Problem 1. (10 points)
Consider a hypothesis testing problem where the hypothesis \( H \) can take the values 0 or 1 with equal probability. The observation \( Y = (Y_1, Y_2) \), when \( H = i \), is given by

\[
Y = c_i + Z,
\]

where \( c_0 = (1, 2) \) and \( c_1 = (2, 1) \) are vectors in \( \mathbb{R}^2 \) and \( Z \) is Gaussian, zero mean, with covariance matrix \( K = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \).

Let \( \alpha \) be a real number and let \( T = (1 - \alpha)Y_1 + \alpha Y_2 \) be a 1-dimensional statistic.

(a) (2 pts) What is the probability distribution of the statistic \( T \) when \( H = i, i = 0, 1? \)

Solution: Defining \( v = (1 - \alpha, \alpha) \) and \( Z = (Z_1, Z_2) \), when \( H = i, T = u_i + W \), where \( u_i = \langle v, c_i \rangle \), and \( W = \langle v, Z \rangle = (1 - \alpha)Z_1 + \alpha Z_2 \). For \( i = 0 \), we have \( u_0 = \langle v, c_0 \rangle = (1 - \alpha) + 2\alpha = 1 + \alpha \), and for \( i = 1 \), we have \( u_1 = \langle v, c_1 \rangle = 2(1 - \alpha) + \alpha = 2 - \alpha \). For either value of \( H = i, T \) is a Gaussian random variable with mean \( u_i \) and variance the same as \( W \), given by \((1 - \alpha)^2 + 2\alpha^2 = \sigma^2 \), say. Hence, when \( H = 0, T \sim \mathcal{N}(1 + \alpha, \sigma^2) \) and when \( H = 1, T \sim \mathcal{N}(2 - \alpha, \sigma^2) \).

(b) (2 pts) Consider the MAP decision rule based only on the statistic \( T \). What is the error probability for this rule?

Solution: With the same notation as in part (a), note that \( W \) (which is the noise associated with the statistic \( T \)) is a zero mean Gaussian with variance \( \sigma^2 = (1 - \alpha)^2 + 2\alpha^2 \). The error probability is thus \( Q \left( \frac{d}{2\sigma} \right) \) where \( d = |u_0 - u_1| = |1 - 2\alpha| \).

(c) (3 pts) Which choice of \( \alpha \) will minimize the error probability in (b)?)

Solution: Minimizing the error probability in (b), is equivalent to maximizing \( \frac{d^2}{\sigma^2} = (1 - 2\alpha)^2/(1 - 2\alpha + 3\alpha^2) \) (since \( |x| \mapsto x^2 \) is an increasing mapping and \( Q(\cdot) \) is a decreasing function). On differentiating, we find that \( \alpha = -1 \) is the maximizer.

(d) (3 pts) Is the statistic \( T \), with the \( \alpha \) of (c), a sufficient statistic?

Solution: Yes. The likelihood ratio for the observation \( y = (y_1, y_2) \) is equal to

\[
\frac{f_{Y|H}(y|0)}{f_{Y|H}(y|1)} = \frac{f_Z(y - c_0)}{f_Z(y - c_1)} = \frac{f_{Z_1}(y_1 - 1)f_{Z_2}(y_2 - 2)}{f_{Z_1}(y_1 - 2)f_{Z_2}(y_2 - 1)}
\]

\[
= \frac{\exp\left(-\frac{(y_1 - 1)^2}{2}\right) \exp\left(-\frac{(y_2 - 2)^2}{4}\right)}{\exp\left(-\frac{(y_1 - 2)^2}{2}\right) \exp\left(-\frac{(y_2 - 1)^2}{4}\right)}
= \exp\left(y_1 - \frac{1}{2} + y_2 - 1 - 2y_1 + 2 - \frac{y_2}{2} + \frac{1}{4}\right)
= \exp\left(\frac{3}{4} - \frac{1}{2}(2y_1 - y_2)\right),
\]

which is a function of \( y \) only through \( 2y_1 - y_2 \), which is exactly the statistic \( T \) with \( \alpha = -1 \).
Remark: In general, when \( Y = c_i + Z \) with \( Z \) Gaussian and covariance \( K \), \( T = \langle v, Y \rangle \) is a statistic, and \( v \) is chosen to maximize the SNR = \( \frac{(v,c_0-c_1)^2}{\langle v, K v \rangle} \), then \( T \) is a sufficient statistic. This problem is only a special case.
**Problem 2. (10 points)**

Consider a communication channel with input \((x_1, x_2)\) in \(\mathbb{R}^2\), and output \((Y_1, Y_2)\) in \(\mathbb{R}^2\) given by

\[
Y_1 = A_1 x_1 - A_2 x_2 + Z_1 \\
Y_2 = A_2 x_1 + A_1 x_2 + Z_2
\]

where \(A_1, A_2, Z_1, Z_2\) are all i.i.d. \(\mathcal{N}(0, 1)\) random variables.

(a) (2 pts) Observe that \((Y_1, Y_2)\) is a Gaussian vector for any given \((x_1, x_2)\). Find its mean and covariance matrix in terms of \(x_1, x_2\).

**Solution:** Since \(A_1, A_2, Z_1, Z_2\) are all zero mean random variables, \((Y_1, Y_2)\) is also zero mean. Since \(A_1, A_2, Z_1, Z_2\) are independent, the variances of \(Y_1\) and \(Y_2\) are given by

\[
\text{Var}(Y_1) = x_1^2 \text{Var}(A_1) + x_2^2 \text{Var}(A_2) + \text{Var}(Z_1) = 1 + x_1^2 + x_2^2, \\
\text{Var}(Y_2) = x_1^2 \text{Var}(A_2) + x_2^2 \text{Var}(A_1) + \text{Var}(Z_2) = 1 + x_1^2 + x_2^2.
\]

Further, the covariance between \(Y_1\) and \(Y_2\) is equal to zero, as

\[
\mathbb{E}[Y_1 Y_2] = x_1 x_2 (\mathbb{E}[A_1^2] - \mathbb{E}[A_2^2]) + (x_1^2 - x_2^2) \mathbb{E}[A_1] \mathbb{E}[A_2] + \mathbb{E}[Z_1] (\ldots) + + \mathbb{E}[Z_2] (\ldots) = 0.
\]

Hence \((Y_1, Y_2) \sim \mathcal{N}((0, 0), (1 + x_1^2 + x_2^2)I_2) = \mathcal{N}((0, 0), (1 + \|x\|^2)I_2)\).

(b) (3 pts) Suppose \(c_1, \ldots, c_m\) are \(m\) vectors in \(\mathbb{R}^2\), and that when the message \(H = i\), the vector \(c_i\) is input to the communication channel above. The receiver, from the observation \((Y_1, Y_2)\) tries to guess the value of \(H\). Show that, no matter how the vectors \(c_1, \ldots, c_m\) are chosen, \(T = Y_1^2 + Y_2^2\) is a sufficient statistic.

**Solution:** The pdf of the observation \(y = (y_1, y_2)\) when \(H = i\) is given by

\[
f_{Y|H}(y|i) = \frac{1}{2\pi(1 + \|c_i\|^2)} \exp \left( -\frac{y_1^2 + y_2^2}{2(1 + \|c_i\|^2)} \right),
\]

which depends only on \(y\) only through \(y_1^2 + y_2^2\), hence \(T = Y_1^2 + Y_2^2\) is a sufficient statistic.

(c) (2 pts) Suppose \(m = 4\), and \(c_1 = (5, 0), c_2 = (0, 5), c_3 = (3, 4), c_4 = (4, 3)\). All four messages are equally likely. What is the probability of error of the MAP decoder?

**Solution:** Since all the \(c_i\)'s have the same norm, the output is independent of the input, thus the probability of error = 3/4 (equivalent to a random guess between the four options).

(d) (3 pts) Consider four designs, all with \(m = 2\), all with equally likely messages:

1. \(c_1 = (0, 0), c_2 = (10, 0)\)
2. \(c_1 = (8, 6), c_2 = (0, 0)\)
3. \(c_1 = (0, 0), c_2 = (5, 0)\)
4. \(c_1 = (8, 6), c_2 = (10, 0)\)

With \(P_1, P_2, P_3, P_4\) denoting the error probability with MAP decoding of these systems, how will \(P_1, \ldots, P_4\) be ordered? What is the value of \(P_4\)?

**Solution:** \(P_4 = P_2 < P_3 < P_4 = 1/2\).
$P_1 = P_2$ because in both cases, one message has norm 0 and the other has norm 10.

$P_1 < P_3$ because the distance between the message constellation is more in design 1.

$P_4 = 1/2$ for the same reason as part (c), and $P_3 < P_4$ because we can definitely do better than a random guess between the two messages in design 3.

Remark: Note that the channel is the “real” equivalent of the “complex” channel $Y = Ax + Z$. Since $A$ and $Z$ are both circularly symmetric, is it not a surprise that $|Y|^2$ is a sufficient statistic and the input influences the output only via $|x|^2$. 
Problem 3. (11 points)
Consider a communication system over an additive white Gaussian noise channel (with noise intensity = 1) with two equally likely messages transmitted via waveforms \( w_0(t) = -w_1(t) = \sqrt{\mathcal{E}} \mathbf{1}\{ |t| < \frac{1}{2} \} \).

(a) (3 pts) At the receiver suppose we pass the received signal \( R(t) \) through a filter with impulse response \( h(t) = \mathbf{1}\{ |t| < \frac{1}{2} \} \), sample the filter output at \( t_0 = 0 \), and, with \( Y \) denoting the value of the sample, decide \( \hat{H} = 1 \) if \( Y < 0 \), \( \hat{H} = 0 \) if \( Y \geq 0 \). Is this receiver optimal? What is the probability of error?

Solution: Just as in part (a), the optimal choice of \( \hat{H} \) is to decide \( \hat{H}(Y_1) \) is to decide 1 if \( Y_1 < 0 \), \( \hat{H}(Y_1) = 0 \) if \( Y_1 \geq 0 \), and since the noise variance is now 2, the error probability is \( Q(\sqrt{\mathcal{E}}) \).

(b) (2 pts) Suppose we choose \( n = 1 \), and \( t_1 = 0 \). What is the distribution of \( Y_1 \) given \( H = i, i = 0,1 ? \)

Solution: Let \( h'(t) = \mathbf{1}\{ |t| < 1 \} \) be the new impulse response. Then \( Y_1 = c_i' + Z' \), where

\[
\begin{align*}
c_i' &= \int_R w_i(t) h'(-t) \, dt = \begin{cases} 
\int_{-\frac{1}{2}}^{\frac{1}{2}} \sqrt{\mathcal{E}} \, dt & \text{if } i = 0 \\
\int_{\frac{1}{2}}^{\frac{3}{2}} -\sqrt{\mathcal{E}} \, dt & \text{if } i = 1 
\end{cases} = \begin{cases} 
\sqrt{\mathcal{E}} & \text{if } i = 0 \\
-\sqrt{\mathcal{E}} & \text{if } i = 1 
\end{cases} \\
Z' &= \int_R N(t) h'(-t) \, dt = \int_{-1}^{1} N(t) \, dt \sim \mathcal{N}(0,2),
\end{align*}
\]

where the last step follows since \( \|h'\|^2 = 2 \). Hence, \( Y_1 \sim \mathcal{N}(\sqrt{\mathcal{E}}, 2) \) for \( i = 0 \) and \( Y_1 \sim \mathcal{N}(-\sqrt{\mathcal{E}}, 2) \) for \( i = 1 \).

(c) (3 pts) What is the optimal choice of \( \hat{H}(Y_1) \) and what is the corresponding probability of error?

Solution: Just as in part (a), the optimal choice of \( \hat{H}(Y_1) \) is to decide \( \hat{H}(Y_1) = 1 \) if \( Y_1 < 0 \), \( \hat{H}(Y_1) = 0 \) if \( Y_1 \geq 0 \), and since the noise variance is now 2, the error probability is \( Q(\sqrt{\mathcal{E}}) \).

(d) (3 pts) Suppose we choose \( n = 2 \), with \( t_1 = -1/2, t_2 = 1/2 \). What is the optimal choice of \( \hat{H}(Y_1, Y_2) \) and what is the probability of error?

Hint: \( Y_1 + Y_2 \) is a sufficient statistic.
Solution: Observe that \( Y_1 + Y_2 = \int_{\mathbb{R}} R(t) \left( h' \left( -\frac{1}{2} - t \right) + h' \left( \frac{1}{2} - t \right) \right) \, dt \), then define

\[
 h''(t) = h' \left( -\frac{1}{2} - t \right) + h' \left( \frac{1}{2} - t \right) = \begin{cases} 
 1 & \text{if } -\frac{3}{2} < t < -\frac{1}{2} \\
 2 & \text{if } -\frac{1}{2} < t < \frac{1}{2} \\
 1 & \text{if } \frac{1}{2} < t < \frac{3}{2} \\
 0 & \text{else}
\end{cases}.
\]

Hence we have \( Y_1 + Y_2 = c_i'' + Z'' \), where

\[
 c_i'' = \int_{\mathbb{R}} w_i(t) h''(t) \, dt
 = \begin{cases} 
 \int_{-\frac{1}{2}}^{\frac{1}{2}} 2\sqrt{E} \, dt & \text{if } i = 0 \\
 \int_{-\frac{1}{2}}^{\frac{1}{2}} -2\sqrt{E} \, dt & \text{if } i = 1
\end{cases} = \begin{cases} 
 2\sqrt{E} & \text{if } i = 0 \\
 -2\sqrt{E} & \text{if } i = 1
\end{cases},
\]

\[
 Z' = \int_{\mathbb{R}} N(t) h''(t) \, dt \sim \mathcal{N}(0, 6),
\]

where the last step follows since \( \|h''\|^2 = 6 \). Thus we have that the optimal choice of \( \hat{H}(Y_1, Y_2) \) is to decide \( \hat{H}(Y_1, Y_2) = 1 \) if \( Y_1 + Y_2 < 0 \), \( \hat{H}(Y_1) = 0 \) if \( Y_1 + Y_2 \geq 0 \), and the error probability is

\[
 Q \left( \frac{2\sqrt{E}}{\sqrt{6}} \right) = Q \left( \sqrt{\frac{2E}{3}} \right).
\]

Remark: Let \( \hat{X}(f) \) denote the Fourier transform of any signal \( X(t) \). The signal that we would ideally like to have as in part (a) is the inverse Fourier transform of \( \hat{R}(f) \hat{h}(f) \), with \( \hat{h}(f) = \text{sinc}(f) \), and then sample the output at \( t = 0 \) to obtain the sufficient statistic \( Y \). However, we are only able to observe the signal \( S(t) \) with Fourier transform \( \hat{S}(f) = \hat{R}(f) \hat{h}'(f) \), where \( \hat{h}' = 2\text{sinc}(2f) \). Therefore, in principle, by passing \( S(t) \) through another filter with frequency response (i.e., the Fourier transform of the impulse response) \( \hat{g}(f) = \frac{\hat{h}(f)}{\hat{h}'(f)} = \frac{1}{2\cos(\pi f)} \), we should be able to obtain the desired signal. Since \( \hat{g}(f) \) is periodic with period 2, the impulse response \( g(t) \) is of the form \( \sum_{n \in \mathbb{Z}} c_n \delta \left( t - \frac{n}{2} \right) \) for some \( c_n \). This is equivalent to sampling \( S(t) \) at the instances \( t = \frac{n}{2} \), scaling them by \( c_n \), and summing the resulting quantities. The parts (c) and (d) look at the error performance when we truncate the sum to \( n = 1 \) and \( n = 2 \) terms respectively. By continuing to an infinite number of terms, we can, in principle, we recover the same error probability as part (a).
Problem 4. (10 points) 
Consider a communication system for the AWGN channel with noise intensity $N_0/2$ with four equally likely messages, and suppose the waveforms $w_1, \ldots, w_4$ have all unit norm and that $\langle w_i, w_k \rangle = \alpha$ for all $i \neq k$.

(a) (2 pts) Express $\|w_1 + w_2 + w_3 + w_4\|^2$ in terms of $\alpha$, and show that $-1/3 \leq \alpha \leq 1$.

Solution: Expanding $\|w_1 + w_2 + w_3 + w_4\|^2$, we have

$$\|w_1 + w_2 + w_3 + w_4\|^2 = \sum_{1 \leq i \leq 4} \|w_i\|^2 + \sum_{1 \leq i \neq j \leq 4} \langle w_i, w_j \rangle = 4 + 12\alpha,$$

and since $4 + 12\alpha = \|w_1 + w_2 + w_3 + w_4\|^2 \geq 0$, we have $\alpha \geq -1/3$. Furthermore, by the Cauchy-Schwarz inequality, we have, for $i \neq k$,

$$\alpha = \langle w_i, w_k \rangle \leq \|w_i\|\|w_k\| = 1.$$ 

(b) (3 pts) Let $\{\tilde{w}_1, \ldots, \tilde{w}_4\}$ be obtained by a translation of $\{w_1, \ldots, w_4\}$ so that the new signal set is of minimal average energy. Do $\tilde{w}_1, \ldots, \tilde{w}_4$ all have the same energy? If so, what is this energy in terms of $\alpha$?

Solution: The minimal average energy set of waveforms $\{\tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4\}$ is formed by subtracting the arithmetic mean $m = \frac{1}{4}(w_1 + w_2 + w_3 + w_4)$ from each waveform, i.e., $\tilde{w}_i = w_i - m$. Then,

$$\|\tilde{w}_i\|^2 = \|w_i - m\|^2 = \|w_i\|^2 + \|m\|^2 - 2\langle w_i, m \rangle$$

$$= 1 + \left\|\frac{1}{4}(w_1 + w_2 + w_3 + w_4)\right\|^2 - 2\left\langle w_i, \frac{1}{4} \sum_{1 \leq j \leq 4} w_j \right\rangle$$

$$= 1 + \frac{1}{16}(4 + 12\alpha) - \frac{2}{4}(1 + 3\alpha) = \frac{3}{4}(1 - \alpha)$$

for all $i$. Hence all the minimal energy waveforms $\tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4$ have the same energy, given by $\frac{3}{4}(1 - \alpha)$.

(c) (2 pts) Is there a common value of $\langle \tilde{w}_i, \tilde{w}_k \rangle$ for $i \neq k$? If so, what is the common value in terms of $\alpha$?

Solution: For any $i \neq k$, we have

$$\langle \tilde{w}_i, \tilde{w}_k \rangle = \langle w_i - m, w_k - m \rangle = \langle w_i, w_k \rangle + \|m\|^2 - \langle w_i, m \rangle - \langle w_k, m \rangle$$

$$= \alpha + \left\|\frac{1}{4}(w_1 + w_2 + w_3 + w_4)\right\|^2 - \left\langle w_i, \frac{1}{4} \sum_{1 \leq j \leq 4} w_j \right\rangle - \left\langle w_k, \frac{1}{4} \sum_{1 \leq j \leq 4} w_j \right\rangle$$

$$= \alpha + \frac{1}{16}(4 + 12\alpha) - \frac{2}{4}(1 + 3\alpha) = \frac{1}{4}(\alpha - 1).$$

(d) (3 pts) Let $c_1, c_2, c_3, c_4$ be the the corners of a regular tetrahedron in $\mathbb{R}^3$, centered at the origin. I.e., (i) $\|c_i\|^2 = A^2$ for all $i$, (ii) $\langle c_i, c_k \rangle = -A^2/3$ for all $i \neq k$. (As a consequence, $c_1 + c_2 + c_3 + c_4 = 0$.) Let $e_{\text{tetra}}(A)$ denote the error probability of the MAP decoder that observes $Y = c_i + Z$ where $Z$ is $N(0, I_3)$ where each of the four $c_i$’s are equally likely. Express the probability of error of the communication system (the system which uses the
waveforms \( w_1, w_2, w_3, w_4 \) described at the start of the problem in terms of \( \alpha, N_0 \) and \( e_{\text{tetra}}(\cdot) \).

Hint: No lengthy computations are needed.

Solution: First observe that the waveform set \( \tilde{W} = \{ \tilde{w}_1, \ldots, \tilde{w}_4 \} \) is an isometric transformation of \( W = \{ w_1, \ldots, w_4 \} \), hence the error probability of the system using \( W \) is identical to that using \( \tilde{W} \). Observe that \( \tilde{W} \) has dimension 3, since \( \tilde{w}_1 + \cdots + \tilde{w}_4 = 0 \), and let \( \Psi = \{ \psi_1, \psi_2, \psi_3 \} \) be an orthonormal basis for the waveform set.

Given the received signal \( R(t) = \tilde{w}_i(t) + N(t) \), where \( N(t) \) is AWGN with noise intensity \( N_0^2 \), we compute the sufficient statistic \( Y = (\langle R, \psi_1 \rangle, \langle R, \psi_2 \rangle, \langle R, \psi_3 \rangle) = c_i + Z \), where \( c_i = (\langle \tilde{w}_i, \psi_1 \rangle, \langle \tilde{w}_i, \psi_2 \rangle, \langle \tilde{w}_i, \psi_3 \rangle) \) and \( Z = (\langle N, \psi_1 \rangle, \langle N, \psi_2 \rangle, \langle N, \psi_3 \rangle) \sim \mathcal{N}(0, N_0^2 I_3) \).

Define \( \tilde{Y} = \frac{Y}{\sqrt{N_0/2}} = \tilde{c}_i + \tilde{Z} \) with \( \tilde{c}_i = \frac{c_i}{\sqrt{N_0/2}} \) and \( \tilde{Z} = \frac{Z}{\sqrt{N_0/2}} \sim \mathcal{N}(0, I_3) \). Setting \( E = \frac{3}{4}(1 - \alpha) \), we have \( \| \tilde{c}_i \|^2 = \frac{1}{N_0/2} \| \tilde{w}_i \|^2 = \frac{2E}{N_0}, \langle \tilde{c}_i, \tilde{c}_k \rangle = \frac{1}{N_0/2} \langle \tilde{w}_i, \tilde{w}_k \rangle = -\frac{2E}{N_0} \) for all \( i \neq k \), i.e., \( \tilde{c}_i \) are corners of a regular tetrahedron in \( \mathbb{R}^3 \) centered at the origin as described in the problem, with \( A = \sqrt{\frac{2E}{N_0}} = \frac{3(1-\alpha)}{2N_0} \). Hence the error probability of the system is \( e_{\text{tetra}} \left( \sqrt{\frac{3(1-\alpha)}{2N_0}} \right) \).

Remark: This problem is related to the simplex conjecture, which states that the optimal choice of \( M \) signal vectors in AWGN, with an average energy constraint but no constraint on the dimension of the signal set, is the vertices of the \( (M - 1) \)-dimensional regular simplex (e.g., regular tetrahedron for \( M = 4 \), equilateral triangle for \( M = 3 \)). A counter example has been shown for \( M \geq 7 \), and hence this conjecture is not true in general. Refer to M. Steiner, “The strong simplex conjecture is false,” in IEEE Transactions on Information Theory, vol. 40, no. 3, pp. 721-731, May 1994 (available online at https://ieeexplore.ieee.org/abstract/document/335884), for more details.