Parallel Additive Gaussian Channels

Let us assume that we have $N$ parallel one-dimensional channels disturbed by noise sources with variances $\sigma_1^2, \ldots, \sigma_N^2$.

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

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

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

Energy Constraint: The total input energy is constrained on average to $E$ the total average energy per channel use:

$$\sum_{n=1}^{N} \bar{x}_n^2 = \sum_{n=1}^{N} E_n = E$$

The capacity of these parallel channels is achieved by the following input energy distribution:

$$\sigma_n^2 + E_n = \mu; \quad \sigma_n^2 < \mu$$

$$E_n = 0; \quad \sigma_n^2 \geq \mu$$

where $\mu$ is the Lagrange multiplier chosen such that $\sum_n E_n = E$.

Then the capacity of this set of parallel channels is given by the following Theorem:

$$C = \sum_{n=1}^{N} \frac{1}{2} \log \left( 1 + \frac{E_n}{\sigma_n^2} \right) = \sum_{n=1}^{N} \frac{1}{2} \log \left( \frac{\mu}{\sigma_n^2} \right)$$

Source: [1], Section 7.5, pp. 343 ff.
**Waterfilling Theorem:**

*Proof:* Let \( x = [x_1, \cdots, x_N] \) and \( y = [y_1, \cdots, y_N] \) and consider

\[
I(x; y) \overset{(1)}{\leq} \sum_{n=1}^{N} I(x_n; y_n) \quad \text{(1) independent } x_n
\]

\[
\overset{(2)}{\leq} \sum_{n=1}^{N} \frac{1}{2} \log \left( 1 + \frac{E_n}{\sigma_n^2} \right) \quad \text{(2) Gaussian distributed } x_n
\]

Since equality can be achieved in both inequalities above the next step is to find the maximizing energy distribution \( E = [E_1, \cdots, E_N] \). We make use of Theorem 1 and identify the following sufficient and necessary conditions:

\[
\frac{\partial f(E)}{\partial E_n} \leq \lambda
\]

\[
\frac{1}{2(\sigma_n^2 + E_n)} \leq \lambda
\]

\[
\sigma_n^2 + E_n \leq \frac{1}{2\lambda} = \mu
\]

This theorem is called the *Waterfilling Theorem*. Its functioning can be visualized in the following figure, where the power levels are in black:
Correlated Parallel Channels (MIMO)

Correlated channels arise from e.g., multiple antenna channels using $N_t$ transmit antennas and $N_r$ receive antennas:

This channel is a **multiple-input multiple-output** (MIMO) channel described by the matrix equation:

$$y = Hx + n$$

- The transmitted signals $x_n$ are complex signals, as are the channel gains $h_{ij}$ and the received signals $y_n$.
- The noise is complex additive Gaussian noise with variance $N_0$ (that is $N_0/2$ in each dimension).
- The path gains $h_{ij}$ are complex gain coefficients modeling a random phase shift and a channel gain. Often these are modeled as *Rayleigh random variables* modeling a scattering-rich or mobile radio transmission environment.

**MIMO Rayleigh Channel:** The $h_{ij}$ are modeled as i.i.d. (or correlated) complex Gaussian random variables with variance $1/2$ in each dimension.
Channel Decomposition

The correlated MIMO channel can be decomposed via the singular value decomposition (SVD):

\[ H = U D V^+ \]

(If \( r < t \))

where \( U \) and \( V \) are unitary matrices, i.e., \( U U^+ = I \), and \( V V^+ = I \). The matrix \( D \) contains the singular values of \( H \), which are the square roots of the eigenvalues of \( HH^+ \) and \( H^+ H \). If \( H \) is square, the singular values are identical to the square of the eigenvalues of \( H \).

The channel equation can now be written in an equivalent form:

\[ y = H x + n = U D V^+ x + n \]

\[ U^+ y = \tilde{y} = D \tilde{x} + \tilde{n} \]

If \( N_t > N_r \) only the first \( N_r \) signals of \( \tilde{x} \) will be received.

If \( N_r > N_t \) the \( N_r - N_t \) "bottom" channels will carry no signal.

This leads to parallel Gaussian channels \( \tilde{y}_n = d_n \tilde{x}_n + \tilde{n}_n \)
The multiplicative factors $d_n$ can be eliminated by multiplying the received signal $y$ with $D^{-1}$. This leads back exactly to the parallel channel problem, and the capacity of the MIMO channel is determined by the waterfilling theorem:

$$C = \sum_{n=1}^{N} \log \left( 1 + \frac{d_n^2 E_n}{2\sigma_n^2} \right) = \sum_{n=1}^{N} \log \left( \frac{d_n^2 \mu}{2\sigma_n^2} \right)$$

Note: The channels are complex, and hence there is no factor $1/2$ and the variance is $2\sigma_n^2$.

This capacity is achieved with the waterfilling power allocation:

$$\frac{2\sigma_n^2}{d_n^2} + E_n = \mu; \quad \sigma_n^2 < \mu$$
$$E_n = 0; \quad \sigma_n^2 \geq \mu$$

**Optimal System:** These considerations lead to the following *optimal* signalling strategy:

1. Perform the SVD of the channel $H \rightarrow U, V, D$.
2. Multiply the input signal vector $\tilde{x}$ with $V$. This is matrix processing.
3. Multiply the output signal $y$ with the matrix processor $U^+$.
4. Use each channel with signal-to-noise ratio $d_n^2 E_n/(2\sigma_n^2)$ independently.
**Symmetric MIMO Capacity**

**Drawback:** the channel $H$ needs to be known at both the transmitter and the receiver so the SVD can be computed.

**Fact:** Channel knowledge is not typically available at the transmitter, and the only choice we have is to distribute the energy uniformly over all component channels. This leads to the **Symmetric Capacity**:

$$C = \sum_{n=1}^{N} \log \left( 1 + \frac{d_n^2 E}{2N_t \sigma^2} \right) = \log \prod_{n=1}^{N} \left( 1 + \frac{d_n^2 E}{2N_t \sigma^2} \right)$$

Noting that the $d_n^2$ are the eigenvalues of $HH^+$, the above formula can be written in terms of matrix eigenvalues, using the fact $\det(M) = \prod \lambda(M)$, and $\det(I + M) = \prod(1 + \lambda(M))$:

$$C = \log \prod_{n=1}^{N} \left( 1 + \frac{d_n^2 E}{2N_t \sigma^2} \right) = \log \det \left( I_{N_r} + \frac{E}{2N_t \sigma^2} HH^+ \right)$$

$$\quad = \log \det \left( I_{N_r} + \frac{E}{2N_t \sigma^2} H^+ H \right)$$

**Discussion:**

The capacity of a MIMO channel is governed by the singular values of $H$, or in the symmetrical case by its eigenvalues. The eigenvalues determine the channel gains of the independent parallel channels.

- Channel $H$ needs to be known at the receiver → Channel Estimation

We will study the behavior of channel eigenvalues later in this course.
MIMO Fading Channels

Assumed that the channel $H$ is known at the receiver

$$I(x; (y, H)) = I(x; H) + I(x; y|H)$$
$$= I(x; y|H)$$
$$= E_H [I(x; y|H = H_0)]$$

We need to average the mutual information over all channel realizations.

$I(x; y|H)$ is maximized if $x$ is **circularly symmetric complex Gaussian** with covariance $Q$, and

$$I(x; (y, H)) = E_H \left[ \log \det \left( I_r + \frac{E}{2N_t\sigma^2}HQH^+ \right) \right]$$

$$I(x; (y, H)) = E_H \left[ \log \det \left( I_r + \frac{E}{2N_t\sigma^2}(HU)D(U^+H^+) \right) \right]$$

- The spectral decomposition of $Q = UD U^+$ produces an equivalent channel $\tilde{H} = HU$ with the same statistics, hence the maximizing $Q$ is diagonal.

- Furthermore, concavity of the function $\log \det()$ shows that $Q = I$, hence the maximizing $Q$ is a multiple of the identity

**Capacity of the MIMO Rayleigh Channel:**

$$C = E_H \left[ \log \det \left( I_r + \frac{E}{2N_t\sigma^2}HH^+ \right) \right]$$

- By the law of large numbers:

$$HH^+ \xrightarrow{N_t \to \infty} N_t I_r \quad \text{and} \quad C = r \log \left( 1 + \frac{E}{2\sigma^2} \right)$$
Evaluation of the Capacity Formula

Following Telatar [2] define the random matrix

\[ W = \begin{cases} \mathbf{H} \mathbf{H}^+ & \text{if } N_r < N_t \\ \mathbf{H}^+ \mathbf{H} & \text{if } N_r \geq N_t \end{cases} \]

\( W \) is an \( m \times m; m = \min(r, t) \) non-negative definite matrix with real, non-negative eigenvalues \( \mu_n = d_n^2 \)

The capacity can be written in terms of these eigenvalues:

\[ C = E_{\{\mu_n\}} \left[ \sum_{n=1}^{m} \log \left( 1 + \frac{E}{2N_t\sigma^2} \mu_n \right) \right] \]

for \( r = t \) symmetric channels, let \( \{\lambda_n\} \) be the eigenvalues of \( \mathbf{H} \), then \( \mu_n = \lambda_n^2 \)

- The eigenvalues of \( W \) follow a Wishart distribution:

\[ p(\mu_1, \ldots, \mu_m) = \alpha \prod_{i=1}^{m} e^{-\mu_i} \mu_i^{n-m} \prod_{i<j} (\mu_i - \mu_j)^2 \]

\[ p(\mu_i, \ldots, \mu_m) = \alpha \prod_{i=1}^{m} e^{-\mu_i} \mu_i^{n-m} \det \begin{bmatrix} 1 & \cdots & 1 \\ \mu_1 & \cdots & \mu_m \\ \vdots & \ddots & \vdots \\ \mu_1^{m-1} & \cdots & \mu_m^{m-1} \end{bmatrix} \]

- \( C \) therefore depends only on the distribution of a single eigenvalue:

\[ C = mE_{\mu} \left[ \log \left( 1 + \frac{\mu E}{2N_t\sigma^2} \right) \right] = \int_{0}^{\infty} mp(\mu) \log \left( 1 + \frac{\mu E}{2N_t\sigma^2} \right) d\mu \]

This integration evaluates to

\[ C = \int_{0}^{\infty} \log \left( 1 + \frac{\mu E}{2N_t\sigma^2} \right) \sum_{k=1}^{m-1} \frac{k!}{(k+n-m)!} \left[ L_{n-m}^{m-1}(\mu) \right]^2 \mu^{n-m} e^{-\mu} d\mu \]

- \( L_{n-m}^{m-1}(x) = \frac{1}{n!} e^x \frac{d^k}{dx^k} (e^{-x} x^{n-m+k}) \) is the Laguerre polynomial of order \( k \)
As the number of antennas $N_t \to \infty$, $N_r \to \infty$, the number of eigenvalues $N(\mu) \to \infty$, and the capacity formula

$$C = E_{\{\mu_i\}} \left[ \sum_{i=1}^{m} \log \left( 1 + \frac{\mu E}{2N_t\sigma^2} \right) \right] \to \int_{0}^{\infty} m \log \left( 1 + \frac{\mu Em}{2N_t\sigma^2} \right) dF_{\mu}(\mu)$$

where $F_{\mu}(\mu)$ is the cumulative distribution (CDF) of the eigenvalues of $W$, which becomes continuous as $m, n \to \infty$

For random matrices like $W$, a general result states

$$\frac{dF_{\mu}(\mu)}{d\mu} \to \begin{cases} \frac{1}{2\pi} \sqrt{\left( \frac{\mu_+}{\mu} - 1 \right) \left( 1 - \frac{\mu_-}{\mu} \right)}; & \text{for } \mu \in [\mu_- , \mu_+] \\ 0; & \text{otherwise} \end{cases}$$

• \( n/m \to \tau \geq 1 \), and \( \mu_\pm = (\sqrt{\tau} \pm 1)^2 \).

In the limit, the capacity of the Rayleigh MIMO is given by

$$\frac{C}{m} = \frac{1}{2\pi} \int_{d_-}^{d_+} \log \left( 1 + \frac{\mu Em}{2N_t\sigma^2} \right) \sqrt{\left( \frac{\mu_+}{\mu} - 1 \right) \left( 1 - \frac{\mu_-}{\mu} \right)} d\mu$$

References
