Problem Set 2 Solutions

Problem 2.1 (Cartesian-product constellations)

(a) Show that if $\mathcal{A}' = \mathcal{A}^K$, then the parameters $N, \log_2 M, E(\mathcal{A}')$ and $K_{\min}(\mathcal{A}')$ of $\mathcal{A}'$ are K times as large as the corresponding parameters of $\mathcal{A}$, whereas the normalized parameters ρ, E_s, E_b and $d^2_{\min}(\mathcal{A})$ are the same as those of $\mathcal{A}$. Verify that these relations hold for $(M \times M)$-QAM constellations.

There are M possibilities for each of the n components of $\mathcal{A}'$, so the total number of possibilities is $|\mathcal{A}'| = M^K$. The number of bits supported by $\mathcal{A}'$ is therefore $\log_2 |\mathcal{A}'| = K \log_2 M$, K times the number for $\mathcal{A}$. The number of dimensions in $\mathcal{A}'$ is KN, so its nominal spectral efficiency is $\rho = (\log_2 M)/N$, the same as that of $\mathcal{A}$.

The average energy of $\mathcal{A}'$ is

$$E_{\mathcal{A}'} = \overline{||\mathbf{a}_1||^2} + \overline{||\mathbf{a}_2||^2} + \cdots + \overline{||\mathbf{a}_K||^2} = KE_{\mathcal{A}}.$$

The average energy per bit is $E_b = E_{\mathcal{A}'}/\log_2 |\mathcal{A}'| = E_{\mathcal{A}}/(\log_2 M)$, and the average energy per two dimensions is $E_s = 2E_{\mathcal{A}'}/KN = 2E_{\mathcal{A}}/N$, which for both are the same as for $\mathcal{A}$.

Two distinct points in $\mathcal{A}'$ must differ in at least one component. The minimum squared distance is therefore the minimum squared distance in any component, which is $d^2_{\min}(\mathcal{A})$.

The number of nearest neighbors to any point $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K) \in \mathcal{A}'$ is the sum of the numbers of nearest neighbors to $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K$, respectively, since there is a nearest neighbor for each such nearest neighbor to each component. The average number $K_{\min}(\mathcal{A}')$ of nearest neighbors to $\mathcal{A}'$ is therefore the sum of the average number of nearest neighbors in each component, which is $K_{\min}(\mathcal{A}') = KK_{\min}(\mathcal{A})$.

An $(M \times M)$-QAM constellation $\mathcal{A}'$ is equal to the 2-fold Cartesian product $\mathcal{A}^2$, where $\mathcal{A}$ is an M-PAM constellation. Therefore all the above results hold with $K = 2$. In particular, the QAM constellation has the same $d^2_{\min}$ as the PAM constellation, but twice the number $K_{\min}$ of nearest neighbors.

(b) Show that if the signal constellation is a Cartesian product $\mathcal{A}^K$, then MD detection can be performed by performing independent MD detection on each of the K components of the received KN-tuple $\mathbf{y} = (y_1, y_2, \ldots, y_K)$. Using this result, sketch the decision regions of a (4×4)-QAM signal set.

Given a received signal $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K)$, to minimize the squared distance $||\mathbf{y}_1 - \mathbf{a}_1||^2 + ||\mathbf{y}_2 - \mathbf{a}_2||^2 + \cdots + ||\mathbf{y}_K - \mathbf{a}_K||^2$ over $\mathcal{A}'$, we may minimize each component $||\mathbf{y}_j - \mathbf{a}_j||^2$ separately, since choice of one component $\mathbf{a}_j$ imposes no restrictions on the choices of other components.

The MD decision regions of a (4×4)-QAM signal set $\mathcal{A}' = \mathcal{A}^2$ are thus simply those of a 4-PAM signal set $\mathcal{A}$, independently for each coordinate. Such decision regions are sketched in the Problem Set 1 solutions, Problem 1.3(b).

(c) Show that if $\Pr(E)$ *is the probability of error for MD detection of* $\mathcal{A}$*, then the probability of error for MD detection of* $\mathcal{A}'$ *is*

$$\Pr(E)' = 1 - (1 - \Pr(E))^K,$$

Show that $\Pr(E)' \approx K \Pr(E)$ *if* $\Pr(E)$ *is small.*

A signal in $\mathcal{A}'$ is received correctly if and only if each component is received correctly. The probability of correct decision is therefore the product of the probabilities of correct decision for each of the components separately, which is $(1-\Pr(E))^K$. The probability of error for $\mathcal{A}'$ is therefore $\Pr(E') = 1-(1-\Pr(E))^K$. When $\Pr(E)$ is small, $(1-\Pr(E))^K \approx 1 - K\Pr(E)$, so $\Pr(E') \approx K\Pr(E)$.

Problem 2.2 ($\Pr(E)$ invariance to translation, orthogonal transformations, or scaling)

Let $\Pr(E \mid \mathbf{a}_j)$ *be the probability of error when a signal* $\mathbf{a}_j$ *is selected equiprobably from an* N*-dimensional signal set* $\mathcal{A}$ *and transmitted over a discrete-time AWGN channel, and the channel output* $\mathbf{Y} = \mathbf{a}_j + \mathbf{N}$ *is mapped to a signal* $\hat{\mathbf{a}}_j \in \mathcal{A}$ *by a minimum-distance decision rule. An error event* E *occurs if* $\hat{\mathbf{a}}_j \neq \mathbf{a}_j$*.* $\Pr(E)$ *denotes the average error probability.*

(a) Show that the probabilities of error $\Pr(E \mid \mathbf{a}_j)$ *are unchanged if* $\mathcal{A}$ *is translated by any vector* $\mathbf{v}$*; i.e., the constellation* $\mathcal{A}' = \mathcal{A} + \mathbf{v}$ *has the same* $\Pr(E)$ *as* $\mathcal{A}$*.*

If all signals are equiprobable and the noise is iid Gaussian, then the optimum detector is a minimum-distance detector.

In this case the received sequence $\mathbf{Y}' = \mathbf{a}'_j + \mathbf{N}$ may be mapped reversibly to $\mathbf{Y} = \mathbf{Y}' - \mathbf{v} = \mathbf{a}_j + \mathbf{N}$, and then an MD detector for $\mathcal{A}$ based on $\mathbf{Y}$ is equivalent to an MD detector for $\mathcal{A}'$ based on $\mathbf{Y}'$. In particular, it has the same probabilities of error $\Pr(E \mid \mathbf{a}_j)$.

(b) Show that $\Pr(E)$ *is invariant under orthogonal transformations; i.e.,* $\mathcal{A}' = U\mathcal{A}$ *has the same* $\Pr(E)$ *as* $\mathcal{A}$ *when* U *is any orthogonal* $N \times N$ *matrix (i.e.,* $U^{-1} = U^T$*).*

In this case the received sequence $\mathbf{Y}' = \mathbf{a}'_j + \mathbf{N}$ may be mapped reversibly to $\mathbf{Y} = U^{-1}\mathbf{Y}' = \mathbf{a}_j + U^{-1}\mathbf{N}$. Since the noise distribution depends only on the squared norm $||\mathbf{n}||^2$, which is not affected by orthogonal transformations, the noise sequence $\mathbf{N}' = U^{-1}\mathbf{N}$ has the same distribution as $\mathbf{N}$, so again the probability of error $\Pr(E)$ is unaffected.

(c) Show that $\Pr(E)$ *is unchanged if both* $\mathcal{A}$ *and* $\mathbf{N}$ *are scaled by* $\alpha > 0$*.*

In this case the received sequence $\mathbf{Y}' = \mathbf{a}'_j + \alpha\mathbf{N}$ may be mapped reversibly to $\mathbf{Y} = \alpha^{-1}\mathbf{Y}' = \mathbf{a}_j + \mathbf{N}$, which again reduces the model to the original scenario.

Problem 2.3 (optimality of zero-mean constellations)

Consider an arbitrary signal set $\mathcal{A} = \{\mathbf{a}_j, 1 \leq j \leq M\}$*. Assume that all signals are equiprobable. Let* $\mathbf{m}(\mathcal{A}) = \frac{1}{M}\sum_j \mathbf{a}_j$ *be the average signal, and let* $\mathcal{A}'$ *be* $\mathcal{A}$ *translated by* $\mathbf{m}(\mathcal{A})$ *so that the mean of* $\mathcal{A}'$ *is zero:* $\mathcal{A}' = \mathcal{A} - \mathbf{m}(\mathcal{A}) = \{\mathbf{a}_j - \mathbf{m}(\mathcal{A}), 1 \leq j \leq M\}$*. Let* $E(\mathcal{A})$ *and* $E(\mathcal{A}')$ *denote the average energies of* $\mathcal{A}$ *and* $\mathcal{A}'$*, respectively.*

(a) Show that the error probability of an MD detector is the same for $\mathcal{A}'$ *as it is for* $\mathcal{A}$*.*

This follows from Problem 2.2(a)

(b) Show that $E(\mathcal{A}') = E(\mathcal{A}) - ||\mathbf{m}(\mathcal{A})||^2$. Conclude that removing the mean $\mathbf{m}(\mathcal{A})$ is always a good idea.

Let A be the random variable with alphabet $\mathcal{A}$ that is equal to $\mathbf{a}_j$ with probability $1/M$ for all j, and let $A' = A - \mathbf{m}(\mathcal{A})$ be the *fluctuation* of A. Then $E(\mathcal{A}) = \overline{||A||^2}$, $\mathbf{m}(\mathcal{A}) = \overline{A}$, and

$$E(\mathcal{A}') = \overline{||A - \overline{A}||^2} = \overline{||A||^2} - 2\langle \overline{A}, \overline{A} \rangle + ||\overline{A}||^2 = \overline{||A||^2} - ||\overline{A}||^2 = E(\mathcal{A}) - ||\mathbf{m}(\mathcal{A})||^2.$$

In other words, the second moment of A is greater than or equal to the variance of A, with equality if and only if the mean of A is zero.

This result and that of part (a) imply that if $\mathbf{m}(\mathcal{A}) \neq \mathbf{0}$, then by replacing $\mathcal{A}$ with $\mathcal{A}' = \mathcal{A} - \mathbf{m}(\mathcal{A})$, the average energy can be reduced without changing the probability of error. Therefore $\mathcal{A}'$ is always preferable to $\mathcal{A}$; *i.e.*, an optimum constellation must have zero mean.

(c) Show that a binary antipodal signal set $\mathcal{A} = \{\pm \mathbf{a}\}$ is always optimal for $M = 2$.

A two-point constellation with zero mean must have $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$, which implies $\mathbf{a}_2 = -\mathbf{a}_1$.

Problem 2.4 (Non-equiprobable signals).

Let $\mathbf{a}_j$ and $\mathbf{a}_{j'}$ be two signals that are not equiprobable. Find the optimum (MPE) pairwise decision rule and pairwise error probability $\Pr\{\mathbf{a}_j \to \mathbf{a}_{j'}\}$.

The MPE rule is equivalent to the *maximum-a-posteriori-probability* (MAP) rule: choose the $\hat{\mathbf{a}} \in \mathcal{A}$ such that $p(\hat{\mathbf{a}} \mid \mathbf{y})$ is maximum among all $p(\mathbf{a}_j \mid \mathbf{y}), \mathbf{a}_j \in \mathcal{A}$. By Bayes' law,

$$p(\mathbf{a}_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{a}_j) p(\mathbf{a}_j)}{p(\mathbf{y})}.$$

The pairwise decision rule is thus to choose $\mathbf{a}_j$ over $\mathbf{a}_{j'}$ if $p(\mathbf{y} \mid \mathbf{a}_j)p(\mathbf{a}_j) > p(\mathbf{y} \mid \mathbf{a}_{j'})p(\mathbf{a}_{j'})$, or *vice versa*. Using the logarithm of the noise pdf, we can write this as

$$\frac{-||\mathbf{y} - \mathbf{a}_j||^2}{2\sigma^2} + \log p(\mathbf{a}_j) > \frac{-||\mathbf{y} - \mathbf{a}_{j'}||^2}{2\sigma^2} + \log p(\mathbf{a}_{j'}),$$

or equivalently

$$||\mathbf{y} - \mathbf{a}_j||^2 < ||\mathbf{y} - \mathbf{a}_{j'}||^2 + K,$$

where $K = 2\sigma^2 \log p(\mathbf{a}_j)/p(\mathbf{a}_{j'})$. Therefore the pairwise MAP rule is equivalent to a minimum-squared-distance rule with a bias K.

Following the development shown in (5.2), we have

$$\begin{array}{rcl} ||\mathbf{y} - \mathbf{a}_j||^2 & - & ||\mathbf{y} - \mathbf{a}_{j'}||^2 = \\ -2\langle \mathbf{y}, \mathbf{a}_j \rangle + ||\mathbf{a}_j||^2 & - & (-2\langle \mathbf{y}, \mathbf{a}_{j'} \rangle + ||\mathbf{a}_{j'}||^2) = \\ 2\langle \mathbf{y}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle & - & 2\langle \mathbf{m}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle, \end{array}$$

where $\mathbf{m}$ denotes the midvector $\mathbf{m} = (\mathbf{a}_j + \mathbf{a}_{j'})/2$.

Therefore $||\mathbf{y} - \mathbf{a}_j||^2 - ||\mathbf{y} - \mathbf{a}_{j'}||^2 < K$ if and only if $2\langle \mathbf{y}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle - 2\langle \mathbf{m}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle < K$. Since the magnitudes of the projections $\mathbf{y}_{|\mathbf{a}_{j'} - \mathbf{a}_j}$ and $\mathbf{m}_{|\mathbf{a}_{j'} - \mathbf{a}_j}$ of $\mathbf{y}$ and $\mathbf{m}$ onto the difference vector $\mathbf{a}_{j'} - \mathbf{a}_j$ are

$$|\mathbf{y}_{|\mathbf{a}_{j'} - \mathbf{a}_j}| = \frac{\langle \mathbf{y}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle}{||\mathbf{a}_{j'} - \mathbf{a}_j||}; \quad |\mathbf{m}_{|\mathbf{a}_{j'} - \mathbf{a}_j}| = \frac{\langle \mathbf{m}, \mathbf{a}_{j'} - \mathbf{a}_j \rangle}{||\mathbf{a}_{j'} - \mathbf{a}_j||},$$

we have $||\mathbf{y} - \mathbf{a}_j||^2 - ||\mathbf{y} - \mathbf{a}_{j'}||^2 < K$ if and only if

$$|\mathbf{y}_{|\mathbf{a}_{j'} - \mathbf{a}_j}| < |\mathbf{m}_{|\mathbf{a}_{j'} - \mathbf{a}_j}| + \frac{K}{2||\mathbf{a}_{j'} - \mathbf{a}_j||} = |\mathbf{m}_{|\mathbf{a}_{j'} - \mathbf{a}_j}| + \frac{\sigma^2 \lambda(\mathbf{a}_j, \mathbf{a}_{j'})}{d(\mathbf{a}_j, \mathbf{a}_{j'})},$$

where $\lambda(\mathbf{a}_j, \mathbf{a}_{j'})$ is the log likelihood ratio $\log p(\mathbf{a}_j)/p(\mathbf{a}_{j'})$.

The conclusion is that the decision boundary is still a hyperplane perpendicular to the difference vector $\mathbf{a}_{j'} - \mathbf{a}_j$, but shifted by $\sigma^2 \lambda(\mathbf{a}_j, \mathbf{a}_{j'})/d(\mathbf{a}_j, \mathbf{a}_{j'})$.

The probability of error is thus the probability that a one-dimensional Gaussian variable of zero mean and variance σ^2 will exceed $d(\mathbf{a}_j, \mathbf{a}_{j'})/2 + \sigma^2 \lambda(\mathbf{a}_j, \mathbf{a}_{j'})/d(\mathbf{a}_j, \mathbf{a}_{j'})$. This is given as always by the Q function

$$\Pr\{\mathbf{a}_j \to \mathbf{a}_{j'}\} = Q\left(\frac{d(\mathbf{a}_j, \mathbf{a}_{j'})}{2\sigma} + \frac{\sigma^2 \lambda(\mathbf{a}_j, \mathbf{a}_{j'})}{d(\mathbf{a}_j, \mathbf{a}_{j'})}\right).$$

Problem 2.5 (UBE for M-PAM constellations).

For an M-PAM constellation $\mathcal{A}$, show that $K_{\min}(\mathcal{A}) = 2(M-1)/M$. Conclude that the union bound estimate of $\Pr(E)$ *is*

$$\Pr(E) \approx 2\left(\frac{M-1}{M}\right) Q\left(\frac{d}{2\sigma}\right).$$

Observe that in this case the union bound estimate is exact. Explain why.

The $M-2$ interior points have 2 nearest neighbors, while the 2 boundary points have 1 nearest neighbor, so the average number of nearest neighbors is

$$K_{\min}(\mathcal{A}) = \frac{1}{M}\left((M-2)(2) + (2)(1)\right) = \frac{2M-2}{M}.$$

In general the union bound estimate is $\Pr(E) \approx K_{\min}(\mathcal{A}) Q(d_{\min}(\mathcal{A})/2\sigma)$. Plugging in $d_{\min}(\mathcal{A}) = d$ and the above expression for $K_{\min}(\mathcal{A})$, we get the desired expression.

For the $M-2$ interior points, the exact error probability is $2Q(d/2\sigma)$. For the 2 boundary points, the exact error probability is $Q(d/2\sigma)$. If all points are equiprobable, then the average $\Pr(E)$ is exactly the UBE given above.

In general, we can see that the union bound is exact for one-dimensional constellations, and only for one-dimensional constellations. The union bound estimate is therefore exact only for equi-spaced one-dimensional constellations; *i.e.,* essentially only for M-PAM.