EE4-65/EE9-SO27 Wireless Communications

Bruno Clerckx

Department of Electrical and Electronic Engineeing, Imperial College London

January 2016

Course Ojectives

- Advanced course on wireless communication and communication theory
  - Provides the 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 MIMO (Multiple Input Multiple Output) and multi-user/multi-cell communications
  - Includes as special cases SISO (Single Input Single Output), MISO (Multiple Input Single Output), SIMO (Single Input Multiple Output)
  - Applications: everywhere in wireless communication networks: 3G, 4G(LTE,LTE-A), (5G?), WiMAX(IEEE 802.16e, IEEE 802.16m), WiFi(IEEE 802.11n), satellite,...+ in other fields, e.g. radar, medical devices, speech and sound processing, ...
- 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
  - Mathematical modelling and analysis of (MIMO-based) wireless communication systems
  - Design (transmitters and receivers) of multi-cell multi-user MIMO wireless communication systems
  - Practical understanding of MIMO applications and performance evaluations

Content

Central question: How to deal with fading and interference in wireless networks?

- Some fundamentals/revision (matrix analysis, probability, information theory)
- **Single link: point to point communications**
  - Fading and Diversity
  - MIMO Channels - Modelling and Propagation
  - Capacity of point-to-point MIMO Channels
  - Space-Time Coding/Decoding over I.I.D. Rayleigh Flat Fading Channels
  - Partial Channel State Information at the Transmitter (CSIT)
- **Multiple links: multiuser communications**
  - Multi-User MIMO - Capacity of Multiple Access Channels (Uplink)
  - Multi-User MIMO - Capacity of Broadcast Channels (Downlink)
  - Multi-User MIMO - Scheduling, Linear Precoding (Downlink)
- **Multiple cells: multiuser multicell communications**
  - Introduction to Multi-Cell MIMO
  - Capacity of Interference Channel
- **Real-World MIMO Wireless Networks**
  - MIMO and Interference Management in 4G and beyond (LTE, LTE-Advanced, WiMAX)

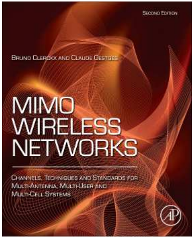
Important Information

- Course webpage: http://www.ee.ic.ac.uk/bruno.clerckx/Teaching.html
- Prerequisite: EE9SC2 Advanced Communication Theory
- Lectures on Tuesday from 14.00 till 16.00
- Exam (written, 3 hours, closed book): 70%; Project (using Matlab): 30%.
- Do the problems in problem sheets (2 types: 1. paper/pencil, 2. matlab)
- Project
  - to be distributed around mid February (details to come later)
  - report to be submitted by end of March (details to come later).

Important Information

- Reference book

  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.

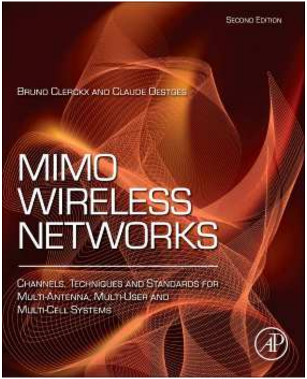


- Another interesting reference on wireless communications (more introductory) "Fundamentals of Wireless Communication," by D. Tse and P. Viswanath, Cambridge University Press, May 2005

Some fundamentals/revisions (matrix analysis, probability)

Reference Book

- 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.



  - Appendix A, B

Matrix properties

- *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 $\mathbf{H}$ $[n_r \times n_t]$: $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$
  - $\mathbf{U}$ $[n_r \times r(\mathbf{H})]$: unitary matrix of left singular vectors
  - $\boldsymbol{\Sigma} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}$: 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 $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H$ with $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^2$.
- $\mathbf{A} = \mathbf{H}^H\mathbf{H}$ 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 $\mathbf{A}$: $\|\mathbf{A}\|_F^2 = \sum_i \sum_j |A(i,j)|^2$
- $\|\mathbf{A}\|_F^2 = \text{Tr}\left\{\mathbf{A}\mathbf{A}^H\right\} = \text{Tr}\left\{\mathbf{A}^H\mathbf{A}\right\}$
- $\text{Tr}\{\mathbf{A}\mathbf{B}\} = \text{Tr}\{\mathbf{B}\mathbf{A}\}$
- $\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$
- *Hadamard's inequality*: $\det(\mathbf{A}) \leq \prod_{k=1}^n \mathbf{A}(k,k)$ if $\mathbf{A} > 0$ (positive definite matrix, all eigenvalues are positive) of size $n \times n$

Matrix properties

- *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.
- $\det(\mathbf{A}_{m\times m} \otimes \mathbf{B}_{n\times n}) = \det(\mathbf{A})^n \det(\mathbf{B})^m$
- $\mathrm{Tr}\{\mathbf{A} \otimes \mathbf{B}\} = \mathrm{Tr}\{\mathbf{A}\}\,\mathrm{Tr}\{\mathbf{B}\}$
- $\mathrm{Tr}\{\mathbf{AB}\} \geq \mathrm{Tr}\{\mathbf{A}\}\,\sigma^2_{min}(\mathbf{B})$ with $\sigma_{min}(\mathbf{B})$ the smallest singular value of $\mathbf{B}$
- $\mathrm{vec}(\mathbf{A})$ converts $[m \times n]$ matrix into $mn \times 1$ vector by stacking the columns of $\mathbf{A}$ on top of one another.
  - $\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\mathrm{vec}(\mathbf{B})$
- $\mathrm{Tr}\left\{\mathbf{ABB}^H\mathbf{A}^H\right\} = \mathrm{vec}\left(\mathbf{A}^H\right)^H \left(\mathbf{I} \otimes \mathbf{BB}^H\right)\mathrm{vec}\left(\mathbf{A}^H\right)$
- $\det(\mathbf{I} + \epsilon\mathbf{A}) = 1 + \epsilon\mathrm{Tr}\{\mathbf{A}\}$ if $\epsilon << 1$

Gaussian random variable

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

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

  Standard Gaussian random variable: $\mu = 0$ and $\sigma^2 = 1$

- *Real Gaussian random vector* $\mathbf{x}$ of dimension n with mean vector $\boldsymbol{\mu} = \mathcal{E}\left\{\mathbf{x}\right\}$ and covariance matrix $\mathbf{R} = \mathcal{E}\left\{\left(\mathbf{x}-\boldsymbol{\mu}\right)\left(\mathbf{x}-\boldsymbol{\mu}\right)^T\right\}$:

$$p\left(\mathbf{x}\right) = \frac{1}{\left(\sqrt{2\pi}\right)^n \sqrt{\det\left(\mathbf{R}\right)}} \exp\left(-\frac{\left(\mathbf{x}-\boldsymbol{\mu}\right)^T \mathbf{R}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)}{2}\right).$$

  Standard Gaussian random vector $\mathbf{x}$ of dimension n: entries are independent and identically distributed (i.i.d.) standard Gaussian random variables $x_1, \ldots, x_n$

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

Gaussian random variable

- *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

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

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

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

  Hence, $\mu = \mathcal{E}\left\{y\right\} = 2\sigma^2$.
- More generally, 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

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

Appendix: Basics of Information Theory

T. Cover and J. Thomas, "Elements of Information Theory," Second Edition, Wiley, 2006.

Discrete Memoryless Channel

Definition

A *discrete* channel is defined as a system consisting of an input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and a probability transition matrix $p(y|x)$ that expresses the probability of observing the output symbols y given that the symbold x is sent.

Definition

The channel is *memoryless* if the probability distribution of the output depends only on the input at that time and is conditionally independent of previous channel inputs or outputs, i.e. if $x_1,...,x_n$ are inputs, and $y_1,...,y_n$ denote the corresponding outputs, for n channel uses, then

$$p(y_1, ..., y_n|x_1, ...x_n) = p(y_1|x_1)...p(y_n|x_n)$$

Example

Binary Symmetric Channel (BSC): x and y take values in 0,1 such that

$$p(y|x) = \begin{cases} 1-p, & y = x, \\ p, & y = 1-x. \end{cases}$$

Entropy

- Entropy is a measure of the average uncertainty of a random variable

Definition

For a discrete random variable X, the entropy $H(X)$ is defined as

$$H(X) = \mathcal{E}\left\{\log_2 \frac{1}{p(X)}\right\} = -\mathcal{E}\left\{\log_2 p(X)\right\} = -\sum_x p(x)\log_2 p(x),$$

where $p(x)$ is the probability mass function of X.

- It is the number of bits on average required to describe the random variable.

Example

Let X be a Bernoulli random variable

$$X = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1-p. \end{cases}$$

Then $H(X) = -p\log_2 p - (1-p)\log_2(1-p)$. For $p = 0, 1$, there is no uncertainty on the value of the RV, so no information gained. For $p = 1/2$, $H(X)$ (uncertainty/information) is maximized.

Entropy

Lemma

$H(X) \geq 0$

Proof: $0 \leq p(x) \leq 1$ such that $\log_2 \frac{1}{p(x)} \geq 0$ $\square$

Definition

The joint entropy $H(X, Y)$ of a pair of discrete random variables X and Y with a joint pmf $p(x, y)$ is defined as

$$H(X, Y) = -\mathcal{E}\left\{\log_2 p(X, Y)\right\} = -\sum_x \sum_y p(x, y) \log_2 p(x, y)$$

Conditional Entropy

- The conditional entropy of a random variable given another is the expected value of the entropies of the conditional distributions, averaged over the conditioning random variable

Definition

The conditional entropy $H(Y|X)$ is defined as

$$\begin{aligned} H(Y|X) &= \sum_x p(x) H(Y|X=x) \\ &= -\sum_x p(x) \sum_y p(y|x) \log_2 p(y|x) \\ &= -\sum_x \sum_y p(x,y) \log_2 p(y|x) \\ &= -\mathcal{E}\left\{\log_2 p(Y|X)\right\} \end{aligned}$$

Joint Entropy

Theorem

Chain rule

$$H(X,Y) = H(X) + H(Y|X)$$

Proof:

$$\begin{aligned}
H(X,Y) &= -\sum_x \sum_y p(x,y)\log_2 p(x,y) = -\sum_x \sum_y p(x,y)\log_2 p(x)p(y|x) \\
&= -\sum_x \sum_y p(x,y)\log_2 p(x) - \sum_x \sum_y p(x,y)\log_2 p(y|x) \\
&= -\sum_x p(x)\log_2 p(x) - \sum_x \sum_y p(x,y)\log_2 p(y|x) \\
&= H(X) + H(Y|X)
\end{aligned}$$

Alternatively,

$$\log_2 p(X,Y) = \log_2 p(X) + \log_2 p(Y|X)$$
$$\mathcal{E}\{\log_2 p(X,Y)\} = \mathcal{E}\{\log_2 p(X)\} + \mathcal{E}\{\log_2 p(Y|X)\}$$

Relative Entropy

- The relative entropy is a measure of the distance between two distributions.

Definition

The relative entropy between two pmf $p(x)$ and $q(x)$ is defined as

$$D(p||q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)} = \mathcal{E}_p \left\{ \log_2 \frac{p(X)}{q(X)} \right\}$$

Theorem

The relative entropy is always nonnegative $D(p||q) \geq 0$ and is zero if and only if $p = q$.

Mutual Information

- The mutual information is a measure of the amount of information that one RV contains about another RV. It is a measure of the dependence between the two RVs.

Definition

For a pair of discrete random variables X and Y with a joint pmf $p(x, y)$ and marginal pmf $p(x)$ and $p(y)$, the mutual information $I(X;Y)$ is the relative entropy between $p(x, y)$ and $p(x)p(y)$

$$I(X;Y) = D(p(x,y)||p(x)p(y)) = \mathcal{E}_{p(x,y)}\left\{\log_2 \frac{p(X,Y)}{p(X)p(Y)}\right\}$$
$$= \sum_x \sum_y p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$

Mutual Information

- The mutual information $I(X;Y)$ is the reduction in the uncertainty of one random variable due to the knowledge of the other

$$
\begin{aligned}
I(X;Y) &= \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \\
&= \sum_{x,y} p(x,y) \log_2 \frac{p(x|y)}{p(x)} \\
&= -\sum_{x,y} p(x,y) \log_2 p(x) + \sum_{x,y} p(x,y) \log_2 p(x|y) \\
&= -\sum_{x} p(x) \log_2 p(x) - \left(-\sum_{x,y} p(x,y) \log_2 p(x|y) \right) \\
&= H(X) - H(X|Y) \\
&= H(Y) - H(Y|X) = I(Y;X)
\end{aligned}
$$

- $I(X;Y) = H(X) + H(Y) - H(X,Y)$.
- $I(X;X) = H(X) - H(X|X) = H(X)$

Mutual Information

Theorem

Nonnegativity of mutual information: For any two random variables X,Y

$$I(X;Y) \geq 0$$

with equality if and only if X *and* Y *are independent*

Theorem

Conditioning reduces entropy: For any two random variables X,Y

$$H(X|Y) \leq H(X)$$

with equality if and only if X *and* Y *are independent*

Proof: $0 \leq I(X;Y) = H(X) - H(X|Y)$ □

Knowing another RV Y can only reduce on the average the uncertainty in X.

Channel Coding Theorem

Theorem

(a) For a DMC with channel transition pmf $p(y|x)$*, we can use i.i.d. inputs with pmf* $p(x)$ *to communicate reliably, as long as the code rate satisfies*

$$R < I(X;Y).$$

(b) The achievable rate can be maximized over the input density $p(x)$ *to obtain the channel capacity*

$$C = \max_{p(x)} I(X;Y).$$

Differential Entropy

Definition

For a continuous random variable X, the differential entropy $h(X)$ is defined as

$$h(X) = \mathcal{E}\left\{\log_2 \frac{1}{p(x)}\right\} = -\mathcal{E}\left\{\log_2 p(x)\right\} = -\int p(x) \log_2 p(x) dx,$$

where $p(x)$ is the probability density function of X.

Caution: $h(X)$ can be negative.

Example

For $X \sim N(\mu, \sigma^2)$, $-\log_2 p(x) = \frac{(x-\mu)^2}{2\sigma^2} \log_2(e) + \frac{1}{2}\log_2(2\pi\sigma^2)$. Thus, $h(X) = -\mathcal{E}\left\{\log_2 p(x)\right\} = \frac{1}{2}\log_2(e) + \frac{1}{2}\log_2(2\pi\sigma^2) = \frac{1}{2}\log_2(2\pi e\sigma^2)$. The mean does not affect the differential entropy.

Theorem

Consider a RV with zero mean and variance σ^2. Then $h(X) \leq \frac{1}{2}\log_2(2\pi e\sigma^2)$, with equality iff $X \sim N(0, \sigma^2)$.

AWGN Channel

- Real discrete-time AWGN channel

$$Y = X + N, \qquad N \sim N(0, \sigma^2)$$

where X is power-constrained input $\mathcal{E}\left\{X^2\right\} \leq E_s$

- The channel transition density is given by

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

AWGN Channel Capacity

Theorem

The capacity of the real AWGN channel is

$$C = \max_{p(x):\mathcal{E}\left\{X^2\right\}\leq E_s} I(X;Y) = \frac{1}{2}\log_2(1+\frac{E_s}{\sigma^2}).$$

Proof: Consider $Y = X + N$, with $N \sim N(0,\sigma^2)$ and $\mathcal{E}\left\{X^2\right\} \leq E_s$. Given $X = x$, $h(Y|X=x) = h(N)$, so that $h(Y|X) = h(N)$ and

$$I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(N).$$

Maximizing $I(X;Y)$ comes to maximize $h(Y)$. Since X and N are independent, $\mathcal{E}\left\{Y^2\right\} = \mathcal{E}\left\{X^2\right\} + \mathcal{E}\left\{N^2\right\} \leq E_s + \sigma^2$. We now know that

$$h(Y) \leq \frac{1}{2}\log_2(2\pi e(E_s+\sigma^2))$$

and equality is achieved iff $Y \sim N(0, E_s+\sigma^2)$. $Y \sim N(0, E_s+\sigma^2)$ is achieved if the input distribution is $X \sim N(0, E_s)$, independent of the noise. We then get

$$I(X;Y) = h(Y) - h(N) = \frac{1}{2}\log_2(2\pi e(E_s+\sigma^2)) - \frac{1}{2}\log_2(2\pi e\sigma^2) = \frac{1}{2}\log_2(1+\frac{E_s}{\sigma^2}).$$

Jensen's inequality

Theorem

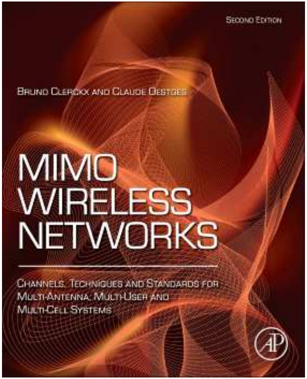
If f is a convex function and X is a random variable,

$$\mathcal{E}\left\{f(X)\right\} \geq f(\mathcal{E}\left\{X\right\}).$$

Fading and Diversity

Reference Book

- 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.



  - Chapter 1
    - Section: 1.2, 1.3, 1.4, 1.5
    - Appendix A, B

Space-Time Wireless Channels: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.
- The transmitted signal is then represented by

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

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

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

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

- 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

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

Example

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

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

- If $T_s >> \tau_{max}$,
  - $\mathsf{h}_B(t,\tau)$ 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

$$\mathsf{y}_k = \sqrt{E_s}\mathsf{h}_k c_k + \mathsf{n}_k$$

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

Path-Loss and Shadowing

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

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

where

  - *path-loss* Λ_0: a real-valued deterministic attenuation term modeled as $\Lambda_0 \propto R^\eta$ where η designates the path-loss exponent and R the distance between Tx and Rx.
  - *shadowing* S: a real-valued random additional attenuation term, which, for a given range, depends on the specific location of the transmitter and the receiver and modeled as a lognormal variable, i.e., $10 \log_{10}(S)$ is a zero-mean normal variable of given standard deviation σ_S.
  - *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

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

where $|_{\mathsf{dB}}$ indicates the conversion to dB, and L_0 is the deterministic path-loss at a reference distance R_0, and Λ is generally known as the path-loss.

Path-Loss and Shadowing

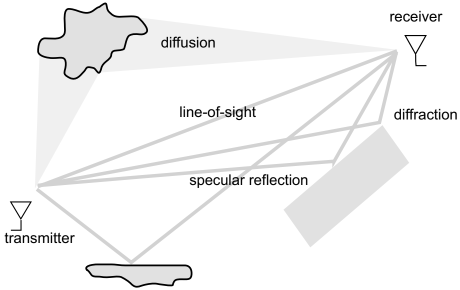
- Path loss models are identical for both single- and multi-antenna systems.

- 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
$$y = \sqrt{E_s} h c + n,$$
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 then defined as $\rho \triangleq E_s/\sigma_n^2$.

Fading

- Multipaths



- Assuming that the signal reaches the receiver via a large number of paths of similar energy,
  - 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

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

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

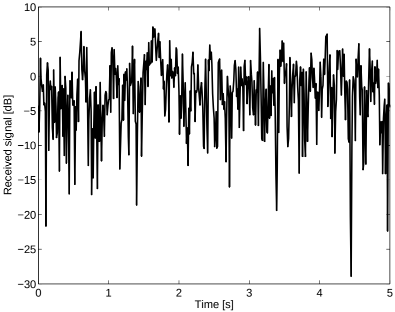
  whose first two moments are

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

- The *phase* of h is uniformly distributed over $[0, 2\pi)$

Fading

- Illustration of the typical received signal strength of a Rayleigh fading channel over a certain time interval



  - The signal level randomly fluctuates, with some sharp declines of power and instantaneous received SNR known as *fades*.
  - 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 $n \sim N(0, \sigma_n^2)$,

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

  and

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

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

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

Diversity in Multiple Antennas Wireless Systems

- 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

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

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

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

  - 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

$$\begin{aligned}\bar{P} &= \int_0^\infty \mathcal{Q}\big(\sqrt{2\rho}s\big) p_s(s)\, ds \\ &= \frac{1}{2}\left(1 - \sqrt{\frac{\rho}{1+\rho}}\right) \\ &\overset{(\rho\nearrow)}{\cong} \frac{1}{4\rho}\end{aligned}$$

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

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

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

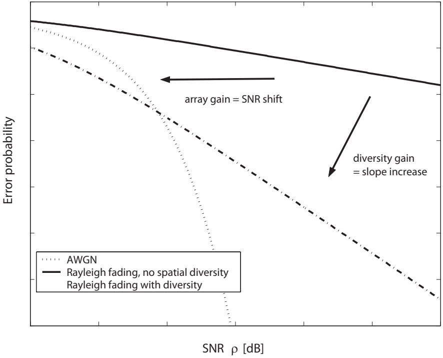
- *Diversity Gain*: increase in the error rate slope as a function of the SNR. Defined as the negative slope of the log-log plot of the average error probability $\bar{P}$ versus SNR

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

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

Diversity in Multiple Antennas Wireless Systems

- 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

- 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

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

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

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

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

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

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

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

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

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

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

  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,

$$z = \mathbf{g}\mathbf{y} = \sum_{n=1}^{n_r} g_n y_n = \sqrt{E_s}\mathbf{g}\mathbf{h}c + \mathbf{g}\mathbf{n}$$

  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
  - $\mathbf{h} \triangleq [h_1, \ldots, h_{n_r}]^T$
- 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 $g_n = e^{-j\phi_n}$.
  - Mean value of the output SNR $\bar{\rho}_{out}$ (averaged over the Rayleigh fading):

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

    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.

Receive Diversity via Gain Combining

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

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

  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

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

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

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

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

$$\bar{P} = (4\rho)^{-n_r} \binom{2n_r - 1}{n_r}.$$

  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

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

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

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

- *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 $\mathbf{y} = \sqrt{E_s}\mathbf{h}c + \mathbf{n}_i$.
  - 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.,

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

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

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

  where $\mathbf{R}_{\mathbf{n}_i} = \mathcal{E}\left\{\mathbf{n}_i \mathbf{n}_i^H\right\}$ is the correlation matrix of the combined noise plus interference signal $\mathbf{n}_i$.
  - 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 $\mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-H/2}$.
  - The Signal to Interference plus Noise Ratio (SINR) at the output of the MMSE combiner simply writes as

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

  - 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

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 $\mathbf{h}_i$.
The received signal model writes as

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

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.

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

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

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 systems exploit diversity at the transmitter through the use of n_t transmit antennas in combination with pre-processing or precoding.

- A significant difference with receive diversity is that the transmitter might not have the knowledge of the MISO channel.
  - 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,
  - 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).

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

Transmit Diversity via Matched Beamforming

- The actual transmitted signal is a vector $\mathbf{c}'$ that results from the multiplication of the signal c by a weight vector $\mathbf{w}$.
- At the receiver, the signal reads as

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

  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 it is also known as *matched beamforming* or *transmit MRC*.
- 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

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

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

Transmit Diversity via Space-Time Coding

- *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 $\mathbf{h} = [h_1 \ h_2]$.
  - 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:

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

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

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

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

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

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

Transmit Diversity via Space-Time Coding

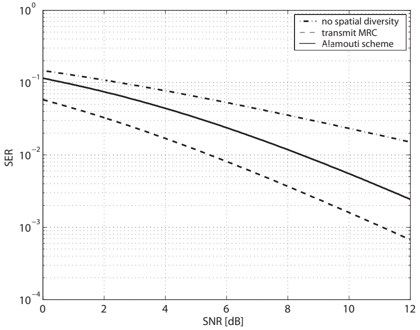
- The mean output SNR (averaged over the channel statistics) is thus equal to

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

  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

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

  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

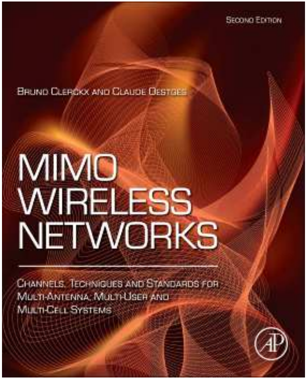
Indirect Transmit Diversity

- 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.

MIMO Systems - Transmission

Reference Book

- 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.



  - Chapter 1
    - Section: 1.2.4, 1.3.2, 1.6

Introduction - 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.
- 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

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

  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, $\mathcal{E}\{\|\mathbf{H}\|_F^2\} = n_t n_r$.
- when Tx has a *perfect channel knowledge*: (dominant and multiple) eigenmode transmission
- when Tx has *no knowledge of the channel*: space-time coding (with $\mathbf{c}'_k = \mathbf{c}_k$)

Space-Time Coding

- 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

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

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

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

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

- Equivalently

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

Space-Time Coding

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

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

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

- Average output SNR

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

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

- Average symbol error rate

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

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

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

- Decompose

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

- Received SNR is maximized by matched filtering, leading to

$$\mathbf{w} = \mathbf{v}_{max}$$
$$\mathbf{g} = \mathbf{u}_{max}^H$$

  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

$$\mathbf{H} = \mathbf{U}_{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_{\mathbf{H}}^H, \qquad \boldsymbol{\Sigma}_{\mathbf{H}} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{r(\mathbf{H})}\}.$$

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

$$\mathbf{g}\mathbf{H}\mathbf{w} = \mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w} = \mathbf{w}^H\mathbf{V}_{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{H}}^2\mathbf{V}_{\mathbf{H}}^H\mathbf{w} = \sum_{i=1}^{r(\mathbf{H})} \sigma_i^2 \left|\mathbf{v}_i^H\mathbf{w}\right|^2 \leq \sigma_{max}^2$$

where $\mathbf{v}_i$ is the i column of $\mathbf{V}_{\mathbf{H}}$ 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 $\mathbf{w} = \mathbf{v}_{max}$,

$$\begin{aligned} \mathbf{g} = \mathbf{w}^H \mathbf{H}^H &= \mathbf{v}_{max}^H \mathbf{V_H} \mathbf{\Sigma_H} \mathbf{U_H}^H \\ &= \sigma_{max} \mathbf{u}_{max}^H \end{aligned}$$

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

Multiple Eigenmode Transmission

- 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. $\mathbf{c}' = \mathbf{V_H c}$.
  - Rx side: multiply the received vector $\mathbf{y}$ by $\mathbf{G} = \mathbf{U_H^H}$.
  - Overall,

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

    The channel has been decomposed into n_t parallel SISO channels given by $\{\sigma_1, \ldots, \sigma_{n_t}\}$.
- The rate achievable in the MIMO channel is the sum of the SISO channel capacities

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

  where $\{s_1, \ldots, s_{n_t}\}$ is the power allocation on each of the channel eigenmodes.
- The capacity scales linearly in n_t. By contrast, this transmission does not necessarily achieve the full diversity gain of $n_t n_r$ but does at least provide n_r-fold array and diversity gains (still assuming $n_t \leq n_r$).
- In general, the rate scales linearly with the rank of $\mathbf{H}$.

Multiple Eigenmode Transmission

Example

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

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

where $\{s_1, \ldots, s_{r(\mathbf{H})}\}$ is the power allocation on each of the channel eigenmodes.
Two strategies (for a total power constraint $\sum_{k=1}^{r(\mathbf{H})} s_k = 1$):

- Uniform power allocation: $R_u = \sum_{k=1}^{r(\mathbf{H})} \log_2(1 + \rho 1/r(\mathbf{H})\sigma_k^2)$
- 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 $k > 1$, $R_u \approx \log_2(1 + \rho\sigma_1^2/r(\mathbf{H})) \leq R_d$ 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.
- MIMO channels offer *multiplexing gain*: measure of the number of independent streams that can be transmitted in parallel in the MIMO channel. Defined as

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

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

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

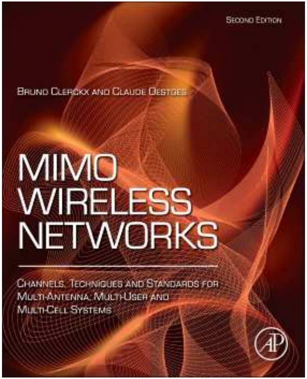
$$R \approx g_s \log_2(\rho)$$

- 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!

Channel Modelling

Reference Book

- 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.



- Chapter 2
  - Section: 2.1.1, 2.1.2, 2.1.3, 2.1.5, 2.2, 2.3.1
- Chapter 3
  - Section: 3.2.1, 3.2.2, 3.4.1

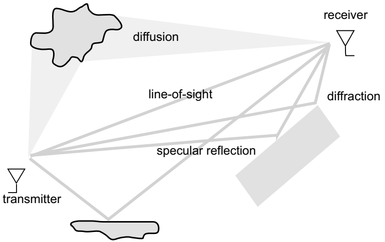
Double-Directional Channel Modeling

- 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
- Neglecting path-loss and shadowing, the time-variant double-directional channel

$$h\left(t,\mathbf{p}_t,\mathbf{p}_r,\tau,\boldsymbol{\Omega}_t,\boldsymbol{\Omega}_r\right)=\sum_{k=0}^{n_s-1}h_k\left(t,\mathbf{p}_t,\mathbf{p}_r,\tau,\boldsymbol{\Omega}_t,\boldsymbol{\Omega}_r\right),$$

  - $\mathbf{p}_t$, $\mathbf{p}_r$: location of Tx and Rx, respectively
  - n_s contributions
  - time t: variation with time (with the motion of the receiver)
  - delay τ: each contribution arrives with a delay proportional to its path length
  - $\boldsymbol{\Omega}_t$, $\boldsymbol{\Omega}_r$: direction of departure (DoD), directions of arrival (DoA). In spherical coordinates (i.e., the azimuth Θ_t and elevation ψ_t) on a sphere of unit radius

$$\boldsymbol{\Omega}_t=\left[\cos\Theta_t\sin\psi_t,\sin\Theta_t\sin\psi_t,\cos\psi_t\right]^T$$



Double-Directional Channel Modeling

- In the case of a plane wave, and considering a fixed transmitter and a mobile receiver,

$$h_k\left(t, \mathbf{p}_t, \mathbf{p}_r, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right) \triangleq \alpha_k\, e^{j\phi_k}\, e^{-j\Delta\omega_k t}\, \delta(\tau - \tau_k)\, \delta(\mathbf{\Omega}_t - \mathbf{\Omega}_{t,k})\, \delta(\mathbf{\Omega}_r - \mathbf{\Omega}_{r,k}),$$

  where
  - α_k is the amplitude of the k^{th} contribution,
  - ϕ_k is the phase of the k^{th} contribution,
  - $\Delta\omega_k$ is the Doppler shift of the k^{th} contribution,
  - τ_k is the time delay of the k^{th} contribution,
  - $\Omega_{t,k}$ is the DoD of the k^{th} contribution,
  - $\Omega_{r,k}$ is the DoA of the k^{th} contribution.
- A more compact notation (all temporal variations are grouped into t)

$$h\left(t, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right) = \sum_{k=0}^{n_s-1} h_k\left(t, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right)$$

- Impulse response of the channel (as in Lecture 1, without path loss/shadowing)

$$h(t,\tau) = \iint h\big(t, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\big)\, d\mathbf{\Omega}_t\, d\mathbf{\Omega}_r$$

- Narrowband transmission (the channel is not frequency selective)

$$h(t) = \iiint h\big(t, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\big)\, d\tau\, d\mathbf{\Omega}_t\, d\mathbf{\Omega}_r$$

Wide-Sense Stationary Uncorrelated Scattering Homogeneous

- Assumption: Wide-Sense Stationary Uncorrelated Scattering Homogeneous (WSSUSH) channels
- Wide-Sense Stationary:
  - Time correlations only depend on the time difference
  - Signals arriving with different Doppler frequencies are uncorrelated
- Uncorrelated Scattering:
  - Frequency correlations only depend on the frequency difference
  - Signals arriving with different delays are uncorrelated
- Homogeneous:
  - Spatial correlation only depends on the spatial difference at both transmit and receive sides
  - Signals departing/arriving with different directions are uncorrelated

Spectra

- Doppler spectrum and coherence time
- Power delay spectrum and delay spread
- Power direction spectrum and angle spread
  - the power-delay joint direction spectrum

$$\mathcal{P}_h(\tau, \boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r) = \mathcal{E}\{\, \left| h(t, \tau, \boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r) \right|^2 \,\},$$

  - the joint direction power spectrum

$$\mathcal{A}(\boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r) = \int \mathcal{P}_h(\tau, \boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r)\, d\tau,$$

  - the transmit direction power spectrum

$$\mathcal{A}_t(\boldsymbol{\Omega}_t) = \int\int \mathcal{P}_h(\tau, \boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r)\, d\tau\, d\boldsymbol{\Omega}_r,$$

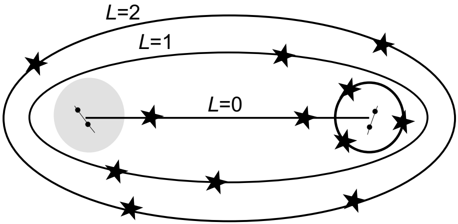
  - the receive direction power spectrum

$$\mathcal{A}_r(\boldsymbol{\Omega}_r) = \int\int \mathcal{P}_h(\tau, \boldsymbol{\Omega}_t, \boldsymbol{\Omega}_r)\, d\tau\, d\boldsymbol{\Omega}_t.$$

Angular Spread

- The channel angle-spreads are defined similarly to the delay-spread
  - delay-spread $\Longleftrightarrow$ channel frequency selectivity
  - angle-spread $\Longleftrightarrow$ channel spatial selectivity

$$
\begin{aligned}
\boldsymbol{\Omega}_{t,M} &= \frac{\int \boldsymbol{\Omega}_t \mathcal{A}_t(\boldsymbol{\Omega}_t)\, d\boldsymbol{\Omega}_t}{\int \mathcal{A}_t(\boldsymbol{\Omega}_t)\, d\boldsymbol{\Omega}_t} \\
\Omega_{t,RMS} &= \sqrt{\frac{\int \left\| \boldsymbol{\Omega}_t - \boldsymbol{\Omega}_{t,M} \right\|^2 \mathcal{A}_t(\boldsymbol{\Omega}_t)\, d\boldsymbol{\Omega}_t}{\int \mathcal{A}_t(\boldsymbol{\Omega}_t)\, d\boldsymbol{\Omega}_t}}
\end{aligned}
$$



The MIMO Channel Matrix

- Convert the double-directional channel to a $n_r \times n_t$ MIMO channel

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

where

$$h_{nm}(t,\tau) \quad \triangleq \quad \iint h_{nm}\left(t,\tau,\mathbf{\Omega}_t,\mathbf{\Omega}_r\right)\, d\mathbf{\Omega}_t \, d\mathbf{\Omega}_r$$

- For narrowband (i.e. same delay for all antennas) balanced (i.e. $|h_{nm}| = |h_{11}|$) arrays and plane wave incidence, $h_{nm}\left(t,\tau,\mathbf{\Omega}_t,\mathbf{\Omega}_r\right)$ is a phase shifted version of $h_{11}\left(t,\tau,\mathbf{\Omega}_t,\mathbf{\Omega}_r\right)$

$$h_{nm}(t,\tau) = \int\int h_{11}\left(t,\tau,\mathbf{\Omega}_t,\mathbf{\Omega}_r\right) e^{-j\mathbf{k}_r^T(\mathbf{\Omega}_r)\left[\mathbf{p}_r^{(n)}-\mathbf{p}_r^{(1)}\right]} e^{-j\mathbf{k}_t^T(\mathbf{\Omega}_t)\left[\mathbf{p}_t^{(m)}-\mathbf{p}_t^{(1)}\right]} d\mathbf{\Omega}_t d\mathbf{\Omega}_r$$

where $\mathbf{k}_t(\mathbf{\Omega}_t)$ and $\mathbf{k}_r(\mathbf{\Omega}_r)$ are the transmit and receive wave propagation 3×1 vectors.

Steering Vectors

- For a transmit ULA oriented broadside to the link axis,
$$e^{-j\mathbf{k}_t^T(\mathbf{\Omega}_t)\cdot\left[\mathbf{p}_t^{(m)}-\mathbf{p}_t^{(1)}\right]}=e^{-j(m-1)\varphi_t(\theta_t)},$$
where $\varphi_t\left(\theta_t\right)=2\pi(d_t/\lambda)\cos\theta_t$, and $d_t=\left\|\mathbf{p}_t^{(m)}-\mathbf{p}_t^{(m-1)}\right\|$ denotes the inter-element spacing of the transmit array.
  - θ_t is defined relatively to the array orientation (so $\theta_t=\pi/2$ corresponds to the link axis for a broadside array).

- Steering vector (expressed here for a ULA)
  - At the transmitter in the relative direction θ_t:
$$\mathbf{a}_t(\theta_t)=\begin{bmatrix}1 & e^{-j\varphi_t(\theta_t)} & \dots & e^{-j(n_t-1)\varphi_t(\theta_t)}\end{bmatrix}^T.$$
  - At the receiver in the relative direction θ_r:
$$\mathbf{a}_r\left(\theta_r\right)=\begin{bmatrix}1 & e^{-j\varphi_r(\theta_r)} & \dots & e^{-j(n_r-1)\varphi_r(\theta_r)}\end{bmatrix}^T.$$

- Under the plane wave and balanced narrowband array assumptions, the MIMO channel matrix can be rewritten as a function of steering vectors as
$$\mathbf{H}(t,\tau)=\int\int h\left(t,\mathbf{p}_t^{(1)},\mathbf{p}_r^{(1)},\tau,\mathbf{\Omega}_t,\mathbf{\Omega}_r\right)\mathbf{a}_r(\mathbf{\Omega}_r)\,\mathbf{a}_t^T(\mathbf{\Omega}_t)\,d\mathbf{\Omega}_t\,d\mathbf{\Omega}_r.$$

A Finite Scatterer MIMO Channel Representation

- The transmitter and receiver are coupled via a finite number of scattering paths with $n_{s,t}$ DoDs at the transmitter and $n_{s,r}$ DoAs at the receiver.
  $\longrightarrow$ Replace the integral by a summation (assume for simplicity 2-D azimuthal propagation)

$$\begin{aligned}\mathbf{H}(t,\tau) &= \sum_{l=1}^{n_{s,t}} \sum_{p=1}^{n_{s,r}} h_{11}^{(l,p)}(t,\tau)\mathbf{a}_r(\theta_r^{(p)})\mathbf{a}_t^T(\theta_t^{(l)}) \\ &= \mathbf{A}_r \mathbf{H}_s(t,\tau)\mathbf{A}_t^T\end{aligned}$$

  where
  - $\mathbf{A}_r$ and $\mathbf{A}_t$ represent the $n_r \times n_{s,r}$ and $n_t \times n_{s,t}$ matrices whose columns are the steering vectors related to the directions of each path observed at Rx and Tx
  - $\mathbf{H}_s(t,\tau)$ is a $n_{s,r} \times n_{s,t}$ matrix whose elements are the complex path gains between all DoDs and DoAs at time instant t and delay τ
- Assume the columns of $\mathbf{A}_t$ are written as $\mathbf{a}_t(\theta_t^{(l)})$, $l = 1, ..., n_{s,t}$. Let us write

$$\mathbf{H} = \underbrace{\mathbf{A}_r \mathbf{H}_s}_{\tilde{\mathbf{H}}_s} \mathbf{A}_t^T = \sum_{l=1}^{n_{s,t}} \tilde{\mathbf{H}}_s(:,l)\mathbf{a}_t^T(\theta_t^{(l)}) = \sum_{l=1}^{n_{s,t}} \mathbf{H}^{(l)},$$

  where $\mathbf{H}^{(l)}$ can be viewed as the channel matrix corresponding to the l^{th} scatterer located in the direction of departure $\theta_t^{(l)}$.

Statistical Properties of the MIMO Channel Matrix

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

$$\mathbf{R} = \mathcal{E}\{\text{vec}(\mathbf{H}^H)\text{vec}(\mathbf{H}^H)^H\}$$

  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).

Example

2x2 MIMO

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

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

Spatial Correlation

- How are these correlations related to the propagation channel?
- Let us consider the case of ULAs and 2-D azimuthal propagation

$$h_{nm}(t) = \int\int h_{11}\left(t, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right) e^{-j(m-1)\varphi_t(\theta_t)} e^{-j(n-1)\varphi_r(\theta_r)} d\theta_t \, d\theta_r$$

where
  - $\varphi_{r,t}\left(\theta_{r,t}\right) = 2\pi(d_{r,t}/\lambda)\cos\theta_{r,t}$,
  - d_r and d_t are the inter-element spacing at the receive/transmit arrays
  - $h_{11}\left(t, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right) \triangleq \int h_{11}\left(t, \tau, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right) d\tau$.
- Correlation between channels h_{nm} and h_{qp}

$$\begin{aligned}\mathcal{E}\left\{h_{nm}h_{qp}^*\right\} &= \mathcal{E}\left\{\int_0^{2\pi}\int_0^{2\pi}\left|h_{11}\left(t, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right)\right|^2 e^{-j(m-p)\varphi_t(\theta_t)} e^{-j(n-q)\varphi_r(\theta_r)} d\theta_t \, d\theta_r\right\} \\ &= \int_0^{2\pi}\int_0^{2\pi}\mathcal{E}\left\{\left|h_{11}\left(t, \mathbf{\Omega}_t, \mathbf{\Omega}_r\right)\right|^2\right\} e^{-j(m-p)\varphi_t(\theta_t)} e^{-j(n-q)\varphi_r(\theta_r)} d\theta_t \, d\theta_r, \\ &= \int_0^{2\pi}\int_0^{2\pi}\mathcal{A}\left(\theta_t, \theta_r\right) e^{-j(m-p)\varphi_t(\theta_t)} e^{-j(n-q)\varphi_r(\theta_r)} d\theta_t \, d\theta_r,\end{aligned}$$

where $\mathcal{A}\left(\theta_t, \theta_r\right)$ is the joint direction power spectrum restricted to the azimuth angles.
- The channel correlation is related to both the antenna spacings and the joint direction power spectrum!

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. The transmit correlation writes as

$$t = \int_0^{2\pi} e^{j2\pi(d_t/\lambda)\cos\theta_t} \mathcal{A}_t(\theta_t) d\theta_t,$$

which only depends on the transmit antenna spacing and the transmit direction power spectrum.

- *isotropic scattering*: very rich scattering environment around the transmitter with a uniform distribution of the energy, i.e. $\mathcal{A}_t(\theta_t) \cong 1/2\pi$

$$t = \frac{1}{2\pi}\int_0^{2\pi} e^{j\varphi_t(\theta_t)} d\theta_t = \frac{1}{2\pi}\int_0^{2\pi} e^{j2\pi(d_t/\lambda)\cos\theta_t} d\theta_t$$
$$= \mathrm{J}_0\left(2\pi\frac{d_t}{\lambda}\right).$$

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

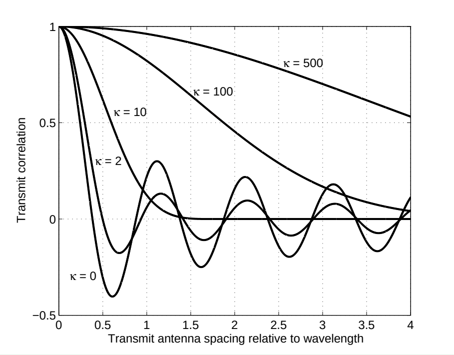
Spatial Correlation

Example

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

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

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



– $\mathcal{A}_t(\theta_t)$ in real-world channels: neither uniform nor a delta.

– isotropic scattering ($\kappa = 0$): first minimum for $d_t = 0.38\lambda$

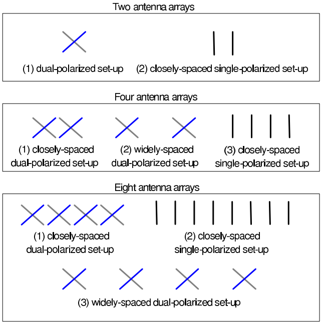
– directional scattering ($\kappa = \infty$): correlation never reaches 0

– 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.

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).
- What about real-world channels? Sometimes significantly deviate from this ideal channel:
  - *limited angular spread and/or reduced array sizes* cause the channels to become *correlated* (channels are not independent 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 (channels are not identically distributed anymore).



Correlated Rayleigh Fading Channels

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

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

  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

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

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

  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

$$\mathbf{R}_t = \mathcal{E}\left\{ \left[\begin{array}{c} h_1^* \\ h_2^* \end{array} \right] \left[\begin{array}{cc} h_1 & h_2 \end{array} \right] \right\} = \left[\begin{array}{cc} \mathcal{E}\left\{ |h_1|^2 \right\} & \mathcal{E}\left\{ h_1^* h_2 \right\} \\ \mathcal{E}\left\{ h_1 h_2^* \right\} & \mathcal{E}\left\{ |h_2|^2 \right\} \end{array} \right] = \left[\begin{array}{cc} 1 & t^* \\ t & 1 \end{array} \right]$$

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

$$\mathbf{R}_t = \left[\begin{array}{cc} 1 & 1 \\ t/|t| & -t/|t| \end{array} \right] \left[\begin{array}{cc} 1+|t| & 0 \\ 0 & 1-|t| \end{array} \right] \left[\begin{array}{cc} 1 & 1 \\ t/|t| & -t/|t| \end{array} \right]^H .$$

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

□

Correlated Rayleigh Fading Channels

Example

Question: Assume the previous example with $|t| \to 1$. Compute the weights of the matched beamformer (or maximum ratio transmission/transmit MRC).
Answer: With matched beamforming, $\mathbf{w} = \mathbf{h}^H / \|\mathbf{h}\|$ where

$$\begin{aligned}
\mathbf{h} &= \mathbf{h}_w \mathbf{R}_t^{1/2} \\
&= \mathbf{h}_w \begin{bmatrix} 1 & 1 \\ t/|t| & -t/|t| \end{bmatrix} \begin{bmatrix} \sqrt{1+|t|} & 0 \\ 0 & \sqrt{1-|t|} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ t/|t| & -t/|t| \end{bmatrix}^H \\
&= 2\mathbf{h}_w \begin{bmatrix} 1 \\ t/|t| \end{bmatrix} \begin{bmatrix} 1 \\ t/|t| \end{bmatrix}^H
\end{aligned}$$

where the last equality comes from the fact that $|t| = 1$. This shows that for high correlation, the channel direction $(\mathbf{h}/\|\mathbf{h}\|)$ is aligned with $\begin{bmatrix} 1 & t^*/|t| \end{bmatrix}$. Hence

$$\mathbf{w} = \mathbf{h}^H / \|\mathbf{h}\| = \begin{bmatrix} 1 \\ t/|t| \end{bmatrix}.$$

Transmission is performed in the direction where all scatterers are located. □

MIMO - An Interpretation using Radiation Patterns

- MIMO system with n_t transmit and n_r receive antennas communicating through a frequency flat-fading channel
- At the k^{th} time instant, the transmitted and received signals are related by

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

  where

  - $\mathbf{c}_k$ is the $n_t \times 1$ transmitted signal vector
  - $\mathbf{y}_k$ 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 $\mathcal{E}\{\mathbf{n}_k\mathbf{n}_l^H\} = \sigma_n^2\mathbf{I}_{n_r}\delta\left(k-l\right)$,

Radiation Patterns

- Decompose the channel $\mathbf{H}_k = \sum_{l=0}^{L-1} \mathbf{H}_k^{(l)} = \sum_{l=0}^{L-1} \mathbf{H}_k^{(l)}(:,1)\,\mathbf{a}_t^T\left(\theta_{t,k}^{(l)}\right)$, where $\mathbf{a}_t\left(\theta_{t,k}^{(l)}\right)$ is the transmit array response in the direction of departure $\theta_{t,k}^{(l)}$.
- Hence,

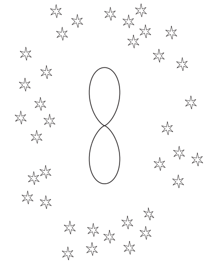
$$\mathbf{H}_k \mathbf{c}_k = \sum_{l=0}^{L-1} \mathbf{H}_k^{(l)}(:,1)\mathbf{a}_t^T\left(\theta_{t,k}^{(l)}\right)\mathbf{c}_k$$

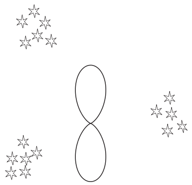
- The original MIMO transmission can be considered as the SIMO transmission of an equivalent codeword, given at the k^{th} time instant by

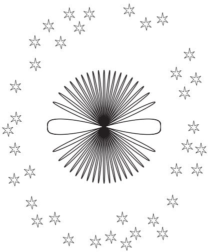
$$\mathbf{a}_t^T \mathbf{c}_k$$

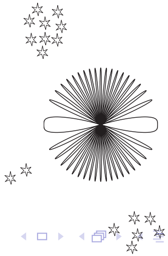
- It may be thought of as an array factor function of the transmitted codewords. At every symbol period,
  - the energy radiated in any direction varies as a function of the transmitted codewords.
  - for a given codeword and omnidirectional antennas, the radiated energy is not uniformly distributed in all directions, but may present maxima and minima in certain directions.

Radiation Patterns







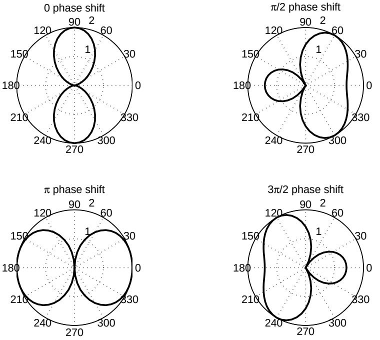


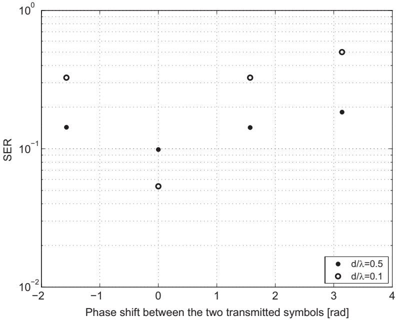
Radiation Patterns

Example

$n_t = 2$: $\mathbf{c}_k = [\ c_1[k] \quad c_2[k]\ ]^T$

$$\mathbf{c}_k^T \mathbf{a}_t(\theta_t) = c_1[k] \underbrace{\left[1 + \frac{c_2[k]}{c_1[k]} e^{-2\pi j \frac{d_t}{\lambda}\cos\theta_t}\right]}_{G_t(\theta_t|\mathbf{c}_k)}$$

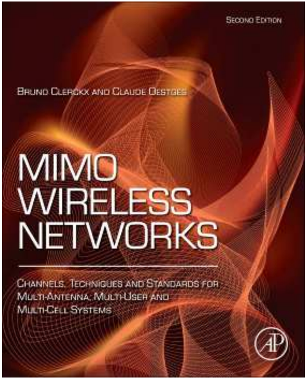




Capacity of point-to-point MIMO Channels

Reference Book

- 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.



  - Chapter 5
    - Section: 5.1, 5.2, 5.3, 5.4.1, 5.4.2 (except "Antenna Selection Schemes"), 5.5.1 - "Kronecker Correlated Rayleigh Channels", 5.5.2, 5.7, 5.8.1 (except Proof of Proposition 5.9 and Example 5.4)

Introduction - 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

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

  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 $\mathbf{Q} = \mathcal{E}\left\{\mathbf{c}'\mathbf{c}'^H\right\}$.
- Short-term power constraint: $\text{Tr}\{\mathbf{Q}\} \leq 1$.
- Long-term power constraint (over a duration $T_p >> T$): $\mathcal{E}\left\{\text{Tr}\{\mathbf{Q}\}\right\} \leq 1$ where the expectation refers here to the averaging over successive codeword of length T.
- 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 MIMO Channels

Proposition

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

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

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

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

Note the difference with SISO capacity.

Capacity and Water-Filling Algorithm

- 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)

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

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

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

$$\mathbf{Q}^\star = \mathbf{V_H} \text{diag}\left\{s_1^\star, \ldots, s_n^\star\right\} \mathbf{V_H}^H,$$

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

Proposition

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

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

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

Water-Filling Algorithm

- Iterative power allocation
  - Order eigenvalues λ_k in decreasing order of magnitude
  - At iteration i, evaluate the constant μ from the power constraint

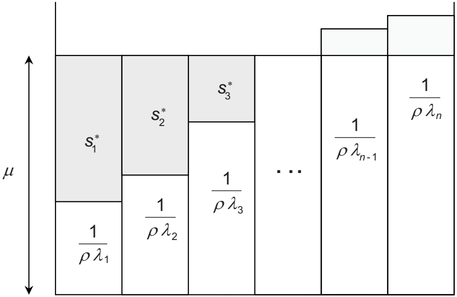
$$\mu(i) = \frac{1}{n-i+1}\left(1 + \sum_{k=1}^{n-i+1} \frac{1}{\rho\lambda_k}\right)$$

  - Calculate power

$$s_k(i) = \mu(i) - \frac{1}{\rho\lambda_k}, \qquad k = 1, \ldots, n-i+1.$$

    If $s_{n-i+1} < 0$, set to 0

  - Iterate till the power allocated on each mode is non negative.



Water-Filling Algorithm

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

$$\mathbf{H} = \begin{bmatrix} a & 0 & a & 0 \\ 0 & b & 0 & b \end{bmatrix}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

   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.

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

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

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

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

- At any SNR
  - lower bound

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

  - upper bound (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)

$$C_{CSIT}(\mathbf{H}) = \sum_{k=1}^{n}\log_2\left[1+\rho s_k^{\star}\lambda_k\right] \overset{(a)}{\leq} n\log_2\left(1+\frac{\rho}{n}\left[\sum_{k=1}^{n}s_k^{\star}\lambda_k\right]\right),$$
$$\leq n\log_2\left[1+\frac{\rho}{n}\lambda_{max}\right].$$

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
- MIMO Capacity with Perfect Transmit Channel Knowledge
  - similar strategy as in deterministic channels: transmit along eigenvectors of channel matrix and allocate power following water-filling
  - short term power constraint: water-filling solution applied over space as in deterministic channels

$$
\begin{aligned}
\bar{C}_{CSIT,ST} &= \mathcal{E}\left\{ \max_{\mathbf{Q}\geq 0:\mathsf{Tr}\{\mathbf{Q}\}=1} \log_2 \det \left[\mathbf{I}_{n_r} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H \right] \right\} \\
&= \sum_{k=1}^{n} \mathcal{E}\left\{ \log_2 \left[1 + \rho s_k^{\star} \lambda_k \right] \right\}.
\end{aligned}
$$

  - Impact on coding strategy? Use a variable-rate code (family of codes of different rates) adapted as a function of the water-filling allocation. No need for the codeword to span many coherence time periods.

MIMO Capacity with Partial Transmit Channel Knowledge

- $\mathbf{H}$ 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 $\mathbf{H}_k$

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

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

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

Definition

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

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

where $\mathbf{Q}$ is the input covariance matrix optimized as to maximize the ergodic mutual information.

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

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

- Low SNR: allocate all the available power to the strongest or dominant eigenmode. Use $\log_2(1+x) \approx x \log_2(e)$ for x small and get

$$\begin{aligned}
\bar{C}_{CSIT,ST} &= \mathcal{E}\left\{ \log_2 \left[1 + \rho\lambda_{max}\right] \right\} \\
&\cong \rho\mathcal{E}\left\{\lambda_{max}\right\} \log_2(e) \\
&\cong \rho n \log_2(e), \qquad N, n \to \infty, N/n >> 0. \\
\bar{C}_{CSIT,LT} &= \mathcal{E}\left\{ \log_2 \left[1 + \rho s^{\star}_{max}\lambda_{max}\right] \right\} \\
&\cong \rho\mathcal{E}\left\{s^{\star}_{max}\lambda_{max}\right\} \log_2(e)
\end{aligned}$$

  Observations: $\bar{C}_{CSIT}$ grows linearly in the minimum number of antennas n.
- High SNR: uniform power allocation on all non-zeros eigenmodes

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

  Observations: $\bar{C}_{CSIT}$ also scales linearly with n. The spatial multiplexing gain is $g_s = n$. MISO fading channels do not offer any multiplexing gain.

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

- Optimal covariance matrix

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.,

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

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

- Low SNR:

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

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

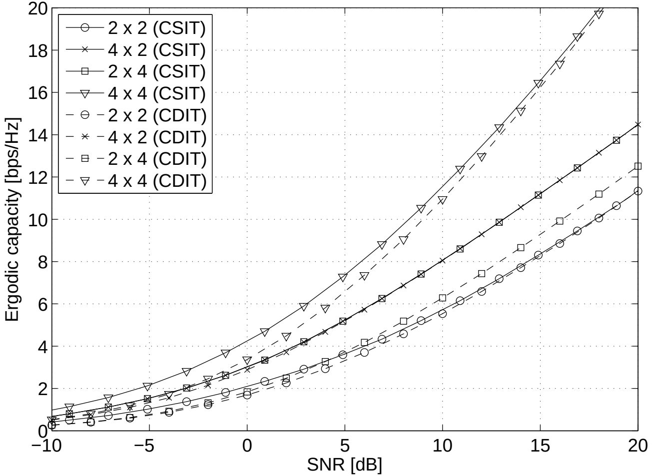
$$\bar{C}_{CDIT} \approx \log_2(n_r\rho)$$

  - 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

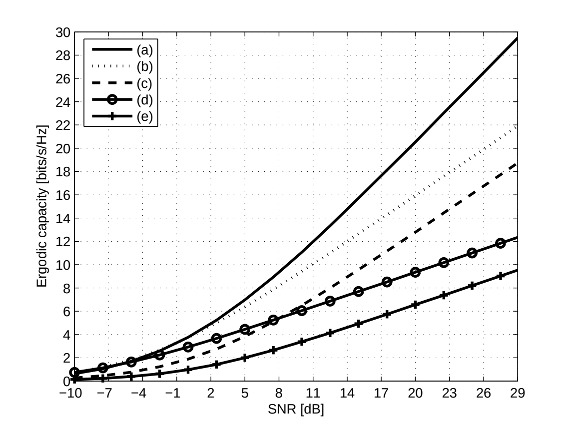
- Ergodic capacity of various $n_r \times n_t$ i.i.d. Rayleigh channels with full (CSIT) and partial (CDIT) channel knowledge at the transmitter.



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 $n_t + n_r = 8$.



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) $g_s = 1$.

I.I.D. Rayleigh Fast Fading Channels

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

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

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

- Low SNR

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

- 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 $n_t = n_r$

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

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 $\min\{n_t, n_r\}$

Correlated Rayleigh Fast Fading Channels: Partial Transmit Channel Knowledge

- Assume the channel covariance matrix is known to the transmitter.

Proposition

In Kronecker correlated Rayleigh fast fading channels, the optimal input covariance matrix can again be expressed as

$$\mathbf{Q} = \mathbf{U}_{\mathbf{R}_t} \mathbf{\Lambda}_{\mathbf{Q}} \mathbf{U}_{\mathbf{R}_t}^H,$$

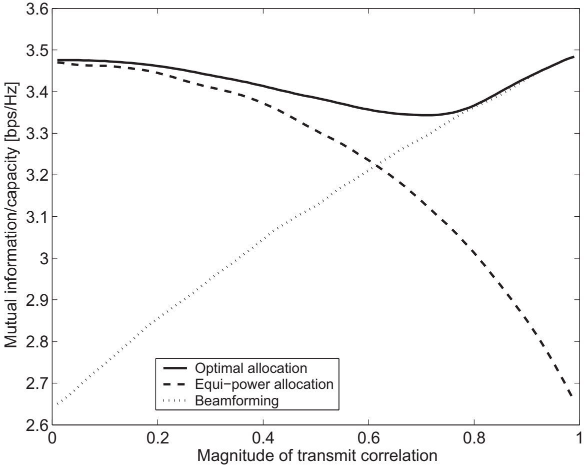
where $\mathbf{U}_{\mathbf{R}_t}$ is a unitary matrix formed by the eigenvectors of $\mathbf{R}_t$ (arranged in such order that they correspond to decreasing eigenvalues of $\mathbf{R}_t$), and $\mathbf{\Lambda}_{\mathbf{Q}}$ is a diagonal matrix whose elements are also arranged in decreasing order.

Power allocation has to be computed numerically. Approximation using Jensen's inequality is possible.

- Spatial correlation: beneficial or detrimental?
  - receive correlations degrade both the mutual information $\bar{\mathcal{I}}_e$ and the capacity with CDIT,
  - transmit correlations always decrease $\bar{\mathcal{I}}_e$ but may increase $\bar{C}_{CDIT}$ at low SNR (irrespective of n_t and n_r) or at higher SNR when $n_t > n_r$ (analogous to the full CSIT case).

Correlated Rayleigh Fast Fading Channels: Partial Transmit Channel Knowledge

- Mutual information of various strategies at 0 dB SNR as a function of the transmit correlation $|t|$ in TIMO. Beamforming refers here to the tranmsission of one stream along the dominant eigenvector of $\mathbf{R}_t$.



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

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

  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

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

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

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

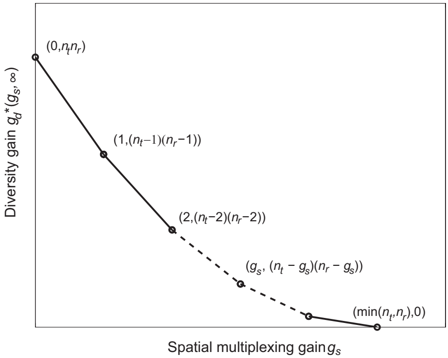
$$\begin{aligned}\lim_{\rho\longrightarrow\infty}\frac{R(\rho)}{\log_2(\rho)} &= g_s\\ \lim_{\rho\longrightarrow\infty}\frac{\log_2\left(P_{out}(R)\right)}{\log_2(\rho)} &= -g_d^\star(g_s,\infty)\end{aligned}$$

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.



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

- 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.

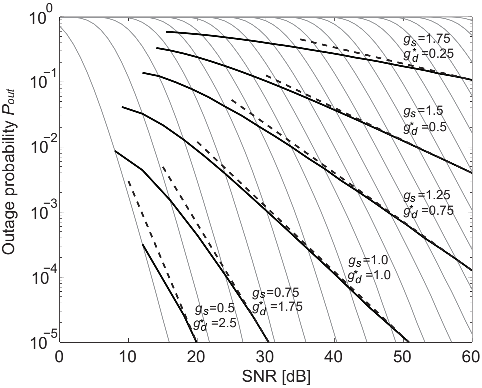


Figure: 2×2 MIMO i.i.d. Rayleigh fading channels

Diversity-Multiplexing Trade-Off of a Scalar Rayleigh Channel h

- 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

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

- At high SNR,

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

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

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

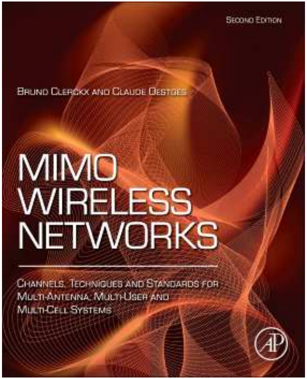
An outage occurs at high SNR when $|h|^2 \leq \rho^{-(1-g_s)}$ with a probability $\rho^{-(1-g_s)}$.

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

Space-Time Coding over I.I.D. Rayleigh Flat Fading Channels

Reference Book

- 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.



- Chapter 6
  - Section: 6.1, 6.2, 6.3 (except "Antenna Selection" in 6.3.2), 6.4.1, 6.4.2 (except the Proofs), 6.5.1, 6.5.2, 6.5.3, 6.5.4, 6.5.8, Figure 7.1

Introduction - Previous Lectures

- Previous lecture
  - Capacity of deterministic MIMO channels

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

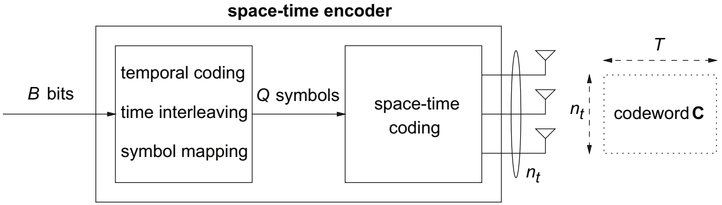
  - Ergodic capacity of fast fading channels
  - Outage capacity and probability of slow fading channels

- 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



- 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).

System Model

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

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

  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

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

  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,

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

  where $\mathcal{Q}(x)$ is the Gaussian $\mathcal{Q}$-function.
- The average PEP, $P(\mathbf{C} \to \mathbf{E})$, 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.

- 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

- Conditional PEP

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

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

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

- Average PEP

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

- This integration is sometimes difficult to calculate. Therefore, alternatives forms of the Gaussian $\mathcal{Q}$-function are used.
  - Craig's formula

$$\mathcal{Q}(x) = \frac{1}{\pi}\int_0^{\pi/2} \exp\left(-\frac{x^2}{2\sin^2(\beta)}\right) d\beta.$$

  - Chernoff bound

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

- We can derive the average PEP as follows

$$\begin{aligned} P\left(\mathbf{C} \rightarrow \mathbf{E}\right) &= \mathcal{E}_{\mathbf{H}_k}\left\{P\left(\mathbf{C} \rightarrow \mathbf{E} \middle| \left\{\mathbf{H}_k\right\}_{k=0}^{T-1}\right)\right\} \\ &= \frac{1}{\pi}\int_0^{\pi/2} M_\Gamma\left(-\frac{1}{2\sin^2\left(\beta\right)}\right)d\beta \\ &\leq M_\Gamma\left(-\frac{1}{2}\right) \end{aligned}$$

with $M_\Gamma\left(\gamma\right)$ moment generating function (MGF) of $\Gamma = \frac{\rho}{2}\sum_{k=0}^{T-1}\left\|\mathbf{H}_k\left(\mathbf{c}_k - \mathbf{e}_k\right)\right\|_F^2$

$$M_\Gamma\left(\gamma\right) \triangleq \int_0^\infty \exp\left(\gamma\Gamma\right)p_\Gamma\left(\Gamma\right)d\Gamma$$

Theorem

The moment generating function of a Hermitian quadratic form in complex Gaussian random variable $y = \mathbf{z}\mathbf{F}\mathbf{z}^H$, where $\mathbf{z}$ is a circularly symmetric complex Gaussian vector with mean $\bar{\mathbf{z}}$ and a covariance matrix $\mathbf{R}_\mathbf{z}$ and $\mathbf{F}$ a Hermitian matrix, is given by

$$M_y\left(s\right) \triangleq \int_0^\infty \exp\left(sy\right)p_y\left(y\right)dy = \frac{\exp\left(s\bar{\mathbf{z}}\mathbf{F}\left(\mathbf{I} - s\mathbf{R}_\mathbf{z}\mathbf{F}\right)^{-1}\bar{\mathbf{z}}^H\right)}{\det\left(\mathbf{I} - s\mathbf{R}_\mathbf{z}\mathbf{F}\right)}$$

Fast Fading MIMO Channels

- Defining

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \cdots & \mathbf{H}_T \end{bmatrix}$$
$$\mathbf{D} = \text{diag}\left\{\mathbf{c}_1 - \mathbf{e}_1, \mathbf{c}_2 - \mathbf{e}_2, \ldots, \mathbf{c}_T - \mathbf{e}_T\right\},$$

  we may write

$$\sum_{k=1}^{T} \|\mathbf{H}_k\left(\mathbf{c}_k - \mathbf{e}_k\right)\|_F^2 = \|\mathbf{H}\mathbf{D}\|_F^2 = \text{Tr}\left\{\mathbf{H}\mathbf{D}\mathbf{D}^H\mathbf{H}^H\right\} = \text{vec}\left(\mathbf{H}^H\right)^H \boldsymbol{\Delta}\, \text{vec}\left(\mathbf{H}^H\right)$$

  where $\boldsymbol{\Delta} = \mathbf{I}_{n_r} \otimes \mathbf{D}\mathbf{D}^H$. This is a hermitian quadratic form of complex Gaussian random variables of the form $\mathbf{z}\mathbf{F}\mathbf{z}^H$ (with $\mathbf{z} = \text{vec}\left(\mathbf{H}^H\right)^H$ and $\mathbf{F} = \boldsymbol{\Delta}$).
- We can use Theorem where the mean $\bar{\mathbf{z}} = \mathbf{0}$ is the zero vector and the covariance matrix is $\mathbf{R}_{\mathbf{z}} = \mathcal{E}\left\{\text{vec}\left(\mathbf{H}^H\right)\text{vec}\left(\mathbf{H}^H\right)^H\right\} = \mathbf{I}_{Tn_rn_t}$ (for i.i.d. channels).
- PEP averaged over i.i.d. Rayleigh fast fading channel (with $\eta = \rho/(4\sin^2\beta)$)

$$\begin{aligned} P\left(\mathbf{C} \rightarrow \mathbf{E}\right) &= \frac{1}{\pi}\int_0^{\pi/2} \left(\det\left(\mathbf{I}_{Tn_rn_t} + \eta\mathbf{R}_{\mathbf{z}}\boldsymbol{\Delta}\right)\right)^{-1} d\beta \\ &= \frac{1}{\pi}\int_0^{\pi/2} \left(\det\left(\mathbf{I}_{Tn_t} + \eta\mathbf{D}\mathbf{D}^H\right)\right)^{-n_r} d\beta \\ &= \frac{1}{\pi}\int_0^{\pi/2} \left(\det\left(\mathbf{I}_T + \eta\mathbf{D}^H\mathbf{D}\right)\right)^{-n_r} d\beta. \end{aligned}$$

- In i.i.d. Rayleigh fast fading channels, average PEP reads as

$$P(\mathbf{C} \to \mathbf{E}) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{k=0}^{T-1} \left(1 + \eta \left\| \mathbf{c}_k - \mathbf{e}_k \right\|^2\right)^{-n_r} d\beta$$

where $\eta = \rho/(4\sin^2 \beta)$.
- Upper bound using the Chernoff bound

$$P(\mathbf{C} \to \mathbf{E}) \leq \prod_{k=0}^{T-1} \left(1 + \frac{\rho}{4} \left\| \mathbf{c}_k - \mathbf{e}_k \right\|^2\right)^{-n_r}.$$

- In the high SNR regime, the average PEP is further upper-bounded by

$$P(\mathbf{C} \to \mathbf{E}) \leq \left(\frac{\rho}{4}\right)^{-n_r l_{\mathbf{C},\mathbf{E}}} \prod_{k \in \tau_{\mathbf{C},\mathbf{E}}} \left\| \mathbf{c}_k - \mathbf{e}_k \right\|^{-2n_r}$$

with $l_{\mathbf{C},\mathbf{E}}$ the effective length of the pair of codewords $\{\mathbf{C}, \mathbf{E}\}$, i.e., $l_{\mathbf{C},\mathbf{E}} = \sharp\tau_{\mathbf{C},\mathbf{E}}$ with $\tau_{\mathbf{C},\mathbf{E}} = \{k \mid \mathbf{c}_k - \mathbf{e}_k \neq 0\}$.
- Diversity gain: $n_r l_{\mathbf{C},\mathbf{E}}$, coding gain: $\prod_{k \in \tau_{\mathbf{C},\mathbf{E}}} \left\| \mathbf{c}_k - \mathbf{e}_k \right\|^{-2n_r}$.

The Distance-Product Criterion

- At high SNR, the error probability is naturally dominated by the worst-case PEP

Design Criterion

(**Distance-product criterion**) *Over i.i.d. Rayleigh fast fading channels,*

1. **distance criterion**: *maximize the minimum effective length* L_{min} *of the code over all pairs of codewords* $\{\mathbf{C}, \mathbf{E}\}$ *with* $\mathbf{C} \neq \mathbf{E}$

$$L_{min} = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E}}} l_{\mathbf{C},\mathbf{E}}$$

2. **product criterion**: *maximize the minimum product distance* d_p *of the code over all pairs of codewords* $\{\mathbf{C}, \mathbf{E}\}$ *with* $\mathbf{C} \neq \mathbf{E}$

$$d_p = \min_{\substack{\mathbf{C},\mathbf{E} \\ \mathbf{C}\neq\mathbf{E} \\ l_{\mathbf{C},\mathbf{E}}=L_{min}}} \prod_{k\in\tau_{\mathbf{C},\mathbf{E}}} \|\mathbf{c}_k - \mathbf{e}_k\|^2$$

- The presence of multiple antennas at the transmitter does not impact the achievable diversity gain $g_d^o(\infty) = n_r L_{min}$ but improves the coding gain $g_c = d_p$.
- The diversity gain is maximized first, and the coding gain is maximized only in a second step.

Slow Fading MIMO Channels

- 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

$$P\left(\mathbf{C}\to\mathbf{E}\right)=\mathcal{E}_{\mathbf{H}}\left\{P\left(\mathbf{C}\to\mathbf{E}|\mathbf{H}\right)\right\}=\frac{1}{\pi}\int_0^{\pi/2}M_\Gamma\left(-\frac{1}{2\sin^2\left(\beta\right)}\right)d\beta$$

  where $M_\Gamma\left(\gamma\right)$ moment generating function (MGF) of $\Gamma=\frac{\rho}{2}\left\|\mathbf{H}\left(\mathbf{C}-\mathbf{E}\right)\right\|_F^2$.
- Note that

$$\left\|\mathbf{H}\left(\mathbf{C}-\mathbf{E}\right)\right\|_F^2=\mathrm{Tr}\left\{\mathbf{H}\tilde{\mathbf{E}}\mathbf{H}^H\right\}=\mathrm{vec}\left(\mathbf{H}^H\right)^H\left(\mathbf{I}_{n_r}\otimes\tilde{\mathbf{E}}\right)\mathrm{vec}\left(\mathbf{H}^H\right)$$

  where $\tilde{\mathbf{E}}=\left(\mathbf{C}-\mathbf{E}\right)\left(\mathbf{C}-\mathbf{E}\right)^H$.
- This is a hermitian quadratic form of complex gaussian random variables of the form $\mathbf{z}\mathbf{F}\mathbf{z}^H$ (with $\mathbf{z}=\mathrm{vec}\left(\mathbf{H}^H\right)^H$ and $\mathbf{F}=\mathbf{I}_{n_r}\otimes\tilde{\mathbf{E}}$) and we can use Theorem where the mean $\bar{\mathbf{z}}=\mathbf{0}$ is the zero vector and the covariance matrix is $\mathbf{R}_{\mathbf{z}}=\mathcal{E}\left\{\mathrm{vec}\left(\mathbf{H}^H\right)\mathrm{vec}\left(\mathbf{H}^H\right)^H\right\}=\mathbf{I}_{n_rn_t}$ (for i.i.d. channels).

- In i.i.d. Rayleigh slow fading channels, average PEP reads as

$$P\left(\mathbf{C} \rightarrow \mathbf{E}\right) = \frac{1}{\pi}\int_0^{\pi/2}\left(\det\left(\mathbf{I}_{n_r n_t} + \eta\mathbf{F}\right)\right)^{-1} d\beta = \frac{1}{\pi}\int_0^{\pi/2}\left[\det\left(\mathbf{I}_{n_t} + \eta\tilde{\mathbf{E}}\right)\right]^{-n_r} d\beta.$$

- Upper bound using the Chernoff bound

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

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

- At high SNR, $\frac{\rho}{4}\lambda_i(\tilde{\mathbf{E}}) >> 1$

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

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

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* $\mathbf{C} \neq \mathbf{E}$

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

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* $\tilde{\mathbf{E}}$,

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

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

Definition

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

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

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

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

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

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

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

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

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, $r(\mathbf{C}-\mathbf{E})\leq 2$. Actually, $r(\mathbf{C}-\mathbf{E})=2$ 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

Example

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

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

What is the diversity gain achieved by this code over slow Rayleigh fading channels?
Answer: Check the rank of its error matrix

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

where $d_k = c_k - e_k$ for $k = 1, ..., 4$. 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), $r\left(\mathbf{C} - \mathbf{E}\right) = 1$. Hence diversity gain of n_r.

Information Theory Motivated Design Methodology: Fast Fading MIMO Channels - Achieving The Ergodic Capacity

- Recall Lecture on ergodic capacity

$$\bar{C} = \max_{\mathbf{Q}:\mathrm{Tr}\{\mathbf{Q}\}=1} \mathcal{E}\left\{\log_2 \det\left(\mathbf{I}_{n_r} + \rho\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\right\}.$$

- Perfect Transmit Channel Knowledge
  - 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 $\lambda_k(\mathbf{Q})$ and be decoded independently of the other streams.
  - The optimal power allocation $\{\lambda_k^\star\}$ based on the water-filling allocation strategy.
  - Capacity achievable using a variable-rate coding strategy ($T = T_{coh}$ is enough as long as the noise can be averaged out).
- Partial Transmit Channel Knowledge
  - 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).

Information Theory Motivated Design Methodology: Slow Fading MIMO Channels Achieving The DMT

- 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!
- Rank-determinant criterion focuses on diversity maximization under fixed rate.
- What if we want to design codes achieving the diversity-multiplexing trade-off?

Definition

A scheme $\{\mathcal{C}(\rho)\}$, i.e., a family of codes indexed by the SNR ρ, is said to achieve a diversity gain $g_d(g_s, \infty)$ and a multiplexing gain g_s at high SNR if

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

$$\lim_{\rho\to\infty} \frac{\log_2(P_e(\rho))}{\log_2(\rho)} = -g_d(g_s, \infty)$$

where $R(\rho)$ is the data rate and $P_e(\rho)$ the average error probability averaged over the additive noise, the i.i.d. channel statistics and the transmitted codewords. The curve $g_d(g_s, \infty)$ is the diversity-multiplexing trade-off achieved by the scheme in the high SNR regime.

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

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

A General Framework for Linear STBCs

- A linear STBC is expressed in its general form as

$$\mathbf{C} = \sum_{q=1}^{Q} \mathbf{\Phi}_q \Re[c_q] + \mathbf{\Phi}_{q+Q} \Im[c_q]$$

where
  - $\mathbf{\Phi}_q$ are complex basis matrices of size $n_t \times T$,
  - c_q stands for the complex information symbol (taken for example from PSK or QAM constellations),
  - Q is the number of complex symbols c_q transmitted over a codeword,
  - $\Re$ and $\Im$ stand for the real and imaginary parts.

Definition

Tall $(T \leq n_t)$ unitary basis matrices are such that $\mathbf{\Phi}_q^H \mathbf{\Phi}_q = \frac{1}{Q}\mathbf{I}_T$ $\forall q = 1, \ldots, 2Q$. *Wide* $(T \geq n_t)$ unitary basis matrices are such that $\mathbf{\Phi}_q \mathbf{\Phi}_q^H = \frac{T}{Qn_t}\mathbf{I}_{n_t}$ $\forall q = 1, \ldots, 2Q$.

Definition

The spatial multiplexing rate of a space-time block code is defined as $r_s = \frac{Q}{T}$. A full rate space-time block code is characterized by $r_s = n_t$.

A General Framework for Linear STBCs

- Average Pairwise Error Probability of STBCs

Proposition

A PSK/QAM based linear STBC consisting of unitary basis matrices minimizes the worst-case PEP averaged over i.i.d. Rayleigh slow fading channels if (sufficient condition) the unitary basis matrices $\{\boldsymbol{\Phi}_q\}_{q=1}^{2Q}$ *satisfy the conditions*

$$\boldsymbol{\Phi}_q \boldsymbol{\Phi}_p^H + \boldsymbol{\Phi}_p \boldsymbol{\Phi}_q^H = \mathbf{0}_{n_t}, \quad q \neq p \text{ for wide } \{\boldsymbol{\Phi}_q\}_{q=1}^{2Q},$$
$$\boldsymbol{\Phi}_q^H \boldsymbol{\Phi}_p + \boldsymbol{\Phi}_p^H \boldsymbol{\Phi}_q = \mathbf{0}_T, \quad q \neq p \text{ for tall } \{\boldsymbol{\Phi}_q\}_{q=1}^{2Q}.$$

Proof: Using Hadamard's inequality and the unitarity of basis matrices,

$$\min_{q=1,\ldots,Q} \min_{d_q} \det\left(\mathbf{I}_{n_t} + \eta \tilde{\mathbf{E}}\right) \leq \det\left(\mathbf{I}_{n_t} + \eta \frac{T}{Q n_t} \mathbf{I}_{n_t} d_{min}^2\right).$$

Equality if unitary basis matrices are skew-hermitian. □

A General Framework for Linear STBCs

- Decoding

Proposition

Applying the space-time matched filter to the output vector decouples the transmitted symbols if and only if the basis matrices are wide unitary

$$\mathbf{\Phi}_q \mathbf{\Phi}_q^H = \frac{T}{Q n_t} \mathbf{I}_{n_t}, \ \forall q = 1, \ldots, 2Q$$

and pairwise skew-hermitian

$$\mathbf{\Phi}_q \mathbf{\Phi}_p^H + \mathbf{\Phi}_p \mathbf{\Phi}_q^H = \mathbf{0}_{n_t}, \ \forall q \neq p.$$

The complexity of ML decoding of linear STBCs grows exponentially with n_t and Q. The decoupling property allows each symbol to be decoded independently of the presence of the other symbols through a simple space-time matched filter.

Example

A code such that $T = 1$, $n_t = 2$, $Q = 2$, $r_s = 2$ with the following **tall** basis matrices

$$\mathbf{\Phi}_1 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ 0 \end{array} \right], \; \mathbf{\Phi}_2 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 0 \\ 1 \end{array} \right], \; \mathbf{\Phi}_3 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} j \\ 0 \end{array} \right], \; \mathbf{\Phi}_4 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 0 \\ j \end{array} \right],$$

This code is called Spatial Multiplexing. Optimal for worst-case PEP min. spectrally-efficient, large decoding complexity.

A General Framework for Linear STBCs

Example

A code such that $T = 2$, $n_t = 2$, $Q = 2$, $r_s = 1$ with the following **wide** basis matrices

$$\begin{aligned} \mathbf{\Phi}_1 &= \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{\Phi}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \\ \mathbf{\Phi}_3 &= \tfrac{1}{\sqrt{2}} \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}, \quad \mathbf{\Phi}_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}, \end{aligned}$$

This code is called Alamouti code. Optimal for worst-case PEP min. not spectrally-efficient, low decoding complexity.

Spatial Multiplexing/V-BLAST/D-BLAST

- 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 $n_t \times 1$.

Example

$$\mathbf{C} = \frac{1}{\sqrt{n_t}} \begin{bmatrix} c_1 & \ldots & c_{n_t} \end{bmatrix}^T = \frac{1}{\sqrt{n_t}} \sum_{q=1}^{n_t} \mathbf{I}_{n_t}(:,q)\,\Re[c_q] + j\mathbf{I}_{n_t}(:,q)\,\Im[c_q].$$

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

ML decoding

- Error probability

$$
\begin{aligned}
P(\mathbf{C} \to \mathbf{E}) &= \frac{1}{\pi} \int_0^{\pi/2} \left[\det\left(\mathbf{I}_{n_t} + \eta \tilde{\mathbf{E}}\right)\right]^{-n_r} d\beta \\
&= \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{\eta}{n_t} \sum_{q=1}^{n_t} |c_q - e_q|^2\right)^{-n_r} d\beta \\
&\leq \left(\frac{\rho}{4n_t}\right)^{-n_r} \left(\sum_{q=1}^{n_t} |c_q - e_q|^2\right)^{-n_r}
\end{aligned}
$$

  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.

ML decoding

- Over slow fading channels, what is the multiplexing-diversity trade-off achieved by SM with ML decoding?

Proposition

For $n_r \geq n_t$, the diversity-multiplexing trade-off at high SNR achieved by Spatial Multiplexing with ML decoding and QAM constellation over i.i.d. Rayleigh fading channels is given by

$$g_d\left(g_s, \infty\right) = n_r \left(1 - \frac{g_s}{n_t}\right), \quad g_s \in [0, n_t]\,.$$

Zero-Forcing (ZF) Linear Receiver

- MIMO ZF receiver acts similarly to a ZF equalizer in frequency selective channels.
- ZF filtering effectively decouples the channel into n_t parallel channels
  - interference from other transmitted symbols is suppressed
  - scalar decoding may be performed on each of these channels
- The complexity of ZF decoding similar to SISO ML decoding, but the inversion step is responsible for the noise enhancement (especially at low SNR).
- Assuming that a symbol vector $\mathbf{C} = 1/\sqrt{n_t}\begin{bmatrix} c_1 & \dots & c_{n_t} \end{bmatrix}^T$ is transmitted, the output of the ZF filter $\mathbf{G}_{ZF}$ is given by

$$\mathbf{z} = \mathbf{G}_{ZF}\mathbf{y} = \begin{bmatrix} c_1 & \dots & c_{n_t} \end{bmatrix}^T + \mathbf{G}_{ZF}\mathbf{n}$$

where $\mathbf{G}_{ZF}$ inverts the channel,

$$\mathbf{G}_{ZF} = \sqrt{\frac{n_t}{E_s}}\mathbf{H}^\dagger$$

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

$$\mathcal{E}\left\{\mathbf{G}_{ZF}\mathbf{n}\left(\mathbf{G}_{ZF}\mathbf{n}\right)^H\right\} = \frac{n_t}{\rho}\mathbf{H}^{\dagger}\left(\mathbf{H}^{\dagger}\right)^H = \frac{n_t}{\rho}\left(\mathbf{H}^H\mathbf{H}\right)^{-1}.$$

- The output SNR on the q^{th} subchannel is thus given by

$$\rho_q = \frac{\rho}{n_t}\frac{1}{\left(\mathbf{H}^H\mathbf{H}\right)^{-1}(q,q)}, \qquad q = 1, \ldots, n_t.$$

- 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 $2\left(n_r - n_t + 1\right)$ 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 $n = \min\left\{n_t,n_r\right\}$ independent data pipes.

Zero-Forcing (ZF) Linear Receiver

- In slow fading channels, what is the diversity-multiplexing trade-off achieved by SM with ZF?

Proposition

For $n_r \geq n_t$, the diversity-multiplexing trade-off achieved by Spatial Multiplexing with QAM constellation and ZF filtering in i.i.d. Rayleigh fading channels is given by

$$g_d\left(g_s, \infty\right) = \left(n_r - n_t + 1\right)\left(1 - \frac{g_s}{n_t}\right), \;\; g_s \in \left[0, n_t\right].$$

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 $\mathbf{H}(:,p)$, $p \neq q$.
- Assume the following system model with $n_r \geq n_t$

$$\begin{aligned}\mathbf{y} &= \mathbf{Hc} + \mathbf{n}, \\ &= \mathbf{h}_q c_q + \sum_{p \neq q} \mathbf{h}_p c_p + \mathbf{n}\end{aligned}$$

  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 $p \neq q$:

$$\begin{aligned}\mathbf{H}_{-q} &= \begin{bmatrix} \dots & \mathbf{h}_p & \dots \end{bmatrix}_{p \neq q}, \\ &= \begin{bmatrix} \mathbf{U}' & \tilde{\mathbf{U}} \end{bmatrix} \mathbf{\Lambda} \mathbf{V}^H\end{aligned}$$

  where $\tilde{\mathbf{U}}$ is the matrix containing the left singular vectors corresponding to the null singular values. Similarly we define

$$\mathbf{c}_{-q} = \begin{bmatrix} \dots & c_p & \dots \end{bmatrix}^T_{p \neq q}.$$

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 $\mathbf{H}_{-q}$

$$\begin{aligned}\tilde{\mathbf{U}}^H\mathbf{y} &= \tilde{\mathbf{U}}^H\mathbf{h}_q c_q + \tilde{\mathbf{U}}^H\mathbf{H}_{-q}\mathbf{c}_{-q} + \tilde{\mathbf{U}}^H\mathbf{n} \\ &= \tilde{\mathbf{U}}^H\mathbf{h}_q c_q + \tilde{\mathbf{U}}^H\mathbf{n}.\end{aligned}$$

- 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

- Filter maximizing the SINR. Minimize the total resulting noise: find $\mathbf{G}$ such that $\mathcal{E}\left\{\left\|\mathbf{G}\mathbf{y} - \left[\begin{array}{ccc} c_1 & \ldots & c_{n_t} \end{array}\right]^T\right\|^2\right\}$ is minimum.
- The combined noise plus interference signal $\mathbf{n}_{i,q}$ when estimating symbol c_q writes as

$$\mathbf{n}_{i,q} = \sum_{p \neq q} \sqrt{\frac{E_s}{n_t}} \mathbf{h}_p c_p + \mathbf{n}.$$

  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

$$\mathbf{g}_{MMSE,q} = \sqrt{\frac{E_s}{n_t}} \mathbf{h}_q^H \left(\sigma_n^2 \mathbf{I}_{n_r} + \sum_{p \neq q} \frac{E_s}{n_t} \mathbf{h}_p \mathbf{h}_p^H\right)^{-1}.$$

- An alternative and popular representation of the MMSE filter can also be written as

$$\mathbf{G}_{MMSE} = \sqrt{\frac{n_t}{E_s}} \left(\mathbf{H}^H \mathbf{H} + \frac{n_t}{\rho} \mathbf{I}_{n_t}\right)^{-1} \mathbf{H}^H = \sqrt{\frac{n_t}{E_s}} \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \frac{n_t}{\rho} \mathbf{I}_{n_r}\right)^{-1}$$

- 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 $n_r - n_t + 1$.

Successive Interference Canceler

- Successively decode one symbol (or more generally one layer/stream) and cancel the effect of this symbol from the received signal.
- Decoding order based on the SINR of each symbol/layer: the symbol/layer with the highest SINR is decoded first at each iteration.
- 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.

- 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 $n_r - n_t + i$.
- 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.

Successive Interference Canceler

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 $\mathbf{H}$.
2. *Recursion:*
   1. *step 1:* extract the q_i^{th} transmitted symbol from the received signal $\mathbf{y}^{(i)}$

      $$\tilde{c}_{q_i} = \mathbf{G}^{(i)}(q_i,:)\,\mathbf{y}^{(i)}$$

      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

$$
\begin{aligned}
\bar{C}_{ZF-SIC} &= \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(1+\rho_q\right)\right\} \\
&\overset{(\rho\nearrow)}{\approx} \min\{n_t,n_r\}\log_2\left(\frac{\rho}{n_t}\right) + \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(\chi^2_{2(n_r-n_t+q)}\right)\right\} = \bar{C}_{CDIT}
\end{aligned}
$$

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

$$\bar{C}_{MMSE-SIC} = \sum_{q=1}^{\min\{n_t,n_r\}} \mathcal{E}\left\{\log_2\left(1+\rho_q\right)\right\} = \mathcal{E}\left\{\log_2 \det\left(\mathbf{I}_{n_r} + \frac{\rho}{n_t}\mathbf{H}\mathbf{H}^H\right)\right\} = \bar{\mathcal{I}}_e,$$

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.

Successive Interference Canceler

- In slow fading channels, what is the diversity-multiplexing trade-off achieved by unordered ZF-SIC?

Proposition

For $n_r \geq n_t$, the diversity-multiplexing trade-off achieved by Spatial Multiplexing with QAM constellation and unordered ZF-SIC receiver over i.i.d. Rayleigh fading channels is given by

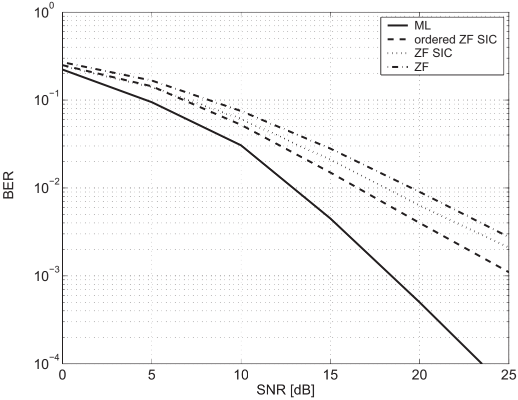
$$g_d\left(g_s, \infty\right) = \left(n_r - n_t + 1\right)\left(1 - \frac{g_s}{n_t}\right), \; g_s \in \left[0, n_t\right].$$

The achieved trade-off is similar to the trade-off achieved by a simple ZF receiver. This comes from the fact that the first layer dominates the error probability since its error exponent is the smallest.

- By increasing the number of receive antennas by 1,
  - with ZF or unordered ZF-SIC, we can *either* accommodate one extra stream with the same diversity order *or* increase the diversity order of every stream by 1,
  - with ML, we can accommodate one extra stream and *simultaneously* increase the diversity order of every stream by 1.

Impact of Decoding Strategy on Error Probability

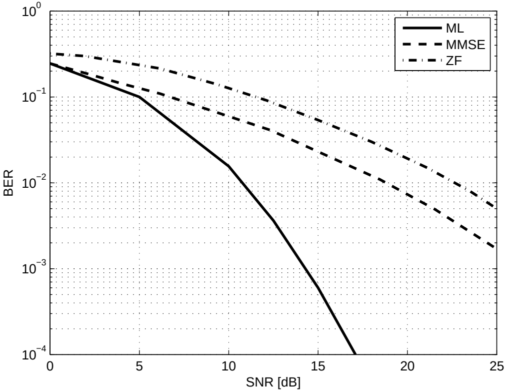
- SM with ML, ordered and non ordered ZF SIC and simple ZF decoding in 2×2 i.i.d. Rayleigh fading channels for 4 bits/s/Hz.



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

- 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
  1. the basis matrices are wide unitary

$$\mathbf{\Phi}_q \mathbf{\Phi}_q^H = \frac{T}{Q n_t} \mathbf{I}_{n_t} \ \forall q = 1 \ldots 2Q$$

  2. the basis matrices are pairwise skew-hermitian

$$\mathbf{\Phi}_q \mathbf{\Phi}_p^H + \mathbf{\Phi}_p \mathbf{\Phi}_q^H = 0, \ q \neq p$$

  or equivalently by this unique property

$$\mathbf{C}\mathbf{C}^H = \frac{T}{Q n_t} \left[\sum_{q=1}^{Q} |c_q|^2 \right] \mathbf{I}_{n_t}.$$

- 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 $r_s = 1$

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

- basis matrices are unitary and skew-hermitian (discussed before).
- $\mathbf{C}\mathbf{C}^H = \frac{1}{2}\left[\,|c_1|^2 + |c_2|^2\,\right]\mathbf{I}_2$.
- $r_s = 1$ 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 ($Q = 3$)

$$\mathbf{C} = \frac{2}{3} \begin{bmatrix} c_1 & -c_2^* & c_3^* & 0 \\ c_2 & c_1^* & 0 & c_3^* \\ c_3 & 0 & -c_1^* & -c_2^* \end{bmatrix}.$$

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

$$\text{choose } c_i \text{ iff } \left|z_1 - \sqrt{\frac{E_s}{2}} \left(|h_1|^2 + |h_2|^2\right) c_i\right|^2 \leq \left|z_1 - \sqrt{\frac{E_s}{2}} \left(|h_1|^2 + |h_2|^2\right) c_k\right|^2 \; \forall i \neq k$$

and similarly for c_2. Independent decoding of symbols c_1 and c_2 is so performed.

Orthogonal Space-Time Block Codes

- Error Probability

$$
\begin{aligned}
P(\mathbf{C} \to \mathbf{E}) &= \frac{1}{\pi} \int_0^{\pi/2} \left[\det\left(\mathbf{I}_{n_t} + \eta \tilde{\mathbf{E}}\right)\right]^{-n_r} d\beta \\
&= \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \eta \frac{T}{Q n_t} \sum_{q=1}^{Q} |c_q - e_q|^2\right)^{-n_r n_t} d\beta \\
&\overset{(\rho \nearrow)}{\leq} \left(\frac{\rho}{4} \frac{T}{Q n_t}\right)^{-n_r n_t} \left(\sum_{q=1}^{Q} |c_q - e_q|^2\right)^{-n_r n_t} .
\end{aligned}
$$

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

- Diversity-Multiplexing Trade-off Achieved by O-STBCs

Proposition

The diversity-multiplexing trade-off at high SNR achieved by O-STBCs using QAM constellations in i.i.d. Rayleigh fading channels is given by

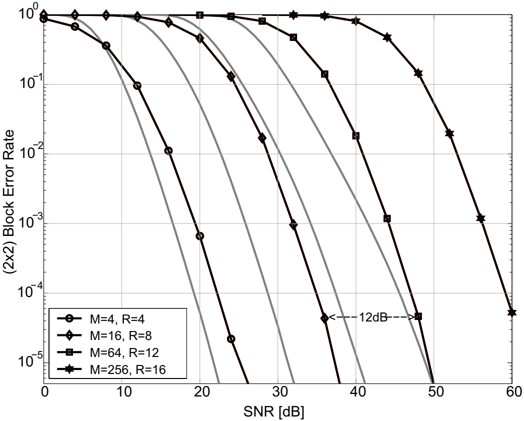
$$g_d\left(g_s,\infty\right)=n_r n_t\left(1-\frac{g_s}{r_s}\right),\ \ g_s\in\left[0,r_s\right].$$

Proposition

The Alamouti code with any QAM constellation achieves the optimal diversity-multiplexing trade-off for two transmit and one receive antennas in i.i.d. Rayleigh fading channels.

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.



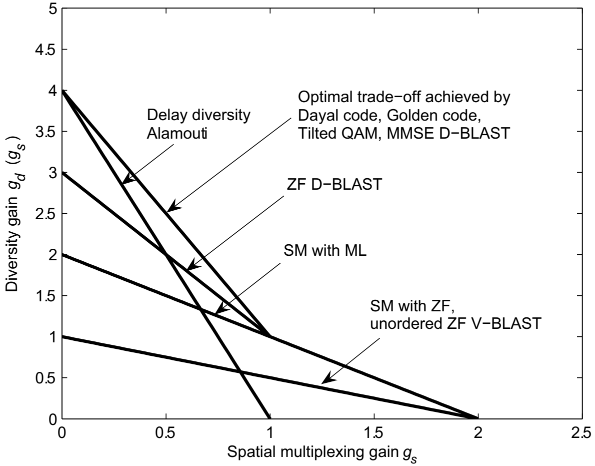
  - full diversity exploited: $g_d\,(g_s = 0, \infty) = g_d^o\,(\infty) = 4$.
  - the growth of the multiplexing gain is slow: 12 dB separate the curves, corresponding to a multiplexing gain $g_s = 1$, i.e., 1 bit/s/Hz increase per 3 dB SNR increase.

Other Code Constructions

- Quasi-Orthogonal Space-Time Block Codes
  - increase the spatial multiplexing rate while still partially enjoying the decoupling properties of O-STBCs
  - use O-STBCs of reduced dimensions as the building blocks of a higher dimensional code
- Linear Dispersion Codes
  - if a larger receiver complexity is authorized, it is possible to relax the skew-hermitian conditions and increase the data rates while still providing transmit diversity.
- Algebraic Space-Time Codes
  - structured codes using algebraic tools
  - many of them are designed to achieve the optimal diversity-multiplexing tradeoff

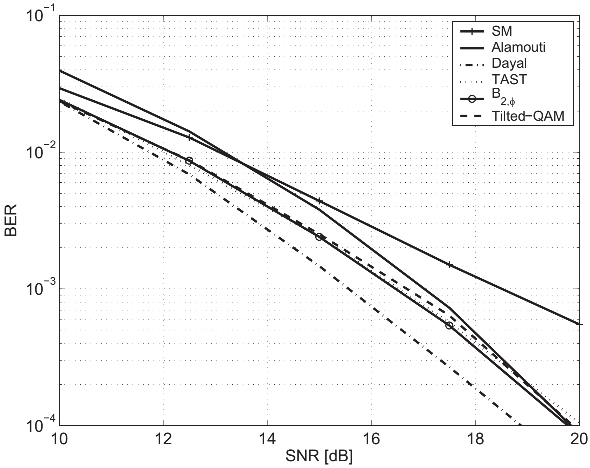
Global Performance Comparison

- Asymptotic diversity-multiplexing trade-off $g_d\,(g_s, \infty)$ achieved by several space-time codes in a 2×2 i.i.d. Rayleigh fading MIMO channel.



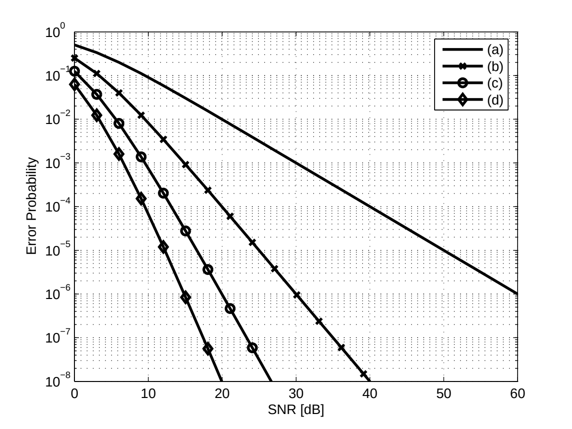
Global Performance Comparison

- 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.



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.

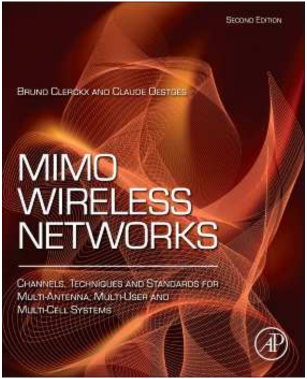
Answer: The simplest approach is to perform
for (a), receive matched filter with $n_r = 1$, $n_t = 1$
for (b), receive matched filter with $n_r = 2$, $n_t = 1$
for (c), receive matched filter with $n_r = 3$, $n_t = 1$
for (d), receive matched filter with $n_r = 4$, $n_t = 1$

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.

MIMO with Partial Channel State Information at the Transmitter

Reference Book

- 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.



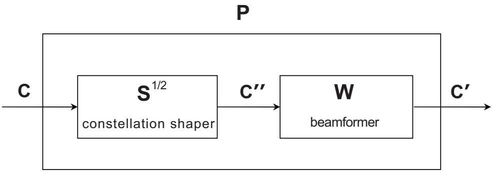
  - Chapter 10
    - Section: 10.1, 10.2.1, 10.5, 10.6.1, 10.6.2, 10.6.3, 10.6.4, 10.9

Introduction

- 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 Channel Statistics at the Transmitter
  - low rate feedback link
  - statistical properties of the channel (correlations, K-factor) vary at a much slower rate than the fading channel itself
  - The receiver estimates the channel stochastic properties and sends them back to the transmitter "once in a while" (if channel reciprocity cannot be exploited)
  - stationary channel: statistics do not change over time
- 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.

System Model

- MIMO system with n_t transmit and n_r receive antennas communicating through a frequency flat slow fading channel.
- The encoder outputs a codeword $\mathbf{C} = [\mathbf{c}_0 \ldots \mathbf{c}_{T-1}]$ of size $n_e \times T$ contained in the codebook $\mathcal{C}$ over T symbol durations.
- Precoder $\mathbf{P}$ $[n_t \times n_e]$ processes the codeword $\mathbf{C}$ and the codeword $\mathbf{C}' = \mathbf{PC}$ $[n_t \times T]$ is transmitted over n_t antennas.



- Linear precoder $\mathbf{P} = \mathbf{W}\mathbf{S}^{1/2}$
  - multi-mode beamformer $\mathbf{W}$ whose columns have a unit-norm
  - constellation shaper $\mathbf{S}^{1/2}$ (if $\mathbf{S}$ real-valued and diagonal, it can be thought of as the power allocation scheme across the modes)
- nomalization: $\mathcal{E}\left\{\text{Tr}\left\{\mathbf{C}'\mathbf{C}'^H\right\}\right\} = T$, $\mathcal{E}\left\{\text{Tr}\left\{\mathbf{C}\mathbf{C}^H\right\}\right\} = T$ and $\text{Tr}\left\{\mathbf{P}\mathbf{P}^H\right\} = n_e$, $\mathcal{E}\left\{\left\|\mathbf{H}\right\|_F^2\right\} = n_t n_r$.

Channel Statistics based Precoding

- Information Theory motivated strategy.

Proposition

In Kronecker correlated Rayleigh fast fading channels, the optimal input covariance matrix can again be expressed as

$$\mathbf{Q} = \mathbf{U}_{\mathbf{R}_t} \mathbf{\Lambda}_{\mathbf{Q}} \mathbf{U}_{\mathbf{R}_t}^H,$$

where $\mathbf{U}_{\mathbf{R}_t}$ is a unitary matrix formed by the eigenvectors of $\mathbf{R}_t$ (arranged in such order that they correspond to decreasing eigenvalues of $\mathbf{R}_t$), and $\mathbf{\Lambda}_{\mathbf{Q}}$ is a diagonal matrix whose elements are also arranged in decreasing order.

Transmit a single stream along the dominant eigenvector of $\mathbf{R}_t$ if very large transmit correlation. Transmit multiple streams with uniform power allocation if very low transmit correltion.

- Error Probability motivated strategy

$$\mathbf{P}^{\star} = \arg\min_{\mathbf{P}} \max_{\tilde{\mathbf{E}} \neq \mathbf{0}} P\left(\mathbf{C} \rightarrow \mathbf{E}\right)$$

  - challenging problem for arbitrary codes
  - focus on O-STBC

Channel Statistics based Precoding

- O-STBCs in Kronecker Rayleigh fading channels

Proposition

In Kronecker Rayleigh fading channels, the optimal precoder minimizing the average PEP/SER is given by $\mathbf{P} = \mathbf{W}\mathbf{S}^{1/2}$ *where*

- $\mathbf{W} = \mathbf{U}'_{\mathbf{R}_t}$ *with* $\mathbf{U}'_{\mathbf{R}_t}$ *the* $n_t \times n_e$ *submatrix of* $\mathbf{U}_{\mathbf{R}_t}$ *containing the* n_e *dominant eigenvector of* $\mathbf{R}_t$*, i.e.,* $\mathbf{R}_t = \mathbf{U}_{\mathbf{R}_t}\mathbf{\Lambda}_{\mathbf{R}_t}\mathbf{U}^H_{\mathbf{R}_t}$*,*
- $\mathbf{S}^{1/2} = \mathbf{D}$*,* $\mathbf{D}$ *being a real-valued diagonal matrix accounting for the power allocation.*

- Power allocation strategy follows the water-filling solution.

Channel Statistics based Precoding

Example

Let us consider the Alamouti O-STBC with two transmit antennas ($n_e = n_t = 2$). Denoting $\mathbf{S} = \text{diag}\{s_1, s_2\}$, the transmitted codewords are proportional, at the first time instant, to

$$\frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}\mathbf{S}^{1/2}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}(:,1)\sqrt{s_1}c_1 + \frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}(:,2)\sqrt{s_2}c_2$$

and, at the second time instant, to

$$\frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}\mathbf{S}^{1/2}\begin{bmatrix} -c_2^* \\ c_1^* \end{bmatrix} = -\frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}(:,1)\sqrt{s_1}c_2^* + \frac{1}{\sqrt{2}}\mathbf{U}_{\mathbf{R}_t}(:,2)\sqrt{s_2}c_1^*.$$

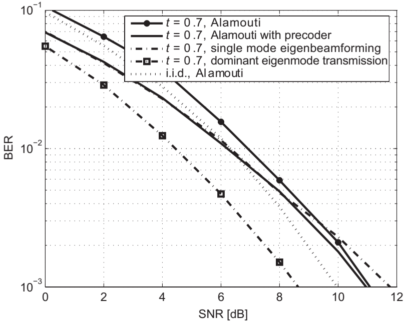
Extreme cases:

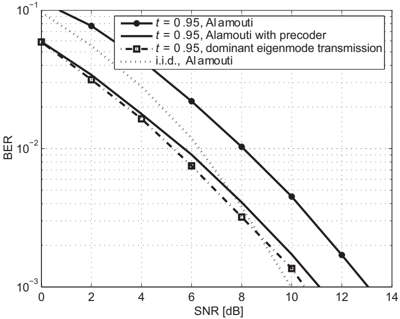
- $s_1 = s_2 = 1$: Alamouti scheme
- $s_1 = 2$, $s_2 = 0$: beamforming in the dominant eigenbeam

The precoder allocates more power to angular directions corresponding to the peaks of the transmit direction power spectrum.

Channel Statistics based Precoding

- Performance of a transmit correlation based precoded Alamouti scheme in 2×2 transmit correlated ($t = 0.7$ and $t = 0.95$) Rayleigh channels





Quantized Precoding: dominant eigenmode transmission

- Assume dominant eigenmode transmission (i.e. beamforming)

$$\begin{aligned} \mathbf{y} &= \sqrt{E_s}\mathbf{H}\mathbf{w}c + \mathbf{n}, \\ z &= \mathbf{g}^H\mathbf{y}, \\ &= \sqrt{E_s}\mathbf{g}^H\mathbf{H}\mathbf{w}c + \mathbf{g}^H\mathbf{n} \end{aligned}$$

  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

$$\mathbf{w}^\star = \arg \max_{\substack{1\leq i\leq n_p \\ \mathbf{w}_i\in\mathcal{W}}} \|\mathbf{H}\mathbf{w}_i\|^2 .$$

Quantized Precoding: distortion function

- How to design the codebook? Need for a distortion function, i.e. measure of the average (over all channel realizations) array gain loss induced by the quantization process

$$d_f = \mathcal{E}_{\mathbf{H}} \left\{ \lambda_{max} - \left\| \mathbf{H}\mathbf{w}^{\star} \right\|^2 \right\}$$

- Upper-bound

$$\begin{aligned} d_f &\leq \mathcal{E}_{\mathbf{H}} \left\{ \lambda_{max} - \lambda_{max} \left| \mathbf{v}_{max}^H \mathbf{w}^{\star} \right|^2 \right\}, \\ &\overset{(a)}{=} \underbrace{\mathcal{E}_{\mathbf{H}} \left\{ \lambda_{max} \right\}}_{\text{quality of the channel}} \underbrace{\mathcal{E}_{\mathbf{H}} \left\{ 1 - \left| \mathbf{v}_{max}^H \mathbf{w}^{\star} \right|^2 \right\}}_{\text{quality of the codebook}} \end{aligned}$$

where $\mathbf{v}_{max}$ is the dominant right singular vector of $\mathbf{H}$. Equality (a) is only valid for i.i.d. Rayleigh fading channels.

Quantized Precoding: Lloyd

- Generalized Lloyd Algorithm:

Algorithm

For the given codebook, find the optimal quantization cells using the nearest neighbor rule. For the so-obtained quantization cells, determine that optimal quantized precoders using the centroid condition. Iterate till convergence.

- Essential conditions:
  - Assume MISO channel.
  - *centroid condition*: the optimal quantized precoder $\mathbf{w}_k$ of any quantization cell $\mathbf{R}_k$ is to be chosen as the dominant eigenvector of $\mathbf{R}_k = \mathcal{E}\left\{\mathbf{h}^H\mathbf{h} \mid \mathbf{h} \in \mathcal{R}_k\right\}$.
  - *nearest neighbor rule*: all channel vectors that are closer to the quantized precoder $\mathbf{w}_k$ are assigned to quantization cell $\mathcal{R}_k$, i.e. $\mathbf{h} \in \mathcal{R}_k$ if $\|\mathbf{h}\|^2 - |\mathbf{h}\mathbf{w}_k|^2 \leq \|\mathbf{h}\|^2 - |\mathbf{h}\mathbf{w}_j|^2$,
- Optimal codebook design for arbitrary fading channels

Quantized Precoding: Grassmannian

- Grassmannian Line Packing

Design Criterion

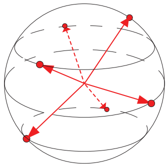
Choose a codebook $\mathcal{W}$ made of n_p unit-norm vectors $\mathbf{w}_i$ $(i = 1, \ldots, n_p)$ such that the minimum distance

$$\delta_{line}(\mathcal{W}) = \min_{1 \leq k < l \leq n_p} \sqrt{1 - \left|\mathbf{w}_k^H \mathbf{w}_l\right|^2},$$

is maximized.

- Problem of packing n_p lines in $\mathbb{C}^{n_t}$ in such a way that the minimum distance between any pair of lines is maximized.
- Close to optimal only for i.i.d. Rayleigh Fading Channels.

Grassmannian codebook



Quantized Precoding: How many bits?

- How many feedback bits $B = \log_2(n_p)$ are required? In i.i.d. channels

$$\bar{C}_{quant} \approx \mathcal{E}_{\mathbf{h}} \left\{ \log_2 \left(1 + \rho \left\| \mathbf{h} \right\|^2 \left(1 - 2^{-\frac{B}{n_t - 1}} \right) \right) \right\},$$

leading to an SNR degradation of $10 \log_{10} \left(1 - 2^{-\frac{B}{n_t - 1}} \right)$ dB relative to perfect CSIT.

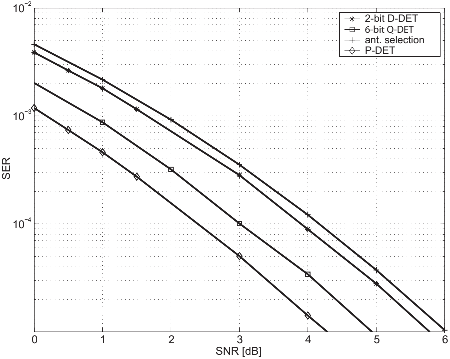
Proposition

In order to maintain a constant SNR or capacity gap between perfect CSIT and quantized feedback, it is not necessary to scale the number of feedback bits as a function of the SNR. The multiplexing gain g_s is not affected by the quality of CSIT.

- Achievable diversity gain?
  - Antenna selection (AS) is a particular case of a quantized precoding whose codebook is chosen as the columns of the identity matrix $\mathbf{I}_{n_t}$.
  - AS achieves a diversity gain of n_t.
  - Sufficient to take a full rank codebook matrix with $n_p \geq n_t$ to extract the full diversity

Quantized Precoding: Evaluations

- SER of a 3×3 MIMO system using 2-bit and 6-bit quantized BPSK-based DET.



Example

Codebook for quantized DET for $n_t = 3$ and $n_p = 4$.

$$\mathcal{W} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{j}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \\ \frac{-j}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{-j}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \\ \frac{j}{\sqrt{3}} \end{bmatrix} \right\}$$

Example

Question: Discuss the validity of the following statement and detail your argument: 'Consider a point-to-point i.i.d. MISO Rayleigh slow fading channel with 4 transmit antennas and 1 receive antenna and a transmission strategy based on partial transmit channel knowledge consisting of transmitting a single stream using quantized precoding. The codebook of precoders is given by

$$\mathcal{W} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

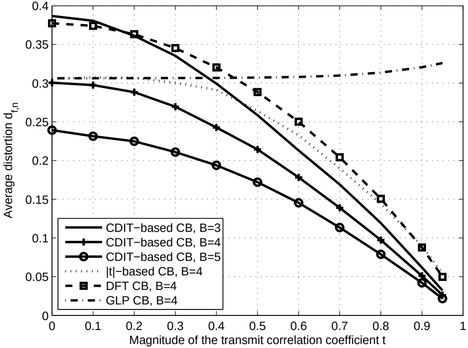
This transmission strategy achieves a diversity gain of 4.'

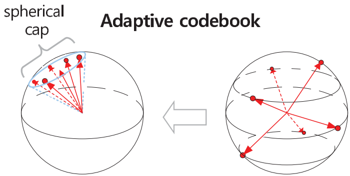
Example

Answers: The codebook only contains 3 precoders. Looking at the precoder, no stream can be transmitted on the fourth antenna. The fourth antenna is therefore useless. This implies that the system effectively looks like a MISO with 3 transmit antenna where transmit antenna selection is performed. With n_r antennas, receive antenna selection achieves a diversity gain of n_r (see lecture notes). Similarly, transmit antenna selection achieves a diversity equal to 3 in this scenario. Hence the proposed strategy will achieve a diversity gain of 3. Hence the statement is wrong.

Quantized Precoding: spatially correlated channels

- Grassmiannian only appropriate for i.i.d. Rayleigh fading.
- Spatial correlation decreases the quantization space compared to i.i.d. channels.
  - e.g. Lloyd, adaptive/CDIT-based codebook, DFT (for uniform linear arrays)
- Normalized average distortion (SNR loss) $d_{f,n} = d_f / \mathcal{E}_{\mathbf{H}} \{\lambda_{max}\}$ as a function of the codebook size $n_p = 2^B$ and the transmit correlation coefficient t with $n_t = 4$.





$$\mathcal{W}_c = \left\{ \frac{\mathbf{R}_t^{1/2}\mathbf{w}_1}{\left\|\mathbf{R}_t^{1/2}\mathbf{w}_1\right\|}, \ldots, \frac{\mathbf{R}_t^{1/2}\mathbf{w}_{n_p}}{\left\|\mathbf{R}_t^{1/2}\mathbf{w}_{n_p}\right\|} \right\}$$

Quantized Precoding: some extensions

- 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 $\mathbf{H}\mathbf{W}_i^{(n_e)}$.

- 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

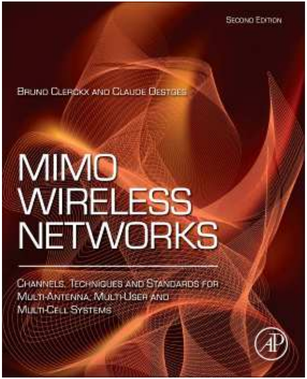
$$R = \sum_{q=1}^{n_e} \log_2 \left(1 + \rho_q \left(\mathbf{H}\mathbf{W}_i^{(n_e)} \right) \right).$$

where ρ_q is the SINR of stream q on at the output of the combiner for the equivalent channel $\mathbf{H}\mathbf{W}_i^{(n_e)}$.

Multi-User MIMO - Multiple Access Channels (Uplink) & Broadcast Channels (Downlink)

Reference Book

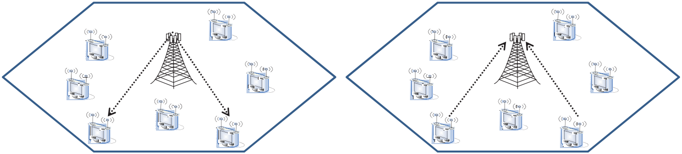
- Bruno Clerckx and Claude Oestges, "MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems," Academic Press (Elsevier), Oxford, UK, Jan 2013.



  - Chapter 12
    - Section: 12.1, 12.2, 12.3, 12.4

Introduction

- So far, we looked at a single link/user. Most systems are multi-user!
- How to deal with multiple users? What is the benefit of MIMO in a multi-user setting?
- MIMO 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.

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

  - $\mathbf{y}_{ul} \in \mathbb{C}^{n_r}$
  - $\mathbf{H}_{ul,q} \in \mathbb{C}^{n_r \times n_{t,q}}$ 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,

$$\mathbf{c}'_{ul} = \left[\mathbf{c}'^{T}_{ul,1}, \ldots, \mathbf{c}'^{T}_{ul,K}\right]^{T},$$

$$\mathbf{H}_{ul} = \left[\Lambda_1^{-1/2}\mathbf{H}_{ul,1}, \ldots, \Lambda_K^{-1/2}\mathbf{H}_{ul,K}\right],$$

  the system model also writes as

$$\mathbf{y}_{ul} = \mathbf{H}_{ul}\mathbf{c}'_{ul} + \mathbf{n}_{ul}.$$

  $\mathbf{H}_{ul}$ 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 Deterministic Channels

- 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.

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

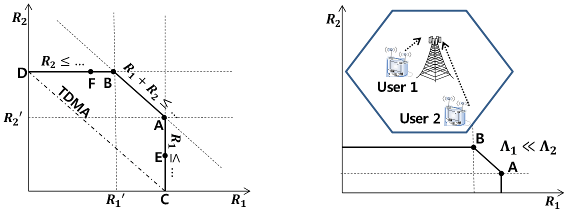
$$C = \max_{(R_1,\ldots,R_K)\in\mathcal{C}} \sum_{q=1}^{K} R_q.$$

Rate Region of MIMO MAC

- For given input covariance matrices $\mathbf{Q}_{ul,1}, \ldots, \mathbf{Q}_{ul,K}$, the achievable rate region is defined by
  1. The rate achievable by a given user q *with a given transmit strategy* $\mathbf{Q}_{ul,q}$ cannot be larger than its achievable rate in a single-user setup, i.e.
  $$R_q \leq \log_2 \det \left[\mathbf{I}_{n_r} + \frac{\Lambda_q^{-1}}{\sigma_n^2} \mathbf{H}_{ul,q} \mathbf{Q}_{ul,q} \mathbf{H}_{ul,q}^H \right], \quad q = 1, \ldots, K$$
  where $\mathbf{Q}_{ul,q} = \mathcal{E}\{\mathbf{c}'_q \mathbf{c}'^H_q\}$ is subject to the power constraint $\mathrm{Tr}\{\mathbf{Q}_{ul,q}\} \leq E_{s,q}$.
  2. The sum of the rates achievable by a subset S of the users should be smaller than the total rate achievable when those users "cooperate" with each other to form a giant array with $n_{t,S} = \sum_{q \in S} n_{t,q}$ transmit antennas subject to their respective power constraints, i.e.
  $$\begin{aligned} \sum_{q \in S} R_q &\leq \log_2 \det \left[\mathbf{I}_{n_r} + \frac{1}{\sigma_n^2} \mathbf{H}_{ul,S} \mathbf{Q}_{ul,S} \mathbf{H}_{ul,S}^H \right] \\ &= \log_2 \det \left[\mathbf{I}_{n_r} + \frac{1}{\sigma_n^2} \sum_{q \in S} \Lambda_q^{-1} \mathbf{H}_{ul,q} \mathbf{Q}_{ul,q} \mathbf{H}_{ul,q}^H \right], \end{aligned}$$
  with $\mathbf{H}_{ul,S} = \left[\Lambda_i^{-1/2} \mathbf{H}_{ul,i}, \ldots, \Lambda_j^{-1/2} \mathbf{H}_{ul,j} \right]_{i,j \in S}$,
  $\mathbf{Q}_{ul,S} = \mathrm{diag}\left\{ \mathbf{Q}_{ul,i}, \ldots, \mathbf{Q}_{ul,j} \right\}_{i,j \in S}$, subject to the constraints $\mathrm{Tr}\{\mathbf{Q}_{ul,q}\} \leq E_{s,q}$.
- The rate region looks like a K-dimensional polyhedron with $K!$ corner points on the boundary.

Rate Region of a Two-User MIMO MAC

- This rate region is a pentagon with two corner points A and B.



(a) Two-user MIMO MAC rate region for fixed $Q_{ul,1}$ and $Q_{ul,2}$

(b) Rate regions with various path losses

- Remarkably, at point A, user 1 can transmit at a rate equal to its single-link MIMO rate and user 2 can simultaneously transmit at a rate $R_2' > 0$ equal to

$$R_2' = \log_2 \det \left[\mathbf{I}_{n_r} + \frac{\Lambda_1^{-1}}{\sigma_n^2} \mathbf{H}_{ul,1} \mathbf{Q}_{ul,1} \mathbf{H}_{ul,1}^H + \frac{\Lambda_2^{-1}}{\sigma_n^2} \mathbf{H}_{ul,2} \mathbf{Q}_{ul,2} \mathbf{H}_{ul,2}^H \right]$$

$$- \log_2 \det \left[\mathbf{I}_{n_r} + \frac{\Lambda_1^{-1}}{\sigma_n^2} \mathbf{H}_{ul,1} \mathbf{Q}_{ul,1} \mathbf{H}_{ul,1}^H \right]$$

$$= \log_2 \det \left[\mathbf{I}_{n_r} + \frac{\Lambda_2^{-1}}{\sigma_n^2} \mathbf{H}_{ul,2} \mathbf{Q}_{ul,2} \mathbf{H}_{ul,2}^H \left(\mathbf{I}_{n_r} + \frac{\Lambda_1^{-1}}{\sigma_n^2} \mathbf{H}_{ul,1} \mathbf{Q}_{ul,1} \mathbf{H}_{ul,1}^H \right)^{-1} \right].$$

Capacity Region of MIMO MAC

- We have assumed so far specific input covariance matrices.
  - A different choice of the beamforming matrix and the power allocation leads to a different transmit strategy $\mathbf{Q}_{ul,q}$ and generally a different shape of the pentagon (or more generally the K-dimensional polyhedron).
  - The trade-off between user rates is therefore affected by the choice of the input covariance matrices.
  - The optimal set of input covariance matrices that maximizes the sum-rate can be found using a generalization of the single-link water-filling solution (Detail in the book).
- The capacity region is equal to the union (over all transmit strategies satisfying the power constraints) of all the K-dimensional polyhedrons.

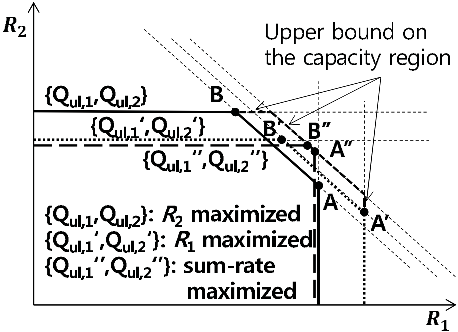
Proposition

The capacity region $\mathcal{C}_{MAC}$ of the Gaussian MIMO MAC for a determinsitic channel $\mathbf{H}_{ul}$ is the union of all achievable rate vectors $(R_1, \ldots, R_K)$ given by

$$\bigcup_{\substack{\mathrm{Tr}\left\{\mathbf{Q}_{ul,q}\right\} \leq E_{s,q} \\ \mathbf{Q}_{ul,q} \geq 0, \forall q}} \left\{ \begin{array}{l} (R_1, \ldots, R_K) : \sum_{q \in S} R_q \leq \\ \log_2 \det \left[\mathbf{I}_{n_r} + \sum_{q \in S} \frac{\Lambda_q^{-1}}{\sigma_n^2} \mathbf{H}_{ul,q} \mathbf{Q}_{ul,q} \mathbf{H}_{ul,q}^H \right], \forall S \subseteq \mathcal{K} \end{array} \right\}.$$

Capacity Region of a Two-User MIMO MAC

- Due to the union of pentagons, the capacity region of the two-user MIMO MAC does not look like a pentagon in general.



(c) Bounds on the capacity region

- However, with a single antenna $(n_{t,q} = 1)$, the capacity region remains a pentagon because a single data is transmitted per user at the full power, i.e. $E_{s,q}$.

Capacity Region of SISO MAC

Corollary

$$\mathcal{C}_{MAC} = \left\{ (R_1, \ldots, R_K) : \sum_{q \in S} R_q \leq \log_2 \left(1 + \sum_{q \in S} \eta_q \left|h_{ul,q}\right|^2 \right), \forall S \subseteq \mathcal{K} \right\}$$

where $\eta_q = \Lambda_q^{-1} E_{s,q} / \sigma_n^2$.

Example

Two-user SISO: $\mathcal{C}_{MAC}$ is the set of all rates pair (R_1, R_2) satisfying to

$$R_q \leq \log_2 \left(1 + \eta_q \left|h_{ul,q}\right|^2\right), q = 1, 2$$
$$R_1 + R_2 \leq \log_2 \left(1 + \eta_1 \left|h_{ul,1}\right|^2 + \eta_2 \left|h_{ul,2}\right|^2\right).$$

$$R_2' = \log_2 \left(1 + \eta_1 \left|h_{ul,1}\right|^2 + \eta_2 \left|h_{ul,2}\right|^2\right) - \log_2 \left(1 + \eta_1 \left|h_{ul,1}\right|^2\right)$$
$$= \log_2 \left(1 + \frac{\eta_2 \left|h_{ul,2}\right|^2}{1 + \eta_1 \left|h_{ul,1}\right|^2}\right) = \log_2 \left(1 + \frac{\Lambda_2^{-1} \left|h_{ul,2}\right|^2 E_{s,2}}{\sigma_n^2 + \Lambda_1^{-1} \left|h_{ul,1}\right|^2 E_{s,1}}\right).$$

Capacity Region of SIMO MAC

Corollary

$$\mathcal{C}_{MAC} = \left\{ \begin{array}{l} (R_1, \ldots, R_K) : \sum_{q \in S} R_q \leq \\ \log_2 \det \left[\mathbf{I}_{n_r} + \sum_{q \in S} \eta_q \mathbf{h}_{ul,q} \mathbf{h}_{ul,q}^H \right], \forall S \subseteq \mathcal{K} \end{array} \right\}$$

where $\eta_q = \Lambda_q^{-1} E_{s,q} / \sigma_n^2$.

Example

Two-user SIMO: $\mathcal{C}_{MAC}$ is the set of all rates pair (R_1, R_2) satisfying to

$$R_q \leq \log_2 \left(1 + \eta_q \left\| \mathbf{h}_{ul,q} \right\|^2 \right) = \log_2 \det \left(\mathbf{I}_{n_r} + \eta_q \mathbf{h}_{ul,q} \mathbf{h}_{ul,q}^H \right), q = 1, 2$$

$$R_1 + R_2 \leq \log_2 \det \left(\mathbf{I}_{n_r} + \eta_1 \mathbf{h}_{ul,1} \mathbf{h}_{ul,1}^H + \eta_2 \mathbf{h}_{ul,2} \mathbf{h}_{ul,2}^H \right).$$

$$\begin{aligned} R_2' &= \log_2 \det \left(\mathbf{I}_{n_r} + \eta_1 \mathbf{h}_{ul,1} \mathbf{h}_{ul,1}^H + \eta_2 \mathbf{h}_{ul,2} \mathbf{h}_{ul,2}^H \right) - \log_2 \det \left(\mathbf{I}_{n_r} + \eta_1 \mathbf{h}_{ul,1} \mathbf{h}_{ul,1}^H \right) \\ &= \log_2 \det \left(\mathbf{I}_{n_r} + \eta_2 \mathbf{h}_{ul,2} \mathbf{h}_{ul,2}^H \left(\mathbf{I}_{n_r} + \eta_1 \mathbf{h}_{ul,1} \mathbf{h}_{ul,1}^H \right)^{-1} \right) \\ &= \log_2 \left(1 + \eta_2 \mathbf{h}_{ul,2}^H \left(\mathbf{I}_{n_r} + \eta_1 \mathbf{h}_{ul,1} \mathbf{h}_{ul,1}^H \right)^{-1} \mathbf{h}_{ul,2} \right) \end{aligned}$$

Achievability of the Capacity Region

- For $n_t = 1$, the SIMO MAC architecture is reminiscent of the Spatial Multiplexing architecture discussed for a single-link MIMO channel.
- We can therefore fully reuse the various receiver architectures derived for single-link MIMO.
- Recall the optimality of the MMSE V-BLAST (also called Spatial Multiplexing with MMSE-SIC receiver)

Proposition

MMSE-SIC is optimal for achieving the corner points of the MIMO MAC rate region.

- The exact corner point that is achieved on the rate 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 MMSE receiver for user 2's stream treating user 1's stream as colored Gaussian interference.

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).
- SISO: both TDMA and SIC exploit a single degree of freedom but TDMA rate region is strictly smaller than the one achievable with SIC.
- SIMO: TDMA incurs a big loss compared to SIMO MAC (with MMSE-SIC) as it only exploits a single degree of freedom despite the presence of $\min\{n_r, K\}$ degrees of freedom achievable with SIMO MAC at high SNR.
- MIMO: As n_t increases, the gap between the TDMA and MIMO MAC rate regions decreases.

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
  - $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$
  - $\mathbf{H}_q \in \mathbb{C}^{n_{r,q} \times n_t}$ 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 $\mathcal{CN}\left(0, \sigma_{n,q}^2 \mathbf{I}_{n_{r,q}}\right)$.
- 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: $\mathrm{Tr}\{\mathbf{Q}\} \leq E_s$.

MIMO BC System Model

- By stacking up the received signal vectors, the noise vectors and the channel matrices of all K users,

$$\mathbf{y} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_K^T\right]^T,$$

$$\mathbf{n} = \left[\mathbf{n}_1^T, \ldots, \mathbf{n}_K^T\right]^T,$$

$$\mathbf{H} = \left[\Lambda_1^{-1/2}\mathbf{H}_1^T, \ldots, \Lambda_K^{-1/2}\mathbf{H}_K^T\right]^T,$$

  the system model also writes as

$$\mathbf{y} = \mathbf{H}\mathbf{c}' + \mathbf{n}.$$

  $\mathbf{H}$ 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 $\mathbf{Q}_q = \mathcal{E}\left\{\mathbf{c}'_q \mathbf{c}'^H_q\right\}$.

Capacity Region of two-user SISO Deterministic 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.
- 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 $s_1 + s_2 = E_s$)

$$R_1 = \log_2\left(1 + \frac{\Lambda_1^{-1} s_1}{\sigma_{n,1}^2} |h_1|^2\right)$$

$$R_2 = \log_2\left(1 + \frac{\Lambda_2^{-1} |h_2|^2 s_2}{\sigma_{n,2}^2 + \Lambda_2^{-1} |h_2|^2 s_1}\right).$$

- *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.

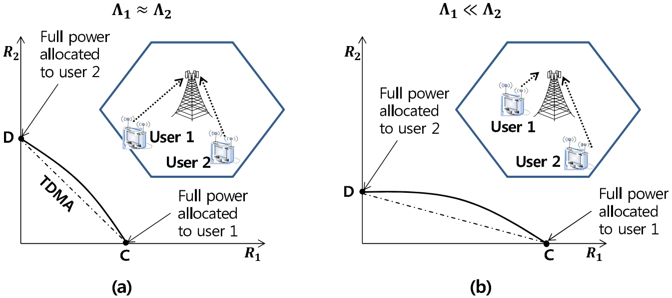
$$R_2 \leq \log_2\left(1 + \frac{\Lambda_1^{-1} |h_1|^2 s_2}{\sigma_{n,1}^2 + \Lambda_1^{-1} |h_1|^2 s_1}\right).$$

- The channel gains normalized w.r.t. their respective noise power should be ordered

$$\frac{\Lambda_2^{-1} |h_2|^2}{\sigma_{n,2}^2} \leq \frac{\Lambda_1^{-1} |h_1|^2}{\sigma_{n,1}^2}.$$

Capacity Region of two-user SISO Deterministic BC

- 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 $s_1 + s_2 = E_s$.



Capacity Region of K-user SISO Deterministic 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

$$\bigcup_{s_q : \sum_{q=1}^{K} s_q = E_s} \left\{ (R_1, \ldots, R_K) : R_q \leq \log_2 \left(1 + \frac{|\mathsf{h}_q|^2 s_q}{1 + |\mathsf{h}_q|^2 \left[\sum_{p=1}^{q-1} s_p \right]} \right), \forall q \right\}.$$

Proposition

The sum-rate capacity of the SISO BC is achieved by allocating the transmit power to the strongest user

$$C_{BC} = \log_2 \left(1 + E_s \max_{q=1,\ldots,K} |\mathsf{h}_q|^2 \right) = \log_2 \left(1 + \max_{q=1,\ldots,K} \eta_q |h_q|^2 \right).$$

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

- Receiver cancellation - *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.
- Transmitter cancellation - *Dirty-Paper Coding* (DPC)
  - Assume a system model $y = hc' + i + n$ with i, n Gaussian interference and noise. Simply subtracting i for transmit signal is not a good idea!

Proposition

If Tx has full (non-causal) knowledge of the interference, the capacity of the dirty paper channel is equal to the capacity of the channel with the interference completely absent.

  - By encoding users in the increasing order of their normalized channel gains, DPC achieves the capacity region of the SISO BC.

Example

Assume $|\mathrm{h}_1|^2 \geq |\mathrm{h}_2|^2$. By treating user 2's signal c_2' as known Gaussian interference at Tx and encoding user 1's signal c_1' using DPC, user 1 can achieve a rate as high as if user 2's signal was absent. User 2 treats user 1's signal as Gaussian noise.

Achievability of the SISO BC Capacity Region

Proposition

With the appropriate cancellation/encoding ordering, Superposition Coding with SIC and Dirty-Paper Coding are both optimal for achieving the SISO BC capacity region.

Proposition

The SISO BC sum-rate capacity is achievable with dynamic TDMA (to the strongest user), Superposition Coding with SIC (with the appropriate cancellation ordering) and Dirty-Paper Coding (with the appropriate encoding ordering).

Capacity Region of MIMO BC and its Achievability

- MAC with multiple Rx antennas provides a tremendous capacity increase compared to suboptimal TDMA. So does BC with multiple Tx antennas!
- MIMO BC difficult problem: users' channels cannot be ranked anymore.
- Assume an increasing encoding order from user 1 to K:
  1. Encode user 1's signal into $\mathbf{c}_1'$.
  2. With full knowledge of $\mathbf{c}_1'$, encode user 2's signal into $\mathbf{c}_2'$ using DPC: $\mathbf{c}_1'$ appears invisible to user 2 but $\mathbf{c}_2'$ appears like a Gaussian interference to user 1.
  3. With full knowledge of user 1 and user 2's signals, encode user 3's signal into $\mathbf{c}_3'$ using DPC.
  4. ... till K users are encoded.
- A given user q sees signals from users $p > q$ as a Gaussian interference but does not see any interference signals from users $p < q$:
  - Covariance of Noise plus Interference at user q: $\sigma_{n,q}^2 \mathbf{I}_{n_{r,q}} + \Lambda_q^{-1} \mathbf{H}_q \left[\sum_{p>q} \mathbf{Q}_p\right] \mathbf{H}_q^H$.
  - With a MMSE receiver that whitens the colored Gaussian interference (same as in MAC)

$$R_q = \log_2 \det \left[\mathbf{I}_{n_{r,q}} + \Lambda_q^{-1} \mathbf{H}_q \mathbf{Q}_q \mathbf{H}_q^H \left(\sigma_{n,q}^2 \mathbf{I}_{n_{r,q}} + \Lambda_q^{-1} \mathbf{H}_q \left[\sum_{p>q} \mathbf{Q}_p \right] \mathbf{H}_q^H \right)^{-1} \right]$$

- Capacity region: Repeat for all covariance matrices $\mathbf{Q}_1, \ldots, \mathbf{Q}_K$ satisfying the sum-power constraint $\sum_q \text{Tr}\{\mathbf{Q}_q\} = E_s$ and all user ordering.
- Only DPC can achieve the MISO/MIMO BC sum-rate capacity.

Comparisons with TDMA

- SISO
  - 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

Proposition

For channels $\mathbf{H}_1, \ldots, \mathbf{H}_K$, SNR η_q, number of receive antennas n_r, the gain of DPC over TDMA is upper-bounded by the minimum between the number of transmit antennas n_t and the number of users K

$$\frac{C_{BC}(\mathbf{H})}{C_{TDMA}(\mathbf{H})} \leq \min\{n_t, K\}.$$

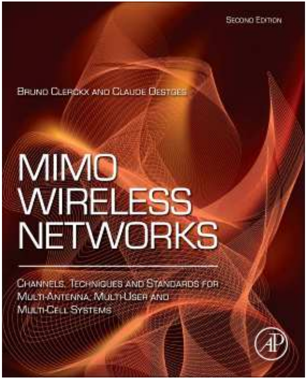
Intuition:

- TDMA exploits at least one spatial dimension with the largest effective SNR among all users.
- DPC exploits up to n_t dimensions. Since the quality of each of those n_t dimensions cannot be larger than the single dimension used in the TDMA lower bound, DPC cannot achieve a rate larger than n_t times the TDMA capacity.

Multi-User MIMO - Scheduling and Precoding (Downlink)

Reference Book

- 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.



  - Chapter 12
    - Section: 12.1,12.5,12.6,12.8

Introduction

- BC: $K >> n_t$, MAC: $K >> n_r \rightarrow$ All users cannot be scheduled at the same time.
  - Which users to schedule?
  - How to account for fairness?

- DPC is optimal in MIMO BC but is very complex to implement.
  - Can we derive suboptimal strategies? Yes, there are various linear and non-linear precoding techniques
  - How to design suboptimal linear precoders?
  - What is the performance of those precoders combined with scheduling?

- What if we do not have perfect channel knowledge at the transmitter to design the precoders in 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
  - $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$
  - $\mathbf{H}_q \in \mathbb{C}^{n_{r,q} \times n_t}$ 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 $\mathcal{CN}\left(0, \sigma_{n,q}^2 \mathbf{I}_{n_{r,q}}\right)$.
- 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: Tr$\{\mathbf{Q}\} \leq E_s$ with $\mathbf{Q} = \mathcal{E}\left\{\mathbf{c}'\mathbf{c}'^H\right\}$.

System Model - Linear 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

$$\mathbf{c}' = \mathbf{P}\mathbf{c} = \mathbf{W}\mathbf{S}^{1/2}\mathbf{c} = \sum_{q \in \mathbf{K}} \mathbf{P}_q \mathbf{c}_q = \sum_{q \in \mathbf{K}} \mathbf{W}_q \mathbf{S}_q^{1/2} \mathbf{c}_q$$

  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 $\mathrm{Tr}\{\mathbf{P}\mathbf{P}^H\} \leq E_s$, 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

$$\begin{aligned}\mathbf{z}_q &= \mathbf{G}_q \mathbf{y}_q, \\ &= \Lambda_q^{-1/2} \mathbf{G}_q \mathbf{H}_q \mathbf{W}_q \mathbf{S}_q^{1/2} \mathbf{c}_q + \sum_{p \in \mathbf{K},\, p \neq q} \Lambda_q^{-1/2} \mathbf{G}_q \mathbf{H}_q \mathbf{W}_p \mathbf{S}_p^{1/2} \mathbf{c}_p + \mathbf{G}_q \mathbf{n}_q.\end{aligned}$$

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.

Multi-User Diversity Gain in SISO

- 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 $\bar{\rho}_{out}$

$$\bar{\rho}_{out} = \mathcal{E}\left\{\eta \max_{q=1,\ldots,K} |h_q|^2\right\} = \eta \sum_{q=1}^{K} \frac{1}{q}.$$

  - SNR gain provided by MU diversity g_m

$$g_m = \frac{\bar{\rho}_{out}}{\eta} = \sum_{q=1}^{K} \frac{1}{q} \overset{K\to\infty}{\cong} \log(K).$$

    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 in SISO

- 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: $\mathcal{E}_x\{\mathcal{F}(x)\} \leq \mathcal{F}(\mathcal{E}_x\{x\})$ 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

- In MU-MIMO, the performance is function of the channel magnitude but also of the spatial directions and properties of the channel matrices.
- MU diversity offers abundant spatial channel directions and allows to appropriately choose users with good channel matrix properties or spatial separations.
- Opportunistic Beamforming: precode multiple streams along the unitary precoding matrix $\mathbf{W}$ (orthogonal beams). For a large number of users, thanks to MU diversity, each beam matches one user channel with a high probability and orthogonality of beams prevents users from experiencing multi-user interference

$$y_q = \Lambda_q^{-1/2}\mathbf{h}_q\mathbf{W}\mathbf{S}^{1/2}\mathbf{c} + n_q \overset{K\to\infty}{=} \Lambda_q^{-1/2} \left\|\mathbf{h}_q\right\| s_q^{1/2} c_q + n_q.$$

  - The terminal only measures the effective channel, i.e. the channel precoded by each beam, and reports the SNR (or CQI) for one or multiple beam(s).
  - Works well only for very large K.

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.

Resource Allocation, Fairness and Scheduling Criteria

- 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 $\mathcal{U}$.

$$\left\{\mathbf{c}'^{\star}, \mathbf{G}^{\star}, \mathbf{K}^{\star}\right\} = \arg \max_{\mathbf{c}', \mathbf{G}, \mathbf{K} \subset \mathcal{K}} \mathcal{U}$$

  where $\mathbf{c}'^{\star}$ is the optimum transmit vector, $\mathbf{G}^{\star}$ denotes the optimum set of receive beamformers, and $\mathbf{K}^{\star} \subset \mathcal{K}$ refers to the optimum 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.
- Those two strategies can be addressed by using two different utility metrics:

$$\left\{\mathbf{c}'^{\star}, \mathbf{G}^{\star}, \mathbf{K}^{\star}\right\} = \arg \max_{\mathbf{c}', \mathbf{G}, \mathbf{K} \subset \mathcal{K}} \sum_{q \in \mathbf{K}} w_q R_q$$

  where
  - rate-maximization approach: $w_q = 1$
  - 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

$$\bar{R}_q\,(k+1) = \begin{cases} (1-1/t_c)\,\bar{R}_q\,(k) + 1/t_c R_q(k), & q \in \mathbf{K}^\star \\ (1-1/t_c)\,\bar{R}_q\,(k)\,, & q \notin \mathbf{K}^\star \end{cases}$$

  where t_c is the scheduling time scale and $\mathbf{K}^\star$ refers to the scheduled user set at time k. The resources should thus be allocated at time instant k as

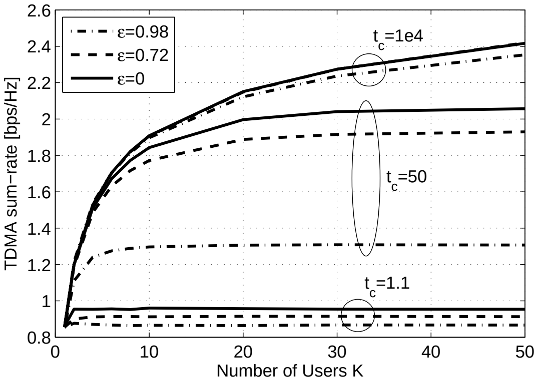
$$\left\{\mathbf{c}'^\star, \mathbf{G}^\star, \mathbf{K}^\star\right\} = \arg \max_{\mathbf{c}', \mathbf{G}, \mathbf{K} \subset \mathcal{K}} \sum_{q \in \mathbf{K}} \gamma_q \frac{R_q(k)}{\bar{R}_q(k)}.$$

- 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.

- 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

$$h_k = \epsilon h_{k-1} + \sqrt{1-\epsilon^2} n_k$$

with ϵ the channel time correlation coefficient.



User Grouping

- Given the presence of K users in the cell, the scheduler for MU-MIMO aims at finding the best scheduled user set among all possible candidates within $\mathcal{K}$.

- The *exhaustive search* is computationally intensive. Assuming a single stream transmission per user and $n_e \leq \min\{n_t, K\}$, a search like (with $R(\mathbf{K}) = \sum_{q \in \mathbf{K}} w_q R_q$)

$$\mathbf{K}^\star = \arg \max_{\substack{\mathbf{K} \subset \mathcal{K} \\ n_e \leq \min\{n_t, K\}}} R(\mathbf{K})$$

  requires to consider a large number of different sets and has a complexity that quickly becomes cumbersome as K increases.

Precoding with Perfect Transmit Channel Knowledge

- Single-link Spatial Multiplexing: Multiple Eigenmode Transmission relies on CSI knowledge at both the transmitter and the receiver and splits the spatial channel equalization between the transmitter and the receiver. As a result, the channel is decoupled into multiple parallel data pipes.
- Unfortunately, this approach cannot be applied to MU-MIMO as the receivers do not cooperate.
- MIMO BC, DPC optimal but extremely complex. Any suboptimal strategies?
- In MU-MIMO where CSI is available at the transmitter, precoding techniques reminiscent of the receiver architectures for SM

precoding	transmitter side	receiver side
Linear	Matched Beamforming (MBF)	MRC
Linear	**Zero-Forcing Beamforming (ZFBF)**	ZF
Linear	Regularized Zero-Forcing Beamforming (R-ZFBF)	MMSE
Non-linear	Tomlinson-Harashima Precoding (THP)	SIC
Non-linear	Vector Perturbation (VP)	sphere decoder

Achievable rate

- Maximum rate achievable by user q with linear precoding is

$$R_q = \sum_{l=1}^{n_{u,q}} \log_2 (1 + \rho_{q,l})$$

where $\rho_{q,l}$ denotes the SINR experienced by stream l of user q

$$\rho_{q,l} = \frac{\Lambda_q^{-1} \left|\mathbf{g}_{q,l}\mathbf{H}_q\mathbf{p}_{q,l}\right|^2}{I_l + I_c + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2} = \frac{\Lambda_q^{-1} \left|\mathbf{g}_{q,l}\mathbf{H}_q\mathbf{w}_{q,l}\right|^2 s_{q,l}}{I_l + I_c + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2}$$

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 $\rho_q = \frac{\Lambda_q^{-1}\left|\mathbf{h}_q\mathbf{w}_q\right|^2 s_q}{\sum_{\substack{p \in \mathbf{K} \\ p \neq q}} \Lambda_q^{-1}\left|\mathbf{h}_q\mathbf{w}_p\right|^2 s_p + \sigma_{n,q}^2}$.

Zero-Forcing Beamforming (ZFBF)

- Most popular MU-MIMO precoder. Assume single receive antenna per user.
- Channel Direction Information (CDI) of user q: $\bar{\mathbf{h}}_q = \mathbf{h}_q / \|\mathbf{h}_q\|$.
- Idea is to force the intra-cell interference I_c to zero: the precoder of a user q, $\mathbf{w}_q$, is chosen such that $\mathbf{h}_p \mathbf{w}_q = 0 \; \forall p \in \mathbf{K} \setminus q$. Only possible if $n_e \leq n_t$!
- Define

$$\mathbf{H} = \left[\Lambda_i^{-1/2} \mathbf{h}_i^T, \ldots, \Lambda_j^{-1/2} \mathbf{h}_j^T \right]_{i,j \in \mathbf{K}}^T = \mathbf{D} \bar{\mathbf{H}}$$

with

$$\mathbf{D} = \mathrm{diag} \left\{ \Lambda_i^{-1/2} \|\mathbf{h}_i\|, \ldots, \Lambda_j^{-1/2} \|\mathbf{h}_j\| \right\}_{i,j \in \mathbf{K}},$$

$$\bar{\mathbf{H}} = \left[\bar{\mathbf{h}}_i^T, \ldots, \bar{\mathbf{h}}_j^T \right]_{i,j \in \mathbf{K}}^T.$$

The ZFBF aims at designing $\mathbf{W} = [\mathbf{w}_i, \ldots, \mathbf{w}_j]_{i,j \in \mathbf{K}}$ such that $\mathbf{HW}$ is diagonal.

- Assuming $n_e \leq n_t$ and $\bar{\mathbf{H}}$ is full rank, the precoders can be chosen as the normalized columns of the right pseudo inverse of $\mathbf{H}$

$$\mathbf{F} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H \right)^{-1} = \bar{\mathbf{F}} \mathbf{D}^{-1} = \bar{\mathbf{H}}^H \left(\bar{\mathbf{H}} \bar{\mathbf{H}}^H \right)^{-1} \mathbf{D}^{-1}.$$

Transmit precoder $\mathbf{w}_q$ for user $q \in \mathbf{K}$: $\mathbf{w}_q = \mathbf{F}(:,q) / \|\mathbf{F}(:,q)\| = \bar{\mathbf{F}}(:,q) / \|\bar{\mathbf{F}}(:,q)\|$ where $\mathbf{F}(:,q)$ is to be viewed as the column of $\mathbf{F}$ corresponding to user q.

Zero-Forcing Beamforming (ZFBF)

- Assuming that $\mathbf{c} = \begin{bmatrix} c_i, \ldots, c_j \end{bmatrix}^T_{i,j \in \mathbf{K}}$, the received signal of user $q \in \mathbf{K}$ is

$$\mathbf{y}_q = \Lambda_q^{-1/2} \mathbf{h}_q \mathbf{w}_q s_q^{1/2} c_q + n_q = d_q c_q + n_q,$$

  with $d_q = \Lambda_q^{-1/2} \mathbf{h}_q \mathbf{w}_q s_q^{1/2} = \Lambda_q^{-1/2} \frac{\|\mathbf{h}_q\|}{\|\bar{\mathbf{F}}(:,q)\|} s_q^{1/2}$.
  Observations: MU-MIMO channel with ZFBF is split into n_e parallel (non-interfering) channels.
- The rate achievable by user q is given by

$$R_q = \log_2\left(1 + d_q^2/\sigma_{n,q}^2\right).$$

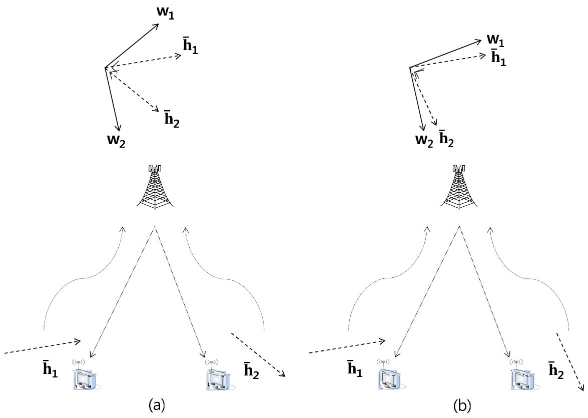
  d_q^2 is low if $\mathbf{H}$ is badly conditioned but would get larger if users' CDI are orthogonal or quasi-orthogonal.
  - reminiscent of the loss caused by noise enhancement incurred by the linear ZF
- For large K, better conditioning of matrix $\mathbf{H}$ through the use of user grouping.
- By uniformly allocating the power across user streams $s_q = E_s/n_e$ and by choosing $n_e = \tilde{n} = \min\{n_t, K\}$, $d_q^2/\sigma_{n,q}^2 = \alpha_q^2 \eta_q / n_e$ with $\alpha_q^2 = |\mathbf{h}_q \mathbf{w}_q|^2 = \|\mathbf{h}_q\|^2 / \|\bar{\mathbf{F}}(:,q)\|^2$

$$C_{BF}(\mathbf{H}) = \sum_{q=1}^{\min\{n_t,K\}} \log_2\left(1 + \alpha_q^2 \frac{\eta_q}{\tilde{n}}\right).$$

  At high SNR with $\eta_q = \eta$, $C_{BF}(\mathbf{H}) \approx \min\{n_t, K\} \log_2(\eta_q)$. The multiplexing gain $\min\{n_t, K\}$ is achieved (same as with DPC).

Zero-Forcing Beamforming (ZFBF)

- Illustration of ZFBF precoding for a two-user scenario: (a) non-orthogonal user set, (b) quasi-orthogonal user set.



Block Diagonalization (BD)

- Extension of ZFBF to multiple receive antennas and multiple streams per user.
- Constraints on the transmit filters targeting user $q \in \mathbf{K}$

$$\Lambda_p^{-1/2}\mathbf{H}_p\mathbf{W}_q = \mathbf{0}, \forall p \neq q, p \in \mathbf{K}$$

- Denoting $\tilde{\mathbf{K}}_q = \mathbf{K} \setminus q$ of size $\tilde{K}_q = \sharp\tilde{\mathbf{K}}_q$, the interference space $\tilde{\mathbf{H}}_q \in \mathbb{C}^{n_r\tilde{K}_q \times n_t}$ is

$$\tilde{\mathbf{H}}_q = \left[\begin{array}{ccc} \ldots & \Lambda_p^{-1/2}\mathbf{H}_p^T & \ldots \end{array} \right]_{p \in \tilde{\mathbf{K}}_q}^T.$$

- BD filter design forces $\mathbf{W}_q$ to lie in the null space of $\tilde{\mathbf{H}}_q$: null space of $\tilde{\mathbf{H}}_q$ to be strictly larger than 0 $\rightarrow r\left(\tilde{\mathbf{H}}_q\right) < n_t$.
- An orthogonal basis of the null space of $\tilde{\mathbf{H}}_q$ is obtained by taking its SVD

$$\tilde{\mathbf{H}}_q = \tilde{\mathbf{U}}_q\tilde{\mathbf{\Lambda}}_q \left[\begin{array}{cc} \tilde{\mathbf{V}}_q & \tilde{\mathbf{V}}_q' \end{array} \right]^H$$

  where $\tilde{\mathbf{V}}_q'$ refers to the eigenvectors of $\tilde{\mathbf{H}}_q$ associated with the null singular values.
- Assuming the zero-interference constraint is possible for all users in $\tilde{\mathbf{K}}_q$ and that $r\left(\mathbf{H}_q\tilde{\mathbf{V}}_q'\right) = n_{u,q}$, $\mathbf{W}_q$ writes as a linear combination of columns of $\tilde{\mathbf{V}}_q'$ as

$$\mathbf{W}_q = \tilde{\mathbf{V}}_q'\mathbf{A}_q$$

  with some $n_{u,q} \times n_{u,q}$ unitary matrix $\mathbf{A}_q$.

Block Diagonalization (BD)

- Multi-user interference is eliminated and each user experiences an equivalent single-user MIMO channel $\tilde{\mathbf{H}}_{eq,q} = \mathbf{H}_q \tilde{\mathbf{V}}'_q$, for which the optimal solution is obtained by transmitting along the $n_{u,q}$ dominant right singular vectors of $\tilde{\mathbf{H}}_{eq,q}$

$$\tilde{\mathbf{H}}_{eq,q} = \begin{bmatrix} \tilde{\mathbf{U}}_{eq,q} & \tilde{\mathbf{U}}'_{eq,q} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{eq,q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{V}}_{eq,q} & \tilde{\mathbf{V}}'_{eq,q} \end{bmatrix}^H$$

  where $\tilde{\mathbf{V}}_{eq,q}$ refers to the $n_{u,q}$ dominant right singular vectors.
- The final beamformer for user q writes as

$$\mathbf{W}_q = \tilde{\mathbf{V}}'_q \tilde{\mathbf{V}}_{eq,q}.$$

- Applying $\mathbf{G}_q = \tilde{\mathbf{U}}_{eq,q}^H$, the equivalent channel of each user is

$$\mathbf{z}_q = \mathbf{G}_q \mathbf{y}_q = \Lambda_q^{-1/2} \tilde{\mathbf{\Lambda}}_{eq,q} \mathbf{S}_q^{1/2} \mathbf{c}_q + \mathbf{G}_q \mathbf{n}_q.$$

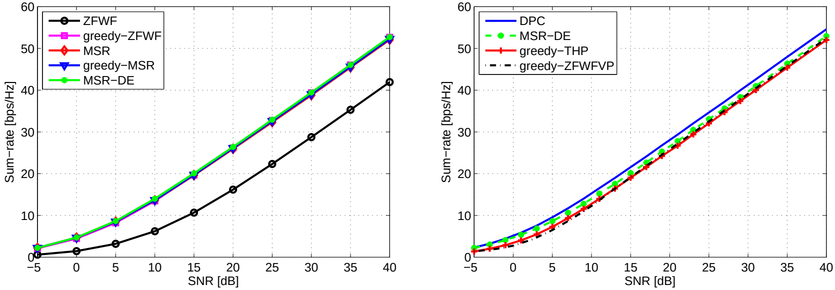
- Achievable sum-rate (with $\tilde{\lambda}_{eq,q,m}$ diagonal entries of $\tilde{\mathbf{\Lambda}}_{eq,q}^2$)

$$\sum_{q \in \mathbf{K}} \sum_{m=1}^{n_{u,q}} \log_2 \left(1 + s_{q,m} \frac{\Lambda_q^{-1} \tilde{\lambda}_{eq,q,m}}{\sigma_{n,q}^2} \right)$$

  For a sum-power constraint $\sum_{q \in \mathbf{K}} \sum_{m=1}^{n_{u,q}} s_{q,m} = E_s$, the optimal power allocation is obtained by water-filling.

Global Performance Comparison

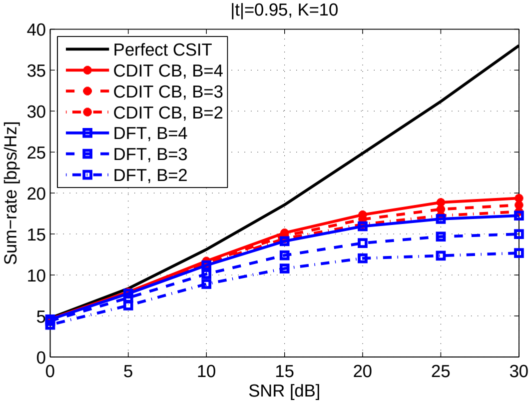
- Sum-rate of linear (left) and non-linear (right) MU-MIMO precoders vs SNR in $n_t = 4, K = 20$ i.i.d. Rayleigh fading channels



Observations: ZFBF without user selection (ZFWF) performs poorly. ZFBF with user selection (greedy-ZFWF) is a competitive strategy for MU-MIMO broadcast channels, in terms of both performance and complexity.
Keep in mind the assumptions: perfect CSIT, the same average SNR for all users and a max-rate scheduler (i.e. there is no fairness issue involved here).

Precoding with Partial Transmit Channel Knowledge

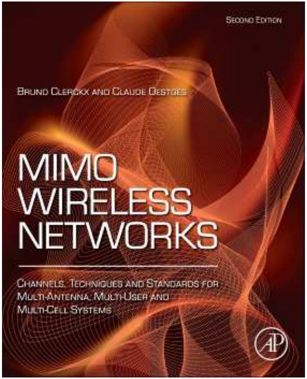
- If imperfect CSIT (e.g. quantized feedback), residual interference term in the SINR does not vanish with SNR, therefore inducing a ceiling effect as SNR increases.
- Performance of channel statistics-based codebook (CDIT-CB) and DFT codebook with Greedy user selection for $B = 2, 3, 4$, $n_t = 4$, $|t| = 0.95$ and $K = 10$.



Introduction to Multi-Cell MIMO

Reference Book

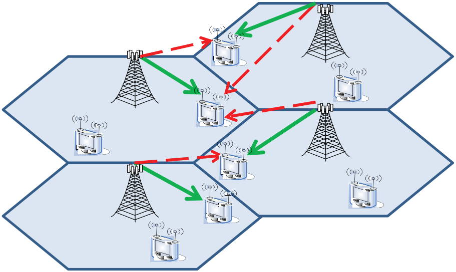
- Bruno Clerckx and Claude Oestges, "MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems," Academic Press (Elsevier), Oxford, UK, Jan 2013.



  - Chapter 13
    - Section: 13.1, 13.2, 13.3

Introduction

- 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.



Cellular network

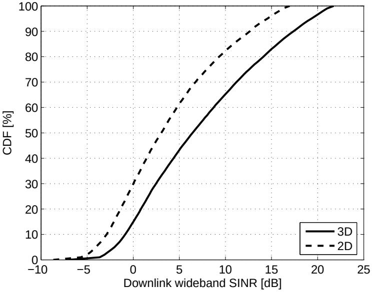
desired signal
interference

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.



Classical Inter-Cell Interference Mitigation

- Divide-and-conquer approach:
  - fragmenting the network area into small zones independently controlled from each other
  - making progressively use of advanced error correction coding, link adaptation, frequency selective scheduling and lately single-user and multi-user MIMO in each of those zones.

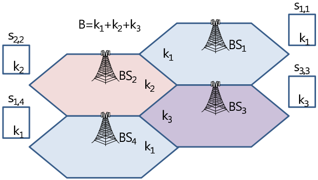


Figure: Frequency Reuse Partitioning.

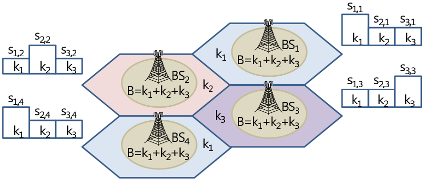


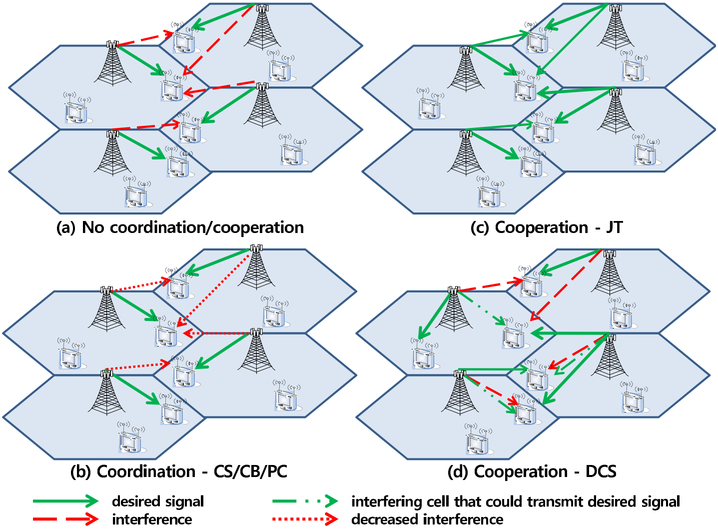
Figure: Static Fractional Frequency Reuse.

Towards Multi-Cell MIMO: Coordination and Cooperation

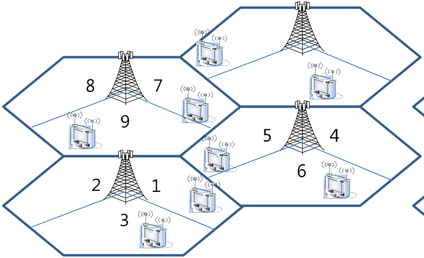
- Jointly allocate resources across the whole network (and not for each cell independently) and use the antennas of multiple cells to improve the received signal quality at the mobile terminal and to reduce the co-channel interferences.

- Two categories:
  - Coordination: No data sharing (user data is available at a single transmsitter) - CSI sharing. Modelled by an Interference Channel and Interfering Broadcast/Multiple Access Channel
  - Cooperation: Data sharing (user data is available at multiple transmsitters) - CSI sharing. Modelled by a Broadcast Channel (for Downlink) and Multiple Access Channel (for Uplink)

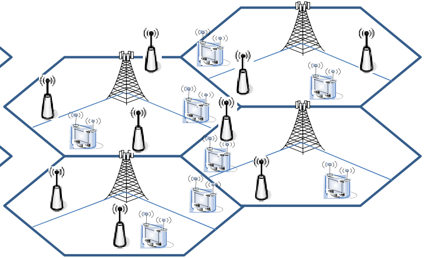
Towards Multi-Cell MIMO: Coordination and Cooperation



Network Deployments



(a) Homogeneous network



(b) Heterogeneous network

System Model - Interference Channel

- Interfering (broadcast/multiple access) channel
  - For each transmitter i (one per cell), the intended receivers (i.e. users) are in cell i.
  - Each receiver (i.e. user) is only interested in what is being sent by the corresponding transmitter.
  - Transmitters and receivers do not cooperate but only coordinate their transmissions by sharing CSI information. In the downlink, one transmitter does not have access to the codewords sent by other transmitters and cannot perform DPC. In the uplink, one receiver never has access to other received signals and cannot perform SIC.
- 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

$$\mathbf{y}_q = \Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{c}_i' + \underbrace{\sum_{j \neq i} \Lambda_{q,j}^{-1/2}\mathbf{H}_{q,j}\mathbf{c}_j'}_{\text{inter-cell interference}} + \mathbf{n}_q$$

  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

$$\mathbf{c}_i' = \mathbf{P}_i \mathbf{c}_i = \mathbf{W}_i \mathbf{S}_i^{1/2} \mathbf{c}_i = \sum_{q \in \mathbf{K}_i} \mathbf{P}_{q,i} \mathbf{c}_{q,i} = \sum_{q \in \mathbf{K}_i} \mathbf{W}_{q,i} \mathbf{S}_{q,i}^{1/2} \mathbf{c}_{q,i}$$

  where

  - $\mathbf{c}_i$ 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 $\mathbf{Q}_i = \mathcal{E}\{\mathbf{c}_i' \mathbf{c}_i'^H\}$ subject to the transmit power constraint $\mathrm{Tr}\{\mathbf{Q}_i\} \leq E_{s,i}$.

Linear Precoding

- The received signal $\mathbf{y}_q \in \mathbb{C}^{n_{r,q}}$ of user $q \in \mathbf{K}_i$

$$\mathbf{y}_q = \Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{P}_{q,i}\mathbf{c}_{q,i} + \underbrace{\sum_{p\in\mathbf{K}_i,\, p\neq q} \Lambda_{q,i}^{-1/2}\mathbf{H}_{q,i}\mathbf{P}_{p,i}\mathbf{c}_{p,i}}_{\text{intra-cell (multi-user) interference}}$$

$$+ \underbrace{\sum_{j\neq i}\sum_{l\in\mathbf{K}_j} \Lambda_{q,j}^{-1/2}\mathbf{H}_{q,j}\mathbf{P}_{l,j}\mathbf{c}_{l,j}}_{\text{inter-cell interference}} + \mathbf{n}_q.$$

- Apply a receive combiner to stream l of user q in cell i

$$z_{q,l} = \mathbf{g}_{q,l}\mathbf{y}_q = \Lambda_{q,i}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}c_{q,i,l} + \underbrace{\sum_{m\neq l} \Lambda_{q,i}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}c_{q,i,m}}_{\text{inter-stream interference}}$$

$$+ \underbrace{\sum_{p\in\mathbf{K}_i,\, p\neq q} \Lambda_{q,i}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{P}_{p,i}\mathbf{c}_{p,i}}_{\text{intra-cell (multi-user) interference}} + \underbrace{\sum_{j\neq i}\sum_{l\in\mathbf{K}_j} \Lambda_{q,j}^{-1/2}\mathbf{g}_{q,l}\mathbf{H}_{q,j}\mathbf{P}_{l,j}\mathbf{c}_{l,j}}_{\text{inter-cell interference}} + \mathbf{g}_{q,l}\mathbf{n}_q.$$

Achievable Rate

- By treating all interference as noise, the maximum rate achievable by user q in cell i with linear precoding is

$$R_{q,i} = \sum_{l=1}^{n_{u,q}} \log_2 \left(1 + \rho_{q,l}\right).$$

- The quantity $\rho_{q,l}$ denotes the SINR experienced by stream l of user-q and writes as

$$\rho_{q,l} = \frac{S}{I_l + I_c + I_o + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2}.$$

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

$$\begin{aligned}
S &= \Lambda_{q,i}^{-1} \left|\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right|^2, \\
I_l &= \sum_{m \neq l} \Lambda_{q,i}^{-1} \left|\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}\right|^2, \\
I_c &= \sum_{p \in \mathbf{K}_i, p \neq q} \sum_{m=1}^{n_{u,p}} \Lambda_{q,i}^{-1} \left|\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{p,i,m}\right|^2, \\
I_o &= \sum_{j \neq i} \Lambda_{q,j}^{-1} \left\|\mathbf{g}_{q,l}\mathbf{H}_{q,j}\mathbf{P}_j\right\|^2.
\end{aligned}$$

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: $I_l + I_c + I_o + \left\|\mathbf{g}_{q,l}\right\|^2 \sigma_{n,q}^2 = \mathbf{g}_{q,l}\mathbf{R}_{\mathbf{n_i}}\mathbf{g}_{q,l}^H$ where

$$\begin{aligned}\mathbf{R}_{\mathbf{n_i}} = \sum_{m\neq l} \Lambda_{q,i}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,m}\right)^H \\ + \sum_{p\in\mathbf{K}_i, p\neq q}\sum_{m=1}^{n_{u,p}} \Lambda_{q,i}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{p,i,m}\left(\mathbf{H}_{q,i}\mathbf{p}_{p,i,m}\right)^H \\ + \sum_{j\neq i}\Lambda_{q,j}^{-1}\mathbf{H}_{q,j}\mathbf{P}_j\left(\mathbf{H}_{q,j}\mathbf{P}_j\right)^H + \sigma_{n,q}^2\mathbf{I}_{n_{r,q}}\end{aligned}$$

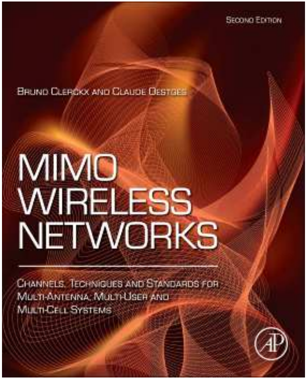
  is the covariance matrix of the noise plus interference.
- MMSE combiner for stream l: $\mathbf{g}_{q,l} = \Lambda_{q,i}^{-1/2}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right)^H \mathbf{R}_{\mathbf{n_i}}{}^{-1}$
- SINR $\rho_{q,l}$ experienced by stream l of user-q

$$\rho_{q,l} = \frac{\Lambda_{q,i}^{-1}\left|\mathbf{g}_{q,l}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right|^2}{\mathbf{g}_{q,l}\mathbf{R}_{\mathbf{n_i}}\mathbf{g}_{q,l}^H} = \Lambda_{q,i}^{-1}\left(\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}\right)^H\mathbf{R}_{\mathbf{n_i}}{}^{-1}\mathbf{H}_{q,i}\mathbf{p}_{q,i,l}.$$

Capacity of the Interference Channel

Reference Book

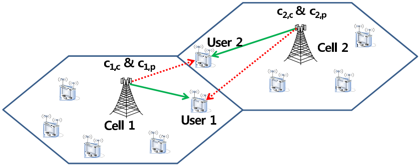
- Bruno Clerckx and Claude Oestges, "MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems," Academic Press (Elsevier), Oxford, UK, Jan 2013.



  - Chapter 13
    - Section: 13.4

SISO Interference Channel

- What is the capacity region of the two-user SISO IC?



- $\eta_{q,i} = \Lambda_{q,i}^{-1} E_{s,i} / \sigma_{n,q}^2$
  - long-term SNR when user q is served by cell i
  - long-term INR (interference to noise ratio) when q is a victim user of cell i
- $\tilde{\eta}_{q,i} = \eta_{q,i} \left| h_{q,i} \right|^2$ can be thought of as an instantaneous SNR or INR
- two-user SISO IC: transmitter 1 (i.e. cell 1) communicates with user 1 and transmitter 2 (i.e. cell 2) with user 2
  - achievable rate region function of $\tilde{\eta}_{1,1}, \tilde{\eta}_{2,2}, \tilde{\eta}_{1,2}, \tilde{\eta}_{2,1}$
  - symmetric SISO IC characterized by $\tilde{\eta}_{1,1} = \tilde{\eta}_{2,2} = \tilde{\eta}_d$ and $\tilde{\eta}_{1,2} = \tilde{\eta}_{2,1} = \tilde{\eta}_c$
  - symmetric rate: $R_{sym} = \max_{(R_1, R_2) \in \mathcal{C}_{IC}} \min \{R_1, R_2\}$ where R_1 and R_2 are the rates achievable by user 1 and 2 respectively in the two-user SISO IC and $\mathcal{C}_{IC}$ is the capacity region of the SISO IC

Degrees of Freedom - Multiplexing Gain

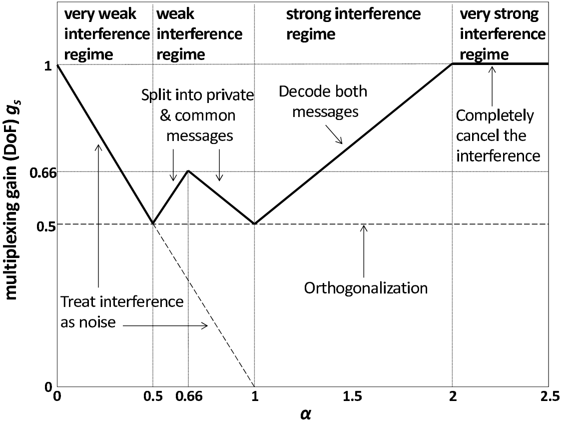


Figure: Achievable multiplexing gain per user of the two-user Gaussian SISO IC (α = INR/SNR).

Very Strong Interference Regime

- Can the capacity region, under some interference conditions, become a square only determined by the inequalities $R_i \leq \log_2 (1 + \tilde{\eta}_{i,i}), i = 1, 2$?
  - i.e. each transmitter can communicate with its receiver at a rate equal to the one achievable without any interference
- Very strong interference regime conditions: $\tilde{\eta}_{1,2} \geq \tilde{\eta}_{2,2} + \tilde{\eta}_{1,1}\tilde{\eta}_{2,2}$ and $\tilde{\eta}_{2,1} \geq \tilde{\eta}_{1,1} + \tilde{\eta}_{1,1}\tilde{\eta}_{2,2}$
- The interference is so strong that each user performs SIC by decoding the interfering message first and subtracting it from the received signal before decoding its own message.
- Each transmitter can communicate with its receiver at a rate $R_i = \log_2 \left(1 + \tilde{\eta}_{i,i}\right)$ for $i = 1, 2$, as in the absence of any interference.
- The symmetric rate (symmetric capacity) simply writes as

$$R_{sym} = \log_2 (1 + \tilde{\eta}_d) .$$

Very Strong Interference Regime

- The very strong interference conditions can be viewed from an angle that is reminiscent of the SIC behavior in SISO BC.
- When user 1 decodes user 2's signal in the very strong interference regime, it treats its own signal as noise. Hence, for user 1 to be able to cancel correctly user 2's signal, the interfering channel between transmitter 2 and user 1 has to be strong enough to support R_2, i.e.

$$R_2 \leq \log_2\left(1 + \frac{\Lambda_{1,2}^{-1}\left|h_{1,2}\right|^2 E_{s,2}}{\sigma_{n,1}^2 + \Lambda_{1,1}^{-1}\left|h_{1,1}\right|^2 E_{s,1}}\right) = \log_2\left(1 + \frac{\tilde{\eta}_{1,2}}{1+\tilde{\eta}_{1,1}}\right).$$

Given that user 2 wants to receive its message at a rate $R_2 = \log_2\left(1+\tilde{\eta}_{2,2}\right)$, this puts the constraints

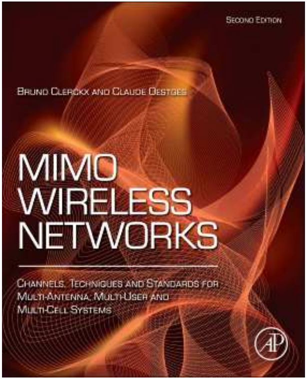
$$\log_2\left(1+\tilde{\eta}_{2,2}\right) \leq \log_2\left(1 + \frac{\tilde{\eta}_{1,2}}{1+\tilde{\eta}_{1,1}}\right),$$

which equivalently writes as $\tilde{\eta}_{1,2} \geq \tilde{\eta}_{2,2} + \tilde{\eta}_{1,1}\tilde{\eta}_{2,2}$. The other condition is obtained similarly by looking at user 2's requirement to decode user 1's message correctly.

Real-World MIMO Wireless Networks

Reference Book

- 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.



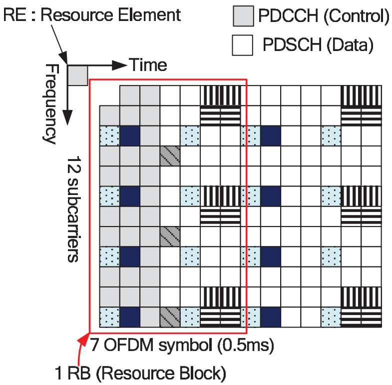
  - Chapter 14

System Requirements

- peak rate
  - highest theoretical throughput achievable with SU-MIMO spatial multiplexing but are typically not achieved in practical deployments.
  - e.g. 8x8 Spatial multiplexing with 8 streams transmission
- cell average spectral efficiency
  - average spectral efficiency of a cell (with K users).
  - much more representative of throughput encountered in practice
- cell edge user spectral efficiency
  - spectral efficiency achieved by at least 95% of the users in the network.
  - much more representative of throughput encountered in practice

Frame Structure

- Multiplexing/Access
  - DL: OFDM
  - UL: DFT-Spread OFDM (SC-FDM)
- Frame structure
  - OFDMA/SC-FDMA create a time-frequency grid composed of time-frequency resources
  - A resource block (RB) is formed by 12 consecutive REs in the frequency domain for a duration of 7 OFDM symbols in the time domain.
  - A subframe is formed of 14 consecutive OFDM/SC-FDM symbols.
  - Scheduling and data transmission is performed at the RB-level with the minimum scheduling unit consisting of two RBs within one subframe.
  - First 3 symbols used to carry control information.



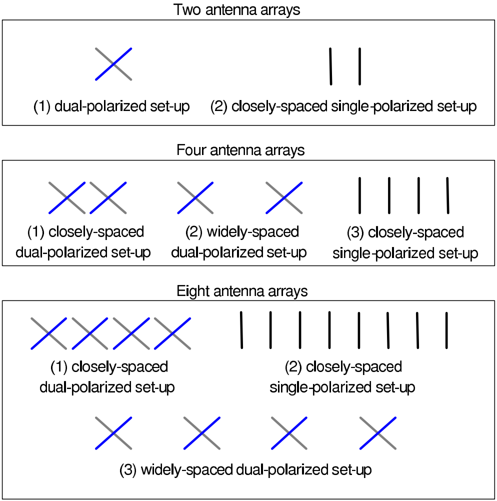
Key Downlink Technologies

- Antenna configurations: 2, 4 or 8 transmit antennas and a minimum of 2 receive antennas
- LTE Rel. 8 (finalized in Dec 2008):
  - Up to 4x4 (up to 4 stream transmission)
  - Transmit diversity (to protect against fading) using Orthogonal Space-Frequency Block Coding (O-SFBC) for 2Tx, non-orthogonal SFBC for 4TX
  - Open-loop (for high speed) Spatial Multiplexing with rank adaptation based on predefined precoders
  - Closed-loop (for low speed) Spatial Multiplexing based on codebook precoding
  - Stone-age MU-MIMO based on common reference signals (CRS)
- LTE Rel. 9 (finalized in Dec 2009):
  - Up to 4x4 (up to 4 stream transmission)
  - Introduction of demodulation reference signals (DM-RS)
  - Enhancement of MU-MIMO to support ZFBF-like precoding
- LTE-A Rel. 10 (finalized mid 2011):
  - Up to 8x8 (up to 8 stream transmission)
  - New channel measurement reference signals (CSI-RS)
  - New feedback mechanisms for 8Tx (dual codebook $\mathbf{W}_1\mathbf{W}_2$ structure)
  - HetNet - eICIC
- LTE-A Rel. 11 (finalized in Dec 2012):
  - Coordinated Multi-Point Transmission/Reception (CoMP) for Homogeneous (Macro) and heterogeneous (pico, DAS) networks
    - Dynamic cell/point selection combined with dynamic ON/OFF blanking

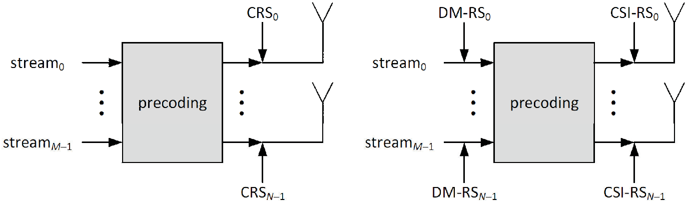
Key Uplink Technologies

- Antenna configurations: 1, 2 or 4 transmit antennas in the uplink with a minimum of 2 receive antennas
- LTE Rel. 8 (finalized in Dec 2008):
  - single antenna transmission and transmit antenna selection
  - MU-MIMO
- LTE-A Rel. 10 (finalized mid 2011):
  - Spatial Multiplexing with codebook
  - Transmit diversity (for control channels)
- LTE-A Rel. 11 (finalized in Dec 2012):
  - Coordinated Multi-Point Transmission/Reception (CoMP)

Antenna Deployments



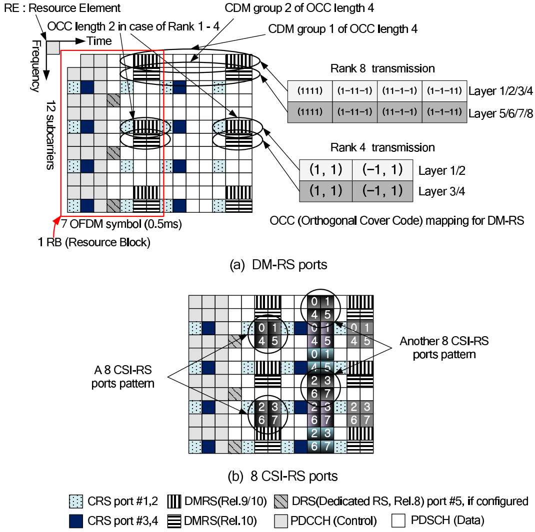
Reference Signals



CRS-based **DM-RS-based**

Dedicated RS (DRS)	**Common RS (CRS)**
For demodulation	For demodulation and measurement
Targets a specific terminal	Shared among a group of terminals
Terminal specifically precoded	Commonly non-precoded
Overhead proportional to the number of transmitted streams	Overhead proportional to the number of transmit antennas
Sent in RBs where data is present	Sent in all RBs
Channel estimation less flexible	Channel estimation more flexible

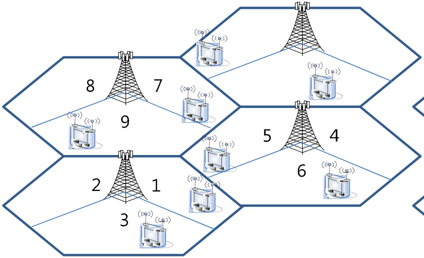
Reference Signals



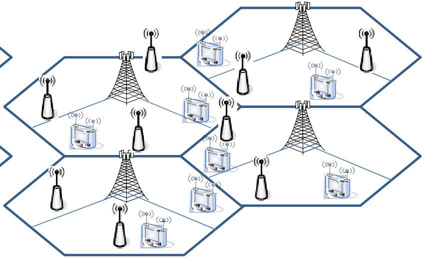
Channel State Information (CSI) feedback

- Three main feedback information:
  - Rank Indicator (RI): the preferred number of streams (denoted as layers in LTE) a user would like to receive
  - Precoding Matrix Indicator (PMI): the preferred precoder in the codebook
  - Channel Quality Indicator (CQI): the rate achievable with each stream (used to perform link adaptation)
- Open-Loop relies only on RI and CQI
  - High mobility or limited CSI feedback prevent the use of PMI
- Closed-Loop (Spatial Multiplexing and MU-MIMO) rely on RI, CQI and PMI
  - If Spatial Multiplexing, the actual precoder is the same as the one selected by the user (PMI)
  - If MU-MIMO based on CRS, the actual precoder is the same as the one selected by the user (PMI)
  - If MU-MIMO based on DM-RS, the actual precoder (e.g. ZFBF) is computed based on the one selected by the user (PMI).

Network Deployments



(a) Homogeneous network



(b) Heterogeneous network

Beyond LTE-A: Massive Multi-Cell and Massive Multi-Antenna Networks