c_k-e_k=0$ except for one index k (in order to keep $\mathbf{C}\neq\mathbf{E}$), e.g. $k=1$,

$$\mathbf{C}-\mathbf{E}=\frac{1}{\sqrt{2}}\begin{bmatrix} c_1-e_1 & 0 & \dots & 0 & 0 \\ 0 & c_1-e_1 & 0 & \dots & 0 \end{bmatrix},$$

the rank is equal to 2. Hence diversity gain of $2n_r$.

The Rank-Determinant Criterion

$n_t = 2$ $\quad \frac{Q}{T} = 2 \quad r(\hat{E}) = \min(n_t, T) \le 2$

$T = 2$

Example

$Q = 4$

Question: Assume that c_1, c_2, c_3 and c_4 are constellation symbols taken from a unit average energy QAM constellation. Consider the Linear Space-Time Block Code, characterized by codewords

$C = \begin{bmatrix} c_1+c_3 & c_2+c_4 \\ c_2-c_4 & c_1-c_3 \end{bmatrix}$ each entry has 4 options

$\Rightarrow$ totally $4^4 = 256$

full rate code?

full rank

$$\mathbf{C} = \frac{1}{2}\begin{bmatrix} c_1 + c_3 & c_2 + c_4 \\ c_2 - c_4 & c_1 - c_3 \end{bmatrix}.$$

hard to detect with ML.

$\hat{C} = \arg\min \| y - H \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \|^2$ 16 complexity

What is the diversity gain achieved by this code over slow Rayleigh fading channels?

$r(\hat{E}) \to 1 \Rightarrow g_d = n_r$

Answer: Check the rank of its error matrix

Alamonti vs. mentioned:

rank deficient can happen.

$\begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$ $\quad \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$$\mathbf{C} - \mathbf{E} = \frac{1}{2}\begin{bmatrix} d_1 + d_3 & d_2 + d_4 \\ d_2 - d_4 & d_1 - d_3 \end{bmatrix}$$

g_d ✓ $\quad$ r ✓ $\quad$ decouple ✓

g_d ✗ $\quad$ r ✗ $\quad$ decouple ✗

where $d_k = c_k - e_k$ for $k = 1, \ldots, 4$. This code is rank deficient. It is easily seen that by taking two codewords $\mathbf{C}$ and $\mathbf{E}$ such that $d_3 = d_4 = 0$ and $d_1 = d_2 = d$ (which is encountered for any constellations), $r(\mathbf{C} - \mathbf{E}) = 1$. Hence diversity gain of n_r.

Space-Time Block Coding (STBC)

- STBCs can be seen as a mapping of Q symbols (complex or real) onto a codeword $\mathbf{C}$ of size $n_t \times T$.
- Codewords are uncoded in the sense that no error correcting code is contained in the STBC.
- Linear STBCs are by far the most widely used
  - Spread information symbols in space and time in order to improve either the diversity gain, either the spatial multiplexing rate ($r_s = \frac{Q}{T}$) or both the diversity gain and the spatial multiplexing rate.
  - Pack more symbols into a given codeword, i.e., increase Q, to increase the data rate.

Example

Alamouti code: $n_t = 2$, $Q = 2$, $T = 2$, $r_s = 1$

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}.$$

SM $\begin{cases} r_s = n_t \\ g_d = n_r \end{cases}$

- Spatial Multiplexing (SM), also called V-BLAST, is a full rate code ($r_s = n_t$) that consists in transmitting independent data streams on each transmit antenna.
- In uncoded transmissions, we assume one symbol duration ($T = 1$) and codeword $\mathbf{C}$ is a symbol vector of size $n_t \times 1$.

$C - \tilde{E} = |$ $\quad (C-\tilde{E})^H = —$ $\quad \tilde{E} = \ulcorner$ $\quad r(\tilde{E}) = 1$ $\quad$ # eigenvalue $= 1$ $\quad$ value $=$ norm$(C-\tilde{E})$

Example

$$\mathbf{C} = \frac{1}{\sqrt{n_t}} \left[\begin{array}{ccc} c_1 & \ldots & c_{n_t} \end{array} \right]^T.$$

Each element c_q is a symbol chosen from a given constellation.

diversity gain $= r(\tilde{E}) \cdot n_r$

$r(\tilde{E}) = 1 \Downarrow g_d = n_r$

achieve capacity when fast-fading.

high complexity at Rx when $n_t \nearrow$

QPSK: 2×2 bps/Hz $= 4$ bps/Hz (higher rate)

$g_d = r(\tilde{E}) = 1$ (low diversity gain)

$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ 8 bits $\quad \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$ 8 bits

c_1, c_2: 4 bits each $\Rightarrow$ 16-QAM.

for same modulation (QPSK): 2×1 bps/Hz $= 2$ bps/Hz (lower rate)

$g_d = r(\tilde{E}) = 2$ (high div. gain)

ML decoding

- Error probability

{diversity gain = $-n_r$

achieves ergodic capacity

Spatial multiplexing + ML :

$$P(\mathbf{C} \to \mathbf{E}) \le \left(\frac{\rho}{4n_t}\right)^{-n_r} \left(\sum_{q=1}^{n_t} |c_q - e_q|^2\right)^{-n_r}$$

The SNR exponent is equal to n_r. Due to the lack of coding across transmit antennas, no transmit diversity is achieved and only receive diversity is exploited.

- Over fast fading channels, we know that it is not necessary to code across antennas to achieve the ergodic capacity.

Proposition

Spatial Multiplexing with ML decoding and equal power allocation achieves the ergodic capacity of i.i.d. Rayleigh fast fading channels.

$\arg\min_{\underline{c}} \|\underline{y} - H\underline{c}\|^2$

Zero-Forcing (ZF) Linear Receiver

2x2 $y = [h_1\ h_2]\begin{bmatrix}c_1\\c_2\end{bmatrix} + n$ $y = h_1c_1 + h_2c_2 + n$ coloured noise

- MIMO ZF receiver acts similarly to a ZF equalizer in frequency selective channels.

$g_1^T y = g_1^T h_1 c_1 + g_1^T h_2 c_2 + g_1^T n$

noise enhancement

- ZF filtering effectively decouples the channel into n_t parallel channels

row vect.   forcing interf from c_2 to be zero

  - interference from other transmitted symbols is suppressed
  - scalar decoding may be performed on each of these channels

$g_2^T y = g_2^T h_1 c_1 + g_2^T h_2 c_2 + g_2^T n$ $z = \begin{bmatrix}z_1\\z_2\end{bmatrix} = \begin{bmatrix} g_1^T h_1 & 0 \\ 0 & g_2^T h_2\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix} + \begin{bmatrix} g_1^T n \\ g_2^T n\end{bmatrix}$

- The complexity of ZF decoding similar to SISO ML decoding, but the inversion step is responsible for the noise enhancement (especially at low SNR).

2x2 Only 1 (dir) vector orthogonal to h_1/h_2

$z = H^{-1}y = H^{-1}HC + H^{-1}n$ more c_i

- Assuming that a symbol vector $\mathbf{C} = 1/\sqrt{n_t}\left[c_1 \quad \dots \quad c_{n_t}\right]^T$ is transmitted, the output of the ZF filter $\mathbf{G}_{ZF}$ is given by

more options

SNR

if $h_1 \cdot h_2$ orthogonal matched filter maximise SNR

$g_1 = h_1^H \Rightarrow$ orthogonal to h_2: no interf.

$$\mathbf{z} = \mathbf{G}_{ZF}\mathbf{y} = \left[\, c_1 \quad \dots \quad c_{n_t} \,\right]^T + \mathbf{G}_{ZF}\mathbf{n}$$

$n_r < n_t$: cannot avoid interf. for all!

where $\mathbf{G}_{ZF}$ inverts the channel,

$n_r \geq n_t$: $z = H^\dagger HC + H^\dagger n$

$= (H^H H)^{-1} H^H H C + H^{-1} n$

$n_r \neq n_t$

$\boxed{\text{ZF: decouple to SISO}}$

if $h_1 \cdot h_2$ closer. always no interference. project maximise SNR but smaller.

$$\mathbf{G}_{ZF} = \sqrt{\frac{n_t}{E_s}}\,\mathbf{H}^\dagger$$

with $\mathbf{H}^\dagger = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H$ denoting the Moore-Penrose pseudo inverse.

Zero-Forcing (ZF) Linear Receiver

- Covariance matrix of the noise at the output of the ZF filter

noise enhancement

$$\mathcal{E}\left\{\mathbf{G}_{ZF}\mathbf{n}\left(\mathbf{G}_{ZF}\mathbf{n}\right)^{H}\right\} = \frac{n_t}{\rho}\mathbf{H}^{\dagger}\left(\mathbf{H}^{\dagger}\right)^{H} = \frac{n_t}{\rho}\left(\mathbf{H}^{H}\mathbf{H}\right)^{-1}.$$

- The output SNR on the q^{th} subchannel is thus given by

ZF: pro: lower complexity
con: diversity gain $= n_r - n_t + 1$
underdetermined if $n_t > n_r$.

$$\rho_q = \frac{\rho}{n_t}\frac{1}{\left(\mathbf{H}^{H}\mathbf{H}\right)^{-1}(q,q)}, \qquad q = 1, \ldots, n_t.$$

- Inversion leads to noise enhancement. Severe degradation at low SNR.
- Assuming that the channel is i.i.d. Rayleigh distributed, ρ_q is a χ^2 random variable with $2(n_r - n_t + 1)$ degrees of freedom, denoted as $\chi^2_{2(n_r-n_t+1)}$. The average PEP on the q^{th} subchannel is thus upper-bounded by

$$P\left(c_q \rightarrow e_q\right) \leq \left(\frac{\rho}{4n_t}\right)^{-(n_r-n_t+1)} \left|c_q - e_q\right|^{-2(n_r-n_t+1)}.$$

The lower complexity of the ZF receiver comes at the price of a diversity gain limited to $n_r - n_t + 1$. Clearly, the system is undetermined if $n_t > n_r$.

Zero-Forcing (ZF) Linear Receiver

- In fast fading channels, the average maximum achievable rate $\bar{C}_{ZF}$ is equal to the sum of the maximum rates achievable by all layers

$$\bar{C}_{ZF} = \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(1+\rho_q\right)\right\}$$

$$\overset{(\rho\nearrow)}{\approx} \min\left\{n_t, n_r\right\} \log_2\left(\frac{\rho}{n_t}\right) + \min\left\{n_t, n_r\right\} \mathcal{E}\left\{\log_2\left(\chi^2_{2(n_r-n_t+1)}\right)\right\}.$$

Note the difference with

$$\bar{C}_{CDIT} \approx n \log_2\left(\frac{\rho}{n_t}\right) + \sum_{k=1}^{n} \mathcal{E}\left\{\log_2(\chi^2_{2(N-n+k)})\right\}.$$

Spatial Multiplexing in combination with a ZF decoder allows for transmitting over $n = \min\{n_t, n_r\}$ independent data pipes.

Zero-Forcing (ZF) Linear Receiver

- ZF receiver maximizes the SNR under the constraint that the interferences from all other layers are nulled out.
  - For a given layer q, the ZF combiner $\mathbf{g}_q$ is such that this layer is detected through a projection of the output vector $\mathbf{y}$ onto the direction closest to $\mathbf{H}(:,q)$ within the subspace orthogonal to the one spanned by the set of vectors $\mathbf{H}(:,p)$, $p \neq q$.

→ maximise gain

→ suppress interf.

- Assume the following system model with $n_r \geq n_t$

$n_r = 4$, $n_t = 3$, $q = 1$

$H = [h_1\, h_2\, h_3]$ 4×3

$H_{-1} = [h_2\, h_3]$ 4×2 $= \underset{4\times4}{U}\,\underset{4\times2}{\Lambda}\,\underset{2\times2}{V^H}$

$y = h_1 c_1 + H_{-1}\begin{bmatrix} c_2 \\ c_3 \end{bmatrix} + n$

$$\mathbf{y} = \mathbf{H}\mathbf{c} + \mathbf{n},$$

$$= \mathbf{h}_q c_q + \underbrace{\sum_{p\neq q} \mathbf{h}_p c_p}_{\text{interf}} + \mathbf{n}$$

$\# = n_t - 1$

proj. $\tilde{U}^H y = \tilde{U}^H h_1 c_1 + \tilde{U}^H [U' | \tilde{U}] \begin{bmatrix} \Lambda \\ 0 \end{bmatrix} [2\times2] \begin{bmatrix} c_2 \\ c_3 \end{bmatrix} + \tilde{U}^H n$

AWGN enhanced

zero block

where $\mathbf{h}_q$ is the q^{th} column of $\mathbf{H}$.

- Let us build the following $n_r \times (n_t - 1)$ matrix by collecting all h_p with $p \neq q$:

$$\mathbf{H}_{-q} = \begin{bmatrix} \dots & \mathbf{h}_p & \dots \end{bmatrix}_{p\neq q},$$

$$= \begin{bmatrix} \mathbf{U}' & \tilde{\mathbf{U}} \end{bmatrix} \mathbf{\Lambda} \mathbf{V}^H$$

$[0_{2\times2} | I_2] \begin{bmatrix} \Lambda \\ 0 \end{bmatrix} = 0_{2\times2}$

$n_t > n_r$?

- $C_{ZF} < C_{tot}$
- ML ✓ (achieve capacity)
- ZF, MMSE ✗ (underdetermined)

where $\tilde{\mathbf{U}}$ is the matrix containing the left singular vectors corresponding to the null singular values. Similarly we define

$h_1^H \tilde{U} \tilde{U}^H y = h_1^H \tilde{U} \underbrace{\tilde{U}^H h_1}_{\text{new channel}} c + n'$: MRC!

$$\mathbf{c}_{-q} = \begin{bmatrix} \dots & c_p & \dots \end{bmatrix}^T_{p\neq q}.$$

Zero-Forcing (ZF) Linear Receiver

- By multiplying by $\tilde{\mathbf{U}}^H$, we project the output vector onto the subspace orthogonal to the one spanned by the columns of $\mathbf{H}_{-q}$

Rx +1 nr+nt

4 [ ] [ ] 3 Tx

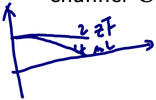
$$\tilde{\mathbf{U}}^H\mathbf{y} = \tilde{\mathbf{U}}^H\mathbf{h}_q c_q + \tilde{\mathbf{U}}^H\mathbf{H}_{-q}\mathbf{c}_{-q} + \tilde{\mathbf{U}}^H\mathbf{n}$$
$$= \tilde{\mathbf{U}}^H\mathbf{h}_q c_q + \tilde{\mathbf{U}}^H\mathbf{n}.$$

nr-nt+1

=nr-(nt-1) # interf

4 2 # antennas to cancel interf

- To maximize the SNR, noting the noise is still white, we match to the effective channel $\tilde{\mathbf{U}}^H\mathbf{h}_q$ such that



$$z = \left(\tilde{\mathbf{U}}^H\mathbf{h}_q\right)^H \tilde{\mathbf{U}}^H\mathbf{h}_q c_q + \left(\tilde{\mathbf{U}}^H\mathbf{h}_q\right)^H \tilde{\mathbf{U}}^H\mathbf{n}$$

and the ZF combiner is $g_q = \left(\tilde{\mathbf{U}}^H\mathbf{h}_q\right)^H \tilde{\mathbf{U}}^H = \mathbf{h}_q^H\tilde{\mathbf{U}}\tilde{\mathbf{U}}^H$.

Minimum Mean Squared Error (MMSE) Linear Receiver

ZF → SNR MMSE → SINR

- Filter maximizing the SINR. Minimize the total resulting noise: find $\mathbf{G}$ such that $\mathcal{E}\{\|\mathbf{G}\mathbf{y} - [\begin{array}{ccc} c_1 & \ldots & c_{n_t} \end{array}]^T\|^2\}$ is minimum.
- The combined noise plus interference signal $\mathbf{n}_{i,q}$ when estimating symbol c_q writes as

MMSE: balance interf and noise.

whitening noise + matched filter.

$$\mathbf{n}_{i,q} = \sum_{p \neq q} \sqrt{\frac{E_s}{n_t}} \mathbf{h}_p c_p + \mathbf{n}.$$

MMSE { low SNR: matched filter ↓ high SNR: ZF.

The covariance matrix of $\mathbf{n}_{i,q}$ reads as

$$\mathbf{R}_{\mathbf{n}_{i,q}} = \mathcal{E}\left\{\mathbf{n}_{i,q}\mathbf{n}_{i,q}^H\right\} = \sigma_n^2 \mathbf{I}_{n_r} + \sum_{p \neq q} \frac{E_s}{n_t} \mathbf{h}_p \mathbf{h}_p^H$$

and the MMSE combiner for stream q is given by

$$\underbrace{\mathbf{g}_{MMSE,q}}_{\text{branch}} = \sqrt{\frac{E_s}{n_t}} \mathbf{h}_q^H \left(\underbrace{\sigma_n^2 \mathbf{I}_{n_r}}_{\text{noise}} + \underbrace{\sum_{p \neq q} \frac{E_s}{n_t} \mathbf{h}_p \mathbf{h}_p^H}_{\text{interf}}\right)^{-1}$$

large dim ↓ complexity.

- An alternative and popular representation of the MMSE filter can also be written as

$$\underbrace{\mathbf{G}_{MMSE}}_{\text{entity}} = \sqrt{\frac{n_t}{E_s}} \left(\mathbf{H}^H\mathbf{H} + \overbrace{\boxed{\frac{n_t}{\rho}\mathbf{I}_{n_t}}}^{\text{noise}}\right)^{-1} \mathbf{H}^H = \sqrt{\frac{n_t}{E_s}} \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \frac{n_t}{\rho}\mathbf{I}_{n_r}\right)^{-1}$$

SNR ↑ → noise ↓ → ZF.

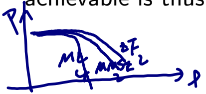
- Bridge between matched filtering at low SNR and ZF at high SNR.

Minimum Mean Squared Error (MMSE) Linear Receiver

- The output SINR on the q^{th} subchannel (stream) is given by

$$\rho_q = \frac{E_s}{n_t}\mathbf{h}_q^H \left(\sigma_n^2 \mathbf{I}_{n_r} + \sum_{p \neq q} \frac{E_s}{n_t}\mathbf{h}_p\mathbf{h}_p^H \right)^{-1} \mathbf{h}_q.$$

- At high SNR, the MMSE filter is practically equivalent to ZF and the diversity achievable is thus limited to $n_r - n_t + 1$.



diversity gain : high SNR → noise ↓↓
→ ZF

Successive Interference Canceler

ZF. MMSE: decouple to independent streams. $n_r-(n_t-1)$

SIC: decode one, subtract ...

- Successively decode one symbol (or more generally one layer/stream) and cancel the effect of this symbol from the received signal. # interf
- Decoding order based on the SINR of each symbol/layer: the symbol/layer with the highest SINR is decoded first at each iteration. less prob. in error
- SM with (ordered) SIC is generally known as V-BLAST, and ZF and MMSE V-BLAST refer to SM with respectively ZF-SIC and MMSE-SIC receivers.

$y = [h_1\ h_2]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + n$ ML → $\hat{c}_1$ $z = y - h_1\hat{c}_1 = h_2 c_2 + n_2$

$\boxed{g_1} y = g_1 [h_1\ h_2]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + g_1 n = g_1 h_1 c_1 + \underbrace{g_1 h_2 c_2 + g_1 n}_{\text{interf+noise}}$

$\boxed{C_i \text{ sees interference from } C_{i+1}, C_{i+2} \ldots}$

C_1 $g_d=1$, C_2 2, C_3 3, C_4 4, C_5 5

based on MMSE

- The diversity order experienced by the decoded layer is increased by one at each iteration. Therefore, the symbol/layer detected at iteration i will achieve a diversity of $n_r - n_t + i$. less interf.

C_1: div. gain $= 1$ (interf reduces g_d)

C_2: div. gain $= 2$ (h_2: 2x1, MRC)

- Major issue: error propagation
  - The error performance is mostly dominated by the weakest stream.
  - Non-ordered SIC: diversity order approximately $n_r - n_t + 1$.
  - Ordered SIC: performance improved by reducing the error propagation caused by the first decoded stream. The diversity order remains lower than n_r.

WiMAX → ML

LTE → MMSE

Successive Interference Canceler

SINR

1. *Initialization:* $i \longleftarrow 1$, $\mathbf{y}^{(1)} = \mathbf{y}, \mathbf{G}^{(1)} = \mathbf{G}_{ZF}(\mathbf{H}), q_1 \overset{(*)}{=} \arg\min_j \left\|\mathbf{G}^{(1)}(j,:)\right\|^2$
   where $\mathbf{G}_{ZF}(\mathbf{H})$ is defined as the ZF filter of the matrix $\mathbf{H}$.

   $SINR = \frac{\sigma_s^2}{|f|^2 \sigma_n^2}$
2. *Recursion:*
   1. *step 1:* extract the q_i^{th} transmitted symbol from the received signal $\mathbf{y}^{(i)}$
   $$\tilde{c}_{q_i} = \mathbf{G}^{(i)}(q_i,:)\,\mathbf{y}^{(i)}$$
   $(C_{q_i} \pm C_i)$

      where $\mathbf{G}^{(i)}(q_i,:)$ is the q_i^{th} row of $\mathbf{G}^{(i)}$;
   2. *step 2:* slice $\tilde{c}_{q_i}$ to obtain the estimated transmitted symbol $\hat{c}_{q_i}$;
   3. *step 3:* assume that $\hat{c}_{q_i} = c_{q_i}$ and construct the received signal
   $$\begin{aligned} \mathbf{y}^{(i+1)} &= \mathbf{y}^{(i)} - \sqrt{\frac{E_s}{n_t}}\mathbf{H}(:,q_i)\,\hat{c}_{q_i} \\ \mathbf{G}^{(i+1)} &= \mathbf{G}_{ZF}\left(\mathbf{H}_{\overline{q_i}}\right) \\ i &\longleftarrow i+1 \\ q_{i+1} &\overset{(*)}{=} \arg\min_{j \notin \{q_1,\ldots,q_i\}} \left\|\mathbf{G}^{(i+1)}(j,:)\right\|^2 \end{aligned}$$
      where $\mathbf{H}_{\overline{q_i}}$ is the matrix obtained by zeroing columns $q_1, \ldots, q_i$ of $\mathbf{H}$. Here $\mathbf{G}_{ZF}\left(\mathbf{H}_{\overline{q_i}}\right)$ denotes the ZF filter applied to $\mathbf{H}_{\overline{q_i}}$.

Successive Interference Canceler

- In fast fading channels, the maximum rate achievable with ZF-SIC

$$\bar{C}_{ZF-SIC} = \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(1+\rho_q\right)\right\}$$

$$\stackrel{(\rho\nearrow)}{\approx} \min\{n_t,n_r\}\log_2\left(\frac{\rho}{n_t}\right) + \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(\chi^2_{2(n_r-n_t+q)}\right)\right\} = \bar{C}_{CDIT}$$

The loss that was observed with ZF filtering is now compensated because the successive interference cancellation improves the SNR of each decoded layer.

Proposition

Spatial Multiplexing with ZF-SIC (ZF V-BLAST) and equal power allocation achieves the ergodic capacity of i.i.d. Rayleigh fast fading MIMO channels at asymptotically high SNR.

This only holds true when error propagation is neglected.

Successive Interference Canceler

- MMSE-SIC does better for any SNR

$$\bar{C}_{MMSE-SIC} = \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2(1+\rho_q)\right\} = \mathcal{E}\left\{\log_2 \det\left(\mathbf{I}_{n_r} + \frac{\rho}{n_t}\mathbf{H}\mathbf{H}^H\right)\right\} = \bar{\mathcal{I}}_e,$$

Proposition

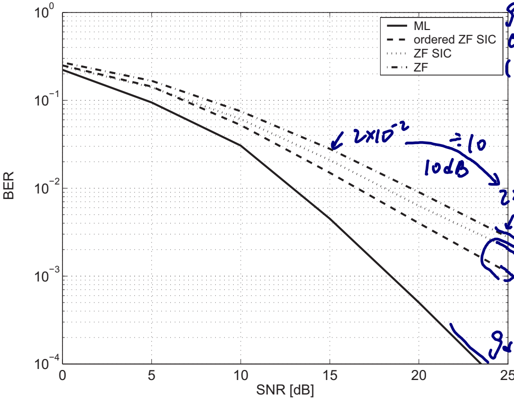
Spatial Multiplexing with MMSE-SIC (MMSE V-BLAST) and equal power allocation achieves the ergodic capacity for all SNR in i.i.d. Rayleigh fast fading MIMO channels.

Result also valid for a deterministic channel.

- Diversity achieved similar to that of ZF/MMSE receiver. This comes from the fact that the first layer dominates the error probability since its error exponent is the smallest.

Impact of Decoding Strategy on Error Probability

- SM with ML, ordered and non ordered ZF SIC and simple ZF decoding in 2×2 i.i.d. Rayleigh fading channels for 4 bits/s/Hz.



$g_d >> 1$ $g_d > 1$ $g_d = 1$

ordered > unordered > ZF

(closed to 1, BER dominated by worse branch)

$g_d = n_r - n_t + 1 = 1$

$g_d > 1$ $(C_1 = 1 . C_2 = 2)$

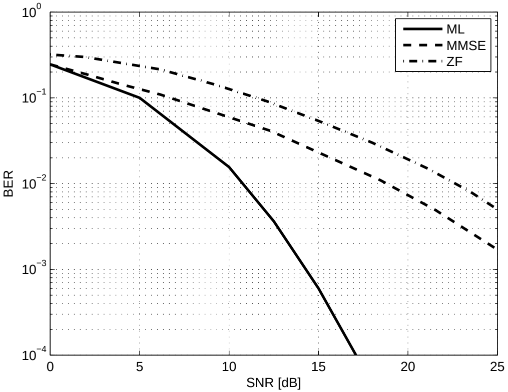
$g_d' > g_d$ $(C_{q_1} = 1, C_{q_2} = 2)$

$g_d = n_r = 2$

The slope of the ML curve approaches 2. ZF only achieves a diversity order of $n_r - n_t + 1 = 1$.

Impact of Decoding Strategy on Error Probability

- SM with ML, ZF and MMSE in i.i.d. Rayleigh slow fading channels with $n_t = n_r = 4$ and QPSK.



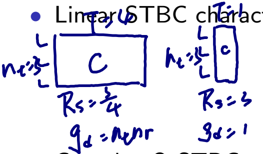
Orthogonal Space-Time Block Codes

STBC → diversity. reliability. $C = n_t \overset{T}{[\quad]}$ orthogonal: $CC^H = \alpha I$ decouple, $H_{eff}^H H_{eff} = I$

↳ decouple available (low-complexity)

- O-STBC vs. SM
  - Remarkable properties which make them extremely easy to decode: MIMO ML decoding decouples into several SIMO ML decoding
  - Achieve a full-diversity of $n_t n_r$.
  - Much smaller spatial multiplexing rate than SM.

- Linear STBC characterized by the two following properties



$$\mathbf{C}\mathbf{C}^H = \frac{T}{Qn_t}\left[\sum_{q=1}^{Q}|c_q|^2\right]\mathbf{I}_{n_t}.$$

- Complex O-STBCs with $r_s = 1$ only exist for $n_t = 2$. For larger n_t, codes exist with $r_s \leq 1/2$. For some particular values of $n_t > 2$, complex O-STBCs with $1/2 < r_s < 1$ have been developed. This is the case for $n_t = 3$ and $n_t = 4$ with $r_s = 3/4$.

Orthogonal Space-Time Block Codes

Example

Alamouti code: complex O-STBC for $n_t = 2$ with a spatial multiplexing rate $r_s = 1$

$g_d = |h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2 = cp$

$$\mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}.$$

- $\mathbf{C}\mathbf{C}^H = \frac{1}{2}\left[\,|c_1|^2 + |c_2|^2\,\right]\mathbf{I}_2$.
- $r_s = 1$ since two symbols are transmitted over two symbol durations.

Example

For $n_t = 3$, a complex O-STBC expanding on four symbol durations ($T = 4$) and transmitting three symbols on each block ($Q = 3$)

LTE-A: $n_t = 2, 4, 8$

Alamouti

comb. of 2x Alamouti

$$\mathbf{C} = \frac{2}{3} \begin{bmatrix} c_1 & -c_2^* & c_3^* & 0 \\ c_2 & c_1^* & 0 & c_3^* \\ c_3 & 0 & -c_1^* & -c_2^* \end{bmatrix}.$$

The spatial multiplexing rate r_s is equal to $3/4$.

Orthogonal Space-Time Block Codes

Proposition

O-STBCs enjoy the decoupling property.

Example

Assume a MISO transmission based on the Alamouti code

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} = \sqrt{\frac{E_s}{2}} \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix} + \begin{bmatrix} n_1 & n_2 \end{bmatrix}$$

or equivalently

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \sqrt{\frac{E_s}{2}} \underbrace{\begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}}_{\mathbf{H}_{eff}} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}.$$

Orthogonal Space-Time Block Codes

Example

Applying the space-time matched filter $\mathbf{H}_{eff}^{H}$ to the received vector decouples the transmitted symbols

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{H}_{eff}^{H} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \sqrt{\frac{E_s}{2}} \left[|h_1|^2 + |h_2|^2 \right] \mathbf{I}_2 \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \mathbf{H}_{eff}^{H} \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}.$$

Expanding the original ML metric

$$\left| y_1 - \sqrt{\frac{E_s}{2}} \left(h_1 c_1 + h_2 c_2 \right) \right|^2 + \left| y_2 - \sqrt{\frac{E_s}{2}} \left(-h_1 c_2^* + h_2 c_1^* \right) \right|^2$$

and making use of z_1 and z_2, the decision metric for c_1 is

choose c_i iff $\left| z_1 - \sqrt{\frac{E_s}{2}} \left(|h_1|^2 + |h_2|^2 \right) c_i \right|^2 \leq \left| z_1 - \sqrt{\frac{E_s}{2}} \left(|h_1|^2 + |h_2|^2 \right) c_k \right|^2 \; \forall i \neq k$

and similarly for c_2. Independent decoding of symbols c_1 and c_2 is so performed.

Orthogonal Space-Time Block Codes

- Error Probability

$$P\left(\mathbf{C} \rightarrow \mathbf{E}\right) \overset{(\rho \nearrow)}{\leq} \left(\frac{\rho}{4}\frac{T}{Qn_t}\right)^{-n_r n_t} \left(\sum_{q=1}^{Q} \left|c_q - e_q\right|^2\right)^{-n_r n_t}.$$

Full diversity gain of $n_t n_r$.

Orthogonal Space-Time Block Codes

- O-STBCs are not capacity efficient $\mathcal{I}_{O-STBC}\left(\mathbf{H}\right) \leq \mathcal{I}_e\left(\mathbf{H}\right)$
  - mutual information of MIMO channel

$$\mathcal{I}_e\left(\mathbf{H}\right) = \log_2\left(1 + \frac{\rho}{n_t}\left\|\mathbf{H}\right\|_F^2 + \ldots + \left(\frac{\rho}{n_t}\right)^{r(\mathbf{H})}\prod_{k=1}^{r(\mathbf{H})}\lambda_k\left(\mathbf{H}\mathbf{H}^H\right)\right).$$

  - mutual information of MIMO channel transformed by the O-STBC

$$\mathcal{I}_{O-STBC}\left(\mathbf{H}\right) = \frac{Q}{T}\log_2\left(1 + \frac{\rho T}{Q n_t}\left\|\mathbf{H}\right\|_F^2\right).$$

Proposition

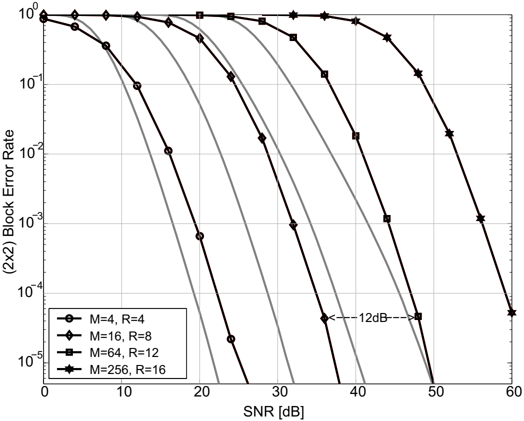
For a given channel realization $\mathbf{H}$, the mutual information achieved by any O-STBC is always upper-bounded by the channel mutual information with equal power allocation $\mathcal{I}_e$. Equality occurs if and only if both the rank of the channel and the spatial multiplexing rate of the code are equal to one.

Corollary

The Alamouti scheme is optimal with respect to the mutual information when used with only one receive antenna.

Orthogonal Space-Time Block Codes

- Block error rate for 4 different rates $R = 4, 8, 12, 16$ bits/s/Hz in 2×2 i.i.d. slow Rayleigh fading channels.



  - full diversity exploited: $g_d\,(g_s = 0, \infty) = g_d^o\,(\infty) = 4$.
  - the growth of the multiplexing gain is slow: 12 dB separate the curves, corresponding to a multiplexing gain $g_s = 1$, i.e., 1 bit/s/Hz increase per 3 dB SNR increase.

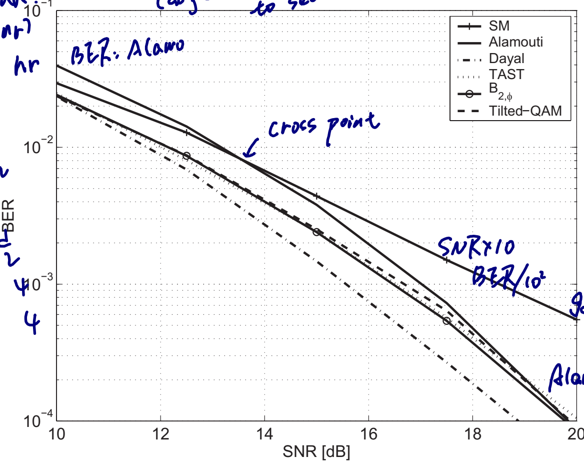
Global Performance Comparison

- Bit error rate (BER) of several space-time block codes in i.i.d. slow Rayleigh fading channels with $n_t = 2$ and $n_r = 2$ in a 4-bit/s/Hz transmission. ML decoding is used.

full-rate full-rank: → complex Rx: codebook larger to search.

rank = min(nt.nr)

gd: min(nt.nr) nr



BER: Alamo

cross point

SNR×10 BER/10²

gd=1

Alamouti: gd=4 SNR×10 BER/10⁴

nt=2 [C] T=2

Q=4 gd= nr·2

T=2

Rs=2

Rs d

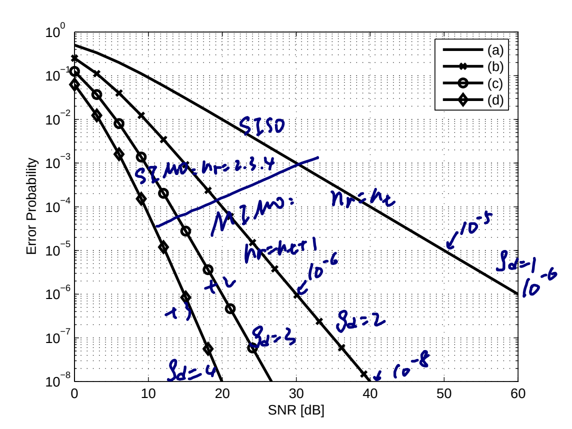
SM(ML) 2 2

Alamouti: 1 4

full-rate full-rank 2 4

Example

Question: Here is the average Error Probability of one scheme (i.e., one transmission and reception strategy) vs. SNR for point-to-point channels with i.i.d. Rayleigh slow fading and four different antenna configurations (a) to (d). The CSI is unknown to the transmitter.



Example

Question: What is the diversity gain (at high SNR) achieved by that scheme in each antenna configuration? Provide your reasoning.

Answer: The diversity gain is the slope at high SNR of the error curve vs. the SNR on a log-log scale, i.e. $-\lim_{\rho\to\infty}\frac{\log(P_e(\rho))}{\log(\rho)}$ with ρ being the SNR.
For (a), the diversity gain is 1 as the error rate decreases by 10^{-1} when the SNR is increased from 50dB to 60dB.
For (b), the diversity gain is 2 as the error rate decreases by 10^{-2} when the SNR is increased from 30dB to 40dB.
For (c), the diversity gain is 3 as the error rate decreases by 10^{-3} when the SNR is increased from roughly 26dB to 36dB.
For (d), the diversity gain is 4 as the error rate decreases by 10^{-4} when the SNR is increased from 10dB to 20dB.

Example

Question: For each scenario (a) to (d), identify an antenna configuration (i.e., n_t and n_r) and the corresponding transmission/reception strategy that can achieve such diversity gain. Provide your reasoning.

Answer: The simplest approach is to perform
for (a), receive matched filter with $n_r = 1$, $n_t = 1$
for (b), receive matched filter with $n_r = 2$, $n_t = 1$
for (c), receive matched filter with $n_r = 3$, $n_t = 1$
for (d), receive matched filter with $n_r = 4$, $n_t = 1$

Alternative strategies are possible, for instance selection combining at the receiver for all 4 cases. We could also perform transmit diversity based on space-time coding and use O-STBC for (b),(c),(d) to achieve diversity order of 2 with $n_t = 2$ and $n_r = 1$, 3 with $n_t = 3$ and $n_r = 1$, 4 with $n_t = 4$ and $n_r = 1$, respectively. Alternatively, we could as well use Spatial Multiplexing with ZF receiver and transmit two streams over two transmit antennas and use 2,3,4,5 receive antennas for a,b,c,d respectively.

Quantized Precoding

Limited Feedback

- full CSIT
  - array and diversity gain
  - lower system complexity (parallel virtual transmissions)
  - hardly achievable (especially when the channel varies rapidly), costly in terms of feedback
- Exploiting a Limited Amount of Feedback at the Transmitter
  - codebook of precoding matrices, i.e., a finite set of precoders, designed off-line and known to both the transmitter and receiver.
  - The receiver estimates the best precoder as a function of the current channel and feeds back only the index of this best precoder in the codebook.

precoder (known at Tx/Rx)

$C = \{w_1, w_2, \dots, w_N\}$

① estimate CSI
② find optimum beamforming vector in C
③ send back $\log_2 N$ bits indicating which vector to use

$BS \underset{DL}{\overset{UL}{\rightleftarrows}} UE$

f | DL | UL | (t_1, t_2) TDD $\rightarrow h_{DL} = h_{UL}$ (every 10 ms. consume resources)

t; f_1 DL, f_2 UL FDD $\longrightarrow h_{DL} \neq h_{UL}$ (limited feedback)

① pilot
② quantised CSI

Quantized Precoding: Dominant Eigenmode Transmission

- Assume dominant eigenmode transmission (i.e. beamforming)

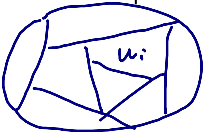
$n_r \times n_t \; n_t \times 1$
$y = Hwc + n$
$z = gy = gHwc + gn, \; g = Hw^H$
need H at TX

$$
\begin{aligned}
\mathbf{y} &= \sqrt{E_s}\mathbf{H}\mathbf{w}c + \mathbf{n}, \\
z &= \mathbf{g}^H \mathbf{y}, \\
&= \sqrt{E_s}\mathbf{g}^H \mathbf{H}\mathbf{w}c + \mathbf{g}^H \mathbf{n}
\end{aligned}
$$

  where $\mathbf{g}$ and $\mathbf{w}$ are respectively $n_r \times 1$ and $n_t \times 1$ vectors.
- Assuming MRC, the optimal beamforming vector $\mathbf{w}$ that maximizes the SNR is given by

$$
\mathbf{w}^{\star} = \arg \max_{\mathbf{w} \in \mathcal{C}_w} \|\mathbf{H}\mathbf{w}\|^2
$$

  with $\mathcal{C}_w$ set of unit-norm vectors. The best precoder is the dominant right singular vector of $\mathbf{H}$.
- Reduce the number of feedback bits: limit the space $\mathcal{C}_w$ over which $\mathbf{w}$ can be chosen to a codebook called $\mathcal{W}$. The receiver evaluates the best precoder $\mathbf{w}^{\star}$ among all unit-norm precoders $\mathbf{w}_i \in \mathcal{W}$ (with $i = 1, \ldots, n_p$) such that



set of H
$\Downarrow$ map
Codebook$\{w_i\}$

$$
\mathbf{w}^{\star} = \arg \max_{\substack{1 \le i \le n_p \\ \mathbf{w}_i \in \mathcal{W}}} \|\mathbf{H}\mathbf{w}_i\|^2 .
$$

MIMO: $y = \underline{\underline{HW}}c + \cdots$

Codebook 1 (DFT) $\{w_1, w_2 \ldots w_{16}\}$ $S_1 = P$ [$n_e = 1$]

Codebook 2 (MET) $\{W_1, W_2 \ldots W_{16}\}$ $S_1 = \frac{P}{2}$ $S_2 = \frac{P}{2}$ (no WF) [$n_e = 2$]

- Extension to other kinds of channel models (e.g. spatial/time correlation, polarization), transmission strategies (e.g. O-STBCs, SM), reception strategies (e.g. MRC, ZF, MMSE, ML), criteria (e.g. error rate or transmission rate)
- Quantized precoding for SM with rank adaptation and rate maximization

$$\mathbf{W}^\star = \arg\max_{n_e} \max_{\mathbf{W}_i^{(n_e)} \in \mathcal{W}_{n_e}} R.$$

The codebooks $\mathcal{W}_{n_e}$ are defined for ranks $n_e = 1, \ldots, \min\{n_t, n_r\}$. Rate is computed on the equivalent precoded channel $\mathbf{H}\mathbf{W}_i^{(n_e)}$. (effective channel)

- Uniform power allocation and joint ML decoding

$$R = \log_2 \det\left[\mathbf{I}_{n_e} + \frac{\rho}{n_e}\left(\mathbf{W}_i^{(n_e)}\right)^H \mathbf{H}^H \mathbf{H}\mathbf{W}_i^{(n_e)}\right].$$

- With other types of receivers/combiner

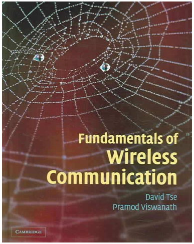
$$R = \sum_{q=1}^{n_e} \log_2\left(1 + \rho_q\left(\mathbf{H}\mathbf{W}_i^{(n_e)}\right)\right).$$

where ρ_q is the SINR of stream q on at the output of the combiner for the equivalent channel $\mathbf{H}\mathbf{W}_i^{(n_e)}$.

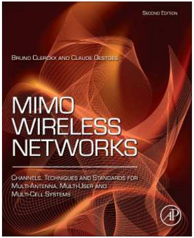
Part 3: Multiuser Communications

Capacity of Multiuser Channels
Fairness, Scheduling and Precoding
Multiuser Multicell Communications

Reference Book



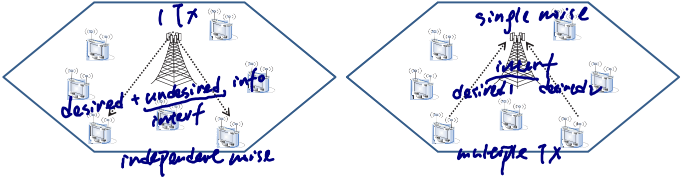
Chapter 12 - Sections 12.1, 12.2.1, 12.3.1, 12.5.1, 12.5.2 Chapter 13 - Section 13.2



Capacity of Multiuser Channels

BC vs MAC

- So far, we looked at a single link/user. Most systems are multi-user!
- How to deal with multiple users?
- Broadcast Channel (BC) and Multiple Access Channel(MAC)



(a) Broadcast Channel - Downlink (b) Multiple Access Channel - Uplink

Differences between BC and MAC:

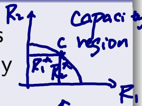
- there are multiple independent receivers (and therefore multiple independent additive noises) in BC while there is a single receiver (and therefore a single noise term) in MAC.
- there is a single transmitter (and therefore a single transmit power constraint) in BC while there are multiple transmitters (and therefore multiple transmit power constraints) in MAC.
- the desired signal and the interference (originating from the co-scheduled signals) propagate through the same channel in the BC while they propagate through different channels in the MAC.

- In a multi-user setup, given that all users share the same spectrum, the rate achievable by a given user q, denoted as R_q, will depend on the rate of the other users R_p, $p \neq q$ $\rightarrow$ Trade-off between rates achievable by different users!
- The capacity region $\mathcal{C}$ formulates this trade-off by expressing the set of all user rates $(R_1, \ldots, R_K)$ that are simultaneously achievable.

Definition

The capacity region $\mathcal{C}$ of a channel $\mathbf{H}_{ul}$ is the set of all rate vectors $(R_1, \ldots, R_K)$ such that simultaneously user 1 to user K can reliably communicate at rate R_1 to rate R_K, respectively.

rates achievable for independent users



Any rate vector not in the capacity region is not achievable (i.e. transmission at those rates will lead to errors).

Definition

The sum-rate capacity C of a capacity region $\mathcal{C}$ is the maximum achievable sum of rates

may not fair

$$C = \max_{(R_1, \ldots, R_K) \in \mathcal{C}} \sum_{q=1}^{K} R_q.$$

SISO MAC System Model

- Uplink multi-user SISO transmission
  - total number of K users ($\mathcal{K} = \{1, \ldots, K\}$) distributed in a cell,
  - 1 transmit antennas at each mobile terminal
  - 1 receive antenna at the base station
- Received signal (we drop the time dimension)

Tx1
Rx
Tx2

path loss can be different

$$y_{ul} = \sum_{q=1}^{K} \Lambda_q^{-1/2} h_{ul,q} c'_{ul,q} + n_{ul}$$

where

- $h_{ul,q}$ models the small scale time-varying fading process and Λ_q^{-1} refers to the large-scale fading accounting for path loss and shadowing
- n_{ul} is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_n^2\right)$.

- Power constraint: $\mathcal{E}\left\{|c'_{ul,q}|^2\right\} \leq E_{s,q}$.

user-dependent

SISO MAC System Model

- By stacking up the transmit signal vectors and the channel matrices of all K users,

$$\mathbf{c}'_{ul} = \left[c'_{ul,1}, \ldots, c'_{ul,K}\right]^T,$$
$$\mathbf{h}_{ul} = \left[\Lambda_1^{-1/2} h_{ul,1}, \ldots, \Lambda_K^{-1/2} h_{ul,K}\right],$$

the system model also writes as

$$\overset{1}{y_{ul}} = \overset{1\times K}{\mathbf{h}_{ul}} \overset{K\times 1}{\mathbf{c}'_{ul}} + \overset{1}{n_{ul}}.$$

$E_s \mathcal{E}\{\Lambda^{-1}|h|^2\}$

$E_s \Lambda^{-1} \underbrace{\mathcal{E}\{|h|^2\}}_{1}$

$y = \Lambda^{-1} h c + n$ $\quad$ $y = \sqrt{E_s} h c + n$

$c \to E_s$ (power) $\quad$ $c \to 1$ (normalised)

different user can have different Λ

- Long term SNR of user q defined as $\eta_q = E_{s,q}\Lambda_q^{-1}/\sigma_n^2$.
- Note on the notations: the dependence on the path loss and shadowing is made explicit in order to stress that the co-scheduled users experience different path losses and shadowings and therefore receive power.
- The receiver (i.e. the BS in a UL scenario) has always perfect knowledge of the CSI.

Capacity Region of a Two-User SISO MAC

- The capacity region is a pentagon with two corner points A and B.

$C = R_1 + R_2$

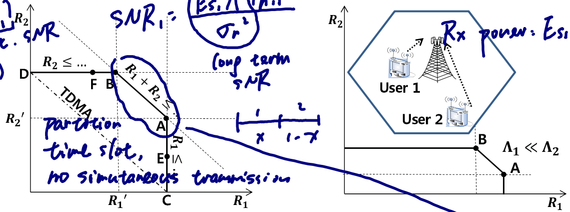
$R_1 \le \log_2(1+SNR_1)$ inst. SNR

$R_2 \le \log_2(1+SNR_2)$

$SNR_1 = \frac{E_{s,1}\Lambda_1^{-1}|h_1|^2}{\sigma_n^2}$ long term SNR

partition time slot, no simultaneous transmission

Rx power: $E_{s,1}\Lambda_1^{-1}|h_1|^2 + E_{s,2}\Lambda_2^{-1}|h_2|^2$



(a) Two-user MIMO MAC rate region for fixed $Q_{ul,1}$ and $Q_{ul,2}$

(b) Rate regions with various path losses

decode usr 1 first | decode usr 2 first

x | $1-x$

- $\mathcal{C}_{MAC}$ is the set of all rates pair (R_1, R_2) satisfying to

shape ~ long term SNR

$$R_q \le \log_2\left(1 + \eta_q \left|h_{ul,q}\right|^2\right), q = 1, 2$$ Tx

$$R_1 + R_2 \le \log_2\left(1 + \eta_1 \left|h_{ul,1}\right|^2 + \eta_2 \left|h_{ul,2}\right|^2\right).$$ Rx

- Remarkably, at point A, user 1 can transmit at a rate equal to its single-link SISO rate and user 2 can simultaneously transmit at a rate $R_2' > 0$ equal to by SIC.

first decode 2 → interf-free for 1.

$$R_2' = \log_2\left(1 + \eta_1 \left|h_{ul,1}\right|^2 + \eta_2 \left|h_{ul,2}\right|^2\right) - \log_2\left(1 + \eta_1 \left|h_{ul,1}\right|^2\right)$$

$$= \log_2\left(1 + \frac{\eta_2 \left|h_{ul,2}\right|^2}{1 + \eta_1 \left|h_{ul,1}\right|^2}\right) = \log_2\left(1 + \frac{\Lambda_2^{-1}\left|h_{ul,2}\right|^2 E_{s,2}}{\sigma_n^2 + \Lambda_1^{-1}\left|h_{ul,1}\right|^2 E_{s,1}}\right).$$

Rx power 2

interf.

noise + Rx power 1

Capacity Region of SISO MAC

Proposition

$$\mathcal{C}_{MAC} = \{ (R_1, \dots, R_K) : \sum_{q \in S} R_q \leq \log_2 \left(1 + \sum_{q \in S} \eta_q \left|h_{ul,q}\right|^2 \right), \forall S \subseteq \mathcal{K} \}$$

Subset!

where $\eta_q = \Lambda_q^{-1} E_{s,q} / \sigma_n^2$.

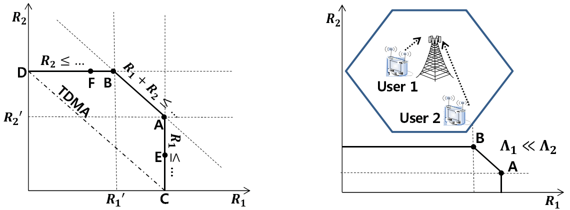
$R_1 \leq \cdots$
$R_2 \leq \cdots$
$R_3 \leq \cdots$

$R_1 + R_2 \leq \cdots$
$R_2 + R_3 \leq \cdots$
$R_3 + R_1 \leq \cdots$

$R_1 + R_2 + R_3 \leq \cdots$

Achievability of the Capacity Region

- SIC is optimal for achieving the corner points of the SISO MAC rate region.
- The exact corner point that is achieved on the capacity region depends on the stream cancellation ordering:
  - Point A, user 2 is canceled first (i.e. all streams from user 2) such that user 1 is left with the Gaussian noise and can achieve a rate equal to the single-link bound.
  - Assuming $n_t = 1$, $R_2' = \log_2(1 + \rho_q)$ where ρ_q is the SINR of the receiver for user 2's stream treating user 1's stream as Gaussian interference.

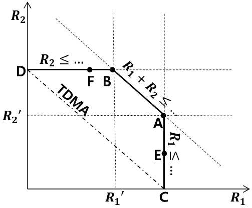


(a) Two-user MIMO MAC rate region for fixed $Q_{ul,1}$ and $Q_{ul,2}$

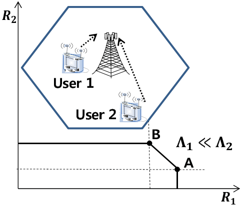
(b) Rate regions with various path losses

Comparisons with TDMA

- TDMA allocates the time resources in an orthogonal manner such that users are never transmitting at the same time (line D-C in the rate region).
- TDMA rate region is strictly smaller than the one achievable with SIC.



(a) Two-user MIMO MAC rate region for fixed $Q_{ul,1}$ and $Q_{ul,2}$



(b) Rate regions with various path losses

SISO BC System Model

- Downlink multi-user SISO transmission
  - total number of K users ($\mathcal{K} = \{1, \ldots, K\}$) distributed in a cell,
  - 1 receive antenna at each mobile terminal
  - 1 transmit antenna at the base station
- Received signal (we drop the time dimension)

$$y_q = \Lambda_q^{-1/2} h_q c' + n_q$$

different users

  where
  - h_q models the small scale time-varying fading process and Λ_q^{-1} refers to the large-scale fading accounting for path loss and shadowing
  - n_q is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_{n,q}^2\right)$.
- Transmit power constraint: $\mathcal{E}\left\{|c'|^2\right\} = E_s$.

only one constraint for the combined symbol

SISO BC System Model

- By stacking up the received signal vectors, the noise vectors and the channel matrices of all K users,

$$\mathbf{y} = [y_1, \ldots, y_K]^T,$$
$$\mathbf{n} = [n_1, \ldots, n_K]^T,$$
$$\mathbf{h} = \left[\Lambda_1^{-1/2} h_1, \ldots, \Lambda_K^{-1/2} h_K\right]^T,$$

the system model also writes as

$$\mathbf{y} = \mathbf{h}c' + \mathbf{n}.$$

antennas of different users.

- SNR of user q defined as $\eta_q = E_s \Lambda_q^{-1} / \sigma_{n,q}^2$. cannot be utilised together (MRC)
- Perfect instantaneous channel state information (CSI) at the Tx and all Rx.
- Generally speaking, c' is written as the superposition of statistically independent signals c'_q

$$c' = \sum_{q=1}^{K} c'_q.$$

Capacity Region of two-user Deterministic SISO BC

- In *two-user* SISO MAC, point A was obtained by canceling user 2's signal first such that user 1 is left with Gaussian noise.

capacity is strong enough to decode usr 2.

- Let us apply the same philosophy to the SISO BC:
  - transmit $c' = c'_1 + c'_2$, with power of c'_q denoted as s_q
  - user 1 cancels user 2's signal c'_2 so as to be left with its own Gaussian noise
  - user 2 decodes its signal by treating user 1's signal c'_1 as Gaussian noise.
- Achievable rates of such strategy (with sum-power constraint $s_1 + s_2 = E_s$)

s_1 s_2
[c_1 c_2]
[x]

└ Rx1 decode usr 2 first → no interf $R_1 = \log_2\left(1 + \frac{\Lambda^{-1} s_1}{\sigma_n^2} |h_1|^2\right)$

└ Rx2 decode usr2 first → interf as noise $R_2 = \left(\log_2\left(1 + \frac{\Lambda_2^{-1}|h_2|^2 s_2}{\sigma_{n,2}^2 + \Lambda_2^{-1}|h_2|^2 s_1}\right)\right)$

no privacy issue: encrypted symbol

$$R_1 = \log_2\left(1 + \frac{\Lambda_1^{-1} s_1}{\sigma_{n,1}^2} |h_1|^2\right)$$

$$R_2 = \log_2\left(1 + \frac{\Lambda_2^{-1} |h_2|^2 s_2}{\sigma_{n,2}^2 + \Lambda_2^{-1} |h_2|^2 s_1}\right)$$

. UE2 decode c_2

s_1 travel through same (2) channel with s_2.

- *Careful!* For user 1 to be able to correctly cancel user 2's signal, user 1's channel has to be good enough to support R_2, i.e.

UE1 step 1 decode UE2 $c_1 . c_2$ step 2 decode UE1 c_1

$$R_2 \le \log_2\left(1 + \frac{\Lambda_1^{-1} |h_1|^2 s_2}{\sigma_{n,1}^2 + \Lambda_1^{-1} |h_1|^2 s_1}\right)$$

. UE1 decode c_2

treat c_1 as interf c_2 as desired

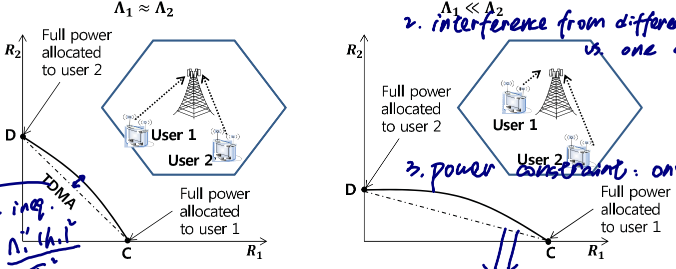
- The channel gains normalized w.r.t. their respective noise power should be ordered

UE2 step 1 decode UE2 treat c_1 as noise. (UE2 decoded by both) channel 2 channel 1 strong user can decode other user

$$\frac{\Lambda_2^{-1} |h_2|^2}{\sigma_{n,2}^2} \le \frac{\Lambda_1^{-1} |h_1|^2}{\sigma_{n,1}^2}.$$

- If the ordering condition is satisfied, the above strategy achieves the boundary of the capacity region of the two-user SISO BC for any power allocation s_1 and s_2 satisfying $s_1 + s_2 = E_s$.
- The capacity region is given by the union of all rate pairs (R_1, R_2) over all power allocations s_1 and s_2 satisfying $s_1 + s_2 = E_s$. BC vs MAC:

1. multiple noise vs. one noise
2. interference from different channel vs. one channel
3. power constraint: one vs. many



(a) (b)

gain from the ineq. $\frac{\Lambda_0^{-1}|h_0|^2}{\sigma_{n_0}^2} \le \frac{\Lambda_1^{-1}|h_1|^2}{\sigma_{n_1}^2}$

BC: all to the best

MAC: simutananeously

Capacity Region of K-user Deterministic SISO BC

- Define $\mathsf{h}_q = \Lambda_q^{-1/2} h_q / \sigma_{n,q}$. Assume $|\mathsf{h}_1|^2 \geq |\mathsf{h}_2|^2 \geq \ldots \geq |\mathsf{h}_K|^2$.

Proposition

With the ordering $|\mathsf{h}_1|^2 \geq |\mathsf{h}_2|^2 \geq \ldots \geq |\mathsf{h}_K|^2$, the capacity region $\mathcal{C}_{BC}$ of the Gaussian SISO BC is the set of all achievable rate vectors $(R_1, \ldots, R_K)$ given by

$$\bigcup_{s_q : \sum_{q=1}^{K} s_q = E_s} \left\{ (R_1, \ldots, R_K) : R_q \leq \log_2 \left(1 + \frac{|\mathsf{h}_q|^2 s_q}{1 + |\mathsf{h}_q|^2 \left[\sum_{p=1}^{q-1} s_p \right]} \right), \forall q \right\}.$$

interf from usr 1 ... q-1

Proposition

The sum-rate capacity of the SISO BC is achieved by allocating the transmit power to the strongest user

channel 1

channel 2

no interf.

$$\bar{C}_{BC} = \log_2 \left(1 + E_s \max_{q=1,\ldots,K} |\mathsf{h}_q|^2 \right) = \log_2 \left(1 + \max_{q=1,\ldots,K} \eta_q |h_q|^2 \right).$$

R_2 R_1 R_1+R_2

Recall that the MAC sum-rate capacity is obtained with all users simultaneously transmitting at their respective full power.

Achievability of the SISO BC Capacity Region

- *Superposition coding with SIC and appropriate ordering*
- User ordering: decode and cancel out weaker users signals before decoding their own signal.
- The weakest user decodes only the coarsest constellation. The strongest user decodes and subtracts all constellation points in order to decode the finest constellation.

Achievability of the SISO BC Capacity Region

Proposition

With the appropriate cancellation/encoding ordering, Superposition Coding with SIC is optimal for achieving the SISO BC capacity region.

Proposition

The SISO BC sum-rate capacity is achievable with dynamic TDMA (to the strongest user) and Superposition Coding with SIC (with the appropriate cancellation ordering).

Comparisons with TDMA

- Similarly to MAC, TDMA rate region is contained in the BC capacity region.
- The gap between the BC capacity region and the TDMA rate region increases proportionally with the asymmetry between users normalized channel gains.
- TDMA achieves the sum-rate capacity of SISO BC.

MIMO MAC System Model

- Uplink multi-user MIMO (MU-MIMO) transmission
  - total number of K users ($\mathcal{K} = \{1, \ldots, K\}$) distributed in a cell,
  - $n_{t,q}$ transmit antennas at mobile terminal q (we simply drop the index q and write n_t if $n_{t,q} = n_t \ \forall q$)
  - n_r receive antenna at the base station
- Received signal (we drop the time dimension)

$$\mathbf{y}_{ul} = \sum_{q=1}^{K} \Lambda_q^{-1/2} \mathbf{H}_{ul,q} \mathbf{c}'_{ul,q} + \mathbf{n}_{ul}$$

where
  - $\mathbf{y}_{ul}$ $[n_r \times 1]$
  - $\mathbf{H}_{ul,q}$ $[n_r \times n_{t,q}]$ models the small scale time-varying fading process and Λ_q^{-1} refers to the large-scale fading accounting for path loss and shadowing
  - $\mathbf{n}_{ul}$ is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_n^2 \mathbf{I}_{n_r}\right)$.
- User q's input covariance matrix is defined as the covariance matrix of the transmit signal of user q as $\mathbf{Q}_{ul,q} = \mathcal{E}\left\{\mathbf{c}'_{ul,q} {\mathbf{c}'_{ul,q}}^H\right\}$.
- Power constraint: $\mathrm{Tr}\{\mathbf{Q}_{ul,q}\} \leq E_{s,q}$.

MIMO MAC System Model

- By stacking up the transmit signal vectors and the channel matrices of all K users,

$$\begin{aligned}
\mathbf{c}'_{ul} &= \left[\mathbf{c}'^{T}_{ul,1}, \ldots, \mathbf{c}'^{T}_{ul,K}\right]^{T}, \\
\mathbf{H}_{ul} &= \left[\Lambda_1^{-1/2}\mathbf{H}_{ul,1}, \ldots, \Lambda_K^{-1/2}\mathbf{H}_{ul,K}\right], \\
\mathbf{y}_{ul} &= \mathbf{H}_{ul}\mathbf{c}'_{ul} + \mathbf{n}_{ul},
\end{aligned}$$

  $\mathbf{H}_{ul}$ is assumed to be full-rank as it would be the case in a typical user deployment.
- Long term SNR of user q defined as $\eta_q = E_{s,q}\Lambda_q^{-1}/\sigma_n^2$.
- Note on the notations: the dependence on the path loss and shadowing is made explicit in order to stress that the co-scheduled users experience different path losses and shadowings and therefore receive power.
- We assume that the receiver (i.e. the BS in a UL scenario) has always perfect knowledge of the CSI, but we will consider strategies where the transmitters have perfect or partial knowledge of the CSI.
- Capacity region of MIMO MAC can be characterized.

MIMO BC System Model

- Downlink multi-user MIMO (MU-MIMO) transmission
  - total number of K users ($\mathcal{K} = \{1, \ldots, K\}$) distributed in a cell,
  - $n_{r,q}$ receive antennas at mobile terminal q (we simply drop the index q and write n_r if $n_{r,q} = n_r \ \forall q$)
  - n_t transmit antenna at the base station
- Received signal (we drop the time dimension)

$$\mathbf{y}_q = \Lambda_q^{-1/2} \mathbf{H}_q \mathbf{c}' + \mathbf{n}_q$$

  where
  - $\mathbf{y}_q [n_{r,q} \times 1]$
  - $\mathbf{H}_q \ [n_{r,q} \times n_t]$ models the small scale time-varying fading process and Λ_q^{-1} refers to the large-scale fading accounting for path loss and shadowing
  - $\mathbf{n}_q$ is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_{n,q}^2 \mathbf{I}_{n_{r,q}}\right)$.
- The input covariance matrix is defined as the covariance matrix of the transmit signal as $\mathbf{Q} = \mathcal{E}\left\{\mathbf{c}'\mathbf{c}'^H\right\}$.
- Power constraint: Tr$\{\mathbf{Q}\} \leq E_s$.

MIMO BC System Model

- By stacking up the received signal vectors, the noise vectors and the channel matrices of all K users,

$$\mathbf{y} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_K^T\right]^T,$$
$$\mathbf{n} = \left[\mathbf{n}_1^T, \ldots, \mathbf{n}_K^T\right]^T,$$
$$\mathbf{H} = \left[\Lambda_1^{-1/2}\mathbf{H}_1^T, \ldots, \Lambda_K^{-1/2}\mathbf{H}_K^T\right]^T,$$
$$\mathbf{y} = \mathbf{H}\mathbf{c}' + \mathbf{n},$$

  $\mathbf{H}$ is assumed to be full-rank as it would be the case in a typical user deployment.
- SNR of user q defined as $\eta_q = E_s\Lambda_q^{-1}/\sigma_{n,q}^2$.
- Perfect instantaneous channel state information (CSI) at the Tx and all Rx.
- Generally speaking, $\mathbf{c}'$ is written as the superposition of statistically independent signals $\mathbf{c}_q'$

$$\mathbf{c}' = \sum_{q=1}^{K} \mathbf{c}_q'.$$

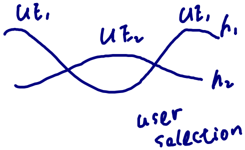
  The input covariance matrix of user q is defined as $\mathbf{Q}_q = \mathcal{E}\left\{\mathbf{c}_q'\mathbf{c}_q'^H\right\}$.
- Capacity region of MIMO BC can be characterized.

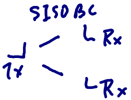
Fairness, Scheduling and Precoding

deliver to single/ multiple user ?

Multi-User Diversity

- In single-link systems, channel fading is viewed as a source of unreliability mitigated through diversity techniques (e.g. space-time coding).
- In multi-user communications, fading is viewed as a source of randomization that can be exploited!
- Multi-User (MU) diversity is a form of selection diversity among users provided by independent time-varying channels across the different users.
- Provided that the BS is able to track the user channel fluctuations (based on feedback), it can schedule transmissions to the users with favorable channel fading conditions, i.e. near their peaks, to improve the total cell throughput.
- Recall that MU diversity was already identified as part of the sum-rate maximization in SISO BC.





more users => higher capacity
$g_m \sim \log K$

Multi-User Diversity Gain

- Assume that the fading distribution of the K users are independent and identically ($\Lambda_q^{-1} = \Lambda^{-1}$ and channel gains h_q are drawn from the same) Rayleigh distributed and that users experience the same average SNR $\eta_q = \eta$ ($\sigma_{n,q}^2 = \sigma_n^2$) $\forall q$:

$$\mathbf{y}_q = \Lambda^{-1/2} h_q c' + \mathbf{n}_q.$$

- Assume MU-SISO where one user is scheduled at a time in a TDMA manner: select the user with the largest channel gain.
- Mathematically same as antenna selection diversity.
- Average SNR gain
  - Average SNR after user selection $\bar{\rho}_{out}$

amplitude: Rayleigh

i.i.d. Rayleigh max = $\Sigma \frac{1}{q}$

$$\bar{\rho}_{out} = \mathcal{E}\left\{\eta \max_{q=1,\ldots,K} |h_q|^2\right\} = \eta \sum_{q=1}^{K} \frac{1}{q}.$$

i.i.d.

  - SNR gain provided by MU diversity g_m

$$g_m = \frac{\bar{\rho}_{out}}{\eta} = \sum_{q=1}^{K} \frac{1}{q} \overset{K\to\infty}{\cong} \log(K).$$

~ number of receive antennas

g_m is of the order of $\log(K)$ and hence the gain of the strongest user grows as $\log(K)$!

- Heavily relies on CSIT (partial or imperfect feedback impacts the performance) and independent user fading distributions (correlated fading or LOS are not good for MU diversity)

Multi-User Diversity Gain

- Sum-rate capacity

$$\bar{C}_{TDMA} = \mathcal{E}\left\{C_{TDMA}\right\} = \mathcal{E}\left\{\log_2\left(1 + \eta \max_{q=1,\ldots,K} |h_q|^2\right)\right\}.$$

  - low SNR

$$\bar{C}_{TDMA} \approx \mathcal{E}\left\{\max_{q=1,\ldots,K} |h_q|^2\right\} \eta \log_2(e) \approx g_m C_{awgn}.$$

    Observations: capacity of the fading channel $\log(K)$ times larger than the AWGN capacity.

  - high SNR (Use Jensen's inequality: $\mathcal{E}_x\left\{\mathcal{F}(x)\right\} \leq \mathcal{F}\left(\mathcal{E}_x\left\{x\right\}\right)$ if $\mathcal{F}$ concave)

$$\begin{aligned}
\bar{C}_{TDMA} &\approx \log_2(\eta) + \mathcal{E}\left\{\log_2\left(\max_{q=1,\ldots,K} |h_q|^2\right)\right\}, \\
&\approx C_{awgn} + \mathcal{E}\left\{\log_2\left(\max_{q=1,\ldots,K} |h_q|^2\right)\right\}, \\
&\overset{(a)}{\leq} C_{awgn} + \log_2\left(\mathcal{E}\left\{\max_{q=1,\ldots,K} |h_q|^2\right\}\right), \\
&= C_{awgn} + \log_2(g_m).
\end{aligned}$$

    Observations: capacity of a fading channel is larger than the AWGN capacity by a factor roughly equal to $\log_2(g_m) \approx \log\log(K)$.

- Fading channels are significantly more useful in a multi-user setting than in a single-user setting

Multi-User Diversity

- Few fundamental differences with classical spatial/time/frequency diversity:
  - Diversity techniques, like space-time coding, mainly focus on improving reliability by decreasing the outage probability in slow fading channels. MU diversity on the other hand increases the data rate over time-varying channels.
  - Classical diversity techniques mitigate fading while MU diversity exploits fading.
  - MU diversity takes a system-level view while classical diversity approaches focus on a single-link. This system-level view becomes increasingly important as we shift from single-cell to multi-cell scenarios.

Fairness and Scheduling

- Sum-rate maximization in SISO BC: pick the strongest user. Is that fair?
- An appropriate scheduler should allocate resources (time, frequency, spatial, power) to the users in a fair manner while exploiting the MU diversity gain.
- Goal of the resource allocation strategy at the scheduler: maximize the utility metric $\mathcal{U}$.

  $R_1 = \log_2(1 + \frac{E_s |h_1|^2}{\sigma_{n1}^2})$
  $R_2 = \log_2(1 + \frac{E_s |h_2|^2}{\sigma_{n2}^2})$

$$q^\star = \arg\max_{q \in \mathcal{K}} \mathcal{U}$$

  where $q^\star$ refers to the optimum user (or more generally subset of users) to be scheduled.
- Two major kinds of resource allocation strategies:
  - *rate-maximization policy*: maximizes the sum-rate - no fairness among users
  - *fairness oriented policy*, commonly relying on a *proportional fair* (PF) metric: maximizes a weighted sum-rate and guarantees fairness among users.

  max rate: always UE1

- Those two strategies can be addressed by using two different utility metrics:

  weighted rate: UE1 > UE2

$$q^\star = \arg\max_{q \in \mathcal{K}} \underbrace{w_q R_q}_{\text{weighted rate}}$$

UE1 

  where
  - rate-maximization approach: $w_q = 1$

  average performance UE2 

  - proportional fair approach: $w_q = \frac{\gamma_q}{\bar{R}_q}$ ($\bar{R}_q$ is the long-term average rate of user q and γ_q is the Quality of Service (QoS) of each user).

Practical Proportional Fair Scheduling

- The long-term average rate $\bar{R}_q$ of user q is updated using an exponentially weighted low-pass filter such that the estimate of $\bar{R}_q$ at time $k+1$, denoted as $\bar{R}_q\,(k+1)$, is function of the long-term average rate $\bar{R}_q\,(k)$ and of the current rate $R_q(k)$ at current time instant k as outlined by

$R_q > \bar{R}_q$: $\begin{cases} \bar{R}_q(k+1) \nearrow \Rightarrow \frac{R_q(k+1)}{\bar{R}_q(k+1)} \searrow . \; k \checkmark \\ \bar{R}_q(k+1) \searrow \Rightarrow \frac{R_q(k+1)}{\bar{R}_q(k+1)} \nearrow . \; k \times \end{cases}$

need a transition from 0 to stable!

$$\bar{R}_q\,(k+1) = \begin{cases} (1 - 1/t_c)\,\bar{R}_q\,(k) + 1/t_c R_q(k), & q \text{ scheduled at time } k \\ (1 - 1/t_c)\,\bar{R}_q\,(k)\,, & q \text{ not scheduled at time } k \end{cases}$$

where t_c is the scheduling time scale. User to be scheduled at time instant k is

t_c large: $\bar{R}$ update in small scale $\Rightarrow$ tend to max rate $t_c \to \infty$: max rate

$$q^\star = \arg\max_q \gamma_q \frac{R_q(k)}{\bar{R}_q(k)}.$$

t_c small: $\bar{R}$ update in large scale $\Rightarrow$ tend to max fairness $t_c \to 0$: max fairness

- The scheduling time scale t_c is a design parameter of the system that highly influences the user fairness and the performance
  - Very large t_c: assuming all users experience identical fading statistics and have the same QoS, the PF scheduler is equivalent to the rate-maximization scheduler, i.e. users contributing to the highest sum-rate are selected.
  - Small t_c: assuming all users have the same QoS, the scheduler divides the available resources equally among users (*Round-Robin* scheduling). No MU diversity is exploited.

Proportional Fair Scheduling

t_c large: time long enough to see peaks

t_c small: no MU diversity

- Sum-rate of SISO TDMA with PF scheduling at SNR=0 dB as a function of the number of users K, the scheduling time scale t_c and the channel model

t_c medium: time not long enough for fast changing channel to see peaks

$$h_k = \epsilon h_{k-1} + \sqrt{1-\epsilon^2} n_k$$

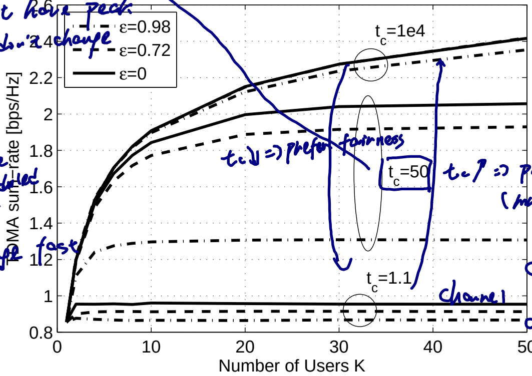
how quickly the channel change

with ϵ the channel time correlation coefficient.

$\varepsilon = 0 \Rightarrow$ channel doesn't change
$\Rightarrow$ users don't have peak
$\Rightarrow$ scheduling don't change

$\varepsilon = 0$: channel doesn't change

$\varepsilon \to 1$: fast changing channel



$t_c \downarrow \Rightarrow$ prefer fairness

$t_c \uparrow \Rightarrow$ prefer rate (multiuser div. gain)

avg rate of scheduled user

$\varepsilon \to 1 \Rightarrow$ channel change fast

$\mathcal{CN}(0,1)$
$\downarrow$ path loss

channel correlated

Precoding

- Downlink multi-user MIMO (MU-MIMO) transmission
  - total number of K users ($\mathcal{K} = \{1, \ldots, K\}$) distributed in a cell,
  - $n_{r,q}$ receive antennas at mobile terminal q (we simply drop the index q and write n_r if $n_{r,q} = n_r \ \forall q$)
  - n_t transmit antenna at the base station

MU MIMO : interf !

- Received signal (we drop the time dimension)

Tx

Rx1

Rx2

transmit at the same time.

$$\mathbf{y}_q = \Lambda_q^{-1/2} \mathbf{H}_q \mathbf{c}' + \mathbf{n}_q$$

where

- $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$
- $\mathbf{H}_q \in \mathbb{C}^{n_{r,q} \times n_t}$ models the small scale time-varying fading process and Λ_q^{-1} refers to the large-scale fading accounting for path loss and shadowing
- $\mathbf{n}_q$ is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_{n,q}^2 \mathbf{I}_{n_{r,q}}\right)$.

intercell interf.

Tx2

Tx1

Tx

Intransit interf.

- Long term SNR of user q defined as $\eta_q = E_s \Lambda_q^{-1} / \sigma_{n,q}^2$.
- Generally speaking, $\mathbf{c}'$ is written as the superposition of statistically independent signals $\mathbf{c}'_q$

$$\mathbf{c}' = \sum_{q=1}^{K} \mathbf{c}'_q.$$

- Power constraint: $\mathrm{Tr}\{\mathbf{Q}\} \leq E_s$ with $\mathbf{Q} = \mathcal{E}\left\{\mathbf{c}'\mathbf{c}'^H\right\}$.

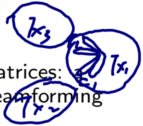
Precoding

- *scheduled user set*, denoted as $\mathbf{K} \subset \mathcal{K}$, is the set of users who are actually scheduled (with a non-zero transmit power) by the transmitter at the time instant of interest.
- The transmitter serves users belonging to $\mathbf{K}$ with n_e data streams and user $q \in \mathbf{K}$ is served with $n_{u,q}$ data streams ($n_{u,q} \leq n_e$). Hence, $n_e = \sum_{q \in \mathbf{K}} n_{u,q}$.
- Linear Precoding

c_1 [:] (2 beamformers)

$$\mathbf{c}' = \mathbf{P}\mathbf{c} = \underbrace{\mathbf{W}}_{\text{beamforming}}\underbrace{\mathbf{S}^{1/2}}_{\text{power}}\mathbf{c} = \sum_{q \in \mathbf{K}} \mathbf{P}_q \mathbf{c}_q = \sum_{q \in \mathbf{K}} \mathbf{W}_q \mathbf{S}_q^{1/2} \mathbf{c}_q$$

c_2 [:] (2 beamformers)



where
  - $\mathbf{c}$ is the symbol vector made of n_e unit-energy independent symbols.
  - $\mathbf{P} \in \mathbb{C}^{n_t \times n_e}$ is the precoder subject to $\text{Tr}\{\mathbf{P}\mathbf{P}^H\} \leq E_s$, made of two matrices: a power control diagonal matrix denoted as $\mathbf{S} \in \mathbb{R}^{n_e \times n_e}$ and a transmit beamforming matrix $\mathbf{W} \in \mathbb{C}^{n_t \times n_e}$.
  - $\mathbf{P}_q \in \mathbb{C}^{n_t \times n_{u,q}}$, $\mathbf{W}_q \in \mathbb{C}^{n_t \times n_{u,q}}$, $\mathbf{S}_q \in \mathbb{R}^{n_{u,q} \times n_{u,q}}$, and $\mathbf{c}_q \in \mathbb{C}^{n_{u,q}}$ are user q's sub-matrices and sub-vector of $\mathbf{P}$, $\mathbf{W}$, $\mathbf{S}$, and $\mathbf{c}$, respectively.
- The received signal $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$ is shaped by $\mathbf{G}_q \in \mathbb{C}^{n_{u,q} \times n_{r,q}}$ and the filtered received signal $\mathbf{z}_q \in \mathbb{C}^{n_{u,q}}$ at user q writes as

$$\begin{aligned}\mathbf{z}_q &= \mathbf{G}_q \mathbf{y}_q, \\ &= \Lambda_q^{-1/2} \mathbf{G}_q \mathbf{H}_q \mathbf{W}_q \mathbf{S}_q^{1/2} \mathbf{c}_q + \sum_{p \in \mathbf{K},\, p \neq q} \Lambda_q^{-1/2} \mathbf{G}_q \mathbf{H}_q \mathbf{W}_p \mathbf{S}_p^{1/2} \mathbf{c}_p + \mathbf{G}_q \mathbf{n}_q.\end{aligned}$$

Achievable Rate

- Maximum rate achievable by user q with linear precoding is

sum of streams (by beamformers)

$$R_q = \sum_{l=1}^{n_{u,q}} \log_2 (1 + \rho_{q,l})$$

where $\rho_{q,l}$ denotes the SINR experienced by stream l of user q

combiner precoder

$$\rho_{q,l} = \frac{\Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{p}_{q,l}\right|^2}{I_l + I_c + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2} = \frac{\Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{w}_{q,l}\right|^2 s_{q,l}}{I_l + I_c + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2}$$

with $\mathbf{p}_{q,l} = \mathbf{w}_{q,l} s_{q,l}$ (resp. $\mathbf{g}_{q,l}$) the precoder (resp. combiner) attached to stream l of user q, I_l the inter-stream interference and I_c the intra-cell interference (also called multi-user interference)

$$I_l = \sum_{m \neq l} \Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{p}_{q,m}\right|^2 = \sum_{m \neq l} \Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{w}_{q,m}\right|^2 s_{q,m},$$

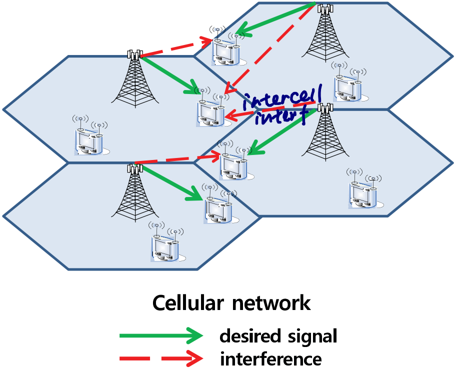
$$I_c = \sum_{\substack{p \in \mathbf{K} \\ p \neq q}} \sum_{m=1}^{n_{u,p}} \Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{p}_{p,m}\right|^2 = \sum_{\substack{p \in \mathbf{K} \\ p \neq q}} \sum_{m=1}^{n_{u,p}} \Lambda_q^{-1} \left|\mathbf{g}_{q,l} \mathbf{H}_q \mathbf{w}_{p,m}\right|^2 s_{p,m}.$$

- If $n_r = 1$, the SINR of user q simply reads as $\rho_q = \frac{\Lambda_q^{-1} \left|\mathbf{h}_q \mathbf{w}_q\right|^2 s_q}{\sum_{p \in \mathbf{K}, p \neq q} \Lambda_q^{-1} \left|\mathbf{h}_q \mathbf{w}_p\right|^2 s_p + \sigma_{n,q}^2}$.
- Various designs: MBF, ZFBF, R-ZFBF, BD, ...

Multiuser Multicell Communications

Interference

- Current wireless networks primarily operate using a frequency reuse 1 (or close to 1), i.e. all cells share the same frequency band
- Interference is not only made of intra-cell (i.e. multi-user interference), but also of inter-cell (i.e. multi-cell) interference.
- Cell edge performance is primarily affected by the inter-cell interference.

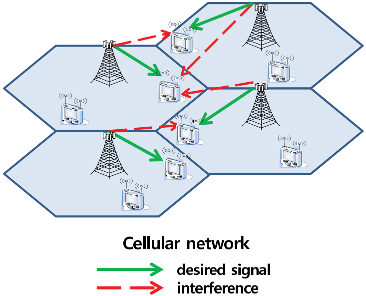


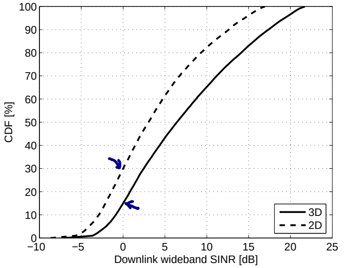
Wideband/long-term SINR

- For user q in cell i, the wideband/long-term SINR is commonly evaluated by ignoring the effect of fading but only account for path loss and shadowing

$$SINR_{w,q} = \frac{\Lambda_{q,i}^{-1} E_{s,i}}{\sigma_{n,q}^2 + \sum_{j \neq i} \Lambda_{q,j}^{-1} E_{s,j}}.$$

- Provides a rough estimate of the network performance. Function of major propagation mechanisms (path loss, shadowing, antenna radiation patterns,...), base stations deployment and user distribution.
- CDF of $SINR_{w,q}$ in a frequency reuse 1 network (cells share the same frequency band) with 2D and 3D antenna patterns in urban macro deployment.





Multi-User Multi-Cell Network

- General downlink multi-cell multi-user MIMO network with a total number of K_T users distributed in n_c cells.
- K_i users in every cell i, $n_{t,i}$ transmit antennas at BS i, $n_{r,q}$ receive antennas at mobile terminal q.
- The received signal of a given user q in cell i is

$$\mathbf{y}_q = \underbrace{\Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{c}_i'}_{\text{info of cell } i} + \underbrace{\sum_{j\neq i} \Lambda_{q,j}^{-1/2}\mathbf{H}_{q,j}\mathbf{c}_j'}_{\text{inter-cell interference}} + \mathbf{n}_q$$

where

- $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$,
- $\mathbf{n}_q$ is a complex Gaussian noise $\mathcal{CN}\left(0, \sigma_{n,q}^2 \mathbf{I}_{n_{r,q}}\right)$,
- $\Lambda_{q,i}^{-1}$ refers to the path-loss and shadowing between transmitter i and user q,
- $\mathbf{H}_{q,i} \in \mathbb{C}^{n_{r,q} \times n_{t,i}}$ models the MIMO fading channel between transmitter i and user q.

Linear Precoding

- *scheduled user set* of cell i, denoted as $\mathbf{K}_i$, as the set of users who are actually scheduled by BS i at the time instant of interest

- Transmit $n_{e,i}$ streams in each cell i using MU-MIMO linear precoding

$$\mathbf{c}'_i = \mathbf{P}_i \mathbf{c}_i = \underbrace{\mathbf{W}_i}_{\text{beamform dir.}} \underbrace{\mathbf{S}_i^{1/2}}_{\text{power}} \mathbf{c}_i = \sum_{q \in \mathbf{K}_i} \mathbf{P}_{q,i} \mathbf{c}_{q,i} = \sum_{q \in \mathbf{K}_i} \mathbf{W}_{q,i} \mathbf{S}_{q,i}^{1/2} \mathbf{c}_{q,i}$$

where

$P_i = W_i S_i^{\frac{1}{2}}$

  - $\mathbf{c}_i$ is the symbol vector made of $n_{e,i}$ unit-energy independent symbols
  - $\mathbf{P}_i \in \mathbb{C}^{n_{t,i} \times n_{e,i}}$ is the precoder made of two matrices, namely a power control diagonal matrix denoted as $\mathbf{S}_i \in \mathbb{R}^{n_{e,i} \times n_{e,i}}$ and a transmit beamforming matrix $\mathbf{W}_i \in \mathbb{C}^{n_{t,i} \times n_{e,i}}$.
  - $\mathbf{P}_{q,i} \in \mathbb{C}^{n_{t,i} \times n_{u,q}}$, $\mathbf{W}_{q,i} \in \mathbb{C}^{n_{t,i} \times n_{u,q}}$, $\mathbf{S}_{q,i} \in \mathbb{R}^{n_{u,q} \times n_{u,q}}$, and $\mathbf{c}_{q,i} \in \mathbb{C}^{n_{u,q}}$ are user q's sub-matrices and sub-vector of $\mathbf{P}_i$, $\mathbf{W}_i$, $\mathbf{S}_i$, and $\mathbf{c}_i$, respectively.
  - The input covariance matrix at cell i is $\mathbf{Q}_i = \mathcal{E}\{\mathbf{c}'_i \mathbf{c}'^H_i\}$ subject to the transmit power constraint $\text{Tr}\{\mathbf{Q}_i\} \leq E_{s,i}$.

Linear Precoding

- The received signal $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$ of user $q \in \mathbf{K}_i$

\# col = \# streams

$$\mathbf{y}_q = \underbrace{\Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{P}_{q,i}\mathbf{c}_{q,i}}_{\text{by cell } i \text{ to user } q} + \underbrace{\sum_{p\in\mathbf{K}_i,\, p\neq q} \Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{P}_{p,i}\mathbf{c}_{p,i}}_{\text{intra-cell (multi-user) interference}}$$

by cell i to user $p \neq q$

$$+ \underbrace{\sum_{j\neq i}\sum_{l\in\mathbf{K}_j} \Lambda_{q,j}^{-1/2}\mathbf{H}_{q,j}\mathbf{P}_{l,j}\mathbf{c}_{l,j}}_{\text{inter-cell interference}} + \mathbf{n}_q.$$

by other cells $j \neq i$ to user l

$$C_{q,i} = \begin{bmatrix} c_{q,i,1} \\ \vdots \\ c_{q,i,l} \end{bmatrix} \leftarrow \text{stream } l$$

- Apply a receive combiner to stream l of user q in cell i

cov. matrix of interf + noise
MMSE

$$z_{q,l} = \mathbf{g}_{q,l}\mathbf{y}_q = \underbrace{\underbrace{\Lambda_{q,i}^{-1/2}}\underbrace{\mathbf{g}_{q,l}}_{g}\underbrace{\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}}}_{h}\underbrace{c_{q,i,l}}_{c} + \left[\underbrace{\sum_{m\neq l} \Lambda_{q,i}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}c_{q,i,m}}_{\text{inter-stream interference}}\right.$$

user q.
stream l.

$$\overbrace{\left. + \underbrace{\sum_{p\in\mathbf{K}_i,\, p\neq q} \Lambda_{q,i}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{P}_{p,i}\mathbf{c}_{p,i}}_{\text{intra-cell (multi-user) interference}} + \underbrace{\sum_{j\neq i}\sum_{l\in\mathbf{K}_j} \Lambda_{q,j}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,j}\mathbf{P}_{l,j}\mathbf{c}_{l,j}}_{\text{inter-cell interference}}\right]}^{I} + \overbrace{\mathbf{g}_{q,l}\mathbf{n}_q}^{n}.$$

$z = g\,h\,c + I + n$

Achievable Rate

- By treating all interference as noise, the maximum rate achievable by user q in cell i with linear precoding is

scheduler $q^* = \arg\max \gamma_q \frac{R_q(k)}{R_q(k+1)}$

$$R_{q,i} = \sum_{l=1}^{n_{u,q}} \log_2 (1 + \rho_{q,l}) .$$

- The quantity $\rho_{q,l}$ denotes the SINR experienced by stream l of user-q and writes as

$$\rho_{q,l} = \frac{S}{I_l + I_c + I_o + \underbrace{\|\mathbf{g}_{q,l}\|^2 \sigma^2_{n,q}}_{\text{noise after combining}}} .$$

where S refers to the received signal power of the intended stream, I_l the inter-stream interference, I_c the intra-cell interference (i.e. interference from co-scheduled users) and I_o the inter-cell interference and they write as

$$S = \Lambda^{-1}_{q,i} \left| \mathbf{g}_{q,l} \mathbf{H}_{q,i} \mathbf{p}_{q,i,l} \right|^2 ,$$

precoder from codebook

inter-stream
$$I_l = \sum_{m \neq l} \Lambda^{-1}_{q,i} \left| \mathbf{g}_{q,l} \mathbf{H}_{q,i} \mathbf{p}_{q,i,m} \right|^2 ,$$

intra-cell
$$I_c = \sum_{p \in \mathbf{K}_i, p \neq q} \sum_{m=1}^{n_{u,p}} \Lambda^{-1}_{q,i} \left| \mathbf{g}_{q,l} \mathbf{H}_{q,i} \mathbf{p}_{p,i,m} \right|^2 ,$$

inter-cell
$$I_o = \sum_{j \neq i} \Lambda^{-1}_{q,j} \left\| \mathbf{g}_{q,l} \mathbf{H}_{q,j} \mathbf{P}_j \right\|^2 .$$

Achievable Rate

Example

Given the precoders in all cells, what is the SINR of stream l of user-q in cell i?

- Noise plus interference: $I_l + I_c + I_o + \|\mathbf{g}_{q,l}\|^2 \sigma_{n,q}^2 = \mathbf{g}_{q,l}\mathbf{R}_{\mathbf{n_i}}\mathbf{g}_{q,l}^H$ where

$$\mathbf{R}_{\mathbf{n_i}} = \sum_{m\neq l} \underbrace{\Lambda_{q,i}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}\right)^H}_{\text{inter stream}}$$

$$+ \sum_{p\in\mathbf{K}_i, p\neq q} \sum_{m=1}^{n_{u,p}} \underbrace{\Lambda_{q,i}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{p,i,m}\left(\mathbf{H}_{q,i}\mathbf{p}_{p,i,m}\right)^H}_{\text{inter user}}$$

$$+ \sum_{j\neq i} \underbrace{\Lambda_{q,j}^{-1}\mathbf{H}_{q,j}\mathbf{P}_j\left(\mathbf{H}_{q,j}\mathbf{P}_j\right)^H}_{\text{inter cell}} + \sigma_{n,q}^2\mathbf{I}_{n_{r,q}}$$

  is the covariance matrix of the noise plus interference.

- MMSE combiner for stream l: $\mathbf{g}_{q,l} = \Lambda_{q,i}^{-1/2}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right)^H \mathbf{R}_{\mathbf{n_i}}{}^{-1}$
- SINR $\rho_{q,l}$ experienced by stream l of user-q

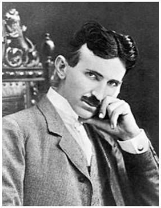
$$\rho_{q,l} = \frac{\Lambda_{q,i}^{-1}\left|\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right|^2}{\mathbf{g}_{q,l}\mathbf{R}_{\mathbf{n_i}}\mathbf{g}_{q,l}^H} = \Lambda_{q,i}^{-1}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right)^H \mathbf{R}_{\mathbf{n_i}}{}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}.$$

Future of Wireless

Wireless is More than just Communications

Radio waves carry both energy and information

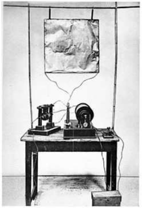
Wireless Power Transmission (WPT)



Tesla 1901

0G

Wireless Information Transmission (WIT)

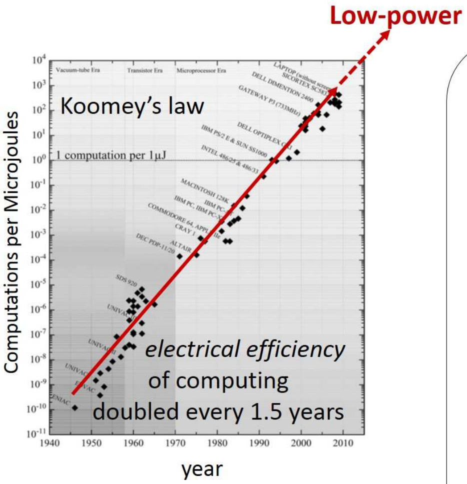


Marconi 1896

5G

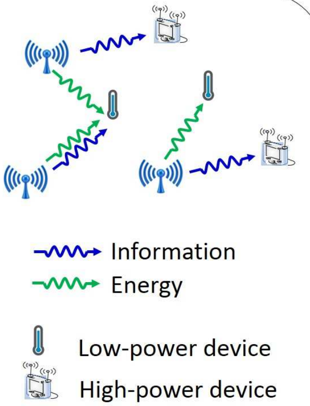
Unified Wireless Information and Power Transmission (WIPT)

In 20 Years from Now ... **Trillions** of **Low-Power** Devices



2017: 5 Billions phones

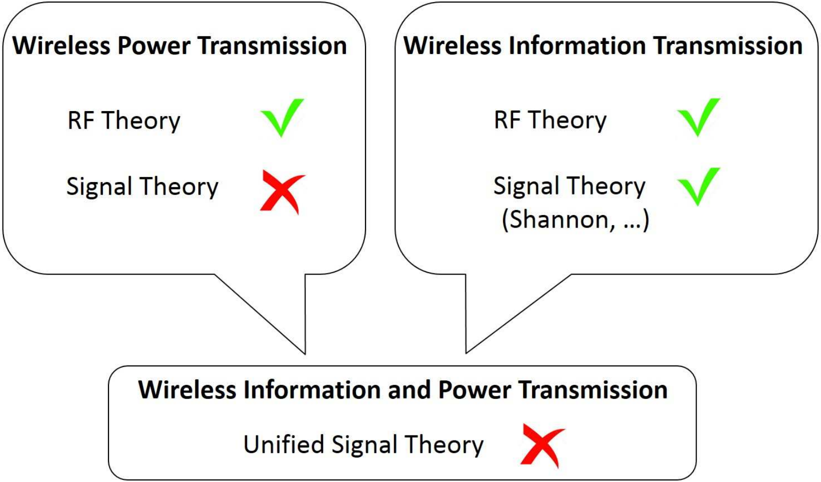
2035: **Trillions** IoT devices



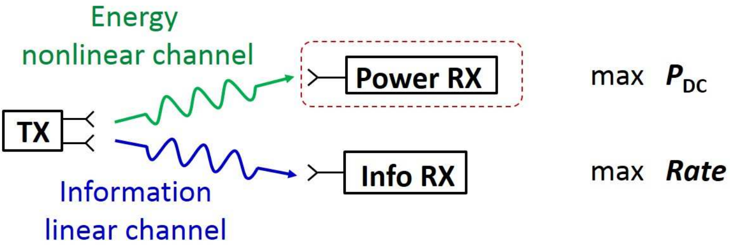
Future of Wireless:

Communicate and Energize

A **Missing** Signal Theory of Wireless Transmission



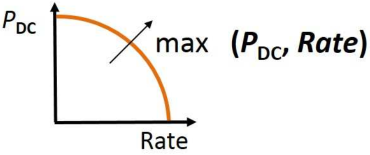
Signal Theory and Design



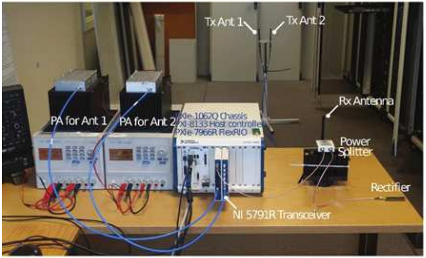
→ **New communications and signal design for WPT**

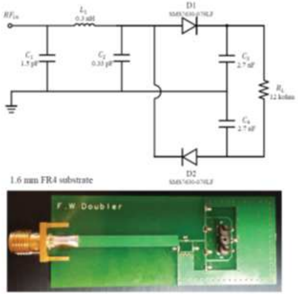
→ **Novel and unified signal theory and design for WIPT**

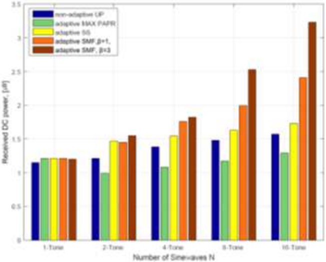
→ **Fundamental tradeoff**



System Design, Prototyping and Experimentation







Machine learning for Wireless Communications and Power Transfer

16-symbols modulation for different values of energy