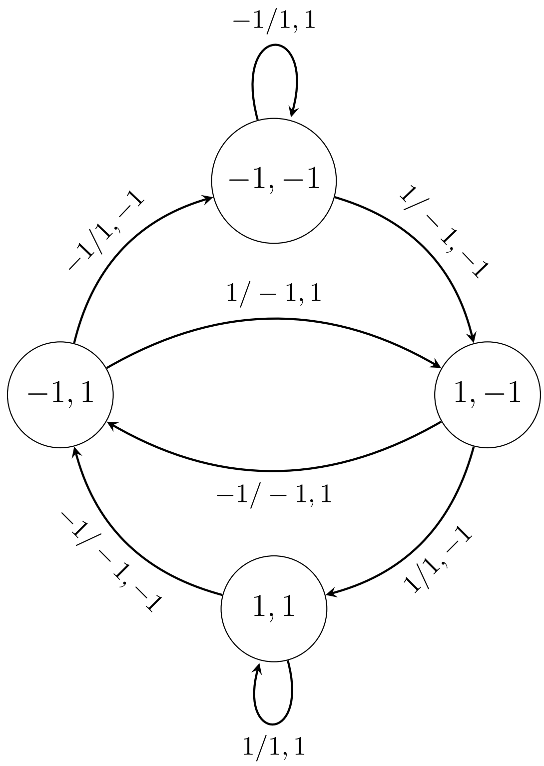
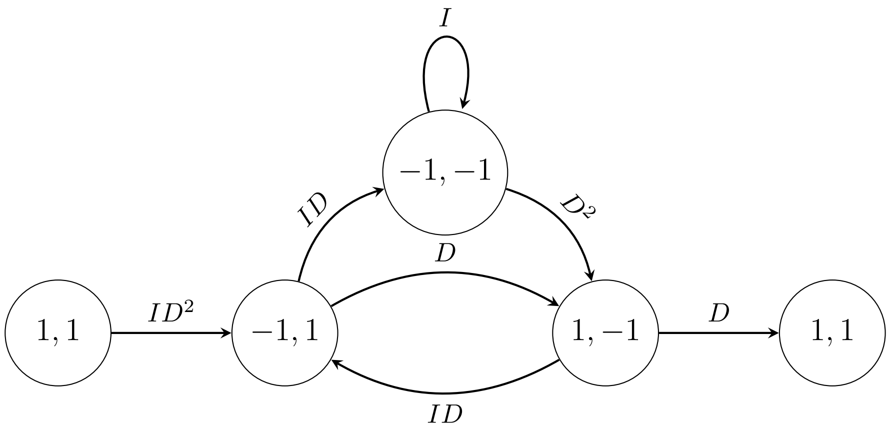
SOLUTION 1.

(a) The state diagram and detour flow graph are shown here. The states are labeled as (b_{j-1}, b_{j-2}) and the transitions with $b_j/x_{2j-1}, x_{2j}$.





(b) The output to $-1, -1, \dots, -1$ will be $-1, -1, 1, -1, 1, 1, \dots, 1$.

(c) Given the observation $y = (y_1, \ldots, y_{2n})$, the ML codeword is given by $\arg\max_{x\in\mathcal{C}} p(y|x)$ where $\mathcal{C}$ represents the set of codewords (i.e., the set of all possible paths on the trellis). Alternately, the ML codeword is given by

$$\arg\max_{x\in\mathcal{C}} \sum_{i=1}^{2n} \log p(y_i|x_i).$$

Hence, a branch metric for the BSC is

$$\log p(y_i|x_i) = \begin{cases} \log \epsilon & \text{if } y_i \neq x_i, \\ \log(1-\epsilon) & \text{if } y_i = x_i. \end{cases}$$

The decoder chooses the path with the largest metric.

(d) The channel output will be $1, 1, \ldots, 1$ (length $2n$ all-1 sequence). The decoder will clearly choose the path corresponding to the all-1 input sequence on the trellis and, hence, decode the maximum likelihood transmitted input sequence as $1, 1, \ldots, 1$ (length n all-1 sequence).

Solution 2.

(a) Firstly we have

$$\psi_{\mathcal{F}}(f) = e^{-\pi f^2}(1 - e^{-\mathrm{j}\pi f}) = 2\mathrm{j}e^{-\mathrm{j}\pi f/2}e^{-\pi f^2}\sin\left(\frac{\pi}{2}f\right).$$

Moreover, $\mathbb{E}[X_i] = 0$ and

$$K_X[k] = \mathbb{E}\left[X_{i+k}X_i^*\right] = \mathcal{E}\mathbb{1}\{k=0\}.$$

Thus,

$$S_X(f) = |\psi_{\mathcal{F}}(f)|^2 \sum_k K_X[k]e^{-\mathrm{j}2\pi kf} = \boxed{4\mathcal{E}e^{-2\pi f^2}\sin^2\left(\frac{\pi}{2}f\right)}.$$

Therefore, $S_X(f) = 0$ for $\forall f = 2m,\ m \in \mathbb{Z}$.

(b) It is easy to check that still $\mathbb{E}[X_i] = 0$ and

$$\begin{aligned} K_X[k] &= \mathbb{E}\left[X_{i+k}X_i^*\right] \\ &= s^2\left(\mathbb{E}[D_{i+k}D_i] + \alpha\mathbb{E}[D_{i+k}D_{i-2}] + \alpha\mathbb{E}[D_{i+k-2}D_i] + \alpha^2\mathbb{E}[D_{i+k-2}D_{i-2}]\right) \\ &= s^2\left((1+\alpha^2)\mathbb{1}\{k=0\} + \alpha\mathbb{1}\{k=-2\} + \alpha\mathbb{1}\{k=2\}\right). \end{aligned}$$

In particular,

$$K_X[0] = \mathbb{E}[X_i^2] = s^2(1+\alpha^2) = \mathcal{E} \quad \Longrightarrow \quad \boxed{s = \pm\sqrt{\frac{\mathcal{E}}{1+\alpha^2}}}.$$

Therefore,

$$\boxed{S_X(f) = 4\frac{\mathcal{E}}{1+\alpha^2}e^{-2\pi f^2}\sin^2\left(\frac{\pi}{2}f\right)\left((1+\alpha^2) + 2\alpha\cos(4\pi f)\right)}.$$

Finally, if $|\alpha| \neq 1$, since $(1+\alpha^2) + 2\alpha\cos(4\pi f)$ has no real zeros for f, $S_X(f) = 0$ for $\forall f = 2m,\ m \in \mathbb{Z}$. However, if $\alpha = 1$, solving $(1+\alpha^2) + 2\alpha\cos(4\pi f) = 0$ for f gives additional zeros at $f = \frac{2m+1}{4},\ m \in \mathbb{Z}$. Similarly if $\alpha = -1$ there will be additional nulls at frequencies $f = \frac{m}{2},\ m \in \mathbb{Z}$.

(c) Since $d, d' \in \{-1, +1\}$, we can express $f(d, d')$ as

$$f(d, d') = \frac{1}{2}(d - d').$$

Once again we have $\mathbb{E}[X_i] = 0$ and

$$\begin{aligned} K_X[k] &= \mathbb{E}\left[X_{i+k}X_i^*\right] \\ &= \frac{s^2}{4}\left(\mathbb{E}[D_{i+k}D_i] - \mathbb{E}[D_{i+k}D_{i-1}] - \mathbb{E}[D_{i+k-1}D_i] + \mathbb{E}[D_{i+k-1}D_{i-1}]\right) \\ &= \frac{s^2}{4}\left(2\mathbb{1}\{k=0\} - \mathbb{1}\{k=-1\} - \mathbb{1}\{k=1\}\right). \end{aligned}$$

Thus

$$K_X[0] = \frac{s^2}{2} = \mathcal{E} \quad \Longrightarrow \quad \boxed{s = \pm\sqrt{2\mathcal{E}}},$$

and

$$\boxed{S_X(f) = 4\mathcal{E}e^{-2\pi f^2}\sin^2\left(\frac{\pi}{2}f\right)\left(1 - \cos\left(2\pi f\right)\right)}.$$

(d) Using the precoder of (c) $S_X(f) = 0$ for $\forall f = m,\ m \in \mathbb{Z}$ (thus, in particular $S_X(1) = 0$). Using the precoding proposed in (b) we have

$$S_X(1) = 4\frac{\mathcal{E}}{1+\alpha^2}e^{-2\pi}\left(1+\alpha\right)^2 = 0 \quad \Longleftrightarrow \quad \boxed{\alpha = -1}.$$

Solution 3.

(a) Based on Nyquist's criterion we know that $\boxed{B \geq \frac{1}{2}}$.

(b) If $B = \frac{1}{2}$, in order for $\psi(t)$ to be unit-norm and orthogonal to its 1-shifts we must have $|\psi_{\mathcal{F}}(f)|^2 = \mathbb{1}\{-\frac{1}{2} \leq f \leq \frac{1}{2}\}$. Therefore,

$$\psi_{\mathcal{F}}(f) = e^{-\mathrm{j}2\pi f t_0}\mathbb{1}\{-\frac{1}{2} \leq f \leq \frac{1}{2}\} \quad \Leftrightarrow \quad \psi(t) = \operatorname{sinc}\left(t - t_0\right)$$

Finally solving for $\psi(0) = 0$ gives $t_0 = 0$. Thus $\boxed{\psi(t) = \operatorname{sinc}(t)}$.

(c) If $\theta = 0$,

$$\begin{aligned} \Re\{R_E(t)\} &= \Re\{w_E(t)\} + N_R(t), \\ \Im\{R_E(t)\} &= \Im\{w_E(t)\} + N_I(t), \end{aligned}$$

where $N_R(t)$ and $N_I(t)$ are independent white Gaussian noise processes of power spectral density $\frac{N_0}{2}$.

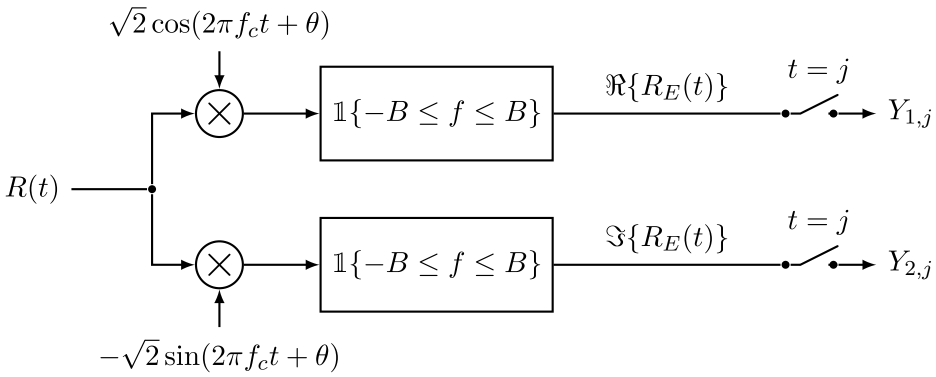
A sufficient statistic to estimate X_j from the received signal is obtained by computing the (complex-valued) inner products

$$Y_j = \langle R_E(t), \psi(t-j)\rangle,$$

or equivalently, pairs of real-valued inner products

$$Y_{1,j} = \langle \Re\{R_E(t)\}, \psi(t-j)\rangle \quad \text{and} \quad Y_{2,j} = \langle \Im\{R_E(t)\}, \psi(t-j)\rangle.$$

To this end, one in principle has to filter the outputs of the down-converter using matched filters of impulse response $\psi^*(-t)$ and sample the outputs of the filters at times $t = j, j \in \mathbb{Z}$. However, in this problem we see that a filter with impulse response $\psi^*(-t)$ is nothing but a low-pass filter with frequency response $\mathbb{1}\{-\frac{1}{2} \leq f \leq \frac{1}{2}\}$ which is already included in the down-converter. Thus, it is sufficient to sample the output of the down-converters directly to obtain the desired sufficient statistics.

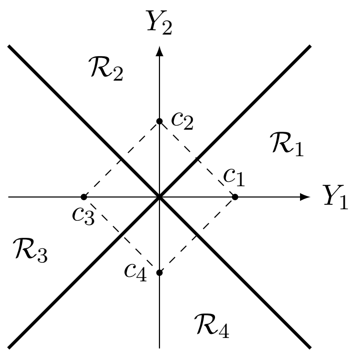


(d) We have the following hypothesis testing problem:

$$\text{under } H = i : \qquad Y = c_i + Z,$$

where $Z \sim \mathcal{N}(0, \frac{N_0}{2} I_2)$ and $c_1 = [1, 0]$, $c_2 = [0, 1]$, $c_3 = [-1, 0]$, and $c_4 = [0, -1]$.

For an AWGN setting, the ML decision rule will be the minimum distance decision rule with the following decision regions:



This is a 4-PSK constellation and the probability of error of an ML decoder for such a constellation is

$$\boxed{P_e = 2Q\left(\frac{1}{\sqrt{N_0}}\right) - Q\left(\frac{1}{\sqrt{N_0}}\right)^2}.$$

(e) Using the trigonometric identity $\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ we can see that the output of the top modulator, in presence of the phase difference, is

$$R(t)\cos(\theta) \times \sqrt{2}\cos(2\pi f_c t) - R(t)\sin(\theta) \times \sqrt{2}\sin(2\pi f_c t).$$

Thus, as the low-pass filter is a linear system, the output of the top low-pass filter is:

$$\Re\{R_E(t)\} = \Re\{w_E(t)\}\cos(\theta) + \Im\{w_E(t)\}\sin(\theta) + \cos(\theta)N_R(t) + \sin(\theta)N_I(t).$$

Similarly, we can show that the output of the bottom low-pass filter is:

$$\Im\{R_E(t)\} = \Im\{w_E(t)\}\cos(\theta) - \Re\{w_E(t)\}\sin(\theta) + \cos(\theta)N_I(t) - \sin(\theta)N_R(t).$$

Therefore, the observable $Y = [Y_1, Y_2]$ (under $H = i$) is now equal to

$$Y = R_\theta c_i + R_\theta Z$$

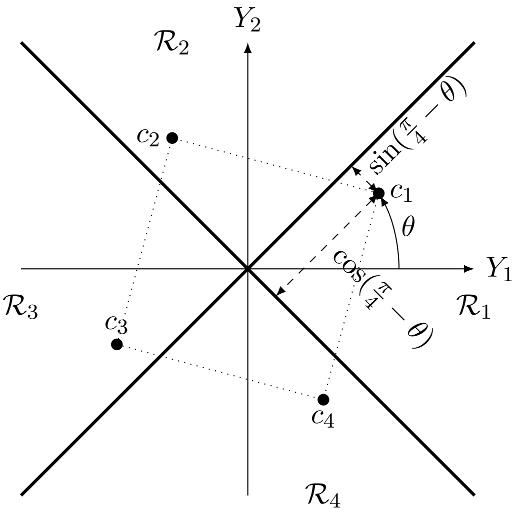
where

$$R_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

is the rotation matrix, the codewords c_i are as in part (d), and $Z \sim \mathcal{N}(0, \frac{N_0}{2}I_2)$. Moreover, we know that $V = R_\theta Z$ has the same statistics as Z. Thus, we can write the observable Y as

$$\text{under } H = i: \qquad Y = R_\theta c_i + V$$

with $V \sim \mathcal{N}(0, \frac{N_0}{2}I_2)$.



Using the above diagram, we can see that the probability of error of the receiver is

$$\boxed{P_e = Q\left(\frac{\sin(\frac{\pi}{4} - \theta)}{\sqrt{N_0/2}}\right) + Q\left(\frac{\cos(\frac{\pi}{4} - \theta)}{\sqrt{N_0/2}}\right) - Q\left(\frac{\sin(\frac{\pi}{4} - \theta)}{\sqrt{N_0/2}}\right) Q\left(\frac{\cos(\frac{\pi}{4} - \theta)}{\sqrt{N_0/2}}\right)}.$$

(f) If $|\theta| > \frac{\pi}{4}$, the constellation will be rotated in such a way that each codeword will be moved out of its corresponding decision region (e.g. c_1 will be moved to the decision region $\mathcal{R}_2$, c_2 to $\mathcal{R}_3$, ...). Therefore, in the absence of noise the decoder always decodes the sent codeword incorrectly (error probability is 1). As the noise variance increases the error probability decreases (there is a higher chance for the noise to move the observable Y into the correct decision region). In particular if the noise variance goes to infinity the observable Y will be a point chosen on the $\mathbb{R}^2$ plane uniformly at random. Thus, with probability $\frac{1}{4}$ it will be in the decoding region corresponding to the transmitted codeword which means the error probability will be decreased to $\frac{3}{4}$.