Tutorials
Information Theory and Coding
SS2020

Institut für Nachrichtentechnik
Technische Universität Hamburg

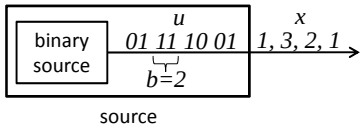
Contents

Chapter 1

Source Coding

1.1 Binary Source and Huffman Code

A binary source emits statistical independent binary data symbols $u \in \{0, 1\}$. The probability of a bit being "0 " is p_0, the probability of a bit "1 " is $p_1 = 1 - p_0$. Before compression, $b = 2$ data symbols u are grouped to a symbol x as depicted in the following figure:



Solve the following problems a)-c)for all parameter combinations

- $b = 2$ and $b = 3$
- $p_0 = 1/2$, $p_0 = 1/4$ and $p_0 = 2/5$.

a) Determine the probabilities $P_X(x)$ of all possible source symbols x.

b) Determine the Entropy $H(X)$ of the source.

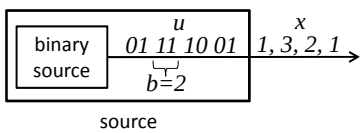
c) State all possible Huffman Codes and

   (1) calculate the average codeword length $\mu_L = E\{L\}$,
   (2) calculate the variance σ_L^2 of the codeword length,
   (3) compare the average codeword length $\mu_L = E\{L\}$ to the entropy $H(X)$ of the source.

d) For $b = 2$, there exist two generally different Huffman-Codes. Determine the probability p_0, for which it is switched between both variants.

1.2 Fano Code

A binary source emits statistical independent binary data symbols $u \in \{0, 1\}$. The probability of a bit being "0 " is p_0, the probability of a bit "1 " is $p_1 = 1 - p_0$. Before compression, $b = 2$ data symbols u are grouped to a symbol x as depicted in the following figure:



Solve the following problems a)-c)for all parameter combinations

- $b = 2$ and $b = 3$

- $p_0 = 1/2$, $p_0 = 1/4$ and $p_0 = 2/5$.

a) State all possible Fano Codes and

   (1) calculate the average codeword length $\mu_L = E\{L\}$,
   (2) calculate the variance σ_L^2 of the codeword length,
   (3) compare the average codeword length $\mu_L = E\{L\}$ to the entropy $H(X)$ of the source.

1.3 Huffman Code and Relative Entropy

Consider a 4-ary random variable X with realizations $x \in \{a, b, c, d\}$. The true probability distribution $p(x)$ is given in the following table:

x	$P(X = x)$
a	1/2
b	1/4
c	1/8
d	1/8

a) Determine the entropy $H(X)$ of the random variable X.

b) Construct a Huffman-Code for binary encoding of the realizations x. Make sure that all steps in the construction of the Huffman code are clearly given. Represent the mapping of realizations x to codewords c of the Huffman code in a table.

c) Determine the average codeword length of the Huffman code from b).

Construction of a Huffman code requires knowledge of the statistics of the random variable X. Often, the probability distribution $p(x)$ of the random variable X is not exactly known but has to be estimated. Assume, that a probability distribution $q(x)$ according to the following table has been estimated instead of the true probability distribution $p(x)$:

x	$q(x)$
a	1/8
b	1/8
c	1/4
d	1/2

d) Construct a Huffman code for X based on the estimated probabilities $q(x)$. Represent the mapping of realizations x to codewords **w** of the Huffman code in a table.

e) Determine the average codeword length of the Huffman code from d) and compare to your result from c).

Assume that the source emits the symbol sequence b, c, a, c, d.

f) Determine the coded bit sequence.

g) Assume that due to a transmission error, the fourth bit in the coded sequence has been corrupted. Decode the erroneous bit sequence. Which fundamental problem of the Huffman code do you identify?

The distance of two probability distributions $p(x)$ and $q_X(x)$ can be measured by the *relative entropy*

$$D(p||q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)}$$

which is also called the Kullback Leibler distance.

h) Determine the relative entropy $D(p||q)$ of the probability distributions $p(x)$ and $q(x)$.

i) Interpret the meaning of the relative entropy for compression in view of your results for problems a)-e).

j) Show, that the relative entropy $D(p||q)$ is always non-negative.

> Help: Jensen's inequality: $E\{f(X)\} \leq f(E\{X\})$ for a concave function $f(\cdot)$ and a random variable X, where $E\{\cdot\}$ denotes expectation.

k) Derive the relation between $p(x)$ and $q(x)$ for which the relative entropy is zero, i.e. $D(p||q) = 0$.

l) Is the relative entropy in general a true distance in the sense that it is symmetric, i.e. does $D(p||q) = D(q||p)$ hold for arbitrary probability distributions $p(x)$ and $q(x)$?

n) Assume that a sequence of b independent realizations of X, i.e. $(x_1, x_2, \ldots, x_b)$, is mapped onto a distinct hyper symbol S. Compute the ratio between the entropy of a single symbol $H(X)$ and the entropy of a hyper symbol $H(S)$.

Chapter 2

Entropy and Mutual Information

2.1 Mutual Information

Show that the mutual information between two random variables X and Y can be expressed by

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

where $H(Y)$ and $H(Y|X)$ denote the entropy of Y and the conditional entropy of Y given X, respectively.

2.2 Entropy of a Continuous Gaussian Random Variable

Show that the entropy $h(X)$ of a continuous Gaussian random variable X is given by

$$h(X) = \frac{1}{2}\log_2(2\pi e\sigma_X^2).$$

2.3 Chain Rules of Entropy and Mutual Information

Chain rules play an important role in information theory.

a) Proof the chain rule of entropy, i.e.

$$H(X_1, X_2, ..., X_N) = \sum_{i=1}^{N} H(X_i|X_{i-1}, ..., X_1).$$

b) Proof the chain rule for mutual information, i.e.

$$I(X_1, X_2, ..., X_N; Y) = \sum_{i=1}^{N} I(X_i; Y|X_{i-1}, ..., X_1).$$

2.4 Maximum Entropy

In this problem, we will prove, that the entropy of a discrete random variable X with N different realizations x and probability distribution $p(x)$ is upper bounded by

$$H(X) \leq \log_2 N$$

We will make use of the relative entropy (also called the Kullback Leibler divergence)

$$0 \leq D(p||q) = \sum_x p(x)\log_2\frac{p(x)}{q(x)},$$

which is a measure for the distance of two probability distributions $p(x)$ and $q(x)$. The relative entropy $D(p||q)$ is always non-negative.

a) Determine the probability distribution $q(x)$ of a uniformly distributed random variable X, which can take N different realizations x.

b) Determine the entropy $H_u(X)$ of a uniformly distributed random variable X, which can take N different realizations x.

c) Determine the relative entropy $D(p||q)$ between a probability distribution $p(x)$ and a uniform probability distribution $q(x)$ depending on the entropy $H(X)$ of a random variable with probability distribution $p(x)$.

d) Use the result from c) in order to prove that $H(X) \leq \log_2 N$.

e) Which probability distribution delivers the maximum possible entropy of a random variable X with N different realizations?

2.5 Conditional Entropy

Show, that $H(X|Y) = 0$ holds for the conditional entropy $H(X|Y)$, if the random variable X is a function of the random variable Y.

Help: $0 \cdot \log_2 0 = 0$ (for convention, since $x \cdot \log_2 x \to 0$ for $x \to 0$).

Chapter 3

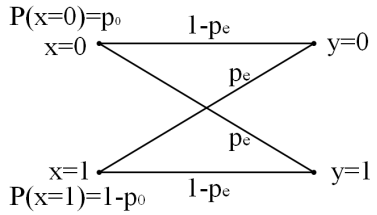
Channel Models and Channel Capacity

3.1 Capacity of the Binary Symmetrical Channel (BSC)

Determine the channel capacity for a binary symmetrical channel (BSC) using the general equation for the mutual information $I(X;Y)$. The transmit symbols from the binary alphabet $x \in \{0,1\}$ have the probabilities

$$P(x) = \begin{cases} p_0 & \text{if} \quad x = 0 \\ 1 - p_0 & \text{if} \quad x = 1 \end{cases}$$

The error probability of the BSC is denoted by p_e.



a) Determine the probabilities $P(y)$ of the output symbols $y = 0$ and $y = 1$.

b) Divide the equation for the mutual information $I(X;Y)$ into two parts such that the first part depends only on the probability $P(y)$ of the channel output symbols y.

c) Determine the first term (entropy $H(Y)$) from b) as a function of p_0.

d) Determine the second term (conditional entropy $H(Y|X)$) for the given probability distribution $P(x)$ and the given channel transition probabilities $P(y|x)$ depending on p_0 and p_e.

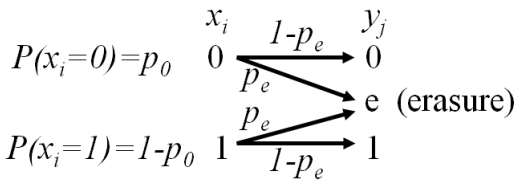
e) Derive an expression for the mutual information $I(X;Y)$ depending on p_0 and p_e.

f) Show that the mutual information is maximized for equally likely channel input symbols x.

g) Determine the channel capacity C of the BSC depending on the channel error probability p_e.

3.2 Binary Erasure Channel

The binary erasure channel is defined in the following figure:



The transmit symbols x_i are taken from the set $x_i \in \{0, 1\}$ with probabilities $P(x_i = 0) = p_0$ and $P(x_i = 1) = 1 - p_0$.
The observed symbols y_j at the channel output are taken from the set $y_i \in \{0, 1, e\}$, where 'e' indicates an erasure, i.e. the respective bit is lost.
The channel is symmetric and does not cause corruption of bits. Correct reception of a bit is achieved with probability $1 - p_e$. An erasure occurs with probability p_e.
The channel capacity of the binary erasure channel shall be derived in the following.

a) Determine the probabilities $P(y_j)$ of the channel output symbols y_j depending on p_0 and p_e for all possible realizations of y_j.

b) Determine the entropy $H(Y)$ of the channel output depending on p_0 and p_e. Write $H(Y)$ as a weighted sum of the binary entropy functions $H_b(p_0)$ and $H_b(p_e)$.

c) Determine the probabilities $P(y_j|x_i)$ that y_j is observed given x_i was transmitted for all combinations of realizations x_i and y_j depending on p_0 and p_e.

d) Determine the joint probabilities $P(x_i, y_j)$ for all combinations of realizations x_i and y_j depending on p_0 and p_e.

e) Determine the mutual information $I(X; Y)$ of channel input and output depending on p_0 and p_e.

f) Determine the input symbol probability p_0 which maximizes the mutual information $I(X; Y)$. Explain your solution clearly.

g) Determine the channel capacity C of the binary erasure channel. Give reasons for your solution.

h) Assume that the erasure probability is $p_e = 0.25$. What is the maximum rate R_{max} of an error correcting code which theoretically allows error free transmission through the binary erasure channel? Give reasons for your solution.

3.3 Capacity of the Binary Input AWGN Channel

Shannon's famous equation

$$C = \frac{1}{2} \log_2 \left(1 + \frac{2E_s}{N_0} \right) \tag{3.1}$$

does not put restrictions on the transmit symbol alphabet. The capacity achieving transmit symbol distribution $f_X(x)$ turns out to be a Gaussian distribution. In practice, the transmit symbols can often not be chosen arbitrarily but are determined by the modulation scheme. In case of BPSK modulation, the transmit symbols are taken from the alphabet $X \in \{-1, +1\}$. Such an AWGN channel with binary input is called a *binary-input AWGN channel*.
Show that the capacity of a binary input AWGN channel with input symbols $X \in \{\pm 1\}$ and noise variance σ_N^2 is given by

$$C = -\frac{1}{\sqrt{8\pi\sigma_N^2}} \int_{-\infty}^{\infty} \left(e^{-\frac{(y+1)^2}{2\sigma_N^2}} + e^{-\frac{(y-1)^2}{2\sigma_N^2}} \right) \log_2 \left[\frac{1}{\sqrt{8\pi\sigma_N^2}} \left(e^{-\frac{(y+1)^2}{2\sigma_N^2}} + e^{-\frac{(y-1)^2}{2\sigma_N^2}} \right) \right] dy$$
$$- \frac{1}{2} \log_2 \left(2\pi e \sigma_N^2 \right)$$

Hint: As you saw in previous tasks (binary symmetrical and binary erasure channel) the capacity was obtained for input symbols with equal probability. You can assume that equally probable input symbols do maximize the mutual information for the binary AWGN channel, too.

3.4 Wireless Transmission

Consider an *AWGN* channel with two-sided noise power spectral density $N_0/2$, bandwidth $B = 5$ MHz and a received signal power P_X of -80 dBm. The receiver is characterized by a noise figure F of 10 dB. The receiver operates at a temperature of 21 °C.

a) Determine the noise power P_N.

b) Determine the signal to noise power ratio (SNR) in dB.

c) Determine the theoretically achievable data rate C^* in bit/s, i.e. the data rate which can be transmitted error-free in theory.

d) What is the capacity-achieving probability distribution $f_X(x)$ of the transmit symbols X?

e) Which code rate R and which codeword length N have to be applied for capacity-achieving transmission?

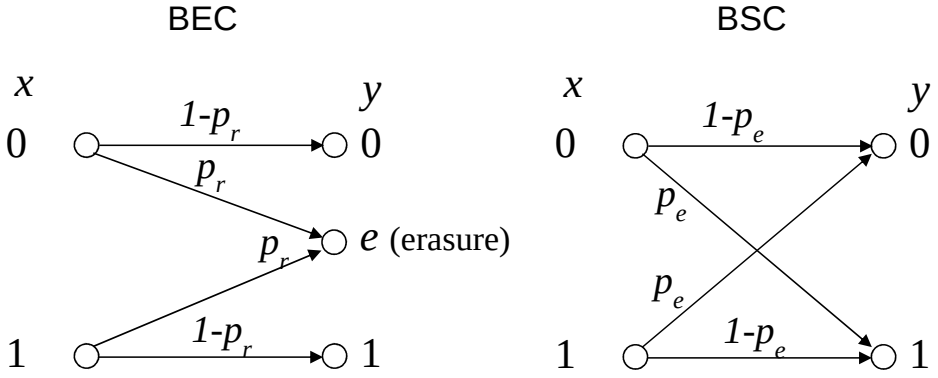
f) Which data rate could be theoretically achieved if the bandwidth was doubled to $2B$?

g) Which transmit power would be required in order to obtain the same achievable data rate as in f) but with the original bandwidth B?

h) Which bandwidth and which transmit power would be required in order to achieve the same data rate as in c) but with uncoded transmission (no channel coding) using $BPSK$ modulation at a target bit error probability of 10^{-5}.

3.5 BEC vs. BSC

The binary erasure channel (BEC) and the binary symmetrical channel (BSC) should be compared.



The transmit symbols x_i are taken from the set $\{0, 1\}$ with probabilities $P(x_i = 0) = p_0$ and $P(x_i = 1) = p_1$. The probability for a transmission error in the BSC is p_e, the probability for an erasure in the BEC is p_r.
The channel capacity of the BSC is

$$C_{\mathrm{BSC}}(p_e) = 1 - H_b(p_e)$$

with the binary entropy function $H_b(p) = -(1-p)log_2(1-p) - p\log_2(p)$.
The mutual information of the BEC is

$$I_{\mathrm{BEC}}(X, Y) = (1 - p_r)H_b(p_0)\,.$$

a) Is the BSC a discrete memoryless channel? Why?

b) Which propabilities p_0 and p_1 maximize the mutual information of the BEC?

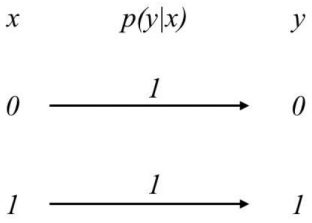
c) Determine the channel capacity $C_{\mathrm{BEC}}(p_r)$ of the BEC.

d) Show that the channel capacity of the BEC is larger than the one of the BSC for $p_e = p_r < \frac{1}{2}$ (error probability and erasure probability are equal)

e) Give an vivid explanation of the result from task d)

3.6 Capacity of the Noiseless Binary Channel

The noiseless binary channel is given by the following state transition diagram



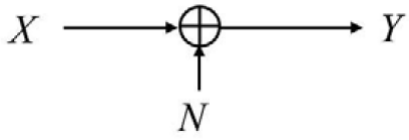
a) Determine the channel capacity by intuition. Explain your solution clearly and answer in complete sentences.

b) Derive the channel capacity formally. Make sure that all steps in your derivation are clearly given.

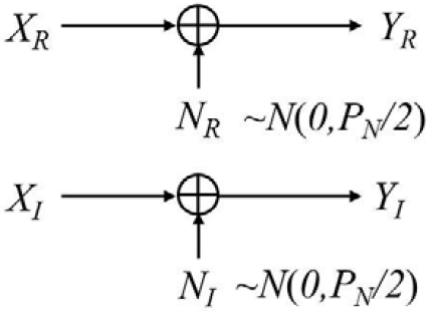
c) Determine the capacity achieving probability distribution $p(x)$ of the channel input symbols.

3.7 Capacity of a Complex AWGN Channel

We consider a discrete-time complex AWGN channel as depicted in the following figure:



The complex noise N is circularly symmetric white and Gaussian with power P_N.The complex AWGN channel could be the equivalent baseband channel of a bandpass transmission scheme. The complex AWGN channel can be represented by two independent real AWGN subchannels (real part and imaginary part) as depicted in the following figure:



The additive white Gaussian noise in real part and imaginary part is uncorrelated, i.e. $E\{N_R \cdot N_I\} = E\{N_R\} \cdot E\{N_I\}$ and has the same power $E\{N_R^2\} = E\{N_I^2\} = \frac{P_N}{2}$.
The total transmit power of the complex AWGN channel is given by P_X.

a) Determine the optimum allocation of the transmit power to the two parallel subchannels (real part and imaginary part).

b) State the equation for the channel capacity of one of the two subchannels, e.g. for the real part, depending on its input power P_R.

c) Derive the equation for the channel capacity of the complex AWGN channel starting from your result from b).

d) Sketch qualitatively the channel capacity C of the complex AWGN channel vs. the SNR in dB. Label the axes completely.

e) Sketch qualitatively in your plot from d) the capacity CQPSK of the complex AWGN channel, when the transmit symbols are restricted to QPSK.

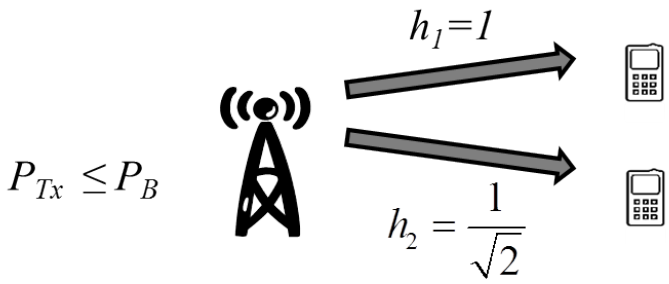
f) Discuss, if it makes sense from an information theory point of view to apply QPSK modulation at low SNR or high SNR, respectively.

Chapter 4

Multi-User Information Theory

4.1 Capacity Region of Downlink TDMA and FDMA in Wireless Communications

Consider the following downlink transmission scenario in a wireless system:



A base station serves two users with independent data. The available bandwidth is $B = 10$ MHz and the maximum transmit power of the base station is $P_B = 10$ W. Each user faces an AWGN channel with noise power spectral density $N_0 = 10^{-20}$ W/Hz in the complex equivalent baseband. The channel coefficients h_k of the two users k, $k \in \{1, 2\}$ are given by

$$h_1 = 1 \qquad h_2 = \frac{1}{\sqrt{2}}$$

The received power at user k is given by $|h_k|^2 P_{\mathbf{Tx}}$, where $P_{\mathbf{Tx}}$ is the transmit power of the base station. The achievable rates with different transmission strategies shall be analyzed in the following problems.

a) Determine the single user capacity C_1 for user 1, i.e. the capacity which is achieved if only user 1 is served.

b) Determine the single user capacity C_2 for user 2, i.e. the capacity which is achieved if only user 2 is served.

c) Sketch the achievable rate region for time division multiple access (TDMA) with power constraint.

d) Determine the achievable rate pair $R_{1,\mathrm{TDMA}}, R_{2,\mathrm{TDMA}}$ for TDMA, if the channel is allocated to both users for the same fraction of the total transmission time.

e) Determine the achievable rate pair $R_{1,\mathrm{FDMA}}, R_{2,\mathrm{FDMA}}$ for FDMA, if the same fraction of the channel bandwidth is allocated to the two users.

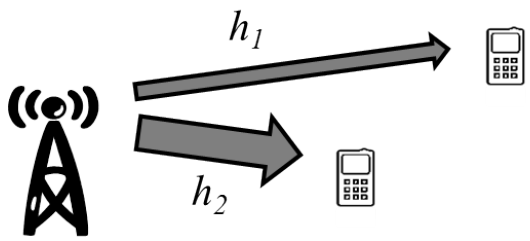
f) Determine the achievable sum rate for the TDMA and FDMA schemes in d) and e) where the same fraction of resources for both users is allocated.

g) Which TDMA and FDMA transmission strategy maximizes the sum rate?

h) Determine the maximum sum rate according to your solution from g).

4.2 Capacity Region of Superposition Coding in a Broadcast Channel

Consider a wireless broadcast channel, where a base station serves two users:



The available bandwidth is denoted B and the maximum transmit power of the base station is P_B. The received signal at user k is given by

$$y_k = h_k x + n_k, \tag{4.1}$$

where h_k is the channel coefficient of user k, x is the transmit signal and n_k is additive white Gaussian noise with the same power P_N at both users.
The base station applies superposition coding strategy such that the transmit signal is a superposition of the signals s_k intended for the two users $k \in \{1, 2\}$:

$$x = s_1 + s_2, \tag{4.2}$$

where s_k is contained in x with power P_k and

$$P_1 + P_2 \leq P_B.$$

Under the assumption that

$$|h_1|^2 < |h_2|^2,$$

the achievable rates for the superposition coding strategy have been shown to be given by the following expressions:

$$R_1 = B \log_2 \left(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + P_N} \right), \tag{4.3}$$

$$R_2 = B \log_2 \left(1 + \frac{|h_2|^2 P_2}{P_N} \right). \tag{4.4}$$

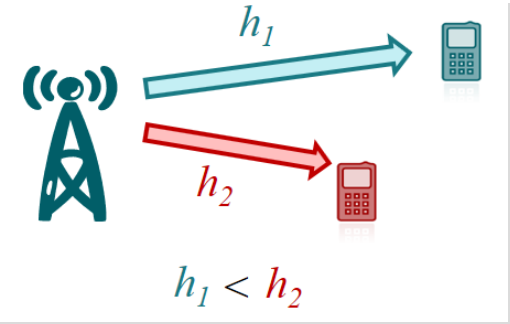
a) Which power allocation P_1, P_2 in equation (4.2) corresponds to a single user transmission?

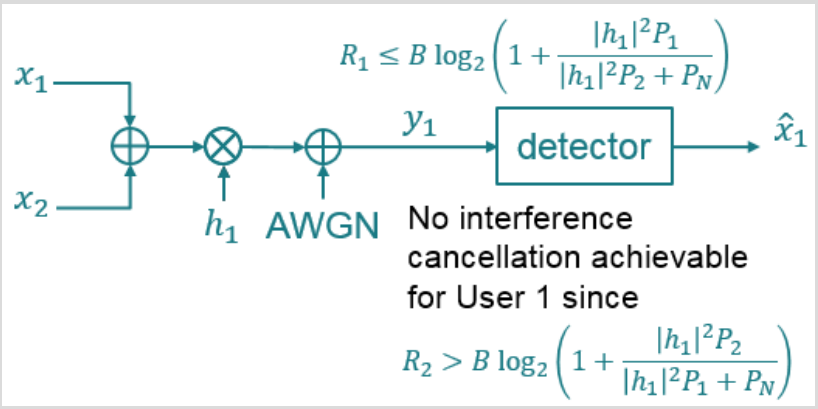
$P_1 \leq P_B, P_2 = 0 \Rightarrow$ only user 1 is served
$P_2 \leq P_B, P_1 = 0 \Rightarrow$ only user 2 is served

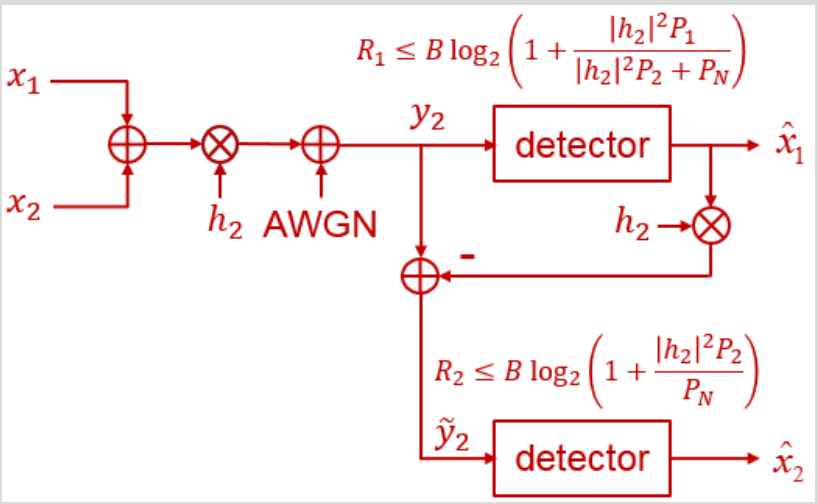
b) Give an interpretation of equations (4.3) and (4.4): Which transmission and detection strategies can achieve the rates according to (4.3) and (4.4)? Answer in complete sentences!

User 2 transmits at its single user capacity. User 1 transmits at a rate which is supported, if the signal of user 2 occurs as additional AWGN.

User 2 experiences the better channel quality. Hence, user 2 can detect for user 1 and subtract it from its received signal. Then, it can detect its own signal as if the signal for user 1 was not present. User 1 experiences the worse channel. It cannot detect the signal for user 2 as the rate of user 2 exceeds the channel capacity of user 1, i.e. $R_2 > C_1$. Hence, user 1 detects its signal while treating the signal for user 2 as additional additive white Gaussian noise (AWGN) with power $|h_1|^2 P_2$. The signal for user 2 can be regarded as independent white Gaussian noise, since the optimum transmit signal in an AWGN channel is Gaussian distributed and consequently, both s_1 and s_2 are Gaussian distributed.







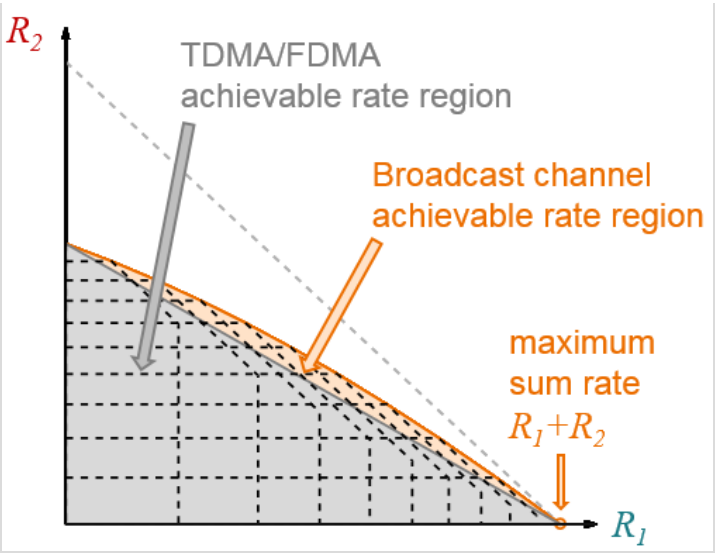
c) Determine an equation for the achievable rate region of superposition coding.

The achievable rate region for a Broadcast channel is equal to the union of all corresponding dual MAC capacity regions.

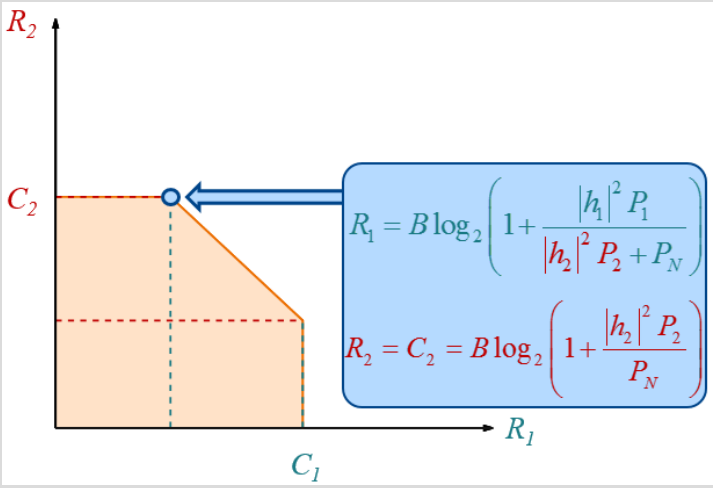
$$C_{SC} = \bigcup_{P_1, P_2 : P_1 + P_2 \leq P_B} C_{\text{dual MAC}}(P_1, P_2)$$

It has been stated in the task description, that the achievable rates R_1 and R_2 for the superposition coding strategy are given by equation 4.3 and 4.4. Further, a detection strategy has been developed in task b) which achieves these rates.

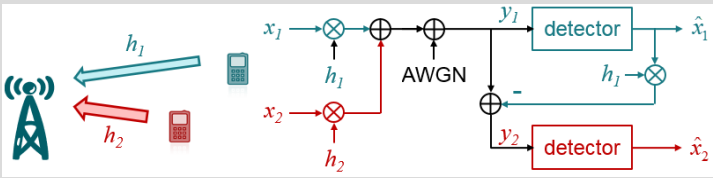
In following figure the union of dual MAC capacity regions (dashed lines) leads to the capacity region of the broadcast channel (solid orange line).



In case of $h_1 < h_2$ the orange line is determined by the connection of upper corners from all dual MAC capacity regions. As an example the subsequent figure shows this corner point.



For a dual MAC the upper corner point of the capacity region can be achieved through interference cancellation techniques. Notice that the resulting rates R_1 and R_2 correspond to the equations 4.3 and 4.4 of the task description, even though the detector for the dual MAC is slightly different compared to broadcast detector of task b).



Thus, we can simply take the union of all rate pairs resulting from combinations of P_1 and P_2 with $P_1 + P_2 \leq P_{\mathrm{BC}}$ to define the achievable rate region as

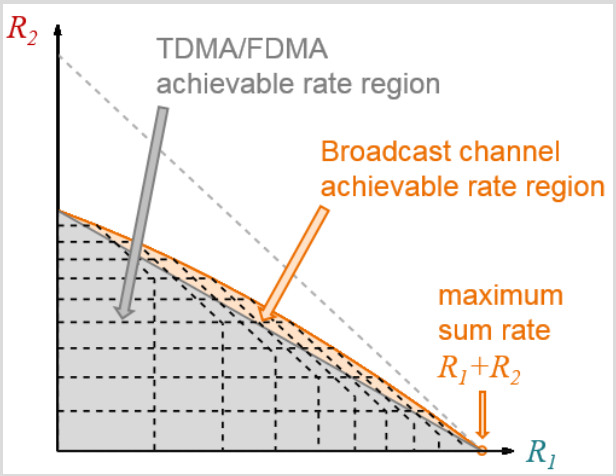
$$C_{SC} = \bigcup_{P_1,P_2:P_1+P_2\leq P_B} \left[R_1 = B\log_2\left(1+\frac{|h_1|^2 P_1}{|h_1|^2 P_2 + P_N}\right), R_2 = B\log_2\left(1+\frac{|h_2|^2 P_2}{P_N}\right)\right]$$

d) Is superposition coding an optimum transmission scheme in the sense that it can achieve the maximum possible rate region? Give clear reasons and answer in complete sentences!

The ultimate rate region which is achievable for a given power constraint without prior restriction of the multiple access scheme is called the capacity region. The rate region in c) is equal to the capacity region of the broadcast channel, since it is the union of the dual MAC capacity regions for all power allocations P_1, P_2. Therefore, superposition coding is an optimum transmission strategy.

e) Which transmission strategy maximizes the sum rate of superposition coding in case of unequal channel quality, i.e. $|h_1|^2 \neq |h_2|^2$?

In case of $|h_1|^2 \neq |h_2|^2$, the sum rate is maximized, if all resources are allocated to the user with the better channel quality. I.e., if $|h_k| > |h_l|$, only user k should be served with $P_k = P_B, P_l = 0$. User k can then transmit at its single user capacity.

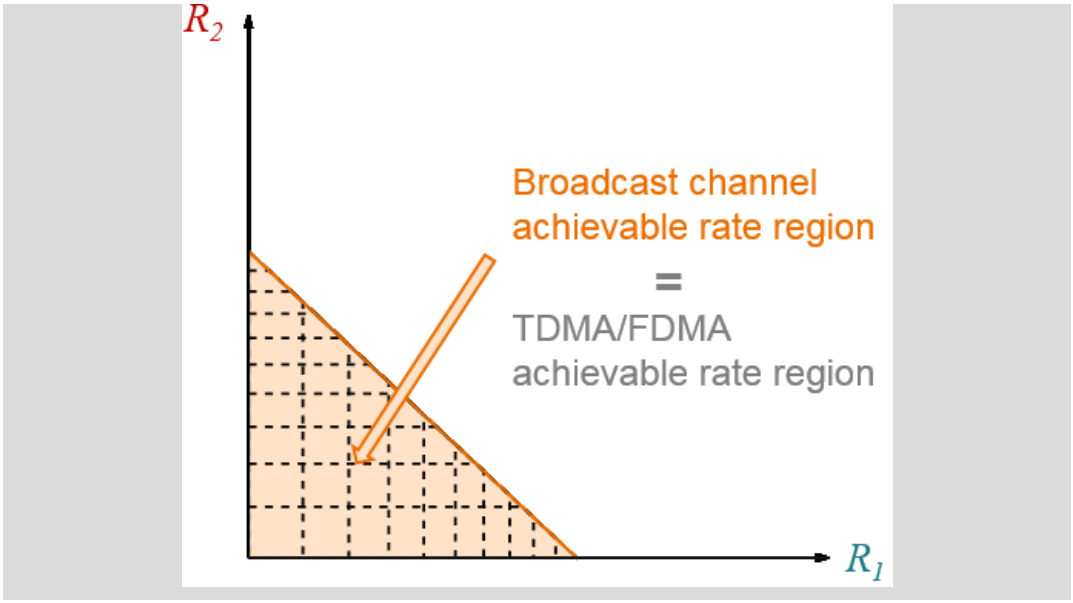


f) Determine an equation for the maximum sum rate according to your solution from e).

$$\max\{R_1 + R_2\} = \max\left\{B \log_2 \frac{|h_1|^2 P_B}{P_N}, B \log_2 \frac{|h_2|^2 P_B}{P_N}\right\}$$

g) Under which condition for $|h_1|^2$ and $|h_2|^2$ does superposition coding provide a larger achievable rate region than time division multiple access (TDMA) ?

For $|h_1| = |h_2|$, the broadcast channel capacity region is equal to the achievable rate region of TDMA. Hence, superposition coding provides a larger achievable rate region than TDMA only if $|h_1| \neq |h_2|$.


R_2
Broadcast channel
achievable rate region
=
TDMA/FDMA
achievable rate region
R_1

Chapter 5

Decoding Principles

5.1 Hard Decision and Soft Decision, Maximum Likelihood (ML) and Maximum A-Posterior (MAP) Decoding

Consider a rate $R = \frac{1}{3}$ repetition code. The information bit sequence $\mathbf{u} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$ shall be transmitted. For transmission, we use BPSK modulation, where the mapping of a code bit x_i to a transmit symbol d_i is given by

$$x_i = 0 \longrightarrow d_i = +1$$
$$x_i = 1 \longrightarrow d_i = -1$$

a) Determine the encoded bit sequence $\mathbf{x}$. How many codewords are transmitted?

b) Determine the sequence $\mathbf{d}$ of BPSK modulated transmit symbols.

The BPSK symbols are transmitted through an AWGN channel. The received sequence is given by

$$\mathbf{y} = \begin{bmatrix} +0.3 & -0.1 & +2.0 & +0.1 & +0.2 & -2.0 & +0.1 & +0.2 & -0.2 \end{bmatrix}^T .$$

c) Determine for each information bit the decoding result in case of

1) hard decision maximum likelihood (ML) decoding. Check, if decoding errors occur.
2) soft decision maximum likelihood (ML) decoding. Check, if decoding errors occur.

Assume that the source bits $u_k = 0$ and $u_k = 1$ are equally likely, i.e. $P(u_k = 0) = P(u_k = 1) = 0.5$.

d) Determine for each information bit the decoding result in case of

1) hard decision maximum a-posteriori (MAP) dedoding. Check, if decoding errors occur and compare to the result of problem c).
2) soft decision maximum a-posteriori (MAP) decoding. Check, if decoding errors occur and compare to the result of problem c).

Assume now that the source emits a bit $u_k = 1$ with higher probability than a bit $u_k = 0$. More precisely, we have $P(u_k = 0) = \frac{1}{4}$ and $P(u_k = 1) = \frac{3}{4}$.

e) Determine for each information bit the decoding result in case of soft decision maximum a-posteriori (MAP) decoding for signal to noise power ratios (SNRs) of 0 dB, 5 dB, 10 dB and 100 dB. Check, if decoding errors occur and compare to the results of problems c) and d).

f) Hard decision decoding can be viewed as receiving data from a binary symmetric channel (BSC) with error probability p_e. Consider hard decision MAP decoding with source statistics $P(u_k = 0) = \frac{1}{4}$ and $P(u_k = 1) = \frac{3}{4}$. Determine the range of channel error probabilities p_e, for which a MAP decoder will decode $\hat{u} = 0$, given that the sequence $\hat{\mathbf{y}} = \begin{bmatrix} +1 & -1 & +1 \end{bmatrix}^T$ was received.

5.2 Introduction to Log-Likelihood Ratios

Consider log-likelihood ratios (L-values) of bits $u_k \in \{\pm 1\}$.

a) What is the probability $P(u_k = +1)$ if

   (1) $L(u_k) = 0$
   (2) $L(u_k) = \log_e(2)$
   (3) $L(u_k) = -1$

b) A bit $u_k = +1$ is two times as likely as $u_k = -1$. Determine the log-likelihood ratio $L(u_k)$.

c) Consider two bits with log-likelihood ratios $L(u_1) = -0.2$ and $L(u_2) = 20$. Determine the log-likelihood ratio $L(u_1 \oplus u_2)$ of the XOR combination $u_1 \oplus u_2$.

d) Show that the probability $P(u_k = +1)$ can be expressed as

$$P(u_k = \pm 1) = \frac{e^{\frac{L(u_k)}{2}}}{1 + e^{L(u_k)}} e^{u_k \frac{L(u_k)}{2}}$$

5.3 Soft-Output Decoder

a) A source generates statistical independent binary data $u_k \in \{\pm 1\}$, where $u_k = +1$ occurs with twice the probability of $u_k = -1$.

   (1) Determine the log-likelihood ratio $L_a(u_k)$, that represents the statistics of the source.

   (2) Determine the log-likelihood ratio $L(x)$, where the bit

$$x = u_1 \cdot u_2$$

   is the product of two consecutive source bits u_1 and u_2.

b) A soft-output decoder determines a log-likelihood ratio of $L(\hat{u}) = -1$ for the information bit u.

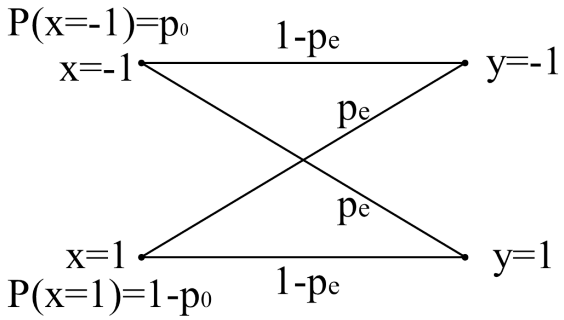
   (1) What is the hard decision $\hat{u}$ on the bit u.
   (2) Determine is the bit error probability P_b of the decoded bit $\hat{u}$? It can be estimated by the soft-output L-value of the decoder.

5.4 LLRs in AWGN Channel

A binary code of rate $R = \frac{1}{2}$ is used for a transmission over an AWGN channel. For the information word u the codeword $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ is transmitted and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ is received.
Show that $L(\hat{u}) = L(u|y_1, y_2) = L(y_1|u) + L(y_2|u) + L_a(u)$ as the channel is memoryless.

5.5 Binary Symmetric Channel, L-Values, ML, MAP

Consider a binary symmetric channel (BSC) with transmit symbols $x_k \in \{-1, +1\}$ and error probability $p_e = 0.1$.



a) Is the BSC a memoryless channel? Why?

b) The uncoded binary data sequence $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = \begin{bmatrix} -1 & -1 \end{bmatrix}^T$ is transmitted through the BSC channel. Determine the probability that

   (1) the sequence $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T = \begin{bmatrix} +1 & -1 \end{bmatrix}^T$ is received.

   (2) the sequence $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T = \begin{bmatrix} -1 & +1 \end{bmatrix}^T$ is received.

   (3) both symbols are received error-free.

c) The sequence $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T = \begin{bmatrix} -1 & +1 \end{bmatrix}^T$ is received from the BSC. Assume that the transmit symbols $x_k = \pm 1$ are statistically independent and equally likely. Determine the probability that

   (1) the sequence $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = \begin{bmatrix} -1 & -1 \end{bmatrix}^T$ was transmitted.

   (2) the sequence $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = \begin{bmatrix} +1 & +1 \end{bmatrix}^T$ was transmitted.

d) Derive the probabilities from c) under the assumption that the probability of a transmit symbol $x_k = +1$ is three times the probability of a transmitted $x_k = -1$.

e) Derive the general equation for an a-posteriori probability (APP) estimate and a maximum likelihood (ML) estimate using log-likelihood ratios.

f) Determine the APP and ML soft-output log-likelihood ratios as well as the respective hard decisions for the received sequence and the source statistics used in d).

Chapter 6

Linear Block Codes

6.1 Systematic $(3, 2, 2)_2$ Block Codes

Consider systematic $(3, 2, 2)_2$ block codes, where the information bits appear in original order as the first code bits.

a) Determine the code rate R.

b) How many distinct $(3, 2, 2)_2$ block codes exist?

c) Are the codes from b) linear codes?

6.2 Binary Code of Length $N = 5$

A binary code of length $N = 5$ is used for the transmission of data symbols. The encoder maps the data symbols to vectors of length 5 bits, where exactly 3 bits have value '1'.

a) How many codewords exist?

b) State all codewords of this code. What is the minimum Hamming distance of the code?

c) Is this code a linear code?

6.3 Single-Parity-Check-Code and Dual Code

Consider a $(N,\ N-1)_2$ Single-Parity-Check-Code (SPC-Code).

a) Determine the code rate R.

b) Determine the minimum Hamming Distance d_{min} depending on N.

c) How many bit errors can at least be detected by a $(N,\ N-1)_2$ SPC-Code?

d) How many bit errors can at least be corrected by a $(N,\ N-1)_2$ SPC-Code?

For the following parts e) - l), consider the special case $N = 4$.

e) State all codewords of the SPC-codes.

f) Determine a generator matrix $\mathbf{G}_{SPC}$ of the SPC-code.

g) Determine a parity check matrix $\mathbf{H}_{SPC}$ of the SPC-code.

h) Determine the product $\mathbf{H}_{SPC}^T \cdot \mathbf{G}_{SPC}$ of the parity check matrix $\mathbf{H}_{SPC}$ and the generator matrix $\mathbf{G}_{SPC}$.

i) Derive the Maximum-Likelihood (ML) decoding rule for the SPC-Code for transmission over an AWGN channel. Simplify the expression as much as possible!

The dual code $\mathcal{C}^\perp$ corresponding to the code $\mathcal{C}$ is defined as the code which is generated by switching the roles of generator matrix and parity check matrix, i.e.

$$\mathbf{G}_{SPC}^\perp = \mathbf{H}_{SPC}\,.$$

j) Determine the code rate $R^{\perp}$ of the dual code of the $(N,\ N-1)_2 = (4,\ 3)_2$ SPC-Code.

k) Which well known block code is this dual code?

l) Determine the minimum Hamming distance $d_{min}^{\perp}$ of the dual code $\mathcal{C}^{\perp}$.

m) Now, consider a general $(N,\ K)_2$ code $\mathcal{C}$. What is the code rate $R_{\mathcal{C}^{\perp}}$ of the dual code $\mathcal{C}^{\perp}$ of $\mathcal{C}$ depending on N and K.

6.4 Existence of an $(N,\ K,\ d_{min})_2$

Consider a linear $(N,\ K,\ d_{min})_2$ block code.

a) How many bit errors can a code with minimum Hamming distance $d_{min} = 5$ always correct?

b) How many bit errors can a code with minimum Hamming distance $d_{min} = 5$ always detect?

c) Can a linear $(10,\ 8,\ 5)_2$ block code exist? Explain your answer!

6.5 Hamming Code

In this task a $(7,\ 4,\ 3)$ Hamming code with the following generator matrix is considered:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

a) Is this Hamming code a systematic code?

b) Determine a parity check matrix $\mathbf{H}$ for this Hamming code.

c) Sketch the Tanner graph characterizing this Hamming code.

d) What is the girth of a Tanner graph? Determine the girth of the graph in task c)

e) Can message-passing decoding be a maximum likelihood decoding in the case of this Hamming code?

f) How many bit errors t can at least be corrected by this code?

g) Determine the probability that more than t errors occur in a BSC with error probability p_e, when taking the error correction capabilities of the code into account.

h) How many bit erasures can at least be corrected in a binary erasure channel (BEC)?

6.6 Reed-Muller Codes of First Order

A $RM(1,\ m)$ Reed-Muller Code of first order is determined by the following generator matrix:

$$\mathbf{G}_{1,m} = \begin{bmatrix} \mathbf{G}_{1,m-1} & \mathbf{0} \\ \mathbf{G}_{1,m-1} & \mathbf{1}_{2^{m-1},1} \end{bmatrix},$$

where $\mathbf{1}_{2^{m-1},1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ is a $(2^{m-1} \times 1)$ vector containing only ones.

For initialization, $\mathbf{G}_{1,1}$ is chosen to be the (2×2) identity matrix $\mathbf{I}_2$.

a) Determine the generator matrix $\mathbf{G}_{1,2}$ of a $RM(1,\ 2)$ code (i.e. $m = 2$).

b) Does the generator matrix from a) represent a systematic encoder? Explain your answer!

c) Determine the code rate R of the $RM(1,\ 2)$ code?

d) State all codewords of the $RM(1,\ 2)$ code.

e) Is the RM(1, 2) code a linear code? Explain your answer!

f) Determine the minimum Hamming distance d_{min} of the $RM(1,\ 2)$ code?

g) How many errors can a $RM(1,\ 2)$ code always detect? Explain your answer!

h) How many errors can a $RM(1,\ 2)$ code always correct? Explain your answer!

i) State the Hamming-weight distribution of the $RM(1,\ 2)$ code.

j) Show that the $RM(1,\ 2)$ code is a Single-Parity-Check code.

k) Derive the generator matrix of a systematic $(4,\ 3,\ 2)_2$- Single-Parity-Check code by elementary column operations on the generator matrix $\mathbf{G}_{1,2}$ of a $RM(1,\ 2)$ code.

l) Determine a parity check matrix $\mathbf{H}$ of a $RM(1,\ 2)$ code.

m) The dual code $\mathcal{C}^{\perp}$ of a code $\mathcal{C}$ is obtained if the roles of generator matrix and parity check matrix are exchanged. Which code is the dual code of a $RM(1,\ 2)$ code?

6.7 Properties of Linear Block Codes

The code $\mathcal{C}$ consists of the following codewords:

$$\mathbf{x}_1 = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix}1\\1\\0\\1\\0\end{bmatrix},\ \mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\\1\end{bmatrix},\ \mathbf{x}_4 = \begin{bmatrix}1\\1\\1\\0\\1\end{bmatrix}.$$

a) Is the code $\mathcal{C}$ linear?

b) Determine the minimum Hamming distance d_{min}.

c) How many bit errors can at least be detected by the code $\mathcal{C}$?

d) How many bit errors can at least be corrected by the code $\mathcal{C}$?

e) Determine the code rate R?

f) Which of the following matrices are parity check matrices for the code $\mathcal{C}$? Give reason.

$$\mathbf{H}_1 = \begin{bmatrix}1&0&0\\1&0&1\\0&1&0\\0&0&1\\0&1&1\end{bmatrix},\quad \mathbf{H}_2 = \begin{bmatrix}1&0&0\\1&1&0\\0&0&1\\0&1&1\\0&1&0\end{bmatrix},\quad \mathbf{H}_3 = \begin{bmatrix}1&0&1\\0&1&1\\1&0&1\\1&1&0\\0&1&1\end{bmatrix},\quad \mathbf{H}_4 = \begin{bmatrix}1&1&0&0\\1&0&0&1\\0&1&1&1\\0&1&0&1\\0&0&1&0\end{bmatrix}$$

g) Determine the product of $\mathbf{H}^T \cdot \mathbf{G}$ for the parity check matrix $\mathbf{H}$ found in f) and the generator matrix $\mathbf{G}$ of the code $\mathcal{C}$.

h) Determine a mapping of the information vectors $\mathbf{u}_i$ to the codewords $\mathbf{x}_i$ of code $\mathcal{C}$ for a non systematic encoder.

6.8 Shortening of Linear Block Codes

Sometimes, it is not possible to find a code of suitable codeword length or information word length for a given application. Therefore, it is desirable to construct a suitable code by shortening or extension of a known code. We will investigate the strategy of code shortening for the example of a $(7, 4, 3)_2$ Hamming code. The parity check matrix $\mathbf{H}$ of the Hamming code is given by

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

a) Determine the codeword length N of the Hamming code.

b) Determine the code rate R of the Hamming code.

c) Determine the minimum Hamming distance $d_{\min}$ of the Hamming code.

d) Determine the number of parity checks which are defined in the parity check matrix $\mathbf{H}$.

e) Derive a systematic generator matrix $\mathbf{G}$ from the parity check matrix $\mathbf{H}$. The information bits shall appear in original order as the first bits of the codeword. Make sure that each step in your derivation is clearly given.

The set of codewords of the $(7, 4, 3)_2$ Hamming code is summarized in the following table:

0000000	*0100101*	*1000011*	*1100110*
0001111	*0101010*	*1001100*	*1101001*
0010110	*0110011*	*1010101*	*1110000*
0011001	*0111100*	*1011010*	*1111111*

In the following problems f)-j), we consider code shortening. The codeword length and the information word length of the original code are denoted by N and K, respectively. For shortening of the code, the shortened code consists only of a subset of the codewords in the table above. For the shortened code, we choose only those codewords, which have M leading zeros. The M leading zeros are then deleted in order to obtain codewords of length $N - M$.

f) Determine the information word length of the shortened code depending on K and M.

g) Determine the code rate R_S of the shortened code depending on N, K and M.

h) What is the impact of code shortening on the minimum Hamming distance. Give clear reasons for your answer.

i) How can a parity check matrix $\mathbf{H}_S$ of the shortened code be obtained from the parity check matrix $\mathbf{H}$ of the original code? What are the dimensions of the parity check matrix $\mathbf{H}_S$ of the shortened code depending on N, K and M ?

j) Consider shortening of the $(7, 4, 3)_2$ Hamming code by $M = 3$. Determine the codewords of the resulting code. Which type of code is obtained?

6.9 Code Extension

Code extension refers to a method, where an overall parity check bit is added to each codeword of a given channel code. We denote the parity check matrix of the linear original code by $\mathbf{H}$. The parity check matrix $\mathbf{H}_e$ of the extended code can be obtained by adding first an all-zeros row and then an all-ones column to the original parity check matrix $\mathbf{H}$:

$$\mathbf{H}_e = \begin{bmatrix} & & & 1 \\ & \mathbf{H} & & \vdots \\ & & & 1 \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

a) Is the code rate R_e of the extended code increased or decreased compared to the code rate R of the original code? Give reasons !

b) Assume that the original code has an odd minimum Hamming weight $w_{\min}$. Determine the minimum Hamming distance $d_{\min,e}$ of the extended code depending on the minimum Hamming distance $d_{\min}$ of the original code.

c) Assume that the original code has an even minimum Hamming weight $w_{\min}$. Determine the minimum Hamming distance $d_{\min,e}$ of the extended code depending on the minimum Hamming distance $d_{\min}$ of the original code.

d) Does code extension make sense regarding the error detection/error correction capabilities for codes with even or odd minimum Hamming distance, respectively? Give reasons!

For the remaining problems, we consider an original code with parity check matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

e) Which type of code is defined by the parity check matrix $\mathbf{H}$.

f) Determine a systematic generator matrix $\mathbf{G}$ for the original code.

g) Determine the minimum Hamming distance $d_{\min}$ of the original code.

h) Determine the parity check matrix $\mathbf{H}_e$ of the extended code.

i) Determine the set of code words of the extended code.

j) Determine the minimum Hamming distance $d_{\min,e}$ of the extended code.

Chapter 7

Low Density Parity Check Codes

7.1 Minimum Hamming Distance of LDPC Codes

Consider an LDPC code which is represented by a parity check matrix $\mathbf{H}$. The minimum bit row weight of $\mathbf{H}$ is given by $d_{b,min}$. Let $c_l \in \{0, 1\}$ denote the code bit which is associated with the minimum row weight $d_{b,min}$. The girth of the Tanner graph is larger than 4.

a) By how many parity check equations is the code bit c_l checked?

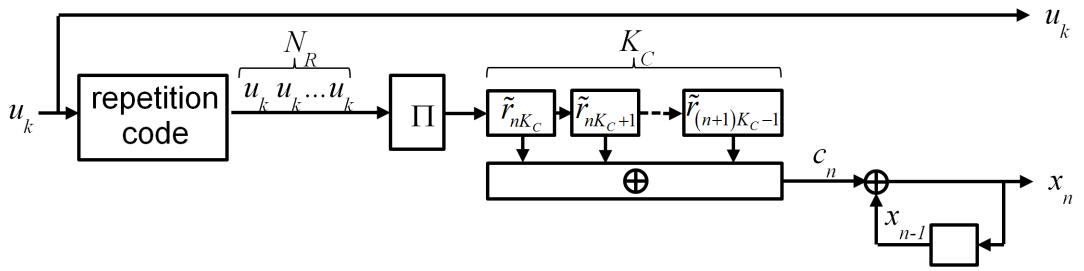
b) Assume a codeword with $c_l = 1$. How many other non-zero code bits must be connected to each of the check nodes, to which c_l is connected? Can those non-zero code bits be the same for different check nodes?

c) Determine the minimum Hamming distance for the LDPC code with girth larger than 4 and minimum row weight $d_{b,min}$.

d) What may be the consequence of increasing the minimum row weight $d_{b,min}$ in the design of LDPC codes with the purpose of increasing the minimum Hamming distance?

7.2 Repeat Accumulate (RA) Code

The block diagram of a repeat accumulate (RA) code is depicted in the following figure:



The information bits and parity bits are denoted by $u_k \in 0, 1$ and $x_n \in 0, 1$, respectively. A parity check matrix of the RA-code is given by

$$\mathbf{H}^T = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

a) Is the RA-code systematic? Give reasons!

b) Is the RA-code regular or irregular? Give reasons!

c) Sketch the Tanner graph which represents the parity check matrix $\mathbf{H}^T$ above.

d) Determine the girth of the Tanner graph.

e) Explain precisely why a short girth is undesired for message passing decoding.

f) Mark a shortest cycle in the Tanner graph from c). Mark also the entries in the parity check matrix $\mathbf{H}^T$ which correspond to the marked cycle.

g) Determine the rates R_R of the repetition coding part within the RA-code.

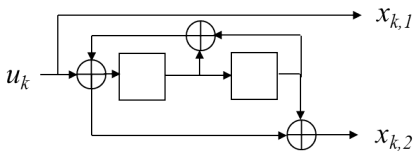
h) Determine the rate R_C of the combiner part within the RA-code.

i) Determine the rate R_A of the accumulator part within the RA-code.

j) Determine the overall code rate R of the RA-code.

7.3 Tanner Graph of a Turbo Code

Consider the following encoder of a convolutional code



a) Is the convolutional encoder systematic ? Give reasons!

b) Determine the rate R of the convolutional code.

c) Determine the code memory M.

d) Determine the generator polynomials of the convolutional code.

e) Describe the generator polynomials in octal representation.

f) Sketch a trellis segment which describes the convolutional code. Label all nodes and transitions completely.

g) Determine for all states of the convolutional encoder the information bit sequence, which has to be fed into the encoder in order to terminate the code in the all zero state.

For the remaining problems, consider that the information bit sequence

$$u_0, u_1, u_2, u_3, u_4$$

of the length $K = 5$ is encoded without termination. The encoder is initialized in the all-zeros-state.

h) Determine the sequence of encoder states depending on $u_0, ..., u_4$ for the complete encoding process.

i) Determine the parity bit sequence $x_{0,2}, x_{1,2}, x_{2,2}, x_{3,2}, x_{4,2}$ depending on $u_0, ..., u_4$.

j) Sketch the Tanner graph of the convolutional code. Label the bit nodes!

k) Is message passing a suitable decoding method for the convolutional code? Give reasons!

l) Determine the parity check matrix $\mathbf{H}^T$ of the convolutional code for an information word length $K = 5$.

For the remaining problems, consider a rate $R = 1/3$ turbo encoder which uses the recursive convolutional code above as constituent codes and an interleaver which is determined by the permutation matrix

$$\mathbf{\Pi} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \text{i.e.} \qquad \begin{bmatrix} \tilde{u}_0 \\ \tilde{u}_1 \\ \tilde{u}_2 \\ \tilde{u}_3 \\ \tilde{u}_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\mathbf{\Pi}} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix},$$

where $\tilde{\mathbf{u}} = [\tilde{u}_0, ..., \tilde{u}_4]^T$ denotes the interleaved information bit sequence.

m) Determine the sequence of interleaved bits depending on $u_0, ..., u_4$

n) Determine the permutation matrix of the deinterleaver.

o) Sketch the block diagram of the turbo encoder.

p) Sketch the Tanner graph which represents the turbo code.

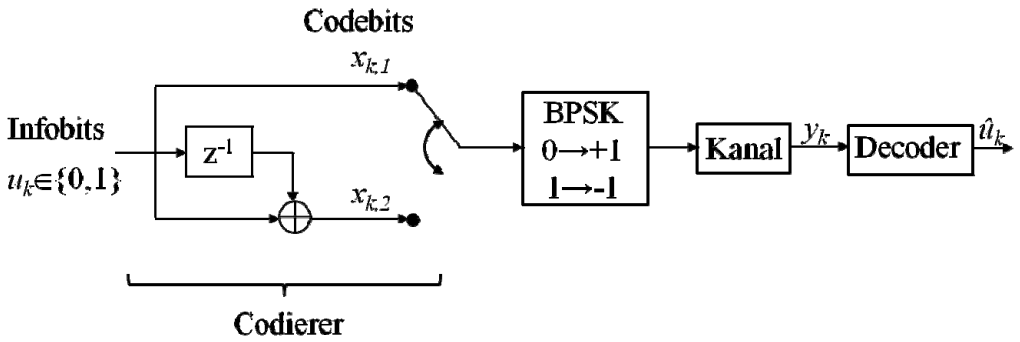
q) Will message passing perform well for the given turbo code? Give reasons.

Chapter 8

Convolutional Code

8.1

The encoder of a convolutional code is given in the following figure. (This code shall be used in all parts of the problem except for part d).)



At the beginning the encoder is in state "0", i.e. the memory element is initialized with "0". The convolutional code is terminated in state "0".

a) Is the encoder systematic? Explain your answer!

b) What is the code rate R of the convolutional code?

c) Determine the generator polynomials $g_n(D)$ of the encoder?

d) Determine the generator polynomials of a recursive encoder, which generates the same code. Sketch the recursive encoder.

e) How many information bits are necessary to terminate the code in a certain state?

f) Sketch the state diagram of the convolutional code. Label the diagram clearly and completely.

g) Sketch a trellis-segment with all possible transitions. Label the diagram clearly and completely.

h) Determine the free distance d_f of the code.

i) Determine the maximum length of an error burst, which allows error-free decoding with hard decision decoding.

j) Consider an information word of length 2 bit, i.e. $\mathbf{u} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T$. (Note: Consider that the code is terminated in state "0"!)

   (1) State all possible codewords $\mathbf{x}$ and assign them in a table to the corresponding information words $\mathbf{u} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T$.

   (2) Is the code linear? Explain your answer!

   (3) Determine a generator matrix $\mathbf{G}$ of a block code, that generates the same code.

   (4) Determine the code rate R_B of the block code?

8.2 Convolutional Code of the GSM-mobile System

For the transmission of speech, the GSM system uses a binary convolutional code with the following generator polynomials:

$$g_1(D) = 1 + D^3 + D^4$$
$$g_2(D) = 1 + D + D^3 + D^4$$

a) Sketch the encoder (shift register).

b) Determine the memory M and the constraint length C of the convolutional code.

c) Determine the code rate R of the convolutional code.

d) Is the encoder systematic? Explain your answer!

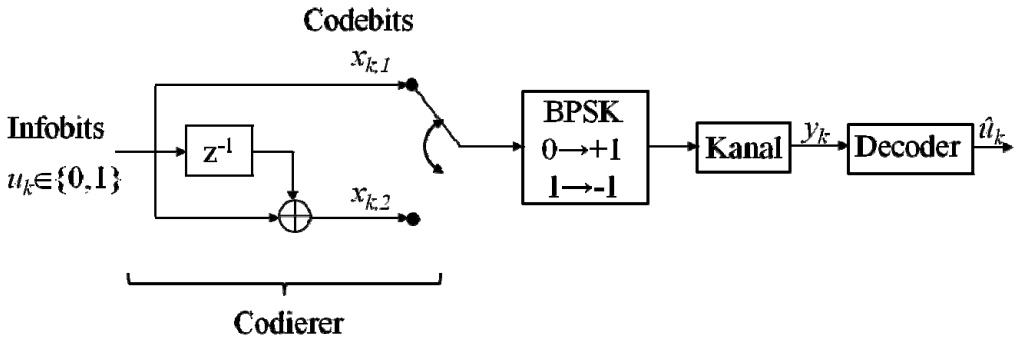
e) Sketch a trellis segment with all possible transitions. Label the figure completely! In particular, define the states of the trellis and the variables which determine the transitions clearly.

f) Determine the metric increment of a Viterbi decoder with soft-decision maximum-likelihood (ML)-decoding for an AWGN-channel.

g) Determine the free distance d_f of the code.

8.3

Consider the convolutional code, which is determined by the following picture.



The encoder is initialized in state "0", i.e. all memory elements contain "0". The convolutional code will be terminated in state "0".

a) State the generator polynomials $g_n(D)$ of the encoder.

b) Describe the generator polynomials of the encoder in octal representation.

c) Sketch a trellis segment with all possible transitions. Label the figure completely.

d) For an AWGN-channel with noise variance σ_N^2, derive the expression for the metric increment of a Viterbi-decoder with

   (1) soft-decision MAP-decoding

   (2) soft-decision maximum-likelihood decoding

   using log-likelihood ratios (L-values). Assume that the code bits $\mathbf{x_k}$ are transmitted with BPSK-modulation, i.e. $x_k \in \{\pm 1\}$.

e) At the input of the decoder, the sequence

$$\mathbf{y} = \begin{bmatrix} +0.5 & -2.0 & -1.0 & +1.5 \end{bmatrix}^T$$

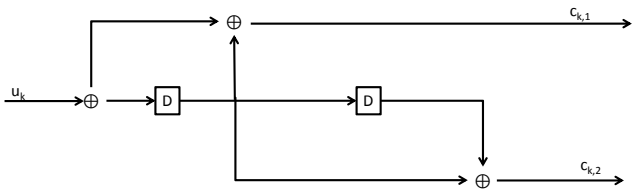
   is observed.

   Performe the Viterbi-algorithm for Soft-Decision Maximum-Likelihood decoding and determine the decoded information sequence. Make sure that all steps are clearly explained.

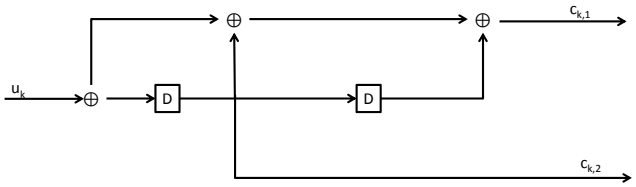
8.4

Consider the following three different shift registers.

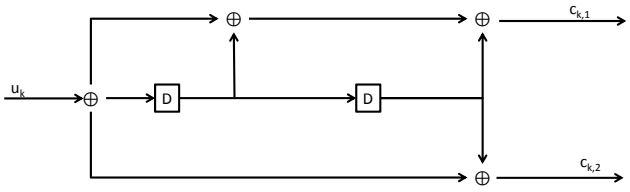
1)



2)



3)



a) Determine the generator polynomials of the convolutional encoder 1) and 2)?

b) Are the shift register 1) and 2) catastrophic encoders? Explain your answer! What is the major problem of catastrophic encoders?

c) Determine the free distance of the convolutional codes 2) and 3)!

d) Which of the shift registers 2) or 3) would you prefer? Explain your answer!

e) What is the rate R_3 of the convolutional code 3)?

f) Determine the constraint length C_3 and the memory M_3 of the convolutional code 3)?

g) Sketch the shift register of the equivalent recursive convolutional encoder, that generates the same code as the shift register 3).

h) State the expression for the metric increment of a Viterbi-decoder with soft-decision maximum-likelihood decoding (ML) decoding for an AWGN channel.

8.5

Consider the following generator polynomials of a convolutional code:

$$g_1(D) = 1 + D + D^2 + D^3$$
$$g_2(D) = 1 + D + D^3$$

a) Determine the code rate R.

b) Determine the memory M and the constraint length C of the convolutional code.

c) Determine the octal representation of the convolutional code.

d) Determine an upper bound on the free distance d_f using the generator polynomials.

e) Check if the encoder is catastrophic or not. What characterizes a catastrophic encoder?

f) Determine the generator polynomials of a recursive systematic encoder, that generates the same code. $g_1(D)$ is supposed to define the feedback.

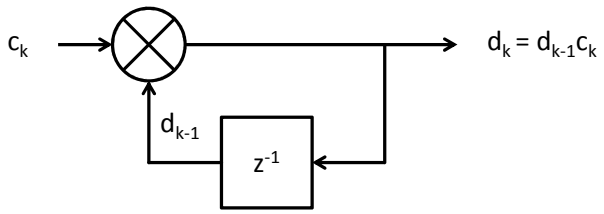
g) Sketch the recursive-systematic encoder of part f) as a shift register.

h) Sketch the encoder of a Turbo-Code, that uses recursive-systematic convolutional codes as component codes.

i) Which fundamental difference between a recursive and a non-recursive convolutional code explains why recursive-systematic convolutional codes are particularly suitable for Turbo-Codes? Give an explanation.

8.6 Differential Modulation with Viterbi Detector

A differential BPSK (D-BPSK) modulator as depicted in the following figure can be viewed as a recursive convolutional encoder with rate $R = 1$. Hint: $c_k \in \{\pm 1\}$.



a) Determine the state transition diagram of the differential modulator.

b) Determine the trellis butterfly of the differential modulator.

c) Determine the free distance of the differential modulation scheme.

d) Sketch the trellis path for the transmit sequence $\mathbf{c} = \begin{bmatrix} +1 & +1 & -1 & -1 & +1 \end{bmatrix}^T$. Assume that the differential encoder is initialized in state $+1$ at the beginning of the sequence. Which sequence $\mathbf{d}$ is transmitted?

e) Derive the metric increments of an APP and an ML Viterbi detector for differential modulation using log-likelihood ratios.

Assume that the sequence $\mathbf{c}$ from d) has been transmitted through an AWGN channel with an SNR of 10 dB. The received sequence is given by

$$\mathbf{y} = \begin{bmatrix} +2.0 & -0.1 & -0.5 & +0.1 & +1.0 \end{bmatrix}^T$$

The source statistics are given by $P(c_k = +1) = \frac{1}{4}$, $P(c_k = -1) = \frac{3}{4}$.

f) Determine the estimated sequence using the Viterbi algorithm for the following cases:

(1) Hard decision maximum likelihood (ML) sequence estimation.

(2) Hard decision maximum a posteriori probability (MAP) sequence estimation.

(3) Soft decision maximum likelihood (ML) sequence estimation.

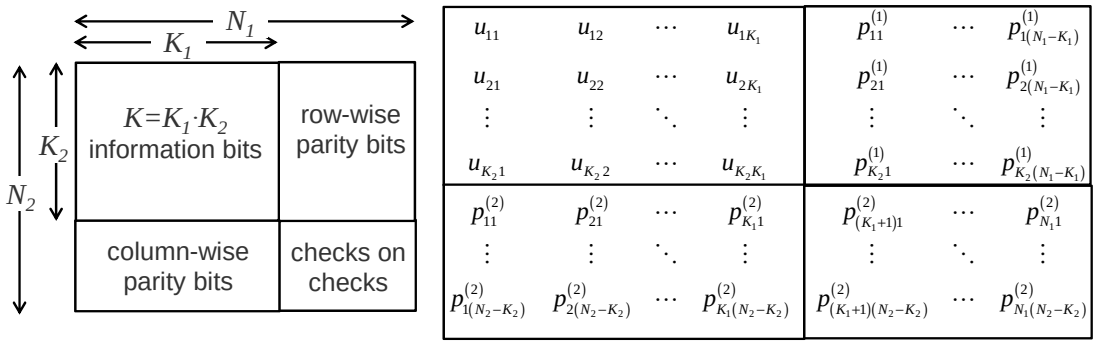
(4) Soft decision maximum a posteriori probability (MAP) sequence estimation.

Chapter 9

Special Block Codes

9.1 Product Code

The idea of a product code is to construct a powerful $(N,\ K)$ code by combining simple systematic linear $(N_i,\ K_i,\ d_{min,i})$ constituent block codes. The principle is depicted in the following figure:



The $K = K_1 \cdot K_2$ information bits are arranged in a $(K_1 \times K_2)$ matrix. Then, we encode first the rows using code 1. In a second step we encode the columns using code 2. The parity bits are appended as indicated in the figure. Furthermore, we encode the parity bits of code 1 column-wise using code 2 ("parity checks on parity checks").

a) What is the code rate R of the product code depending on the parameters N_i and K_i of the constituent codes?

b) Determine the Singleton bound on the minimum Hamming distance d_{min} of the product code depending on the parameters N_i and K_i of the constituent codes.

c) Show, that the minimum Hamming distance d_{min} of the product code is given by the product $d_{min,1} \cdot d_{min,2}$ of the minimum Hamming distances of the constituent codes.

d) Assume that the constituent code i can correct a burst error up to length t_i. What is the maximum length t of a burst error which can be corrected by the product code using a trivial decoding scheme which decodes successively column- and row-wise?

e) Assume that the constituent code i can correct up to t_i single errors. How many single errors can at least be corrected by the product code?

We now assume that both constituent codes are $(7,\ 4,\ 3)_2$ Hamming codes.

f) Determine the parameters $(N,\ K,\ d_{min})_q$ of the product code.

g) How many errors should this product code correct according to the result of e)?

h) Assume that four bit errors have occurred which are located in a square as indicated in the following figure. Is an error correction possible using a trivial decoding scheme which decodes successively column- and row-wise?

We now assume that both constituent codes are $(3,\ 2)_2$ single parity check codes.

i) Determine the parameters $(N,\ K,\ d_{min})_q$ of the product code.

j) Determine the codeword for the info sequence $\mathbf{u} = \begin{bmatrix}0 & 1 & 1 & 1\end{bmatrix}^T$.

k) How many errors can be corrected by a trivial decoding scheme which decodes successively column- and row-wise. Explain, how the errors can be localized.

l) How many candidate codewords have to be taken into consideration by a maximum likelihood (ML) decoder for the overall product code?

9.2 Sum Construction of Linear Block Codes

A powerful linear block code $\mathcal{C}$ is to be constructed of two simple linear block codes $\mathcal{C}_1$ and $\mathcal{C}_2$ with the same codeword length $N_1 = N_2 = N$. $\mathcal{C}_1$ is a $(N,\ K_1,\ d_{min,1})_2$ code with generator matrix $\mathbf{G}_1$, $\mathcal{C}_2$ is a $(N,\ K_2,\ d_{min,2})_2$ code with generator matrix $\mathbf{G}_2$.
We define the following vectors:

Codeword of code $\mathcal{C}_i$: $\mathbf{x}_i = \begin{bmatrix} x_{i,1} \\ \vdots \\ x_{i,N} \end{bmatrix}$, $i = 1,\ 2$.

Information word of code $\mathcal{C}_i$: $\mathbf{u}_i = \begin{bmatrix} u_{i,1} \\ \vdots \\ u_{i,K_i} \end{bmatrix}$, $i = 1,\ 2$.

First we consider the following construction of code $\mathcal{C}$:

$$\mathcal{C} = \left\{ \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \middle| \mathbf{x}_1 \in \mathcal{C}_1, \mathbf{x}_2 \in \mathcal{C}_2 \right\}$$

a) State the codeword length N', the information word length K and the code rate R of code $\mathcal{C}$.

b) Determine the generator matrix $\mathbf{G}$ of code $\mathcal{C}$ dependent on the generator matrices $\mathbf{G}_1$ and $\mathbf{G}_2$.

c) Determine the minimum Hamming distance d_{min} of code $\mathcal{C}$ dependent on $d_{min,1}$ and $d_{min,2}$.

d) Determine the minimum Hamming distance d_{min} of code $\mathcal{C}$ dependent on $d_{min,1}$ and $d_{min,2}$ for the special case, that the code C_2 is a repetition code of rate $R_2 = \frac{1}{N}$.

From now on we consider the following construction of the code $\mathcal{C}$:

$$\mathcal{C} = \left\{ \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_1 \oplus \mathbf{x}_2 \end{bmatrix} \middle| \mathbf{x}_1 \in \mathcal{C}_1, \mathbf{x}_2 \in \mathcal{C}_2 \right\} \tag{9.1}$$

where $\oplus$ is the modulo-2 addition.

e) Determine the generator matrix $\mathbf{G}$ of code $\mathcal{C}$ dependent of the generator matrices $\mathbf{G}_1$ and $\mathbf{G}_2$.

f) Determine the minimum Hamming distance of code $\mathcal{C}$ dependent on $d_{min,1}$ and $d_{min,2}$.

Help:

- First consider the special cases $\mathbf{x}_1 = \mathbf{0}$ and $\mathbf{x}_2 = \mathbf{0}$.
- Then consider all other cases using the triangle inequality: $w_H(\mathbf{x} \oplus \mathbf{y}) \leq w_H(\mathbf{x}) + w_H(\mathbf{y})$.

g) Determine the minimum Hamming distance d_{min} of code $\mathcal{C}$ dependent on $d_{min,1}$ and $d_{min,2}$ for the special case, that the code C_2 is a repetition code with rate $R_2 = \frac{1}{N}$.

h) Compare the results of part g) with the results of part d). Which code construction generates the more powerful code? Explain your answer!

9.3 Random Codes

When Shannon published his famous paper in 1948, channel codes did not exist. Shannon's theorems have been proven based on the idea of random codes. In this problem, we will address binary random codes, which map information words of length K bits to codeword of length N bits. A random code is constructed by randomly choosing the codewords among all possible binary sequences of length N.

a) Construct a random code of rate $R = \frac{1}{2}$ with codeword length $N = 6$. State all the codewords and the mapping of information words to codewords in a table.

b) How many different binary sequences of length N exist?

c) How many different binary sequences of length N with Hamming weight w exist?

d) How many codewords exist for a binary $(N, K)_2$ code ?

e) Assume, that a random code has been constructed. We then randomly choose a binary sequence of length N. Determine the probability, that this randomly chosen sequence is a codeword.

f) Determine the expected number $E\{A_w\}$ of codewords with Hamming weight w for a random $(N, K)_2$ code.

g) Is a random code in general a linear code ? Give reasons!

h) Which minimum Hamming distance $d_{\min}$ can be guaranteed by the random code construction?

i) Determine the minimum Hamming weight $w_{\min}$ of your code from a).

j) Determine the minimum Hamming distance $d_{\min}$ of your code from a).