Problem 1. (8 points)
Suppose $c_1, \dots, c_m$ (all in $\mathbb{R}^n$) are the codewords of a communication system with two receivers. Receiver A observes (A_1, A_2) where $A_1 = c_i + Z_1$, $A_2 = Z_1 + Z_2$, Receiver B observes $B_1 = c_i + Z_1$, $B_2 = c_i - Z_2$, where $Z_1 \in \mathbb{R}^n$ and $Z_2 \in \mathbb{R}^n$ are additive noises.

(a) (2 pts) Show that receiver A and receiver B have the same error probability (assuming both implement the optimal decision rule).
Hint: Show that receiver A can form the observation of receiver B and vice versa.

Solution: From (B_1, B_2) we can compute $(B_1, B_1 - B_2) = (A_1, A_2)$, similarly from (A_1, A_2) we can compute $(A_1, A_1 - A_2) = (B_1, B_2)$. Hence (A_1, A_2) is a sufficient statistic for (B_1, B_2) and (B_1, B_2) is a sufficient statistic for (A_1, A_2), which implies that they must have the same error probability.

For the rest of the problem suppose Z_1 and Z_2 are independent and both are $\mathcal{N}(0, \sigma^2 I_n)$.

(b) (2 pts) Show that $U = (Z_1 - Z_2)/2$ and $W = (Z_1 + Z_2)/2$ are independent.

Solution: Because of Gaussianity it suffices to check that $\mathbb{E}[U_i W_k] = 0$ for all $i, k = 1, \dots, n$. This is obvious for $i \neq k$, since we have $\mathbb{E}[U_i W_k] = \mathbb{E}[U_i]\mathbb{E}[W_k] = 0$. For $i = k$ we get

$$\begin{aligned}\mathbb{E}[U_i W_i] &= \frac{1}{4}\mathbb{E}\big[(Z_{1i} + Z_{2i})(Z_{1i} - Z_{2i})\big] \\ &= \frac{1}{4}\mathbb{E}\big[Z_{1i}^2 - Z_{2i}^2\big] = 0.\end{aligned}$$

(c) (2 pts) Show that for receiver B, $T = (B_1 + B_2)/2$ is a sufficient statistic.

Solution: With $Y = (T, W)$, $f_{Y|H}(y|i) = f_{UW}\big((t - c_i), w\big)$, which, by (b) equals $f_U(t - c_i) f_W(w)$ which is of the form $g_i(t)h(y)$, so we are done by the Fisher-Neyman theorem.

Consider now a receiver C that observes B_1 but with noise Z_1 being $\mathcal{N}(0, \tau^2 I_n)$.

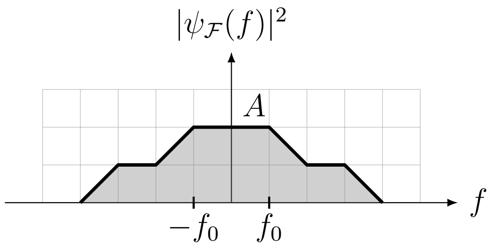
(d) (2 pts) Can you find a τ_0 such that receiver C will perform better/worse than receiver B when τ^2 is less/more than τ_0^2?
Hint: Clearly τ_0 should depend on σ.

Solution: Receiver C observes $c_i + Z$, where $Z \sim \mathcal{N}(0, \tau^2 I_n)$. Recall that for receiver B, $T = c_i + (Z_1 - Z_2)/2$ is a sufficient statistic, and $(Z_1 - Z_2)/2 \sim \mathcal{N}\big(0, \frac{\sigma^2}{2} I_n\big)$. Hence, $\tau_0^2 = \sigma^2/2$ is the critical value of τ^2.

Remarks: The conclusion from part (d) is that observing $c_i + Z$ with $Z \sim \mathcal{N}(0, \tau^2 I_n)$ is equivalent to observing $(c_i + Z_1, c_i - Z_2)$ with Z_1, Z_2 i.i.d. $\mathcal{N}(0, \sigma^2 I_n)$ when $\tau^2 = \sigma^2/2$, which makes sense intuitively.

Problem 2. (10 points)

Suppose a pulse $\psi(t)$ has Fourier transform $\psi_{\mathcal{F}}(f)$, with $|\psi_{\mathcal{F}}(f)|^2$ sketched as below.



(a) (3 pts) Find the smallest positive T and the corresponding $A = |\psi_{\mathcal{F}}(0)|^2$ that will make $\big(\psi_j(t) = \psi(t - jT) : j \in \mathbb{Z}\big)$ an orthonormal collection of waveforms.

Solution: We use the Nyquist criterion: the figure has the "band-edge symmetry property" around $2.5f_0$, so $\sum_k |\psi_{\mathcal{F}}(f - kf_1)|^2 = A$ when $f_1 = 5f_0$. So with $T = 1/f_1 = 1/(5f_0)$, the collection of pulses $\big(\psi_k : k \in \mathbb{Z}\big)$ with $\psi_k(t) = \psi(t - kT)$ an orthogonal collection. To make the collection orthonormal, we also need to choose $A = T$.

With T and A as in (a), let $w(t) = \sum_{j\in\mathbb{Z}} \sqrt{\mathcal{E}_s} X_j \psi(t - jT)$ and $W(t) = w(t + \Theta)$ where Θ is uniform in $[0, T]$ and independent of $\big(X_j : j \in \mathbb{Z}\big)$.

(b) (3 pts) Suppose X_j are i.i.d., with $\Pr(X_j = 1) = \Pr(X_j = -1) = 1/2$. Sketch the power spectral density $S_W(f)$ of $W(t)$. Which (if any) of the values $S_W(0)$, $S_W\big(1/(2T)\big)$, $S_W\big(1/T\big)$ are equal to 0?

Solution: The power spectral density of W is given by

$$S_W(f) = \mathcal{E}_s \frac{|\psi_{\mathcal{F}}(f)|^2}{T} \sum_k K_X[k] \exp(-\mathrm{j}2\pi kfT).$$

Here $K_X[k] = \mathbb{1}\{k = 0\}$, so $S_W(f) = \mathcal{E}_s|\psi_{\mathcal{F}}(f)|^2/T$. Hence, $S_W(0) = \mathcal{E}_s|\psi_{\mathcal{F}}(0)|^2/T = 1$, $S_W\big(1/(2T)\big) = \mathcal{E}_s|\psi_{\mathcal{F}}(2.5f_0)|^2/T$ and $S_W\big(1/T\big) = \mathcal{E}_s|\psi_{\mathcal{F}}(5f_0)|^2/T$. Of the three values, only the last one is zero.

Suppose we are requested to ensure $S_W(0) = 0$. A colleague suggests setting $X_j = B_j - B_{j-1}$ where B_j are i.i.d., with $\Pr(B_j = 0) = \Pr(B_j = 1) = 1/2$.

(c) (2 pts) With $\big(X_j : j \in \mathbb{Z}\big)$ as above, find $K_X[k] = \mathbb{E}[X_j X_{j+k}]$.

Solution: The values of k for which $K_X[k]$ are not zero are 0, 1 and -1, with $K_X[0] = 1/2$, $K_X[1] = K_X[-1] = -1/4$.

(d) (2 pts) Find $S_W(f)$ with this $\big(X_j : j \in \mathbb{Z}\big)$. Does the suggestion of our colleague work?

Solution: Using the values of $K_X[k]$ from (c) in the expression for the power spectral density, we have

$$\begin{aligned} S_W(f) &= \sum_k K_X[k] \exp(-j2\pi kfT) \\ &= \frac{1}{2} - \frac{1}{2}\cos(2\pi fT) = \sin^2(\pi fT), \end{aligned}$$

so $S_W(f) = \mathcal{E}_s \frac{|\psi_{\mathcal{F}}(f)|^2}{T} \sin^2(\pi fT)$, which gives $S_W(0) = 0$. Hence, the suggestion of our college does indeed work.

Remarks: In some cases, it is useful to have $S_W(0) = 0$, since it is difficult to implement an amplifier for "zero" frequencies in practice. This example shows how a simple encoding of X_i's gives us this feature.

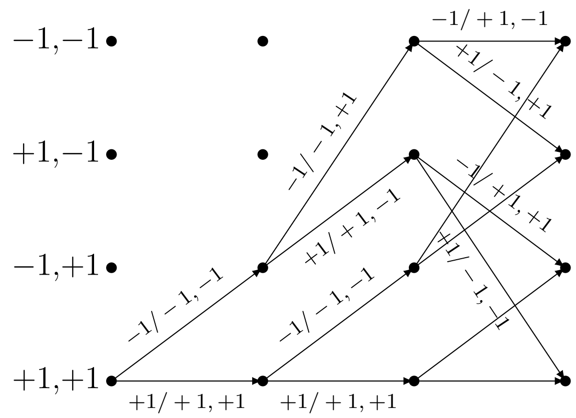
Problem 3. (13 points)
Suppose the bit stream $b_1, b_2, \ldots$ with $b_i = \pm 1$ is encoded by a convolutional encoder to the symbol stream $x_1, x_2, \ldots$ via

$$x_{2j-1} = b_j b_{j-2}, \quad x_{2j} = b_j b_{j-1} b_{j-2}$$

(as in the running-example used in the book and class). In computing the x's we assume $b_0 = b_{-1} = +1$. Recall that $T(I, D)$ for this encoder is $ID^5/(1 - 2ID)$.
Suppose that the receiver receives a sequence y_i where $y_i = x_i$ with probability $1 - p$ and $y_i = 0$ with probability p. Which of these two alternatives happens is chosen independently for each i. Observe that if $y_i \neq 0$, then the receiver is sure that $x_i = y_i$.

(a) (3 pts) Draw a Trellis section that describes the encoder map.

Solution: A Trellis section showing all transitions is given by:

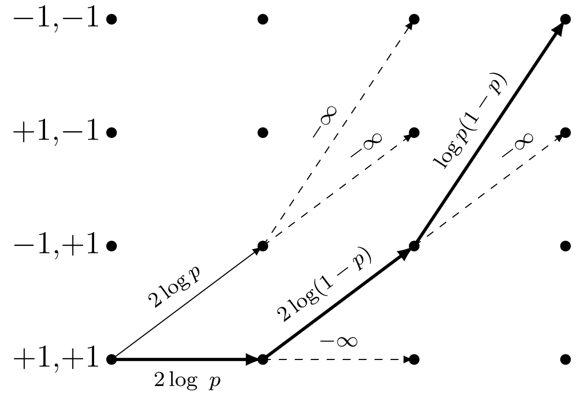


(b) (2 pts) Describe the Viterbi decoder, by providing the following information: given the received sequence $y_1, y_2, \ldots$, we associate the branch labeled (x_{2j-1}, x_{2j}) in the trellis a metric given by _____, and the corresponding path metric is to be *maximized.*

Solution: To the branch (x_{2j-1}, x_{2j}), we associate the metric $\log p_{Y|X}(y_{2j-1}|x_{2j-1}) + \log p_{Y|X}(y_{2j}|x_{2j})$ to ensure that the corresponding path metric is to be maximized, where $p_{Y|X}(y|x) = 1 - p$ if $y = x$ and p if $y = 0$ (and zero otherwise).

(c) (3 pts) Suppose $y_1, y_2, y_3, y_4, y_5, y_6 = 0, 0, -1, -1, 0, +1$. What is the Viterbi-decoded sequence $\hat{b}_1, \hat{b}_2, \hat{b}_3$?

Solution: Labeling each branch of the trellis with the branch metric in (c) for the given y sequence, we have:



Hence, the Viterbi-decoded sequence is $\hat{b}_1, \hat{b}_2, \hat{b}_3 = +1, -1, -1$.

(d) (2 pts) Suppose the true bit sequence $(b_1, \ldots, b_k)$ is the all-plus sequence. Observe that this is encoded as an all-plus x sequence $(x_1, \ldots, x_n)$, and thus the received sequence contains only 0's and +1's. Suppose $b' = (b'_1, \ldots, b'_k)$ an incorrect bit sequence, with encoding $(x'_1, \ldots, x'_n)$. Show that the probability that the Viterbi decoder will decide b' is at most p^d where d is the number of −1's in $(x'_1, \ldots, x'_n)$.

Solution: For the Viterbi decoder to decide b' corresponding to x' with d elements equal to −1, it is necessary that the channel must have "erased" (i.e., changed to a 0) at least d elements. The probability that this occurs is exactly p^d, and hence, the probability that the Viterbi decoder decides b' is at most p^d.

(e) (3 pts) By making use of $T(I, D)$ and (d), find an upper bound to the bit error probability (in terms of p).

Solution: Recall that the bit error probability is given by

$$P_b = \frac{1}{kk_0} \sum_{j=0}^{k-1} \sum_h i(h)\pi(h),$$

where the sum is over all detours h that start at depth j w.r.t. the reference, $i(h)$ is the input distance between the detour h and the reference path, and $\pi(h)$ is the probability that the detour h is taken. This can be upper bounded by $\sum_{i=1}^{\infty} \sum_{d=1}^{\infty} i p^d a(i,d)$, by using part (d) to observe that p^d is an upper bound to probability that the detour with output distance d is taken. Let $A(D) = \frac{\partial}{\partial I} T(I, D)$ evaluated at $I = 1$. Then the required expression $\sum_{i=1}^{\infty} \sum_{d=1}^{\infty} i p^d a(i,d)$ is exactly $A(p) = \frac{p^5}{(1-2p)^2}$.

Remarks: We have seen examples of using the Bhattacharya bound as the "z" parameter when dealing with convolutional codes for binary-input channels. When the channel is a binary erasure channel with erasure probability p (as here), the Bhattacharya bound happens to be exactly p. Here we see (via part (d)), an alternate way to obtain the same upper bound on the bit error probability.

Problem 4. (11 points)

Suppose $\phi(t)$ and $\xi(t)$ are two complex-valued low-pass waveforms, containing frequencies only in the frequency range $[-B, B]$. Suppose $f_0 > B$. Consider the following sequence of operations done on a complex number c to obtain a complex number y:

1. construct the real waveform $w(t) = \sqrt{2}\,\mathrm{Re}\{c\,\phi(t)\exp(\mathrm{j}2\pi f_0 t)\}$

2. construct the complex waveform $r(t) = \sqrt{2}\, w(t) \exp(-\mathrm{j}2\pi f_0 t)$

3. take the inner product of r and ξ to form $y = \langle r, \xi \rangle$.

(a) (2 pts) Show that $y = c\langle \phi, \xi \rangle$.

Solution: Since $\mathrm{Re}(z) = \frac{z+z^*}{2}$, we have

$$w(t) = \frac{1}{\sqrt{2}}\big[c\,\phi(t)\exp(\mathrm{j}2\pi f_0 t) + c^*\phi(t)^*\exp(-\mathrm{j}2\pi f_0 t)\big].$$

Then, $r(t) = \sqrt{2}\, w(t)\exp(-\mathrm{j}2\pi f_0 t)$ is given by

$$r(t) = c\,\phi(t) + c^*\phi(t)^*\exp(-\mathrm{j}4\pi f_0 t).$$

Since ξ is a low-pass waveform with frequency restricted to $[-B, B]$ and $f_0 > B$, we have that the inner product of the second term above with ξ is zero. Hence, $y = \langle r, \xi \rangle = c\langle \phi, \xi \rangle$.

Consider now a transmitter that transforms the message i to a bandpass transmitted waveform $w_i(t)$ as follows (exactly in the way we discussed in class):

$$\big[i\big] \to \big[c_i \in \mathbb{C}^n\big] \to \big[w_{i,E}(t) = \sum_j c_{ij}\phi_j(t)\big] \to \big[w_i(t) = \sqrt{2}\mathrm{Re}\{w_{i,E}(t)\exp(\mathrm{j}2\pi f_0 t)\}\big].$$

Here $\phi_1, \dots, \phi_n$ are complex orthonormal, baseband waveforms (all supported in the frequency range $[-B, B]$), and $f_0 > B$.
The signal $w_i(t)$ is transmitted on an AWGN channel with noise intensity $N_0/2$, the received signal is $R(t)$.
The receiver operates as follows:

$$\big[R(t)\big] \to \big[R(t)\sqrt{2}\exp(-\mathrm{j}2\pi f_0 t)\big] \to \big[Y \in \mathbb{C}^n \text{ where } Y_j = \langle R, \xi_j \rangle\big] \to \big[\text{decision device}\big] \to \big[\hat{i}\big].$$

Note that the receiver forms Y using complex orthonormal baseband basis functions $\xi_1, \dots, \xi_n$ (all in the frequency range $[-B, B]$); had the ξ's been equal to ϕ's, then we would have our optimal receiver, with error probability $p_{\mathrm{opt}}(N_0)$.

(b) (3 pts) Find an $n \times n$ matrix A such that Y can be written in the form $Y = Ac_i + Z$, where Z is $\mathcal{N}_{\mathbb{C}}(0, N_0 I_n)$, i.e., $Z = Z_R + \mathrm{j}Z_I$ with $Z_R, Z_I \sim \mathcal{N}\big(0, \frac{N_0}{2} I_n\big)$ being independent.

Hint: Use (a) and express the entries A_{kj} of the matrix A in terms of the ϕ's and ξ's.

Solution: Observe that $Y_k = \langle R, \xi_k \rangle$, where $R(t) = w_i(t) + N(t)$, i.e.,

$$R(t) = \left(\sum_j \sqrt{2}\mathrm{Re}\{c_{ij}\phi_j(t)\exp(\mathrm{j}2\pi f_0 t)\}\right) + N(t).$$

By part (a), we have that the inner product of R with ξ_k is given by $\left(\sum_j c_{ij}\langle \phi_j, \xi_k \rangle\right) + \langle N, \xi_k \rangle$. This second term is simply a circularly symmetric complex Gaussian random variable $Z \sim \mathcal{N}_{\mathbb{C}}(0, N_0 I_n)$ (since the ξ's form an orthonormal basis). The first term is exactly Ac_i, where the elements of A are given by $A_{kj} = \langle \phi_j, \xi_k \rangle$, and we are done.

(c) (3 pts) Suppose $n = 2$ and $\langle \phi_1, \xi_1 \rangle = 0, \langle \phi_1, \xi_2 \rangle = \mathtt{j}, \langle \phi_2, \xi_1 \rangle = 1, \langle \phi_2, \xi_2 \rangle = 0$. Show how the decision device (that produces $\hat{\imath}$ from Y) can be implemented so that the error probability is equal to $p_{\text{opt}}(N_0)$.

Solution: The optimal receiver would have had $\tilde{Y} = c_i + \tilde{Z}$, where $\tilde{Z} \sim \mathcal{N}_{\mathcal{C}}(0, N_0 I_n)$. By (b), our Y has components $Y_1 = c_{i,2} + Z_1$, and $Y_2 = \mathtt{j}c_{i,1} + Z_2$. Since Z_1, Z_2 are independent and circularly symmetric, $Y' = (-\mathtt{j}Y_2, Y_1) = c_i + (-\mathtt{j}Z_2, Z_1)$ has the same statistics as $\tilde{Y}$, which can be used to construct a $\hat{\imath}$ that has the same statistics (and thus the same error probability) as that of the optimal receiver.

(d) (3 pts) Suppose $n = 2$ and $\langle \phi_1, \xi_1 \rangle = \langle \phi_1, \xi_2 \rangle = \langle \phi_2, \xi_2 \rangle = 1/2$, $\langle \phi_2, \xi_1 \rangle = -1/2$. Show that with the best possible decision device, the error probability will be $p_{\text{opt}}(2N_0)$.

Solution: By part (b), this Y has components $Y_1 = \frac{c_{i,1} - c_{i,2}}{2} + Z_1$ and $Y_2 = \frac{c_{i,1} + c_{i,2}}{2} + Z_1$. This is in a one-to-one correspondence with $Y' = (Y_1 + Y_2, Y_2 - Y_1) = c_i + (Z_1 + Z_2, Z_2 - Z_1)$. Since Z_1, Z_2 are independent and circularly symmetric, this Y' has the same statistics as $\tilde{Y} = c_i + \tilde{Z}$, where $\tilde{Z} \sim \mathcal{N}_{\mathcal{C}}(0, 2N_0 I_n)$, which can be used to construct a $\hat{\imath}$ that has the same statistics (and thus the same error probability) as that of the optimal receiver when used over an AWGN channel of noise intensity N_0.

Remarks: This problem is an example of a "mismatched" receiver — we encode using the orthonomal basis represented by ϕ, but we encode using an orthonormal basis represented by ξ. In part (b), we see an example where this does not affect the error probability, whereas in part (c), we see an example where it does affect the error probability. In fact, the error probability will remain unchanged if and only if the matrix A consisting of the inner products of ϕ's and ξ's performs a transformation that is purely rotative, as in (b) — under the transformation $A = \begin{pmatrix} 0 & 1 \\ \mathtt{j} & 0 \end{pmatrix}$, the vector $(1, 0)^\top$ is mapped to $(\mathtt{j}, 0)^\top$, and $(0, 1)^\top$ is mapped to $(0, 1)^\top$. When $A = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}$ as in (c), however, we see that $(1, 0)^\top$ is mapped to $(1/2, 1/2)^\top$, and $(0, 1)^\top$ is mapped to $(-1/2, 1/2)^\top$, hence the distances are shrunk by a factor of $\frac{1}{\sqrt{2}}$, which is equivalent to an increase in the noise intensity by 2.