FINAL - SOLUTIONS

Wednesday, 26th March, 2014, 11:30-14:30

This exam has 6 problems worth 100 points in total plus 5 bonus points.

NAME: ____

UID: ______________________________

PAGES: _______________

Problem	Max.	Awarded
1	24	
2	10	
3	14+5	
4	19	
5	17	
6	16	
TOTAL	100+5	

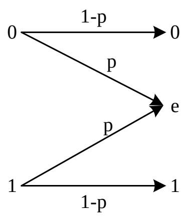
Instructions

- You are allowed to use **<u>2 sheets</u>** of paper for reference. No mobile phones or calculators are allowed in the exam.
- You can attempt the problems in any order as long as it is clear as to which problem is being attempted and which solution to the problem you want us to grade.
- If you are stuck in any part of a problem do not dwell on it, try to move on and attempt it later.
- Write on this sheet in the space provided, the **total number of pages** submitted. The pages in this booklet are numbered already; number the extra pages that you attach. Strike off unused pages.
- The number of pages (including this page) in this booklet is <u>23</u>. Make sure that your booklet has all the pages printed clearly before starting the exam.

Good Luck! Go Bruins!

Some **relations/definitions**:

- P_X denotes the probability mass function of a discrete random variable X whose realizations are denoted x. Bold symbols denote the vector analogs.
- $\forall$ denotes "for all"; $\in$ denotes "element(s) of" (set membership); $\cup$ denotes "union"; $\cap$ denotes "intersection".
- In an n-dimensional space, the n unit vectors are represented as $\mathbf{e}_i, i = 1, 2, \dots, n$ which means for $j = 1, 2, \dots, n$,
$$e_{i,j} = \begin{cases} 1 & i = j, \\ 0 & \text{else.} \end{cases}$$
- The energy of a vector representation $\boldsymbol{x}$ of a signal $x(t)$ can be calculated as $E = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|^2$ where $\|\cdot\|$ denotes the norm or length of a vector.
- The Euclidean distance between two constellation points $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ is $d = \|\boldsymbol{x} - \boldsymbol{y}\|_2 = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$.
- The average energy $\mathcal{E}$ of a constellation of vectors $\boldsymbol{s}_i, i = 0, 1, \dots, M-1$ is given as $\mathcal{E} = \sum_{i=0}^{M-1} P_{\boldsymbol{s}}(\boldsymbol{s}_i)\|\boldsymbol{s}_i\|^2$.
- $Q(a) \triangleq \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\,\mathrm{d}x$, the tail probability of a standard normal distribution which is $\sim \eta(0,1)$. $\sim$ denotes "distributed as".
- The binary erasure channel (BEC) is an abstract channel model which is described as follows. The input is either received correctly or it becomes invalid as an "erasure".



Some **numerical, logarithmic, Q-function, trigonometric** values:

$2 \times 5.375 = 10.75$	$Q(0) = 0.5$	$\sin\left(\frac{\pi}{2}\right) = 1$	$\cos\left(\frac{\pi}{2}\right) = 0$
	$Q(\infty) = 0$	$\sin^2\left(\frac{\pi}{4}\right) = 0.5$	
$\log_{10}\left(\frac{1}{3}\right) = -0.4771$	$Q(-\infty) = 1$	$\sin\left(\frac{\pi}{3}\right) = 0.8660$	$\cos\left(\frac{\pi}{3}\right) = 0.5$
$\log_{10}\left(\frac{1}{2}\right) = -0.3010$	$Q(x) \approx 0, x > 6$	$\sin\left(\frac{2\pi}{3}\right) = 0.8660$	$\cos\left(\frac{2\pi}{3}\right) = -0.5$
$\log_{10}\left(\frac{2}{3}\right) = -0.1761$	$Q(x) \approx 0.5, x \approx 0$	$\sin(\pi + \theta) = -\sin\theta$	$\cos(\pi + \theta) = -\cos\theta$
$\frac{\log\left(\frac{3}{2}\right)}{\log(3)} = 0.3691$	$Q(\sqrt{15}) = 5.375 \times 10^{-5}$	$\sin\left(\frac{\pi}{8}\right) = 0.3827$	$\cos\left(\frac{\pi}{8}\right) = 0.9239$
$\log_{10}(107.5) = 2.0314$	$Q(\sqrt{7.5}) = 0.0031$	$\sin\left(\frac{2\pi}{8}\right) = 0.7071$	$\cos\left(\frac{2\pi}{8}\right) = 0.7071$
$\log_{10}(6200) = 3.7924$	$Q(x) = 1 - Q(-x), \forall x$	$\sin\left(\frac{3\pi}{8}\right) = 0.9239$	$\cos\left(\frac{3\pi}{8}\right) = 0.3827$

Some **identities**:

$\log(a \cdot b) = \log a + \log b$	$2\sin\theta_1\cos\theta_2 = \sin(\theta_1 + \theta_2) + \sin(\theta_1 - \theta_2)$	$e^{j\theta} = \cos\theta + j\sin\theta$
$\log(a/b) = \log a - \log b$	$2\cos\theta_1\cos\theta_2 = \cos(\theta_1 + \theta_2) + \cos(\theta_1 - \theta_2)$	
$k \times \log(a \cdot b) = \log(a \cdot b)^k$	$2\sin\theta_1\sin\theta_2 = \cos(\theta_1 - \theta_2) - \cos(\theta_1 + \theta_2)$	
$\log\left(\frac{1}{a}\right) = -\log a$	$\sin^2(\theta) = \frac{1-\cos(2\theta)}{2}$	
	$\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}$	

Problem 1 (OVERVIEW (24 pts))

In the following, clearly indicate, in your answer, your final choice along with the accompanying justifications/mathematical treatment for a [YES/NO] type question. Also for all the questions, give a clear and supporting mathematical treatment accompanying your answer.

(a) *(Noise-less and noisy observations)* Consider a system which transmits M messages $H_i, i = 0, 1, \ldots, M-1$ as unit vectors along M dimensions i.e. the signaling set, in vector representation, is *[4pts]*

$$\mathbf{s}_i = \mathbf{e}_{i+1}, i = 0, 1, \ldots, M-1.$$

Assume that M is even and denote the transmitted message as

$$\mathbf{X} = \mathbf{s}_i, \text{ if } H = i.$$

A vector noisy channel affects each component of the transmitted vector $\mathbf{X}$ independently; with additive Gaussian noise $N \sim \eta(0, \sigma^2)$ only on the odd numbered dimensions of the vector. The even numbered dimensions are received unaffected. In other words, if $\mathbf{Y}$ is the received message vector, then

$$Y_j = \begin{cases} X_j + N_j & j = 1, 3, 5, \ldots, M-1. \\ X_j & j = 2, 4, 6, \ldots, M. \end{cases} \tag{1}$$

Assuming that the messages at the transmitter are equally likely, design the optimal decoder $d(\mathbf{Y})$ which minimizes the probability of error for this system.

$\rightarrow$ (Solution) Note that the even-numbered dimensions are received unaffected; hence any $\mathbf{y}$ with '1' in one of the dimensions $j : j = 2, 4, \ldots, M$ should be decoded as follows:

$$d(\mathbf{y}) = \mathbf{s}_{j-1} \text{ if } y_j = 1 \text{ for exactly one of } j = 2, 4, \ldots, M. \tag{2}$$

In the case that all the even-numbered dimensions of the received vector $\mathbf{y}$ have a 0 in them, we apply the MAP rule which reduces to the ML rule since the prior probabilities of signals are equal. The decoder can be derived/written as follows. Denote the conditional probability density function of $\mathbf{y}$ conditioned on hypothesis i as $f_{\mathbf{y}|H}(\mathbf{y} \mid i)$.

$$\begin{aligned}
y_j = 0, \forall j = 2, 4, \ldots, M \implies d(\mathbf{y}) &= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\max} \; f_{\mathbf{y}|H}(\mathbf{y} \mid i) && (3) \\
&= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\max} \left[\frac{1}{\sqrt{2\pi\sigma^2}^M} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{M}(y_k - e_{i,k})^2} \right] \\
&= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\max} \; e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{M}(y_k - e_{i,k})^2} \\
&= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\min} \sum_{k=1}^{M}(y_k - e_{i,k})^2 \\
&= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\min} \; -2y_i \\
y_j = 0, \forall j = 2, 4, \ldots, M \implies d(\mathbf{y}) &= \mathbf{s}_{j-1} : j = \underset{i=1,3,\ldots,M-1}{\arg\max} \; y_i. && (4)
\end{aligned}$$

Note that the above is the MAP decoding rule for orthogonal signaling using M-dimensional unit vectors with an exception of presence of noise only in the odd-numbered dimensions.

(b) *(Waveform channel design)* Consider a set of 2 "antipodal" waveforms $s_1(t) = -s_0(t)$ used in a binary signaling scheme. Assuming that the channel adds AWGN noise $n(t)$, the optimal waveform receiver can be implemented using just one correlator and a threshold comparator. [YES/NO] Sketch the optimum receiver structure for your answer if $y(t) = s_i(t) + n(t), i = 0$ or 1 denotes the received signal. *Note:* A complete mathematical derivation is not necessary. *[5pts]*

$\rightarrow$ (Solution) YES. The signaling set is specified to be $s_0(t) = -s_1(t)$. The number of basis functions required to describe this signaling set is just 1 since the 2 signals are linearly dependent on each other. In particular, if we denote the energy of these 2 signals as E, then the basis function required to describe this binary signaling set is

$$\psi_0(t) = \frac{s_0(t)}{\sqrt{E}}. \tag{5}$$

With respect to this basis function, the vector representation of the two signals are:

$$\mathbf{s}_0 = \sqrt{E}, \quad \mathbf{s}_1 = -\sqrt{E}. \tag{6}$$

On an AWGN channel, with prior probabilities of these signals as $P_H(0), 1 - P_H(0)$, where $P_H(0)$ denotes the prior probability of signal $s_0(t)$, the optimum receiver will consist only of one correlator (which can also be implemented as a matched filter) along with a threshold comparator which compares the result of correlation against a threshold 'tr' which is a function of $P_H(0)$. We sketch the receiver as shown in Fig. 1.

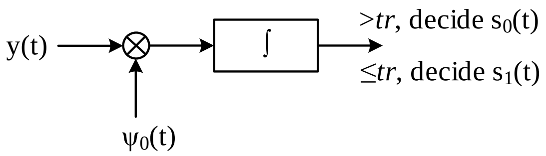


Figure 1: Optimum receiver for antipodal signaling

(c) *(Optimum receiver)* Suppose one of M equiprobable signals $x_i(t), i = 0, 1, \ldots, M-1$ is transmitted during a period of time T over an AWGN channel. Each signal is identical to all others in a sub-interval $[t_1, t_2]$, where $0 < t_1 < t_2 < T$. The optimum receiver can ignore the sub-interval $[t_1, t_2]$. [YES/NO] Support your answer with proper justification. *[4pts]*

$\rightarrow$ (Solution) YES. Denote the vector representation of the signals as $\mathbf{x}_i$ and also denote the received signal as $r(t)$ with its vector representation as $\mathbf{r}$. The optimum receiver over an AWGN channel for equiprobable signaling is the minimum distance decoder (MAP $\Longrightarrow$ ML $\Longleftrightarrow$ Nearest neighbour decoding) and the solution follows as:

$$\hat{H} = \arg\min_i \|\mathbf{r} - \mathbf{x}_i\| \tag{7}$$

$$\Longleftrightarrow \arg\max_i 2\mathbf{r}^T\mathbf{x}_i + \mathbf{x}_i^T\mathbf{x}_i = \arg\max_i \int_0^T (2r(t) + x_i(t))x_i(t)\mathrm{d}t$$

$$\Longleftrightarrow \arg\max_i \int_0^{t_1} (2r(t) + x_i(t))x_i(t)\mathrm{d}t + \int_{t_1}^{t_2} (2r(t) + x_i(t))x_i(t)\mathrm{d}t + \int_{t_2}^{T} (2r(t) + x_i(t))x_i(t)\mathrm{d}t \tag{8}$$

$$\Longleftrightarrow \arg\max_i \int_0^{t_1} (2r(t) + x_i(t))x_i(t)\mathrm{d}t + \int_{t_2}^{T} (2r(t) + x_i(t))x_i(t)\mathrm{d}t, \tag{9}$$

since the value of the integral $\int_{t_1}^{t_2} [\cdot]\ \mathrm{d}t$ in (8) evaluates to the same for every i.

(d) *(Linear block code over BEC)* Consider the following parity-check matrix of a linear block code: *[5pts]*

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

The decoder for the code, on account of having been deployed to decode over a Binary erasure channel (BEC), received the following word:

$$\mathbf{y} = [e\ e\ 0\ e\ 0\ 1\ 0] \tag{11}$$

where e denotes an erasure. The decoder can decode this received word to a valid codeword. [YES/NO] If yes, what is the decoded codeword? If no, demonstrate why not.

$\rightarrow$ (Solution) YES. The requirement for $\mathbf{y}$ to be a valid codeword is that it should satisfy the "parity-check" equation as follows:

$$\mathbf{y}\mathbf{H}^T = \mathbf{0}. \tag{12}$$

By using the above, we write down the 3 equations that we obtain as:

$$y_0 \oplus y_1 = 0, \quad y_0 \oplus y_1 \oplus y_3 \oplus 1 = 0, \quad y_0 \oplus y_3 = 0. \tag{13}$$

For the above, there is a unique solution since the three equations are linearly independent of each other. Namely, the solution is $y_0 = y_1 = y_3 = 1$. Hence the received word can be decoded (uniquely) to the valid codeword

$$\mathbf{c} = [1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 1\ \ 0]\,. \tag{14}$$

(e) *(OFDM - Time-varying channel)* Consider an OFDM system with $N_c = 3$ sub-carriers implemented for the dispersive channel (ISI channel) given by the following z-transform: *[6pts]*

$$H(z) = 1.33 + 0.33z^{-1} + 0.33z^{-2}. \tag{15}$$

Assume that there is AWGN noise added at the receiver. The transmitter employs a certain transmission strategy and uses all the sub-carriers available to it. Denote the probability of symbol error with this transmit scheme as P_{e}. After some time, the channel condition changes as follows:

$$H_{\text{new}}(z) = 1.67 + 0.17z^{-1} + 0.17z^{-2}. \tag{16}$$

The transmitter continues to employ the same transmit strategy. Assuming that the receiver knows the new channel coefficients, the probability of symbol error for the present channel satisfies $P_{\text{new}} < P_{\text{e}}$. [YES/NO] Support your answer with proper justification.

→ (Solution) YES. In the following, we use "channel" and "sub-carrier" interchangeably. Consider the N_c-point DFT of the original channel $h_l = [1.33 \;\; 0.33 \;\; 0.33]$ which is given as

$$\hat{H} = [2 \;\; 1 \;\; 1]. \tag{17}$$

The above can be computed mathematically using the values of $\sin\left(\frac{\pi}{3}\right)$, $\sin\left(\frac{2\pi}{3}\right)$, $\cos\left(\frac{\pi}{3}\right)$ and $\cos\left(\frac{2\pi}{3}\right)$ given in the second page. For OFDM, the parallel AWGN channel representation that we obtain is given as

$$\hat{y}[m] = \hat{H}[m]\hat{x}[m] + \hat{w}[m], \quad m = 1, 2, 3. \tag{18}$$

where $\hat{x}[m]$ are the transmitted data symbols (before IDFT at the transmitter).

First we note that the DFT of the channel coefficients is a vector with non-zero components and hence the transmitter can signal over all the 3 available sub-carriers. Denote the average energy of the constellations used by the transmitter as E_1, E_2, E_3. The problem says that the transmitter uses all of its sub-carriers and hence we make this assumption of energies. The signal-to-noise ratio (SNR) of the individual sub-carriers are given as (assuming, without loss of generality that, $\hat{w}$ is AWGN with statistic $(0, \sigma^2)$ over each channel m)

$$\text{SNR}_1 = \frac{|\hat{H}[1]|^2 E_1}{\sigma^2}, \quad \text{SNR}_2 = \frac{|\hat{H}[2]|^2 E_2}{\sigma^2}, \quad \text{SNR}_3 = \frac{|\hat{H}[3]|^2 E_3}{\sigma^2} \tag{19}$$

$$\text{SNR}_1 = \frac{4 \cdot E_1}{\sigma^2}, \quad \text{SNR}_2 = \frac{1 \cdot E_2}{\sigma^2}, \quad \text{SNR}_3 = \frac{1 \cdot E_3}{\sigma^2}. \tag{20}$$

For the new channel, we similarly find the DFT coefficients to be

$$\hat{H}_{new} = [2 \;\; 1.5 \;\; 1.5], \tag{21}$$

for which the corresponding SNRs are

$$\text{SNR}_{\text{new},1} = \frac{4 \cdot E_1}{\sigma^2}, \quad \text{SNR}_{\text{new},2} = \frac{2.25 \cdot E_2}{\sigma^2}, \quad \text{SNR}_{\text{new},3} = \frac{2.25 \cdot E_3}{\sigma^2}. \tag{22}$$

Since the signal-to-noise ratios of sub-carriers $2, 3$ are better for the new channel compared to the old, the probability of error should satisfy $P_{new} < P_e$.

Problem 2 (TETRAHEDRAL DIE **(10 pts)**)

Two four-sided dice, each labeled with $0, 1, 2, 3$ are in a box. For such a die, once we roll it, we agree to observe the number on the bottom face!

- One die is fair (equiprobable outcomes on a roll). Denote this die f.
- The other die is loaded so that 0 is observed with probability $1/2$, the remaining numbers being equally likely. Denote this die u.

In N rolls of a die d which can be either f or u (the die is either fair or unfair), the probability of observing k_i occurrences of face i, where $k_0 + k_1 + k_2 + k_3 = N$ is given by the multinomial distribution:

$$P_{|d}(k_0, k_1, k_2, k_3 \mid d) = \left[\frac{N!}{k_0!k_1!k_2!k_3!}\right] P_{0|d}^{k_0} P_{1|d}^{k_1} P_{2|d}^{k_2} P_{3|d}^{k_3} \tag{23}$$

where $P_{i|d}, i = 0, 1, 2, 3$ are the probabilities of observing face i, conditioned upon die being $d = u$ or $d = f$.

Objective: The experiment consists of picking a die with equal probability, tossing it $N = 10$ times and observing the number of appearances of each of $0, 1, 2$ and 3. The task is to decide which die was chosen.

(a) Formulate the maximum likelihood (ML) decision rule and show that it reduces to *[8pts]*

$$k_0 \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} t. \tag{24}$$

The expression for t should either be a numerical value or a numerical expression which can further be evaluated. *Note:* Use $\log_{10}$ values provided in the second page if necessary.

(b) Given that 10 rolls produce observations of $k_0 = 5, k_1 = 2, k_2 = 1, k_3 = 2$, what would you conclude? *[2pts]*

(Solution starts here)

(a) $\to$ We start with the MAP rule which is equivalent to the ML rule since the prior probabilities of picking either die from the box are equal. For this, we summarize the given conditional probabilities as follows:

$$P_{i|d}(0 \mid f) = \frac{1}{4}, \quad P_{i|d}(1 \mid f) = \frac{1}{4}, \quad P_{i|d}(2 \mid f) = \frac{1}{4}, \quad P_{i|d}(3 \mid f) = \frac{1}{4}. \tag{25}$$

$$P_{i|d}(0 \mid u) = \frac{1}{2}, \quad P_{i|d}(1 \mid u) = \frac{1}{6}, \quad P_{i|d}(2 \mid u) = \frac{1}{6}, \quad P_{i|d}(3 \mid u) = \frac{1}{6}. \tag{26}$$

The MAP (ML) rule for deciding on the die is written and the solution is obtained as follows:

$$\text{MAP: } P_{|d}(k_0, k_1, k_2, k_3 \mid u) \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} P_{|d}(k_0, k_1, k_2, k_3 \mid f) \tag{27}$$

$$\left(\frac{1}{2}\right)^{k_0}\left(\frac{1}{6}\right)^{k_1}\left(\frac{1}{6}\right)^{k_2}\left(\frac{1}{6}\right)^{k_3} \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} \left(\frac{1}{4}\right)^{10}$$

$$\text{Since } k_1 + k_2 + k_3 = 10 - k_0, \quad \left(\frac{2}{3}\right)^{10-k_0} \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} \left(\frac{1}{2}\right)^{k_0}$$

$$10 \log_{10}\left(\frac{2}{3}\right) \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} k_0 \log_{10}\left(\frac{1}{3}\right)$$

$$k_0 \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} 10 \frac{\log_{10} \frac{3}{2}}{\log_{10} 3} \tag{28}$$

$$\implies k_0 \underset{\text{fair}}{\overset{\text{unfair}}{\gtrless}} 3.691. \tag{29}$$

The result means that k_0 is a sufficient statistic for deciding on the hypothesis testing problem.

(b) Since $k_0 = 5 > 3.691$, we conclude that the die that we picked is most likely the "unfair" die. Note that the qualifier "most likely" is used because the MAP rule does have some probability of error.

Problem 3 (UNKNOWN PRIORS (14+5 pts))

You receive a signal $Y = X + Z$, where Z is Gaussian $\sim (0, 1)$ and the prior distribution P_X of the transmitted signal X as per your belief is given by

$$P_X(x) = \begin{cases} q & x = 1, \\ 1-q & x = -1. \end{cases} \tag{30}$$

(a) Derive the maximum a-posteriori probability (MAP) rule for the hypothesis testing problem of decoding $d(Y) = \hat{X}$ assuming $q = \frac{e}{1+e}$ where "e" refers to the standard exponent. Express the MAP decoding rule as *[3pts]*

$$y \underset{\hat{x}=-1}{\overset{\hat{x}=1}{\gtrless}} t. \tag{31}$$

with an explicit value for t.

(b) Unfortunately, it turns out that the prior distribution of X that you have is incorrect (sorry about that!). Instead, X has the distribution *[6pts]*

$$P_X(x) = \begin{cases} p & x = 1, \\ 1-p & x = -1, \end{cases} \tag{32}$$

where $0 \leq p \leq 1$. You, however, do not know that you've been given the wrong distribution and are using the decision rule from part (a). What is the resulting error probability?

(c) *The worst prior:* So much for wrong prior information; what value of p yields the largest probability of error for the decision rule from part (a)? *[5pts]*

(d) **Bonus:** *(Exact explanation/mathematical argument required; no partial credit)* *[5pts]*

"Minimax" strategy for combating the worst prior: Someone reveals that your value q is incorrect, but you still do not have any knowledge about what the exact prior p is. Assume that whatever value of q you get to choose, the value of p that provides the worst probability of error for that q will be chosen by nature. This looks hopeless really, but there is something that you can still do:

1. Namely, you can minimize the worst that you can do! What should be your choice of q that minimizes this worst-case probability of error of all choices of q that you can assume?
2. Explicitly state this minimum of all worst-case error probabilities.

Justify your answer rigorously.

(Solution starts here)

(a) For binary signaling, the MAP rule (from notes of Lecture 3) is given as:

$$y \underset{\hat{X}=-1}{\overset{\hat{X}=1}{\gtrless}} \frac{\sigma^2}{2} \ln\left(\frac{1-q}{q}\right) \tag{33}$$

$$y \underset{\hat{X}=-1}{\overset{\hat{X}=1}{\gtrless}} -0.5. \tag{34}$$

(b) The probability of error is obtained from the law of total probability as follows:

$$P_e = P_X(X=-1)\mathbb{P}[Z > 0.5] + P_X(X=1)\mathbb{P}[Z < -1.5] \tag{35}$$
$$= (1-p)\cdot Q(0.5) + p\cdot Q(1.5). \tag{36}$$

(c) We see that nature can choose between $0 \leq p \leq 1$. In order to maximize the probability of error, consider (36) in a more generic form so that we can get an intuition of what nature should be doing. The probability of error for a binary signaling system with actual prior probabilities $p, 1-p$ and our beliefs $q, 1-q$ is written as

$$P_e = P_X(X=-1)\cdot B + P_X(X=1)\cdot A \tag{37}$$
$$= (1-p)\cdot B + p\cdot A. \tag{38}$$

where $A = \mathbb{P}[\text{error} \mid X=1], B = \mathbb{P}[\text{error} \mid X=-1]$ are constants that only depend upon our decision rule threshold that we derive if we are given a certain information about the prior probabilities. In this case, our belief about the priors are $q, 1-q$ and constants A, B are $Q(1.5), Q(0.5)$ respectively.

We see that, in order to maximize (38), nature has to put all its possible weight as the multiplier of the maximum value of A, B. Hence in our case, the value of p that yields the worst-case probability of error for us is

$$p = 0 \tag{39}$$

since $B > A$.

(d) **Bonus:**

1. From the more generic analysis as done above, we can conclude that, if we are to minimize the worst-case probability of error, we should give nature no incentive to choose any value of p since it decides on the value of p when it looks at the values A, B. If that is the case, then we should make sure that the values of A and B are equal. This is the only case when nature cannot do better (better for nature means being more adversarial to us) by choosing a particular value of p over any other value of p since if $A = B$,

$$P_e = p\cdot A + (1-p)\cdot A$$
$$= A, \quad \forall\ 0 \leq p \leq 1. \tag{40}$$

Thus, if we are to set our values $A = B$, the following should be done:

$$A = B$$
$$\mathbb{P}[\text{error} \mid X=1] = \mathbb{P}[\text{error} \mid X=-1] \tag{41}$$
$$\mathbb{P}[Z < \text{threshold} - 1] = \mathbb{P}[Z > \text{threshold} + 1] \tag{42}$$
$$\mathbb{P}[Z > 1 - \text{threshold}] = \mathbb{P}[Z > 1 + \text{threshold}] \tag{43}$$

whose only solution is threshold $= 0$. If the optimal value of threshold that minimizes the worst-case probability of error is 0, then our assumption of prior, according to (33), should be

$$P_X(1) = P_X(-1) = \frac{1}{2} \tag{44}$$

which is, unsurprisingly, the underlying assumption of a "maximum likelihood" decoding rule.

2. Once we assume $q = 1 - q = \frac{1}{2}$, we see that the conditional probability of error for both the hypotheses are obtained as

$$A = B = Q(1) \tag{45}$$

using which we obtain the minimum of all worst-case total probability of error as $P_{\mathrm{e}} = Q(1)$.

Problem 4 (M-ARY SIGNALING AND BINARY FADING CHANNEL(**19 pts**))

Consider the M-PSK constellation shown in Fig. 2. The constellation is used by a system over an AWGN binary fading channel wherein the noise components of $\mathbf{n}$ along the two dimensions are independent of each other and are identically $\sim (0, \sigma^2)$. The transmitted point $\mathbf{X} = \mathbf{x}$ is selected equiprobably from one of the constellation points $s_m = \sqrt{E} e^{j\frac{2\pi m}{M}}, m = 0, 1, \ldots, M-1$ and the received point is written as

$$\mathbf{y} = \alpha \mathbf{x} + \mathbf{n}. \tag{46}$$

Throughout this problem, assume that the receiver knows α for every received $\mathbf{y}$ and hence it can decode using a modified AWGN MAP (ML) rule; α is a real-valued random variable.

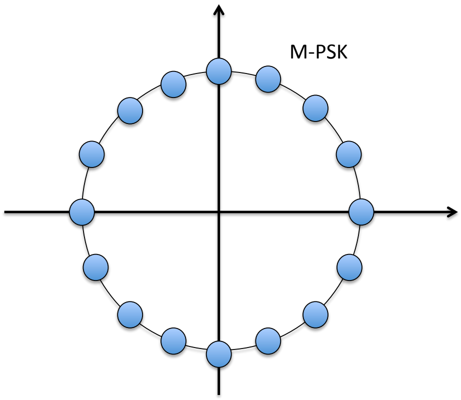


Figure 2: M-PSK

Note: You may find definitions (as a reference) and numerical values on the second page useful while answering the following.

(a) Determine the minimum distance d_{min} of the received constellation ($\alpha\mathbf{X}$) as a function of its average energy and M. For this, first determine the average energy of the received constellation by conditioning on $\alpha = a$ where a is any constant. *[3pts]*

(b) Bound the average probability of symbol error $P_{\mathrm{e}|\alpha}$ using the individual nearest neighbour union bound (INNUB) and express it as a function of the average energy (which is a function of conditioning on the fading parameter $\alpha = a$), M and σ^2. Use the relationship derived in (a) to express d_{min} in terms of the average energy and M. *[4pts]*

(c) Denote $\frac{E}{2\sigma^2} = \mathrm{SNR}$. The binary fading random variable α is distributed as follows: *[5pts]*

$$P_{\alpha^2}(\alpha^2 = a^2) = \begin{cases} \frac{1}{\mathrm{SNR}} & a^2 = 7.5 \times \frac{1}{\mathrm{SNR}}, \\ 1 - \frac{1}{\mathrm{SNR}} & a^2 = 10^3 \times \frac{1}{\mathrm{SNR}}. \end{cases} \tag{47}$$

What is the probability of symbol error P_{e}? *Note:* The probability of symbol error P_{e} is written using the law of total probability as

$$P_{\mathrm{e}} = \sum_{\alpha} P_{\mathrm{e}|\alpha} P(\alpha).$$

(d) Given a symbol error rate requirement of $P_{\mathrm{e}} \leq 10^{-6}$, which constellation among $M = 4$ and $M = 2$ is better and (approximately) by how many dBs? *Hint:* Compare the minimum required operating SNR for achieving the given symbol error rate requirement using the bound obtained. *Note:* A quantity "x" in decibels (dB) is $10 \cdot \log_{10} x$. *[7pts]*

(Solution starts here)

(a) The average energy is given as $\sum_{m=0}^{M-1} P_H(m)\|\boldsymbol{s}_m\|^2$. It is given that the signals are equiprobable (which is intuitive to expect in a real-time system design). The energy of every point is given by the square of norm of the signal point. Conditioning on $\alpha = a$, the received constellation is the transmitted constellation scaled by a factor of 'a'. Hence,

$$\|\boldsymbol{s}_m\|^2 = a^2 E \|e^{\frac{j2\pi m}{M}}\|^2 = a^2 E. \tag{48}$$

Given the fact that every signal point is equiprobable, the received average energy, conditioned on $\alpha = a$, is equal to $a^2 E$.

The minimum distance of the constellation can be computed trigonometrically. Consider a part of the constellation shown in Figure 3.

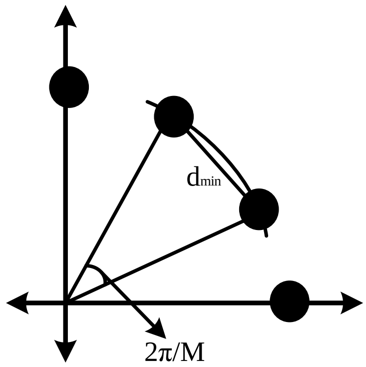


Figure 3: Minimum distance of M-PSK

The length of the arc between any two nearest points is a close approximation to the minimum distance d_{min} as shown in the figure, but an exact computation is possible by applying law of sines to the isosceles triangle. The radius of the circle, conditioning on $\alpha = a$, is $a\sqrt{E}$, seen from the complex representation of the signal points as $re^{j\theta}$:

$$\begin{aligned}\frac{d_{min}}{\sin\frac{2\pi}{M}} =& \frac{a\sqrt{E}}{\sin\frac{\pi - \frac{2\pi}{M}}{2}} \\ d_{min} =& a\sqrt{E}\left(\frac{\sin\frac{2\pi}{M}}{\cos\frac{\pi}{M}}\right) \\ =& 2a\sqrt{E}\sin\left(\frac{\pi}{M}\right).\end{aligned} \tag{49}$$

(b) For the individual nearest neighbour union bound (INNUB), we see that every point has 2 neighbours, both at the distance d_{min} which implies that the bound, due to the equiprobable nature of all the signal points, reduces to simply calculating it for just one signal point as follows:

$$\begin{aligned}P_{\mathrm{e}|\alpha=a} \leq& \sum_{j\in\mathcal{N}_{ml}(m)} Q\left(\frac{\|\boldsymbol{s}_m - \boldsymbol{s}_j\|}{2\sigma}\right) \\ =& 2Q\left(\frac{d_{min}}{2\sigma}\right) \\ =& 2Q\left(\sqrt{a^2\sin^2\left(\frac{\pi}{M}\right)\frac{E}{\sigma^2}}\right).\end{aligned} \tag{50}$$

(c) By denoting $\frac{E}{2\sigma^2} = \text{SNR}$, (50) is used to compute the total probability of error (by using law of total probabilities) as:

$$P_\text{e} \leq \frac{1}{\text{SNR}} 2Q\left(\sqrt{2\sin^2\left(\frac{\pi}{M}\right)\cdot 7.5}\right) + \left(1 - \frac{1}{\text{SNR}}\right) 2Q\left(\sqrt{2\sin^2\left(\frac{\pi}{M}\right)\cdot 10^3}\right). \quad (51)$$

(d) First we note that for values of M that are of interest to us, the second term in (51) is approximately equal to 0. Hence, we proceed to find the minimum required SNR, firstly for $M = 2$, as follows:

$$\begin{aligned} P_{\text{e},M=2} \leq & \frac{1}{\text{SNR}} 2Q\left(\sqrt{2\sin^2\left(\frac{\pi}{M}\right)\cdot 7.5}\right) \\ = & \frac{1}{\text{SNR}} 2Q\left(\sqrt{2\sin^2\left(\frac{\pi}{2}\right)\cdot 7.5}\right) \\ = & \frac{1}{\text{SNR}} 2Q\left(\sqrt{15}\right) \quad (52) \\ P_\text{e} \leq 10^-6 \implies \text{SNR}_{M=2} \geq & 2Q\left(\sqrt{15}\right) 10^6 \quad (53) \\ \implies \text{SNR}_{M=2} \geq & 107.5 \text{ or } 20.314\text{dB}. \quad (54) \end{aligned}$$

Similarly, we can obtain a bound on the required $\text{SNR}_{M=4}$ as

$$\text{SNR}_{M=4} \geq 37.924\text{dB} \quad (55)$$

which leads us to conclude that the constellation with $M = 2$ is approximately 17.6dB superior to the constellation with $M = 4$.

Note: The emphasis here though is that with a fading channel like this, the probability of error scales linearly (inversely proportional) as seen in (52) with respect to SNR. This suggests that the probability of error drops very slowly (linearly) with increasing SNR as compared to the AWGN channel case where it would drop exponentially (as an argument of the Q-function). This is one of the foremost reasons why design for wireless systems is a comparatively challenging task than design for wired systems where fading is not much of a concern.

(This page left intentionally blank)

Problem 5 (Convolutional Code and OFDM(**17 pts**))

Consider a convolutional code described by the state diagram of Fig. 4. An encoder structure is shown in Fig. 5. The state diagram is complete but the structure of the encoder is incomplete.

→State labels in the state diagram are given in the order b_{j-1}, b_{j-2} i.e. for example, state label "10" corresponds to $b_{j-1} = 1, b_{j-2} = 0$.

→Outputs in the state diagram are given in the order $c_{2j-1}c_{2j}$ i.e. for example, if the encoder is in state "01" and the input is $b_j = 1$, then the output is $c_{2j-1} = 1, c_{2j} = 0$ which is represented as "10".

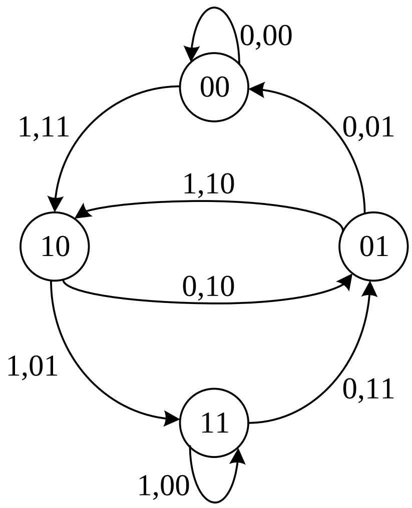


Figure 4: Complete state diagram

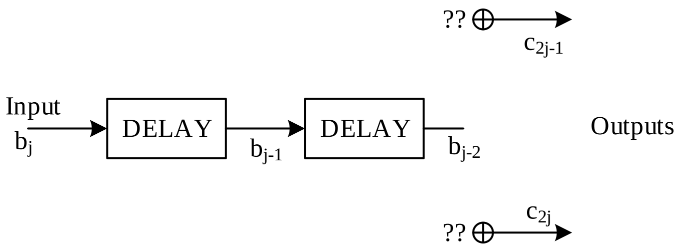


Figure 5: Incomplete encoder structure

(a) Complete the design of this encoder by finding the inputs to the XOR adders which output c_{2j} and c_{2j-1} i.e. copy the incomplete encoder structure and then add to it the necessary connections to the adders such that the outputs shown in the state diagram are generated. [*7pts*]

(b) A four-state rate-1/2 convolutional code (like the one above) is used for transmission over an ISI channel by employing OFDM. We map the output of the convolutional code (using BPSK) as: [*10pts*]

$$\hat{x}_n = \begin{cases} -1 & c_n = 1, \\ 1 & c_n = 0. \end{cases} \tag{56}$$

Premise of the problem: The encoder starts from the all-zero state; which means the delay elements are refreshed to 0's before the first input. We encode 2 bits u_1, u_2 to obtain c_1, c_2, c_3, c_4. At the end, we add two redundant bits u_3, u_4 as input to the convolutional encoder to bring it back to all zeros state.

$\rightarrow$Thus, a total of 8 symbols $\hat{x}_n, n = 1, 2, \ldots, 8$ are transmitted (this is our actual coded symbols mapped using BPSK) corresponding to the codeword sequence $c_1, c_2, \ldots, c_8$. The ISI channel h_l is given by its z-transform as follows:

$$H(z) = 1.5 - 0.5z^{-4}. \tag{57}$$

At the receiver, complex AWGN noise w_n with i.i.d real and imaginary components $\sim \eta(0, \sigma^2)$ is added to the received symbols y_n which is of length $N_c + 2L - 2$.

The problem: Assuming that the transmitter, using $N_c = 8$ sub-carriers, employs required cyclic prefix as it transmits using OFDM, derive the "metric" that needs to be considered at the decoder of the convolutional code to retrieve the maximum likelihood sequence $x'_1, x'_2, \ldots, x'_8$ corresponding to the transmitted sequence $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_8$.

Hint: To do this, first assume that you have processed the received OFDM symbols y_n i.e. assume that you have removed cyclic prefix and the last $L-1$ symbols, where L denotes the number of coefficients in the channel impulse response. Perform the required length DFT of the channel impulse response and obtain a sequence of the N_c data symbols $\hat{x}_n$ which are weighted according to the N_c-point DFT of the channel coefficients plus the AWGN noise. This is the parallel AWGN channel representation obtained by using OFDM. Your task now is to find the metric to be implemented.

Note: For this problem, use the following (standard) definition of a DFT; the DFT V_k of an N_c point sequence $v_n, n = 0, 1, \ldots, N_c - 1$ is:

$$V_k = \sum_{n=0}^{N_c-1} v_n e^{-\frac{j2\pi kn}{N_c}}, k = 0, 1, \ldots, N_c - 1.$$

Provide rigorous explanation for every conceptual step that you take.

(Solution starts here)

(a) From the outputs shown in the state diagram, we can construct the following truth table:

b_j	b_{j-1}	b_{j-2}	c_{2j-1}	c_{2j}
0	0	0	0	0
0	0	1	0	1
0	1	0	1	0
0	1	1	1	1
1	0	0	1	1
1	0	1	1	0
1	1	0	0	1
1	1	1	0	0

From this we can see that $c_{2j-1} = b_j \oplus b_{j-1}$ and $c_{2j} = b_j \oplus b_{j-2}$. The completed encoder structure is shown in Fig. 6.

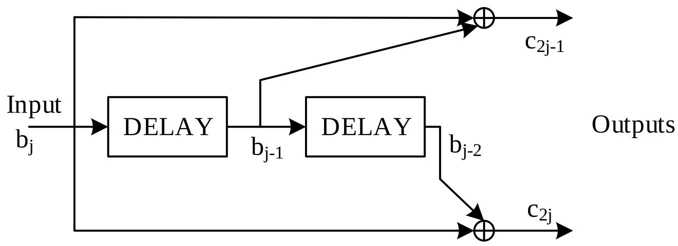


Figure 6: Completed encoder structure

(b) *Analysis of OFDM system:* We know that the parallel AWGN channel representation of an OFDM system is

$$\tilde{Y}[k] = \tilde{H}[k]\hat{x}[k] + \tilde{W}[k], \quad k = 1, 2, \ldots, N_c$$

where $\tilde{Y}_k$, $\tilde{H}_k$ and $\tilde{W}_k$ are the N_c-point DFT's of the output from the channel, the channel impulse response and the AWGN noise. The first step is to find the N_c-point DFT of the channel impulse response. From the z-transform of the channel, we note that the impulse response of the channel is

$$h_n = [1.5 \;\; 0 \;\; 0 \;\; 0 \;\; -0.5 \;\; 0 \;\; 0 \;\; 0]\,. \tag{58}$$

Thus the DFT of the channel response is:

$$\begin{aligned} \tilde{H}_k &= \sum_{n=0}^{N_c-1} h_n e^{-\frac{j2\pi kn}{N_c}}, k = 0, 1, \ldots N_c - 1 \\ &= 1.5 - 0.5e^{-j\pi k}, k = 0, 1, \ldots 7 \\ &= 1.5 - (0.5)(-1)^k, k = 0, 1, \ldots 7 \\ &= [1 \;\; 2 \;\; 1 \;\; 2 \;\; 1 \;\; 2 \;\; 1 \;\; 2]. \end{aligned} \tag{59}$$

For $j \in \{1, 2, 3, 4\}$, we now have (substituting the values of DFT's of channel impulse response in our parallel AWGN channel representation)

$$\tilde{Y}[2j] = 2 \cdot \hat{x}[2j] + \tilde{W}[2j] \tag{60}$$

$$\tilde{Y}[2j-1] = 1 \cdot \hat{x}[2j-1] + \tilde{W}[2j-1]. \tag{61}$$

Metric: We now proceed to derive the metric for this parallel weighted AWGN channel for convolutional decoding. We are looking for the maximum likelihood sequence $x_1', x_2', \ldots, x_8'$. First we write the conditional PDF corresponding to the received sequence corresponding to the parallel AWGN channel representation as follows:

$$f_{\tilde{\mathbf{Y}}|\hat{\mathbf{x}}}(\tilde{\mathbf{y}} \mid \hat{\mathbf{x}}) = \frac{1}{\sqrt{2\pi\sigma^2}^{N_c}} e^{-\frac{1}{2\sigma^2}\sum_{j=1,3,5,7}\left(\tilde{Y}[j]-\hat{x}[j]\right)^2 + \sum_{j=2,4,6,8}\left(\tilde{Y}[j]-2\cdot\hat{x}[j]\right)^2}. \tag{62}$$

The maximum likelihood decoding rule picks the $\hat{\mathbf{x}}$ that maximizes (62). By making appropriate manipulations (and since $\hat{x}^2 = 1, \forall \hat{x}$) the appropriate sequence $\mathbf{x}'$ among all possible sequences $\hat{\mathbf{x}}$ is the one that minimizes the metric

$$\sum_{j=1}^{4} 2 \cdot \hat{x}[2j]\tilde{Y}[2j] + \hat{x}[2j-1]\tilde{Y}[2j-1].$$

Therefore, the result of accounting for the channel response in this case is a weighting of every other term by a factor of 2 while looking for the maximum likelihood received sequence. A very intuitive explanation would go as follows: at some time instants our constellations are changed (scaled by the DFT of the channel impulse response) and at other time instants we receive one of the points from the transmitted constellation. In order to account for this scaling, we scale our constellation appropriately for different time instants to look for the best correlation metric.

(This page left intentionally blank)

Problem 6 (Baseband and Passband Signaling **(16 pts)**)

Let the transmitted bandpass (or passband) signal be given by

$$x(t) = a\cos\left(2\pi f_c t + \frac{\pi}{2}\right) - b\sin\left(2\pi f_c t + \frac{\pi}{2}\right) \tag{63}$$

and $a \in \{1, 2\}, b \in \{1, 2\}$.

(a) Find the baseband equivalent $x_b(t) = x_1(t) + jx_2(t)$ for the transmitted signal where "j" denotes imaginary component. *[6pts]*

(b) Find the vector representation of the baseband signal and draw the corresponding signal constellation. *[3pts]*

(c) If $a = \begin{cases} 1 & \text{w.p } \frac{1}{2} \\ 2 & \text{w.p } \frac{1}{2} \end{cases}$ and $b = \begin{cases} 1 & \text{w.p } \frac{1}{2} \\ 2 & \text{w.p } \frac{1}{2} \end{cases}$, find the average energy of the baseband signal. *[2pts]*

(d) Is this a minimum energy configuration? If not, how will you modify the constellation so that it is of minimum energy? Draw the minimum energy signal constellation if your answer was no. *[5pts]*

(Solution starts here)

(a) We first find the real component of $x_b(t)$:

$$\begin{aligned}
x_1(t) &= \text{LPF}\left\{\sqrt{2}x(t)\cos(2\pi f_c t)\right\} && (64)\\
&= (\sqrt{2})\text{LPF}\left\{a\cos\left(2\pi f_c t + \frac{\pi}{2}\right)\cos(2\pi f_c t) - b\sin\left(2\pi f_c t + \frac{\pi}{2}\right)\cos(2\pi f_c t)\right\}\\
&= \left(\frac{\sqrt{2}}{2}\right)\text{LPF}\left\{a\cos\left(\frac{\pi}{2}\right) + a\cos\left(4\pi f_c t + \frac{\pi}{2}\right) - b\sin\left(4\pi f_c t + \frac{\pi}{2}\right) - b\sin\left(\frac{\pi}{2}\right)\right\}\\
&= -\frac{b}{\sqrt{2}}. && (65)
\end{aligned}$$

Similarly, we find $x_2(t)$:

$$\begin{aligned}
x_2(t) &= \text{LPF}\left\{-\sqrt{2}x(t)\sin(2\pi f_c t)\right\} && (66)\\
&= (-\sqrt{2})\text{LPF}\left\{a\cos\left(2\pi f_c t + \frac{\pi}{2}\right)\sin(2\pi f_c t) - b\sin\left(2\pi f_c t + \frac{\pi}{2}\right)\sin(2\pi f_c t)\right\}\\
&= -\left(\frac{\sqrt{2}}{2}\right)\text{LPF}\left\{a\sin\left(4\pi f_c t + \frac{\pi}{2}\right) - a\sin\left(\frac{\pi}{2}\right) - b\cos\left(\frac{\pi}{2}\right) + b\cos\left(4\pi f_c t + \frac{\pi}{2}\right)\right\}\\
&= \frac{a}{\sqrt{2}}. && (67)
\end{aligned}$$

Therefore, we have that

$$x_b(t) = \frac{-b + ja}{\sqrt{2}}\mathbb{1}_t.$$

(b) The constellation is clearly made of the points $\frac{1}{\sqrt{2}} \times \{(-2, 1), (-1, 2), (-1, 1), (-2, 2)\}$.

(c) Since each point is equally probable, the average energy is simply the average of the energies of each point in the constellation and thus

$$\begin{aligned} E_{avg} &= \frac{1}{4}(1 + 2.5 + 2.5 + 4) \\ &= 2.5. \end{aligned}$$

(d) This is not a minimum-energy constellation. If we shift the constellation by $\left(\frac{3}{2\sqrt{2}}, -\frac{3}{2\sqrt{2}}\right)$, then the new constellation would be centered at zero and would consist of the points $\frac{1}{2\sqrt{2}} \times \{(1,-1),(1,1),(-1,-1),(-1,1)\}$. Since the points are equiprobable, and the centroid of the new constellation is at zero, this new constellation is the minimum-energy configuration.

(This page left intentionally blank)