Handout 21 Principles of Digital Communications
Project (Theory) April 30, 2024

PROBLEM 1. Let $E_1, \ldots, E_k$ be events with $\Pr(E_i) = p_i$. Let $E = \bigcup_i E_i$ be the union of the events. We know, by the union bound, that $\Pr(E) \leq \sum_i p_i$. By noting that the probability of any event is at most 1, we can trivially improve the bound to $\Pr(E) \leq \min\{1, \sum_i p_i\}$. For the rest of this problem, assume that the events $E_1, \ldots, E_k$ are independent.

(a) With A^c denoting the complement of an event A, show that $\Pr(E^c) \leq \exp(-\sum_i p_i)$.
Hint: $1 - x \leq \exp(-x)$.

(b) For $s \geq 0$, sketch the functions $1 - \exp(-s)$, and $\min\{1, s\}$. Show that $1 - \exp(-s) \geq (1 - 1/e)\min\{1, s\}$ for $s \geq 0$.
Hint: Consider the two cases (i) $s \in [0, 1]$, and (ii) $s > 1$.

(c) Combine your results in (a) and (b) to show that, when $E_1, \ldots, E_k$ are independent,

$$\left(1 - \frac{1}{e}\right) \min\{1, s\} \leq \Pr\left(\bigcup_i E_i\right) \leq \min\{1, s\},$$

with $s = \sum_i \Pr(E_i)$.

Moral of the story: For independent events, the trivially improved union bound, $\min\{1, \sum_i \Pr(E_i)\}$ is not only an upper bound to the probability of their union, but also a constant factor approximation to it.

PROBLEM 2. Suppose we design a communication system to send a k-bit message in the following way:

Step 1: We represent a message by a binary sequence $(b_1, \ldots, b_k)$, each b_i in $\{0, 1\}$.

Step 2: Pick two vectors v_0 and v_1 in $\mathbb{R}^r$.

Step 3: The codeword for the message $(b_1, \ldots, b_k)$ is then given by the vector $c = (v_{b_1}, \ldots, v_{b_k})$ (in $\mathbb{R}^n$ with $n = kr$). For example, let $v_0 = (1, 2, 3)$ and $v_1 = (-1, -3, -2)$ in $\mathbb{R}^3$, then the codeword for the 3-bit message $(0, 0, 1)$ is $(\underbrace{1, 2, 3}_{0}, \underbrace{1, 2, 3}_{0}, \underbrace{-1, -3, -2}_{1})$.

Step 4: The vector c is transmitted and received as $Y = c + Z$ where Z is $\mathcal{N}(0, \sigma^2 I_n)$. Write $Y = (Y_1, \ldots, Y_k)$ where each Y_i is in $\mathbb{R}^r$, similarly write $Z = (Z_1, \ldots, Z_k)$ where each Z_i is in $\mathbb{R}^r$.

(a) Assuming all 2^k messages are equally likely, show that the procedure: "*for each* $i = 1, \ldots, k$, *let* $\hat{b}_i = \arg\min_{b \in \{0,1\}} \|Y_i - v_b\|$ *and estimate the transmitted message as* $(\hat{b}_1, \ldots, \hat{b}_k)$" minimizes the probability of error.

(b) With $d^2 = \|v_0 - v_1\|^2$, what is $\Pr\left(\hat{b}_i \neq b_i\right)$ (i.e., the probability that the ith bit of the message is received incorrectly)? What is Pr(error) (i.e., the probability that some bit is received incorrectly)? How does Pr(error) compare with $\min\left\{1, kQ(\frac{d}{2\sigma})\right\}$?

(c) With $d^2 = \|v_0 - v_1\|^2$, consider a new system where v_0 and v_1 are replaced by the scalars $d/2$ and $-d/2$. The codewords of the new system $\left(\pm\frac{d}{2}, \ldots, \pm\frac{d}{2}\right)$ are now in $\mathbb{R}^k$ instead of $\mathbb{R}^{kr}$. What can you say about the average energy $\mathcal{E}$, average energy per bit $\mathcal{E}_b$, the bit error probabilities, and the message error probability of the new system in terms of the corresponding quantities of the original system?

(d) Suppose we need to send k bits (e.g., $k = 100$) using a system as above, and we require the message error probability to be at most α (e.g., $\alpha = 10^{-2}$). Suppose a_1 and a_2 satisfy $Q(a_1) = \frac{\alpha}{k}$ and $(1 - 1/e)\, Q(a_2) = \frac{\alpha}{k}$. Show that if $d/(2\sigma) < a_2$, the error probability requirement cannot be met. What will happen if $d/(2\sigma) \geq a_1$?

Moral: (1) If the message is sent 'bit by bit', as in the system described in the beginning of the problem, one may as well use the simpler system in (c). (2) In a system designed as above, the minimal possible value of $(d/2\sigma)^2$ lies between a_2^2 and a_1^2. (Note that $(d/2\sigma)^2$ equals $\mathcal{E}_b/\sigma^2$.)

PROBLEM 3. Consider a communication system with $2n$ equally likely codewords $\pm\sqrt{\mathcal{E}}e_j$, $j = 1, \ldots, n$ where $e_1, \ldots, e_n$ are the unit coordinate vectors in $\mathbb{R}^n$. The receiver receives $Y = c + Z$ where c is one of these codewords and Z is $\mathcal{N}(0, \sigma^2 I_n)$. As the system is sending $k = \log_2(2n)$ bits, the choice $\mathcal{E} = \sigma^2 A \log_2(2n)$ results in an energy per bit $\mathcal{E}_b$ satisfying $\mathcal{E}_b/\sigma^2 = A$.

The MAP rule for this setup is given by the following: find the j for which $|Y_j|$ is largest, and decide that the codeword $\text{sign}(Y_j)\sqrt{\mathcal{E}}e_j$ was transmitted.

Consider the following alternative decoding method. Pick a threshold $t = \alpha\sqrt{\mathcal{E}}$ with $0 \leq \alpha < 1$. If there is exactly one j for which $|Y_j| > t$, decide that the codeword $\text{sign}(Y_j)\sqrt{\mathcal{E}}e_j$ was transmitted. If there is no j for which $|Y_j| > t$ or several j's for which $|Y_j| > t$, then the decoder declares an error. Note that the error probability of the MAP decoder is upper bounded by the error probability of this (suboptimal) decoder, so any upper bound on the error probability of this decoder also upper bounds the probability of error of the MAP rule.

(a) Show that the probability of error (either by declaring an error, or by deciding on a wrong codeword) of this decoder satisfies

$$\begin{aligned}\Pr(\text{error}) &\leq Q\left((1-\alpha)\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right) + 2(n-1)Q\left(\alpha\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right) \\ &< Q\left((1-\alpha)\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right) + 2^k Q\left(\alpha\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right).\end{aligned}$$

(b) Recall that $\mathcal{E} = kA\sigma^2$. Show that the probability of error is further upper bounded by

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

Also show that if $A > 2\ln 2$ there is an $0 < \alpha < 1$ for which the probability of error approaches zero as k gets large.
Hint: Use (a) and $Q(x) \le \frac{1}{2}\exp(-x^2/2)$.

Moral: If we are given an energy budget in the form energy/bit $= \mathcal{E}_b$, and if this budget satisfies $\mathcal{E}_b/\sigma^2 > 2\ln 2$, then we can, by taking k large enough, meet any desired error probability requirement.

(c) Suppose that $A > 2\ln 2$. Show that

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

Hint: Use (b) and consider the choice $\alpha = \frac{1}{2}\left(1 + \frac{2\ln 2}{A}\right)$. Don't forget to verify that $\alpha < 1$.

(d) For $A = 4, 6, 8, 10, 12$, what are the values $k(A)$ of k that will make the upper bound to the error probability in (c) less than 10^{-3}?

(e) For each of the five values of A in (d), consider a bit-by-bit communication system (à la Problem 2 above) with $\mathcal{E}_b/\sigma^2 = A$ that sends a $k(A)$-bit message. Find the message error probabilities of these systems.

PROBLEM 4. Suppose $c_1, \ldots, c_m$ are codewords in $\mathbb{R}^n$ and all messages are equally likely. When codeword i is sent, the receiver receives $Y = (Y_1, Y_2)$ in $\mathbb{R}^{2n}$ with either

$$(1)\ Y_1 = c_i + Z,\ Y_2 = \tilde{Z}, \quad \text{or} \quad (2)\ Y_1 = \tilde{Z},\ Y_2 = c_i + Z,$$

with the two cases being equally probable. Here Z and $\tilde{Z}$ are independent, Z is $\mathcal{N}(0, \sigma^2 I_n)$, and $\tilde{Z}$ is $\mathcal{N}(0, \tau^2 I_n)$. If the receiver had "side information" telling it which of (1) and (2) occurred, then it could have decoded the message i based on the part of Y that equals $c_i + Z$. But the receiver does not have such information.
Let $H = (i, b)$ where the binary value b indicates which of (1) and (2) took place.

(a) Consider the following rule to decide the value of H from the observation (y_1, y_2). Find $i_1 = \arg\min \|y_1 - c_i\|$, let $i_2 = \arg\min \|y_2 - c_i\|$. Let $d_1 = \frac{\|y_1 - c_{i_1}\|^2}{\sigma^2} + \frac{\|y_2\|^2}{\tau^2}$ and $d_2 = \frac{\|y_2 - c_{i_2}\|^2}{\sigma^2} + \frac{\|y_1\|^2}{\tau^2}$. Decide

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

Does this rule minimize $\Pr\left(\hat{H} \ne H\right)$?

(b) Let $\hat{i}$ be the first component of $\hat{H}$, i.e., $\hat{i} = i_1$ if $d_1 < d_2$ and $\hat{i} = i_2$ else. Does this rule minimize $\Pr\left(\hat{i} \ne i\right)$?

Let $\hat{i}_{\text{o}}(y_1, y_2, b)$ be the MAP estimator of a receiver that somehow has access to the side information as mentioned above, i.e., it is the decision made from the observation (y_1, y_2, b).

(c) Let $\hat{b}$ be the second component of $\hat{H}$ as above, i.e., $\hat{b} = 1$ if $d_1 < d_2$ and $\hat{b} = 2$ else. Justify the following inequalities:

$$\begin{aligned}
\Pr\left(\hat{i}_{\mathrm{o}} \neq i\right) &\overset{(c_0)}{\leq} \Pr\left(\hat{i} \neq i\right) \\
&\overset{(c_1)}{\leq} \Pr\left(\hat{H} \neq H\right) \\
&\overset{(c_2)}{=} \Pr\left(\hat{b} \neq b\right) + \Pr\left(\hat{b} = b \text{ and } \hat{i} \neq i\right) \\
&\overset{(c_3)}{=} \Pr\left(\hat{b} \neq b\right) + \Pr\left(\hat{b} = b \text{ and } \hat{i}_{\mathrm{o}} \neq i\right) \\
&\overset{(c_4)}{\leq} \Pr\left(\hat{b} \neq b\right) + \Pr\left(\hat{i}_{\mathrm{o}} \neq i\right).
\end{aligned}$$

Moral: The message error probabilities of the receiver with and without side information differ at most by $\Pr\left(\hat{b} \neq b\right)$. If $\Pr\left(\hat{b} \neq b\right)$ is small, then not much is lost by not having the side information about the channel state.

(d) Suppose that $H = (i, 1)$. Show that $\hat{b} \neq 1$ only if there exists $i' \in \{1, \ldots, m\}$ with

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

Hint: How does the left-hand side compare to d_1?

(e) From now on, suppose $\sigma = \tau$. Use the union bound to upper bound $\Pr\left(\hat{b} \neq 1 \mid H = (i, 1)\right)$ by $\sum_{i'=1}^{m} Q\left(\sqrt{\frac{\|c_i\|^2 + \|c_{i'}\|^2}{4\sigma^2}}\right)$.

(f) Assume that $\|c_i\| = \sqrt{\mathcal{E}}$ for all $i \in \{1, \ldots, m\}$ and $\frac{\mathcal{E}_b}{\sigma^2} > 4 \ln 2$ where $\mathcal{E}_b$ is the energy per bit. Use (c) and (e) to show that $\Pr\left(\hat{i} \neq i\right) - \Pr\left(\hat{i}_{\mathrm{o}} \neq i\right)$ approaches 0 as m grows.
Hint: What happens to $\Pr\left(\hat{b} \neq b\right)$ as m grows?