Bit- and power-loading

EIT 140, tom<AT>eit.lth.se
Model of multicarrier system

\[ Z_{1} \sim \mathcal{N}(0, \sigma_{Z1}^{2} \mathbf{I}) \]
\[ Z_{2} \sim \mathcal{N}(0, \sigma_{Z2}^{2} \mathbf{I}) \]
\[ Z_{N} \sim \mathcal{N}(0, \sigma_{ZN}^{2} \mathbf{I}) \]

- Includes
  - IDFT
  - CP extension
  - channel
  - CP removal
  - DFT
- For large \( N \): parallel independent flat complex-valued subchannels
Spectral allocation of information and power

- Straightforward transmission over subchannels yields a lousy performance since the worst subchannel, \textit{i.e.}, the subchannel with the highest probability of error, determines the performance of the whole system.

- Single-carrier system: the information is 'spread' over the occupied bandwidth; equaliser at the receiver 'fishes out' the portions with high signal-to-noise ratio (SNR)

- Multicarrier system: we have the opportunity to spectrally allocate power and information—however, we must do it well in order to achieve good performance on \textit{all} subchannels by one or both of the following means:
  - ensure reliability of transmission through \textit{coding}
  - adjust the amount of information transmitted over each subchannel according to quality of the subchannels through \textit{bit-/power-loading}
A single subchannel is memoryless, discrete-time, real-valued with a real-valued scalar $H$ and additive, real-valued, white, Gaussian noise (AWGN) of zero-mean and variance $\sigma_Z^2$:

$$Y = H \cdot X + Z, \quad H \text{ real-valued, } \ Z \sim \mathcal{N}(0, \sigma_Z^2)$$

$$Z \sim \mathcal{N}(0, \sigma_Z^2)$$
Capacity of real-valued subchannel

Differential entropy of the noise:

\[ h(Z) = h(Y|X) = \frac{1}{2} \log(2\pi e \sigma_Z^2) \]

The channel input is a continuous random variable \( X \), with finite power:

\[ \mathbb{E}(X^2) \leq P_X \]

The capacity of this channel is

\[
C = \max_{f_X(x) : \mathbb{E}(X^2) \leq P_X} I(X;Y) \\
= \max_{f_X(x) : \mathbb{E}(X^2) \leq P_X} \{ h(Y) - h(Y|X) \} = \max_{f_X(x) : \mathbb{E}(X^2) \leq P_X} \{ h(Y) \} - h(Z) \\
= \frac{1}{2} \log_2 (2\pi e (H^2 \sigma_X^2 + \sigma_Z^2)) - \frac{1}{2} \log_2 (2\pi e \sigma_Z^2) \\
= \frac{1}{2} \log_2 \left( 1 + \frac{H^2 \sigma_X^2}{\sigma_Z^2} \right) \text{ bit per channel use} \]
Capacity of real-valued subchannel, cont’d

- Capacity-achieving input probability density function (PDF) is Gaussian $X \sim \mathcal{N}(0, \sigma_X^2)$ with $\sigma_X^2 = P_X$
- Output PDF, $Y \sim \mathcal{N}(0, \sigma_Y^2)$, with $\sigma_Y^2 = H^2 \sigma_X^2 + \sigma_Z^2$ and $\sigma_Z^2 = P_Z$
- Ratio $H^2 / \sigma_Z^2$ is a property of the subchannel and determines its quality
Parallel channels

\[ Z_1 \sim \mathcal{N}(0, \sigma_{Z_1}^2) \]
\[ X_1 \xrightarrow{H_1} Y_1 \]
\[ Z_2 \sim \mathcal{N}(0, \sigma_{Z_2}^2) \]
\[ X_2 \xrightarrow{H_2} Y_2 \]
\[ \vdots \]
\[ Z_N \sim \mathcal{N}(0, \sigma_{Z_N}^2) \]
\[ X_N \xrightarrow{H_N} Y_N \]

- \( N \) real-valued, parallel, independent subchannels
- Goal: find the best distribution of \( P^{(\text{tot})}_X \) over the \( N \) subchannels such that the right-hand side of

\[
I(X_1 X_2 \ldots X_N; Y_1 Y_2 \ldots Y_N) \leq \frac{1}{2} \sum_{i=1}^{N} \log_2 \left( 1 + \frac{H_i^2 P_{X_i}}{P_{Z_i}} \right)
\]

is maximised under the constraint

\[
\sum_{i=1}^{N} P_{X_i} \leq P^{(\text{tot})}_X
\]
Parallel channels—waterfilling solution

Constrained optimisation problem

\[ P^{(\text{opt})}_{X_i} = \arg \max_{P_{X_i}} \frac{1}{2} \sum_{i=1}^{N} \log_2 \left( 1 + \frac{H_i^2 P_{X_i}}{P_{Z_i}} \right) \quad \text{subject to} \quad \sum_{i=1}^{N} P_{X_i} \leq P^{(\text{tot})}_{X} \]

for \( i = 1, 2, \ldots, N \) is solved

\[ \left( P^{(\text{tot})}_{X} - \sum_{i=1}^{N} P_{X_i} \right) \lambda \]

where the real-valued \( \lambda \) is the so-called Lagrange multiplier, which yields

\[ \frac{1}{2} \sum_{i=1}^{N} \log_2 \left( 1 + \frac{H_i^2 P_{X_i}}{P_{Z_i}} \right) + \left( P^{(\text{tot})}_{X} - \sum_{i=1}^{N} P_{X_i} \right) \lambda \]
Setting the partial derivatives with respect to $P_{X_i}$ equal to 0 yields

\[
\frac{\partial}{\partial P_{X_i}} \left( \frac{1}{2} \sum_{i=1}^{N} \log_2 \left( 1 + \frac{H_i^2 P_{X_i}}{P_{Z_i}} \right) + \left( P_{X}^{(\text{tot})} - \sum_{i=1}^{N} P_{X_i} \right) \lambda \right) = 0
\]

\[
\frac{H_i^2}{P_{Z_i} + H_i^2 P_{X_i}} \frac{1}{2 \ln(2)} - \lambda = 0
\]

Eventually, we obtain

\[
P_{X_i} = \tilde{\lambda} - \frac{P_{Z_i}}{H_i^2}, \quad \text{with } \tilde{\lambda} = \frac{1}{2\lambda \ln(2)}
\]

Since $P_{X_i} \geq 0$ must hold, the optimum power allocation is given by

\[
P_{X_i} = \max\{\tilde{\lambda} - \frac{P_{Z_i}}{H_i^2}, 0\} \quad \text{with } \tilde{\lambda} \text{ chosen such that } \sum_{i=1}^{N} \max\{\tilde{\lambda} - \frac{P_{Z_i}}{H_i^2}, 0\} = P_{X}^{(\text{tot})}
\]
Complex-valued subchannel

Subchannel model in complex-valued notation:

\[ Y = H \cdot X + Z, \quad H \text{ complex-valued, } \begin{bmatrix} \text{Re}(Z) \\ \text{Im}(Z) \end{bmatrix} \sim \mathcal{N}(0, \sigma_Z^2 I), \]

Subchannel model in matrix notation \((H = Ae^{j\phi}, A \geq 0)\):

\[
\begin{bmatrix}
\text{Re}(Y) \\
\text{Im}(Y)
\end{bmatrix} =
\begin{bmatrix}
A \cos \phi & -A \sin \phi \\
A \sin \phi & A \cos \phi
\end{bmatrix}
\begin{bmatrix}
\text{Re}(X) \\
\text{Im}(X)
\end{bmatrix}
+ \begin{bmatrix}
\text{Re}(Z) \\
\text{Im}(Z)
\end{bmatrix}
\]

Singular value decomposition of \(H\) yields

\[ H = USV, \text{ where } UU^H = I, VV^H = I, S = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \]

Remember that \(S\) contains the non-negative square roots of the eigenvalues of \(HH^H\), which is given by

\[ HH^H = \begin{bmatrix} A^2 & 0 \\ 0 & A^2 \end{bmatrix} \]
Multiplying

\[ y = USV \underbrace{x + n}_{H} \]  

with \( U^H \) from the left, we obtain

\[ U^Hy = S \underbrace{Vx}_{\tilde{x}} + U^Hn, \]

(2)

- \( \tilde{n} \) has the same mean vector and the same covariance matrix as \( n \)
- \( \tilde{x} \) has the same mean vector and the same covariance matrix as \( x \)
- (1) and (2) have the same capacity
- (2) is just a set of 2 real-valued parallel channels \( \rightarrow \) waterfilling!
Complex-valued subchannel, cont’d

However, since

$$S = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$$

and

$$\begin{bmatrix} \text{Re}(Z) \\ \text{Im}(Z) \end{bmatrix} \sim \mathcal{N}(0, \sigma_Z^2 I)$$

the two channels have identical subchannel SNR $\rightarrow$ waterfilling simplifies to equal distribution of available power.

The capacity of a complex-valued subchannel is thus

$$C = \log_2 \left( 1 + \frac{|H|^2 P_X}{P_Z} \right) \text{ bit per channel use}$$

- For complex-valued subchannels, we perform waterfilling with $|H_i|$
- QAM-constellation implements equal power-distribution over the two dimensions
Waterfilling solution: system of linear equations

Solving the waterfilling problem is equivalent to solving the following set of $N+1$ linear equations

$$P_{X_1} = \tilde{\lambda} - \frac{P_{Z_1}}{|H_1|^2}$$

$$P_{X_2} = \tilde{\lambda} - \frac{P_{Z_2}}{|H_2|^2}$$

$$\vdots$$

$$P_{X_N} = \tilde{\lambda} - \frac{P_{Z_N}}{|H_N|^2}$$

$$P_{X_1} + P_{X_2} + \ldots + P_{X_N} = P_X^{(\text{tot})}$$

under the constraint $P_{X_i} \geq 0$, $i = 1, \ldots, N$
Waterfilling solution: example

$|H(f)|^2$: squared magnitude response of channel

$H(z) = 1 + 0.8z^{-1} + z^{-2}$

$P_Z(f)$: noise PSD (zero-mean, variance 0.2 per subchannel ($N = 32$))

$SNR(f) = |H(f)|^2 / P_Z(f)$: resulting channel SNR
Waterfilling solution: example, cont’d

Solution for $P_{X}^{(\text{tot})} = 1$ (upper plot) and for $P_{X}^{(\text{tot})} = 4$ (lower plot):
Waterfilling algorithm: matrix-based implementation

**Step 1:** Set up a system of $N + 1$ equations.

**Step 2:** Solve the system of equations.

**Step 3:** In case all power values are non-negative, the problem is solved. In case $P_{X_i} < 0$ for a set of indices $i \in \mathcal{I} = \{i_1, \ldots, i_K\}$, set up a new system of equations: start from the previous system, omit the $K$ equations $P_{X_i} = \tilde{\lambda} - \frac{P_{Z_i}}{|H_i|^2}$, $i \in \mathcal{I}$ and omit the power values $P_{X_i}$, $i \in \mathcal{I}$ in the last equation, i.e.,

$$\sum_{i \not\in \mathcal{I}} P_{X_i} = P^{(\text{tot})}_X,$$

which results in a new system of $N + 1 - K$ equations. Proceed with Step 2.
Waterfilling algorithm: alternative implementation

**Observation 1:** The waterfilling level for $K$ subchannels of an arbitrary set $\mathcal{I} = \{i_1, \ldots, i_K\}$, in case only these subchannels are used, is given by

$$
\tilde{\lambda} = \frac{P_X^{(\text{tot})} + P_{Zi_1}|H_{i_1}|^2 + \ldots + P_{Zi_K}|H_{i_K}|^2}{K}
$$

Check it out: insert the $K$ equations

$$P_{Xi_j} = \tilde{\lambda} - \frac{P_{Zi_j}}{|H_{i_j}|^2} \quad j = 1, \ldots, K$$

into

$$\sum_{k=1}^{K} P_{Xi_k} = P_X^{(\text{tot})}$$

**Observation 2:** If a subchannel No. $i$ with a certain subchannel SNR $\frac{|H_i|^2}{P_{Zi}}$ is not used because its subchannel SNR is too low, all subchannels with lower SNR are omitted as well $\rightarrow$ process subchannels in sorted order.
Waterfilling algorithm: alternative implementation

**Step 1:** Sort the subchannels such that \( \frac{|H_{i_1}|^2}{P_{Z_{i_1}}} \geq \frac{|H_{i_2}|^2}{P_{Z_{i_2}}} \geq \ldots \geq \frac{|H_{i_N}|^2}{P_{Z_{i_N}}} \); set \( k = 0 \) and \( \lambda \sum = P_X^{(\text{tot})} \).

**Step 2:** Set \( k = k + 1 \) and compute the waterfilling level for the set of users \( \{i_1, \ldots, i_k\} \): \( \lambda \sum = \lambda \sum + \frac{P_{Z_{i_k}}}{|H_{i_k}|^2} \), \( \lambda = \lambda \sum / k \).

**Step 3:** If the power of the subchannel with the lowest subchannel SNR is still larger than zero, i.e., if \( \lambda - \frac{P_{Z_{i_k}}}{|H_{i_k}|^2} > 0 \), go to Step 2. Otherwise proceed with Step 4.

**Step 4:** Undo the increment corresponding to subchannel No. \( i_k \):
\[
\lambda \sum = \lambda \sum - \frac{P_{Z_{i_k}}}{|H_{i_k}|^2}, \quad \lambda = \lambda \sum / (k - 1);
\]
Compute the subchannel power values:
\[
P_{X_{ij}} = \begin{cases} 
\lambda - \frac{P_{Z_{ij}}}{|H_{ij}|^2}, & j = 1, \ldots, k - 1 \\
0, & \text{otherwise}
\end{cases}
\]
Practical aspects of bit and power loading

Two principally different types of constraints:

- Constraint on total power (transmitted on all subchannels)
- Constraint on power spectral density (PSD), which is equivalent to a per-subchannel power constraint if the number of subchannels is large

  - In case the total power is the active constraint, we can do waterfilling with both a total power and a PSD-constraint
  - In case the PSD is the active constraint, there is no need for waterfilling → new problem: find an adequate constellation size for each subchannel such that a prescribed probability of error can be guaranteed
Practical aspects of bit and power loading

- Bit loading in practice usually requires an integer number of bits per subchannel.
- Waterfilling solution will, in general, return a non-integer number of bits per subchannel.
- Finite granularity issue gave rise to the development of several special algorithms (Chow algorithm, Levin-Campello algorithm).

![Bar chart showing bit loading per subchannel index.](chart.png)

- $\text{bit per subchannel}$
- $P_{X_k}$
- $k$
- $\frac{N}{2} + 1$
Gap analysis

Instead of using the capacity formula

\[ c_i = \log_2 \left( 1 + \frac{|H_i|^2 P_{Xi}}{P_{Zi}} \right), \]

we compute the per-subchannel data rate as

\[ b_i = \log_2 \left( 1 + \frac{|H_i|^2 P_{Xi}}{P_{Zi} \Gamma} \right) \]

- \( \Gamma \) denotes the “SNR gap”
- \( \Gamma \) is chosen such that uncoded transmission achieves \( P_s \)
- for QAM, \( \Gamma \) is almost independent of the constellation size
- Example: \( \Gamma = 8.8 \) dB for \( P_s = 10^{-6} \)
- not exact, but simple and robust approach to assigning bits
“Greedy” principle: minimize the cost to achieve a certain sub-goal

Application of bit-loading:

Step 1: find subcarrier $k$ that requires least energy, denoted $\epsilon_k$, to increase $b_k$ by 1

Step 2: find subcarrier $\ell$ that saves most energy, denoted $\epsilon_\ell$, when decreasing $b_\ell$ by 1

Step 3: if more energy is saved than spent ($\epsilon_\ell > \epsilon_k$), move a bit from subcarrier $\ell$ to subcarrier $k$ and goto Step 1

every iteration improves the energy-efficiency while keeping the total number of bits constant and eventually yields a so-called efficient distribution
Rate-adaptive loading (also referred to as fixed-margin loading) aims at maximizing the total datarate under a total power constraint:

\[
\begin{align*}
\text{maximize} & \quad \sum_{P_{X_i}} b_i \\
\text{subject to} & \quad \sum_{P_{X_i}} = \text{const}
\end{align*}
\]

- **b** is *energy-tight* if no additional bit can be added without violating the energy constraint.

Levin-Campello algorithm for rate-adaptive loading:

1. Choose any **b**
2. Make the result of Step 1 efficient
3. Make the result of Step 2 energy-tight
(Performance) Margin adaptive loading (also referred to as power-minimization loading or margin-maximization loading) aims at minimizing the total power under a total datarate constraint:

\[
\begin{align*}
\text{minimize} & \quad \sum_{b_i} P x_i \\
\text{subject to} & \quad \sum b_i = \text{const}
\end{align*}
\] (3)

- \(b\) is *bit-tight* if it achieves the target number of bits

Levin-Campello algorithm for margin-adaptive loading:

- Step 1: choose any \(b\)
- Step 2: make the result of Step 1 efficient
- Step 3: make the result of Step 2 bit-tight
Practical aspects of bit and power loading

- Provision of the maximum rate implies that the rate is variable—in practice, often a minimum rate is required rather than a maximum rate. An algorithm for bit and power loading could focus on minimising the required power per symbol to deliver the required rate.
- Finally, we should remind ourselves that any kind of waterfilling, bit-loading or power-loading requires knowledge of the channel at the transmitter.
Summary

- We worked our way through the capacity of...
  - Real-valued subchannel
  - Complex-valued subchannel
  - $N$ parallel real-valued subchannels
  - $N$ parallel complex-valued subchannel

- Waterfilling algorithms:
  - Matrix-based version
  - Alternative version (involving sorting of subchannels)

- Example

- Practical aspects of waterfilling