Assignment 2, due: Mar. 23, 2026 (in class, hard copy + pdf via email)

Note 1: Late submissions will not be accepted. Please also keep an extra copy of your submission for your own record.

Note 2: Submission format: can be hand-written or printed (handwriting must be readable), letter-sized paper, stapled, cover page must include course code/title, assignment number, your name and student number, date of submission, all pages numbered. Please include all details/steps in your answers (e.g. derivations, analytical steps, etc.) not just final results. All symbols used must be clearly defined.

Note 3: Before doing the assignment, read relevant chapters/sections of the textbook (or other recommended books), study carefully all examples, attempt some end of chapter problems. Extra references are given below. No submission is needed for this part (this is very helpful for your studies though).

1. Consider the system transmitting binary message m over a Gaussian additive noise channel:

$$y = m + \xi \tag{1}$$

where $m = 0$ or 1, equiprobably, and ξ is Gaussian noise of zero mean and variance (power) σ_0^2, independent of m.

   (a) Find a posteriori (conditional) message probabilities $P(m = 0|y)$ and $P(m = 1|y)$ given Rx measurement (observation) y, do all required calculations and express them in a neat closed form. What are their values when $y \to \infty$ or $y \to -\infty$? Why? Find the values of y such that a priori and a posteriori message probabilities are the same, i.e. $P(m|y) = P(m)$ (so that the measurement y does not provide us any new information). Provide analytical answers for this part, not simulation-based, but the latter can be used to validate the analysis.

   (b) Plot $P(m = 0|y)$ and $P(m = 1|y)$ in the same figure as functions of y for $\sigma_0 = 1$; use sufficiently large range of y so that the tendencies are clearly visible for both. How do they compare to the a priori message probabilities $P_0 = P(m = 0)$ and $P_1 = P(m = 1)$? Based on your plot, identify decision regions Ω_0 and Ω_1 and explain the operation of optimal Rx. Also explain when the Rx makes errors and how error probabilities can be determined graphically.

   (c) To see the impact of noise variance, repeat the plot of the previous part for $\sigma_0 = 0.5$ and $\sigma_0 = 0.2$. Comment on the impact of σ_0 on the variation of $P(m|y)$ with y and on the Rx operation and, in particular, on error probability.

   (d) Find σ_0 for the conditional probabilities $P(m|y)$ plotted in slide 25 of Lecture 6. For this setup, the Rx measures $y = 0$ and, based on this measurement, makes its decision $\widehat{m}$. Find $\widehat{m}$ and the probability that this decision is incorrect. Repeat for $y = 0.5$. Explain the results you have obtained.

2. Consider the system transmitting 1 bit using bipolar 2-PAM modulation over the following additive-noise channel:

$$y = s + \xi \tag{2}$$

where $p(\xi) = \alpha e^{-|\xi|}$ is the pdf of noise ξ; the transmitted signal $s = A$ if $m = 1$ and $s = -A$ if $m = 0$; m represents random bit being transmitted; $P_1 = P_0$.

   (a) Find constant α (so that $p(\xi)$ is a valid pdf).

   (b) Find optimal decision rule (that minimizes the error probability/BER). Do all required calculations and express the rule in a neat closed-form format (as simple as possible). Compare this rule with that of Gaussian noise channel and comment/explain the result of such comparison.

   (c) Find the error probability for this optimal rule. Again, do all required calculations and express the result in a neat closed-form (as simple as possible). Using this, find the average number of errors when a block of N bits is transmitted.

   (d) Now, compare the error probability P_e of this channel with that for the Gaussian noise channel of the same variance (i.e. the noise variances in both channels must be the same for fair comparison). Plot $P_e(A)$ for both channels in the same figure (use proper line format so that the BER is clearly visible for both cases).

   (e) Based on the last part, which noise is better? Explain why.

3. In this question, we will use Monte-Carlo method [5] of simulations (in Matlab) to validate the analytical BER of the previous question (this is a very powerful method that can be applied to most problems involving random variables and probabilities and even to some problems beyond that, e.g. computing integrals or estimating the value of π). For this, we generate many (N) random and independent realizations of m and ξ,

$$y_i = s_i + \xi_i, \ i = 1...N \tag{3}$$

where s_i corresponds to m_i, apply the optimal decision rule $D\{\cdot\}$ to each y_i to find estimated bit $\widehat{m}_i$

$$\widehat{m}_i = D\{y_i\} \tag{4}$$

and, by comparing it with the original transmitted bit m_i, decide whether an error took place,

$$e_i = \begin{cases} 0, & \text{if } \widehat{m}_i = m_i \\ 1, & \text{otherwise.} \end{cases} \tag{5}$$

Then, we count the total number N_e of errors,

$$N_e = \sum_{i=1}^{N} e_i \tag{6}$$

and find empirical BER (error probability) $\widehat{P}_e$ as the ratio

$$\widehat{P}_e = \frac{N_e}{N} = \frac{1}{N}\sum_{i=1}^{N} e_i \tag{7}$$

Note that this also represents the empirical BER when the block of N bits is transmitted (which is very much relevant to practical systems and explains why Monte-Carlo simulation method is extensively used in practice). For the estimated empirical BER to be accurate, N has to be sufficiently large.

(a) Find out how large is large enough. Hint: set $A = 2$ and plot $\widehat{P}_e(N)$ versus N for $N = 1, 2, 3...$ (for each value of N, the whole set of m_i and ξ_i has to be generated afresh, e.g. m_1 for $N = 2$ and $N = 3$ will be independent of each other). Note the value of N after which $\widehat{P}_e(N)$ almost does not change anymore with N. To make this experiment valid, you have to run it several times and select largest N (across all runs) as "large enough". Plot $\widehat{P}_e(N)$ for 3 representative runs and indicate "large enough" N for each.

(b) Note that "large enough" N depends on A, i.e. on the true error probability P_e. To determine this relationship, repeat the previous part for several values of A and thus P_e, e.g. for such A that $P_e = 10^{-1},\ 10^{-2},\ 10^{-3}$ and find respective "large enough" N. Can you express "large enough" N as a function of P_e or, equivalently, formulate this "large enough" rule in terms of N_e? This "large enough" rule can be approximate or even "rough", e.g. "order of magnitude" (up to a factor of 10 on the larger side); high accuracy is not needed here, but the respective $\widehat{P}_e(N)$ should be accurate (say, 10%). From these experiments, validate the following conclusions: (i) "largest "large enough" N corresponds to smallest P_e (or largest A)" and (ii) using larger N will not "hurt".

(c) Now, plot $\widehat{P}_e(A)$ and $P_e(A)$ in the same figure for the range of A corresponding to the BER range $[1/2,\ 10^{-3}]$ (using "large enough" N for the empirical BER). Do they agree with each other? Hint: for this experiment to be valid, new sets of m_i and ξ_i has to be generated afresh for each new value of A.

(d) Finally, repeat the last part for Gaussian noise channel thus validating the analytical BER we discussed in class (this should also serve as a validation of your simulation codes).

4. Optional (bonus) question: apply the law of large numbers (LLN) to (7) to argue (more or less rigorously) that

$$\widehat{P}_e(A) \ \rightarrow\ P_e(A) \text{ as } N\ \rightarrow\ \infty \tag{8}$$

This serves as a theoretical foundation for the applicability of Monte-Carlo method to the above problems and also tells us that N must be "large enough".

5. Optional (bonus) question: While LLN tells us that $\widehat{P}_e\ \rightarrow P_e$ as $N\ \rightarrow\ \infty$, this is not possible in practice where N always stays finite. Thus, a question emerges "how

large N is large enough?". This can be addressed in a theoretical way to complement simulation-based studies above as follows.

(i) Show that $\widehat{P}_e$ in (7) is an unbiased estimator of P_e, that is

$$\mathbb{E}\{\widehat{P}_e\} = P_e \tag{9}$$

where $\mathbb{E}\{\cdot\}$ is the expectation with respect to all random variables, i.e. all e_i or, equivalently, all m_i and ξ_i.

(ii) Find the standard deviation $\sigma_{\widehat{P}_e}$ of $\widehat{P}_e$ and use it to determine block length N such that the estimation accuracy is 10% (with respect to the true BER P_e), i.e. $\sigma_{\widehat{P}_e} = 0.1P_e$. Find the respective average number $\overline{N}_e$ of errors. Show that the estimation accuracy improves as N increases and that $\sigma_{\widehat{P}_e} \to 0$ as $N \to \infty$ (the latter implies that any desired accuracy can be achieved by using large-enough N). Does this agree with your simulation-based studies above?

6. Optional (bonus) question: In the above questions, we considered equiprobable messages, $P_0 = P_1$. In practice, this probabilities may be unequal. For the additive-noise Gaussian channel, justify the following statement: "equally-likely messages are most difficult to communicate". Hints: (i) find error probability when $P_0 \neq P_1$, and (ii) show that this error probability is always upper bounded by that for equally-likely messages (for any noise variance or message amplitude). Thus, equiprobable messages indeed represent worst-case scenario: if the system works well for equiprobable messages, it will also work well for unequal-probability messages. In fact, this conclusion holds for a broad class of channels.

For all relevant questions, please include Matlab codes in your submission.

References

[1] J.M. Wozencraft, I.M. Jacobs, Principles of Communication Engineering, Wiley: New York, 1965. Ch. 2 (especially "A Communication Example"), Ch. 4.1-4.2.

[2] S. Haykin, Digital Communication Systems, Wiley, 2014. Ch. 3.11–3.15

[3] R.E. Ziemer, W.H. Tranter, Principles of Communications, Wiley, 2009. Ch. 10.1.

[4] A. Lapidoth, A Foundation in Digital Communication, Cambridge University Press, 2017. Ch. 20.1-20.10 (the latter section is especially relevant)

[5] `https://en.wikipedia.org/wiki/Monte_Carlo_method`

Note on plagiarism

Plagiarism (presenting somebody's else solution/report as your own, where ”somebody” means any source, including Internet (Google, ChatGPT, etc.), “cut-and-paste” from a student to a student, other forms of “borrowing” the material) is absolutely unacceptable and will be penalized. Each student must submit his own solutions. If two (or more) identical or almost identical sets of solutions are found (including those from Internet), each student involved receives 0 (zero) for that particular assignment. If this happens twice, the students involved receive 0 (zero) for the entire assignment component of the course in the marking scheme and the case will be send to the Dean's office for further investigation and further penalties.