3F4 Data Transmission 2024
Crib

1. (a) The encoder has rate $R = 1/2$ because it emits 2 code digits for every input digit.

(b) The octal description is $(6, 5)_8$.

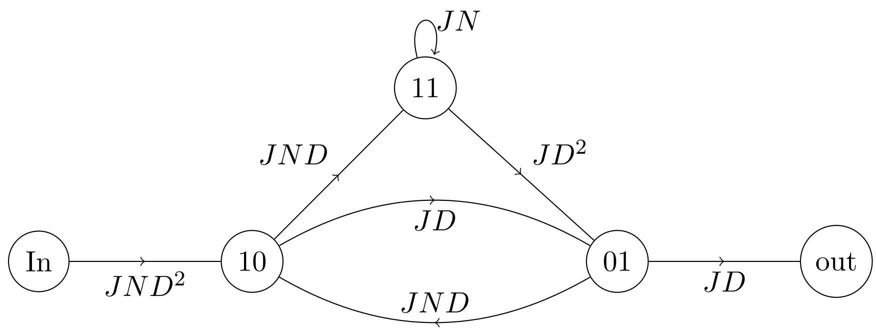
(c) Statement (i) is compatible: it describes another input sequence with the same probability as the decoded sequence. There can be ties in the Viterbi algorithm.

Statement (ii) is incompatible: if sequence $0, 0, 0, 1$ has a higher a-posteriori probability than the sequence $1, 1, 0, 1$, the Viterbi algorithm would have picked it.

Statement (iii) is compatible: although the winning sequence in the Viterbi algorithm starts with a 1, it is possible that the probability of a zero at the start of the sequence is larger than the probability of a one. The Viterbi algorithm is not bit error optimal.

(d) The encoder polynomials are $1 + D$ and $1 + D^2$. In binary, $(1 + D)^2 = 1 + (1 + 1)D + D^2 = 1 + D^2$ hence the greatest common divisor of the two polynomials $1 + D$ is not of the form D^ℓ and the encoder is catastrophic as per Massey and Sain's theorem.

(e) We draw the state diagram with the zero state split into an "input" and "output" state, to help us count the detours from the all zero code sequence:



We then compute the transfer function either using Mason's gain formula (which can be done by inspection) or by setting up a system of equations and resolving it (which requires a long algebraic derivation) to obtain

$$
\begin{aligned}
T(J, N, D) &= \frac{J^3ND^4(1 - JN) + J^4N^2D^6}{1 - JN - J^2ND^2 - J^3N^2D^4 + J^3N^2D^2} \\
&= J^3ND^4(1 - JN + JND^2)\left(1 + JN + J^2ND^2 + J^3N^2D^4 - J^3N^2D^2 + O(J^2N^2D^0)\right)
\end{aligned}
$$

The lowest degree term of this equation in terms of D is D^4, hence $d_{\text{free}} = 4$. One can easily verify that the input sequence $1, 1, 0$ gives the interleaved output sequence $1, 1, 1, 0, 0, 1$ (i.e., $X^{(1)} = 1, 1, 0$ and $X^{(2)} = 1, 0, 1$), which has weight 4.

(f) A systematic encoder can never be catastrophic. Catastrophic encoders map an infinite weight input sequence to a finite weight code sequence. Since half of the outputs bits at the output of this encoder are the input bits, any infinite weight input sequence would automatically result in an infinite weight output sequence.

(g) We observe in the diagram that $X^{(1)}(D) = U(D)$ and

$$\begin{cases} V(D) & = U(D) + DV(D) \\ X^{(2)}(D) = V(D) + D^2V(D) \end{cases}$$

and hence, carefully manipulating these equations (remember that $1+1=0$), we get

$$X^{(2)}(D) = (1+D^2)V(D) = \frac{1+D^2}{1+D}U(D) = (1+D)U(D).$$

Despite the appearance of being a 4 state encoder, this encoder in fact is a two state encoder where the sequence $X^{(2)}$ consists of the sum of two consecutive inputs, i.e.,

$$X_k^{(2)} = U_k + U_{k-1} = X_k^{(1)} + X_{k-1}^{(1)}.$$

(h) We note that the code sequences of the two encoders are the same because both encoders must satisfy $X_k^{(2)} = X_k^{(1)} + X_{k-1}^{(1)}$ (with initial condition $X_{-1}^{(1)} = 0$) based on our response to the previous question.

The code sequences in (i) satisfy this constraints and are hence valid output sequences for both encoders.

The code sequences in (ii) are valid output sequences for the encoder in Fig. 1(a). Despite what we said above about code sequences being the same for both encoders, this is not a valid output for the encoder in Fig. 1(b) because of the zero termination bits which should have resulted in $X_6^{(1)} = X_7^{(1)} = 0$, which is not the case for this example.

The sequences if (iii) are not valid output sequences for any of the two encoders because, for example, $X_1^{(2)} = 0 \neq X_1^{(1)} + X_0^{(1)} = 1$.

2. (a) The signal set is orthogonal and of dimension $K = 4$. An orthonormal basis is simply the same signal set, divided by $\frac{2}{A\sqrt{T}}$. Thus, the amplitude of each of the pulses of the orthonormal basis becomes $\frac{2}{\sqrt{T}}$.

(b) The signal set has four symbols, each labelled with 2 bits. Therefore, the rate is $\frac{2}{T}$ bits/s.

(c) We know that $E_s = 2E_b$ and that each symbol has the same energy $E_s = A^2$. Thus,

$$\boldsymbol{s}_1 = (\sqrt{2E_b}, 0, 0, 0), \boldsymbol{s}_2 = (0, \sqrt{2E_b}, 0, 0), \boldsymbol{s}_3 = (0, 0, \sqrt{2E_b}, 0), \boldsymbol{s}_4 = (0, 0, 0, \sqrt{2E_b}) \qquad (1)$$

(d) The distance between each pair is the same, hence the minimum distance d_{min} is

$$d_{\text{min}} = 4E_b. \qquad (2)$$

(e) The optimal receiver projects the received $y(t)$ signal onto each of the elements of the orthonormal basis and then performs detection based on those projections $y_1, \ldots, y_K$. Detection is based on the MAP rule (since we have non-equiprobable symbols). The output of the detector is the symbol $\boldsymbol{x} \in \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_M\}$ such that

$$\begin{aligned} \hat{\boldsymbol{x}} &= \arg\max_{i \in \{1,\ldots,M\}} P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{s}_i) P_{\boldsymbol{X}}(\boldsymbol{s}_i) && (3) \\ &= \arg\max_{i \in \{1,\ldots,M\}} \frac{1}{(\sqrt{\pi N_0})^M} e^{-\frac{y_1^2+\cdots+y_M^2}{N_0}} e^{-\frac{E_s}{N_0}} e^{\frac{2\sqrt{E_s} y_i}{N_0}} \cdot P_{\boldsymbol{X}}(\boldsymbol{s}_i) && (4) \\ &= \arg\max_{i \in \{1,\ldots,M\}} \frac{2\sqrt{E_s}}{N_0} y_i + \log P_{\boldsymbol{X}}(\boldsymbol{s}_i) && (5) \end{aligned}$$

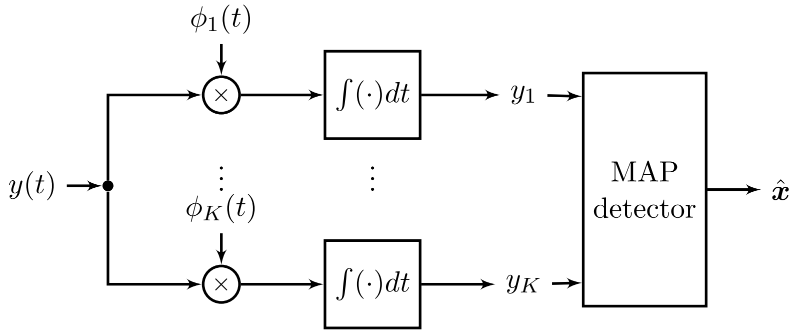


Figure 1: Bank of cross-correlators.

(f) Let the AWGN be $n(t)$ and $n'(t) = n(t) - \sum_{k=1}^{K} n_k\phi_k(t)$ be the noise outside the signal space, where n_k is the noise projection onto the k-th element of the orthonormal basis of the signal space ϕ_k. The property states that $n'(t)$ and $y_1, \ldots, y_K$ are independent, which can be shown by showing that $\mathbb{E}[n'(t)n_k]$ are uncorrelated, and since they are Gaussian, they are independent.

(g) Assuming $\boldsymbol{s}_1$ was transmitted, an error occurs if y_1 is not the largest entry in vector $(y_1, y_2, \ldots, y_M)$. Therefore,

$$\begin{aligned} p_e &= \mathbb{P}\big[\{Y_1 \le Y_2\} \cup \{Y_1 \le Y_3\} \cup \cdots \cup \{Y_1 \le Y_M\}\big] && (6)\\ &= \mathbb{P}\big[\{\sqrt{E}_\text{s} + N_1 \le N_2\} \cup \{\sqrt{E}_\text{s} N_1 \le N_3\} \cup \cdots \cup \{\sqrt{E}_\text{s} + N_1 \le N_M\}\big] && (7)\\ &\le \mathbb{P}\big[\sqrt{E}_\text{s} + N_1 \le N_2\big] + \cdots + \mathbb{P}\big[\sqrt{E}_\text{s} + N_1 \le N_M\big] && (8)\\ &= (M-1)\mathbb{P}\big[N_2 - N_1 \ge \sqrt{E}_\text{s}\big] && (9) \end{aligned}$$

(h) Since $N_2 - N_1$ is a Gaussian random variable with zero mean and variance N_0, the probability term in (9) is a Gaussian tail function. We thus have that

$$p_e \le (M-1)Q\left(\sqrt{\frac{E_\text{s}}{N_0}}\right). \tag{10}$$

3. (a) The dimension of the signal space is $K = 2$.

(b) The points are $Ae^{j\phi}, Ae^{j(\pi-\phi)}, Ae^{j(\pi+\phi)}, Ae^{j(2\pi-\phi)}$. Thus, the vector representation is

$$(A\cos\theta, A\sin\theta), (-A\cos\theta, A\sin\theta), (-A\cos\theta, -A\sin\theta), (A\cos\theta, -A\sin\theta)$$

The average energy per bit is $E_b = \frac{A^2}{2}$, as every point has the same energy.

(c) The optimal decision region are the $x-y$ axes themselves.

(d) i. The received SNR is $\frac{2|h|^2A^2}{N_0}$.

ii. The error probability is

$$p_e = p_{e,r} + p_{e,i} - p_{e,r}p_{e,i}$$

where $p_{e,r}, p_{e,i}$ are the error probabilities for the real/imaginary axes, respectively. Thus if we let $n = n_r + jn_i$

$$p_{e,r} = \mathbb{P}[hA\cos\theta + n_r < 0] = Q\left(\sqrt{\frac{2h^2A^2\cos^2\theta}{N_0}}\right) = Q\left(\sqrt{4h^2\cos^2\theta\frac{E_b}{N_0}}\right) \tag{11}$$

$$p_{e,i} = \mathbb{P}[hA\sin\theta + n_i < 0] = Q\left(\sqrt{\frac{2h^2A^2\sin^2\theta}{N_0}}\right) = Q\left(\sqrt{4h^2\sin^2\theta\frac{E_b}{N_0}}\right) \tag{12}$$

Finally

$$p_e = Q\left(\sqrt{4h^2\cos^2\theta\frac{E_b}{N_0}}\right) + Q\left(\sqrt{4h^2\sin^2\theta\frac{E_b}{N_0}}\right) - Q\left(\sqrt{4h^2\cos^2\theta\frac{E_b}{N_0}}\right)Q\left(\sqrt{4h^2\sin^2\theta\frac{E_b}{N_0}}\right)$$

iii. At high SNR we have that the error probability at receiver 1 is approximated by the estimate provided by the union bound, which is in turn dominated by the error probability in the imaginary component since $\sin\theta < \cos\theta$ for $\theta < \frac{\pi}{4}$

$$p_{e,1} \approx Q\left(\sqrt{4h_1^2\cos^2\theta\frac{E_b}{N_0}}\right) + Q\left(\sqrt{4h_1^2\sin^2\theta\frac{E_b}{N_0}}\right) \approx Q\left(\sqrt{4h_1^2\sin^2\theta\frac{E_b}{N_0}}\right) \quad \text{at high SNR} \tag{13}$$

$$p_{e,2} = Q\left(\sqrt{4h_2^2\cos^2\theta\frac{E_b}{N_0}}\right) \tag{14}$$

We find θ by equating the terms inside the Q-functions,

$$4h_2^2\cos^2\theta = 4h_1^2\sin^2\theta \quad \Longrightarrow \quad \theta = \arctan\frac{h_2}{h_1} = \arctan 0.25 = 14 \text{ deg.}$$

4. (a) The error probability can be expressed as

$$p_e = \frac{1}{M}\sum_{m=1}^{M}\mathbb{P}\left[\bigcup_{m'\neq m}\{\hat{\boldsymbol{x}}(\boldsymbol{Y}) = \boldsymbol{s}_{m'}\}\middle|\boldsymbol{X} = \boldsymbol{s}_m\right] \tag{15}$$

$$\leq \frac{1}{M}\sum_{m=1}^{M}\sum_{m'\neq m}\mathbb{P}[\hat{\boldsymbol{x}}(\boldsymbol{Y}) = \boldsymbol{s}_{m'}|\boldsymbol{X} = \boldsymbol{s}_m] \tag{16}$$

where the last line follows from the union bound and where $\mathbb{P}[\hat{\boldsymbol{x}}(\boldsymbol{Y}) = \boldsymbol{s}_{m'}|\boldsymbol{X} = \boldsymbol{s}_m] = P(\boldsymbol{s}_m \to \boldsymbol{s}_{m'})$ is the probability of deciding in favour of $\boldsymbol{s}_{m'}$ when $\boldsymbol{s}_m$ was sent.

(b) We have that

$$\mathbb{P}[\hat{\boldsymbol{x}}(\boldsymbol{Y}) = \boldsymbol{s}_j|\boldsymbol{X} = \boldsymbol{s}_i] = \mathbb{P}[\|\boldsymbol{Y} - \boldsymbol{s}_j\|^2 < \|\boldsymbol{Y} - \boldsymbol{s}_i\|^2|\boldsymbol{X} = \boldsymbol{s}_i] \tag{17}$$

$$= \mathbb{P}[\|\boldsymbol{s}_i - \boldsymbol{s}_j + \boldsymbol{N}\|^2 \leq \|\boldsymbol{N}\|^2|\boldsymbol{X} = \boldsymbol{s}_i] \tag{18}$$

$$= \mathbb{P}\left[Z_{i,j} \leq -\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2|\boldsymbol{X} = \boldsymbol{s}_i\right] \tag{19}$$

where

$$Z_{i,j} = 2\sum_{k=1}^{K}(s_{i,k} - s_{j,k})N_k.$$

Since $Z_{i,j}$ is a weighted sum of independent Gaussian random variables, it is Gaussian. We proceed to calculate the mean and variance. The mean is

$$\mathbb{E}[Z_{i,j}] = 2\sum_{k=1}^{K}(s_{i,k} - s_{j,k})\mathbb{E}[N_k] = 0. \tag{20}$$

The variance is equal to

$$\mathbb{E}[Z_{i,j}^2] = \mathbb{E}\left[2\sum_{k=1}^{K}(s_{i,k} - s_{j,k})N_k \cdot 2\sum_{\ell=1}^{K}(s_{i,\ell} - s_{j,\ell})N_\ell\right] \tag{21}$$

$$= 4\sum_{k=1}^{K}\sum_{\ell=1}^{K}(s_{i,k} - s_{j,k})(s_{i,\ell} - s_{j,\ell})\mathbb{E}[N_k N_\ell] \tag{22}$$

$$= 2N_0\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2 \tag{23}$$

Thus the pairwise error probability is

$$\mathbb{P}\left[Z_{i,j} \leq -\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2 | \boldsymbol{X} = \boldsymbol{s}_i\right] = \mathbb{P}\left[\frac{Z_{i,j}}{\sqrt{2N_0\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}} \leq -\frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{\sqrt{2N_0\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}}\right] \tag{24}$$

$$= Q\left(\sqrt{\frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{2N_0}}\right). \tag{25}$$

(c) At SNR, since the Q-function decays exponentially and the terms in sum in (16) only depend on the distance between pairs of constellation points, the overall sum in (16) is dominated by the terms at minimum distance.

(d) In distance vector protocols nodes broadcast shortest paths to destinations to their neighbours. For example RIP, which uses Distributed Bellman-Ford. In link-state protocols the nodes broadcast distances to their neighbours to whole network, e.g., OSPF, which uses Dijkstra.

In link-state protocols each node is able to build a model of the entire network and use the very efficient Dijksta algorithm to find shortest paths. They require the ability for each node to broadcast information to every other node in the network, which implies a significant communication overhead. Even though Dijkstra is efficient the fact that it has to be run at all nodes is a computational burden. The advantage is that correct routes are calculated immediately; when the link between C and D fails C broadcasts to A,B and D that it no longer has a direct connection to D. All nodes then run Dijkstra independently and arrive a correct set of shorted paths. The advantage of distance vector protocols is that the communication only has to be local. Each node maintains a table of shortest paths to each other node in the network, which it shares with its neighbours. Convergence can be slow however, as in the example. Before the failure A has routes of 4 to D, 20 to B and 2 to C. When the link C-¿D fails C recalculates it's routes, giving a route of 6 to D, via 8, which it advertises to A. A now calculates its shortest path to D to be 8, via C and then C calculates a shortes path of 10 to D, via A etc This continues for a few cycles until A calculates a route of 20 to D via C at which point the path of 19 via B is preferred. Until this point packets are sent back and forth between A and C.