Principles Of Digital Communication
A Top-Down Approach
Problem Manual

Bixio Rimoldi
School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland

 Version July 1, 2016

Contents

1 Introduction and objectives

Problem 1. *(Probabilities of basic events)*

Assume that X_1 and X_2 are independent random variables that are uniformly distributed in the interval $[0, 1]$. Compute the probability of the following events.

Hint: For each event, identify the corresponding region inside the unit square.

(a) $0 \leq X_1 - X_2 \leq \frac{1}{3}$

(b) $X_1^3 \leq X_2 \leq X_1^2$

(c) $X_2 - X_1 = \frac{1}{2}$

(d) $\left(X_1 - \frac{1}{2}\right)^2 + \left(X_2 - \frac{1}{2}\right)^2 \leq \left(\frac{1}{2}\right)^2$

(e) Given that $X_1 \geq \frac{1}{4}$, compute the probability that $\left(X_1 - \frac{1}{2}\right)^2 + \left(X_2 - \frac{1}{2}\right)^2 \leq \left(\frac{1}{2}\right)^2$

Problem 2. *(Basic probabilities)*

(a) A box contains m white and n black balls. Suppose k balls are drawn. Find the probability of drawing at least one white ball.

(b) We have two coins; the first is fair and the second is two-headed. We pick one of the coins at random, toss it twice and obtain heads both times. Find the probability that the coin is fair.

Problem 3. *(Conditional distribution)*

Assume that X and Y are random variables with joint density function

$$f_{X,Y}(x, y) = \begin{cases} A, & 0 \leq x < y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Are X and Y independent?

(b) Find the value of A.

(c) Find the density function of Y. Do this first by arguing informally using a sketch of $f_{X,Y}(x, y)$, then compute the density formally.

(d) Find $\mathbb{E}[X|Y = y]$.

Hint: Try to find it from a sketch of $f_{X,Y}(x, y)$.

(e) The $\mathbb{E}[X|Y = y]$ found in part (d) is a function of y, call it $f(y)$. Find $\mathbb{E}[f(Y)]$.

(f) Find $\mathbb{E}[X]$ (from the definition) and compare it to the $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[f(Y)]$ that you have found in part (e).

PROBLEM 4. *(Playing darts)*

Assume that you are throwing darts at a target. We assume that the target is one-dimensional, i.e. that the darts all end up on a line. The "bull's eye" is in the center of the line, and we give it the coordinate 0. We assume that the position X_1 of a dart that lands on the target is a random variable that has a Gaussian distribution with variance σ_1^2 and mean 0. Assume now that there is a second target, which is further away. If you throw a dart to that target, the position X_2 has a Gaussian distribution with variance σ_2^2 (where $\sigma_2^2 > \sigma_1^2$) and mean 0. You play the following game: You toss a "coin" which gives you $Z = 1$ with probability p and $Z = 0$ with probability $1 - p$ for some fixed $p \in [0, 1]$. If $Z = 1$, you throw a dart onto the first target. If $Z = 0$, you aim for the second target instead. Let X be the relative position of the dart with respect to the center of the target that you have chosen.

(a) Write down X in terms of X_1, X_2 and Z.

(b) Compute the variance of X. Is X Gaussian?

(c) Let $S = |X|$ be the score, which is given by the distance of the dart to the center of the target (that you picked using the coin). Compute the average score $\mathbb{E}[S]$.

PROBLEM 5. *(Uncorrelated vs. independent random variables)*

(a) Let X and Y be two continuous real-valued random variables with joint probability density function f_{XY}. Show that if X and Y are independent, they are also uncorrelated.

(b) Consider two independent and uniformly distributed random variables $U \in \{0, 1\}$ and $V \in \{0, 1\}$. Assume that X and Y are defined as follows: $X = U + V$ and $Y = |U - V|$. Are X and Y independent? Compute the covariance of X and Y. What do you conclude?

PROBLEM 6. *(Monty Hall)*

Assume you are participating in a quiz show. You are shown three boxes that look identical from the outside, except they have labels 0, 1, and 2, respectively. Only one of them contains one million Swiss francs, the other two contain nothing. You choose one box at random with a uniform probability. Let A be the random variable that denotes your choice, $A \in \{0, 1, 2\}$.

(a) What is the probability that the box A contains the money?

(b) The quizmaster of course knows which box contains the money. He opens one of the two boxes that you did not choose, being careful not to open the one that contains the money. Now, you know that the money is either in A (your first choice) or in B (the only other box that could contain the money). What is the probability that B contains the money?

(c) If you are now allowed to change your mind, i.e. choose B instead of sticking with A, would you do it?

2 Receiver design for discrete-time observations: First layer

PROBLEM 1. *(Hypothesis testing: Uniform and uniform)*

Consider a binary hypothesis testing problem in which the hypotheses $H = 0$ and $H = 1$ occur with probability $P_H(0)$ and $P_H(1) = 1 - P_H(0)$, respectively. The observable Y takes values in $\{0,1\}^{2k}$, where k is a fixed positive integer. When $H = 0$, each component of Y is 0 or 1 with probability $\frac{1}{2}$ and components are independent. When $H = 1$, Y is chosen uniformly at random from the set of all sequences of length $2k$ that have an equal number of ones and zeros. There are $\binom{2k}{k}$ such sequences.

(a) What is $P_{Y|H}(y|0)$? What is $P_{Y|H}(y|1)$?

(b) Find a maximum-likelihood decision rule for H based on y. What is the single number you need to know about y to implement this decision rule?

(c) Find a decision rule that minimizes the error probability.

(d) Are there values of $P_H(0)$ such that the decision rule that minimizes the error probability always chooses the same hypothesis regardless of y? If yes, what are these values, and what is the decision?

PROBLEM 2. *(The "Wetterfrosch")*

Let us assume that a "weather frog" bases his forecast of tomorrow's weather entirely on today's air pressure. Determining a weather forecast is a hypothesis testing problem. For simplicity, let us assume that the weather frog only needs to tell us if the forecast for tomorrow's weather is "sunshine" or "rain". Hence we are dealing with binary hypothesis testing. Let $H = 0$ mean "sunshine" and $H = 1$ mean "rain". We will assume that both values of H are equally likely, i.e. $P_H(0) = P_H(1) = \frac{1}{2}$. For the sake of this exercise, suppose that on a day that precedes sunshine, the pressure may be modeled as a random variable Y with the following probability density function:

$$f_{Y|H}(y|0) = \begin{cases} A - \frac{A}{2}y, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, the pressure on a day that precedes a rainy day is distributed according to

$$f_{Y|H}(y|1) = \begin{cases} B + \frac{B}{3}y, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

The weather frog's purpose in life is to guess the value of H after measuring Y.

(a) Determine A and B.

(b) Find the *a posteriori probability* $P_{H|Y}(0|y)$. Also find $P_{H|Y}(1|y)$.

(c) Show that the implementation of the decision rule $\hat{H}(y) = \arg\max_i P_{H|Y}(i|y)$ reduces to

$$\hat{H}_\theta(y) = \begin{cases} 0, & \text{if } y \le \theta \\ 1, & \text{otherwise} \end{cases}$$

for some threshold θ and specify the threshold's value.

(d) Now assume that the weather forecaster does not know about hypothesis testing and arbitrarily chooses the decision rule $\hat{H}_\gamma(y)$ for some arbitrary $\gamma \in \mathbb{R}$. Determine, as a function of γ, the probability that the decision rule decides $\hat{H} = 1$ given that $H = 0$. This probability is denoted $\Pr\{\hat{H}(Y) = 1 | H = 0\}$.

(e) For the same decision rule, determine the probability of error $P_e(\gamma)$ as a function of γ. Evaluate your expression at $\gamma = \theta$.

(f) Using calculus, find the γ that minimizes $P_e(\gamma)$ and compare your result to θ.

PROBLEM 3. *(Hypothesis testing in Laplacian noise)*

Consider the following hypothesis testing problem between two equally likely hypotheses. Under hypothesis $H = 0$, the observable Y is equal to $a + Z$ where Z is a random variable with Laplacian distribution

$$f_Z(z) = \frac{1}{2} e^{-|z|}$$

Under hypothesis $H = 1$, the observable is given by $-a + Z$. You may assume that a is positive.

(a) Find and draw the density $f_{Y|H}(y|0)$ of the observable under hypothesis $H = 0$, and the density $f_{Y|H}(y|1)$ of the observable under hypothesis $H = 1$.

(b) Find the decision rule that minimizes the probability of error.

(c) Compute the probability of error of the optimal decision rule.

PROBLEM 4. *(Poisson parameter estimation)*

Two hypotheses $H = 0$ and $H = 1$ occur with probabilities $P_H(0) = p_0$ and $P_H(1) = 1 - p_0$, respectively. The observable Y takes values in the set of non-negative integers. Under hypothesis $H = 0$, Y is distributed according to a Poisson law with parameter λ_0, i.e.

$$P_{Y|H}(y|0) = \frac{\lambda_0^y}{y!} e^{-\lambda_0}$$

Under hypothesis $H = 1$,

$$P_{Y|H}(y|1) = \frac{\lambda_1^y}{y!} e^{-\lambda_1}$$

This is a model for the reception of photons in optical communication.

(a) Derive the MAP decision rule by indicating likelihood and log-likelihood ratios.

Hint: The direction of an inequality changes if both sides are multiplied by a negative number.

(b) Derive an expression for the probability of error of the MAP decision rule.

(c) For $p_0 = \frac{1}{3}$, $\lambda_0 = 2$ and $\lambda_1 = 10$, compute (using a computer) the probability of error of the MAP decision rule.

(d) Repeat (c) with $\lambda_1 = 20$ and comment.

Problem 5. *(Lie detector)*

You are asked to develop a "lie detector" and analyze its performance. Based on the observation of brain-cell activity, your detector has to decide if a person is telling the truth or is lying. For the purpose of this exercise, the brain cell produces a sequence of spikes. For your decision you may use only a sequence of n consecutive inter-arrival times $Y_1, Y_2, \ldots, Y_n$. Hence Y_1 is the time elapsed between the first and second spike, Y_2 the time between the second and third, etc. We assume that, a priori, a person lies with some known probability p. When the person is telling the truth, $Y_1, \ldots, Y_n$ is an i.i.d. sequence of exponentially distributed random variables with intensity α, $(\alpha > 0)$, i.e.

$$f_{Y_i}(y) = \alpha e^{-\alpha y}, \quad y \geq 0$$

When the person lies, $Y_1, \ldots, Y_n$ is i.i.d. exponentially distributed with intensity β, $(\alpha < \beta)$.

(a) Describe the decision rule of your lie detector for the special case $n = 1$. Your detector should be designed so as to minimize the probability of error.

(b) What is the probability $P_{L|T}$ that your lie detector says that the person is lying when the person is telling the truth?

(c) What is the probability $P_{T|L}$ that your test says that the person is telling the truth when the person is lying.

(d) Repeat (a) and (b) for a general n.
hintWhen $Y_1, \ldots, Y_n$ is a collection of i.i.d. random variables that are exponentially distributed with parameter $\alpha > 0$, then $Y_1 + \cdots + Y_n$ has the probability density function of the Erlang distribution, i.e.

$$f_{Y_1+\cdots+Y_n}(y) = \frac{\alpha^n}{(n-1)!} y^{n-1} e^{-\alpha y}, \quad y \geq 0$$

Problem 6. *(Fault detector)*

As an engineer, you are required to design the test performed by a fault detector for a "black-box" that produces a sequence of i.i.d. binary random variables $\ldots, X_1, X_2, X_3, \ldots$. Previous experience shows that this "black-box" has an a priori failure probability of $\frac{1}{1025}$. When the "black-box" works properly, $p_{X_i}(1) = p$. When it fails, the output symbols are equally likely to be 0 or 1. Your detector has to decide based on the observation of the past 16 symbols, i.e. at time k the decision will be based on $X_{k-16}, \ldots, X_{k-1}$.

(a) Describe your test.

(b) What does your test decide if it observes the sequence 0101010101010101? Assume that $p = 0.25$.

Problem 7. *(Multiple choice exam)*

You are taking a multiple choice exam. Question number 5 allows for two possible answers. According to your first impression, answer 1 is correct with probability $\frac{1}{4}$ and answer 2 is correct with probability $\frac{3}{4}$. You would like to maximize your chance of giving the correct answer and you decide to have a look at what your neighbors on the left and right have to say. The neighbor on the left has answered $\hat{H}_L = 1$. He is an excellent student who has a record of being correct 90% of the time when asked a binary question. The neighbor on the right has answered $\hat{H}_R = 2$. He is a weaker student who is correct 70% of the time.

(a) You decide to use your first impression as a prior and to consider $\hat{H}_L$ and $\hat{H}_R$ as observations. Formulate the decision problem as a hypothesis testing problem.

(b) What is your answer $\hat{H}$?

Problem 8. *(MAP decoding rule: Alternative derivation)*

Consider the binary hypothesis testing problem where H takes values in $\{0, 1\}$ with probabilities $P_H(0)$ and $P_H(1)$. The conditional probability density function of the observation $Y \in \mathbb{R}$ given $H = i$, $i \in \{0, 1\}$ is given by $f_{Y|H}(\cdot|i)$. Let $\mathcal{R}_i$ be the decoding region for hypothesis i, i.e. the set of y for which the decision $\hat{H} = i$, $i \in \{0, 1\}$.

(a) Show that the probability of error is given by

$$P_e = P_H(1) + \int_{\mathcal{R}_1} \big(P_H(0) f_{Y|H}(y|0) - P_H(1) f_{Y|H}(y|1)\big)\, dy$$

Hint: Note that $\mathbb{R} = \mathcal{R}_0 \cup \mathcal{R}_1$ and $\int_{\mathbb{R}} f_{Y|H}(y|i) dy = 1$ for $i \in \{0, 1\}$.

(b) Argue that P_e is minimized when

$$\mathcal{R}_1 = \big\{y \in \mathbb{R} : P_H(0) f_{Y|H}(y|0) < P_H(1) f_{Y|H}(y|1)\big\}$$

i.e. for the MAP rule.

Problem 9. *(Independent and identically distributed vs. first-order Markov)*

Consider testing two equally likely hypotheses $H = 0$ and $H = 1$. The observable $Y = (Y_1, \ldots, Y_k)^\mathsf{T}$ is a k-dimensional binary vector. Under $H = 0$ the components of the vector Y are independent uniform random variables (also called Bernoulli $\left(\frac{1}{2}\right)$ random variables). Under $H = 1$, the component Y_1 is also uniform, but the components Y_i, $2 \le i \le k$, are distributed as follows:

$$P_{Y_i|Y_1,\ldots,Y_{i-1}}(y_i|y_1, \ldots, y_{i-1}) = \begin{cases} 3/4, & \text{if } y_i = y_{i-1} \\ 1/4 & \text{otherwise} \end{cases}$$

(a) Find the decision rule that minimizes the probability of error.
Hint: Write down a short sample sequence $(y_1, \ldots, y_k)$ and determine its probability under each hypothesis. Then generalize.

(b) Give a simple sufficient statistic for this decision. (For the purpose of this question, a sufficient statistic is a function of y with the property that a decoder that observes y can not achieve a smaller error probability than a MAP decoder that observes this function of y.)

(c) Suppose that the observed sequence alternates between 0 and 1 except for one string of ones of length s, i.e. the observed sequence y looks something like

$$y = 0101010111111\ldots111111010101$$

What is the least s such that we decide for hypothesis $H = 1$?

PROBLEM 10. *(SIMO channel with Laplacian noise)*

One of the two signals $c_0 = -1$, $c_1 = 1$ is transmitted over the channel shown in Figure 1a. The two noise random variables Z_1 and Z_2 are statistically independent of the transmitted signal and of each other. Their density functions are

$$f_{Z_1}(\alpha) = f_{Z_2}(\alpha) = \frac{1}{2}e^{-|\alpha|}$$

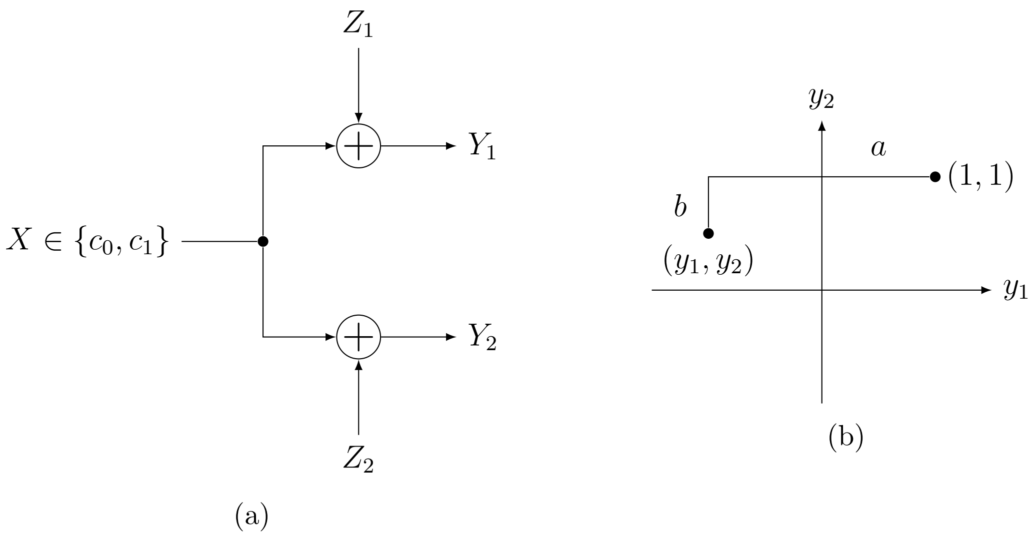


Figure 1

(a) Derive a maximum likelihood decision rule.

(b) Describe the maximum likelihood decision regions in the (y_1, y_2) plane. Describe also the "either choice" regions, i.e. the regions where it does not matter if you decide for c_0 or for c_1.
Hint: Use geometric reasoning and the fact that for a point (y_1, y_2) as shown in Figure 1b, $|y_1 - 1| + |y_2 - 1| = a + b$.

(c) A receiver decides that c_1 was transmitted if and only if $(y_1 + y_2) > 0$. Does this receiver minimize the error probability for equally likely messages?

(d) What is the error probability of the receiver in (c)?
Hint: One way to do this is to use the fact that if $W = Z_1 + Z_2$, then $f_W(w) = \frac{e^{-\omega}}{4}(1 + \omega)$ for $\omega > 0$ and $f_W(-\omega) = f_W(\omega)$.

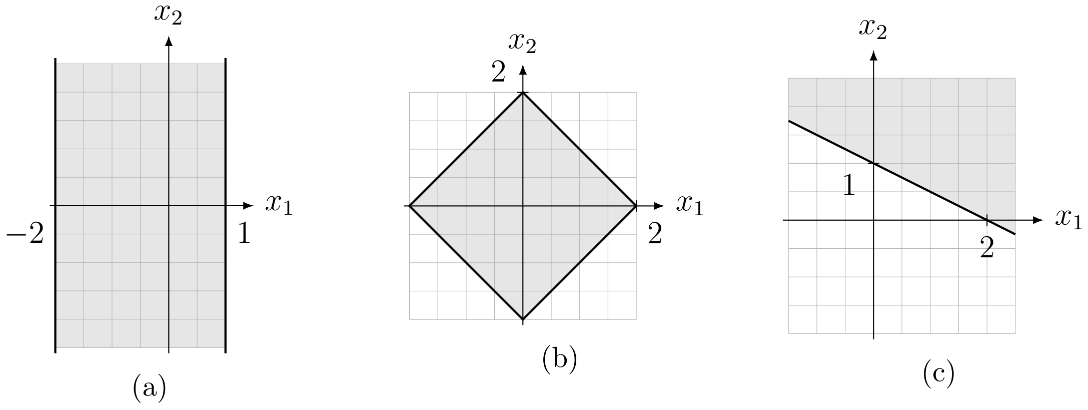


Figure 2

Problem 11. *(Q function on regions)*

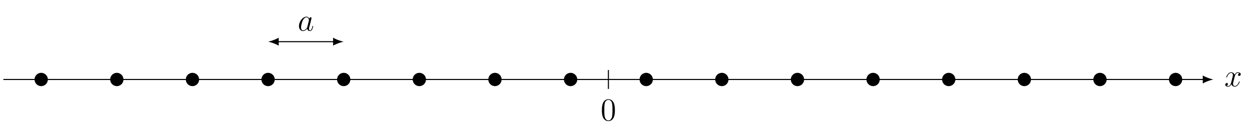
Let $X \sim \mathcal{N}(0, \sigma^2 I_2)$. For each of the three diagrams shown in Figure 2, express the probability that X lies in the shaded region. You may use the Q function when appropriate.

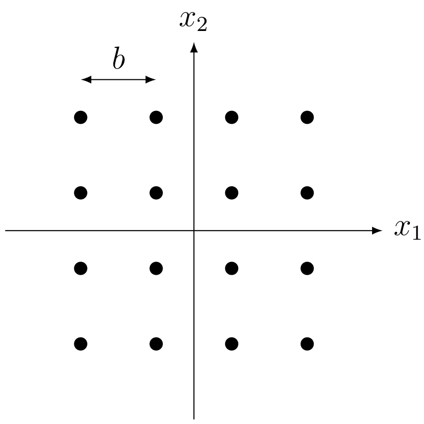
Problem 12. *(Properties of the Q function)*

Prove properties (a) through (d) of the Q function defined in Section 2.3.
Hint: For property (d), multiply and divide inside the integral by the integration variable and integrate by parts. By upper- and lower-bounding the resulting integral, you will obtain the lower and upper bound.

Problem 13. *(16-PAM vs. 16-QAM)*

The two signal constellations shown below are used to communicate across an additive white Gaussian noise channel. Let the noise variance be σ_2. Each point represents a codeword c_i for some i. Assume each codeword is used with the same probability.





(a) For each signal constellation, compute the average probability of error P_e as a function of the parameters a and b, respectively.

(b) For each signal constellation, compute the average energy per symbol $\mathcal{E}$ as a function of parameters a and b, respectively:

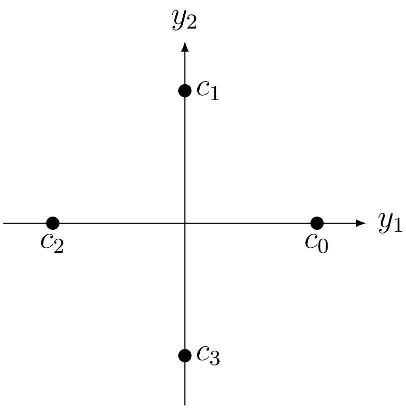
$$\mathcal{E} = \sum_{i=1}^{16} P_H(i) \|c_i\|^2$$

In the next chapter it will become clear in what sense $\mathcal{E}$ relates to the energy of the transmitted signal (see Example 3.2 and the discussion that follows).

(c) Plot P_e versus $\frac{\mathcal{E}}{\sigma^2}$ for both signal constellations and comment.

Problem 14. *(QPSK decision regions)*

Let $H \in \{0, 1, 2, 3\}$ as assume that when $H = i$ you transmit the codeword c_i shown below. Under $H = i$, the receiver observes $Y = c_i + Z$.



(a) Draw the decoding regions assuming that $Z \sim \mathcal{N}(0, \sigma^2 I_2)$ and that $P_H(i) = 1/4$, $i \in \{0, 1, 2, 3\}$.

(b) Draw the decoding regions (qualitatively) assuming $Z \sim \mathcal{N}(0, \sigma^2 I_2)$ and $P_H(0) = P_H(2) > P_H(1) = P_H(3)$. Justify your answer.

(c) Assume again that $P_H(i) = 1/4$, $i \in \{0, 1, 2, 3\}$ and that $Z \sim \mathcal{N}(0, K)$, where $K = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 4\sigma^2 \end{pmatrix}$. How do you decode now?

Problem 15. *(Antenna array)*

The following problem relates to the design of multi-antenna systems. Consider the binary equiprobable hypothesis testing problem:

$$H = 0 : Y_1 = A + Z_1, \quad Y_2 = A + Z_2$$

$$H = 1 : Y_1 = -A + Z_1, \quad Y_2 = -A + Z_2$$

where Z_1, Z_2 are independent Gaussian random variables with different variances $\sigma_1^2 \neq \sigma_2^2$, that is, $Z_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $Z_2 \sim \mathcal{N}(0, \sigma_2^2)$. $A > 0$ is a constant.

(a) Show that the decision rule that minimizes the probability of error (based on the observable Y_1 and Y_2) can be stated as

$$\sigma_2^2 y_1 + \sigma_1^2 y_2 \underset{1}{\overset{0}{\gtrless}} 0$$

(b) Draw the decision regions in the (Y_1, Y_2) plane for the special case where $\sigma_1 = 2\sigma_2$.

(c) Evaluate the probability of the error for the optimal detector as a function of σ_1^2, σ_2^2 and A.

Problem 16. *(Multi-antenna receiver)*

Consider a communication system with one transmitter and n receiver antennas. The receiver observes the n-tuple $Y = (Y_1, \dots, Y_n)^\mathsf{T}$ with

$$Y_k = Bg_k + Z_k, \quad k = 1, 2, \dots, n$$

where $B \in \{\pm 1\}$ is a uniformly distributed source bit, g_k models the gain of antenna k and $Z_k \sim \mathcal{N}(0, \sigma^2)$. The random variables $B, Z_1, \dots, Z_n$ are independent. Using n-tuple notation the model becomes

$$Y = Bg + Z,$$

where Y, g, and Z are n-tuples.

(a) Suppose that the observation Y_k is weighted by an arbitrary real number w_k and combined with the other observations to form

$$V = \sum_{k=1}^{n} Y_k w_k = \langle Y, w \rangle,$$

where w is an n-tuple. Describe the ML receiver for B given the observation V. (The receiver knows g and of course knows w.)

(b) Give an expression for the probability of error P_e.

(c) Define $\beta = \frac{|\langle g, w \rangle|}{\|g\| \|w\|}$ and rewrite the expression for P_e in a form that depends on w only through β.

(d) As a function of w, what are the maximum and minimum values for β and how do you choose w to achieve them?

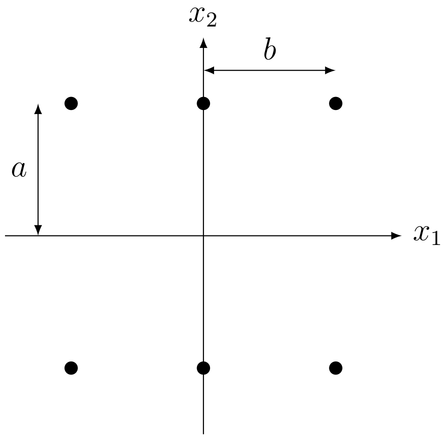
(e) Minimize the probability of error over all possible choices of w. Could you reduce the error probability further by doing ML decision directly on Y rather than of V? Justify your answer.

(f) How would you choose w to minimize the error probability if Z_k had variance σ_k^2, $k = 1, \dots, n$?

Hint: With a simple operation at the receiver you can transform the new problem into the one you have already solved.

Problem 17. *(Signal constellation)*

The signal constellation shown below is used to communicate across the AWGN channel of noise variance σ^2. Assume that the six signals are used with equal probability.



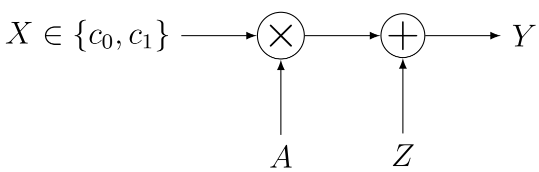
(a) Draw the boundaries of the decision regions.

(b) Compute the average probability of error, P_e, for this signal constellation.

(c) Compute the average energy per symbol for this signal constellation.

Problem 18. *(Hypothesis testing and fading)*

Consider the following communication problem where there are two equiprobable hypotheses. When $H = 0$, we transmit $c_0 = -b$, where b is an arbitrary but fixed positive number. When $H = 1$, we transmit $c_1 = b$. The channel is as shown below, where $Z \sim \mathcal{N}(0, \sigma^2)$ represents noise, $A \in \{0, 1\}$ represents a random attenuation (fading) with $P_A(0) = \frac{1}{2}$, and Y is the channel output. The random variables H, A, and Z are independent.



(a) Find the decision rule that the receiver should implement to minimize the probability of error. Sketch the decision regions.

(b) Calculate the probability of error P_e, based on the above decision rule.

Problem 19. *(MAP decoding regions)*

To communicate across an additive white Gaussian noise channel, an encoder uses the codewords c_i, $i = \{0, 1, 2\}$, shown below:

$$
\begin{aligned}
c_0 &= (1, 0)^\mathsf{T} \\
c_1 &= (-1, 0)^\mathsf{T} \\
c_2 &= (-1, 1)^\mathsf{T}
\end{aligned}
$$

(a) Draw the decoding regions of an ML decoder.

(b) Now assume that codeword i is used with probability $P_H(i)$, where $P_H(0) = P_H(1) = \frac{1}{4}$, $P_H(2) = \frac{1}{2}$ and that the receiver performs a MAP decision. Adjust the decoding regions accordingly. (A qualitative illustration suffices.)

(c) Finally, assume that the noise variance increases (same variance in both components). Update the decoding regions of the MAP decision rule. (Again, a qualitative illustration suffices.)

Problem 20. *(Sufficient statistic)*
Consider a binary hypothesis testing problem specified by:

$$H = 0 : \begin{cases} Y_1 = Z_1 \\ Y_2 = Z_1 Z_2 \end{cases}$$

$$H = 1 : \begin{cases} Y_1 = -Z_1 \\ Y_2 = -Z_1 Z_2, \end{cases}$$

where Z_1, Z_2, and H are independent random variables. Is Y_1 a sufficient statistic?

Problem 21. *(More on sufficient statistic)*
We have seen that if $H \to T(Y) \to Y$, then the probability of error P_e of a MAP decoder that decides on the value of H upon observing both $T(Y)$ and Y is the same as that of a MAP decoder that observes only $T(Y)$. It is natural to wonder if the contrary is also true, specifically if the knowledge that Y does not help reduce the error probability that we an achieve with $T(Y)$ implies $H \to T(Y) \to Y$. Here is a counter-example. Let the hypothesis H be either 0 or 1 with equal probability. (The distribution of H is critical in this example.) Let the observable Y take four values with conditional probabilities

$$P_{Y|H}(y|0) = \begin{cases} 0.4 & \text{if } y = 0 \\ 0.3 & \text{if } y = 1 \\ 0.2 & \text{if } y = 2 \\ 0.1 & \text{if } y = 3 \end{cases} \qquad P_{Y|H}(y|1) = \begin{cases} 0.1 & \text{if } y = 0 \\ 0.2 & \text{if } y = 1 \\ 0.3 & \text{if } y = 2 \\ 0.4 & \text{if } y = 3 \end{cases}$$

and $T(Y)$ is the function

$$T(y) = \begin{cases} 0 & \text{if } y = \{0, 1\} \\ 1 & \text{if } y = \{2, 3\} \end{cases}$$

(a) Show that the MAP decoder $\hat{H}(T(y))$ that decides based on $T(y)$ is equivalent to the MAP decoder $\hat{H}(y)$ that operates based on y.

(b) Compute the probabilities $Pr\{Y = 0 | T(Y) = 0, H = 0\}$ and $Pr\{Y = 0 | T(Y) = 0, H = 1\}$. Is it true that $H \to T(Y) \to Y$?

Problem 22. *(Fisher–Neyman factorization theorem)*

Consider the hypothesis testing problem where the hypothesis is $H \in \{0, 1, \dots, m-1\}$, the observable is Y, and $T(Y)$ is a function of the observable. Let $f_{Y|H}(y|i)$ be given for all $i \in \{0, 1, \dots, m-1\}$. Suppose that there are positive functions $g_0, g_1, \dots, g_{m-1}, h$ so that for each $i \in \{0, 1, \dots, m-1\}$ one can write

$$f_{Y|H}(y|i) = g_i(T(y))h(y) \tag{1}$$

(a) Show that when the above conditions are satisfied, a MAP decision depends on the observable Y only through $T(Y)$. In other words, Y itself is not necessary.
Hint: Work directly with the definition of a MAP decision rule.

(b) Show that $T(Y)$ is a sufficient statistic, that is $H \to T(Y) \to Y$.
Hint: Start by observing the following fact: Given a random variable Y with probability density function $f_Y(y)$ and given an arbitrary event $\mathcal{B}$, we have

$$f_{Y|Y \in \mathcal{B}} = \frac{f_Y(y)\mathbb{1}\{y \in \mathcal{B}\}}{\int_{\mathcal{B}} f_Y(y)dy} \tag{2}$$

Proceed by defining $\mathcal{B}$ to be the event $\mathcal{B} = \{y : T(y) = t\}$ and make use of (2) applied to $f_{Y|H}(y|i)$ to prove that $f_{Y|H,T(Y)}(y|i,t)$ is independent of i.

(c) (Example 1) Under hypothesis $H = i$, let $Y = (Y_1, Y_2, \dots, Y_n)$, $Y_k \in \{0, 1\}$, be an independent and identically distributed sequence of coin tosses such that $P_{Y_k|H}(1|i) = p_i$. Show that the function $T(y_1, y_2, \dots, y_n) = \sum_{k=1}^n y_k$ fulfills the condition expressed in equation (1). Notice that $T(y_1, y_2, \dots, y_n)$ is the number of 1s in y.

(d) (Example 2) Under hypothesis $H = i$, let the observable Y_k be Gaussian distributed with mean m_i and variance 1; that is

$$f_{Y_k|H}(y|i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-m_i)^2}{2}},$$

and $Y_1, Y_2, \dots, Y_n$ be independently drawn according to this distribution. Show that the sample mean $T(y_1, y_2, \dots, y_n) = \frac{1}{n}\sum_{k=1}^n y_k$ fulfills the condition expressed in (1).

Problem 23. *(Irrelevance and operational irrelevance)*

Let the hypothesis H be related to the observables (U, V) via the channel $P_{U,V|H}$ and for simplicity assume that $P_{U|H}(u|h) > 0$ and $P_{V|U,H}(v|u,h) > 0$ for every $h \in \mathcal{H}$, $v \in \mathcal{V}$, and $u \in \mathcal{U}$. We say that V is operationally irrelevant if a MAP decoder that observes (U, V) achieves the same probability of error as one that observes only U, and this is true regardless of P_H. We now prove that irrelevance and operational irrelevance imply one another. We have already proved that irrelevance implies operational irrelevance. Hence it suffices to show that operational irrelevance implies irrelevance or, equivalently, that if V is not irrelevant, then it is not operationally irrelevant. We will prove the latter statement. We begin with a few observations that are instructive. By definition, V irrelevant means $H \to U \to V$. Hence V irrelevant is equivalent to the statement that, conditioned on U, the random variables H and V are independent. This gives us one intuitive explanation about why V is operationally irrelevant when $H \to U \to V$. Once we observe that $U = u$, we can restate the hypothesis testing problem in terms of a hypothesis H and an observable V that are independent (conditioned on $U = u$) and because of independence, from V we learn nothing about H. But if V is not irrelevant, then there is at least a u, call it $u^\star$, for which

H and V are not independent condition on $U = u^\star$. It is when such a u is observed that we should be able to prove that V affects the decision. This suggests that the problem we are trying to solve is intimately related to the simpler problem that involves the hypothesis H and the observable V and the two are not independent. We begin with this problem and then we generalize.

(a) Let the hypothesis be $H \in \mathcal{H}$ (of yet unspecified distribution) and let the observable $V \in \mathcal{V}$ be related to H via an arbitrary but fixed channel $P_{V|H}$. Show that if V is not independent of H then there are distinct elements $i, j \in \mathcal{H}$ and distinct elements $k, l \in \mathcal{V}$ such that

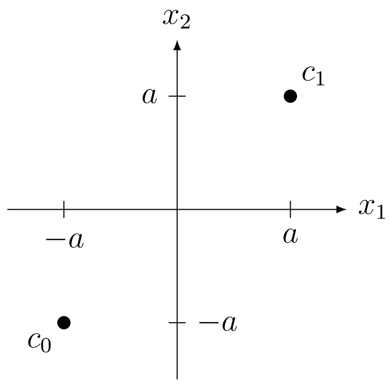
$$P_{V|H}(k|i) > P_{V|H}(k|j)$$
$$P_{V|H}(l|i) < P_{V|H}(l|j)$$

Hint: For every $h \in \mathcal{H}$, $\sum_{v \in \mathcal{V}} P_{V|H}(v|h) = 1$.

(b) Under the condition of part (a), show that there is a distribution P_H for which the observable V affects the decision of a MAP decoder.

(c) Generalize to show that if the observables are U and V, and $P_{U,V|H}$ is fixed so that $H \to U \to V$ does not hold, then there is a distribution on H for which V is not operationally irrelevant.
Hint: Argue as in parts (a) and (b) for the case $U = u^\star$, where $u^\star$ is as described above.

Problem 24. *(Antipodal signaling)*
Consider the signal constellation shown below:



Assume that the codewords c_0 and c_1 are used to communicate over the discrete-time AWGN channel. More precisely:

$$H = 0: \quad Y = c_0 + Z,$$
$$H = 1: \quad Y = c_1 + Z,$$

where $Z \sim \mathcal{N}(0, \sigma^2 I_2)$. Let $Y = (Y_1, Y_2)^\mathsf{T}$.

(a) Argue that Y_1 is not a sufficient statistic.

(b) Give a different signal constellation with two codewords $\tilde{c}_0$ and $\tilde{c}_1$ such that, when used in the above communication setting, Y_1 is a sufficient statistic.

Problem 25. *(Is it a sufficient statistic?)*

Consider the following hypothesis testing problem:

$$\begin{aligned} H = 0: \quad & Y = c_0 + Z, \\ H = 1: \quad & Y = c_1 + Z, \end{aligned}$$

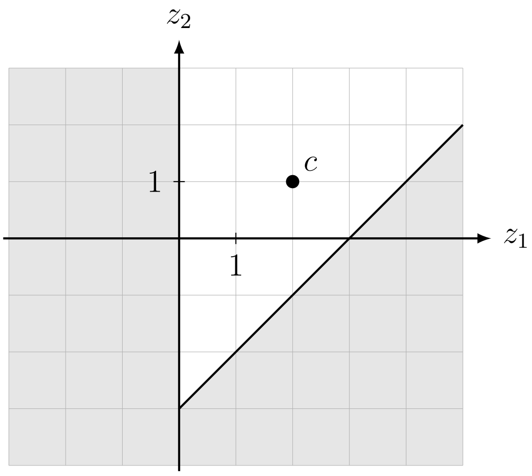
where $c_0 = -c_1 = (1,1)^\mathsf{T}$ and $Z \sim \mathcal{N}(0, \sigma^2 I_2)$.

(a) Can the error probability of an ML decoder that observes $Y = (Y_1, Y_2)^\mathsf{T}$ be lower than that of an ML decoder that observes $Y_1 + Y_2$?

(b) Argue whether or not $H \to (Y_1 + Y_2) \to Y$ forms a Markov chain.
Hint: Y is in a one-to-one relationship with $(Y_1 + Y_2, Y_1 - Y_2)$.
Hint: Argue that the random variables $Z_1 + Z_2$ and $Z_1 - Z_2$ are statistically independent.

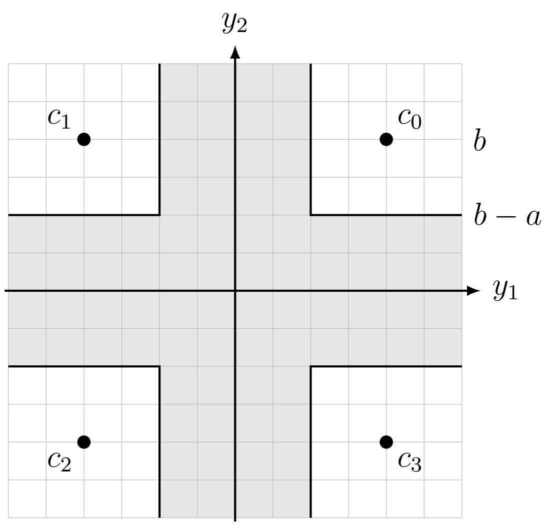
Problem 26. *(Union bound)*

Let $Z \sim \mathcal{N}(c, \sigma^2 I_2)$ be a random vector that takes values in $\mathbb{R}^2$, where $c = (2,1)^\mathsf{T}$. Find a non-trivial upper bound to the probability that Z is in the shaded region of the figure below.



Problem 27. *(QAM with erasure)*

Consider a QAM receiver that outputs a special symbol δ (called *erasure*) whenever the observation falls in the shaded area shown in the figure below, and does minimum-distance decoding otherwise. (This is neither a MAP nor an ML receiver.)



Assume that $c_0 \in \mathbb{R}^2$ is transmitted and that $Y = c_0 + N$ is received where $N \sim \mathcal{N}(0, \sigma^2 I_2)$. Let P_{0i}, $i = \{0,1,2,3\}$ be the probability that the receiver outputs $\hat{H} = i$ and let $P_{0\delta}$ be the probability that it outputs δ. Determine $P_{00}, P_{01}, P_{02}, P_{03}$, and $P_{0\delta}$.

Comment: If we choose $b - a$ large enough, we can make sure that the probability of error is very small (we say that an error occured if $\hat{H} = i$, $i \in \{0,1,2,3\}$ and $H \neq \hat{H}$). When $\hat{H} = \delta$, the receiver can ask for a retransmission of H. This requires a feedback channel from the receiver to the transmitter. In most practical applications, such a feedback channel is available.

PROBLEM 28. *(Repeat codes and Bhattacharyya bound)*

Consider two equally likely hypotheses. Under hypothesis $H = 0$, the transmitter sends $c_0 = (1, \ldots, 1)^\mathsf{T}$ and under $H = 1$ it sends $c_1 = (-1, \ldots, -1)^\mathsf{T}$, both of length n. The channel model is AWGN with variance σ^2 in each component. Recall that the probability of error for an ML receiver that observes the channel output $Y \in \mathbb{R}^n$ is

$$P_e = Q\left(\frac{\sqrt{n}}{\sigma}\right)$$

Suppose now that the decoder has access only to the sign of Y_i, $1 \le i \le n$, i.e. it observes

$$W = (W_1, \ldots, W_n) = (\operatorname{sign}(Y_1), \ldots, \operatorname{sign}(Y_n))$$

(a) Determine the MAP decision rule based on the observable W. Give a simple sufficient statistic.

(b) Find the expression for the probability of error $\tilde{P}_e$ of the MAP decoder that observes W. You may assume that n is odd.

(c) Your answer to (b) contains a sum that cannot be expressed in closed form. Express the Bhattacharyya bound on $\tilde{P}_e$.

(d) For $n = 1, 3, 5, 7$, find the numerical values of P_e, $\tilde{P}_e$, and the Bhattacharyya bound on $\tilde{P}_e$.

PROBLEM 29. *(Tighter union Bhattacharyya bound: Binary case)*

In this problem we derive a tighter version of the *union Bhattacharyya bound* for binary hypotheses. Let

$$\begin{aligned} H = 0: &\quad Y \sim f_{Y|H}(y|0) \\ H = 1: &\quad Y \sim f_{Y|H}(y|1) \end{aligned}$$

The MAP decision rule is

$$\hat{H}(y) = \arg\max_i P_H(i) f_{Y|H}(y|i),$$

and the resulting probability of error is

$$P_e = P_H(0) \int_{\mathcal{R}_1} f_{Y|H}(y|0)dy + P_H(1) \int_{\mathcal{R}_0} f_{Y|H}(y|1)dy$$

(a) Argue that

$$P_e = \int_y \min\left\{P_H(0) f_{Y|H}(y|0), P_H(1) f_{Y|H}(y|1)\right\} dy$$

(b) Prove that for $a, b \geq 0$, $\min(a, b) \leq \sqrt{ab} \leq \frac{a+b}{2}$. Use this to prove the tighter version of the Bhattacharyya bound, i.e.

$$P_e \leq \frac{1}{2} \int_y \sqrt{f_{Y|H}(y|0) f_{Y|H}(y|1)} dy$$

(c) Compare the above bound to (2.19) when there are two equiprobable hypotheses. How do you explain the improvement by a factor $\frac{1}{2}$?

Problem 30. *(Tighter union Bhattacharyya bound: M-ary case)*

In this problem we derive a tight version of the union bound for M-ary hypotheses. Let us analyze the following M-ary MAP detector:

$$\hat{H}(y) = \min\left\{i : P_H(i) f_{Y|H}(y|i) = \max_j \left\{P_H(j) f_{Y|H}(y|j)\right\}\right\}$$

Let

$$\mathcal{B}_{i,j} = \begin{cases} y : P_H(j) f_{Y|H}(y|j) \geq P_H(i) f_{Y|H}(y|i), & j < i \\ y : P_H(j) f_{Y|H}(y|j) > P_H(i) f_{Y|H}(y|i), & j > i \end{cases}$$

(a) Verify that $\mathcal{B}_{i,j} = \mathcal{B}_{j,i}^c$.

(b) Given $H = i$, the detector will make an error if and only if $y \in \bigcup_{j:j\neq i} \mathcal{B}_{i,j}$. The probability of error $P_e = \sum_{i=0}^{M-1} P_e(i) P_H(i)$. Show that:

$$\begin{aligned} P_e &\leq \sum_{i=0}^{M-1} \sum_{j>i} \left[Pr\left\{Y \in \mathcal{B}_{i,j} | H = i\right\} P_H(i) + Pr\left\{Y \in \mathcal{B}_{j,i} | H = j\right\} P_H(j)\right] \\ &= \sum_{i=0}^{M-1} \sum_{j>i} \left[\int_{\mathcal{B}_{i,j}} f_{Y|H}(y|i) P_H(i) dy + \int_{\mathcal{B}_{i,j}^c} f_{Y|H}(y|j) P_H(j) dy\right] \\ &= \sum_{i=0}^{M-1} \sum_{j>i} \left[\int_y \min\left\{f_{Y|H}(y|i) P_H(i), f_{Y|H}(y|j) P_H(j)\right\} dy\right] \end{aligned}$$

To prove the last part, go back to the definition of $\mathcal{B}_{i,j}$.

(c) Hence show that:

$$P_e \leq \sum_{i=0}^{M-1} \sum_{j>i} \left[\left(\frac{P_H(i) + P_H(j)}{2}\right) \int_y \sqrt{f_{Y|H}(y|i) f_{Y|H}(y|j)} dy\right]$$

Hint: For $a, b \geq 0$, $\min(a, b) \leq \sqrt{ab} \leq \frac{a+b}{2}$.

PROBLEM 31. *(Applying the tight Bhattacharyya bound)*

As an application of the *tight Bhattacharyya bound* (Exercise 29), consider the following binary hypothesis testing problem

$$\begin{aligned} H = 0: &\quad Y \sim \mathcal{N}\left(-a, \sigma^2\right) \\ H = 1: &\quad Y \sim \mathcal{N}\left(+a, \sigma^2\right), \end{aligned}$$

where the two hypotheses are equiprobable.

(a) Use the tight Bhattacharyya bound to derive a bound on P_e.

(b) We know that the probability of error for this binary hypothesis testing problem is $Q\left(\frac{a}{\sigma}\right) \leq \frac{1}{2}\exp\left(-\frac{a^2}{2\sigma^2}\right)$, where we have used the result $Q(x) \leq \frac{1}{2}\exp\left(-\frac{x^2}{2}\right)$. How do the two bounds compare? Comment on the result.

PROBLEM 32. *(Bhattacharyya bound for DMCs)*

Consider a discrete memoryless channel (DMC). This is a channel model described by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and a transition probability[1] $P_{Y|X}(y|x)$. When we use this channel to transmit an n-tuple $x \in \mathcal{X}^n$, the transition probability is

$$P_{Y|X}(y|x) = \prod_{i=1}^{n} P_{Y|X}(y_i|x_i)$$

So far, we have come across two DMCs, namely the BSC (binary symmetric channel) and the BEC (binary erasure channel). The purpose of this problem is to see that for DMCs, the Bhattacharyya bound takes a simple form, in particular when the channel input alphabet $\mathcal{X}$ contains only two letters.

(a) Consider a transmitter that sends $c_0 \in \mathcal{X}^n$ and $c_1 \in \mathcal{X}^n$ with equal probability. Justify the following chain of (in)equalities:

$$\begin{aligned} P_e &\overset{(a)}{\leq} \sum_y \sqrt{P_{Y|X}(y|c_0)P_{Y|X}(y|c_1)} \\ &\overset{(b)}{=} \sum_y \sqrt{\prod_{i=1}^{n} P_{Y|X}(y_i|c_{0,i})P_{Y|X}(y_i|c_{1,i})} \\ &\overset{(c)}{=} \sum_{y_1,\ldots,y_n} \prod_{i=1}^{n} \sqrt{P_{Y|X}(y_i|c_{0,i})P_{Y|X}(y_i|c_{1,i})} \\ &\overset{(d)}{=} \sum_{y_1} \sqrt{P_{Y|X}(y_1|c_{0,1})P_{Y|X}(y_1|c_{1,1})} \\ &\quad \cdots \sum_{y_n} \sqrt{P_{Y|X}(y_n|c_{0,n})P_{Y|X}(y_n|c_{1,n})} \\ &\overset{(e)}{=} \prod_{i=1}^{n} \sum_y \sqrt{P_{Y|X}(y|c_{0,i})P_{Y|X}(y|c_{1,i})} \\ &\overset{(f)}{=} \prod_{a\in\mathcal{X}, b\in\mathcal{X}, a\neq b} \left(\sum_y \sqrt{P_{Y|X}(y|a)P_{Y|X}(y|b)}\right)^{n(a,b)}, \end{aligned}$$

[1] Here we are assuming that the output alphabet is discrete. Otherwise we use densities instead of probabilities.

where $n(a, b)$ is the number of positions i in which $c_{0,i} = a$ and $c_{1,i} = b$.

(b) The Hamming distance $d_H(c_0, c_1)$ is defined as the number of positions in which c_0 and c_1 differ. Show that for a binary input channel, i.e. when $\mathcal{X} = \{a, b\}$, the *Bhattacharyya bound* becomes

$$P_e \leq z^{d_H(c_0,c_1)},$$

where

$$z = \sum_y \sqrt{P_{Y|X}(y|a) P_{Y|X}(y|b)}$$

Notice that z depends only on the channel, whereas its exponent depends only on c_0 and c_1.

(c) Evaluate the channel parameter z for the following.

(i) The binary input Gaussian channel described by the densities

$$f_{Y|X}(y|0) = \mathcal{N}\left(-\sqrt{E}, \sigma^2\right)$$
$$f_{Y|X}(y|1) = \mathcal{N}\left(\sqrt{E}, \sigma^2\right)$$

(ii) The binary symmetric channel (BSC) with $\mathcal{X} = \mathcal{Y} = \{\pm 1\}$ and transition probabilities described by

$$P_{Y|X}(y|x) = \begin{cases} 1-\delta, & \text{if } y = x \\ \delta, & \text{otherwise} \end{cases}$$

(iii) The binary erasure channel (BEC) with $\mathcal{X} = \{\pm 1\}$, $\mathcal{Y} = \{-1, E, 1\}$, and transition probabilities given by

$$P_{Y|X}(y|x) = \begin{cases} 1-\delta, & \text{if } y = x, \\ \delta, & \text{if } y = E, \\ 0, & \text{otherwise} \end{cases}$$

Problem 33. *(Bhattacharyya bound and Laplacian noise)*

Assuming two equiprobable hypotheses, evaluate the Bhattacharyya bound for the following (Laplacian noise) setting:

$$H = 0: \quad Y = -a + Z$$
$$H = 1: \quad Y = +a + Z,$$

where $a \in \mathbb{R}_+$ is a constant and Z is a random variable of probability density function $f_Z(z) = \frac{1}{2} \exp(-|z|)$, $z \in \mathbb{R}$.

Problem 34. *(Dice tossing)*

You have two dice, one fair and one biased. A friend tells you that the biased die produces a 6 with probability $\frac{1}{4}$, and produces the other values with uniform probabilities. You do not know a priori which of the two is a fair die. You choose with uniform probabilities one of the two dice, and perform n consecutive tosses. Let $Y_i \in \{1, \ldots, 6\}$ be the random variable modeling the ith experiment and let $Y = (Y_1, \ldots, Y_n)$.

(a) Based on the observable Y, find the decision rule to determine whether the die you have chosen is biased. Your rule should maximize the probability that the decision is correct.

(b) Identify a sufficient statistic $S \in \mathbb{N}$.

(c) Find the Bhattacharyya bound on the probability of error. You can either work with the observable $(Y_1, \dots, Y_n)$ or with $(Z_1, \dots, Z_n)$, where Z_i indicates whether the ith observation is a 6 or not. Yet another alternative is to work with S.

Hint: Depending on the approach, the following may be useful: $\sum_{i=0}^{n} \binom{n}{i} x^i = (1+x)^n$ for $n \in \mathbb{N}$.

PROBLEM 35. *(ML receiver and union bound for orthogonal signaling)*

Let $H \in \{1, \dots, m\}$ be uniformly distributed and consider the communication problem described by:

$$H = i: \quad Y = c_i + Z, \quad Z \sim \mathcal{N}\left(0, \sigma^2 I_m\right),$$

where $c_1, \dots, c_m$, $c_i \in \mathbb{R}^m$, is a set of constant-energy orthogonal codewords. Without loss of generality we assume

$$c_i = \sqrt{\mathcal{E}} e_i,$$

where e_i is the ith unit vector in $\mathbb{R}^m$, i.e. the vector that contains 1 at position i and 0 elsewhere, and $\mathcal{E}$ is some positive constant.

(a) Describe the maximum-likelihood decision rule.

(b) Find the distances $\|c_i - c_j\|$, $i \neq j$.

(c) Using the union bound and the Q function, upper bound the probability $P_e(i)$ that the decision is incorrect when $H = i$.

PROBLEM 36. *(Uniform polar to Cartesian)*

Let R and Φ be independent random variables. R is distributed uniformly over the unit interval, Φ is distributed uniformly over the interval $[0, 2\pi)$.

(a) Interpret R and Φ as the polar coordinates of a point in the plane. It is clear that the point lies inside (or on) the unit circle. Is the distribution of the point uniform over the unit disk? Take a guess !

(b) Define the random variables

$$X = R\cos\Phi$$
$$Y = R\sin\Phi$$

Find the joint distribution of the random variables X and Y by using the Jacobian determinant.

(c) Does the result of part (b) support or contradict your guess from part (a)? Explain.

Problem 37. *(Real-valued Gaussian random variables)*

For the purpose of this exercise, two zero-mean real-valued Gaussian random variables X and Y are called *jointly Gaussian* if and only if their joint density is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x,y)\Sigma^{-1}(x,y)^{\mathsf{T}}\right),$$

where (for zero-mean random vectors) the so-called covariance matrix Σ is

$$\Sigma = \mathbb{E}\left[(X,Y)^{\mathsf{T}}(X,Y)\right] = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}$$

(a) Show that if X and Y are zero-mean jointly Gaussian random variables, then X is a zero-mean Gaussian random variable, and so is Y.

(b) Show that if X and Y are independent zero-mean Gaussian random variables, then X and Y are zero-mean jointly Gaussian random variables.

(c) However, if X and Y are Gaussian random variables but *not* independent, then X and Y are not necessarily jointly Gaussian. Give an example where X and Y are Gaussian random variables, yet they are *not* jointly Gaussian.

(d) Let X and Y be independent Gaussian random variables with zero mean and variance σ_X^2 and σ_Y^2, respectively. Find the probability density function of $Z = X + Y$.

Observe that no computation is required if we use the definition of jointly Gaussian random variables given in Appendix 2.10.

Problem 38. *(Correlation vs. independence)*

Let Z be a random variable with probability density function

$$f_Z(z) = \begin{cases} 1/2, & -1 \le z \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Also, let $X = Z$ and $Y = Z^2$.

(a) Show that X and Y are uncorrelated.

(b) Are X and Y independent?

(c) Now let X and Y be jointly Gaussian, zero mean, uncorrelated with variances σ_X^2 and σ_Y^2, respectively. Are X and Y independent? Justify your answer.

Problem 39. *(Data-storage channel)*

The process of storing and retrieving binary data on a thin-film disk can be modeled as transmitting binary symbols across an additive white Gaussian noise channel where the noise Z has a variance that depends on the transmitted (stored) binary symbol X. The noise has the following input-dependent density:

$$f_Z(z) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{z^2}{2\sigma_1^2}} & \text{if } X = 1 \\ \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{z^2}{2\sigma_0^2}} & \text{if } X = 0, \end{cases}$$

where $\sigma_1 > \sigma_0$. The channel inputs are equally likely.

(a) On the same graph, plot the two possible output probability density functions. Indicate, qualitatively, the decision regions.

(b) Determine the optimal receiver in terms of σ_0 and σ_1.

(c) Write an expression for the error probability P_e as a function of σ_0 and σ_1.

PROBLEM 40. *(A simple multiple-access scheme)*

Consider the following very simple model of a multiple-access scheme. There are two users. Each user has two hypotheses. Let $\mathcal{H}_1 = \mathcal{H}_2 = \{0, 1\}$ denote the respective set of hypotheses and assume that both users employ a uniform prior. Further, let X_1 and X_2 be the respective signals sent by user one and two. Assume that the transmissions of both users are independent and that $X_1 \in \{\pm 1\}$ and $X_2 \in \{\pm 2\}$ where X_1 and X_2 are positive if their respective hypotheses is zero and negative otherwise. Assume that the receiver observes the signal $Y = X_1 + X_2 + Z$, where Z is a zero-mean Gaussian random variable with variance σ^2 and is independent of the transmitted signal.

(a) Assume that the receiver observes Y and wants to estimate both H_1 and H_2. Let $\hat{H}_1$ and $\hat{H}_2$ be the estimates. What is the generic form of the optimal decision rule?

(b) For the specific set of signals given, what is the set of possible observations, assuming that $\sigma^2 = 0$? Label these signals by the corresponding (joint) hypotheses.

(c) Assuming now that $\sigma^2 > 0$, draw the optimal decision regions.

(d) What is the resulting probability of correct decision? That is, determine the probability $Pr\left\{\hat{H}_1 = H_1, \hat{H}_2 = H_2\right\}$.

(e) Finally, assume that we are interested only by the transmission of user two. Describe the receiver that minimizes the error probability and determine $Pr\left\{\hat{H}_2 = H_2\right\}$.

PROBLEM 41. *(Data-dependent noise)*

Consider the following binary Gaussian hypothesis testing problem with data-dependent noise. Under hypothesis $H = 0$ the transmitted signal is $c_0 = -1$ and the received signal is $Y = c_0 + Z_0$, where Z_0 is zero-mean Gaussian with variance one. Under hypothesis $H = 1$ the transmitted signal is $c_1 = 1$ and the received signal is $Y = c_1 + Z_1$, where Z_1 is zero-mean Gaussian with variance σ^2. Assume that the prior is uniform.

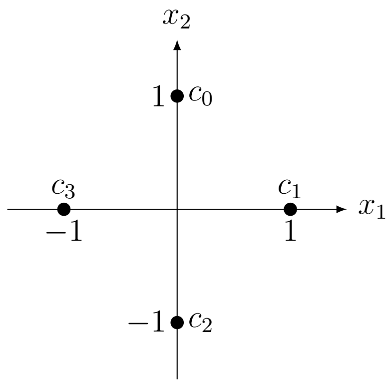
(a) Write the optimal decision rule as a function of the parameter σ^2 and the received signal Y.

(b) For the value $\sigma^2 = e^4$ compute the decision regions.

(c) Give expressions as simple as possible for the error probabilities $P_e(0)$ and $P_e(1)$.

Problem 42. *(Correlated noise)*

Consider the following communication problem. The message is represented by a uniformly distributed random variable H, that takes values in $\{0,1,2,3\}$. When $H = i$ we send c_i, where $c_0 = (0,1)^\mathsf{T}, c_1 = (1,0)^\mathsf{T}, c_2 = (0,-1)^\mathsf{T}, c_3 = (-1,0)^\mathsf{T}$ (see figure below). When $H = i$, the receiver observes the vector $Y = c_i + Z$, where Z is a zero-mean Gaussian random vector of covariance matrix $\Sigma = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}$.



(a) In order to simplify the decision problem, we transform Y into $\hat{Y} = BY = Bc_i + BZ$, where B is a 2-by-2 invertible matrix, and use $\hat{Y}$ as a sufficient statistic. Find a B such that BZ is a zero-mean Gaussian random vector with independent and identically distributed components.
Hint: If $A = \frac{1}{4}\begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}$, then $A\Sigma A^\mathsf{T} = I$, with $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(b) Formulate the new hypothesis testing problem that has $\hat{Y}$ as the observable and depict the decision regions.

(c) Give an upper bound to the error probability in this decision problem.

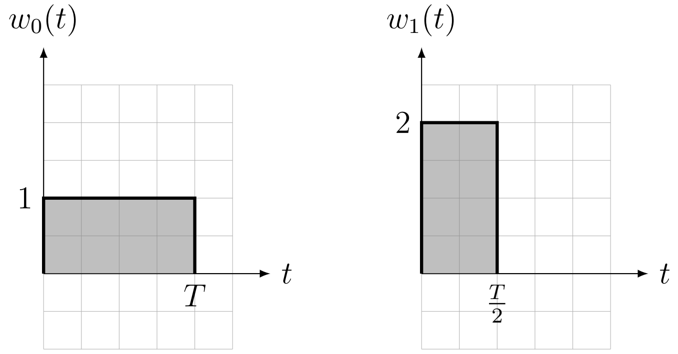
3 Receiver design for the continuous-time AWGN channel: Second layer

Problem 1. *(Gram–Schmidt procedure on tuples)*

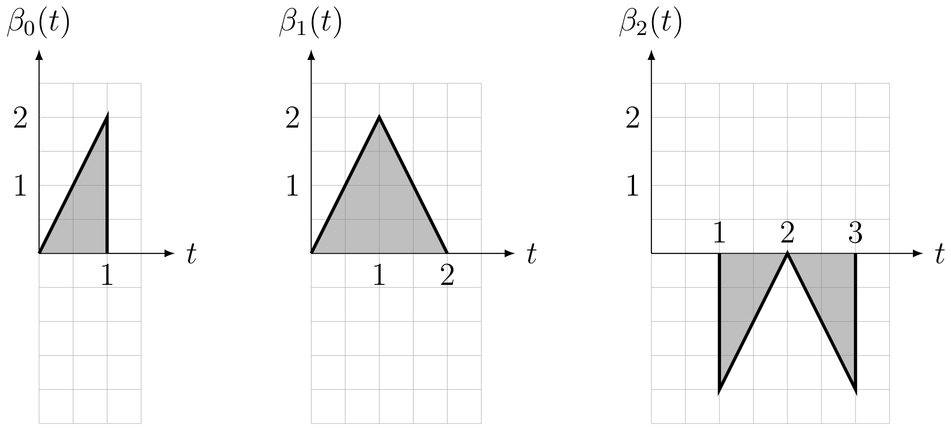
By means of the Gram–Schmidt orthonormalization procedure, find an orthonormal basis for the subspace spanned by the four vectors $\beta_1 = (1, 0, 1, 1)^\mathsf{T}$, $\beta_2 = (2, 1, 0, 1)^\mathsf{T}$, $\beta_3 = (1, 0, 1, -2)^\mathsf{T}$, and $\beta_4 = (2, 0, 2, -1)^\mathsf{T}$.

Problem 2. *(Gram–Schmidt procedure on two waveforms)*

Use the Gram–Schmidt procedure to find an orthonormal basis for the vector space spanned by the functions shown below.



Problem 3. *(Gram–Schmidt procedure on three waveforms)*



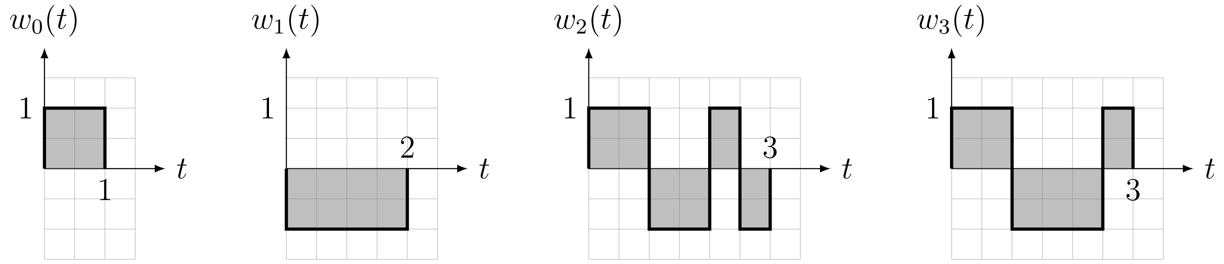
(a) By means of the Gram–Schmidt procedure, find an orthonormal basis for the space spanned by the above waveforms.

(b) In your chosen orthonormal basis, let $w_0(t)$ and $w_1(t)$ be represented by the codewords $c_0 = (3, -1, 1)^\mathsf{T}$ and $c_1 = (-1, 2, 3)^\mathsf{T}$ respectively. Plot $w_0(t)$ and $w_1(t)$.

(c) Compute the (standard) inner products $\langle c_0, c_1 \rangle$ and $\langle w_0, w_1 \rangle$ and compare them.

(d) Compute the norms $\|c_0\|$ and $\|w_0\|$ and compare them.

Problem 4. *(Orthonormal expansion)*

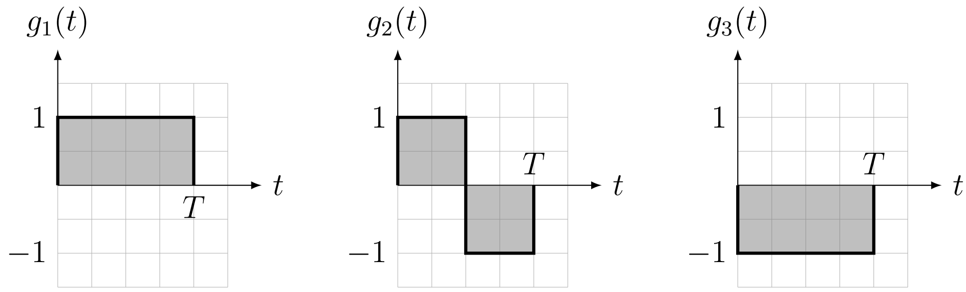


For the signal set shown above, do the following.

(a) Use the Gram–Schmidt (GS) procedure to find an orthonormal basis $\psi_1(t), \ldots, \psi_n(t)$.
Hint: No need to work out the intermediary steps of the GS procedure. The purpose of this exercise is to check, with hardly any calculation, your understanding of what the GS procedure does.

(b) Find the codeword $c_i \in \mathbb{R}^n$ that describes $w_i(t)$ with respect to your orthonormal basis. (No calculation needed.)

Problem 5. *(Noise in regions)*
Let $N(t)$ be white Gaussian noise of power spectral density $\frac{N_0}{2}$ and $g_1(t)$, $g_2(t)$, $g_3(t)$ be the waveforms shown below:



For $i = 1, 2, 3$, let $Z_i = \int N(t) g_i^*(t) dt$, $Z = (Z_1, Z_2)^\mathsf{T}$, and $U = (Z_1, Z_3)^\mathsf{T}$.

(a) Determine the norm $\|g_i\|$, $i = 1, 2, 3$.

(b) Are Z_1 and Z_2 independent? Justify your answer.

(c) Find the probability P_a that Z lies in the square of Figure 3a.

(d) Find the probability P_b that Z lies in the square of Figure 3b.

(e) Find the probability Q_a that U lies in the square of Figure 3a.

(f) Find the probability Q_b that U lies in the square of Figure 3c.

Problem 6. *(Two-signals error probability)*
The two signals shown below are used to communicate one bit across the continuous-time AWGN channel of power spectral density $\frac{N_0}{2} = 6 \frac{W}{Hz}$. Write an expression for the error probability of an ML receiver.

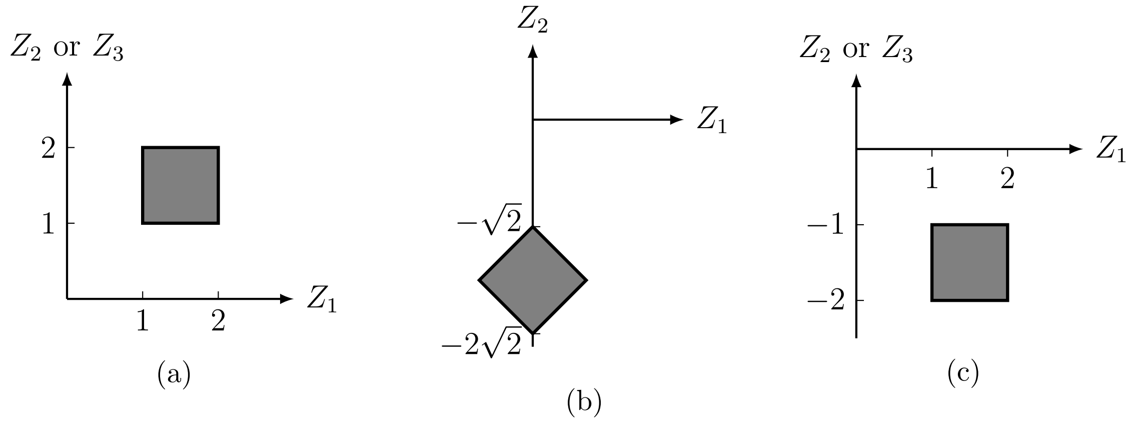
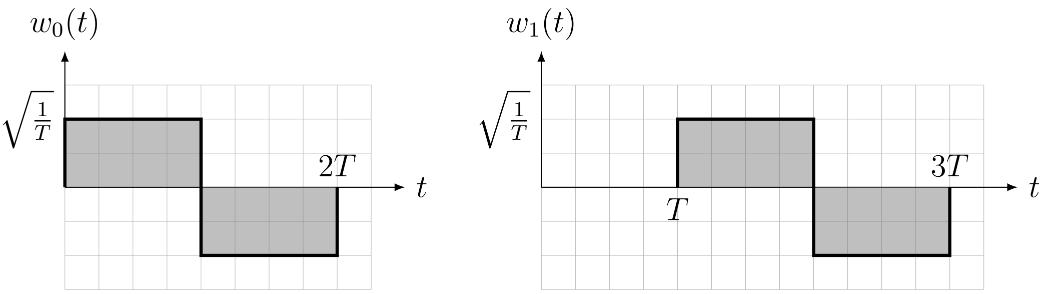


Figure 3



PROBLEM 7. *(On-off signaling)*

Consider the binary hypothesis testing problem specified by:

$$\begin{aligned} H = 0: \quad & R(t) = w(t) + N(t) \\ H = 1: \quad & R(t) = N(t), \end{aligned}$$

where $N(t)$ is additive white Gaussian noise of power spectral density $\frac{N_0}{2}$ and $w(t)$ is the signal shown in Figure 4a.

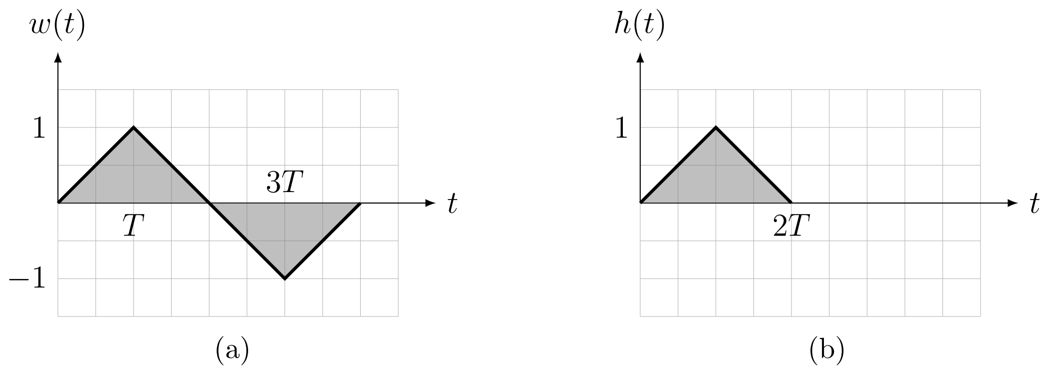


Figure 4

(a) Describe the maximum likelihood receiver for the received signal $R(t)$, $t \in \mathbb{R}$.

(b) Determine the error probability for the receiver you described in (a).

(c) Sketch a block diagram of your receiver of part (a) using a filter with impulse response $h(t)$ (or a scaled version thereof) shown in Figure 4b.

Problem 8. *(QAM receiver)*

Let the channel output be

$$R(t) = W(t) + N(t),$$

where $W(t)$ has the form

$$W(t) = \begin{cases} X_1\sqrt{\frac{2}{T}}\cos 2\pi f_c t + X_2\sqrt{\frac{2}{T}}\sin 2\pi f_c t & 0 \le t \le T \\ 0 & \text{otherwise,} \end{cases}$$

$2f_cT \in \mathbb{Z}$ is a constant known to the receiver, $X = (X_1, X_2)$ is a uniformly distributed random vector that takes values in

$$\left\{\sqrt{\mathcal{E}}(1,1), \sqrt{\mathcal{E}}(-1,1), \sqrt{\mathcal{E}}(-1,-1), \sqrt{\mathcal{E}}(1,-1)\right\}$$

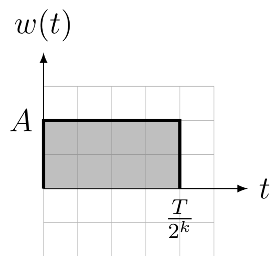
for some known constant $\mathcal{E}$, and $N(t)$ is white Gaussian noise of power spectral density $\frac{N_0}{2}$.

(a) Specify a receiver that, based on the channel output $R(t)$, decides on the value of the vector X with least probability of error.

(b) Find the error probability of the receiver you have specified.

Problem 9. *(Signaling scheme example)*

Let the message H be uniformly distributed over the message set $\mathcal{H} = \left\{0, 1, 2, \ldots, 2^k - 1\right\}$. When $H = i \in \mathcal{H}$, the transmitter sends $w_i(t) = w(t - i\frac{T}{2^k})$, where $w(t)$ is shown below:



The channel output is $R(t) = w_i(t) + N(t)$, where $N(t)$ denotes white Gaussian noise of power spectral density $\frac{N_0}{2}$.
Sketch a block diagram of a receiver that, based on $R(t)$, decides on the value of H with least probability of error.

Problem 10. *(Matched filter implementation)*

In this problem, we consider the implementation of matched filter receivers. In particular, we consider frequency-shift keying (FSK) with the following signals:

$$w_j(t) = \begin{cases} \sqrt{\frac{2}{T}}\cos 2\pi \frac{n_j}{T} t & 0 \le t \le T \\ 0 & \text{otherwise,} \end{cases}$$

where $n_j \in \mathbb{Z}$ and $0 \le j \le m-1$. Thus, the communication scheme consists of m signals $w_j(t)$ of different frequencies $\frac{n_j}{T}$.

(a) Determine the impulse response $h_j(t)$ of a causal matched filter for the signal $w_j(t)$. Plot $h_j(t)$ and specify the sampling time.

(b) Sketch the matched filter receiver. How many matched filters are needed?

(c) Sketch the output of the matched filter with impulse response $h_j(t)$ when the input is $w_j(t)$.

PROBLEM 11. *(Matched filter intuition)*

In this problem, we develop further intuition about matched filters. You may assume that all waveforms are real-valued. Let $R(t) = \pm w(t) + N(t)$ be the channel output, where $N(t)$ is additive white Gaussian noise of power spectral density $\frac{N_0}{2}$ and $w(t)$ is an arbitrary but fixed pulse. Let $\phi(t)$ be a unit-norm but otherwise arbitrary pulse, and consider the receiver operation

$$Y = \langle R, \phi \rangle = \langle w, \phi \rangle + \langle N, \phi \rangle$$

The signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR} \triangleq \frac{|\langle w, \phi \rangle|^2}{\mathbb{E}[|\langle N, \phi \rangle|^2]}$$

Notice that the SNR remains the same if we scale $\phi(t)$ by a constant factor. Notice also that

$$\mathbb{E}[|\langle N, \phi \rangle|^2] = \frac{N_0}{2}$$

(a) Use the Cauchy–Schwarz inequality to give an upper bound on the SNR. What is the condition for equality in the Cauchy–Schwarz inequality? Find the $\phi(t)$ that maximizes the SNR. What is the relationship between the maximizing $\phi(t)$ and the signal $w(t)$?

(b) Let us verify that we would get the same result using a pedestrian approach. Instead of waveforms we consider tuples. So let $c = (c_1, c_2)^\mathsf{T} \in \mathbb{R}^2$ and use calculus (instead of the Cauchy–Schwarz inequality) to find the $\phi = (\phi_1, \phi_2)^\mathsf{T} \in \mathbb{R}^2$ that maximizes $\langle c, \phi \rangle$ subject to the constraint that ϕ has unit norm.

(c) Verify with a picture (convolution) that the output at time T of a filter with input $w(t)$ and impulse response $h(t) = w(T - t)$ is indeed $\langle w, w \rangle = \int_{-\infty}^{\infty} w^2(t)dt$.

PROBLEM 12. *(Two receive antennas)*

Consider the following communication problem. The message is represented by a uniformly distributed random variable $X \in \{\pm 1\}$. The transmitter sends $Xw(t)$ where $w(t)$ is a unit-energy pulse known to the receiver. There are two channels with output $R_1(t)$ and $R_2(t)$ such that

$$\begin{aligned} R_1(t) &= X\beta_1 w(t - \tau_1) + N_1(t) \\ R_2(t) &= X\beta_2 w(t - \tau_2) + N_2(t), \end{aligned}$$

where $\beta_1, \beta_2, \tau_1, \tau_2$ are constants known to the receiver and $N_1(t)$ and $N_2(t)$ are white Gaussian noise of power spectral density $\frac{N_0}{2}$. We assume that $N_1(t)$ and $N_2(t)$ are independent of each other and independent of X. We also assume that $\int w(t-\tau_1)w(t-\tau_2)dt = \gamma$, where $-1 \leq \gamma \leq 1$.

(a) Describe an ML receiver for X that observes both $R_1(t)$ and $R_2(t)$. Determine its probability of error in terms of the Q function, β_1, β_2, γ, and $\frac{N_0}{2}$.

(b) Repeat part (a) assuming that the receiver has access only to the sum-signal $R(t) = R_1(t) + R_2(t)$.

Problem 13. *(Receiver)*

The signal set

$$
\begin{aligned}
w_0(t) &= \operatorname{sinc}^2(t) \\
w_1(t) &= \sqrt{2}\operatorname{sinc}^2(t)\cos(4\pi t)
\end{aligned}
$$

is used to communicate across the AWGN channel of noise power spectral density $\frac{N_0}{2}$.

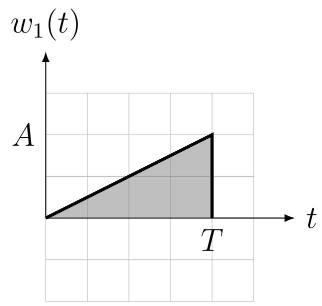
(a) Sketch a block diagram of an ML receiver for the above signal set. (No need to worry about filter causality.)

(b) Determine the error probability of your receiver assuming that $w_0(t)$ and $w_1(t)$ are equally likely.

(c) If you keep the same receiver, but use $w_0(t)$ with probability $\frac{1}{3}$ and $w_1(t)$ with probability $\frac{2}{3}$, does the error probability increase, decrease, or stay the same?

Problem 14. *(ML receiver with single causal filter)*

Let $w_1(t)$ be as shown below and let $w_2(t) = w_1(t - T_d)$, where $T_d \geq T$ is a fixed number known to the receiver.



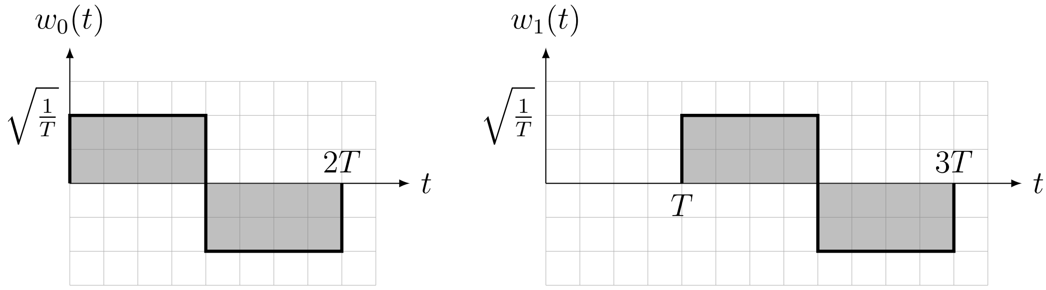
One of the two pulses is selected at random and transmitted across the AWGN channel of noise power spectral density $\frac{N_0}{2}$.

(a) Describe an ML receiver that decides which pulse was transmitted. The n-tuple former must contain a single causal matched filter. Finally, draw the matched filter impulse response.

(b) Express the error probability of the receiver in (a) in terms of A, T, T_d, N_0. Consider both cases $T_d \geq T$ and $T_d < T$.

PROBLEM 15. *(Delayed signals)*

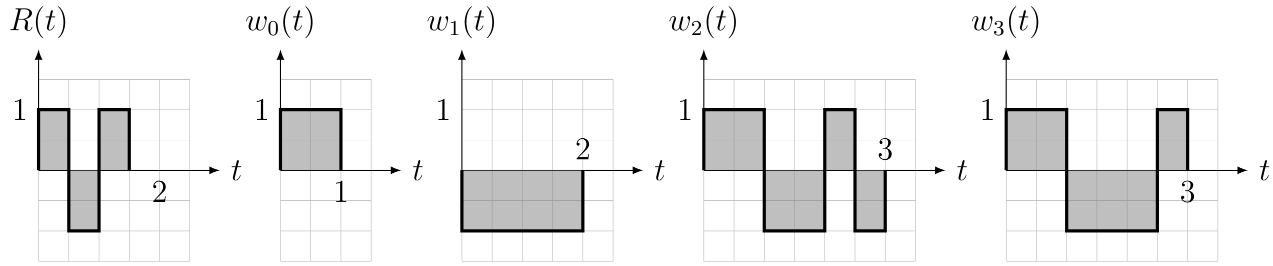
One of the two signals shown below is selected at random and is transmitted over the additive white Gaussian noise channel of power spectral density $\frac{N_0}{2}$.



Draw a block diagram of a maximum likelihood receiver that uses a single matched filter and express its error probability.

PROBLEM 16. *(ML decoder for AWGN channel)*

The signal $R(t)$ is fed to an ML receiver designed for a transmitter that uses the four signals $\{w_0(t), w_1(t), w_2(t), w_3(t)\}$ shown below to communicate across the AWGN channel. Determine the receiver output $\hat{H}$.



PROBLEM 17. *(AWGN channel and sufficient statistic)*

Let $\mathcal{W} = \{w_0(t), w_1(t)\}$ be the signal constellation used to communicate an equiprobable bit across an additive Gaussian noise channel. In this exercise, we verify that the projection of the channel output onto the inner product space $\mathcal{V}$ spanned by $\mathcal{W}$ is not necessarily a sufficient statistic, unless the noise is white.

Let $\psi_1(t), \psi_2(t)$ be an orthonormal basis for $\mathcal{V}$. We choose the additive noise to be $N(t) = Z_1\psi_1(t) + Z_2\psi_2(t) + Z_3\psi_3(t)$ for some normalized $\psi_3(t)$ that is orthogonal to $\psi_1(t)$ and $\psi_2(t)$, and choose Z_1, Z_2, Z_3 to be zero-mean jointly Gaussian random variables of identical variance σ^2. Let $c_i = (c_{i,1}, c_{i,2}, 0)^\mathsf{T}$ be the codeword associated to $w_i(t)$ with respect to the extended orthonormal basis $\psi_1(t), \psi_2(t), \psi_3(t)$. There is a one-to-one correspondence between the channel output $R(t)$ and $Y = (Y_1, Y_2, Y_3)^\mathsf{T}$, where $Y_i = \langle R, \psi_i \rangle$. In terms of Y, the hypothesis testing problem is

$$H = i : Y = c_i + Z, \quad i = \{0, 1\},$$

where we have defined $Z = (Z_1, Z_2, Z_3)^\mathsf{T}$.

(a) As a warm-up exercise, let us first assume that Z_1, Z_2, Z_3 are independent. Use the Fisher–Neyman factorization theorem to show that $(Y_1, Y_2)^\mathsf{T}$ is a sufficient statistic.

(b) Now assume that Z_1 and Z_2 are independent, but $Z_3 = Z_2$. Prove that in this case $(Y_1, Y_2)^\mathsf{T}$ is *not* a sufficient statistic.

(c) To check a specific case, consider $c_0 = (1, 0, 0)^\mathsf{T}$ and $c_1 = (0, 1, 0)^\mathsf{T}$. Determine the error probability of an ML receiver that observes $(Y_1, Y_2)^\mathsf{T}$ and that of another ML receiver that observes $(Y_1, Y_2, Y_3)^\mathsf{T}$.

PROBLEM 18. *(Mismatched receiver)*

Let a channel output be

$$R(t) = cXw(t) + N(t), \tag{3}$$

where $c > 0$ is some deterministic constant, X is a uniformly distributed random variable that takes values in $\{-3, -1, 1, 3\}$, $w(t) = \mathbb{1}_{[0,1)}(t)$, and $N(t)$ is white Gaussian noise of power spectral density $\frac{N_0}{2}$.

(a) Describe the receiver that, based on the channel output $R(t)$, decides on the value of X with least probability of error.

(b) Find the error probability of the receiver in part (a).

(c) Suppose now that you still use the receiver in part (a), but that the received signal is actually

$$R(t) = \frac{3}{4} cXw(t) + N(t),$$

i.e. you were unaware that the channel was attenuating the signal. What is the probability of error now?

(d) Suppose now that you still use the receiver in part (a) and that $R(t)$ is according to (3), but that the noise is colored. In fact, $N(t)$ is a zero-mean stationary Gaussian noise process of auto-covariance function

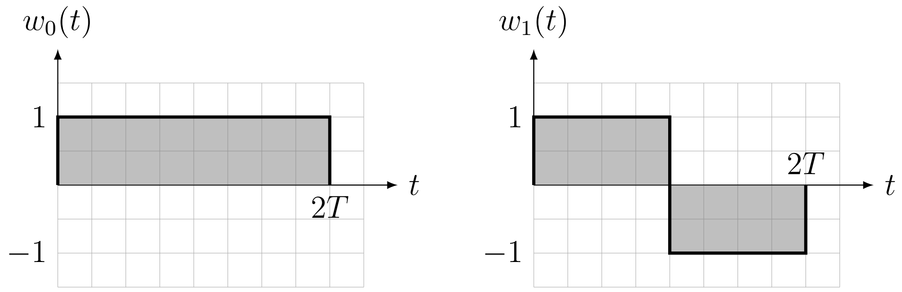
$$K_N(\tau) = \frac{1}{4\alpha} e^{-|\tau|/\alpha},$$

where $0 < \alpha < \infty$ is some deterministic real parameter. What is the probability of error now?

4 Signal design trade-offs

Problem 1. *(Signal translation)*

Consider the signals $w_0(t)$ and $w_1(t)$ shown below, used to communicate 1 bit across the AWGN channel of power spectral density $\frac{N_0}{2}$.



(a) Determine an orthonormal basis $\{\psi_0(t), \psi_1(t)\}$ for the space spanned by $\{w_0(t), w_1(t)\}$ and find the corresponding codewords c_0 and c_1. Work out two solutions, one obtained via Gram–Schmidt and one in which $\psi_1(t)$ is a delayed version of $\psi_0(t)$. Which of the two solutions would you choose if you had to implement the system?

(b) Let X be a uniformly distributed binary random variable that takes values in $\{0, 1\}$. We want to communicate the value of X over an additive white Gaussian noise channel. When $X = 0$, we send $w_0(t)$, and when $X = 1$, we send $w_1(t)$. Draw the block diagram of an ML receiver based on a single matched filter.

(c) Determine the error probability P_e of your receiver as a function of T and N_0.

(d) Find a suitable waveform $v(t)$ such that the signals $\tilde{w}_0(t) = w_0(t) - v(t)$ and $\tilde{w}_1(t) = w_1(t) - v(t)$ have minimum energy. Plot the resulting waveforms.

(e) What is the name of the signaling scheme that uses signals such as $\tilde{w}_0(t)$ and $\tilde{w}_1(t)$? Argue that one obtains this kind of signaling scheme independently of the initial choice of $w_0(t)$ and $w_1(t)$.

Problem 2. *(Orthogonal signal sets)*

Consider a set $\mathcal{W} = \{w_0(t), \dots, w_{m-1}(t)\}$ of mutually orthogonal signals with squared norm $\mathcal{E}$, each used with equal probability.

(a) Find the minimum-energy signal set $\tilde{\mathcal{W}} = \{\tilde{w}_0(t), \dots, \tilde{w}_{m-1}(t)\}$ obtained by translating the original set.

(b) Let $\tilde{\mathcal{E}}$ be the average energy of a signal picked at random within $\tilde{\mathcal{W}}$. Determine $\tilde{\mathcal{E}}$ and the energy saving $\mathcal{E} - \tilde{\mathcal{E}}$.

(c) Determine the dimension of the inner product space spanned by $\tilde{\mathcal{W}}$.

Problem 3. *(Suboptimal receiver for orthogonal signaling)*

This exercise takes a different approach to the evaluation of the performance of block-orthogonal signaling (Example 4.6). Let the message $H \in \{1, \ldots, m\}$ be uniformly distributed and consider the communication problem described by

$$H = i: \quad Y = c_i + Z, \quad Z \sim \mathcal{N}(0, \sigma^2 I_m),$$

where $Y = (Y_1, \ldots, Y_m)^{\mathsf{T}} \in \mathbb{R}^m$ is the received vector and $\{c_1, \ldots, c_m\} \subset \mathbb{R}^m$ is the codebook consisting of constant-energy codewords that are orthogonal to each other. Without loss of essential generality, we can assume

$$c_i = \sqrt{\mathcal{E}} e_i,$$

where e_i is the ith unit vector in $\mathbb{R}^m$, i.e. the vector that contains 1 at position i and 0 elsewhere, and $\mathcal{E}$ is some positive constant.

(a) Describe the statistic of Y_j for $j = 1, \ldots, m$ given that $H = 1$.

(b) Consider a suboptimal receiver that uses a threshold $t = \alpha\sqrt{\mathcal{E}}$ where $0 < \alpha < 1$. The receiver declares $\hat{H} = i$ if i is the only integer such that $Y_i \geq t$. If there is no such i or there is more than one index i for which $Y_i \geq t$, the receiver declares that it cannot decide. This will be viewed as an error. Let $E_i = \{Y_i \geq t\}$ and describe, in words, the meaning of the event

$$E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_m^c$$

(c) Find an upper bound to the probability that the above event *does not* occur when $H = 1$. Express your result using the Q function.

(d) Now let $m = 2^k$ and let $\mathcal{E} = k\mathcal{E}_b$ for some fixed energy per bit $\mathcal{E}_b$. Prove that the error probability goes to 0 as $k \to \infty$, provided that $\frac{\mathcal{E}_b}{\sigma^2} > \frac{2\ln 2}{\alpha^2}$. (Notice that because we can choose α^2 as close to 1 as we wish, if we insert $\sigma^2 = \frac{N_0}{2}$, the condition becomes $\frac{\mathcal{E}_b}{N_0} > \ln 2$, which is a weaker condition than the one obtained in Example 4.6.)

Hint: Use $m - 1 < m = e^{\ln m}$ and $Q(x) < \frac{1}{2}e^{-\frac{x^2}{2}}$.

Problem 4. *(Receiver diagrams)*

For each signaling method discussed in Section 4.4, draw the block diagram of an ML receiver.

Problem 5. *(Bit-by-bit on a pulse train)*

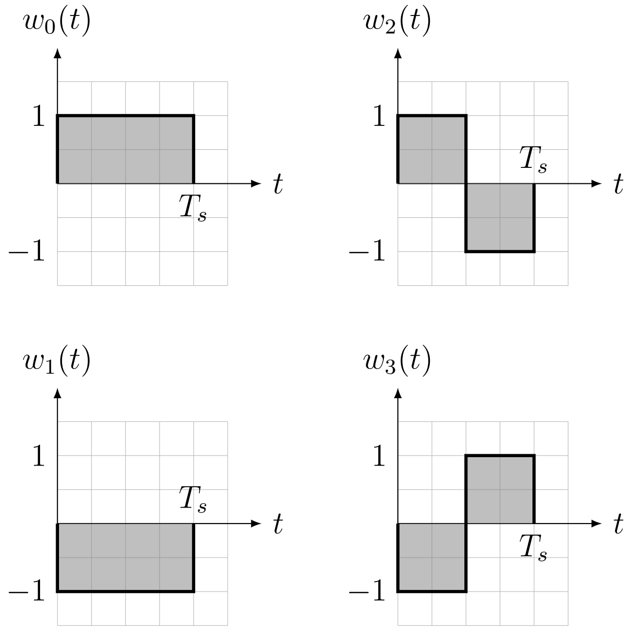
A communication system uses bit-by-bit on a pulse train to communicate at 1 Mbps using a rectangular pulse. The transmitted signal is of the form

$$\sum_j B_j \mathbb{1}_{[0,T_s)}(t - jT_s),$$

where $B_j \in \{\pm b\}$. Determine the value of b needed to achieve bit-error probability $P_b = 10^{-5}$ knowing that the channel corrupts the transmitted signal with additive white Gaussian noise of power spectral density $\frac{N_0}{2} = 10^{-2}$ W/Hz.

Problem 6. *(Bit-error probability)*

A discrete memoryless source produces bits at a rate 10^6 bps. The bits, which are uniformly distributed and i.i.d., are grouped into pairs. Each pair is mapped into a distinct waveform and sent over the AWGN channel of noise power spectral density $\frac{N_0}{2}$. Specifically, the first two bits are mapped into one of the four waveforms shown below with $T_s = 2\times 10^{-6}$ seconds, the next two bits are mapped onto the same set of waveforms delayed by T_s, etc.



(a) Describe an orthonormal basis for the inner product space $\mathcal{W}$ spanned by $w_i(t)$, $i = 0, \dots, 3$ and plot the signal constellation in $\mathbb{R}^n$, where n is the dimensionality of $\mathcal{W}$.

(b) Determine an assignment between pairs of bits and waveforms such that the bit-error probability is minimized and derive an expression for P_b.

(c) Draw a block diagram of the receiver that achieves the above P_b using a single causal filter.

(d) Determine the energy per bit $\mathcal{E}_b$ and the power of the transmitted signal.

Problem 7. *(m-ary frequency-shift keying)*

m-ary frequency-shift keying (m-FSK) is a signaling method that uses signals of the form

$$w_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos\left(2\pi\left(f_c + i\Delta f\right) t\right) \mathbb{1}_{[0,T)}(t), \quad i = 0, \dots, m-1,$$

where $\mathcal{E}, T, f_c, \Delta f$ are fixed parameters, with $\Delta f \ll f_c$.

(a) Determine the average energy $\mathcal{E}$. (You can assume $f_c T \in \mathbb{N}$.)

(b) Assuming $f_c T \in \mathbb{N}$, find the smallest value of Δf that makes $w_i(t)$ orthogonal to $w_j(t)$ when $i \neq j$.

(c) In practice the signals $w_i(t)$, $i = 0, \dots, m-1$ can be generated by changing the frequency of a single oscillator. In passing from one frequency to another, a phase shift θ is introduced. Again, assuming $f_c T \in \mathbb{N}$, determine the smallest value of Δf that ensures orthogonality between $\cos\left(2\pi\left(f_c + i\Delta f\right)t + \theta_i\right)$ and $\cos\left(2\pi\left(f_c + j\Delta f\right)t + \theta_j\right)$ whenever $i \neq j$, regardless of θ_i and θ_j.

(d) Sometimes we do not have complete control over f_c either, in which case it is not possible to set $f_c T \in \mathbb{N}$. Argue that if we choose $f_c T \gg 1$, then for all practical purposes the signals will be orthogonal to one another if the condition found in part (c) is met.

(e) Give an approximate value for the bandwidth occupied by the signal constellation. How does the WT product behave as a function of $k = \log_2(m)$?

Problem 8. *(Packing rectangular pulses)*

This exercise is an interesting variation to Example 4.9. Let $\psi(t) = \frac{1}{\sqrt{T_s}}\mathbb{1}_{\left[-\frac{T_s}{2}, \frac{T_s}{2}\right]}(t)$ be a normalized rectangular pulse of duration T_s and let $\psi_{\mathcal{F}}(f) = \sqrt{T_s}\,\mathrm{sinc}(T_s f)$ be its Fourier transform. The collection $\{\psi_l(t)\}_{l=1}^{n}$, where $\psi_l(t) = \psi(t - lT_s)$, forms an orthonormal set. (This is obvious from the time domain.) It has dimension n by construction.

(a) For the set $\mathcal{G}$ spanned by the above orthonormal basis, determine the relationship between n and WT.

(b) Compare with Example 4.9 and explain the difference.

Problem 9. *(Time- and frequency-limited orthonormal sets)*

Complement Example 4.9 and Problem 8 with similar examples in which the shifts occur in the frequency domain. The corresponding time-domain signals can be complex-valued.

Problem 10. *(Root-mean-square bandwidth)*

The *root-mean-square* bandwidth (abbreviated *rms bandwidth*) of a finite-energy lowpass signal $g(t)$ is defined by

$$B_{rms} = \left[\frac{\int_{\mathbb{R}} f^2 |g_{\mathcal{F}}(f)|^2 df}{\int_{\mathbb{R}} |g_{\mathcal{F}}(f)|^2 df}\right]^{\frac{1}{2}},$$

where $|g_{\mathcal{F}}(f)|^2$ is the energy spectral density of the signal. Correspondingly, rms duration of the signal is defined by

$$T_{rms} = \left[\frac{\int_{\mathbb{R}} t^2 |g(t)|^2 dt}{\int_{\mathbb{R}} |g(t)|^2 dt}\right]^{\frac{1}{2}}$$

We want to show that, with the above definitions and assuming that $|g(t)| \to 0$ faster than $1/\sqrt{|t|}$ as $|t| \to \infty$, the time-bandwidth product satisfies

$$T_{rms} B_{rms} \geq \frac{1}{4\pi}$$

(a) Use the Cauchy–Schwarz inequality and the fact that for any $c \in \mathbb{C}$, $c + c^* = 2\Re\{c\} \leq 2|c|$ to prove that

$$\left\{\int_{\mathbb{R}} [g_1^*(t)g_2(t) + g_1(t)g_2^*(t)]\, dt\right\}^2 \leq 4\int_{\mathbb{R}} |g_1(t)|^2 dt \int_{\mathbb{R}} |g_2(t)|^2 dt$$

(b) Insert $g_1(t) = tg(t)$ and $g_2(t) = \frac{dg(t)}{dt}$ in the inequality above to show that

$$\left[\int_{\mathbb{R}} t\frac{d}{dt}[g(t)g^*(t)]\, dt\right]^2 \leq 4\int_{\mathbb{R}} t^2|g(t)|^2 dt \int_{\mathbb{R}} \left|\frac{dg(t)}{dt}\right|^2 dt$$

(c) Integrate the left-hand side by parts and use the fact that $|g(t)| \to 0$ faster than $1/\sqrt{|t|}$ as $|t| \to \infty$ to obtain

$$\left[\int_{\mathbb{R}} |g(t)|^2 dt\right]^2 \leq 4\int_{\mathbb{R}} t^2|g(t)|^2 dt \int_{\mathbb{R}} \left|\frac{dg(t)}{dt}\right|^2 dt$$

(d) Argue that the above is equivalent to

$$\int_{\mathbb{R}} |g(t)|^2 dt \int_{\mathbb{R}} |g_{\mathcal{F}}(f)|^2 df \leq 4\int_{\mathbb{R}} t^2|g(t)|^2 dt \int_{\mathbb{R}} 4\pi^2 f^2 |g_{\mathcal{F}}(f)|^2 df$$

(e) Complete the proof to obtain $T_{rms}B_{rms} \geq \frac{1}{4\pi}$.

(f) As a special case, consider a Gaussian pulse defined by $g(t) = e^{-\pi t^2}$. Show that for this signal $T_{rms}B_{rms} = \frac{1}{4\pi}$, i.e. the above inequality holds with equality.
Hint: $e^{-\pi t^2} \overset{\mathcal{F}}{\longleftrightarrow} e^{-\pi f^2}$

Problem 11. *(Real basis for complex space)*
Let $\mathcal{G}$ be a complex inner product space of finite-energy waveforms with the property that $g(t) \in \mathcal{G}$ implies $g^*(t) \in \mathcal{G}$.

(a) Let $\mathcal{G}_R$ be the subset of $\mathcal{G}$ that contains only real-valued waveforms. Argue that $\mathcal{G}_R$ is a real inner product space.

(b) Prove that if $g(t) = a(t) + \mathrm{j}b(t)$ is in $\mathcal{G}$, then both $a(t)$ and $b(t)$ are in $\mathcal{G}_R$.

(c) Prove that if $\{\psi_1(t), \ldots, \psi_n(t)\}$ is an orthonormal basis for the real inner product space $\mathcal{G}_R$, then it is also an orthonormal basis for the complex inner product space $\mathcal{G}$.

Comment: In this exercise we have shown that we can always find a real-valued orthonormal basis for an inner product space $\mathcal{G}$ such that $g(t) \in \mathcal{G}$ implies $g^*(t) \in \mathcal{G}$. An equivalent condition is that if $g(t) \in \mathcal{G}$, then also the inverse Fourier transform of $g_{\mathcal{F}}^*(-f)$ is in $\mathcal{G}$. The set $\mathcal{G}$ of complex-valued finite-energy waveforms that are strictly time-limited to $\left(-\frac{T}{2}, \frac{T}{2}\right)$ and bandlimited to $(-B, B)$ (for any of the bandwidth definitions given in Appendix 4.9) fulfills the stated conjugacy condition.

Problem 12. *(Average energy of PAM)*

Let U be a random variable uniformly distributed in $[-a, a]$ and let S be a discrete random variable independent of U and uniformly distributed over the PAM constellation $\{\pm a, \pm 3a, \ldots, \pm(m-1)a\}$, where m is an even integer. Let $V = S + U$.

(a) Find the distribution of V.

(b) Find the variance of U and that of V.

(c) Use part (b) to determine the variance of S. Justify your steps.
Comment: By finding the variance of S, we have found the average energy of the PAM constellation used with uniform distribution.

Problem 13. *(Bandwidth)*

Verify the following statements.

(a) The absolute bandwidth of $\operatorname{sinc}\left(\frac{t}{T_s}\right)$ is $B = \frac{1}{2T_s}$.

(b) The 3-dB bandwidth of an RC lowpass filter is $B = \frac{1}{2\pi RC}$.
Hint: The impulse response of an RC lowpass filter is $h(t) = \frac{1}{RC}\exp\left(-\frac{t}{RC}\right)\mathbb{1}_{\mathbb{R}_+}(t)$. The squared magnitude of its Fourier transform is $|h_{\mathcal{F}}(f)|^2 = \frac{1}{1+(2\pi RCf)^2}$.

(c) The η-bandwidth of an RC lowpass filter is $B = \frac{1}{2\pi RC}\tan\left(\frac{\pi}{2}(1-\eta)\right)$.

(d) The zero-crossing bandwidth of $\mathbb{1}_{\left[-\frac{T_s}{2}, \frac{T_s}{2}\right]}(t)$ is $B = \frac{2}{T_s}$.

(e) The equivalent noise bandwidth of an RC lowpass filter is $B = \frac{1}{4RC}$.

(f) The RMS bandwidth of $h(t) = e^{-\pi t^2}$ is $B = \frac{1}{\sqrt{4\pi}}$.
Hint: $h_{\mathcal{F}}(f) = e^{-\pi f^2}$.

Problem 14. *(Antipodal signaling and Rayleigh fading)*

Consider using antipodal signaling, i.e. $w_0(t) = -w_1(t)$, to communicate 1 bit across a Rayleigh fading channel that we model as follows. When $w_i(t)$ is transmitted the channel output is

$$R(t) = Aw_i(t) + N(t),$$

where $N(t)$ is white Gaussian noise of power spectral density $\frac{N_0}{2}$ and A is a random variable of probability density function

$$f_A(a) = 2ae^{-a^2}\mathbb{1}_{\mathbb{R}_+}(a)$$

We assume that, unlike the transmitter, the receiver knows the realization of A. We also assume that the receiver implements a maximum likelihood decision, and that the signal's energy is $\mathcal{E}_b$.

(a) Describe the receiver.

(b) Determine the error probability conditioned on the event $\{A = a\}$.

(c) Determine the unconditional error probability P_f. (The subscript stands for fading.)

(d) Compare P_f to the error probability P_e achieved by an ML receiver that observes $R(t) = mw_i(t) + N(t)$, where $m = \mathbb{E}[A]$. Comment on the different behavior of the two error probabilities. For each of them, find the $\frac{\mathcal{E}_b}{N_0}$ value necessary to obtain the error probability 10^{-5}.

Hint: Use $\frac{1}{2}\exp\left(-\frac{1}{2}x^2\right)$ as an approximation of $Q(x)$.

Problem 15. *(Non-white Gaussian noise)*

Consider the following transmitter/receiver design problem for an additive non-white Gaussian noise channel.

(a) Let the hypothesis H be uniformly distributed in $\mathcal{H} = \{0, \dots, m-1\}$ and when $H = i$, $i \in \mathcal{H}$, let $w_i(t)$ be the channel input. The channel output is then

$$R(t) = w_i(t) + N(t),$$

where $N(t)$ is Gaussian noise of power spectral density $G(f)$, where we assume that $G(f) \neq 0$ for all f. Describe a receiver that, based on the channel output $R(t)$, decides on the value of H with least probability of error.

Hint: Find a way to transform this problem into one that you can solve.

(b) Consider the setting as in part (a), except that now you get to design the signal set with the restrictions that $m = 2$ and that the average energy cannot exceed $\mathcal{E}$. We also assume that $|G(f)|^2$ is constant in the interval $[a, b], a < b$, where it also achieves its global minimum. Find two signals that achieve the smallest possible error probability under an ML decoding rule.

Problem 16. *(Continuous-time AWGN capacity)*

To prove the formula for the capacity C of the continuous-time AWGN channel of noise power spectral density $\frac{N_0}{2}$ when signals are power-limited to P and frequency-limited to $\left(-\frac{W}{2}, \frac{W}{2}\right)$, we first derive the capacity C_d for the discrete-time AWGN channel of noise variance σ^2 and symbols constrained to average energy not exceeding $\mathcal{E}_s$. The two expressions are:

$$C_d = \frac{1}{2}\log_2\left(1 + \frac{\mathcal{E}_s}{\sigma^2}\right) \text{ [bits per channel use]}$$

$$C = \frac{W}{2}\log_2\left(1 + \frac{P}{W\frac{N_0}{2}}\right) \text{ [bps]}$$

To derive C_d we need tools from information theory. However, going from C_d to C using the relationship $n = WT$ is straightforward. To do so, let $\mathcal{G}_\eta$ be the set of all signals that are frequency-limited to $\left(-\frac{W}{2}, \frac{W}{2}\right)$ and time-limited to $\left(-\frac{T}{2}, \frac{T}{2}\right)$ at level η. We choose η small enough that for all practical purposes all signals of $\mathcal{G}_\eta$ are strictly frequency-limited to $\left(-\frac{W}{2}, \frac{W}{2}\right)$ and strictly time-limited to $\left(-\frac{T}{2}, \frac{T}{2}\right)$. Each waveform in $\mathcal{G}_\eta$ is represented by an n-tuple and as $T \to \infty$, $n \to WT$. Complete the argument assuming $n = WT$ and without worrying about convergence issues.

PROBLEM 17. *(Energy efficiency of single-shot PAM)*

This exercise complements what we have learned in Example 4.3. Consider using the m-PAM constellation

$$\{\pm a, \pm 3a, \pm 5a, \dots, \pm(m-1)a\}$$

to communicate across the discrete-time AWGN channel of noise variance $\sigma^2 = 1$. Our goal is to communicate at some level of reliability, say with error probability $P_e = 10^{-5}$. We are interested in comparing the energy needed by PAM versus the energy needed by a system that operates at channel capacity, namely at $\frac{1}{2}\log_2\left(1 + \frac{\mathcal{E}_s}{\sigma^2}\right)$ bits per channel use.

(a) Using the capacity formula, determine the energy per symbol $\mathcal{E}_s^C(k)$ needed to transmit k bits per channel use. (The subscript C stands for channel capacity.) At any rate below capacity, it is possible to make the error probability arbitrarily small by increasing the codeword length. This implies that there is a way to achieve the desired error probability at energy per symbol $\mathcal{E}_s^C(k)$.

(b) Using single-shot m-PAM, we can achieve an arbitrarily small error probability by making the parameter a sufficiently large. As the size m of the constellation increases, the edge effects become negligible, and the average error probability approaches $2Q\left(\frac{a}{\sigma}\right)$, which is the probability of error conditioned on an interior point being transmitted. Find the numerical value of the parameter a for which $2Q\left(\frac{a}{\sigma}\right) = 10^{-5}$.
Hint: Use $\frac{1}{2}\exp\left(-\frac{1}{2}x^2\right)$ as an approximation of $Q(x)$.

(c) Having fixed the value of a, we can use equation (4.1) to determine the average energy $\mathcal{E}_s^P(k)$ needed by PAM to send k bits at the desired error probability. (The superscript P stands for PAM.) Find and compare the numerical values of $\mathcal{E}_s^P(k)$ and $\mathcal{E}_s^C(k)$ for $k = 1, 2, 4$.

(d) Find $\lim_{n\to\infty} \frac{\mathcal{E}_s^C(k+1)}{\mathcal{E}_s^C(k)}$ and $\lim_{n\to\infty} \frac{\mathcal{E}_s^P(k+1)}{\mathcal{E}_s^P(k)}$.

(e) Comment on PAM's efficiency in terms of energy per bit for small and large values of k. Comment also on the relationship between this exercise and Example 4.3.

5 Symbol-by-symbol on a pulse train: Second layer revisited

PROBLEM 1. *(Sampling and reconstruction)*

Here we use the picket fence miracle to investigate practical ways to approximate sampling and/or reconstruction. We assume that for some positive B, $s(t)$ satisfies $s_{\mathcal{F}}(f) = 0$ for $f \notin [-B, B]$. Let T be such that $0 < T \leq \frac{1}{2B}$.

(a) As a reference, review Example 5.15 of Appendix 5.15.

(b) To generate the intermediate signal $s(t)E_T(t)$ of Example 5.15, we need an electrical circuit that produces δ Diracs. Such a circuit does not exist. As a substitute for $\delta(t)$, we use a rectangular pulse of the form $\frac{1}{T_w}\mathbb{1}_{[-\frac{T_w}{2},\frac{T_w}{2})}(t)$, where $0 < T_w \leq T$ and the scaling by $\frac{1}{T_w}$ is to ensure that the integral over the substitute pulse and that over $\delta(t)$ give the same result, namely 1. The intermediate signal at the input of the reconstruction filter is then $[s(t)E_T(t)] \star [\frac{1}{T_w}\mathbb{1}_{[-\frac{T_w}{2},\frac{T_w}{2})}(t)]$. (We can generate this signal without passing through $E_T(t)$.) Express the Fourier transform $y_{\mathcal{F}}(f)$ of the reconstruction filter output.

(c) In the so-called zero-order interpolator, the reconstructed approximation is the step-wise signal $[s(t)E_T(t)] \star \mathbb{1}_{[-\frac{T}{2},\frac{T}{2})}(t)$. This is the intermediate signal of part (b) with $T_w = T$. Express its Fourier transform.

Comment: There is no interpolation filter in this case.

(d) In the first-order interpolator, the reconstructed approximation consists of straight lines connecting the values of the original signal at the sampling points. This can be written as $[s(t)E_T(t)] \star p(t)$, where $p(t)$ is the triangular-shape waveform

$$p(t) = \begin{cases} \frac{T-|t|}{T}, & t \in [-T, T] \\ 0, & \text{otherwise} \end{cases}$$

Express the Fourier transform of the reconstructed approximation.

Compare $s_{\mathcal{F}}(f)$ to the Fourier transform of the various reconstructions you have obtained.

PROBLEM 2. *(Sampling and projections)*

We have seen that the reconstruction formula of the sampling theorem can be rewritten in such a way that it becomes an orthonormal expansion (expression (5.3)). If $\psi_j(t)$ is the jth element of an orthonormal set of functions used to expand $w(t)$, then the jth coefficient c_j equals the inner product $\langle w, \psi_j \rangle$. Explain why we do not need to explicitly perform an inner product to obtain the coefficients used in the reconstruction formula (5.3).

PROBLEM 3. *(Properties of the self-similarity function)*

Prove the following properties of the self-similarity function (5.5). Recall that the self-similarity function of an $\mathcal{L}_2$ pulse $\xi(t)$ is $R_\xi(\tau) = \int \xi(t+\tau)\xi^*(t)dt$.

(a) Value at zero:

$$R_\xi(\tau) \leq R_\xi(0) = \|\xi\|^2, \quad \tau \in \mathbb{R}$$

(b) Conjugate symmetry:

$$R_\xi(-\tau) = R_\xi^*(\tau), \quad \tau \in \mathbb{R}$$

(c) Convolution representation:

$$R_\xi(\tau) = \xi(\tau) \star \xi^*(-\tau), \quad \tau \in \mathbb{R}$$

Comment: The convolution between $a(t)$ and $b(t)$ can be written as $(a \star b)(t)$ or as $a(t) \star b(t)$. Both versions are used in the literature. We prefer the first version, but in the above case the second version does not require the introduction of a name for $\xi^*(-\tau)$.

(d) Fourier relationship:

$$R_\xi(\tau) \text{ is the inverse Fourier transform of } |\xi_{\mathcal{F}}(f)|^2$$

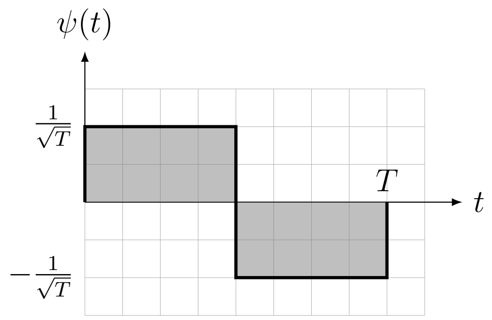
Comment: The fact that $\xi_{\mathcal{F}}(f)$ is in $\mathcal{L}_2$ implies that $|\xi_{\mathcal{F}}(f)|^2$ is in $\mathcal{L}_1$. The Fourier inverse of an $\mathcal{L}_1$ function is continuous. Hence $R_\xi(\tau)$ is continuous.

PROBLEM 4. *(Power spectrum: Manchester pulse)*

Derive the power spectral density of the random process

$$X(t) = \sum_{i \in \mathbb{Z}} X_i \psi(t - iT - \Theta),$$

where $\{X_i\}_{i \in \mathbb{Z}}$ is an i.i.d. sequence of uniformly distributed random variables taking values in $\{\pm\sqrt{\mathcal{E}}\}$, Θ is uniformly distributed in the interval $[0, T]$, and $\psi(t)$ is the so-called Manchester pulse shown below:



The Manchester pulse guarantees that $X(t)$ has at least one transition per symbol, which facilitates the clock recovery at the receiver.

PROBLEM 5. *(Nyquist's criterion)*

For each function $|\psi_{\mathcal{F}}(f)|^2$ in Figure 5, indicate whether the corresponding pulse $\psi(t)$ has unit norm and/or is orthogonal to its time-translates by multiples of T. The function in Figure 5d is $\text{sinc}^2(fT)$.

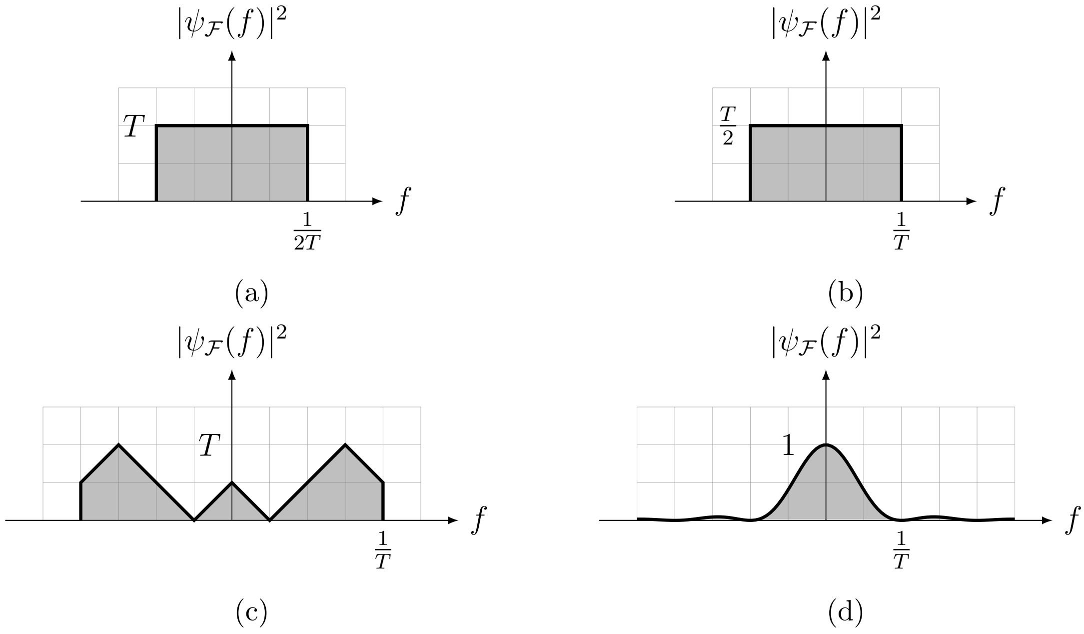


Figure 5

Problem 6. *(Nyquist pulse)*
A communication system uses signals of the form

$$\sum_{l\in\mathbb{Z}} s_l p(t-lT),$$

where $\{s_l\}_{l\in\mathbb{Z}}$ takes values in some symbol alphabet and $p(t)$ is a finite-energy pulse. The transmitted signal is first filtered by a channel impulse response $h(t)$ and then corrupted by additive white Gaussian noise of power spectral density $\frac{N_0}{2}$. The receiver front end is a filter of impulse response $q(t)$.

(a) Neglecting the noise, show that the front-end filter output has the form

$$y(t) = \sum_{l\in\mathbb{Z}} s_l g(t-lT),$$

where $g(t) = (p \star h \star q)(t)$ and $\star$ denotes convolution.

(b) The necessary and sufficient (time-domain) condition that $g(t)$ has to fulfill so that the samples of $y(t)$ satisfy $y(lT) = s_l$, $l \in \mathbb{Z}$, is

$$g(lT) = \delta_l$$

A function that fulfills this condition is called a Nyquist pulse of parameter T. Prove the following theorem:

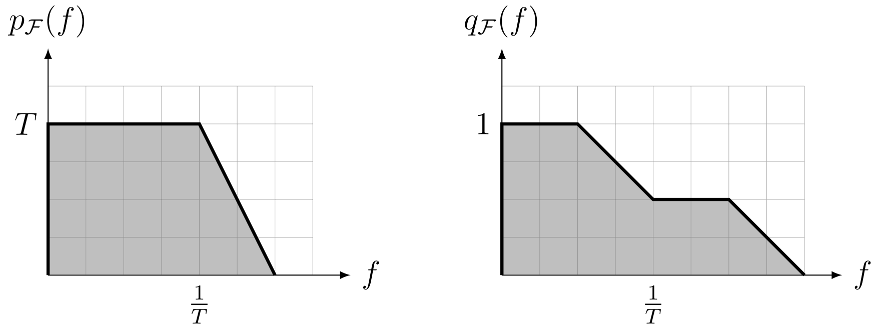
Theorem 1. (Nyquist criterion for Nyquist pulses) The $\mathcal{L}_2$ pulse $g(t)$ is a Nyquist pulse (of parameter T) if and only if its Fourier transform $g_{\mathcal{F}}(f)$ fulfills Nyquist's criterion (with parameter T), i.e.,

$$\operatorname{l.i.m.}\sum_{l\in\mathbb{Z}} g_{\mathcal{F}}\left(f-\frac{l}{T}\right) = T, \quad t \in \mathbb{R}$$

(c) Prove Theorem 5.6 as a corollary to the above theorem.

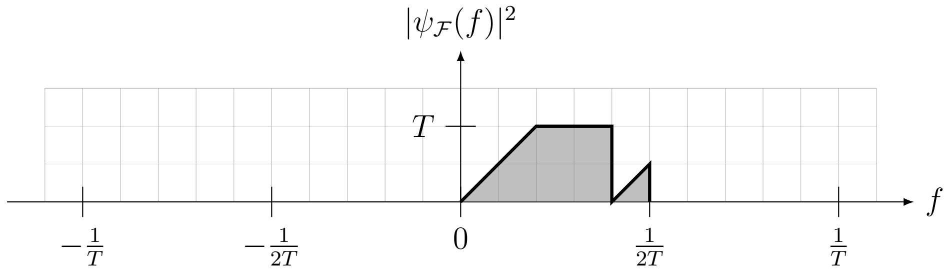
Hint: $\delta_l = \int \psi(t - lT)\psi^*(t)dt$ if and only if the self-similarity function $R_\psi(\tau)$ is a Nyquist pulse of parameter T.

(d) Let $p(t)$ and $q(t)$ be real-valued with Fourier transform as shown below, where only positive frequencies are plotted (both functions being even). The channel frequency response is $h_{\mathcal{F}}(f) = 1$. Determine $y(kT)$, $k \in \mathbb{Z}$.



Problem 7. *(Pulse orthogonal to its T-spaced time translates)*

The figure below shows *part* of the plot of a function $|\psi_{\mathcal{F}}(f)|^2$, where $\psi_{\mathcal{F}}(f)$ is the Fourier transform of some pulse $\psi(t)$.



Complete the plot (for positive and negative frequencies) and label the ordinate, knowing that the following conditions are satisfied:

- For every pair of integers k, l, $\int \psi(t - kT)\psi(t - lT)dt = \delta_{k-l}$;
- $\psi(t)$ is real-valued;
- $\psi_{\mathcal{F}}(f) = 0$ for $|f| > \frac{1}{T}$.

Problem 8. *(Nyquist criterion via picket fence miracle)*

Give an informal proof of 5.6 (Nyquist criterion for orthonormal pulses) using the picket fence miracle (Appendix 5.15).

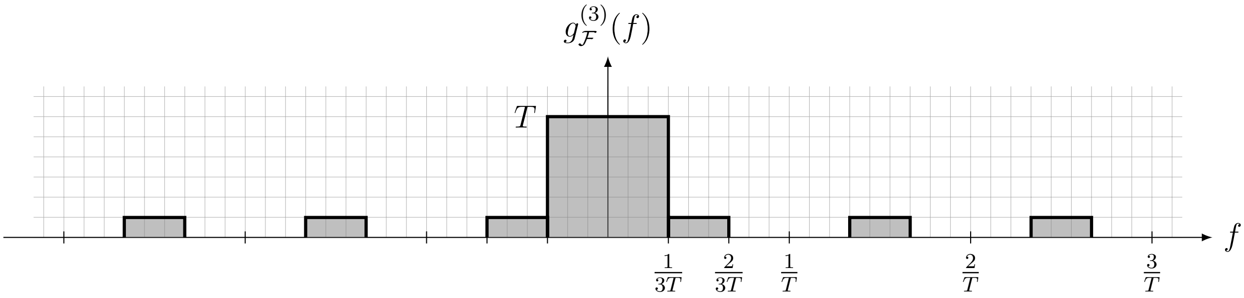
Hint: A function $p(t)$ is a Nyquist pulse of parameter T if and only if $p(t)E_T(t) = \delta(t)$.

PROBLEM 9. *(Peculiarity of Nyquist's criterion)*

Let

$$g_{\mathcal{F}}^{(0)}(f) = T\mathbb{1}_{[-\frac{1}{3T},\frac{1}{3T}]}(f)$$

be the central rectangle in the figure below, and for every positive integer n, let $g_{\mathcal{F}}^{(n)}(f)$ consist of $g_{\mathcal{F}}^{(0)}(f)$ plus $2n$ smaller rectangles of height $\frac{T}{2n}$ and width $\frac{1}{3T}$, each placed in the middle of an interval of the form $[\frac{l}{T}, \frac{l+1}{T}]$, $l = \{-n, \dots, n-1\}$. The figure below shows $g_{\mathcal{F}}^{(3)}(f)$.



(a) Show that for every $n \geq 1$, $g_{\mathcal{F}}^{(n)}(f)$ fulfills Nyquist's criterion with parameter T.
Hint: It is sufficient that you verify that Nyquist's criterion is fulfilled for $f \in [0, \frac{1}{T}]$. Towards that end, first check what happens to the central rectangle when you perform the operation $\sum_{l\in\mathbb{Z}} g_{\mathcal{F}}^{(n)}(f - \frac{l}{T})$. Then see how the small rectangles fill in the gaps.

(b) As n goes to infinity, $g_{\mathcal{F}}^{(n)}(f)$ converges to $g_{\mathcal{F}}^{(0)}(f)$. (It converges for every f and it converges also in $\mathcal{L}_2$, i.e. $\lim_{n\to\infty} \|g_{\mathcal{F}}^{(n)}(f) - g_{\mathcal{F}}^{(0)}(f)\|^2 = 0$.) Peculiar is that the limiting function $g_{\mathcal{F}}^{(0)}(f)$ fulfills Nyquist's criterion with parameter $T^{(0)} \neq T$. What is $T^{(0)}$?

(c) Suppose that we use symbol-by-symbol on a pulse train to communicate across the AWGN channel. To do so, we choose a pulse $\psi(t)$ such that $|\psi_{\mathcal{F}}(f)|^2 = g_{\mathcal{F}}^{(n)}(f)$ for some n, and we choose n sufficiently large that $\frac{T}{2n}$ is much smaller than the noise power spectral density $\frac{N_0}{2}$. In this case, we can argue that our bandwidth B is only $\frac{1}{3T}$. This means a 30% reduction with respect to the minimum absolute bandwidth $\frac{1}{2T}$. This reduction is non-negligible if we pay for the bandwidth we use. How do you explain that such a pulse is not used in practice?
Hint: What do you expect $\psi(t)$ to look like?

(d) Construct a function $g_{\mathcal{F}}(f)$ that looks like the figure above in the shown interval except for the heights of the rectangles. Your function should have infinitely many smaller rectangles on each side of the central rectangle and (like $g_{\mathcal{F}}^{(n)}(f)$) shall satisfy Nyquist's criterion.
Hint: One such construction is suggested by the infinite geometric series $\sum_{i\geq 1} 2^{-i}$, which adds to 1.

PROBLEM 10. *(Raised-cosine expression)*

Let T be a positive integer. Following the steps below, derive the raised-cosine function $|\psi_{\mathcal{F}}(f)|^2$ of roll-off factor $\beta \in (0, 1]$. (It is recommended to plot the various functions.)

(a) Let $p(f) = \cos(f)$, defined over the domain $f \in [0, \pi]$, be the starting point for what will become the right-hand side roll-off edge.

(b) Find constants c and d so that $q(f) = cp(f) + d$ has range $[0, T]$ over the domain $[0, \pi]$.

(c) Find a constant e so that $r(f) = q(ef)$ has domain $[0, \frac{\beta}{T}]$.

(d) Find a constant g so that $s(f) = r(f - g)$ has domain $[\frac{1}{2T} - \frac{\beta}{2T}, \frac{1}{2T} + \frac{\beta}{2T}]$.

(e) Write an expression for the function $|\psi_{\mathcal{F}}(f)|^2$ that has the following properties:

- it is T for $f \in [0, \frac{1}{2T} - \frac{\beta}{2T})$;
- it equals $s(f)$ for $f \in [\frac{1}{2T} - \frac{\beta}{2T}, \frac{1}{2T} + \frac{\beta}{2T}]$;
- it is 0 for $f \in (\frac{1}{2T} + \frac{\beta}{2T}, \infty)$;
- it is an even function.

PROBLEM 11. *(Peculiarity of the sinc pulse)*

Let $\{U_k\}_{k=0}^{n}$ be and i.i.d. sequence of uniformly distributed bits taking value in $\{\pm 1\}$. Prove that for certain values of t and for n sufficiently large, $s(t) = \sum_{k=0}^{n} U_k \mathrm{sinc}(t-k)$ can become larger than any given constant.

Hint: The series $\sum_{k\geq 1} \frac{1}{k}$ diverges, and so does $\sum_{k\geq 1} \frac{1}{k-a}$ for any given constant $a \in (0, 1)$.

Comment: This implies that the eye diagram of $s(t)$ is closed.

PROBLEM 12. *(Matched filter basics)*

Let

$$w(t) = \sum_{k=1}^{K} d_k \psi(t - kT)$$

be a transmitted signal where $\psi(t)$ is a real-valued pulse that satisfies

$$\int \psi(t)\psi(t - kT)dt = \begin{cases} 0, & k \in \mathbb{Z}^* \\ 1, & k = 0, \end{cases}$$

and $d_k \in \{\pm 1\}$.

(a) Suppose that $w(t)$ is filtered at the receiver by the matched filter with impulse response $\psi(-t)$. Show that the filter output $y(t)$ sampled at $mT, m \in \mathbb{Z}$, yields $y(mT) = d_m$, for $1 \leq m \leq K$.

(b) Now suppose that the (noiseless) channel outputs the input plus a delayed and scaled replica of the input. That is, the channel output is $w(t) + \rho w(t - T)$ for some T and some $\rho \in [-1, 1]$. At the receiver, the channel output is filtered by $\psi(-t)$. The resulting waveform $\tilde{y}(t)$ is again sampled at multiples of T. Determine the samples $\tilde{y}(mT)$, for $1 \leq m \leq K$.

(c) Suppose that the kth received sample is $Y_k = d_k + \alpha d_{k-1} + Z_k$, where $Z_k \sim \mathcal{N}(0, \sigma^2)$ and $0 \leq \alpha < 1$ is a constant. Note that d_k and d_{k-1} are realizations of independent random variables that take on the values $\{\pm 1\}$ with equal probability. Suppose that the receiver decides $\hat{d}_k = 1$ if $Y_k > 0$, and decides $\hat{d}_k = -1$ otherwise. Find the probability of error for this receiver.

PROBLEM 13. *(Communication link design)*

Specify the block diagram for a digital communication system that uses twisted copper wires to connect devices that are 5 km apart from each other. The cable has an attenuation of 16 dB/km. You are allowed to use the spectrum between -5 and 5 MHz. The noise at the receiver input is white and Gaussian, with power spectral density $\frac{N_0}{2} = 4.2 \times 10^{-21}$ W/Hz. The required bit-rate is 40 Mbps (megabits per second) and the bit-error probability should be less that 10^{-5}. Be sure to specify the symbol alphabet and the waveform former of the system you propose. Give precise values or bounds for the bandwidth used, the power of the channel input signal, the bit rate, and the error probability. Indicate which bandwidth definition you use.

PROBLEM 14. *(Differential encoding)*

For many years, telephone companies built their networks on twisted pairs. This is a twisted pair of copper wires invented by Alexander Graham Bell in 1881 as a means to mitigate the effect of electromagnetic interference. In essence, an alternating magnetic field induces an electric field in a loop. This applies also to the loop created by two parallel wires connected at both ends. If the wire is twisted, the electric field components that build up along the wire alternate polarity and tend to cancel out one another. If we swap the two contacts at one end of the cable, the signal's polarity at one end is the opposite of that on the other end. Differential encoding is a technique for encoding the information in such a way that the decoding process is not affected by polarity. The differential encoder takes the data sequence $\{D_i\}_{i=1}^n$, here assumed to have independent and uniformly distributed components taking value in $\{0, 1\}$, and produces the symbol sequence $\{X_i\}_{i=1}^n$ according to the following encoding rule:

$$X_i = \begin{cases} X_{i-1}, & D_i = 0 \\ -X_{i-1}, & D_i = 1, \end{cases}$$

where $X_0 = \sqrt{\mathcal{E}}$ be convention. Suppose that the symbol sequence is used to form

$$X(t) = \sum_{i=1}^{n} X_i \psi(t - iT),$$

where $\psi(t)$ is normalized and orthogonal to its T-spaced time-translates. The signal is sent over the AWGN channel of power spectral density $\frac{N_0}{2}$ and at the receiver is passed through the matched filter of impulse response $\psi^*(-t)$. Let Y_i be the filter output at time iT.

(a) Determine $R_X[k]$, $k \in \mathbb{Z}$, assuming an infinite sequence $\{X_i\}_{i\in\mathbb{Z}}$.

(b) Describe a method to estimate D_i from Y_i and Y_{i-1}, such that the performance is the same if the polarity of Y_i is inverted for all i. We ask for a simple decoder, not necessarily ML.

(c) Determine (or estimate) the error probability of your decoder.

PROBLEM 15. *(Mixed questions)*

(a) Consider the signal $x(t) = \cos(2\pi t)\text{sinc}^2(t)$. Assume that we sample $x(t)$ with sampling period T. What is the maximum T that guarantees signal recovery?

(b) You are given a pulse $p(t)$ with spectrum $p_{\mathcal{F}}(f) = \sqrt{T(1-|f|T)}$, $|f| \leq \frac{1}{T}$. What is the value of $\int p(t)p(t-3T)dt$?

Problem 16. *(Properties of the Fourier transform)*

Prove the following properties of the Fourier transform. The sign $\stackrel{\mathcal{F}}{\Longleftrightarrow}$ relates Fourier transform pairs, with the function on the right being the Fourier transform of that on the left. The Fourier transforms of $v(t)$ and $w(t)$ are denoted $v_{\mathcal{F}}(f)$ and $w_{\mathcal{F}}(f)$ respectively.

(a) Linearity:

$$\alpha v(t) + \beta w(t) \stackrel{\mathcal{F}}{\Longleftrightarrow} \alpha v_{\mathcal{F}}(f) + \beta w_{\mathcal{F}}(f)$$

(b) Time-shifting:

$$v(t-t_0) \stackrel{\mathcal{F}}{\Longleftrightarrow} v_{\mathcal{F}}(f)e^{-\mathrm{j}2\pi f t_0}$$

(c) Frequency-shifting:

$$v(t)e^{\mathrm{j}2\pi f_0 t} \stackrel{\mathcal{F}}{\Longleftrightarrow} v_{\mathcal{F}}(f-f_0)$$

(d) Convolution in time:

$$(v \star w)(t) \stackrel{\mathcal{F}}{\Longleftrightarrow} v_{\mathcal{F}}(f)w_{\mathcal{F}}(f)$$

(e) Time scaling by $\alpha \neq 0$:

$$v(\alpha t) \stackrel{\mathcal{F}}{\Longleftrightarrow} \frac{1}{|\alpha|} v_{\mathcal{F}}\left(\frac{f}{\alpha}\right)$$

(f) Conjugation:

$$v^*(t) \stackrel{\mathcal{F}}{\Longleftrightarrow} v_{\mathcal{F}}^*(-f)$$

(g) Time-frequency duality:

$$v_{\mathcal{F}}(t) \stackrel{\mathcal{F}}{\Longleftrightarrow} v(-f)$$

(h) Parseval's relationship:

$$\int v(t)w^*(t)dt \stackrel{\mathcal{F}}{\Longleftrightarrow} \int v_{\mathcal{F}}(f)w_{\mathcal{F}}^*(f)df$$

Comment: As a mnemonic, notice that the above can be written as $\langle v, w\rangle = \langle v_{\mathcal{F}}, w_{\mathcal{F}}\rangle$.

(i) Correlation:

$$\int v(\lambda + t)w^*(\lambda)d\lambda \stackrel{\mathcal{F}}{\Longleftrightarrow} v_{\mathcal{F}}(f)w_{\mathcal{F}}^*(f)$$

Hint: Use Parseval's relationship on the expression on the right and interpret the result.

6 Convolutional coding and Viterbi decoding: First layer revisited

Problem 1. *(Power spectral density)*

Consider the random process

$$X(t) = \sum_{i \in \mathbb{Z}} X_i \sqrt{\mathcal{E}_s} \psi(t - iT_s - T_0),$$

where T_s and $\mathcal{E}_s$ are fixed positive numbers, $\psi(t)$ is some unit-energy function, T_0 is a uniformly distributed random variable taking values in $[0, T_s)$, and $\{X_i\}_{i \in \mathbb{Z}}$ is the output of the convolutional encoder described by

$$X_{2n} = B_n B_{n-2}$$
$$X_{2n+1} = B_n B_{n-1} B_{n-2}$$

with i.i.d. input sequence $\{B_i\}_{i \in \mathbb{Z}}$ taking values in $\{\pm 1\}$.

(a) Express the power spectral density of $X(t)$ for a general $\psi(t)$.

(b) Plot the power spectral density of $X(t)$ assuming that $\psi(t)$ is a unit-norm rectangular pulse of width T_s.

Problem 2. *(Power spectral density: Correlative encoding)*

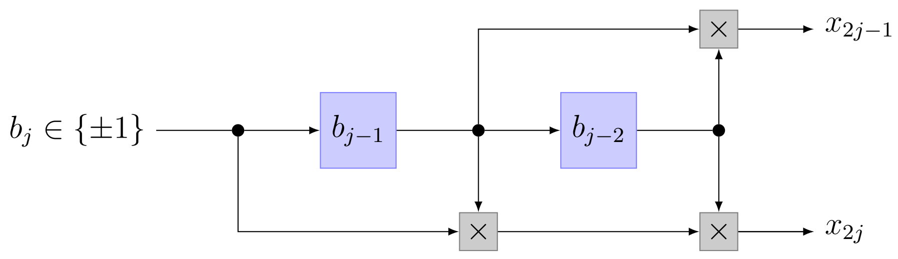
Repeat Problem 1 using the encoder

$$X_i = B_i - B_{i-1}$$

Compare this exercise to Exercise 4 of Chapter 5.

Problem 3. *(Viterbi algorithm)*

An output sequence $(x_1, \ldots, x_{10})$ from the convolutional encoder shown below is transmitted over the discrete-time AWGN channel. The initial and final state of the encoder is $(1, 1)$. Using the Viterbi algorithm, find the maximum likelihood information sequence $(\hat{b}_1, \hat{b}_2, \hat{b}_3, 1, 1)$, knowing that b_1, b_2, b_3 are drawn independently and uniformly from $\{\pm 1\}$ and that the channel output $(y_1, \ldots, y_{10}) = (1, 2, -1, 4, -2, 1, 1, -3, -1, -2)$. (It is for convenience that we are choosing integers rather than real numbers.)



Problem 4. *(Inter-symbol interference)*

From the decoder's point of view, inter-symbol interference (ISI) can be modeled as follows:

$$Y_i = X_i + Z_i \tag{4}$$

$$X_i = \sum_{j=0}^{L} B_{i-j} h_j, \quad i \in \mathbb{N} \tag{5}$$

where B_i is the ith information bit, $h_0, \ldots, h_L$ are coefficients that describe the inter-symbol interference, and Z_i is zero-mean, Gaussian, of variance σ^2, and statistically independent of everything else. Relationship (5) can be described by a trellis, and the ML decision rule can be implemented by the Viterbi algorithm.

(a) Draw the trellis that describes all sequences of the form $X_1, \ldots, X_6$ resulting from information sequences of the form $B_1, \ldots, B_5, 0$, $B_i \in \{0, 1\}$, assuming

$$h_i = \begin{cases} 1, & i = 0 \\ -2, & i = 1 \\ 0, & \text{otherwise} \end{cases}$$

To determine the initial state, you may assume that the preceding information sequence terminated with 0. Label the trellis edges with the input/output symbols.

(b) Specify a metric $f(x_1, \ldots, x_6) = \sum_{i=1}^{6} f(x_i, y_i)$ whose minimization or maximization with respect to the valid $x_1, \ldots, x_6$ leads to a maximum likelihood decision. Specify if your metric needs to be minimized or maximized.

(c) Assume $y_1, \ldots, y_6 = \{2, 0, -1, 1, 0, -1\}$. Find the maximum likelihood estimate of the information sequence $B_1, \ldots, B_5$.

Problem 5. *(Linearity)*

In this exercise, we establish in what sense the encoder of Figure 6.2 in the book is linear.

(a) For this part you might want to review the axioms of a field. Consider the set $\mathcal{F}_0 = \{0, 1\}$ with the following addition and multiplication tables.

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

(The addition in $\mathcal{F}_0$ is the usual addition over $\mathbb{R}$ with the result taken modulo 2. The multiplication is the usual multiplication over $\mathbb{R}$ and there is no need to take the modulo 2 operation since the result is automatically in $\mathcal{F}_0$.) $(\mathcal{F}_0, +, \times)$ form a binary field denoted by $\mathbb{F}_2$. Now consider $\mathcal{F}_- = \{\pm 1\}$ and the following addition and multiplication tables.

$+$	1	-1
1	1	-1
-1	-1	1

$\times$	1	-1
1	1	1
-1	1	-1

(The addition in $\mathcal{F}_-$ is the usual multiplication over $\mathbb{R}$.) Argue that $(\mathcal{F}_-, +, \times)$ form a binary field as well.

Hint: The second set of operations can be obtained from the first set via the transformation $T : \mathcal{F}_0 \to \mathcal{F}_-$ that sends 0 to 1 and 1 to -1. Hence, by construction, for $a, b \in \mathcal{F}_0$, $T(a+b) = T(a) + T(b)$ and $T(a \times b) = T(a) \times T(b)$. Be aware of the double meaning of + and $\times$ in the previous sentence.

(b) For this part you might want to review the notion of a vector space. Let $(\mathcal{F}_0, +, \times)$ be as defined in (a). Let $\mathcal{V} = \mathcal{F}_0^\infty$ be the set of infinite sequences taking values in $\mathcal{F}_0$. Does $(\mathcal{V}, \mathcal{F}_0, +, \times)$ form a vector space? (Addition of vectors and multiplication of a vector with a scalar is done component-wise.) Repeat using $\mathcal{F}_-$.

(c) For this part you might want to review the notion of linear transformation. Let $f : \mathcal{V} \to \mathcal{V}$ be the transformation that sends an infinite sequence $b \in \mathcal{V}$ to an infinite sequence $x \in \mathcal{V}$ according to

$$\begin{aligned} x_{2j-1} &= b_{j-1} + b_{j-2} + b_{j-3} \\ x_{2j} &= b_j + b_{j-2}, \end{aligned}$$

where + is the one defined over the field of scalars implicit in $\mathcal{V}$. Argue that this f is linear.

Comment: When $\mathcal{V} = \mathcal{F}_-^\infty$, this encoder is the one used throughout Chapter 6, with the only difference that in the chapter we multiply over $\mathbb{R}$ rather than adding over $\mathcal{F}_-$, but this is just a matter of notation, the result of the two operations on the elements of $\mathcal{F}_-$ being identical. The standard way to describe a convolutional encoder is to choose $\mathcal{F}_0$ and the corresponding addition, namely addition modulo 2. See Problem 12 for the reason we opt for a non-standard description.

Problem 6. *(Independence of the distance profile from the reference path)*

We want to show that $a(i, d)$ does not depend on the reference path. Recall that in Section 6.4.1 we define $a(i, d)$ as the number of detours that leave the reference path at some arbitrary but fixed trellis depth j and have input distance i and output distance d with respect to the reference path.

(a) Let b and $\bar{b}$, both in $\{\pm 1\}^\infty$, be two infinite-length input sequences to the encoder of Figure 6.2 and let f be the encoding map. The encoder is linear in the sense that componentwise product over the reals $b\bar{b}$ is also a valid input sequence and the corresponding output sequence is $f(b\bar{b}) = f(b)f(\bar{b})$ (see Problem 5). Argue that the distance between b and $\bar{b}$ equals the distance between $b\bar{b}$ and the all-one input sequence. Similarly, argue that the distance between $f(b)$ and $f(\bar{b})$ equals the distance between $f(b\bar{b})$ and the all-one output sequence (which is the output to the all-one input sequence).

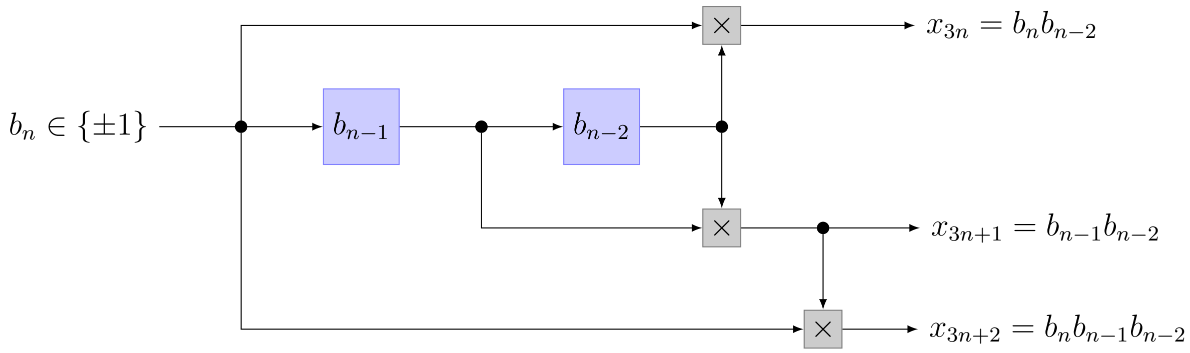
(b) Fix an arbitrary reference path and an arbitrary detour that splits from the reference path at time 0. Let b and $\bar{b}$ be the corresponding input sequences. Because the detour starts at time 0, $b_i = \bar{b}_i$ for $i < 0$ and $b_0 \neq \bar{b}_0$. Argue that $\bar{b}$ uniquely defines a detour $\tilde{b}$ that splits from the all-one path at time 0 and such that:

   (i) the distance between b and $\bar{b}$ is the same as that between $\tilde{b}$ and the all-one input sequence;

   (ii) the distance between $f(b)$ and $f(\bar{b})$ is the same as that between $f(\tilde{b})$ and the all-one output sequence.

(c) Conclude that $a(i, d)$ does not depend on the reference path.

PROBLEM 7. *(Rate* $1/3$ *convolutional code)*



For the convolutional encoder shown above, do the following:

(a) Draw the state diagram and the detour flow graph.

(b) Suppose that the serialized encoder output symbols are scaled so that the resulting energy per bit is $\mathcal{E}_b$ and are sent over the discrete-time AWGN channel of noise variance $\sigma^2 = \frac{N_0}{2}$. Derive an upper bound to the bit-error probability assuming that the decoder implements the Viterbi algorithm.

PROBLEM 8. *(Rate* $2/3$ *convolutional code)*

The following equations describe the output sequence of a convolutional encoder that in each epoch takes $k_0 = 2$ input symbols from $\{\pm 1\}$ and outputs $n_0 = 3$ symbols from the same alphabet.

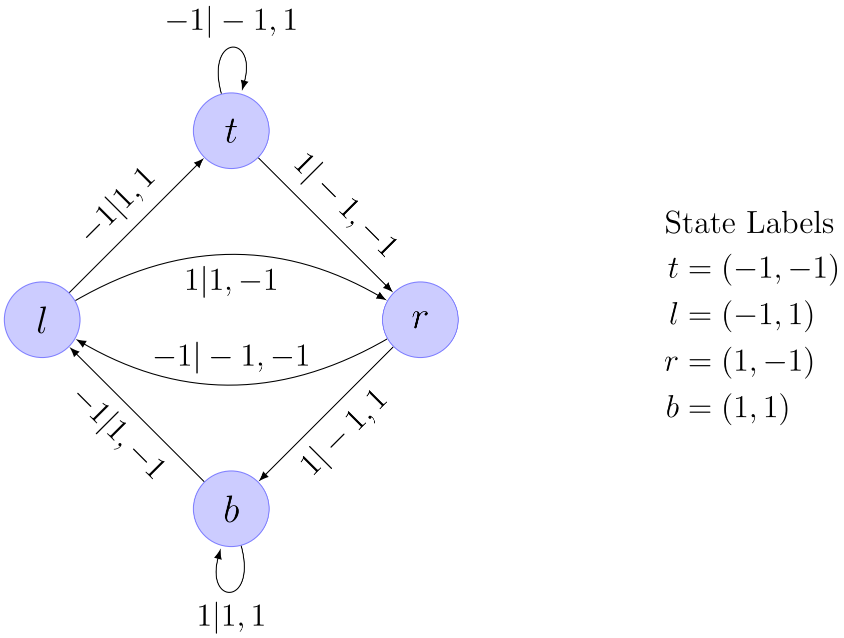
$$\begin{aligned} x_{3n} &= b_{2n} b_{2n-1} b_{2n-2} \\ x_{3n+1} &= b_{2n+1} b_{2n-2} \\ x_{3n+2} &= b_{2n+1} b_{2n} b_{2n-2} \end{aligned}$$

(a) Draw an implementation of the encoder based on delay elements and multipliers.

(b) Draw the state diagram.

(c) Suppose that the serialized encoder output symbols are scaled so that the resulting energy per bit is $\mathcal{E}_b$ and are sent over the discrete-time AWGN channel of noise variance $\sigma^2 = \frac{N_0}{2}$. Derive an upper bound to the bit-error probability assuming that the decoder implements the Viterbi algorithm.

PROBLEM 9. *(Convolutional encoder, decoder, and error probability)*



For the convolutional code described by the state diagram shown above:

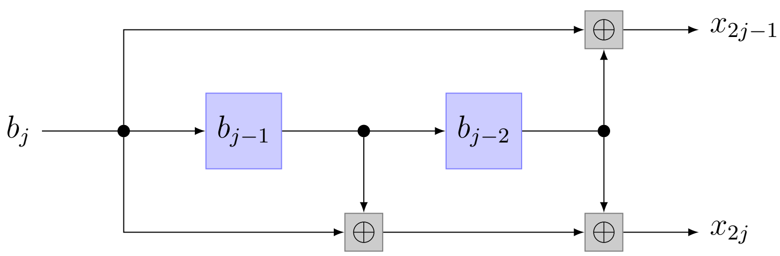
(a) draw the encoder;

(b) as a function of the energy per bit $\mathcal{E}_b$, upper bound the bit-error probability of the Viterbi algorithm when the scaled encoder output sequence is transmitted over the discrete-time AWGN channel of noise variance $\sigma^2 = \frac{N_0}{2}$.

PROBLEM 10. *(Viterbi for the binary erasure channel)*
Consider the convolutional encoder shown below with inputs and outputs over $\{0,1\}$ and addition modulo 2. Its output is sent over the binary erasure channel described by

$$\begin{aligned} P_{Y|X}(0|0) &= P_{Y|X}(1|1) = 1-\epsilon \\ P_{Y|X}(?|0) &= P_{Y|X}(?|1) = \epsilon \\ P_{Y|X}(1|0) &= P_{Y|X}(0|1) = 0, \end{aligned}$$

where $0 < \epsilon < \frac{1}{2}$.



(a) Draw a trellis section that describes the encoder map.

(b) Derive the branch metric and specify whether a maximum likelihood decoder chooses the path with largest or smallest path metric.

(c) Suppose that the initial encoder state is $(0,0)$ and that the channel output is $\{0,?,?,1,0,1\}$. What is the most likely information sequence?

(d) Derive an upper bound to the bit-error probability.

Problem 11. *(Bit-error probability)*

In the process of upper bounding the bit-error probability, in Section 6.4.2 we make the following step:

$$\begin{aligned}\mathbb{E}[\Omega_j] &\le \sum_{i=1}^{\infty}\sum_{d=1}^{\infty} i\, Q\left(\sqrt{\frac{\mathcal{E}_s d}{\sigma^2}}\right) a(i,d) \\ &\le \sum_{i=1}^{\infty}\sum_{d=1}^{\infty} i z^d a(i,d)\end{aligned}$$

(a) Instead of upper bounding the Q function as done above, use the results of Section 6.4.1 to substitute $a(i,d)$ and d with explicit functions of i and get rid of the second sum. You should obtain

$$P_b \le \sum_{i=1}^{\infty} i\, Q\left(\sqrt{\frac{\mathcal{E}_s(i+4)}{\sigma^2}}\right) 2^{i-1}$$

(b) Truncate the above sum to the first five terms and evaluate it numerically for $\frac{\mathcal{E}_s}{\sigma^2}$ between 2 and 6 dB. Plot the results and compare to Figure 6.8 of the book.

Problem 12. *(Standard description of a convolutional encoder)*

Consider the two encoders of Figure 6, where the map $T : \mathcal{F}_0 \to \mathcal{F}_-$ sends 0 to 1 and 1 to -1. Show that the two encoders produce the same output when their inputs are related by $b_j = T(\bar{b}_j)$.

Hint: For $a, b \in \mathcal{F}_0$, $T(a+b) = T(a) \times T(b)$, where addition is modulo 2 and multiplication is over $\mathbb{R}$.

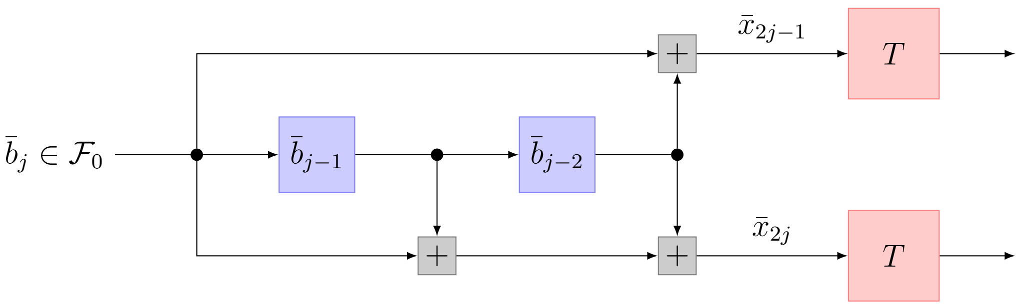
Comment: The encoder of Figure 6b is linear over the field $\mathcal{F}_-$ (see Problem 5), whereas the encoder of Figure 6a is linear over $\mathcal{F}_0$ only if we omit the output map T. The comparison of the two figures should explain why in this chapter we have opted for the description of 6b even though the standard description of a convolutional encoder is as in 6a.

Problem 13. *(Trellis with antipodal signals)*

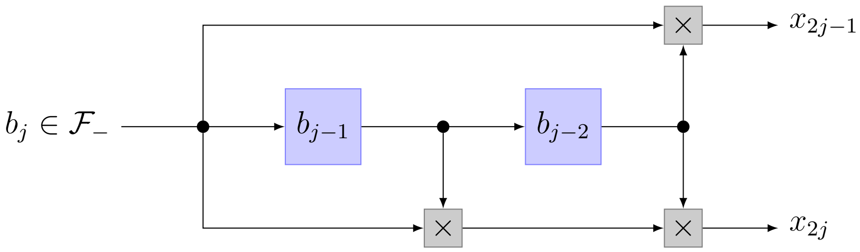
Figure 7a shows a trellis section with the output symbols x_{2j-1}, x_{2j} of a convolutional encoder. Notice how branches that are mirror-images of each other have antipodal output symbols (symbols that are the negative of each other). The purpose of this exercise is to see that when the trellis has this particular structure and codewords are sent through the discrete-time AWGN channel, the maximum likelihood sequence detector further simplifies (with respect to the Viterbi algorithm).

Figure 7b shows two consecutive trellis sections labeled with the branch metric. Notice that the mirror symmetry of 7a implies the same kind of symmetry for 7b. The maximum likelihood path is the one that has the largest path metric. To avoid irrelevant complications we assume that there is only one path that maximizes the path metric.

(a) Let $\sigma_j \in \{\pm 1\}$ be the state visited by the maximum likelihood path at depth j. Suppose that a genie informs the decoder that $\sigma_{j-1} = \sigma_{j+1} = 1$. Write down the necessary and sufficient condition for the maximum likelihood path to go through $\sigma_j = 1$.



(a) Conventional description. Addition is modulo 2.



(b) Description used in the book. Multiplication is over $\mathbb{R}$.

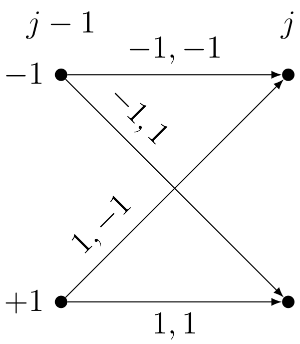
Figure 6

(b) Repeat for the remaining three possibilities of σ_{j-1} and σ_{j+1}. Does the necessary and sufficient condition for $\sigma_j = 1$ depend on the value of σ_{j-1} and σ_{j+1}?

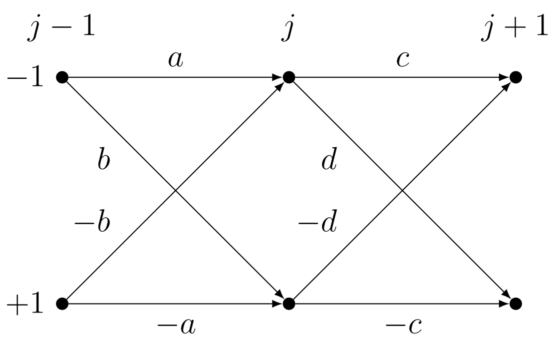
(c) The branch metric for the branch with output symbols x_{2j-1}, x_{2j} is

$$x_{2j-1}y_{2j-1} + x_{2j}y_{2j},$$

where y_j is x_j plus noise. Using the result of the previous part, specify a maximum likelihood sequence decision for $\sigma_j = 1$ based on the observation $y_{2j-1}, y_{2j}, y_{2j+1}, y_{2j+2}$.



(a)



(b)

Figure 7

Problem 14. *(Timing error)*

A transmitter sends

$$X(t) = \sum_i B_i \psi(t - iT),$$

where $\{B_i\}_{i\in\mathbb{Z}}$, $B_i \in \{\pm 1\}$, is a sequence of independent and uniformly distributed bits and $\psi(t)$ is a centered and unit-energy rectangular pulse of width T. The communication channel between the transmitter and the receiver is the AWGN channel of power spectral density $\frac{N_0}{2}$. At the receiver, the channel output $Z(t)$ is passed through a filter matched to $\psi(t)$, and the output is sampled, ideally at times $t_k = kT$, $k \in \mathbb{Z}$.

(a) Consider that there is a timing error, i.e. the sampling time is $t_k = kT - \tau$ where $\frac{\tau}{T} = \frac{1}{4}$. Ignoring the noise, express the matched filter output observation w_k at time $t_k = kT - \tau$ as a function of the bit values b_k and b_{k-1}.

(b) Extending to the noisy case, let $r_k = w_k + z_k$ be the kth matched filter output observation. The receiver is not aware of the timing error. Compute the resulting error probability.

(c) Now assume that the receiver knows the timing error τ (same τ as above) but it cannot correct for it. (This could be the case if the timing error becomes known once the samples are taken.) Draw and label four sections of a trellis that describes the noise-free sampled matched filter output for each input sequence b_1, b_2, b_3, b_4. In your trellis, take into consideration the fact that the matched filter is "at rest" before $x(t) = \sum_{i=1}^{4} b_i \psi(t - iT)$ enters the filter.

(d) Suppose that the sampled matched filter output consists of $\{2, 0.5, 0, -1\}$. Use the Viterbi algorithm to decide on the transmitted bit sequence.

Problem 15. *(Simulation)*

The purpose of this exercise is to determine, by simulation, the bit-error probability of the communication system studied in Chapter 6. For the simulation, we recommend using `MATLAB`, as it has high-level functions for the various tasks, notably for generating a random information sequence, for doing convolutional encoding, for simulating the discrete-time AWGN channel, and for decoding by means of the Viterbi algorithm. Although the actual simulation is on the discrete-time AWGN channel, we specify a continuous-time setup. It is part of your task to translate the continuous-time specifications into what you need for the simulation. We begin with the uncoded version of the system of interest.

(a) By simulation, determine the minimum obtainable bit-error probability P_b of bit-by-bit on a pulse train transmitted over the AWGN channel. Specifically, the channel input signal has the form

$$X(t) = \sum_j X_j \psi(t - jT),$$

where the symbols are i.i.d. and take value in $\{\pm\sqrt{\mathcal{E}_s}\}$, the pulse $\psi(t)$ has unit norm and is orthogonal to its T-spaced time translates. Plot P_b as a function of $\frac{\mathcal{E}_s}{\sigma^2}$ in the range from 2 to 6 dB, where σ^2 is the noise variance. Verify your results with Figure 6.8 of the book.

(b) Repeat with the symbol sequence being the output of the convolutional encoder of Figure 6.2 multiplied by $\sqrt{\mathcal{E}_s}$. The decoder shall implement the Viterbi algorithm. You can once again verify your results with Figure 6.8 of the book.

7 Passband communication via up/down conversion: Third layer

PROBLEM 1. *(Lifting up)*

Let $p(t)$ be real-valued and frequency-limited to $[-B, B]$, where $0 < B < f_c$ for some f_c. Without making any calculations, argue that $p(t)\sqrt{2}\cos(2\pi f_c t)$ and $p(t)\sqrt{2}\sin(2\pi f_c t)$ are orthogonal to each other and have the same norm as $p(t)$.

PROBLEM 2. *(Bandpass filtering in baseband)*

We want to implement a passband filter of impulse response $h(t) = \sqrt{2}\Re\{h_E(t)e^{\mathrm{j}2\pi f_c t}\}$ using baseband filters, where $h_E(t)$ is frequency-limited to $[-B, B]$ and $0 < B < f_c$.

(a) Draw the block diagram of an implementation of the filter of impulse response $h(t)$, based on a filter of impulse response $h_E(t)$ (possible scaled). Your implementation can use an up-converter, a down-converter, and shall behave like the filter of impulse response $h(t)$ for all (passband) input signals of bandwidth not exceeding $2B$ and center frequency f_c.

(b) Draw the box diagram of an implementation that uses only real-valued signals.

PROBLEM 3. *(Equivalent representations)*

A real-valued passband signal $x(t)$ can be written as $x(t) = \sqrt{2}\Re\{x_E(t)e^{\mathrm{j}2\pi f_c t}\}$, where $x_E(t)$ is the baseband-equivalent signal (complex-valued in general) with respect to the carrier frequency f_c. Also, a general complex-valued signal $x_E(t)$ can be written in terms of two real-valued signals, either as $x_E(t) = u(t) + \mathrm{j}v(t)$ or as $x_E(t) = \alpha(t)\exp(\mathrm{j}\beta(t))$.

(a) Show that a real-valued passband signal $x(t)$ can always be written as

$$x_{EI}(t)\cos(2\pi f_c t) - x_{EQ}(t)\sin(2\pi f_c t)$$

and relate $x_{EI}(t)$ and $x_{EQ}(t)$ to $x_E(t)$.

Comment: This formula can be used at the sender to produce $x(t)$ without doing complex-valued operations. The signals $x_{EI}(t)$ and $x_{EQ}(t)$ are called the in-phase and the quadrature components respectively.

(b) Show that a real-valued passband signal $x(t)$ can always be written as

$$a(t)\cos(2\pi f_c t + \theta(t))$$

and relate $a(t)$ and $\theta(t)$ to $x_E(t)$.

Comment: This explains why sometimes people make the claim that a passband signal is modulated in amplitude and in phase.

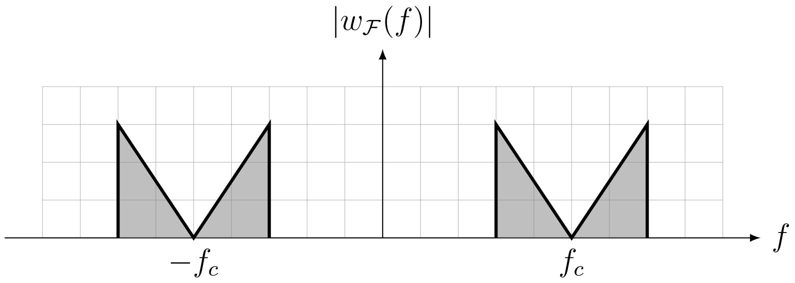
(c) Use part (b) to find the baseband-equivalent of the signal

$$x(t) = A(t)\cos(2\pi f_c t + \varphi),$$

where $A(t)$ is a real-valued lowpass signal. Verify your answer with Example 7.9 where we assumed $\varphi = 0$.

Problem 4. *(Passband)*

Let f_c be a positive carrier frequency and consider an arbitrary real-valued function $w(t)$ whose Fourier transform is shown below:



(a) Argue that there are two different functions, $a_1(t)$ and $a_2(t)$, such that, for $i = \{1, 2\}$,

$$w(t) = \sqrt{2}\Re\{a_i(t)\exp(\mathrm{j}2\pi f_c t)\}$$

This shows that, without some constraint on the input signal, the operation performed by the circuit of Figure 7.4b is not reversible, even in the absence of noise. This was already pointed out in the discussion preceding Lemma 7.8.

(b) Argue that if we limit the input of Figure 7.4b to signals $a(t)$ such that $a_{\mathcal{F}}(f) = 0$ for $f < -f_c$, then the circuit of Figure 7.4a will retrieve $a(t)$ when fed with the output of Figure 7.4b.

(c) Find an example showing that the condition of part (b) is necessary. (Can you find and example with a real-valued $a(t)$?)

(d) Argue that if we limit the input of Figure 7.4b to signals $a(t)$ that are real-valued, then the input of Figure 7.4b can be retrieved from the output.

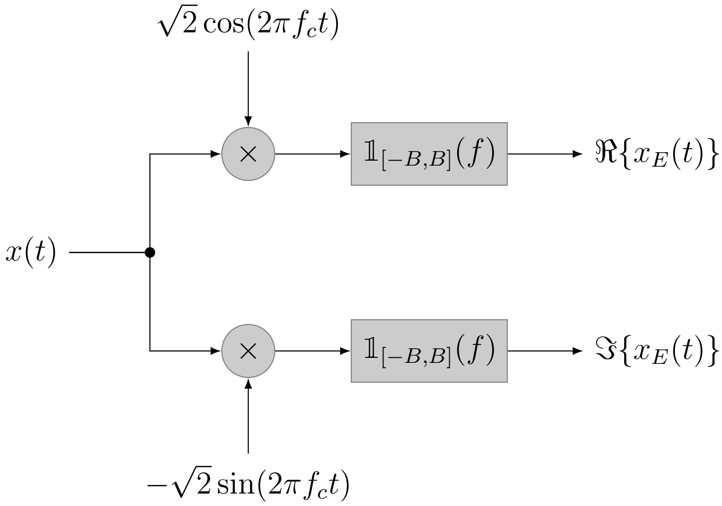
Comment: We are not claiming that the circuit of Figure 7.4a will retrieve $a(t)$.

Hint: You may argue in the time domain or in the frequency domain. If you argue in the time domain, you can assume that $a(t)$ is continuous. If you argue in the frequency domain, you can assume that $a(t)$ has finite bandwidth.

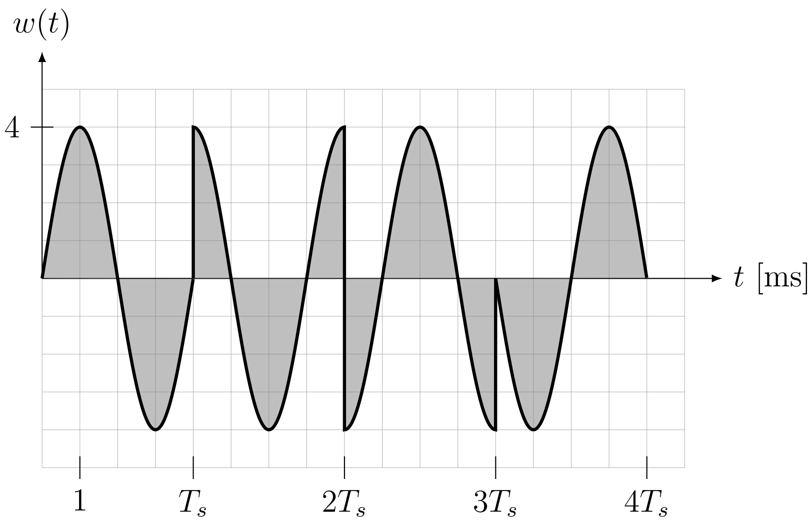
Problem 5. *(From passband to baseband via real-valued operations)*

Let the signal $x_E(t)$ be bandlimited to $[-B, B]$ and let $x(t) = \sqrt{2}\Re\{x_E(t)e^{\mathrm{j}2\pi f_c t}\}$, where $0 < B < f_c$. Show that the circuit shown below recovers the real and imaginary part of $x_E(t)$ when fed with $x(t)$. (The two boxes are ideal lowpass filters of cutoff frequency B.)

Comment: The circuit uses only real-valued operations.



PROBLEM 6. *(Reverse engineering)*



The figure above shows a toy passband signal. (Its carrier frequency is unusually low with respect to its symbol rate.) Specify the three layers of a transmitter that generates the given signal, namely the following:

(a) The carrier frequency f_c used by the up-converter.

(b) The orthonormal basis used by the waveform former to produce the baseband-equivalent signal $w_E(t)$.

(c) The symbol alphabet, seen as a subset of $\mathbb{C}$.

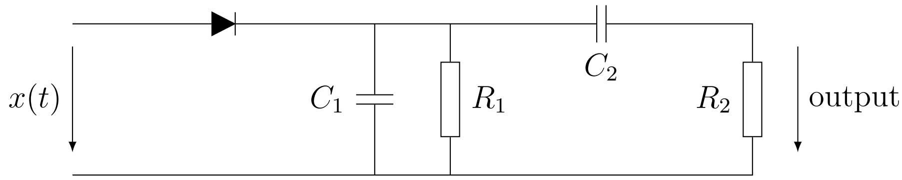
(d) An encoding map, the encoder input sequence that leads to $w(t)$, the bit rate, the encoder output sequence, and the symbol rate.

PROBLEM 7. *(AM receiver)*
Let $x(t) = (1+mb(t))\sqrt{2}\cos(2\pi f_c t)$ be an AM modulated signal as described in Example 7.10. We assume that $1+mb(t) > 0$, that $b(t)$ is bandlimited to $[-B, B]$, and that $f_c > 2B$.

(a) Argue that the envelope of $|x(t)|$ is $(1 + mb(t))\sqrt{2}$ (a drawing will suffice).

(b) Argue that with a suitable choice of components, the output of the figure below is essentially $b(t)$.

Hint: Draw, qualitatively, the voltage on top of R_1 and that on top of R_2.



(c) As an alternative approach, prove that if we pass the signal $|x(t)|$ through an ideal lowpass filter of cutoff frequency f_0, we obtain $1+mb(t)$ scaled by some factor. Specify a suitable interval for f_0.

Hint: Expand $|\cos(2\pi f_c t)|$ as a Fourier series. No need to find explicit values for the Fourier series coefficients.

PROBLEM 8. *(Alternative down-converter)*

Assuming all the $\psi_l(t)$ are bandlimited to $[-B, B]$ and that $0 < B < f_c$, show that the n-tuple former output remains unchanged if we substitute the down-converter of Figure 7.8b with the block diagram of Figure 7.4a.

PROBLEM 9. *(Real-valued implementation)*

Draw a block diagram for the implementation of the transmitter and receiver of Figure 7.8 by means of real-valued operations. Unlike in Figure 7.9, do not assume that the orthonormal basis is real-valued.

PROBLEM 10. *(Circular symmetry)*

(a) Suppose X and Y are real-valued i.i.d. random variables with probability density function $f_X(s) = f_Y(s) = c\exp(-|s|^\alpha)$, where α is a parameter and $c = c(\alpha)$ is the normalizing factor.

  (i) Draw the contour of the joint density function for $\alpha = \frac{1}{2}$, $\alpha = 1$, $\alpha = 2$, and $\alpha = 3$.

  Hint: For simplicity, draw the set of points (x, y) for which $f_{X,Y}(x, y) = c^2(\alpha)e^{-1}$.

  (ii) For which value of α is the joint density function invariant under rotation? What is the corresponding distribution?

(b) In general we can show that if X and Y are i.i.d. random variables and $f_{X,Y}(x, y)$ is circularly symmetric, then X and Y are Gaussian. Use the following steps to prove this.

  (i) Show that if X and Y are i.i.d. and $f_{X,Y}(x, y)$ is circularly symmetric, then $f_X(x)f_Y(y) = \psi(r)$ where ψ is a univariate function and $r = \sqrt{x^2 + y^2}$.

(ii) Take the partial derivative with respect to x and y to show that

$$\frac{f_X'(x)}{xf_X(x)} = \frac{\psi'(r)}{r\psi(r)} = \frac{f_Y'(y)}{yf_Y(y)}$$

(iii) Argue that the only way for the above equalities to hold is that they be equal to a constant value, i.e. $\frac{f_X'(x)}{xf_X(x)} = \frac{\psi'(r)}{r\psi(r)} = \frac{f_Y'(y)}{yf_Y(y)} = -\frac{1}{\sigma^2}$.

(iv) Integrate the above equations and show that X and Y should be Gaussian random variables.

Problem 11. *(Real-valued versus complex-valued constellation)*

Consider 2-PAM and 4-QAM. The source produces i.i.d. and uniformly distributed source bits taking value in $\{\pm 1\}$, and the constellations are $\{\pm 1\}$ and $\{\pm 1 \pm \mathtt{j}\}$ respectively. For 2-PAM, the mapping between the source bits and the channel symbols is the obvious one, i.e. bit b_i is mapped into symbol $s_i = \sqrt{\mathcal{E}_s} b_i$. For the 4-QAM constellation, pairs of bits are mapped into a symbol according to

$$b_{2i}, b_{2i+1} \rightarrow s_i = \sqrt{\mathcal{E}_s}(b_{2i} + \mathtt{j} b_{2i+1})$$

The symbols are mapped into a signal via symbol-by-symbol on a pulse train, where the pulse is real-valued, normalized, and orthogonal to its shifts by multiples of T. The channel adds white Gaussian noise of power spectral density $\frac{N_0}{2}$. The receiver implements an ML decoder. For the two systems, determine (if possible) and compare the following.

(a) The bit-error rate P_b.

(b) The energy per symbol $\mathcal{E}_s$.

(c) The variance σ^2 of the noise seen by the decoder.
Comment: When the symbols are real-valued, the decoder disregards the imaginary part of Y. In this case, what matters is the variance of the real part of the noise.

(d) The symbol-to-noise power ratio $\frac{\mathcal{E}_s}{\sigma^2}$. Write them also as a function of the power P and N_0.

(e) The bandwidth.

(f) The expression for the signals at the output of the waveform former as a function of the bit sequence produced by the source.

(g) The bit rate R.

Summarize, by comparing the two systems from a user's point of view.

Problem 12. *(Smoothness of bandlimited signals)*

We show that a continuous signal of small bandwidth cannot vary much over a small interval. (This fact is used in Problem 13.) Let $w(t)$ be a finite-energy continuous-time passband signal and let $w_E(t)$ be its baseband-equivalent signal. We assume that $w_E(t)$ is bandlimited to $[-B, B]$ for some positive B.

(a) Show that the baseband-equivalent of $w(t-\tau)$ can be modeled as $w_E(t-\tau)e^{\mathrm{j}\phi}$ for some ϕ.

(b) Let $h_{\mathcal{F}}(f)$ be the frequency response of the ideal lowpass filter, i.e. $h_{\mathcal{F}}(f) = \mathbb{1}_{[-B,B]}(f)$. Show that

$$w_E(t+\tau) - w_E(t) = \int w_E(\xi)\left[h(t+\tau-\xi) - h(t-\xi)\right] d\xi$$

(c) Use the Cauchy–Schwarz inequality to prove that

$$|w_E(t+\tau) - w_E(t)|^2 \leq 2\mathcal{E}_w\left[\mathcal{E}_h - R_h(\tau)\right],$$

where

$$R_h(\tau) = \int h(\xi+\tau)h(\xi)d\xi$$

is the self-similarity function of $h(t)$, $\mathcal{E}_h = R_h(0)$ is the energy of $h(t)$, and $\mathcal{E}_w = R_w(0)$ is the energy of $w(t)$.

(d) Show that $R_h(\tau) = (h \star h)(\tau)$.

(e) Show that $R_h(\tau) = h(\tau)$.

(f) Put things together to derive the upper bound

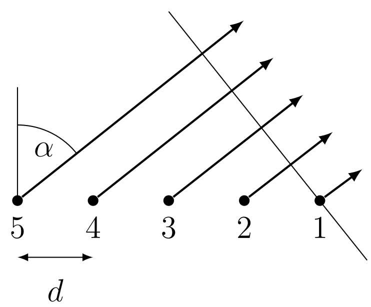
$$|w_E(t+\tau) - w_E(t)| \leq \sqrt{2\mathcal{E}_w\left[\mathcal{E}_h - h(\tau)\right]} = \sqrt{4B\mathcal{E}_w\left(1 - \mathrm{sinc}(2B\tau)\right)}$$

Verify that the bound is tight for $\tau = 0$.

(g) Using part (a) and (f), conclude that if τ is small compared to $\frac{1}{B}$, the baseband-equivalent of $w(t-\tau)$ can be modeled as $w_E(t)e^{\mathrm{j}\phi}$.

Problem 13. *(Antenna array)*

Assume that a transmitter uses an L-element antenna array as shown below for $L = 5$:



The receiving antenna is located in the direction pointed by the arrows, far enough that we can approximate the wavefront as being a straight line. Let $\beta_i w_E(t)$ be the baseband-equivalent signal transmitted at antenna element $i \in \{1, \ldots, L\}$ with respect to some carrier frequency f_c. We assume that each antenna element irradiates isotropically. (More realistically, you can picture each dot as a dipole seen from above and we are interested in the radiation pattern in the plane perpendicular to the dipoles.)

(a) Argue that the received baseband-equivalent signal (at the matched filter input) can be written as

$$r_E(t) = \sum_{i=1}^{L} w_E(t - \tau_i)\beta_i\alpha_i,$$

where $\alpha_i = e^{-\mathrm{j}2\pi f_c \tau_i}$ with $\tau_i = T + i\tau$ for some T and τ. Express τ as a function of the the antenna array's geometric parameters α and d.

(b) We assume that $w_E(t)$ is a continuous bandlimited signal, which implies that for a sufficiently small τ_i, $w_E(t - \tau_i)$ is essentially $w_E(t)$ (see Problem 12). We assume that all τ_i are small enough to satisfy this condition, but that $|f_c\tau_i|$ is not negligible with respect to 1, where f_c is the carrier frequency. Under these assumptions, we model the received baseband-equivalent signal as

$$r_E(t) = \sum_{i=1}^{L} w_E(t)\beta_i\alpha_i$$

plus noise. Choose the vector $\beta = (\beta_1, \dots, \beta_L)^\mathsf{T}$ that maximizes the energy $\int |r_E(t)|^2 dt$, subject to the constraint $\|\beta\| = 1$.

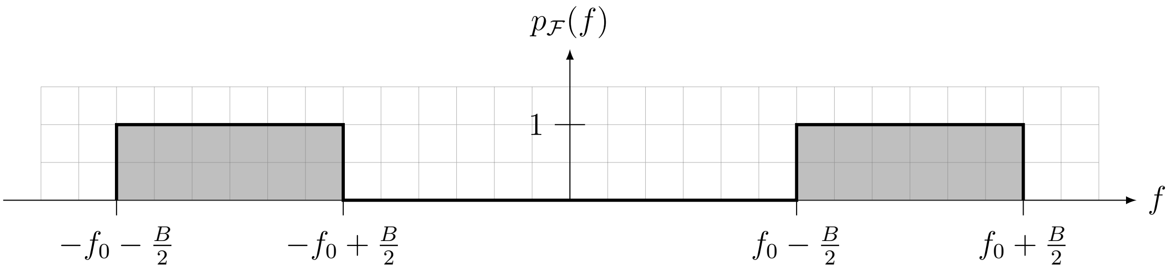
Hint: Use the Cauchy–Schwarz inequality.

(c) Let $\mathcal{E}$ be the energy of $w_E(t)$. Determine the energy of $r_E(t)$ as a function of L when β is selected as in part (b).

(d) In the above problem the received energy grows monotonically with L although $\|\beta\| = 1$ implies that the transmitted energy is constant. Does this violate energy conservation or some other fundamental law of physics?

Problem 14. *(Bandpass pulses)*

Let $p(t)$ be the pulse whose Fourier transform is shown below:



(a) What is the expression for $p(t)$?

(b) Determine the constant c so that $\psi(t) = cp(t)$ has unit energy.

(c) Assume that $f_0 - \frac{B}{2} = B$ and consider the infinite set of functions $\{\psi(t - lT)\}_{l\in\mathbb{Z}}$. Do they form an orthonormal set for $T = \frac{1}{2B}$? (Explain.)

(d) Determine all possible values of $f_0 - \frac{B}{2}$ so that $\{\psi(t - lT)\}_{l\in\mathbb{Z}}$ forms an orthonormal set for $T = \frac{1}{2B}$.

Problem 15. *(Bandpass sampling)*

The Fourier transform of a real-valued signal $w(t)$ satisfies the conjugacy constraint $w_{\mathcal{F}}(f) = w_{\mathcal{F}}^*(-f)$. Hence if $w_{\mathcal{F}}(f)$ is nonzero in some interval (f_L, f_H), $0 \le f_L < f_H$, then it is nonzero also in the interval $(-f_H, -f_L)$. This fact adds a complication to the extension of the sampling theorem to real-valued bandpass signals. Let $\mathcal{D}^+ = (f_L, f_H), \mathcal{D}^- = (-f_H, -f_L)$, and let $\mathcal{D} = \mathcal{D}^- \cup \mathcal{D}^+$ be the passband frequency range of interest. Define $\mathcal{W}$ to be the set of $\mathcal{L}_2$ signals $w(t)$ that are continuous and for which $w_{\mathcal{F}}(f) = 0$ for $f \notin \mathcal{D}$.

(a) Assume $T > 0$ such that

$$\left\{\frac{n}{2T}\right\}_{n\in\mathbb{Z}} \cap \mathcal{D} = \emptyset \tag{6}$$

The above means that $\mathcal{D}^+ \subset [\frac{l}{2T}, \frac{l+1}{2T}]$ for some integer l. Define

$$h_{\mathcal{F}}(f) = \mathbb{1}_{\{|f|\in[\frac{l}{2T},\frac{l+1}{2T}]\}}(f) \quad \text{and} \quad \tilde{w}_{\mathcal{F}}(f) = \sum_{k\in\mathbb{Z}} w_{\mathcal{F}}\left(f - \frac{k}{T}\right),$$

where the latter is the periodic extension of $w_{\mathcal{F}}(f)$. Prove that for all $f \in \mathbb{R}$,

$$w_{\mathcal{F}}(f) = \tilde{w}_{\mathcal{F}}(f) h_{\mathcal{F}}(f)$$

Hint: Write $w_{\mathcal{F}}(f) = w_{\mathcal{F}}^-(f) + w_{\mathcal{F}}^+(f)$ where $w_{\mathcal{F}}^-(f) = 0$ for $f \ge 0$ and $w_{\mathcal{F}}^+(f) = 0$ for $f < 0$ and consider the support of $w_{\mathcal{F}}^+(f - \frac{k}{2T})$ and that of $w_{\mathcal{F}}^-(f - \frac{k}{2T})$, $k \in \mathbb{Z}$.

(b) Prove that when (6) holds, we can write

$$w(t) = \sum_{k\in\mathbb{Z}} T w(kT) h(t - kT),$$

where

$$h(t) = \frac{1}{T}\text{sinc}\left(\frac{t}{2T}\right)\cos(2\pi f_c t)$$

is the inverse Fourier transform of $h_{\mathcal{F}}(f)$ and $f_c = \frac{l}{2T} + \frac{1}{4T}$ is the center frequency of the interval $[\frac{l}{2T}, \frac{l+1}{2T}]$.

Hint: Neglect convergence issues, use the Fourier series to write

$$\tilde{w}_{\mathcal{F}}(f) = \text{l.i.m.} \sum_k A_k e^{j2\pi f T k}$$

and use the result of part (a).

(c) Argue that if (6) is not true, then we can find two distinct signals $a(t)$ and $b(t)$ in $\mathcal{W}$ such that $a(nT) = b(nT)$ for all integers n. Hence (6) is necessary and sufficient to guarantee reconstruction from samples taken every T seconds.

Hint: Show that we can choose $a(t)$ and $b(t)$ in $\mathcal{W}$ such that $\tilde{a}_{\mathcal{F}}(f) = \tilde{b}_{\mathcal{F}}(f)$, where tilde denotes periodic extension as in the definition of $\tilde{w}_{\mathcal{F}}(f)$.

(d) As an alternative characterization, show that (6) is true if and only if

$$\lfloor 2Tf_L \rfloor + 1 = \lceil 2Tf_H \rceil$$

(e) Show that the largest T, denoted $T_{\max}$, that satisfies (6) is

$$T_{\max} = \frac{\lfloor \frac{f_H}{f_H - f_L} \rfloor}{2f_H}$$

Hence $\frac{1}{T_{\max}}$ is the smallest sampling rate that permits the reconstruction of the bandpass signal from its samples.

(f) As an alternative characterization, show that $h(t)$ can be written as

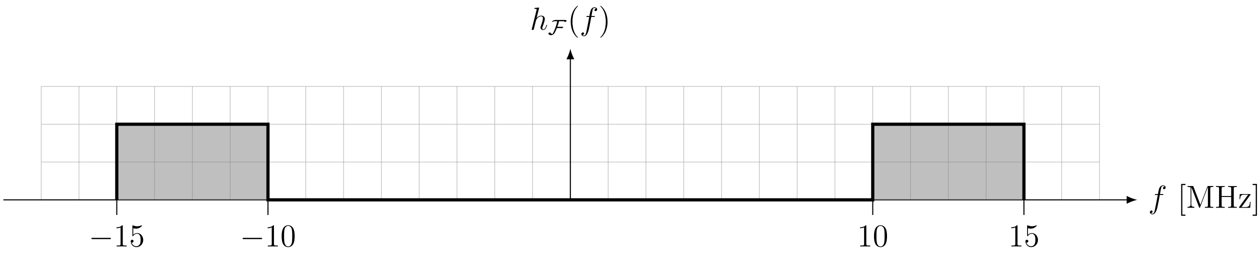
$$h(t) = \left(\frac{l+1}{T}\right) \text{sinc}\left(\frac{(l+1)t}{T}\right) - \frac{l}{T}\text{sinc}\left(\frac{lt}{T}\right)$$

Hint: Rather than manipulating the right side of

$$h(t) = \frac{1}{T}\text{sinc}\left(\frac{t}{2T}\right)\cos(2\pi f_c t),$$

start with a suitable description of $h_{\mathcal{F}}(f)$.

(g) As an application to (6), let $w(t)$ be a continuous finite-energy signal at the output of a filter of impulse response $h_{\mathcal{F}}(f)$ as shown below:



For which of the following sampling frequencies f_s is it possible to reconstruct $w(t)$ from its samples taken every $T = \frac{1}{f_s}$ seconds: $f_s = \{12, 14, 16, 18, 24\}$ MHz?

8 Additional exercises

8.1 Introduction and objectives

Problem 1. *(Gaussian to unit circle)*

Let Z_1 and Z_2 be i.i.d. zero-mean Gaussian random variables, i.e., the pdf of Z_i, $i = \{1, 2\}$ is

$$f_Z(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}$$

for some $\sigma > 0$. Define

$$X := \frac{Z_1}{\sqrt{Z_1^2 + Z_2^2}} \quad \text{and} \quad Y := \frac{Z_2}{\sqrt{Z_1^2 + Z_2^2}}$$

Prove that (X, Y) is a uniformly chosen point on the unit circle.

Problem 2. *(Point on sphere surface)*

Assume you pick a point on the surface of the unit sphere (i.e. the sphere centered at the origin with radius 1) uniformly at random and (X, Y, Z) denotes its coordinates (in 3D space). Compute $\mathbb{E}[X^2]$.

Problem 3. *(Bolt factory)*

In a bolt factory, machines A, B, C manufacture respectively 25%, 35%, and 40% of the total production. Machines A, B, C have defect rates of 5%, 4%, and 2% respectively. A bolt is drawn at random from the produce and is found defective. What are the probabilities that it was manufactured by each machine?

8.2 Receiver design for discrete-time observations: First layer

Problem 1. *(Alternative "Wetterfrosch")*

A TV "weather frog" bases his weather forecast for tomorrow entirely on today's air pressure, which is thus his observable Y. Here, we consider an ambitious weather frog who wants to distinguish three kinds of weather. This means that tomorrow's weather is represented by a random variable H which takes on value 0 if the sun shines tomorrow, 1 if it rains or 2 if the weather is unstable. We assume that the three hypotheses are a priori equally likely, i.e. $P_H(0) = P_H(1) = P_H(2) = 1/3$.

Measurements over several years have led to the following estimate of the probability density function of today's air pressure provided that the sun shines tomorrow,

$$f_{Y|H}(y|0) = \begin{cases} A(1-2y), & 0 \le y \le 0.5 \\ 0, & \text{otherwise} \end{cases}$$

The estimate of the probability density function of today's air pressure provided that it rains tomorrow, is

$$f_{Y|H}(y|1) = \begin{cases} B\left(1 + \frac{y}{2}\right), & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Finally, the estimate of the probability density function of today's air pressure provided that the weather is unstable tomorrow, is

$$f_{Y|H}(y|2) = \begin{cases} C, & 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

The weather frog's goal is to guess the value of H after measuring Y.

(a) Determine A, B and C.

(b) Write down the optimal decision rule (i.e. the rule that minimizes the probability of a wrong forecast) in general terms.

(c) For all values y, draw in one graph $f_{Y|H}(y|0)$, $f_{Y|H}(y|1)$ and $f_{Y|H}(y|2)$. Show on the graph the decision regions corresponding to the optimal decision rule. If we let $\hat{H}(y)$ denote the frog's forecast for a value y of the measurement, can the decision rule be written in the form

$$\hat{H}(y) = \begin{cases} 0, & \text{if } y \leq \theta_1 \\ 2, & \text{if } \theta_1 \leq y < \theta_2 \\ 1, & \text{if } y \geq \theta_2 \end{cases}$$

where θ_1 and θ_2 are some thresholds? If so, determine the values θ_1 and θ_2.

(d) Find the probability of a wrong forecast knowing that tomorrow's weather is unstable, i.e., determine the probability that the decision $\hat{H}$ is different from 2 knowing that, in reality, $H = 2$. This probability is denoted $P_e(2)$.

(e) If we assume that, instead of using the optimal rule, our weather frog always decides that tomorrow's weather is sunny, what will be his probability of error (probability of a wrong forecast)? Explain.

Problem 2. *(Football)*

Consider four teams A, B, C, D playing in a football tournament. There are two rounds in the competition. In the first round there are two matches and the winners progress to play in the final. In the first round A plays against one of the other three teams with equal probability $\frac{1}{3}$ and the remaining two teams play against each other. The probability of A winning against any team depends on the number of red cards r team A got in the previous match. The probabilities of winning for A against B, C, D denoted by p_b, p_c, p_d are $p_b = \frac{0.5}{(1+r)}$, $p_c = p_d = \frac{0.6}{1+r}$. In a match against B, team A will get 1 red card and in a match against C or D, team B will get 2 red cards.
We assume that initially A has 0 red cards and the other teams receive no red cards in the entire tournament. Moreover, teams B, C, D have equal chances to win against each other.

Is betting on team A as the winner a good choice?

Problem 3. *(Minimum-energy signals)*

Consider a given signal constellation consisting of vectors $\{s_1, s_2, \dots, s_m\}$. Let signal s_i occur with probability p_i. In this problem, we study the influence of moving the origin of the coordinate system of the signal constellation. That is, we study the properties of the signal constellation $\{s_1 - a, s_2 - a, \dots, s_m - a\}$ as a function of a.

(a) Draw a sample signal constellation, and draw its shift by a sample vector a.

(b) Does the average error probability P_e depend on the value of a? Explain.

(c) The average energy per symbol depends on the value of a. For a given signal constellation $\{s_1, s_2, \dots, s_m\}$ and given signal probabilities p_i, prove that the value of a that minimizes the average energy per symbol is the centroid (the center of gravity) of the signal constellation, i.e.,

$$a = \sum_{i=1}^{m} p_i s_i$$

Hint: First prove that if X is a real-valued zero-mean random variable and $b \in \mathbb{R}$, then $\mathbb{E}\left[X^2\right] \leq \mathbb{E}\left[(X-b)^2\right]$ with equality iff $b = 0$. Then extend your proof to vectors and consider $X = s - \mathbb{E}[s]$ where $s = s_i$ with probability p_i.

PROBLEM 4. *(One-bit over a binary channel with memory)*

Consider communicating one bit via n uses of a binary channel with memory. The channel output Y_i at time instant i is given by

$$Y_i = X_i \oplus Z_i \qquad i = 1, \dots, n$$

where X_i is the binary channel input, Z_i is the binary noise and $\oplus$ represents modulo-2 addition. The noise sequence is generated as follows: Z_1 is generated from the distribution $Pr\{Z_1 = 1\} = p$ and for $i > 1$,

$$Z_i = Z_{i-1} \oplus N_i$$

where $N_2, \dots, N_n$ are i.i.d. with $Pr\{N_i = 1\} = p$. Let $\left(X_1^{(0)}, \dots, X_n^{(0)}\right)$ and $\left(X_1^{(1)}, \dots, X_n^{(1)}\right)$ denote the codewords (the sequence of symbols sent on the channel) corresponding to the message being 0 and 1 respectively.

(a) Consider the following operation by the receiver. The receiver creates the vector $\left(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n\right)^\mathsf{T}$ where $\hat{Y}_1 = Y_1$ and for $i = 2, 3, \dots, n$, $\hat{Y}_i = Y_i \oplus Y_{i-1}$. Argue that the vector created by the receiver is a sufficient statistic.

Hint: Show that $(Y_1, Y_2, \dots, Y_n)^\mathsf{T}$ can be reconstructed from $\left(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n\right)^\mathsf{T}$.

(b) Write down $\left(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n\right)^\mathsf{T}$ for each of the hypotheses. Notice the similarity with the problem of communicating one bit via n uses of a binary symmetric channel.

(c) How should the receiver choose the codewords $\left(X_1^{(0)}, \dots, X_n^{(0)}\right)$ and $\left(X_1^{(1)}, \dots, X_n^{(1)}\right)$ so as to minimize the probability of error?

Hint: When communicating one bit via n uses of a binary symmetric channel, the probability of error is minimized by choosing two codewords that differ in each component.

PROBLEM 5. *(Gaussian vs. Laplacian noise)*

Consider the following binary hypothesis testing problem. The hypotheses are equally likely and the observable $Y = (Y_1, \dots, Y_n)^\mathsf{T}$ is a n-dimensional real vector whose components are:

$$H_0 : Y_k = Z_k \qquad \text{versus} \qquad H_1 : Y_k = 2A + Z_k, \quad k = 1, \dots, n,$$

where $A > 0$ is a positive constant and $Z_1, \ldots, Z_n$ is an i.i.d. noise sequence. In each of following cases, show that the MAP decision rule reduces to

$$\hat{H}(y_1, \ldots, y_n) = \begin{cases} 0 & \text{if } \sum_{k=1}^{n} \phi(y_i - A) < 0, \\ 1 & \text{otherwise,} \end{cases}$$

and find the function $\phi(\cdot)$.

(a) If Z_k are i.i.d. Gaussian noise samples with zero mean and variance σ^2, i.e.

$$f_{Z_k}(z_k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z_k^2}{2\sigma^2}}, \quad k = 1, \ldots, n.$$

(b) If Z_k are i.i.d. Laplacian noise samples with variance σ^2. That is,

$$f_{Z_k}(z_k) = \frac{1}{\sigma\sqrt{2}} e^{-\frac{\sqrt{2}}{\sigma}|z_k|}, \quad k = 1, \ldots, n.$$

Plot the noise densities for cases (a) and (b) for the same value of σ (take $\sigma = 1$ for convenience). Explain intuitively the difference of the functions $\phi(\cdot)$ that you found in (a) and (b).

PROBLEM 6. *(Antipodal signaling)*

Consider the binary hypothesis testing problem where the hypotheses are equally likely and the observable $Y = (Y_1, \ldots, Y_n)^\mathsf{T}$ is a n-dimensional real vector with components defined as

$$H_0 : Y_k = -A + Z_k \qquad \text{versus} \qquad H_1 : Y_k = A + Z_k, \quad k = 1, \ldots, n,$$

where $A > 0$ is a positive constant and $Z_1, \ldots, Z_n$ are i.i.d. Gaussian noise samples with variance σ^2. Find the decision rule that minimizes the probability of error. Compare your answer with that of Problem 5 part (a). What can you conclude?

PROBLEM 7. *(Ternary hypothesis testing)*

Consider the ternary hypothesis testing problem

$$H_0 : Y = c_0 + Z, \qquad H_1 : Y = c_1 + Z, \qquad H_2 : Y = c_2 + Z,$$

where $Y = [Y_1, Y_2]^\mathsf{T}$ is the two-dimensional observation vector, $c_0 = \sqrt{\mathcal{E}}[1, 0]^\mathsf{T}$, $c_1 = \frac{1}{2}\sqrt{\mathcal{E}}[-1, \sqrt{3}]^\mathsf{T}$, $c_2 = \frac{1}{2}\sqrt{\mathcal{E}}[-1, -\sqrt{3}]^\mathsf{T}$, and $Z = [Z_1, Z_2]^\mathsf{T} \sim \mathcal{N}(0, \sigma^2 I_2)$.

(a) Assuming the three hypotheses are equally likely, draw the optimal decision regions in the (Y_1, Y_2) plane.

(b) Assume now that the apriori probabilities for the hypotheses are $\Pr\{H = 0\} = \frac{1}{2}$, $\Pr\{H = 1\} = \Pr\{H = 2\} = \frac{1}{4}$. Draw the decision regions in the (L_1, L_2) plane where

$$L_i := \frac{f_{Y|H}(Y|i)}{f_{Y|H}(Y|0)}, \quad i = 1, 2.$$

PROBLEM 8. *(Gaussian and uniform distribution)*

A symbol $X \in \{+1, -1\}$ with $\Pr\{X = +1\} = p$ is transmitted through two channels simultaneously. The outputs of the channels are

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= X + Z_2. \end{aligned}$$

Here Z_1 is a random variable that *depends* on X:

$$\begin{aligned} f_{Z_1|X}(z_1|X = +1) &\sim \mathcal{U}[-1, 1] \\ f_{Z_1|X}(z_1|X = -1) &\sim \mathcal{U}[-2, 2], \end{aligned}$$

(where $\mathcal{U}[a, b]$ denotes the uniform distribution on the interval $[a, b]$). The random variable $Z_2 \sim \mathcal{N}(0, 1)$ is independent of both X and Z_1.

(a) Consider a receiver that only observes Y_2. Describe the MAP rule this receiver should implement to estimate X.

(b) Consider now a receiver that observes (Y_1, Y_2). Show that $T = (U, Y_2)$ with

$$U = \begin{cases} -1 & Y_1 < 0 \\ 0 & 0 \le Y_1 \le 1 \\ +1 & Y_1 > 1 \end{cases}$$

is a sufficient statistic to estimate X.

(c) Sketch the decision regions that minimize the probability of error to estimate X on the (y_1, y_2) plane.

(d) Express the probability of error in terms of p and the Q function.

PROBLEM 9. *(Poisson sufficient statistics)*

Consider the hypothesis testing problem where the hypothesis is $H \in \{0, 1, \dots, m-1\}$, and the observable is Y.

(a) Suppose under hypothesis $H = i$, $Y = (Y_1, \dots, Y_n)$ is an i.i.d. sequence of Poisson random variables with parameter $\lambda_i > 0$. That is,

$$P_{Y_k|H}(y_k|i) = \frac{\lambda_i^{y_k}}{(y_k)!} e^{-\lambda_i}, \qquad y_k \in \{0, 1, 2, \dots\}$$

Show that the *sample mean* $T(y_1, \dots, y_n) = \frac{1}{n}\sum_{i=1}^{n} y_i$ is a sufficient statistic.

(b) Suppose under hypothesis $H = i$ the observable $Y = (Y_1, \dots, Y_n)$ is described as

$$Y_k = \theta_i + Z_k, \qquad k = 1, 2, \dots, n,$$

where $Z_k,\ k = 1, 2, \dots, n$ are i.i.d. Exponential random variables with rate $\lambda_i > 0$, i.e.,

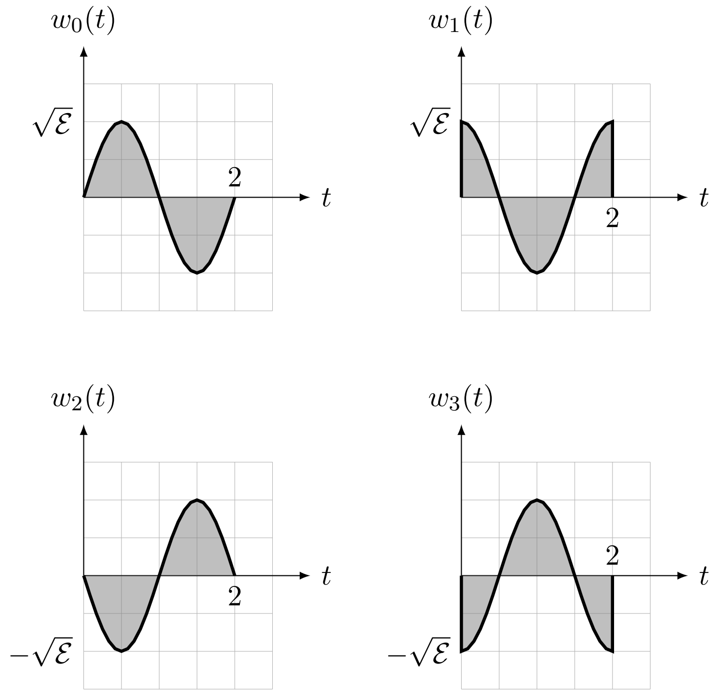
$$f_{Z_k|H}(z_k|i) = \begin{cases} \lambda_i e^{-\lambda_i z_k} & \text{if } z_k \ge 0 \\ 0 & \text{otherwise.} \end{cases}$$

Show that the two-dimensional vector $T(y_1, \dots, y_n) = \left(\min\{y_1, y_2, \dots, y_n\}, \frac{1}{n}\sum_{k=1}^{n} y_k\right)$ is a sufficient statistic.

8.3 Receiver design for the continuous-time AWGN channel: Second layer

Problem 1. *(Phase Shift Keying (PSK))*

Consider the four sinusoid waveforms $w_k(t)$, $k = 0, 1, 2, 3$ represented in the figure below.



(a) Determine an orthonormal basis for the signal space spanned by these waveforms.
Hint: No lengthy calculations needed.

(b) Determine the codewords c_i, $i = 0, 1, 2, 3$ representing the waveforms.

(c) Assume a transmitter sends w_i to communicate a digit $i \in \{0, 1, 2, 3\}$ across a continuous-time AWGN channel of power spectral density $\frac{N_0}{2}$. Write an expression for the error probability of the ML receiver in terms of $\mathcal{E}$ and N_0.

Problem 2. *(On-Off signaling)*

The received signal $R(t)$ in a communication system is given by

$$R(t) = \begin{cases} w(t) + N(t) & \text{if 1 is sent} \\ N(t) & \text{if 0 is sent,} \end{cases}$$

where $N(t)$ is white Gaussian noise of power spectral density $\frac{N_0}{2}$ and $w(t)$ is as shown in Figure 8

At the receiver, the signal $R(t)$ is passed through a filter with impulse response $h(t)$ and the output of the filter is sampled at time t_0 to yield a decision statistic Y. A maximum likelihood decision rule is then used based on Y to decide if 1 or 0 was sent.

(a) For $h(t) = w(4 - t)$, find the error probability if $t_0 = 3$.

(b) Can the error probability in part (a) be improved by choosing t_0 differently?

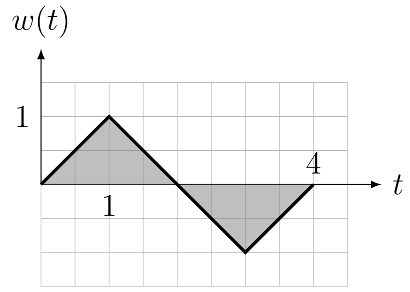


Figure 8

(c) Find the error probability using the following filter with $t_0 = 2$:

$$h(t) = \begin{cases} 1 & 0 \le t \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(d) Can you reduce the error probability in part (c) by sampling the filter output multiple times?

8.4 Signal design trade-offs

PROBLEM 1. *(Exact-energy of PAM)*

In this problem you will compute the average energy $\mathcal{E}(m)$ of m-ary PAM. Throughout the problem, $m \in 2\mathbb{N}$ (positive even integers).

(a) Let U and V be two uniformly distributed discrete random variables that take values in $\mathcal{U} = \{1, 3, \dots, (m-1)\}$ and $\mathcal{V} = \{1, 3, \dots, (m-1)\}$ respectively. Argue that $\mathbb{E}[U^2] = \mathbb{E}[V^2]$.

(b) Consider

$$g(m) = \begin{cases} \sum_{i \in \mathcal{U}} i^2 & m \in 2\mathbb{N} \\ 0 & m = 0 \end{cases}$$

The difference $p(m) = g(m+2) - g(m)$ is a polynomial in m of degree 2. Find $p(m)$.

(c) Even though we are interested in evaluating $g(\cdot)$ only at positive even integers m, our aim is to find a function $g : \mathbb{R} \to \mathbb{R}$. Assuming that such a function exists and has a second derivative, take the second derivative on both sides of $p(m) = g(m+2) - g(m)$ and find a function $g''(m)$ that fulfills the resulting recursion. Then integrate twice and find a general expression for $g(m)$. It will depend on two parameters introduced by the integration.

(d) If you could not solve part (c), you may continue assuming that $g(m)$ has the general form $g(m) = \frac{1}{6}m^3 + am + b$ for some real-valued a and b. Determine $g(0)$ and $g(2)$ directly from the definition of $g(m)$ given in part (b) and use those values to determine a and b.

(e) Express $\mathbb{E}[V^2]$ in terms of the expression you have found for $g(m)$ and verify it for $m = 2, 4, 6$.
Hint: Recall that $\mathbb{E}[V^2] = \mathbb{E}[U^2]$.

(f) More generally, let S be uniformly distributed in $\{\pm d, \pm 3d, \ldots, \pm(m-1)d\}$, where d is an arbitrary positive number and define $\mathcal{E}(d,m) = \mathbb{E}[S^2]$. Use your results found thus far to determine a simple expression for $\mathcal{E}(d,m)$.

(g) Let T be uniformly distributed in $[-md, md]$. Computing $\mathbb{E}[T^2]$ is straightforward, and one expects $\mathbb{E}[S^2]$ to be close to $\mathbb{E}[T^2]$ when m is large. Determine $\mathbb{E}[T^2]$ and compare the result obtained via this continuous approximation to the exact value of $\mathbb{E}[S^2]$.

Problem 2. *(PAM signals)*

Consider using the signal set

$$s_i(t) = s_i \phi(t), \quad i = \{0, 1, \ldots, m-1\},$$

where $\phi(t)$ is a unit-energy waveform, $s_i \in \{\pm\frac{1}{2}d, \pm\frac{3}{2}d, \ldots, \pm\frac{m-1}{2}d\}$, and $m \geq 2$ is an even integer.

(a) Assuming that all signals are equally likely, determine the average energy $\mathcal{E}_s$ as a function of m.
Hint: $\sum_{i=0}^{n} i^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$.

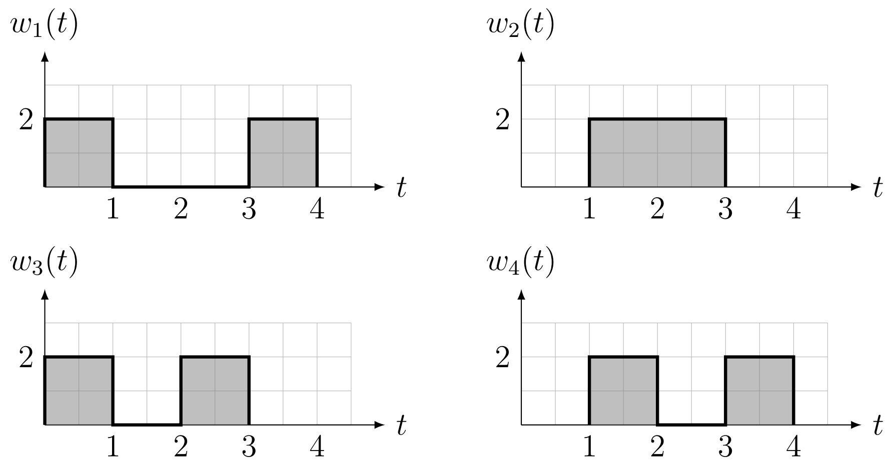
(b) Draw a block diagram for the ML receiver, assuming that the channel is AWGN with power spectral density $\frac{N_0}{2}$.

(c) Give an expression for the error probability.

(d) For large values of m, the probability of error is essentially independent of m, but the energy is not. Let k be the number of bits you send every time you transmit $s_i(t)$ for some i, and rewrite $\mathcal{E}_s$ as a function of k. For large values of k, how does the energy behave when k increases by 1?

Problem 3. *(Shifted signals)*

Consider the signal set shown below. Each signal is equally likely to be chosen for transmission over an AWGN channel with power spectral density $\frac{N_0}{2}$.



(a) Represent the signal set using the four basis signals given by $\psi_1(t) = \psi(t)$, $\psi_2(t) = \psi(t-1)$, $\psi_3(t) = \psi(t-2)$, $\psi_4(t) = \psi(t-3)$, where

$$\psi(t) = \begin{cases} 1 & 0 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) Use the union bound to find an upper bound to the error probability for the optimal receiver.

(c) Transform the four signals by a translation in order to obtain a minimum energy signal set. Sketch the new signal set $\{\tilde{w}_1(t), \tilde{w}_2(t), \tilde{w}_3(t), \tilde{w}_4(t)\}$.

(d) Use the Gram–Schmidt procedure to find an orthogonal basis for $\{\tilde{w}_1(t), \tilde{w}_2(t), \tilde{w}_3(t), \tilde{w}_4(t)\}$.

(e) Find the exact error probability of an optimal receiver designed for $\{\tilde{w}_1(t), \tilde{w}_2(t), \tilde{w}_3(t), \tilde{w}_4(t)\}$.

(f) Based on your answer to (e), what can you say about the error probability of the receiver in (b)?

8.5 Symbol-by-symbol on a pulse train: Second layer revisited

8.6 Convolutional coding and Viterbi decoding: First layer revisited

8.7 Passband communication via up/down conversion: Third layer