PDC Project - Theory

Yoonseo Ko (325515), Lilly-Flore Celma (346446), Maximilien Gridel (346611), Mathis Krause (328110)

Contents

Problem 1 **3**
(a) . . . 3
(b) . . . 3
(c) . . . 4

Problem 2 **5**
(a) . . . 5
(b) . . . 5
(c) . . . 5
(d) . . . 6

Problem 3 **7**
(a) . . . 7
(b) . . . 7
(c) . . . 8
(d) . . . 9
(e) . . . 9

Problem 4 **10**
(a) . . . 10
(b) . . . 10
(c) . . . 11
(d) . . . 11
(e) . . . 12
(f) . . . 13

Problem 1

(a)

We know that:

$$E = \bigcup_{i=1}^{k} E_i = E_1 \cup \cdots \cup E_k, \quad E^c = \overline{E_1 \cup \cdots \cup E_k} = E_1^c \cap \cdots \cap E_k^c = \bigcap_{i=1}^{k} E_i^c$$

$Pr(E^c) = Pr(\bigcap_{i=1}^k E_i^c) = Pr(E_1^c) \cdot \ Pr(E_2^c) \cdots \ Pr(E_k^c)$, since $E_1, \cdots, E_k$ are independent,

$= \prod_{i=1}^k Pr(E_i^c) = \prod_{i=1}^k (1 - p_i)$, since $E_1, \cdots, E_k$ are the events with $Pr(E_i) = p_i$.

By using the hint provided: $1 - x \leq exp(-x)$,

We have: $1 - p_i \leq exp(-p_i) \Leftrightarrow \prod_{i=1}^k (1 - p_i) \leq \prod_{i=1}^k exp(-p_i)$, by the exponential rule: $\prod_{i=1}^k exp(-p_i) = exp(-\sum_{i=1}^k p_i)$

Therefore,

$$\prod_{i=1}^{k}(1 - p_i) \leq exp(-\sum_{i=1}^{k} p_i) \Leftrightarrow Pr(E^c) \leq exp(-\sum_{i=1}^{k} p_i)$$

(b)

Let us sketch the functions $1 - exp(-s)$ and $\min\{1, s\}$:

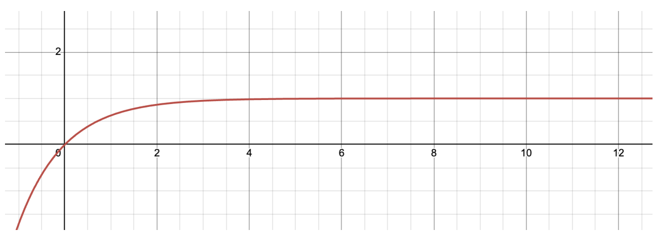

Figure 1: $1 - exp(-s)$ function, where $s \geq 0$

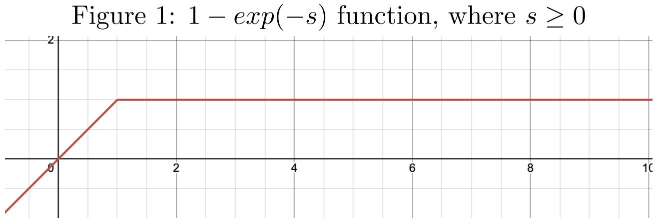

Figure 2: $\min\{1, s\}$ function, where $s \geq 0$

Let us consider the 2 cases: (i) for $s \in [0, 1]$ and (ii) for $s > 1$

- For $s \in [0, 1]$: $\min\{1, s\} = s$,

  By using Taylor series, $exp(s) = 1 + s + \frac{s^2}{2!} + \frac{s^3}{3!} + \cdots \Leftrightarrow 1 - exp(-s) = s - \frac{s^2}{2} + \frac{s^3}{6} - \cdots \approx s - \frac{s^2}{2} + \frac{s^3}{6}$,

  and $(1 - \frac{1}{e}) \min\{1, s\} = (1 - \frac{1}{e})s = s - \frac{s}{e}$

  Therefore,

$$s - \frac{s^2}{2} + \frac{s^3}{6} > s - \frac{s}{e} \Leftrightarrow 1 - exp(-s) \geq (1 - \frac{1}{e}) \min\{1, s\}$$

- For $s > 1$: $\min\{1, s\} = 1$

  We have:
$$-s < -1 \iff exp(-s) < exp(-1) \iff -exp(-s) > -exp(-1)$$
$$\iff 1 - exp(-s) > 1 - exp(-1) \iff (1 - exp(-s)) \min\{1, s\} > (1 - exp(-1)) \min\{1, s\}$$
  , since $\min\{1, s\} = 1$, for $s > 1$.

  Therefore,
$$1 - exp(-s) \geq (1 - \frac{1}{e}) \min\{1, s\}$$

(c)

We have that:
$$s = \sum_{i=1} Pr(E_i) = \sum_i p_i$$

By applying (a), $Pr(E^c) \leq exp(-s)$, and $Pr(E) = 1 - Pr(E^c)$

Thus, we obtain a lower bound:
$$1 - Pr(E^c) = Pr(E) = Pr(\bigcup_i E_i) \geq 1 - exp(-s)$$

By union bound, we know that:
$$Pr(E) \leq \sum_i p_i = s$$

By applying (b), we have:
$$(1 - \frac{1}{e}) \min\{1, s\} \leq 1 - exp(-s) \leq Pr(E) \leq \sum_i p_i = s \leq \min\{1, s\} \leq 1$$

Thus, we can conclude that:
$$(1 - \frac{1}{e}) \min\{1, s\} \leq Pr(\bigcup_i E_i) \leq \min\{1, s\}$$

Problem 2

(a)

We have 2^k equally likely messages. For i, $P_i(\text{message sent}) = \frac{1}{2^k}$.

We can now derive from the $\hat{\text{MAP}}$ rule to the ML rule.

In AWGN:

$$Y_i = c_i + Z = \begin{cases} v_0 + Z & \text{,if } b_i = 0 \\ v_1 + Z & \text{,if } b_i = 1 \end{cases}$$

We aim to minimize the probability of decoding $\hat{\text{b}}$ incorrectly given Y. Each component $\hat{b_i}$ is decoded by construction as the minimum distance between Y_i and v_0, v_1. $\iff \arg\min_{b\in\{0,1\}} \|Y_i - v_b\|$
Or mathematically,

$$f_{Y|b}(y|0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{||y - v_0||^2}{2\sigma^2},\ f_{Y|b}(y|1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{||y - v_1||^2}{2\sigma^2}$$

$$\Lambda = \frac{f_Y(0)}{f_Y(1)} \iff 1 \geq \frac{\|y - v_0\|}{\|y - v_1\|}$$

$$\iff \|y - v_o\|^2 \geq \|y - v_1\|^2$$

$$\iff \arg\min_{b\in\{0,1\}} \|y - v_b\|$$

(b)

We have here a receiver that makes a binary hypothesis decoding. Due to the symmetry of the problem, $P(\hat{b_i} \neq b_i)$ is the same for all i.
Also, $P(\hat{b_i} \neq b_i = 0) = P(\hat{b_i} \neq b_i = 1) = P(\hat{b_i} \neq b_i)$ with probability, $Q(\frac{d}{2\sigma})$, where $d = ||v_0 - v_1||$.
So,

$$P(error) = 1 - P(correct)$$

$$1 - \prod_{i=1}^{k} P(\hat{b_i} = b_i), where P(\hat{b_i} = b_i) = 1 - P(\hat{b_i} \neq b_i)$$

$$1 - (1 - Q(\frac{d}{2\sigma}))^k$$

, by independence of each bit.

To compare, we can use the union bound: i.e. the probability that at least one of the events happens is no greater than the sum of the probabilities of the individual events. $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$.

$$Pr(error) \leq kPr(\text{bit error}) = kQ(\frac{d}{2\sigma})$$

The function $0 \leq Q(x) \leq 1$ is decreasing and Pr(error) is at most 1.

Therefore $\min\{1, kQ(\frac{d}{2\sigma})\}$ is an upper bound to the P(error)

(c)

Since it is still in AWGN channel, the codewords have the same distances. Thus, the bit error and the total error maintain the same.
The average energy can be computed as follows:

$$\varepsilon\ = \frac{1}{2} \cdot\ (\frac{d}{2})^2 + \frac{1}{2} \cdot\ (-\frac{d}{2})^2 = k * \frac{d^2}{4}$$

so transmission result in energy ,

$$\varepsilon_b\ = \frac{d^2}{4}$$

In the original system we had as the transmission is bit by bit

$$\varepsilon \ = \varepsilon_b = \frac{1}{2}(||v_0||^2 + ||v_1||^2)$$

This implies the original system has the same energy of the new if v_0+v_1 =0 otherwise it might use more energy. So if we are in a bit by bit you could use the simpler system.

(d)

We want here $\Pr(\text{error}) \leq \alpha$ where $\alpha = 10^{-2}$.

First case if $d/(2\sigma) < a_2$:
The Q-function is monotonically decreasing hence we have that $Q\left(\frac{d}{2\sigma}\right) > Q(a_2)$ and $Q(a_2) = \frac{\alpha}{k} * (\frac{e}{e-1})$.
Inserting all in the first equation of error probability

$$\Pr(\text{error}) = 1 - (1 - Q\left(\frac{d}{2\sigma}\right))^k > 1 - (1 - Q(a_2))^k = 1 - \left(1 - \frac{\alpha}{k}\frac{e}{e-1}\right)^k$$

The function is strictly decreasing , and we consider k >0 .
So we try to lower bound our error probability by something greater than α to show error probability requirement cannot be met.
First of all a quick graphic confirms our intuition that it will indeed never be met.

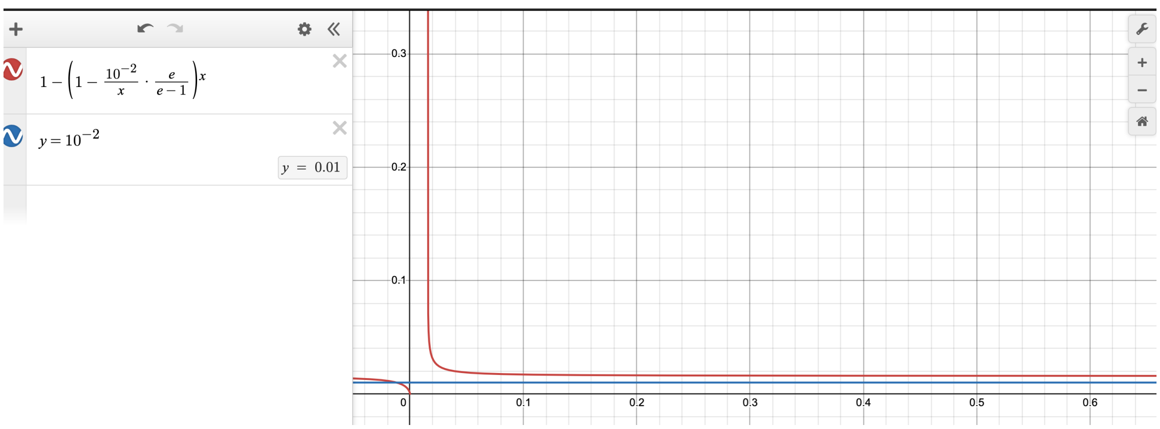

Figure 3: Probability error function vs α

Mathematically we can prove it using the bounds found in Exercise 1-c) or the limits as k the number of bits growths. Recall

$$\left(1 + \frac{x}{n}\right)^n \rightarrow e^x \quad \text{as} \quad n \rightarrow \infty$$

So here with $A = \frac{e}{100(e-1)}$ we end up with

$$\Pr(\text{error}) > \lim_{k\to\infty} 1 - \left(1 - \frac{A}{k}\right)^k = 1 - e^{-A} \approx 0.0157$$

Hence we will always have $\Pr(\text{error}) > \alpha$, so the error probability requirement cannot be met.

The second case is easier to compute. So if $d/(2\sigma) \geq a_1$:
We have that $Q\left(\frac{d}{2\sigma}\right) \leq Q(a_1) = \frac{\alpha}{k}$.
Using the upper bound error expression 2-(b) we found:

$$\Pr(\text{error}) \leq k \cdot Q\left(\frac{d}{2\sigma}\right) \leq k \cdot Q(a_1) = k \cdot \frac{\alpha}{k} = \alpha$$

So if $d/(2\sigma) \geq a_1$, the message error probability is met.
As the moral hinted us, we can say that the minimal possible value for $\frac{d}{2\sigma}$ to achieve a meet a specific message error probability α lies between a_1 and a_2 , which when squared match the SNR per bit formula.

Problem 3

(a)

There are 2n codewords. The decision will be wrong only if, either $|Y_1| \leq t$, or, for some j = 2,...,n, $|Y_j| > t$.

The union bound is:

$$Pr(error) \leq Pr(|Y_1| \leq t) + \sum_{i=2}^{n} Pr(|Y_i| > t)$$

$$Pr(error) \leq Pr(Z_1 > (1-\alpha)\sqrt{E}) + \sum_{i=2}^{n} Pr(|Z_i| > \alpha\sqrt{E})$$

$$Pr(error) \leq Q((1-\alpha)\sqrt{\frac{E}{\sigma^2}}) + \sum_{i=2}^{n} 2Q(\alpha\sqrt{\frac{E}{\sigma^2}})$$

$$Pr(error) \leq Q((1-\alpha)\sqrt{\frac{E}{\sigma^2}}) + 2(n-1)Q(\alpha\sqrt{\frac{E}{\sigma^2}})$$

Recall:

$$k = \log_2(2n)$$

Moreover,

$$2(n-1) < 2n = 2^{\log_2(2n)}$$

Therefore, we can write:

$$Pr(error) < Q\left((1-\alpha)\sqrt{\frac{E}{\sigma^2}}\right) + 2^k Q\left(\alpha\sqrt{\frac{E}{\sigma^2}}\right)$$

(b)

Recall that

$$Q(x) \leq \frac{1}{2}\exp(\frac{-x^2}{2})$$

Also since

$$E = kA\sigma^2$$

Then, the probability of error is further upper bounded by

$$\frac{1}{2}\exp\left(\frac{-(1-\alpha)^2\frac{kA\sigma^2}{\sigma^2}}{2}\right) + \frac{1}{2}\exp\left(\frac{-\alpha^2\frac{kA\sigma^2}{\sigma^2}}{2}\right)$$

$$\frac{1}{2}\exp\left(\frac{-(1-\alpha)^2 kA}{2}\right) + 2^k\frac{1}{2}\exp\left(\frac{-\alpha^2 kA}{2}\right)$$

$$\frac{1}{2}\exp\left(\frac{-(1-\alpha)^2 kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-\alpha^2 kA}{2} + k\ln(2)\right)$$

$$\frac{1}{2}\exp\left(\frac{-(1-\alpha)^2 kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\alpha^2 A - 2\ln(2))}{2}\right)$$

If $A > 2\ln(2)$, then we can $A = 2\ln(2) + \epsilon$ for $\epsilon > 0$.

Conditions on α :

- first term: $\frac{-(1-\alpha)^2 \frac{kA\sigma^2}{\sigma^2}}{2}$, as k grows, this term will decrease for $0 < \alpha < 1$.
- second term: $\frac{-k(\alpha^2 A - 2\ln(2))}{2}$, as k grows, the term will decrease for $\alpha^2 A > 2\ln(2) \Leftrightarrow \alpha^2 > \frac{2\ln(2)}{2\ln(2)+\epsilon} \Leftrightarrow \alpha^2 > \frac{1}{1+\frac{\epsilon}{2\ln(2)}}$.

So, there is an $0 < \alpha < 1$ that satisfies the conditions on the two terms, such that

$$\lim_{k\to\infty} \frac{1}{2}\exp\left(\frac{-(1-\alpha)^2 kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\alpha^2 A - 2\ln(2))}{2}\right) = 0$$

Thus

$$\lim_{k\to\infty} Pr(error) = 0$$

(c)

Take

$$\alpha = \frac{1}{2}\left(1 + \frac{2\ln(2)}{A}\right)$$

Replace it in the expression shown in (b):

$$\frac{1}{2}\exp\left(\frac{-(1-\alpha)^2 kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\alpha^2 A - 2\ln(2)}{2}\right)$$

We have,

$$\frac{1}{2}\exp\left(\frac{-(1-\frac{1}{2}(1+\frac{2\ln(2)}{A}))^2 kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\frac{1}{4}(1+\frac{2\ln(2)}{A})^2 A - 2\ln(2))}{2}\right)$$

$$\frac{1}{2}\exp\left(\frac{-(\frac{1}{4} - \frac{\ln(2)}{A} + \frac{\ln^2(2)}{A^2})kA}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\frac{1}{4}(1+\frac{4\ln(2)}{A} + \frac{4\ln^2(2)}{A^2})A - 2\ln(2))}{2}\right)$$

$$\frac{1}{2}\exp\left(\frac{-k(\frac{1}{4}A - \ln(2) + \frac{\ln^2(2)}{A})}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\frac{1}{4} + \ln(2) + \frac{\ln^2(2)}{A}) - 2\ln(2))}{2}\right)$$

Reorganize the exponentials,

$$\frac{1}{2}\exp\left(\frac{-k(\frac{1}{4}A - \ln(2) + \frac{\ln^2(2)}{A})}{2}\right) + \frac{1}{2}\exp\left(\frac{-k(\frac{1}{4}A - \ln(2) + \frac{\ln^2(2)}{A})}{2}\right)$$

Notice that with have the sum of the same exponential, therefore we can rewrite the upper bound as follows:

$$\exp\left(\frac{-k(\frac{1}{4}A - \ln(2) + \frac{\ln^2(2)}{A})}{2}\right)$$

which is equal to

$$\exp\left(-\frac{\frac{1}{4}(1-\frac{2\ln(2)}{A})^2 Ak}{2}\right)$$

We finally get,

$$Pr(error) < \exp\left(-\frac{1}{8}(1-\frac{2\ln(2)}{A})^2 Ak\right)$$

(d)

From the expression of the upper bound obtained in (c),
Let's find the values of k such that the upper bound of the error probability is less than 10^{-3} depending on the values of A.

$$Pr(error) < \exp\left(-\frac{1}{8}(1-\frac{2\ln(2)}{A})^2 Ak\right) < 10^{-3}$$

Let's isolate k,

$$-\frac{1}{8}(1-\frac{2\ln(2)}{A})^2 Ak < -3\ln(10)$$
$$\Leftrightarrow (1-\frac{2\ln(2)}{A})^2 k > \frac{24\ln(10)}{A}$$
$$\Leftrightarrow k > \frac{24\ln(10)}{A} \times \frac{1}{(1-\frac{2\ln(2)}{A})^2}$$

Recall k is an integer.
From the formula obtained above we can compute the values of k for each value of A.

Values of A	k(A) - Values of k
4	k > 32, take $k = 33$
6	k > 15, take $k = 16$
8	k > 10, take $k = 11$
10	k > 7, take $k = 8$
12	k > 5, take $k = 6$

(e)

Consider a bit-by-bit communication system with $\frac{E_b}{\sigma^2} = A$ that sends k(A)-bit message, with the k(A) chosen above.
From Problem 2, we have the bit error probability $P_{\text{bit error}} = Q(\frac{d}{2\sigma})$. Where d is known from the previous problem, because $(\frac{d}{2\sigma})^2 = \frac{E_b}{\sigma^2} = A \Leftrightarrow \frac{d}{2\sigma} = \sqrt{A} \Leftrightarrow d = 2\sigma\sqrt{A}$. Thus, we can simplify the expression of the bit error probability to $P_{\text{bit error}} = Q(\sqrt{A})$. We can write the message error probability, $P_{\text{message error}} = 1-(1-P_{\text{bit error}})^{k(A)} = 1-(1-Q(\sqrt{A}))^{k(A)}$.

Values of A	Error Probability
4	$1-(1-Q(2))^{k(2)}$
6	$1-(1-Q(\sqrt{6}))^{k(6)}$
8	$1-(1-Q(2\sqrt{2}))^{k(8)}$
10	$1-(1-Q(\sqrt{10}))^{k(10)}$
12	$1-(1-Q(2\sqrt{3}))^{k(12)}$

Problem 4

(a)

We want to know if this rule minimize $Pr(\hat{H} \neq H)$:

$$\hat{H} = \begin{cases} (i_1, 1) & \text{if } d_1 < d_2, \\ (i_2, 2) & \text{else} \end{cases}$$

with $d_1 = \frac{\|y_1 - c_{i_1}\|^2}{\sigma^2} + \frac{\|y_2\|^2}{\tau^2}$ and $d_2 = \frac{\|Y_2 - c_{i'}\|^2}{\sigma^2} + \frac{\|Y_1 - c_i + c_i\|^2}{\tau^2}$
Using the MAP rule, we have

$$\Pr((y_1, y_2) \mid H = (i,1)) \underset{\hat{H}=1}{\overset{\hat{H}=2}{\lessgtr}} \Pr((y_1, y_2) \mid H = (i,2))$$

- $\Pr((y_1, y_2) \mid H = (i,1)) = \Pr(y_1 \mid H = (i,1)) \Pr(y_2 \mid H = (i,1))$ since the events Y_1 and Y_2 are independents.
Since we are in the first case, we have $Y_1 = c_i + Z$ and $Y_2 = \tilde{Z}$, thus $Z = Y_1 - c_i$ and $\tilde{Z} = Y_2$, and since Z and $\tilde{Z}$ have Gaussian distribution, we have :
$\Pr(y_1 \mid H = (i,1)) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y_1 - c_i\|^2\right)$ and $\Pr(y_2 \mid H = (i,1)) = \frac{1}{(2\pi\tau^2)^{n/2}} \exp\left(-\frac{1}{2\tau^2}\|y_2\|^2\right)$, so :
$\Pr((y_1, y_2) \mid H = (i,1)) = \frac{1}{(2\pi\sigma^2\tau^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y_1 - c_i\|^2 - \frac{1}{2\tau^2}\|y_2\|^2\right)$

- $\Pr((y_1, y_2) \mid H = (i,2)) = \Pr(y_1 \mid H = (i,2)) \Pr(y_2 \mid H = (i,2))$:
With the same reasoning, we have :
$\Pr(y_1 \mid H = (i,2)) = \frac{1}{(2\pi\tau^2)^{n/2}} \exp\left(-\frac{1}{2\tau^2}\|y_1\|^2\right)$ and $\Pr(y_2 \mid H = (i,2)) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y_2 - c_i\|^2\right)$, so :
$\Pr((y_1, y_2) \mid H = (i,2)) = \frac{1}{(2\pi\sigma^2\tau^2)^{n/2}} \exp\left(-\frac{1}{2\tau^2}\|y_1\|^2 - \frac{1}{2\sigma^2}\|y_2 - c_i\|^2\right)$
Therefore the MAP-rule gives :

$$\frac{1}{(2\pi\sigma^2\tau^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y_1 - c_i\|^2 - \frac{1}{2\tau^2}\|y_2\|^2\right) \underset{\hat{H}=1}{\overset{\hat{H}=2}{\lessgtr}} \frac{1}{(2\pi\sigma^2\tau^2)^{n/2}} \exp\left(-\frac{1}{2\tau^2}\|y_1\|^2 - \frac{1}{2\sigma^2}\|y_2 - c_i\|^2\right)$$

We simplify this inequalities with the log and we get :

$$\frac{\|y_1 - c_i\|^2}{\sigma^2} + \frac{\|y_2\|^2}{\tau^2} \underset{\hat{H}=2}{\overset{\hat{H}=1}{\lessgtr}} \frac{\|y_2 - c_i\|^2}{\sigma^2} + \frac{\|y_1\|^2}{\tau^2}$$

We recognize in the first part d_1 and in the other d_2, therefore we have that :

$$\hat{H} = \begin{cases} (i_1, 1) & \text{if } d_1 < d_2, \\ (i_2, 2) & \text{else} \end{cases}$$

This decision rule is exactly the given rule in the problem statement. Hence the given rule indeed minimizes $Pr(\hat{H} = H)$.

(b)

$\hat{\imath}$ is the first component of $\hat{H}$.
In this question we want to determine if the following rule minimize $Pr(\hat{\imath} \neq i)$:

$$\hat{\imath} = \begin{cases} i_1 & \text{if } d_1 < d_2, \\ i_2 & \text{else} \end{cases}$$

To see this, let's consider the two possible cases for b:

1: If b = 1, then $Y_1 = c_i + Z$ and $Y_2 = \tilde{Z}$
The optimal decision rule for estimating i based on (Y_1, Y_2) is equivalent to the optimal rule for estimating i based on Y_1

alone, since Y_2 does not contain any information about i in this case. The given rule chooses $i_1 = argmin\|y_1 - c_i\|$, which is the maximum likelihood (ML) estimate of i based on Y_1. Since the ML estimate minimizes the probability of error in general, the given rule also minimizes the probability of error for estimating i in this case.

2: If b = 2, then $Y_1 = \tilde{Z}$ and $Y_2 = c_i + Z$
The optimal decision rule for estimating i based on (Y_1, Y_2) is equivalent to the optimal rule for estimating i based on Y_2 alone, since Y_1 does not contain any information about i in this case. The given rule chooses $i_2 = argmin\|y_2 - c_i\|$, which is the maximum likelihood (ML) estimate of i based on Y_2. Therefore, the given rule also minimizes the probability of error for estimating i in this case.

Since the given rule minimizes the probability of error for estimating i in both possible cases for b, it also minimizes the overall probability of error for estimating i.

(c)

$\hat{i}_o(y_1, y_2, b)$ is the MAP estimator of a receiver that somehow has access to the side information, i.e, it is the decision made from the observation (y_1, y_2, b).

C0 : $\Pr(\hat{i}_o \neq i) \overset{(c_0)}{\leq} \Pr(\hat{i} \neq i)$

$\hat{i}_o(y_1, y_2, b)$ is the MAP estimator of a receiver that somehow has access to the side information. Hence $\hat{i}_o$ potentially enabling a better estimate of i compare to $\hat{i}$ which does not have this information and should estimate b and i. The knowledge of B reduces ambiguity about which part of Y contains noise and which part contains the signal plus noise, leading to a potential reduction in errors in estimating i.

C1 : $\Pr(\hat{i}_o \neq i) \overset{(c_1)}{\leq} \Pr(\hat{H} \neq H)$

The event $H \neq \hat{H}$ include errors in estimating both the codeword index i and the scenario b. Therefore, an error in $\hat{i}$ is a subset of the event leading to $H \neq \hat{H}$. Hence, the probability of a specific error in i is bounded above by the probability of any error in $\hat{H}$.

C2 : $\Pr(\hat{H} \neq H) \overset{(c_2)}{=} \Pr(\hat{b} \neq b) + \Pr(\hat{b} = b \textbf{ and } \hat{i} \neq i)$

This equality holds because it correctly partitions the probability of $H \neq \hat{H}$ into two mutually exclusive cases : either $\hat{b}$ is correct (irrespective of $\hat{i}$), or $\hat{b}$ is correct but $\hat{i}$ is not. The total error in $\hat{H}$ is thus the sum of these two error probabilities.

C3 : $\Pr(\hat{b} \neq b) + \Pr(\hat{b} = b \textbf{ and } \hat{i} \neq i) \overset{(c_3)}{=} \Pr(\hat{b} \neq b) + \Pr(\hat{b} = b \textbf{ and } \hat{i_o} \neq i)$

This equality holds since the total error in estimating $\hat{H}$ using $\hat{i_o}$ is statistically equivalent to the total error in estimating $\hat{H}$ using $\hat{i}$, since $\Pr(\hat{b} = b \text{ and } \hat{i} \neq i) = \Pr(\hat{b} = b \text{ and } \hat{i_o} \neq i)$, because $\hat{i}_o(y_1, y_2, b)$ is the MAP estimator, thus we don't lose any information.

C4 : $\Pr(\hat{b} \neq b) + \Pr(\hat{b} = b \textbf{ and } \hat{i_o} \neq i) \overset{(c_4)}{\leq} \Pr(\hat{b} \neq b) + \Pr(\hat{i_o} \neq i)$

This inequality holds because the errors where the scenario b is correctly identified but the index i is still misidentified are part of the wider set of all index misidentifications. Therefore summing up the scenario misidentification with all index misidentifications include or equals the sum of the scenario error and the specific case of correct scenario but incorrect index.

(d)

We suppose that H = (i, 1), thus we are in the scenario where $Y_1 = c_i + Z$, and $Y_2 = \tilde{Z}$.
$\hat{b} \neq 1$ means $\hat{b} = 2$, there exists $i' \in \{1, ..., m\}$, such that $d_2 < d_1$.

From the statement and the hint, we can rewrite as follows :

$d_1 = \frac{\|y_1 - c_{i_1}\|^2}{\sigma^2} + \frac{\|y_2\|^2}{\tau^2} = \frac{\|Z\|^2}{\sigma^2} + \frac{\|\tilde{Z}\|^2}{\tau^2}$

$d_2 = \frac{\|Y_2 - c_{i'}\|^2}{\sigma^2} + \frac{\|Y_1 - c_i + c_i\|^2}{\tau^2} = \frac{\|\tilde{Z} - c_{i'}\|^2}{\sigma^2} + \frac{\|c_i + Z\|^2}{\tau^2}$

We obtain the following expression,

$$\frac{\|\tilde{Z}-c_{i'}\|^2}{\sigma^2}+\frac{\|c_i+Z\|^2}{\tau^2}<\frac{\|Z\|^2}{\sigma^2}+\frac{\|\tilde{Z}\|^2}{\tau^2}$$

(e)

In this question we want to use the union bound to upper bound $\Pr\left(\hat{b}\neq 1 \mid H=(i,1)\right)$.
The union bound is expressed as follows : $\Pr(\bigcup_{i=1}^{n} A_i)\leq \sum_{i=1}^{n}\Pr(A_i)$
In our context, each event A_i is defined as the event where the incorrect decision $\hat{b}=2$ is made for a particular i'.
From part (d), we know that $\hat{b}\neq 1$ if there exists $i'\in\{1,\ldots,m\}$ such that:

$$\frac{\|Z\|^2}{\sigma^2}+\frac{\|\tilde{Z}\|^2}{\tau^2}>\frac{\|c_i+Z\|^2}{\tau^2}+\frac{\|\tilde{Z}-c_{i'}\|^2}{\sigma^2}.$$

Simplifying with $\sigma=\tau$, the inequality becomes:

$$\|Z\|^2+\|\tilde{Z}\|^2>\|c_i+Z\|^2+\|\tilde{Z}-c_{i'}\|^2.$$

$$2\|Z\|^2>\|c_i+Z\|^2+\|Z-c_{i'}\|^2.$$

In addition we have :

$$\Pr\left(\hat{b}\neq 1 \mid H=(i,1)\right)=\Pr\left(d_1\geq d_2 \mid H=(i,1)\right)$$
$$=Pr\left(\bigcup_{i'=1}^{m}(\|y_1-c_i\|^2+\|y_2\|^2\geq\|y_1\|^2+\|y_2-c_{i'}\|^2)\right)$$
$$\leq\sum_{i'=1}^{m}\Pr\left(\|y_1-c_i\|^2+\|y_2\|^2\geq\|y_1\|^2+\|y_2-c_{i'}\|^2\right)$$
$$=\sum_{i'=1}^{m}\Pr\left(\|Z\|^2+\|\tilde{Z}\|^2\geq\|c_i+Z\|^2+\|\tilde{Z}-c_{i'}\|^2\right)$$
$$=\sum_{i'=1}^{m}\Pr\left(\|Z\|^2+\|\tilde{Z}\|^2\geq\|c_i\|^2+\|Z\|^2+2\langle Z,c_i\rangle+\|\tilde{Z}\|^2+\|c_{i'}\|^2-2\langle\tilde{Z},c_{i'}\rangle\right)$$

Each term inside the summation represents the probability that:

$$0>\|c_i\|^2+\|c_{i'}\|^2+\langle Z,c_i\rangle-2\langle\tilde{Z},c_{i'}\rangle$$
$$2\langle\tilde{Z},c_{i'}\rangle-2\langle Z,c_i\rangle>\|c_i\|^2+\|c_{i'}\|^2$$

Since Z and $\tilde{Z}$ are Gaussian random variables with mean 0 and variance σ^2, so let's compute the variance of $(2\langle\tilde{Z},c_{i'}\rangle-2\langle Z,c_i\rangle)$,

$\mathrm{Var}(2\langle\tilde{Z},c_{i'}\rangle-2\langle Z,c_i\rangle)=\mathrm{Var}(2\langle\tilde{Z},c_{i'}\rangle)+\mathrm{Var}(2\langle Z,c_i\rangle)-\mathrm{Cov}(2\langle\tilde{Z},c_{i'}\rangle,2\langle Z,c_i\rangle)$
Since Z and $\tilde{Z}$ are independent,
$\mathrm{Var}(2\langle\tilde{Z},c_{i'}\rangle-2\langle Z,c_i\rangle)=\mathrm{Var}(2\langle\tilde{Z},c_{i'}\rangle)+\mathrm{Var}(2\langle Z,c_i\rangle)$
$=4\mathrm{Var}(\langle\tilde{Z},c_{i'}\rangle)+4\mathrm{Var}(\langle Z,c_i\rangle)$
$=4\|c_{i'}\|^2\mathrm{Var}(\tilde{Z})+4\|c_i\|^2\mathrm{Var}(Z)$
$=4\sigma^2(\|c_{i'}\|^2+\|c_i\|^2)$

Now the probability we seek is :

$$\Pr\left(2\langle\tilde{Z},c_{i'}\rangle-2\langle Z,c_i\rangle>\|c_i\|^2+\|c_{i'}\|^2\right)$$

Thus,

$$\Pr\left(2\langle\tilde{Z}, c_{i'}\rangle - 2\langle Z, c_i\rangle > \|c_i\|^2 + \|c_{i'}\|^2\right) = Q\left(\frac{\|c_i\|^2 + \|c_{i'}\|^2}{\sqrt{4\sigma^2(\|c_i\|^2 + \|c_{i'}\|^2)}}\right)$$

Simplifying the argument of the Q-function :

$$\Pr\left(2\langle\tilde{Z}, c_{i'}\rangle - 2\langle Z, c_i\rangle > \|c_i\|^2 + \|c_{i'}\|^2\right) = Q\left(\frac{\|c_i\|^2 + \|c_{i'}\|^2}{\sqrt{4\sigma^2}\sqrt{\|c_i\|^2 + \|c_{i'}\|^2}}\right) = Q\left(\sqrt{\frac{\|c_i\|^2 + \|c_{i'}\|^2}{4\sigma^2}}\right).$$

Therefore we have :

$$\Pr(\hat{b} \neq 1 \mid H = (i, 1)) \leq \sum_{i'=1}^{m} Q\left(\sqrt{\frac{\|c_i\|^2 + \|c_{i'}\|^2}{4\sigma^2}}\right)$$

(f)

From part (e), we derived that : $\Pr(\hat{b} \neq 1 \mid H = (i, 1)) \leq \sum_{i'=1}^{m} Q\left(\sqrt{\frac{\|c_i\|^2+\|c_{i'}\|^2}{4\sigma^2}}\right)$.
Given $\|c_i\|^2 = \epsilon$ for each i, the expression simplifies to :

$$Q\left(\sqrt{\frac{\epsilon + \epsilon}{4\sigma^2}}\right) = Q\left(\sqrt{\frac{\epsilon}{2\sigma^2}}\right)$$

The energy per bit $\epsilon_b = \frac{\epsilon}{\log_2 m}$ is

$$\sqrt{\frac{\epsilon}{2\sigma^2}} = \sqrt{\frac{\epsilon_b \times \log_2 m}{2\sigma^2}}$$

Since, $P(b = 1) = P(b = 2)$, then $P(\hat{b} \neq b) = P(\hat{b} \neq 1 \mid H = (i, 1))$. Thus, an upper bound for $P(\hat{b} \neq b)$ is :

$$P(\hat{b} \neq b) \leq \sum_{i'=1}^{m} Q\left(\sqrt{\frac{\epsilon_b \times \log_2 m}{2\sigma^2}}\right) = m \times Q\left(\sqrt{\frac{\epsilon_b \times \log_2 m}{2\sigma^2}}\right)$$

We will now analyse what we have obtained.

Since $Q(x) \leq \frac{1}{2}e^{-x^2/2}$,

$$m \times Q\left(\sqrt{\frac{\epsilon_b \times \log_2 m}{2\sigma^2}}\right) \leq m \times \frac{1}{2}\exp\left(-\frac{\epsilon_b \times \log_2 m}{4\sigma^2}\right) = \frac{1}{2}\exp\left(\ln(m) - \frac{\epsilon_b \times \ln(m)}{4\sigma^2 \ln(2)}\right)$$

$$= \frac{1}{2}\exp\left(\ln(m) \times (1 - \frac{\epsilon_b}{4\sigma^2 \ln(2)})\right)$$

Assuming $\frac{\epsilon_b}{\sigma^2} > 4\ln 2$,

$$\lim_{m\to\infty} \frac{1}{2}\exp\left(\ln(m) \times (1 - \frac{\epsilon_b}{4\sigma^2 \ln(2)})\right) = 0$$

From (c),

$$\Pr(\hat{i}_o \neq i) - \Pr(\hat{i}_o \neq \hat{i}) \leq P(\hat{b} \neq b) \leq m \times Q\left(\sqrt{\frac{\epsilon_b \times \log_2 m}{2\sigma^2}}\right) \leq \frac{1}{2}\exp\left(\ln(m) \times (1 - \frac{\epsilon_b}{4\sigma^2 \ln(2)})\right)$$

Therefore, $Pr(\hat{i}_o \neq i) - \Pr(\hat{i}_o \neq \hat{i})$ approaches 0 as m increases, assuming $\frac{\epsilon_b}{\sigma^2} > 4\ln 2$.