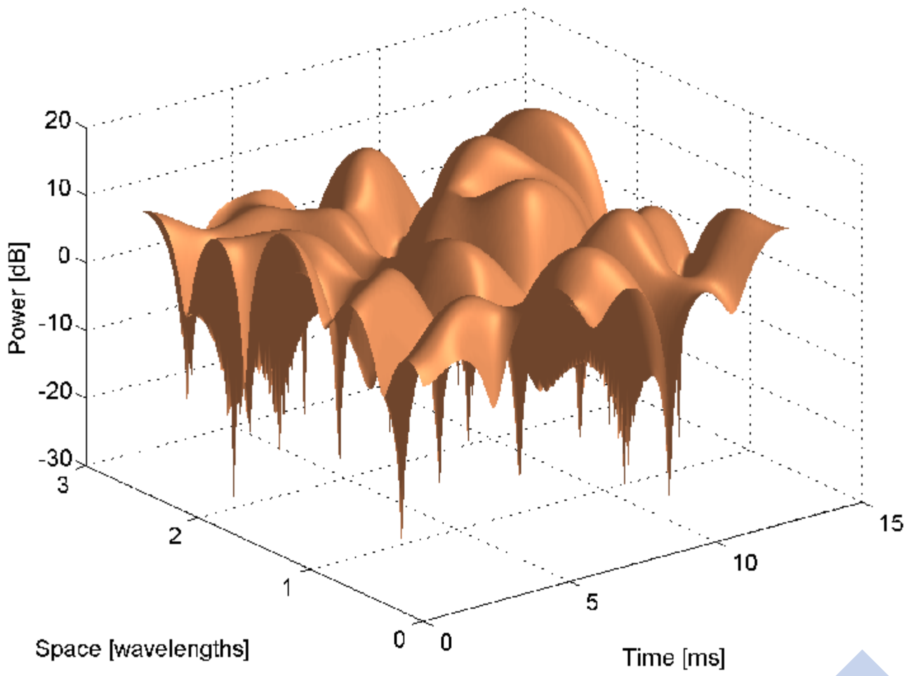
Capacity of Wireless MIMO OFDM Systems

Information theoretic aspects of 5G Systems

Wireless Systems

- 5G systems achieve unprecedented spectral efficiencies by harnessing the diversity offered by space. Antenna diversity at both Tx and Rx allows for the multiplexing of parallel data steams (in the same frequency and time slot)
- In a Multiple-Input Multiple-Output (MIMO) system, the data stream from a single user is sent via n_T separate sub-streams. The number n_T equals the number of transmit antennas.

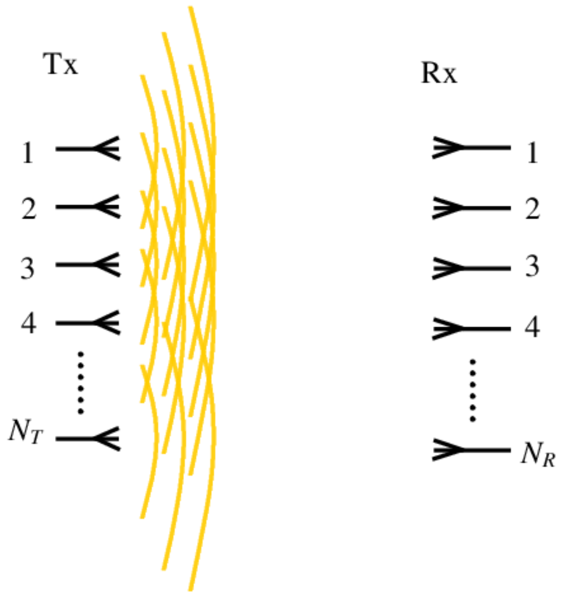


MIMO System Model

The input/output relations of a narrowband, single-user MIMO link is represented by the complex baseband vector model:

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

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



MIMO Channel Model

- The channel matrix $\mathbf{H}$ is random. The receiver has a perfect channel knowledge (e.g. obtained via "training" process).
- The channel is memoryless, i.e., for each use of the channel an independent realisation of $\mathbf{H}$ is drawn.

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

$$h_{ij} = \alpha + \jmath\beta = |h_{ij}| \cdot e^{\jmath\phi_{ij}}$$

In a rich scattering environment without line-of-sight (LOS), individual channels are complex circular Gaussian $h_{ij} \sim \mathcal{CN}(0, \sigma_h^2)$, meaning α and β are iid $\sim \mathcal{N}\left(0, \frac{\sigma_h^2}{2}\right)$, hence the channel gains $|h_{ij}|$ are Rayleigh distributed.

SISO Channel Capacity

$$y = h_{11}x + n$$

The ergodic (mean) capacity of a random channel with $(n_R = n_T = 1)$ and an average transmit power constraint P_T can be expressed as:

$$C = E_H \left\{ \max_{p(x):P \leq P_T} I(X;Y) \right\}$$

Where P is the average power of a single channel codeword transmitted over the channel, and E_H denotes the expectation over all channel realisations. We recall that the Mutual Information is given by :

$$I(X;Y) = H(Y) - H(Y|X)$$

SISO Channel Capacity

The ergodic (mean) capacity of a random complex channel h_{11} and an average transmit power constraint P_T is given by:

$$C = E_H \left\{ \log_2 \left(1 + \rho \cdot |h_{11}|^2 \right) \right\}$$

Where ρ is the average signal-to-noise (SNR) ratio at the receiver antenna. If $|h_{11}|$ is Rayleigh, $|h_{11}|^2$ follows a chi-squared distribution χ_2^2 with two degrees of freedom (exponential distribution).

MIMO Capacity

We follow the same strategy which we used for the capacity derivation of the (scalar and real) Gaussian channel. However, we are dealing now with complex random vectors.

- Complex circular *(n-dim)* Gaussian vector $\mathbf{v} \in \mathbf{C}^n$, denoted as $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{K})$, is zero mean with covariance[1] $E\left\{\mathbf{v}\mathbf{v}^\dagger\right\} = \mathbf{K}$, and has pdf given by $f_\mathbf{v}(\mathbf{v}) = \frac{1}{\pi^n \det(\mathbf{K})} e^{-\mathbf{v}^\dagger \mathbf{K}^{-1} \mathbf{v}}$
- Its differential entropy equals to

$$h(\mathbf{v}) = \log_2(\pi e \det(\mathbf{K}))$$

  For non-circular complex Gaussian vectors the entropy is smaller.

- The differential entropy of a real Gaussian vector $\mathbf{x} \in \mathbf{R}^n$; $f_\mathbf{x}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\mathbf{m})}$ is given by:

$$\frac{1}{2} \log_2(2\pi e \det(\boldsymbol{\Sigma}))$$

[1] $\dagger$ - denotes complex conjugate transpose aka Hermitian transpose

MIMO Capacity

The ergodic (mean) capacity of a MIMO channel with an average transmit power constraint P_T is given by:

$$C = E_H \left\{ \max_{p(x):tr(\mathbf{\Phi}) \leq P_T} I(X;Y) \right\}$$

Where $\mathbf{\Phi} = E\left\{\mathbf{x}\mathbf{x}^\dagger\right\}$ is the covariance matrix of the transmit signal vector $\mathbf{x}$. The total transmit power is limited to P_T across all antennas.

The mutual information can be expanded as:

$$\begin{aligned} I(\mathbf{x};\mathbf{y}) &= h(\mathbf{y}) - h(\mathbf{y} \mid \mathbf{x}) \\ &= h(\mathbf{y}) - h(\mathbf{H}\mathbf{x} + \mathbf{n} \mid \mathbf{x}) \\ &= h(\mathbf{y}) - h(\mathbf{n} \mid \mathbf{x}) \\ &= h(\mathbf{y}) - h(\mathbf{n}) \end{aligned}$$

MIMO Capacity

Assuming the optimal Gaussian distribution for the transmit vector $\mathbf{x}$, the covariance matrix of the received complex vector $\mathbf{y}$ is given by:

$$\begin{aligned} E\left\{\mathbf{y}\mathbf{y}^\dagger\right\} &= E\left\{(\mathbf{H}\mathbf{x}+\mathbf{n})(\mathbf{H}\mathbf{x}+\mathbf{n})^\dagger\right\} \\ &= E\left\{\mathbf{H}\mathbf{x}\mathbf{x}^\dagger\mathbf{H}^\dagger\right\} + E\left\{\mathbf{n}\mathbf{n}^\dagger\right\} \\ &= \mathbf{H}\mathbf{\Phi}\mathbf{H}^\dagger + \mathbf{K}^n \\ &= \mathbf{K}^d + \mathbf{K}^n \end{aligned}$$

The superscripts d and n denote, respectively, the desired part and the noise part of the covariance matrix.

MIMO Capacity

$$h(\mathbf{v}) = \log_2(\pi e \det(\mathbf{K}))$$

The maximum mutual information of a random MIMO channel is then given by[2]:

$$\begin{aligned}
I &= h(\mathbf{y}) - h(\mathbf{n}) \\
&= \log_2\left[\det\left(\pi e\left(\mathbf{K}^d + \mathbf{K}^n\right)\right)\right] - \log_2\left[\det\left(\pi e \mathbf{K}^n\right)\right] \\
&= \log_2\left[\det\left(\mathbf{K}^d + \mathbf{K}^n\right)\right] - \log_2\left[\det\left(\mathbf{K}^n\right)\right] \\
&= \log_2\left[\det\left(\left(\mathbf{K}^d + \mathbf{K}^n\right)\left(\mathbf{K}^n\right)^{-1}\right)\right] \\
&= \log_2\left[\det\left(\mathbf{K}^d\left(\mathbf{K}^n\right)^{-1} + \mathbf{I}_{n_R}\right)\right] \\
&= \log_2\left[\det\left(\mathbf{H}\mathbf{\Phi}\mathbf{H}^\dagger\left(\mathbf{K}^n\right)^{-1} + \mathbf{I}_{n_R}\right)\right]
\end{aligned}$$

---

[2] $\det\left(\mathbf{A}^{-1}\right) = \left(\det\left(\mathbf{A}\right)\right)^{-1}$; $\det\left(\mathbf{AB}\right) = \det\left(\mathbf{A}\right) \cdot \det\left(\mathbf{B}\right)$

MIMO Capacity

- When the transmitter has no knowledge about the channel, a uniform power distribution is used. The transmit covariance matrix is then given by $\mathbf{\Phi} = \frac{P_T}{n_T}\mathbf{I}_{n_T}$.
- Moreover, the uncorrelated noise as seen at each receiver is described by the covariance matrix $\mathbf{K}^n = \sigma^2 \mathbf{I}_{n_R}$

The capacity for "fixed" $\mathbf{H}$ is then:

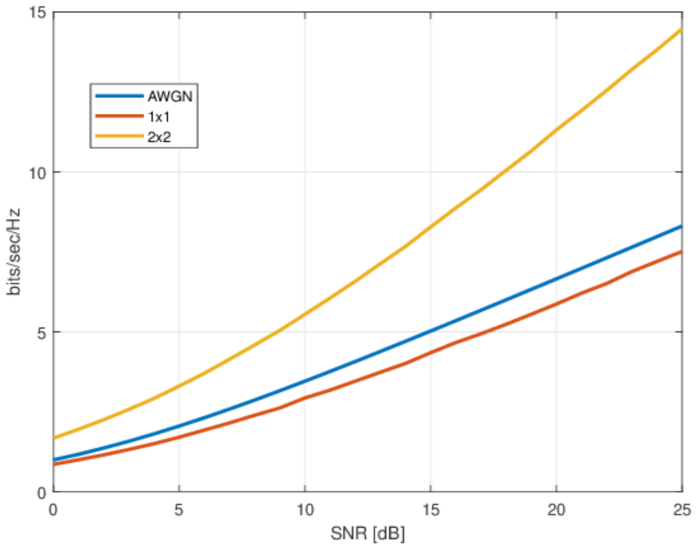
$$C = \log_2\left[\det\left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T}\mathbf{H}\mathbf{H}^\dagger\right)\right]$$

Where the average signal-to-noise (SNR) is $\frac{P_T}{\sigma^2} = \rho$. The ergodic capacity of MIMO channel becomes:

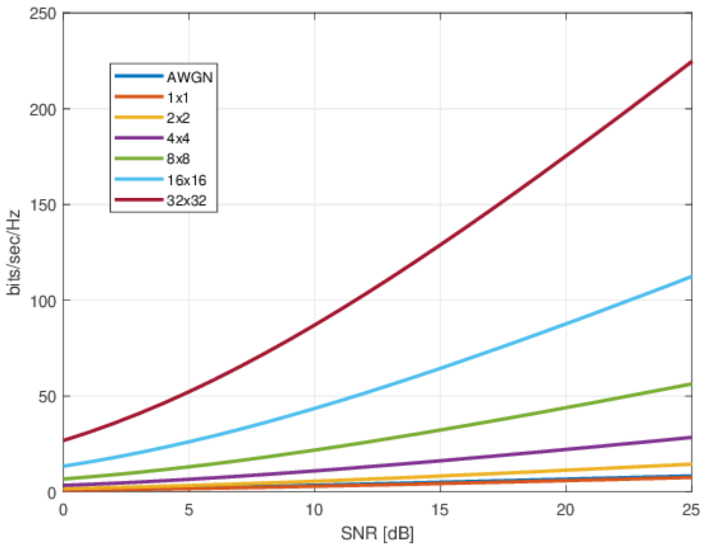
$$C = E_H\left\{\log_2\left[\det\left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T}\mathbf{H}\mathbf{H}^\dagger\right)\right]\right\}$$

MIMO Capacity plots

MIMO channel ergodic capacity plots for
$(n_R = n_T = 1, 2, 4, 8, 16, 32)$ and the AWGN channel.



(a)



(b)

Massive MIMO

Future wireless systems (6G) will deploy arrays with a very large number of transmit antennas. By the law of large numbers, the term $\frac{1}{n_T}\mathbf{H}\mathbf{H}^\dagger \to \mathbf{I}_{n_R}$ as n_T gets large, and n_R is fixed.
Thus the capacity in the limit of large arrays is:

$$C = E_H \left\{ \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T} \mathbf{H}\mathbf{H}^\dagger \right) \right] \right\} =$$
$$E_H \{ \log_2 [\det (\mathbf{I}_{n_R} + \rho \mathbf{I}_{n_R})] \} = E_H \{ \log_2 (1+\rho)^{n_R} \} =$$
$$\mathrm{E}_H \{ n_R \cdot \log_2 (1+\rho) \} = n_R \cdot \log_2(1+\rho)$$

The capacity grows linearly with n_R and there is no fading!!!

Interpretation using SVD decomposition

We can gain further insights into the MIMO channel capacity with a help of SVD. To do so, $\mathbf{H}$ is decomposed using Singular Value Decomposition (SVD). $\mathbf{H} = \mathbf{U\Sigma V}^\dagger$
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 $\{\sigma_1, \sigma_2, \cdots, \sigma_k\}$ on the main diagonal, $k = rank(\mathbf{H}) \leq \min(n_T, n_R)$. Then:

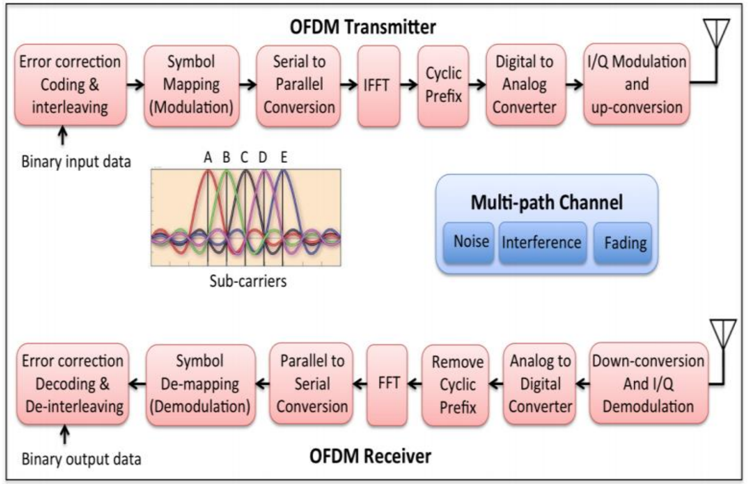
$$\mathbf{HH}^\dagger = \mathbf{U\Sigma V}^\dagger \left(\mathbf{U\Sigma V}^\dagger\right)^\dagger = \mathbf{U\Sigma\Sigma}^\dagger\mathbf{U}^\dagger$$

Interpretation using SVD decomposition

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

The MIMO channel is equivalent to a set of k parallel (non-interfering) Gaussian channels. The capacity is the sum of the k Gaussian channel capacities. (in line 2 we used $\det(\mathbf{I}_{AB} + \mathbf{AB}) = \det(\mathbf{I}_{BA} + \mathbf{BA})$)

OFDM Systems - multiplexing in the frequency domain

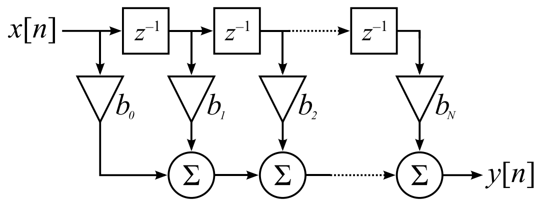


- OFDM systems use Cyclic Prefix and FFT/IFFT matrices $\mathbf{F}, \mathbf{F}^{\dagger}$ (transforms) to overcome channel dispersion and Inter block interference (IBI).
- It can be shown that OFDM can be modelled by an equivalent linear model[2]:

$$\mathbf{z}_i = \mathbf{\Lambda} \mathbf{u}_i + \tilde{\eta}_i$$

$\mathbf{\Lambda} = \mathbf{F}^{\dagger} \widehat{\mathbf{H}} \mathbf{F}$ represents the *eigenvalue decomposition* of $\widehat{\mathbf{H}}$ *circulant* channel matrix, and hence, $\mathbf{\Lambda}$ is a diagonal matrix.

- Therefore, OFDM system is also equivalent to a set of K parallel (non-interfering) Gaussian channels.
- MIMO-OFDM can deliver/multiplex $K \times k$ parallel data streams!

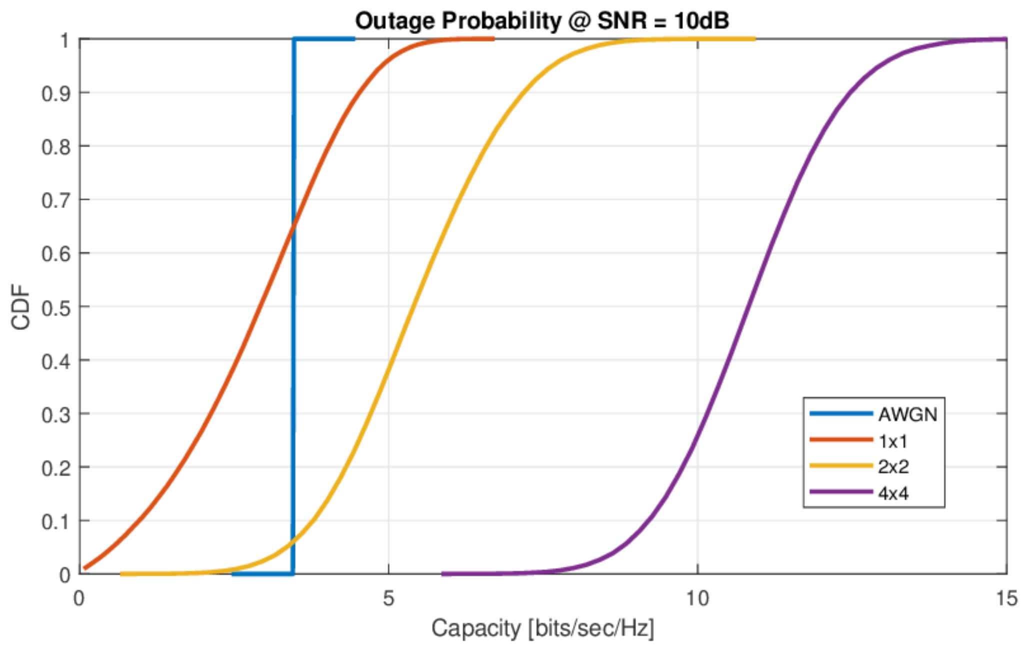


$$\widehat{\mathbf{H}} = \begin{bmatrix} h(0) & 0 & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & h(0) & \cdots & \vdots & 0 & h(L) & \vdots \\ h(L) & \vdots & \ddots & 0 & \vdots & 0 & h(L) \\ 0 & h(L) & \ddots & h(0) & 0 & \vdots & 0 \\ \vdots & 0 & \ddots & \vdots & h(0) & 0 & \vdots \\ 0 & \vdots & \cdots & h(L) & \vdots & h(0) & 0 \\ 0 & 0 & \cdots & 0 & h(L) & \vdots & h(0) \end{bmatrix}$$

[2] the proof (and complete model) is somewhat technical and is on BB

Outage Capacity

- The notion of MIMO capacity we discussed is not in a true sense the "Shannon Capacity". The MIMO channel capacity is a random variable and there is a non-zero P_e at all rates.
- We gain a useful insight by defining *Outage Capacity*, $Pr\{C \leq C_{out}\}$, i.e. the probability that the channel will not offer the capacity (at a given SNR).



- AWGN channel has a true sense "Shannon Capacity", $C_{AWGN} = 3.46 bits/sec/Hz$ at SNR $= 10$dB.

Concluding remarks

- Constrained Mutual Information. In all above calculations we have implicitly assumed Gaussian signalling codebooks. Practical systems use digital modulations, which will constrain the input entropy. Constrained mutual information does not have a closed form expression. However, we know that $C \leq H(X)$.
- We may have access to the channel knowledge at the transmitter (feedback, TDD). We can then optimise wrt power distribution $\mathbf{\Phi}$, which will maximise the mutual information (this is called "water-filling" solution).
- In all calculations we have used *discrete time* model, and hence, the capacity unit was *bits per channel use*. Using Nyquist sampling theorem, we know that each 1Hz of bandwidth can "carry" 2 real samples (or 1 complex) per sec. Therefore, all above capacity results also apply to continuous band-limited channels with the unit "bits/sec/Hertz".

Acknowledgments and further reading

- Gaussian Channels and Parallel Gaussian Channels: "Elements of Information Theory" Thomas M. Cover, Joy A. Thomas.
- Wireless Channel, Ergodic capacity, Outage capacity, "Wireless Communications" by Andrea Goldsmith.
- Several sections are based on "On the Capacity of the MIMO Channel - a Tutorial Introduction" by Bengt Holter (online, available on BB).