ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 17 Principles of Digital Communications
Solutions to Midterm exam Apr. 09, 2025

PROBLEM 1. (12 points)

Consider the following binary hypothesis testing problem. Suppose that under hypothesis $H = i \in \{0, 1\}$, we have

$$(Y_1, Y_2) = c_i + (Z_1, Z_2),$$

where $c_0 = (-1, -1)$ and $c_1 = (+1, +1)$. Define the functions T_1, T_2, T_3 as

$$T_1(Y_1, Y_2) = Y_1 - Y_2, \quad T_2(Y_1, Y_2) = Y_1 + Y_2, \quad T_3(Y_1, Y_2) = \text{sign}(T_2).$$

(a) (3 pts) Suppose Z_1, Z_2 are i.i.d. $\mathcal{N}(0, 1)$. Is T_1 a sufficient statistic? Repeat for T_2 and T_3.

Solution: If Z_1, Z_2 are i.i.d. $\mathcal{N}(0, 1)$, then $f_{Y_1,Y_2|H}(y_1, y_2|i) = \frac{1}{2\pi} \exp\left(-\frac{\|y-c_i\|^2}{2}\right)$, and the likelihood ratio is

$$\frac{f_{Y_1,Y_2|H}(y_1, y_2|1)}{f_{Y_1,Y_2|H}(y_1, y_2|0)} = \exp\left(2(y_1 + y_2)\right).$$

Hence, only T_2 is a sufficient statistic. (It is possible to find values of (y_1, y_2) that have the same values of T_1 and T_3, such as $(-5, -10)$ and $(0, -5)$, as an example, but different values of the likelihood ratio.)

For the rest of the problem, suppose that Z_1, Z_2 are i.i.d. but Laplacian (i.e., each has probability density $f_Z(z) = \frac{1}{2} \exp(-|z|)$).

(b) (2 pts) What are the log likelihood ratios for the observed values $(y_1 = 4, y_2 = 0)$ and $(y_1 = 2, y_2 = 2)$?

Hint: Log likelihood ratio, $\text{LLR}(y_1, y_2)$, is $\ln \frac{f_{Y_1,Y_2|H}(y_1,y_2|1)}{f_{Y_1,Y_2|H}(y_1,y_2|0)}$.

Solution: If Z_1, Z_2 are i.i.d. Laplacian, we have that the log likelihood ratio is

$$\ln \frac{f_{Y_1,Y_2|H}(y_1, y_2|1)}{f_{Y_1,Y_2|H}(y_1, y_2|0)} = |y_1 + 1| - |y_1 - 1| + |y_2 + 1| - |y_2 - 1|,$$

which evaluates to 2 and 4 for the values given, respectively.

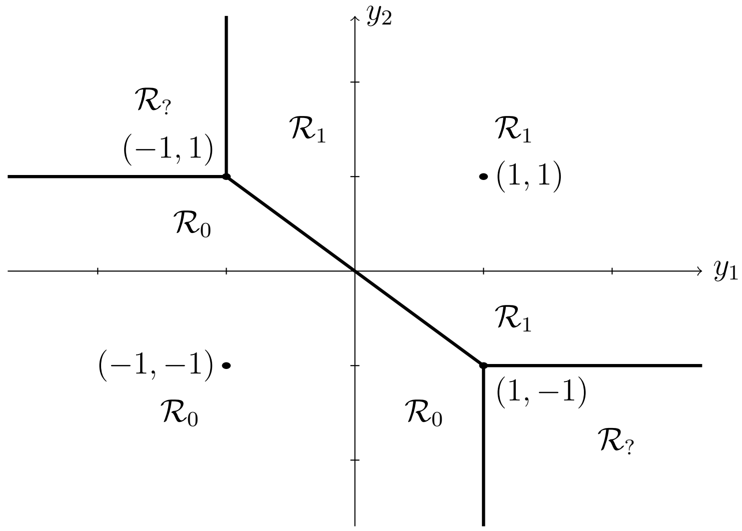
(c) (2 pts) Is T_2 a sufficient statistic? Justify your answer.

Hint: (b) might be useful.

Solution: If T_2 had been a sufficient statistic, then $\text{LLR}(y_1, y_2)$ should have been the same for all (y_1, y_2) with the same value of $T(y_1, y_2) = y_1 + y_2$. But part (b) gives a counter-example to this, and hence, T_2 is not a sufficient statistic.

(d) (3 pts) Show that when $\text{LLR}(y_1, y_2) > 0$ we have $y_1 + y_2 > 0$, and when $\text{LLR}(y_1, y_2) < 0$ we have $y_1 + y_2 < 0$.

Solution: This is identical to problem 2 of problem set 5 (figure with decision regions shown below for reference). We see that $\text{LLR}(y_1, y_2) = 0$ corresponds to the decision region $\mathcal{R}_?$ (using the same notation as the solutions there), and the set of points where $\text{LLR}(y_1, y_2) > 0$, which is $\mathcal{R}_1$, lies completely within the set points such that $y_1 + y_2 > 0$. Similarly, we also see that the set of points where $\text{LLR}(y_1, y_2) < 0$, $\mathcal{R}_0$, lies completely within the set points such that $y_1 + y_2 < 0$.



(e) (2 pts) Suppose the hypotheses are equally likely. One person implements the MAP decision rule $\hat{H}_{\text{MAP}}(Y_1, Y_2)$. Can someone who only observes T_2 implement a decision rule with the same error probability?

Solution: Yes. The MAP rule here coincides with the ML rule, which is to decide 1 if $\text{LLR} > 0$, 0 if $\text{LLR} < 0$, and to choose arbitrarily if $\text{LLR} = 0$. By (d) this can be done based on the value of T_2.

Remarks: We see how the distribution of the noise can affect whether the same functions of the observations remain sufficient statistics or not — T_2 is sufficient when the noise is Gaussian, but it is not when the noise is Laplacian. Nonetheless, we can still evaluate an equivalent decision rule (with the same error probability as the MAP decision rule) using T_2 (even though it is not a sufficient statistic under Laplacian noise).

PROBLEM 2. (16 points)

In a 3-ary hypothesis test with a priori equally likely hypotheses, when the true hypothesis $H = i \in \{0, 1, 2\}$, the observation $Y \in \mathbb{R}^n$ is given by

$$Y = c_i + Z,$$

with $c_i = x_i v$, where $x_0 = -2$, $x_1 = 0$, $x_2 = 2$ are scalars, $v = (1, 1/2, 1/3, \ldots, 1/n) \in \mathbb{R}^n$, and $Z = (Z_1, \ldots, Z_n)$ where $Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0, 1)$.

(a) (3 pts) Let $T_n = Y_1 + \cdots + Y_n$. Consider decision rules based only on the value of T_n. What is the rule that minimizes the error probability? What is the error probability of this rule? [Let $H_n = \sum_{j=1}^{n} \frac{1}{j}$. Write your answer in terms of H_n and the Q-function.]

Solution: The decision rule is to be based on $T_n = \sum_{j=1}^{n} Y_j = x_i \sum_{j=1}^{n} v_{ij} + \sum_{j=1}^{n} Z_j$, which is equal to $H_n x_i + \tilde{Z}$, where $H_n = \sum_{j=1}^{n} j^{-1}$ and $\tilde{Z}$ is a Gaussian random variable with mean 0 and variance n. Thus, the possible values of T_n under the hypotheses $i = 0, 1, 2$ are respectively $-2H_n + \tilde{Z}$, $\tilde{Z}$, and $2H_n + \tilde{Z}$. Hence, the optimal decision rule is the minimum distance rule, which gives

$$\hat{H}(T_n) = \begin{cases} 0 & \text{if } T_n < -H_n, \\ 1 & \text{if } -H_n \le T_n < H_n, \\ 2 & \text{if } H_n \le T_n. \end{cases}$$

Conditioned on i being 0 or 2, the error probability is $P_e(0) = P_e(2) = Q\left(\frac{H_n}{\sqrt{n}}\right)$, and conditioned on $i = 1$, the error probability is $P_e(1) = 2Q\left(\frac{H_n}{\sqrt{n}}\right)$. Hence, the average error probability is

$$P_e = \frac{P_e(0) + P_e(1) + P_e(2)}{3} = \frac{4}{3} Q\left(\frac{H_n}{\sqrt{n}}\right).$$

(b) (2 pts) For which values (if any) of n is T_n a sufficient statistic?

Solution: When $n = 1$, we have $T_1 = Y_1$, and hence, T_1 is clearly a sufficient statistic (it is the entire observation Y). For general n, we can write the likelihood of Y as

$$\begin{aligned} f_{Y_1,\ldots,Y_n|H}(y_1,\ldots,y_n|i) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(\frac{-\|y - c_i\|^2}{2}\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{\|y\|^2 + \|c_i\|^2 - 2\langle y, c_i\rangle}{2}\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{\|y\|^2}{2}\right) \exp\left(\frac{2x_i\langle y, v\rangle - \|c_i\|^2}{2}\right) \\ &= h(y) \exp\left(\frac{2x_i\left(\sum_{j=1}^n \frac{y_j}{j}\right) - \|c_i\|^2}{2}\right), \end{aligned}$$

for some function h of the entire observation $y = (y_1, \ldots, y_n)$. We see that the above term cannot be written in terms of only $y_1 + \cdots + y_n$, hence T_n cannot be a sufficient statistic, by the Fisher–Neyman factorization theorem.

(c) (3 pts) As n gets large, what is the error probability of the decision rule using T_n?
Hint: Facts that might be useful: $H_n = \sum_{j=1}^n j^{-1} \approx \ln(n)$ for large n, $\sum_{j=1}^\infty j^{-2} = \pi^2/6$, $\sum_{j=1}^\infty j^{-3} = 1.2020\ldots$, $\sum_{j=1}^\infty j^{-4} = \pi^4/90$. Other sums available upon request... And $\lim_{n\to\infty} \frac{\ln(n)}{n^a} = 0$ for any $a > 0$.

Solution: From the facts given, we have that $\lim_{n\to\infty} \frac{H_n}{\sqrt{n}} = 0$, and hence the error probability above goes to $\frac{4}{3}Q(0) = \frac{2}{3}$, as Q is a continuous function.

Consider $U_n = Y_1 + \frac{Y_2}{2} + \cdots + \frac{Y_n}{n}$.

(d) (3 pts) Redo (a) with U_n replacing T_n.

Solution: The decision rule is to be based on $U_n = \sum_{j=1}^n \frac{Y_j}{j} = x_i \sum_{j=1}^n \frac{v_{ij}}{j} + \sum_{j=1}^n \frac{Z_j}{j}$. This is equal to $x_i(\sum_{j=1}^n j^{-2}) + Z'$, where Z' is a Gaussian random variable with mean 0 and variance $\sum_{j=1}^n j^{-2}$. Thus, the possible values of U_n under the hypotheses

$i = 0, 1, 2$ are respectively $-2(\sum_{j=1}^n j^{-2}) + Z'$, Z', and $2(\sum_{j=1}^n j^{-2}) + Z'$. Hence, the optimal decision rule is the minimum distance rule, which gives

$$\hat{H}(U_n) = \begin{cases} 0 & \text{if } U_n < -\sum_{j=1}^n j^{-2}, \\ 1 & \text{if } -\sum_{j=1}^n j^{-2} \le U_n < \sum_{j=1}^n j^{-2}, \\ 2 & \text{if } \sum_{j=1}^n j^{-2} \le U_n. \end{cases}$$

Conditioned on i being 0 or 2, the error probability is $P_e(0) = P_e(2) = Q\left(\frac{\sum_{j=1}^n j^{-2}}{\sqrt{\sum_{j=1}^n j^{-2}}}\right) = Q\left(\sqrt{\sum_{j=1}^n j^{-2}}\right)$, and conditioned on $i = 1$, the error probability is $P_e(1) = 2Q\left(\frac{\sum_{j=1}^n j^{-2}}{\sqrt{\sum_{j=1}^n j^{-2}}}\right) = 2Q\left(\sqrt{\sum_{j=1}^n j^{-2}}\right)$. Hence, the average error probability is

$$P_e = \frac{P_e(0) + P_e(1) + P_e(2)}{3} = \frac{4}{3}Q\left(\sqrt{\sum_{j=1}^n j^{-2}}\right).$$

(e) (2 pts) Redo (b) with U_n replacing T_n.

Solution: By the same computation as in part (b), we see that U_n is indeed a sufficient statistic, by the Fisher–Neyman factorization theorem.

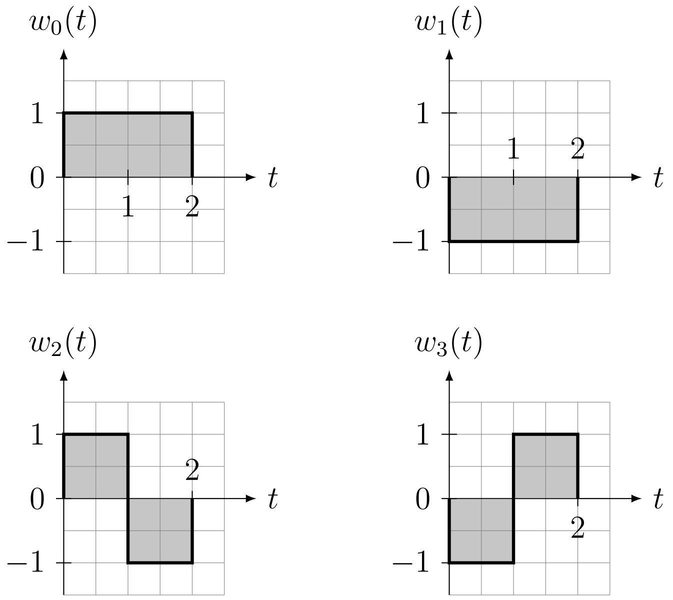
(f) (3 pts) Redo (c) with U_n replacing T_n.

Solution: Using the fact that $\sum_{j=1}^\infty j^{-2} = \frac{\pi^2}{6}$, we have that the above error probability from (d) goes to $\frac{4}{3}Q\left(\frac{\pi}{\sqrt{6}}\right)$.

Remarks: In this problem, U_n is a sufficient statistic and T_n is not. We see that the error probability obtained using T_n is $2/3 \approx 0.67$, which is the same as that of a random guess between three choices. Hence, as n goes to infinity, T_n has no information about the message i. On the other hand, the error probability using U_n is $\frac{4}{3}\left(\frac{\pi}{\sqrt{6}}\right) \approx 0.13$ (it is easy to see that this is certainly smaller than $2/3$, as $\frac{\pi}{\sqrt{6}} > 0$ and $Q(x) < 1/2$ for $x > 0$).

Problem 3. (18 points)

Suppose w_0, w_1, w_2 and w_3 are given in the following waveforms.

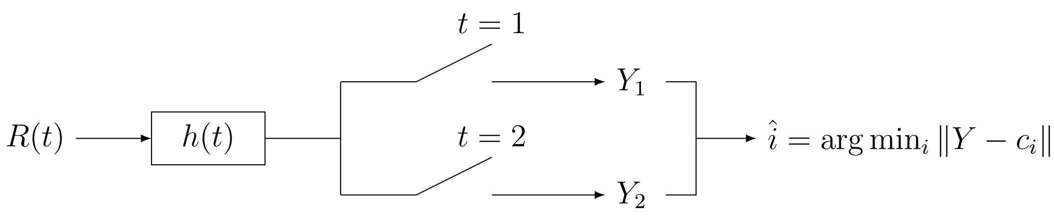


These waveforms are used to convey one of four messages over an AWGN channel with noise intensity σ^2. Assume that the four messages are equally likely.

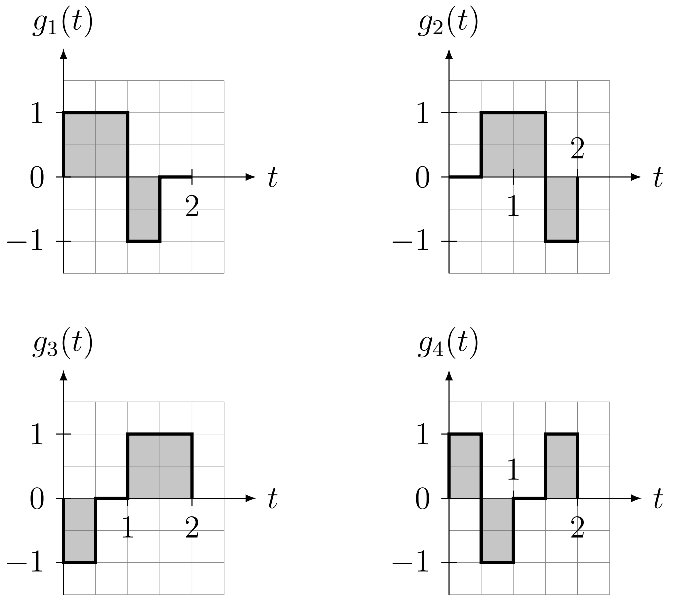
(a) (3 pts) Draw the diagram of an optimal receiver architecture. If your architecture requires computing inner products, use matched filters to do so.

Solution: A basis for the given set of waveforms is $\psi_1(t) = \mathbb{1}_{[0,1)}(t)$ and $\psi_2(t) = \mathbb{1}_{[1,2)}(t)$. We can construct the sufficient statistic $Y = (Y_1, Y_2)$ with $Y_j = \langle R, \psi_j \rangle = \langle w_i, \psi_j \rangle + \langle N, \psi_j \rangle$. Thus, we have $Y = c_i + Z$, where $c_0 = (1, 1)$, $c_1 = (-1, -1)$, $c_2 = (1, -1)$, $c_3 = (-1, 1)$ and $Z = (Z_1, Z_2)$ with independent Gaussian components of mean 0 and variance σ^2. We then decide $\hat{i} = \arg\min_i \|Y - c_i\|$, i.e., $\hat{i} = 0$ if $Y_1 > 0, Y_2 > 0$, $\hat{i} = 1$ if $Y_1 < 0, Y_2 < 0$, $\hat{i} = 2$ if $Y_1 > 0, Y_2 < 0$, and $\hat{i} = 3$ if $Y_1 < 0, Y_2 > 0$.

To compute the inner products with ψ_1 and ψ_2, we can pick a matched filter of impulse response $h(t) = \psi(1 - t) = \mathbb{1}_{[0,1)}(t)$. Sampling the output of the matched filter at times $t = 1, 2$ gives Y_1 and Y_2 respectively. This is shown in the figure below.



Suppose now we are given analog circuits which produce the inner products $U_\ell = \langle R, g_\ell \rangle$, $\ell = 1, 2, 3, 4$ of the received signal $R(t)$ with the waveforms g_1, g_2, g_3, g_4 given as follows:



Note that the g_ℓ's are cyclic shifts of each other in the interval $[0, 2]$.

(b) (3 pts) We are asked to design the best possible receiver *with no further analog circuits* (so: no other analog inner product computations, matched filters, etc.). Does this restriction penalize us (compared to what we are able to do in (a))? If yes, why? If no, why not?

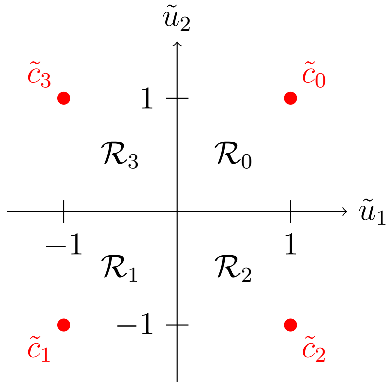
Hint: First plot $4g_1(t) + 3g_2(t) + g_3(t) + 2g_4(t)$ and $4g_3(t) + 3g_4(t) + g_1(t) + 2g_2(t)$.

Solution: By plotting, we observe that $4g_1(t) + 3g_2(t) + g_3(t) + 2g_4(t)$ is equal to $5\mathbb{1}_{[0,1)}(t) = 5\psi_1(t)$. Since the g's are obtained by cyclic right shifts, we also have that $4g_3(t) + 3g_4(t) + g_1(t) + 2g_2(t) = 5\mathbb{1}_{[1,2)}(t) = 5\psi_2(t)$ (this can also seen by plotting it independently). Hence, we can obtain (Y_1, Y_2) as in (a), since $Y_1 = \frac{1}{5}(4U_1 + 3U_2 + U_3 + 2U_4)$ and $Y_2 = \frac{1}{5}(4U_3 + 3U_4 + U_1 + 2U_2)$). Thus there is no penalty when restricted to use only these analog circuits.

Suppose we replace the four waveforms $g_1, \ldots, g_4$ above with only two waveforms $\tilde{g}_1(t) = \mathbb{1}_{[-1,1)}(t)$ and $\tilde{g}_2(t) = \mathbb{1}_{[1,3)}(t)$, with the corresponding inner products with $R(t)$ denoted by $\tilde{U}_1$ and $\tilde{U}_2$. Our receiver is now supposed to base its decision only on $(\tilde{U}_1, \tilde{U}_2)$.

(c) (3 pts) Draw the MAP decision regions for the transmitted message in the $(\tilde{u}_1, \tilde{u}_2)$ plane. Let $\tilde{p}_e(\sigma)$ denote the error probability of this receiver as a function of σ. Similarly, let $p_e(\sigma)$ denote the error probability of the receiver in (a). Find the ratio of $\tilde{p}_e(\sigma/\sqrt{2})$ and $p_e(\sigma)$.

Solution: We have $\tilde{U}_1 = \langle R, \tilde{g}_1 \rangle = \langle w_i, \tilde{g}_1 \rangle + \langle N, \tilde{g}_1 \rangle$ and $\tilde{U}_2 = \langle w_i, \tilde{g}_2 \rangle + \langle N, \tilde{g}_2 \rangle$. Thus, we have that $(\tilde{U}_1, \tilde{U}_2) = \tilde{c}_i + (\tilde{Z}_1, \tilde{Z}_2)$, where $\tilde{c}_0 = (1, 1)$, $\tilde{c}_1 = (-1, -1)$, $\tilde{c}_2 = (1, -1)$, $\tilde{c}_3 = (-1, 1)$, and $\tilde{Z}_1, \tilde{Z}_2$ are independent Gaussian random variables with mean zero and variance $2\sigma^2$. (The variance is twice because $\|\tilde{g}_i\|^2 = 2$.) The constellation and the decision regions are marked below.



This is identical to the scenario in (a), as $\tilde{c}_i = c_i$, except that the variance is now twice. Hence, the error probabilities satisfy $\tilde{p}_e(\sigma/\sqrt{2}) = p_e(\sigma)$, and the desired ratio is 1.

For the rest of the problem we continue with the optimal receiver found in (a).

(d) (3 pts) Conditioned on $i = 0$ being the sent message, find the probability that the receiver in (a) decides $\hat{i} = 1$; repeat for $\hat{i} = 2$, and $\hat{i} = 3$.

Solution: When $i = 0$, we decide $\hat{i} = 1$ if we have $Z_1 < -1$ and $Z_2 < -1$ (using the same notation as in (a)). Hence, the probability of deciding $\hat{i} = 1$ when $i = 0$ is equal to $\Pr(Z_1 < -1, Z_2 < -1) = Q\left(\frac{1}{\sigma}\right)^2$. Similarly, we decide $\hat{i} = 2$ if $Z_1 > -1$ and $Z_2 < -1$, which has probability $Q\left(\frac{1}{\sigma}\right) - Q\left(\frac{1}{\sigma}\right)^2$. We decide $\hat{i} = 3$ if $Z_1 < -1$ and $Z_2 > -1$, which also has probability $Q\left(\frac{1}{\sigma}\right) - Q\left(\frac{1}{\sigma}\right)^2$.

Suppose that the data to be sent is composed of two bits b_1 and b_2, with (b_1, b_2) taking the values `00`, `01`, `10` and `11` with equal probability. We will assign these four possible data values to the messages 0, 1, 2, 3. To transmit a particular bit pair, we send the waveform corresponding to the respective message. This is summarized as follows:

$$(b_1, b_2) \longrightarrow i \longrightarrow w_i(t) \longrightarrow R(t) \longrightarrow \hat{i} \longrightarrow (\hat{b}_1, \hat{b}_2)$$

(e) (3 pts) Consider assigning the data bit values `00`, `01`, `10` and `11` to messages 0, 1, 2, 3 as `00` $\to$ 0, `01` $\to$ 1, `10` $\to$ 2, `11` $\to$ 3. Conditioned on $(b_1, b_2) =$ `00`, what is the expected number of incorrectly received bits $(\hat{b}_1, \hat{b}_2)$? Repeat the question for "`01` is sent", "`10` is sent" and "`11` is sent".

Hint: For the "repeat"s: you don't have to do any calculations.

Solution: The mapping from bit pairs to messages is `00` $\to$ 0, `01` $\to$ 1, `10` $\to$ 2, `11` $\to$ 3. When `00` (message 0) is sent, we make 1 bit error if we receive either message 1 or 2, and two bit errors if we receive message 3. Hence, the expected number of bits that are incorrectly received is

$$\begin{aligned} &1 \cdot \Pr\{\hat{i} = 1 \mid i = 0\} + 1 \cdot \Pr\{\hat{i} = 2 \mid i = 0\} + 2 \cdot \Pr\{\hat{i} = 3 \mid i = 0\} \\ &= 3Q\left(\frac{1}{\sigma}\right) - 2Q\left(\frac{1}{\sigma}\right)^2. \end{aligned}$$

Since the figure is entirely symmetric — we make two bit errors if we end up in a neighbouring region and one error each if we end up in the other neighbouring region or the diametrically opposite region — we get the same error probability even when "`01` is sent", "`10` is sent" and "`11` is sent".

(f) (3 pts) Consider assigning the two-bit values to messages as `00` $\to 0$, `01` $\to 3$, `10` $\to$ 2, `11` $\to 1$. Redo (e).

Solution: With this mapping, we make two bit errors on sending `00` if we receive message 1, and one bit error if we receive either message 2 or 3. Hence, the expected number of incorrectly received bits is

$$
\begin{aligned}
&1 \cdot \Pr\{\hat{i} = 2 \mid i = 0\} + 1 \cdot \Pr\{\hat{i} = 3 \mid i = 0\} + 2 \cdot \Pr\{\hat{i} = 2 \mid i = 0\} \\
&= 2Q\left(\frac{1}{\sigma}\right).
\end{aligned}
$$

Similarly as in (e), we get the same error probability even when "`01` is sent", "`10` is sent" and "`11` is sent" — we make two bit errors when we end up in the diametrically opposite decision region, and one bit error otherwise.

Remarks: In part (b), we see that even though we are restricted to use a specific set of signals to compute inner products, we are still able to implement the optimal decision rule — this is because the signals that we are allowed to use (i.e., $\{g_i\}$) span the orthonormal basis of our waveforms $\{w_i\}$. In part (c), however, the signals that we are allowed to use (i.e., $\{\tilde{g}_i\}$) do not span the orthonormal basis of our waveforms, they instead pick up additional noise, and hence, results in a worse error performance.
In part (d), we see that different incorrect decisions occur with different probabilities, and we can exploit this when assigning bit pairs to messages — the choice of assignment in (e) is worse than that in (f) and has a higher error probability ($3Q - Q^2 > 2Q$, as $Q < 1$), as it places `00` and `11` close to each other, which causes twice as many bit errors as message errors. The assignment in (f) ensures that this does not happen, as messages that are close to each other are assigned bit pairs that differ in only position. This assignment is called a Gray code.