Capacity of multiple-input multiple-output (MIMO) systems in wireless communications

Bengt Holter

Department of Telecommunications
Norwegian University of Science and Technology

Outline

- Introduction
- Channel capacity
  - Single-Input Single-Output (SISO)
  - Single-Input Multiple-Output (SIMO)
  - Multiple-Input Multiple-Output (MIMO)
  - MIMO capacity employing space-time block coding (STBC)
- Outage capacity
  - SISO
  - SIMO
  - MIMO employing STBC
- Summary

MIMO = Multiple-Input Multiple-Output

- Initial MIMO papers
  - I. Telatar, "Capacity of multi-antenna gaussian channels," AT&T Technical Memorandum, jun. 1995
  - G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", Bell Labs Technical Journal, 1996
- MIMO systems are used to (dramatically) increase the capacity and quality of a wireless transmission.
- Increased capacity obtained with spatial multiplexing of transmitted data.
- Increased quality obtained by using space-time coding at the transmitter.

- For a discrete random variable X with alphabet $\mathcal{X}$ and distributed according to the probability mass function $p(x)$, the entropy is defined as

$$H(X) = \sum_{x \in \mathcal{X}} \log_2 \frac{1}{p(x)} \cdot p(x) = -\sum_{x \in \mathcal{X}} \log_2 p(x) \cdot p(x) = \mathcal{E}\left[\log_2 \frac{1}{p(x)}\right]. \quad (1)$$

- The entropy of a random variable is a measure of the uncertainty of the random variable; it is a measure of the amount of information required on the average to describe the random variable.

- With a base 2 logarithm, entropy is measured in bits.

- X follows a gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ when the probability density function (PDF) of X is given by

$$f_X(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}. \tag{2}$$

  where $\Gamma(\cdot)$ is the gamma function ($\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$ $[\Re(\alpha) > 0]$).

- The short hand notation $X \sim \mathcal{G}(\alpha, \beta)$ is used to denote that X follows a gamma distribution with shape parameter α and scale parameter β.

- **Mean**: $\mathcal{E}\{X\} = \mu_x = \alpha \cdot \beta$.

- **Variance**: $\mathcal{E}\{X^2\} = \sigma_x^2 - \mu_x^2 = \alpha \cdot \beta^2$.

SISO = Single-input Single-output

- Representing the input and and output of a memoryless wireless channel with the random variables X and Y respectively, the channel capacity is defined as

$$C = \max_{p(x)} \mathcal{I}(X;Y). \tag{3}$$

- $\mathcal{I}(X;Y)$ denotes the mutual information between X and Y and it is measure of the amount of information that one random variable contains about another random variable.

- According to the definition in (3), the mutual information is maximized with respect to all possible transmitter statistical distributions $p(x)$.

- The mutual information between X and Y can also be written as

$$\mathcal{I}(X;Y) = H(Y) - H(Y|X). \tag{4}$$

- From the equation above, it can be seen that mutual information can be described as the reduction in the uncertainty of one random variable due to the knowledge of the other.

- The mutual information between X and Y will depend on the properties of the wireless channel used to convey information from the transmitter to the receiver.

- For a SISO flat fading wireless channel, the input/output relations (per channel use) can be modelled by the complex baseband notation

$$y = hx + n \tag{5}$$

- y represents a single realization of the random variable Y (per channel use).
- h represents the complex channel between the transmitter and the receiver.
- x represents the transmitted complex symbol.
- n represents complex additive white gaussian noise (AWGN).
- Note that in previous lectures by Prof. Alouini, the channel gain $|h|$ was denoted α. In this presentation, α is used as the shape parameter of a gamma distributed random variable.
- Based on different communication scenarios, $|h|$ may be modelled by various statistical distributions.
- Common multipath fading models are Rayleigh, Nakagami-q (Hoyt), Nakagami-n (Rice), and Nakagami-m.

Capacity with a transmit power constraint

- With an average transmit power constraint P_T, the channel capacity is defined as

$$C = \max_{p(x):P \leq P_T} \mathcal{I}(X;Y). \quad (6)$$

- If each symbol per channel use at the transmitter is denoted by x, the average power constraint can be expressed as $P = \mathcal{E}\{|x|^2\} \leq P_T$.

- Compared to the original definition in (3), the capacity of the channel is now defined as the maximum of the mutual information between the input random variable X and the output random variable Y over all statistical distributions on the input that satisfy the power constraint.

- Since both x and y are continuous upon transmission and reception, the channel is modelled as an amplitude continuous but time discrete channel.

Assumptions

- Perfect channel knowledge at the receiver.
- X is independent of N.
- $N \sim \mathcal{N}(0, \sigma_n^2)$

Mutual information: With $h_d(\cdot)$ denoting differential entropy (entropy of a continuous random variable), the mutual information may be expressed as

$$\begin{aligned} \mathcal{I}(X;Y) &= h_d(Y) - h_d(Y|X) && (7) \\ &= h_d(Y) - h_d(hX + N|X) && (8) \\ &= h_d(Y) - h_d(N|X) && (9) \\ &= h_d(Y) - h_d(N) && (10) \end{aligned}$$

- (9) follows from the fact that since h is assumed perfectly known by the receiver, there is no uncertainty in hX conditioned on X.
- (10) follows from the fact that N is assumed independent of X, i.e., there is no information in X which reduces the uncertainty of N.

Noise differential entropy

- Since N already is assumed to be a complex gaussian random variable, i.e., the noise PDF is given by

$$f_N(n) = \frac{1}{\pi\sigma_n^2} e^{-\frac{n^2}{\sigma_n^2}} \tag{11}$$

- Differential entropy

$$h_d(N) = -\int f_N(n) \log_2 f_N(n) dn \tag{12}$$

- Inserting the noise PDF into (12)

$$
\begin{aligned}
h_d(N) &= -\int f_N(n)\left[-\frac{n^2 \log_2 e}{\sigma_n^2} - \log_2\left(\pi\sigma_n^2\right)\right] dn \\
&= \frac{\log_2 e}{\sigma_n^2}\int n^2 f(n)dn + \log_2\left(\pi\sigma_n^2\right)\int f(n)dn \\
&= \frac{\mathcal{E}\left\{N^2\right\}}{\sigma_n^2}\log_2 e + \log_2\left(\pi\sigma_n^2\right) \\
&= \log_2 e + \log_2\left(\pi\sigma_n^2\right) \\
&= \log_2\left(\pi e\sigma_n^2\right),
\end{aligned}
$$

where $\mathcal{E}\{N^2\} = \sigma_n^2$.

Received signal power

- Since $h_d(N)$ is given, the mutual information $\mathcal{I}(X;Y) = h_d(Y) - h_d(N)$ is maximized by maximizing $h_d(Y)$.

- Since the normal distribution maximizes the entropy over all distributions with the same covariance, $\mathcal{I}(X;Y)$ is maximized when Y is assumed gaussian, i.e., $h_d(Y) = \log_2(\pi e \sigma_y^2)$, where $\mathcal{E}\{Y^2\} = \sigma_y^2$.

- Assuming the optimal gaussian distribution for X, the received average signal power σ_y^2 may be expressed as

$$\begin{aligned} \mathcal{E}\{Y^2\} &= \mathcal{E}\{(hX + N)(h^*X^* + N^*)\} && (13) \\ &= \sigma_x^2 |h|^2 + \sigma_n^2. && (14) \end{aligned}$$

SISO fading channel capacity

$$
\begin{aligned}
C &= h_d(Y) - h_d(N) && (15)\\
&= \log_2(\pi e(\sigma_x^2|h|^2 + \sigma_n^2)) - \log_2(\pi e \sigma_n^2) && (16)\\
&= \log_2\left(1 + \frac{\sigma_x^2}{\sigma_n^2}|h|^2\right) && (17)\\
&= \log_2\left(1 + \frac{P_T}{\sigma_n^2}|h|^2\right), && (18)
\end{aligned}
$$

where it is assumed that $\sigma_x^2 = P_T$.

- Denoting the total received signal-to-noise ratio (SNR) $\gamma_t = \frac{P_T}{\sigma_n^2}|h|^2$, the SISO fadig channel capacity is given by

$$C = \log_2(1 + \gamma_t)$$

- Note that since γ_t is a random variable, the capacity also becomes a random variable.

Nakagami-m fading ⇒ Gamma distributed SNR

- With the assumption that the fading amplitude $|h|$ is a Nakagami-m distributed random variable, the PDF is given by

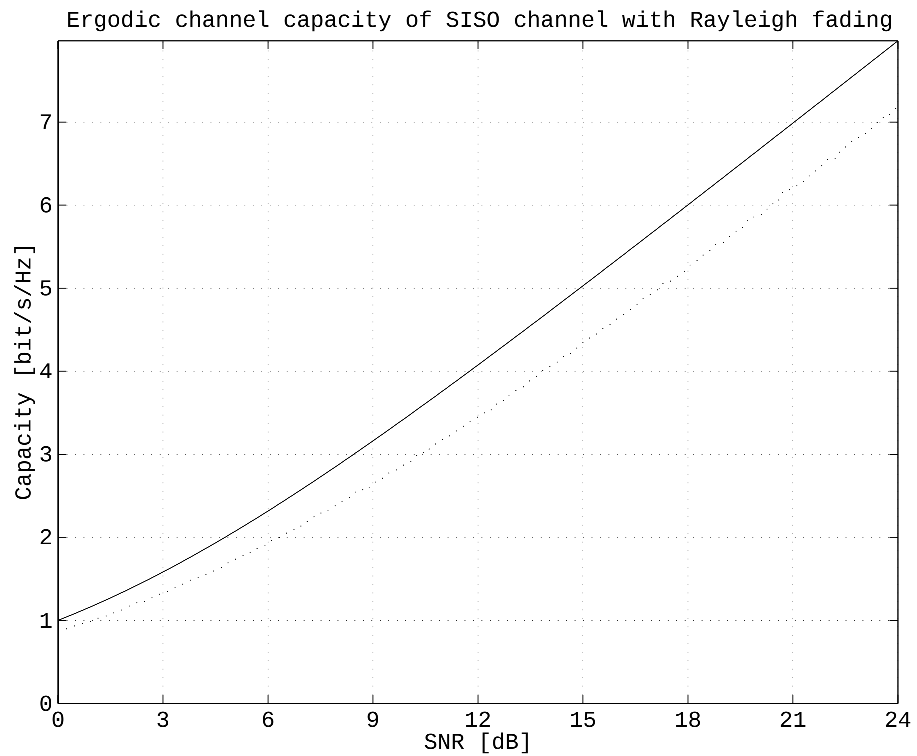
$$f_\alpha(\alpha) = \frac{2m^m |h|^{2m-1}}{\Omega^m \Gamma(m)} \exp\left(\frac{m|h|^2}{\Omega}\right) \tag{19}$$

  where $\Omega = \mathcal{E}\{|h|^2\}$ and m is the Nakagami-m fading parameter which ranges from 1/2 (half Gaussian model) to ∞ (AWGN channel).

- Using transformation of random variables, it can be shown that the overall received SNR γ_t is a gamma distributed random variable $\mathcal{G}(\alpha, \beta)$,

$$f_{\gamma_t}(\gamma_t) = \frac{\gamma_t^{m-1} e^{-\gamma_t/\beta}}{\beta^m \Gamma(m)}, \tag{20}$$

  where $\alpha = m$ and $\beta = \overline{\gamma}_t/m$. In short $\gamma_t \sim \mathcal{G}\left(m, \frac{\overline{\gamma}_t}{m}\right)$ where $\overline{\gamma}_t = \frac{P_T \Omega}{\sigma_n^2}$.



Ergodic capacity of a Rayleigh fading SISO channel (dotted line) compared to the Shannon capacity of a SISO channel (solid line)

3dB increase in SNR $\Rightarrow$ 1 bit/s/Hz capacity increase

SIMO = Single-Input Multiple-Output

- For a SIMO flat fading wireless channel, the input/output relations (per channel use) can be modelled by the complex baseband notation

$$\mathbf{y} = \mathbf{h}x + \mathbf{n} \tag{21}$$

- $\mathbf{y}$ represents a single realization of the multivariate random variable $\mathbf{Y}$ (array repsonse per channel use).

- $\mathbf{h}$ represents the complex channel vector between a single transmit antenna and n_R receive antennas, i.e., $\mathbf{h} = [h_{11}, h_{21}, \ldots, h_{n_R 1}]^{\mathsf{T}}$.

- x represents the transmitted complex symbol per channel use.

- $\mathbf{n}$ represents a complex additive white gaussian noise (AWGN) vector.

Mutual information

- With $h_d(\cdot)$ denoting differential entropy (entropy of a continuous random variable), the mutual information may be expressed as

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= h_d(\mathbf{Y}) - h_d(\mathbf{Y}|\mathbf{X}) && (22)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{hX}+\mathbf{N}|\mathbf{X}) && (23)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{N}|\mathbf{X}) && (24)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{N}) && (25) \end{aligned}$$

- It will be assumed that $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^n)$, where $\mathbf{K}^n = \mathcal{E}\{\mathbf{N}\mathbf{N}^{\mathsf{H}}\}$ is the noise covariance matrix.

- Since the normal distribution maxmizes the entropy over all distributions with the same covariance (i.e. the power constraint), the mutual information is maximized when $\mathbf{Y}$ represents a multivariate Gaussian random variable, i.e., $\mathbf{Y} = \mathcal{N}(\mathbf{0}, \mathbf{K}^y)$ where $\mathbf{K}^y = \mathcal{E}\{\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\}$ is the covariance matrix of the desired signal.

Desired signal covariance matrix

- For a complex gaussian vector $\mathbf{Y}$, the differential entropy is less than or equal to $\log_2 \det(\pi e \mathbf{K}^y)$, with equality if and only if $\mathbf{y}$ is a circularly symmetric complex Gaussian with $\mathcal{E}\{\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\} = \mathbf{K}^y$.

- With the assumption that the signal X is uncorrelated with all elements in $\mathbf{N}$, the received covariance matrix $\mathbf{K}^y$ may be expressed as

$$\begin{aligned} \mathcal{E}\{\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\} &= \mathcal{E}\{(\mathbf{h}X + \mathbf{N})(\mathbf{h}X + \mathbf{N})^{\mathsf{H}}\} && (26) \\ &= \sigma_x^2 \mathbf{h}\mathbf{h}^{\mathsf{H}} + \mathbf{K}^n && (27) \end{aligned}$$

  where $\sigma_x^2 = \mathcal{E}\{X^2\}$.

SIMO fading channel capacity

$$
\begin{aligned}
C &= h_d(\mathbf{Y}) - h_d(\mathbf{N}) && (28)\\
&= \log_2[\det(\pi e(\sigma_x^2 \mathbf{h}\mathbf{h}^{\mathsf{H}} + \mathbf{K}^n))] - \log_2[\det(\pi e \mathbf{K}^n)] && (29)\\
&= \log_2[\det(\sigma_x^2 \mathbf{h}\mathbf{h}^{\mathsf{H}} + \mathbf{K}^n)] - \log_2[\det \mathbf{K}^n] && (30)\\
&= \log_2[\det((\sigma_x^2 \mathbf{h}\mathbf{h}^{\mathsf{H}} + \mathbf{K}^n)(\mathbf{K}^n)^{-1})] && (31)\\
&= \log_2[\det(\sigma_x^2 \mathbf{h}\mathbf{h}^{\mathsf{H}}(\mathbf{K}^n)^{-1} + \mathbf{I}_{n_R})] && (32)\\
&= \log_2[\det(\mathbf{I}_{n_R} + \sigma_x^2 (\mathbf{K}^n)^{-1}\mathbf{h}^{\mathsf{H}}\mathbf{h})] && (33)\\
&= \log_2\left[\left(1 + \frac{P_T}{\sigma_n^2}||\mathbf{h}||^2\right) \cdot \det(\mathbf{I}_{n_R})\right] && (34)\\
&= \log_2\left(1 + \frac{P_T}{\sigma_n^2}||\mathbf{h}||^2\right) && (35)
\end{aligned}
$$

where it is assumed that $\mathbf{K}^n = \sigma_n^2 \mathbf{I}_{n_R}$ and $\sigma_x^2 = P_T$.

- Note that for the SISO fading channel, $\mathbf{K}^n = \sigma_n^2$.

- The capacity formula for the SIMO fading channel could also have been found by assuming maximum ratio combining at the receiver.
- With perfect channel knowledge at the receiver, the optimal weights are given by

$$\mathbf{w}_{opt} = (\mathbf{K}^n)^{-1}\mathbf{h}. \tag{36}$$

- Using these weights together with the assumption that $\mathbf{K}_n = \sigma_n^2 \mathbf{I}_{n_R}$, the overall (instantaneous) SNR γ_t for the current observed channel $\mathbf{h}$ is equal to

$$\gamma_t = \frac{P_T}{\sigma_n^2}||\mathbf{h}||^2. \tag{37}$$

- Thus, since γ_t in this case represents the maximum available SNR, the capacity can be written as

$$C = \log_2(1 + \gamma_t) = \log_2(1 + \frac{P_T}{\sigma_n^2}||\mathbf{h}||^2). \tag{38}$$

Nakagami-m fading $\Rightarrow$ Gamma distributed SNR

- With the assumption that all channel gains in the channel vector $\mathbf{h}$ are independent and indentically distributed (i.i.d.) Nakagami-m random variables (i.e. $m_l = m$), then the overall SNR γ_t is a gamma distributed random variable with shape parameter $\alpha = n_R \cdot m$ and scale parameter $\beta = \overline{\gamma}_l / m$)
- In short, $\gamma_t \sim \mathcal{G}(n_R \cdot m, \overline{\gamma}_l / m)$.
- $\overline{\gamma}_l$ represents the average SNR per receiver branch (assumed equal for all branches in this case)
- Coefficient of variation $\tau = \frac{\sigma_{\gamma_t}}{\mu_{\gamma_t}} = \frac{1}{\sqrt{n_R \cdot m}}$.
- Effective diversity order [Nabar,02]: $N_{div} = \frac{1}{\tau^2} = n_R \cdot m$.

MIMO = Multiple-Input Multiple-Output

- For a MIMO flat fading wireless channel, the input/output relations (per channel use) can be modelled by the complex baseband notation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{39}$$

- $\mathbf{x}$ is the $(n_T \times 1)$ transmit vector.
- $\mathbf{y}$ is the $(n_R \times 1)$ (array response) receive vector.
- $\mathbf{H}$ is the $(n_R \times n_T)$ channel matrix.
- $\mathbf{n}$ is the $(n_R \times 1)$ additive white Gaussian noise (AWGN) vector.

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{1n_T} \\ h_{21} & \cdots & h_{2n_T} \\ \vdots & \ddots & \vdots \\ h_{n_R1} & \cdots & h_{n_Rn_T} \end{bmatrix}$$

Mutual information

- With $h_d(\cdot)$ denoting differential entropy (entropy of a continuous random variable), the mutual information may be expressed as

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= h_d(\mathbf{Y}) - h_d(\mathbf{Y}|\mathbf{X}) &(40)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{HX}+\mathbf{N}|\mathbf{X}) &(41)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{N}|\mathbf{X}) &(42)\\ &= h_d(\mathbf{Y}) - h_d(\mathbf{N}) &(43)\end{aligned}$$

- Assuming $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}^n)$.

- Since the normal distribution maxmizes the entropy over all distributions with the same covariance (i.e. the power constraint), the mutual information is maxmized when $\mathbf{Y}$ represents a multivariate Gaussian random variable.

Desired signal covariance matrix

- With the assumption that $\mathbf{X}$ and $\mathbf{N}$ are uncorrelated, the received covariance matrix $\mathbf{K}^y$ may be expressed as

$$\begin{aligned} \mathcal{E}\{\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\} &= \mathcal{E}\{(\mathbf{H}\mathbf{X}+\mathbf{N})(\mathbf{H}\mathbf{X}+\mathbf{N})^{\mathsf{H}}\} \quad (44) \\ &= \mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}}+\mathbf{K}^n \quad (45) \end{aligned}$$

where $\mathbf{K}^x = \mathcal{E}\{\mathbf{X}\mathbf{X}^{\mathsf{H}}\}$.

MIMO fading channel capacity

$$
\begin{aligned}
C &= h_d(\mathbf{Y}) - h_d(\mathbf{N}) && (46)\\
&= \log_2[\det(\pi e(\mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}} + \mathbf{K}^n))] - \log_2[\det(\pi e\mathbf{K}^n)] && (47)\\
&= \log_2[\det(\mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}} + \mathbf{K}^n)] - \log_2[\det \mathbf{K}^n] && (48)\\
&= \log_2[\det((\mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}} + \mathbf{K}^n)(\mathbf{K}^n)^{-1})] && (49)\\
&= \log_2[\det(\mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}}(\mathbf{K}^n)^{-1} + \mathbf{I}_{n_R})] && (50)\\
&= \log_2[\det(\mathbf{I}_{n_R} + (\mathbf{K}^n)^{-1}\mathbf{H}\mathbf{K}^x\mathbf{H}^{\mathsf{H}})] && (51)
\end{aligned}
$$

- When the transmitter has no knowledge of the channel, it is optimal to evenly distribute the available power P_T among the transmit antennas, i.e., $\mathbf{K}^x = \frac{P_T}{n_T}\mathbf{I}_{n_T}$.
- Assuming that the noise is uncorrelated between branches, the noise covariance matrix $\mathbf{K}^n = \sigma_n^2 \mathbf{I}_{n_R}$.
- The MIMO fading channel capacity can then be written as

$$
C = \log_2\left[\det\left(\mathbf{I}_{n_R} + \frac{P_T}{n_T\sigma_n^2}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)\right]. \qquad (52)
$$

- By the law of large numbers, the term $\frac{1}{n_T}\mathbf{H}\mathbf{H}^{\mathsf{H}} \Rightarrow \mathbf{I}_{n_R}$ as n_T gets large and n_R is fixed. Thus the capacity in the limit of large n_T is

$$C = n_R \cdot \underbrace{\log_2\left(1 + \frac{P_T}{\sigma_n^2}\right)}_{SISO\ capacity} \tag{53}$$

- Further analysis of the MIMO channel capacity is possible by diagonalizing the product matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ either by eigenvalue decomposition or singular value decomposition.

- Eigenvalue decomposition of the matrix product $\mathbf{H}\mathbf{H}^{\mathsf{H}} = \mathbf{E}\Lambda\mathbf{E}^{\mathsf{H}}$:

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma_n^2 n_T} \mathbf{E}\Lambda\mathbf{E}^{\mathsf{H}} \right) \right] \tag{54}$$

  where $\mathbf{E}$ is the eigenvector matrix with orthonormal columns and Λ is a diagonal matrix with the eigenvalues on the main diagonal.

- Singular value decomposition of the channel matrix $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{H}}$:

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma_n^2 n_T} \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^{\mathsf{H}}\mathbf{U}^{\mathsf{H}} \right) \right] \tag{55}$$

  where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices of left and right singular vectors respectively, and $\mathbf{\Sigma}$ is a diagonal matrix with singular values on the main diagonal.

- Using the singular value decomposition approach, the capacity can now be expressed as

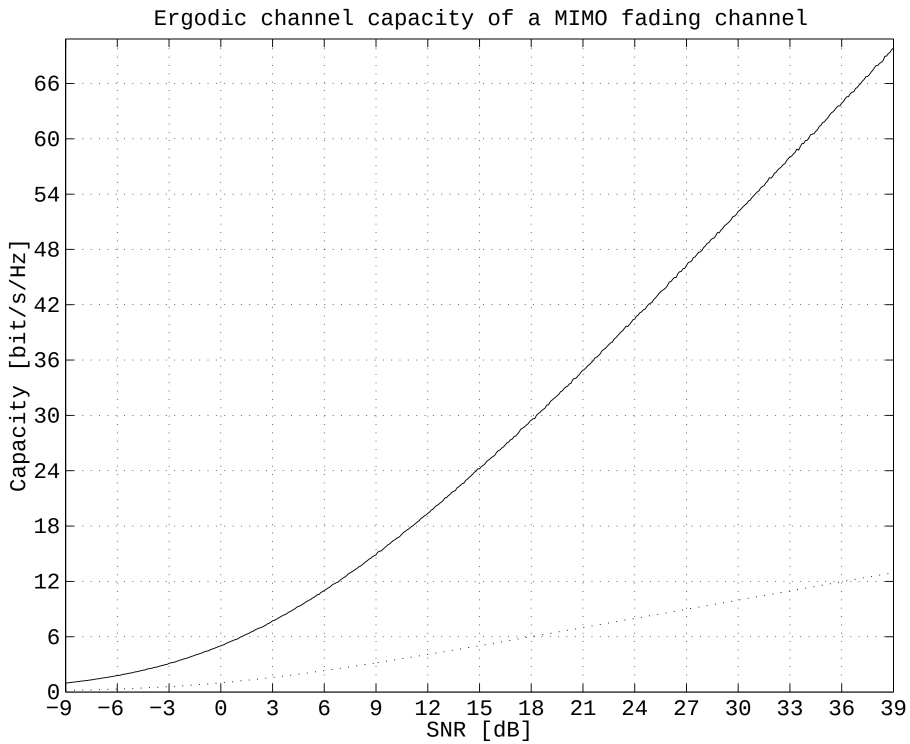
$$
\begin{aligned}
C &= \log_2\left[\det\left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma_n^2 n_T}\mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathsf{H}}\mathbf{U}^{\mathsf{H}}\right)\right] && (56)\\
&= \log_2\left[\det\left(\mathbf{I}_{n_T} + \frac{P_T}{\sigma_n^2 n_T}\mathbf{U}^{\mathsf{H}}\mathbf{U}\boldsymbol{\Sigma}^{2}\right)\right] && (57)\\
&= \log_2\left[\det\left(\mathbf{I}_{n_T} + \frac{P_T}{\sigma_n^2 n_T}\boldsymbol{\Sigma}^{2}\right)\right] && (58)\\
&= \log_2\left[\left(1 + \frac{P_T}{\sigma_n^2 n_T}\sigma_1^2\right)\left(1 + \frac{P_T}{\sigma_n^2 n_T}\sigma_2^2\right)\cdots\left(1 + \frac{P_T}{\sigma_n^2 n_T}\sigma_k^2\right)\right] && (59)\\
&= \sum_{i=1}^{k}\left(1 + \frac{P_T}{\sigma_n^2 n_T}\sigma_i^2\right) && (60)
\end{aligned}
$$

where $k = \text{rank}\{\mathbf{H}\} \leq \min\{n_T, n_R\}$, $\boldsymbol{\Sigma}$ is a real matrix, and $\det(\mathbf{I}_{AB} + \mathbf{AB}) = \det(\mathbf{I}_{BA} + \mathbf{BA})$

- Using the same approach with an eigenvalue decomposition of the matrix product $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, the capacity can also be expressed as

$$C = \sum_{i=1}^{k} \left(1 + \frac{P_T}{\sigma_n^2 n_T} \lambda_i\right) \tag{61}$$

where λ_i are the eigenvalues of the matrix Λ.



The Shannon capacity of a SISO channel (dotted line) compared to the ergodic capacity of a Rayleigh fading MIMO channel (solid line) with $n_T = n_R = 6$

3dB increase in SNR $\Rightarrow$ 6 bits/s/Hz capacity increase!

Transmit diversity

- Antenna diversity techniques are commonly utilized at the base stations due to less constraints on both antenna space and power. In addition, it is more economical to add more complex equipment to the base stations rather than at the remote units.

- To increase the quality of the transmission and reduce multipath fading at the remote unit, it would be beneficial if space diversity also could be utilized at the remote units.

- In 1998, S. M. Alamouti published a paper entitled "A simple transmit diversity technique for wireless communications". This paper showed that it was possible to generate the same diversity order traditionally obtained with SIMO system with a Multiple-Input Single-Output (MISO) system.

- The generalized transmission scheme introduced by Alamouti has later been known as Space-Time Block Codes (STBC).

Alamouti STBC

- With the Alamouti space-time code [Alamouti,1998], two consecutive symbols $\{s_0, s_1\}$ are mapped into a matrix codeword $\mathbf{S}$ according to the following mapping:

$$\mathbf{S} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix}, \qquad (62)$$

- The individual rows represent time diversity and the individual columns space (antenna) diversity.

- Assuming a block fading model, i.e., the channel remains constant for at least T channel uses, the received signal vector $\mathbf{x}$ (array response/per channel use) may be expressed as

$$\mathbf{x}_k = \mathbf{H}\mathbf{s}_k + \mathbf{n}_k,\ k = 1, \ldots, T. \qquad (63)$$

[Alamouti,1998] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE J. Select. Areas Comm., Vol.16, No.8, October 1998

$$\mathbf{x}_k = \mathbf{H}\mathbf{s}_k + \mathbf{n}_k,\ \ k = 1, \ldots, T.$$

- $\mathbf{x}_k \in \mathcal{C}^{n_R}$ denotes the received signal vector per channel use.
- $\mathbf{s}_k \in \mathcal{C}^{n_T}$ denotes the transmitted signal vector (a single row from the matrix codeword $\mathbf{S}$ transposed into a column vector).
- $\mathbf{H} \in \mathcal{C}^{n_R \times n_T}$ denotes the channel matrix with (possibly correlated) zero-mean complex Gaussian random variable entries.
- $\mathbf{n}_k \in \mathcal{C}^{n_R}$ denotes the additive white Gaussian noise where each entry of the vector is a zero-mean complex Gaussian random variable.

- For T consecutive uses of the channel, the received signal may be expressed as

$$\mathbf{X} = \mathbf{H}\mathbf{S} + \mathbf{N}, \tag{64}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T]$ (T consecutive array responses $\Rightarrow$ time responses in n_R branches), $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_T]$, and $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \cdots, \mathbf{n}_T]$.

- For notational simplicity [Hassibi,2001], the already introduced matrices $\mathbf{X}$, $\mathbf{S}$, and $\mathbf{N}$ may be redefined as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T]^\top$, $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_T]^\top$, and $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \cdots, \mathbf{n}_T]^\top$.

- With this new definition of the matrices $\mathbf{X}$, $\mathbf{S}$, and $\mathbf{N}$, time runs vertically and space runs horizontally and the received signal for T channel uses may now be expressed as

$$\mathbf{X} = \mathbf{S}\mathbf{H}^\top + \mathbf{N}. \tag{65}$$

[Hassibi,2001] B. Hassibi, B. M. Hochwald, "High-rate codes that are linear in space and time," 2001

- In [Hassibi,2001], the transpose notation on $\mathbf{H}$ is omitted and $\mathbf{H}$ is just redefined to have dimension $n_T \times n_R$.

- For a 2×2 MIMO channel, equation (65) becomes [Hassibi,2001]

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & -s_1 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} + \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix} \tag{66}$$

- x_{11} and x_{12} represent the received symbols at antenna element no.1 and 2 at time index t and likewise x_{21} and x_{22} represent the received symbols at antenna element no.1 and 2 at time index $t + T_s$

- This can be reorganized [Alamouti,1998] and written as

$$\underbrace{\begin{bmatrix} x_{11} \\ x_{21}^* \\ x_{12} \\ x_{22}^* \end{bmatrix}}_{\triangleq \mathbf{x}} = \underbrace{\begin{bmatrix} h_{11} & h_{21} \\ h_{21}^* & -h_{11}^* \\ h_{12} & h_{22} \\ h_{22}^* & -h_{12}^* \end{bmatrix}}_{\triangleq \mathcal{H}} \underbrace{\begin{bmatrix} s_1 \\ s_2 \end{bmatrix}}_{\triangleq \mathbf{s}} + \underbrace{\begin{bmatrix} n_{11} \\ n_{21}^* \\ n_{12} \\ n_{22}^* \end{bmatrix}}_{\triangleq \mathbf{n}} \tag{67}$$

- With matched filtering at the receiver (perfect channel knowledge):

$$\begin{aligned} \mathbf{y} &= \mathcal{H}^{\mathsf{H}}\mathbf{x} \\ &= \mathcal{H}^{\mathsf{H}}\mathcal{H}\mathbf{s} + \mathcal{H}^{\mathsf{H}}\mathbf{n} \\ &= ||\mathbf{H}||_F^2\mathbf{s} + \mathcal{H}^{\mathsf{H}}\mathbf{n}. \end{aligned} \tag{68}$$

where $||\mathbf{H}||_F^2$ represents the squared Frobenius norm of the matrix $\mathbf{H}$.

$$\begin{aligned} \mathcal{H}^{\mathsf{H}}\mathcal{H} &= \begin{bmatrix} h_{11}^* & h_{21} & h_{12}^* & h_{22} \\ h_{21}^* & -h_{11} & h_{22}^* & -h_{12} \end{bmatrix} \begin{bmatrix} h_{11} & h_{21} \\ h_{21}^* & -h_{11}^* \\ h_{12} & h_{22} \\ h_{22}^* & -h_{12}^* \end{bmatrix} \\ &= \begin{bmatrix} |h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2 & 0 \\ 0 & |h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2 \end{bmatrix} \\ &= ||\mathbf{H}||_F^2 \cdot \mathbf{I}_2. \end{aligned}$$

- This means that the received signals after matched filtering are de-coupled and they can be written individually as

$$y_1 = ||\mathbf{H}||_F^2 s_1 + \mathcal{H}^{\mathsf{H}}\mathbf{n} \qquad (69)$$

$$y_2 = ||\mathbf{H}||_F^2 s_2 + \mathcal{H}^{\mathsf{H}}\mathbf{n} \qquad (70)$$

- In general, the effective channel induced by space-time block coding of complex symbols (before detection) can be represented as [Sandhu,2000]

$$\mathbf{y}_k = ||\mathbf{H}||_F^2 \mathbf{s}_k + \mathcal{H}^{\mathsf{H}}\mathbf{n}. \qquad (71)$$

[Sandhu,2000] S. Sandhu, A. Paulraj, "Space-Time Block Codes: A Capacity Perspective," IEEE Comm. Letter, Vol.4, No.12, December 2000.

- The overall SNR before detection of each symbol is equal to

$$\gamma_t^{mimo} = \frac{||\mathbf{H}||_F^4 |s_k|^2}{\mathcal{E}\{|\mathcal{H}^{\mathsf{H}}\mathbf{n}|^2\}} = \frac{||\mathbf{H}||_F^4 \frac{P_T}{n_T}}{||\mathbf{H}||_F^2 \sigma_n^2} = P||\mathbf{H}||_F^2. \tag{72}$$

  where $P = \frac{P_T}{\sigma_n^2 n_T}$.

- For each transmitted symbol, the effective channel is a scaled AWGN channel with SNR$= P||\mathbf{H}||_F^2$.

- The capacity of a MIMO fading channel using STBC can then be written as

$$C = \frac{K}{T} \cdot \log_2 \left(1 + \frac{P_T}{\sigma_n^2 n_T} ||\mathbf{H}||_F^2 \right). \tag{73}$$

  where $\frac{K}{T}$ in front of the equation denotes the rate of the STBC.

- With the Alamouti STBC, two symbols ($K = 2$) are transmitted in two time slots ($T = 2$), i.e., the Alamouti code is a full rate STBC.

- Assuming uncorrelated channels and that all channel envelopes are i.i.d. Nakagami-m distributed random variables with equal average power $\mathcal{E}\{|h_{ij}|^2\} = \Omega$, the overall SNR may be expressed as a gamma distributed random variable:

$$\gamma_t^{mimo} = \frac{P_T}{n_T \sigma_n^2} \cdot ||\mathbf{H}||_F^2 \quad (74)$$

$$||\mathbf{H}||_F^2 \sim \mathcal{G}(n_T \cdot n_R, \Omega) \quad (75)$$

$$\gamma_t^{mimo} \sim \mathcal{G}(N \cdot m, \overline{\gamma}_l / m) \quad (76)$$

where $N = n_T \cdot n_R$ and $\overline{\gamma}_l = \frac{P_T \Omega}{\sigma_n^2 n_T}$.

- Effective diversity order $N_{div} = \frac{1}{\tau^2} = N \cdot m$.

Capacity summary

- Note that the capacity formulas given below are obtained with the assumption of an average power constraint P_T at the transmitter, uncorrelated equal noise power σ_n^2 in all branches, perfect channel knowledge at the receiver and no channel knowledge at the transmitter.

- **SISO**: $C = \log_2\left(1 + \frac{P_T}{\sigma_n^2}|h|^2\right)$.

- **SIMO**: $C = \log_2\left(1 + \frac{P_T}{\sigma_n^2}||\mathbf{h}||^2\right)$.

- **MIMO**: $C = \log_2\left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma_n^2 n_T}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)$.

- **MIMO with STBC**: $C = \log_2\left(1 + \frac{P_T}{\sigma_n^2 n_T}||\mathbf{H}||_F^2\right)$.

STBC - a capacity perspective

- STBC arec useful since they are able to provide full diversity over the coherent, flat-fading channel.
- In addition, they require simple encoding and decoding.
- Although STBC provide full diversity at a low computational cost, it can be shown that they incur a loss in capacity because they convert the matrix channel into a scalar AWGN channel whose capacity is smaller than the true channel capacity.

S. Sandu, A. Paulraj,"Sapce-time block codes: A capacity perspective," IEEE Communications Letters, Vol.4, No.12, December 2000.

$$
\begin{aligned}
C &= \log_2\left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma_n^2 n_T}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right) \\
&= \log_2 \prod_{i=1}^{k}\left(1 + \frac{P_T}{\sigma_n^2 n_T}\sigma_i^2\right) \\
&= \log_2\left(1 + P\sum_{i=1}^{k}\sigma_i^2 + P^2\sum_{i_1\neq i_2}^{i_1<i_2}\sigma_{i_1}^2\sigma_{i_2}^2 + P^3\sum_{i_1\neq i_2\neq i_3}^{i_1<i_2<i_3}\sigma_{i_1}^2\sigma_{i_2}^2\sigma_{i_3}^2 + \cdots + P^k\prod_{i=1}^{k}\sigma_i^2\right) \\
&= \log_2\left(1 + P||\mathbf{H}||_F^2 + P^2\sum_{i_1\neq i_2}^{i_1<i_2}\sigma_{i_1}^2\sigma_{i_2}^2 + P^3\sum_{i_1\neq i_2\neq i_3}^{i_1<i_2<i_3}\sigma_{i_1}^2\sigma_{i_2}^2\sigma_{i_3}^2 + \cdots + P^k\prod_{i=1}^{k}\sigma_i^2\right) \\
&\geq \log_2\left(1 + P||\mathbf{H}||_F^2\right) \\
&\geq \frac{K}{T}\cdot\log_2\left(1 + P||\mathbf{H}||_F^2\right)
\end{aligned}
$$

- The capacity difference is a function of the channel singular values. This can used to determine under which conditions STBC is optimal in terms of capacity.

- When the channel matrix is a rank one matrix, there is only a single non-zero singular value, i.e., a space-time block code is optimal (with respect to capacity) when it is rate one ($K = T$) **and** it is used over a channel of rank one [Sandhu,2000].

- For the i.i.d. Rayleigh channel with $n_R > 1$, the rank of the channel matrix is greater than one, thus a space- time block code of any rate used over the i.i.d. Rayleigh channel with multiple receive antennas always incurs a loss in capacity.

- A full rate space-time block code used over any channel with one receive antenna is always optimal with respect to capacity.

- Essentially, STBC trades off capacity benefits for low complexity encoding and decoding.

- Note that with spatial multiplexing, the simplification is opposite of STBC. It trades of diversity benefits for lower complexity.

Outage capacity

- Defined as the probability that the instantaneous capacity falls below a certain threshold or target capacity C_{th}

$$P_{out}(C_{th}) = \text{Prob}[C \leq C_{th}] = \int_0^{C_{th}} f_C(C)dC = P_C(C_{th}) \qquad (77)$$

SISO capacity

$$C = \log_2\left(1 + \frac{P_T}{\sigma_n^2} \cdot |h|^2\right) = \log_2\left(1 + \gamma_t^{siso}\right). \qquad (78)$$

- Assuming that $|h|$ is Nakagami-m distributed random variable,
- γ_t^{siso} is a Gamma distributed random variable with shape parameter $\alpha = m$ and scale parameter $\beta = \overline{\gamma}_l/m$.
- $\overline{\gamma}_l = \mathcal{E}\{\gamma_t^{siso}\} = \mathcal{E}\{\frac{P_T|h|^2}{\sigma_n^2}\} = \frac{P_T\Omega}{\sigma_n^2}$.
- $\mathcal{E}\{|h|^2\} = \Omega$.
- $\gamma_t^{siso} \sim \mathcal{G}(m, \overline{\gamma}_l/m)$.

Transformation of random variables

- Let X and Y be continuous random variables with $Y = g(X)$. Suppose g is one-to-one, and both g and its inverse function, g^{-1}, are continuously differentiable. Then

$$f_Y(y) = f_X[g^{-1}(y)] \left| \frac{dg^{-1}(y)}{dy} \right| . \tag{79}$$

- Let $C = g(\gamma_t^{siso}) = \log_2(1 + \gamma_t^{siso})$.

- Then $\gamma_t^{siso} = g^{-1}(C) = 2^C - 1$.

- Capacity PDF

$$f_C(C) = f_{\gamma_t^{siso}}(2^C - 1) \cdot 2^C \ln 2 = \frac{(2^C - 1)^{m-1} e^{-(2^C-1)/\beta}}{\beta^m \Gamma(m)} \cdot 2^C \ln 2 \tag{80}$$

- The SISO outage capacity can be obtained by solving the integral

$$\begin{aligned} P_{out}(C_{th}) &= \int_0^{C_{th}} \frac{(2^C - 1)^{m-1} e^{-(2^C-1)/\beta}}{\beta^m \Gamma(m)} \cdot 2^C \ln 2 \cdot dC && (81) \\ &= 1 - Q\left(m, \frac{(2^{C_{th}} - 1)m}{\overline{\gamma}_l}\right) && (82) \end{aligned}$$

- Q(·, ·) is the normalized complementary incomplete gamma function defined as

$$Q(a, b) = \frac{\Gamma(a, b)}{\Gamma(a)} \quad (83)$$

- $\Gamma(a, b) = \int_b^\infty e^{-t} t^{a-1} dt$.

SIMO capacity

$$C = \log_2\left(1 + \frac{P_T}{\sigma_n^2} \cdot ||\mathbf{h}||^2\right) = \log_2\left(1 + \gamma_t^{simo}\right). \quad (84)$$

- Assuming that every channel gain in the vector $\mathbf{h}$, $|h_l|$, is a Nakagami-m distributed random variable with the same m parameter.
- γ_t^{simo} is a Gamma distributed random variable with shape parameter $\alpha = n_R \cdot m$ and scale parameter $\beta = \overline{\gamma}_l/m$.
- $\gamma_t^{simo} \sim \mathcal{G}(n_R \cdot m, \overline{\gamma}_l/m)$.

Transformation of random variables

- Let $C = g(\gamma_t^{simo}) = \log_2(1 + \gamma_t^{simo})$.
- Then $\gamma_t^{simo} = g^{-1}(C) = 2^C - 1$.
- Capacity PDF

$$f_C(C) = f_{\gamma_t^{simo}}(2^C - 1) \cdot 2^C \ln 2 = \frac{(2^C - 1)^{n_R m - 1} e^{-(2^C - 1)/\beta}}{\beta^{n_R \cdot m} \Gamma(n_R \cdot m)} \cdot 2^C \ln 2 \quad (85)$$

- The SIMO outage capacity can be obtained by solving the integral

$$\begin{aligned} P_{out}(C_{th}) &= \int_0^{C_{th}} \frac{(2^C - 1)^{n_R m - 1} e^{-(2^C - 1)/\beta}}{\beta^{n_R \cdot m} \Gamma(n_R \cdot m)} \cdot 2^C \ln 2 \cdot dC & (86) \\ &= 1 - Q\left(n_R \cdot m, \frac{(2^{C_{th}} - 1)m}{\overline{\gamma}_l}\right) & (87) \end{aligned}$$

MIMO with STBC

$$C = \frac{K}{T}\log_2\left(1 + \frac{P_T}{\sigma_n^2}\cdot ||\mathbf{H}||_F^2\right) = \frac{K}{T}\log_2\left(1 + \gamma_t^{mimo}\right). \qquad (88)$$

- Assuming that every channel gain in the matrix $\mathbf{H}$, $|h_{ij}|$, is a Nakagami-m distributed random variable with the same m parameter.

- γ_t^{mimo} is a Gamma distributed random variable with shape parameter $\alpha = N \cdot m$ ($N = n_T \cdot n_R$) and scale parameter $\beta = \overline{\gamma}_l/(n_T m)$.

- $\gamma_t^{mimo} \sim \mathcal{G}(N \cdot m, \overline{\gamma}_l/(n_T m))$.

Transformation of random variables

- Let $C = g(\gamma_t^{mimo}) = \frac{K}{T}\log_2(1+\gamma_t^{mimo})$.
- Then $\gamma_t^{mimo} = g^{-1}(C) = 2^{(C\cdot T)/K} - 1$.
- Capacity PDF

$$f_C(C) = f_{\gamma_t^{mimo}}(2^{(C\cdot T)/K} - 1)\cdot 2^{(C\cdot T)/K}\frac{K}{T}\ln 2 \quad (89)$$

- The MIMO outage capacity can be obtained by solving the integral

$$\begin{aligned} P_{out}(C_{th}) &= \int_0^{C_{th}} \frac{(2^{(C\cdot T)/K}-1)^{Nm-1}e^{-(2^{(C\cdot T)/K}-1)/\beta}}{\beta^{N\cdot m}\Gamma(N\cdot m)}\cdot 2^{(C\cdot T)/K}\ln 2\cdot dC \\ &= 1 - Q\left(N\cdot m, \frac{(2^{(C_{th}\cdot T)/K}-1)m\cdot n_T}{\overline{\gamma}_l}\right) \end{aligned} \quad (90)$$

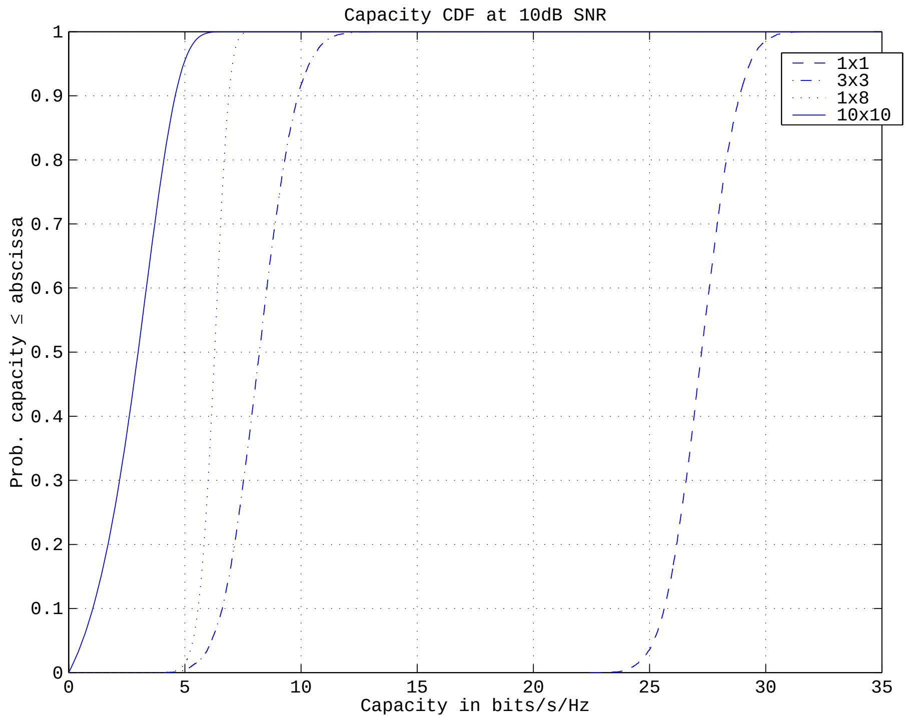
MIMO capacity

- Recall that $C = \sum_{i=1}^{k} \log_2\left(1 + \frac{P_T}{\sigma_n^2}\lambda_i\right)$.
- With the assumption that all eigenvalues are i.i.d random variables and $n_T = n_R$, the maximum capacity can be expressed as $C = n_T \cdot \log_2(1 + \frac{P_T}{\sigma_n^2}\lambda)$.
- Let $C = g(\lambda) = n_T \cdot \log_2(1 + \frac{P_T}{\sigma_n^2}\lambda)$.
- Then $\lambda = g^{-1}(C) = \frac{2^{C/n_T} - 1}{P_T/\sigma_n^2}$.
- Capacity PDF

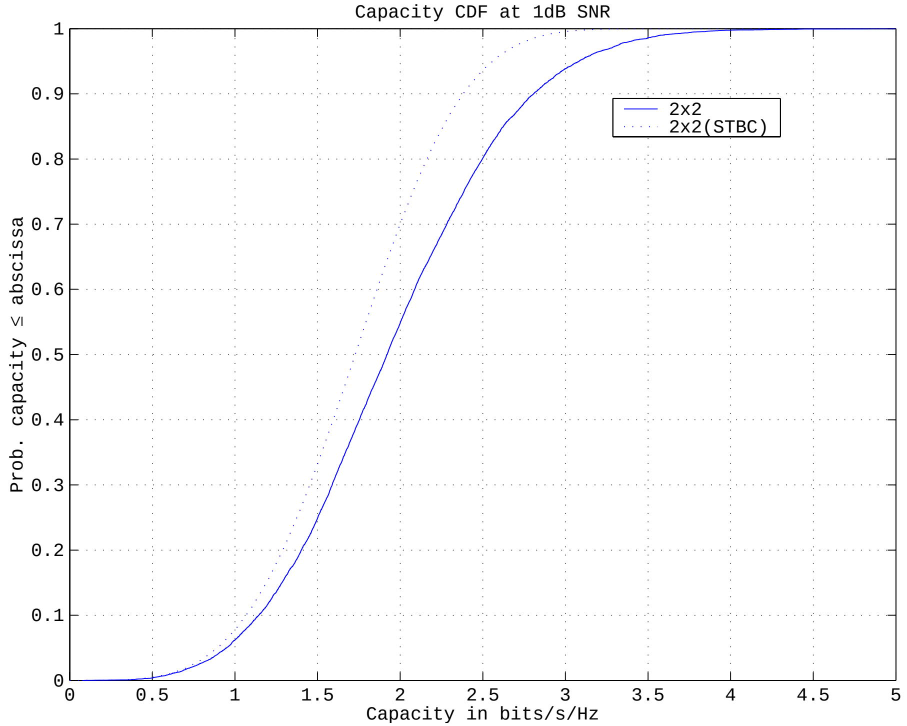
$$f_C(C) = f_\lambda\left(\frac{2^{C/n_T} - 1}{P_T/\sigma_n^2}\right) \cdot 2^{C/n_T} \frac{n_T \sigma_n^2}{P_T} \ln 2. \tag{91}$$

- Need to know the PDF of λ to obtain the capacity PDF.

Outage capacity



Outage capacity of i.i.d. Rayleigh fading channels at 10dB branch SNR



Outage capacity of a 2x2 MIMO Rayleigh fading channel using the Alamouti STBC at the transmitter at 1dB branch SNR

- The capacity formulas of SISO, SIMO and MIMO fading channels have been derived based on maximizing the mutual information between the transmitted and received signal.

- The Alamouti space-time block code has been presented. Although capable of increasing the diversity benefits, the use of STBC trades off capacity for low complexity encoding and decoding.

- By using transformation of random variables, closed-form expressions for the outage capacity for SISO, SIMO and MIMO (STBC at the transmitter) i.i.d. Nakagami-m fading channels were derived.