Problem Set 1 Solutions

Problem 1.1 (Compound interest and dB)

How long does it take to double your money at an interest rate of $P\%$? The bankers' "Rule of 72" estimates that it takes about $72/P$ years; e.g., *at a 5% interest rate compounded annually, it takes about 14.4 years to double your money.*

(a) An engineer decides to interpolate the dB table above linearly for $1 \le 1 + p \le 1.25$; i.e.,

$$\text{ratio or multiplicative factor of } 1 + p \leftrightarrow 4p \text{ dB}.$$

Show that this corresponds to a "Rule of 75;" e.g., *at a 5% interest rate compounded annually, it takes 15 years to double your money.*

If your money compounds at a rate of x dB per year, then since doubling your money corresponds to a 3 dB gain, it will take about $Y = 3/x$ years to double your money.

The engineer's approximation to the dB table is a linear approximation that is exact near $p = 0.25$ ($P = 25\%$), where $p = P/100$. Under this approximation, it will take

$$Y = \frac{3}{4p} = \frac{75}{P}$$

years to double your money. Thus this approximation corresponds to a "Rule of 75." For example, it estimates that it takes $Y = 75/5 = 15$ years to double your money when $P = 5\%$.

(b) A mathematician linearly approximates the dB table for $p \approx 0$ by noting that as $p \to 0$, $\ln(1 + p) \to p$, and translates this into a "Rule of N" for some real number N. What is N? Using this rule, how many years will it take to double your money at a 5% interest rate, compounded annually? What happens if interest is compounded continuously?

As $p \to 0$, we have

$$10 \log_{10}(1 + p) = (10 \log_{10} e) \ln(1 + p) = 4.34 \ln(1 + p) \to 4.34p,$$

where we change the base of the logarithm from 10 to e (recall that $x = 10^{\log_{10} x} = e^{\ln x}$, so $\log_{10} x = (\log_{10} e)(\ln x)$), we read $10 \log_{10} e = 4.34$ from the dB table, and we use $\ln(1 + p) \to p$ as $p \to 0$. Thus we obtain a linear approximation

$$\text{ratio or multiplicative factor of } 1 + p \leftrightarrow 4.34p \text{ dB},$$

which becomes exact as $p \to 0$. This linear approximation translates to the estimate that it takes

$$Y = \frac{10 \log_{10} 2}{(10 \log_{10} e)p} = \frac{3.01}{4.34p} = \frac{69.35}{P}$$

years to double your money, or a "Rule of 69.35." For example, this rule estimates that it takes $Y = 69.35/5 = 13.87$ years to double your money when $P = 5\%$.

A more precise calculation using natural logarithms yields

$$Y = \frac{\ln 2}{p} = \frac{69.31}{P},$$

or a "Rule of 69.31," which estimates that it takes $Y = 69.31/5 = 13.86$ years to double your money when $P = 5\%$.

Remark. Note that this "Rule of 69.31" becomes exact when interest is compounded continuously, so that after Y years your money has increased by a factor of e^{pY}, rather than the factor of $(1+p)^Y$ that you get when interest is compounded annually.

(c) How many years will it actually take to double your money at a 5% interest rate, compounded annually? [Hint: $10 \log_{10} 7 = 8.45$ *dB.] Whose rule best predicts the correct result?*

Since $1.05 = 21/20$, a factor of 1.05 is equivalent to

$$10 \log_{10} 7 + 10 \log_{10} 3 - 10 \log_{10} 10 - 10 \log_{10} 2 = 8.45 + 4.77 - 10 - 3.01 = 0.21 \text{ dB}.$$

Thus it actually takes

$$Y = \frac{3.01}{0.21} = 14.33$$

years to double your money when interest is compounded annually.

We see that this estimate is quite close to the estimate of $Y = 14.4$ years given by the "Rule of 72." Thus the "Rule of 72" is equivalent to a linear approximation of the dB table that is exact near $P = 5\%$. This is the range in which the "Rule of 72" is commonly used. The "Rule of 72" also has the advantage that 72 has many integer divisors (*e.g.,* 2, 3, 4, 6, 8, 9, 12, ...), so that its estimate of Y is an easily calculated integer for many common interest rates. So in this instance the bankers have been rather clever.

Problem 1.2 (Biorthogonal codes)

A $2^m \times 2^m$ $\{\pm 1\}$*-valued Hadamard matrix* H_{2^m} *may be constructed recursively as the* m*-fold tensor product of the* 2×2 *matrix*

$$H_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix},$$

as follows:

$$H_{2^m} = \begin{bmatrix} +H_{2^{m-1}} & +H_{2^{m-1}} \\ +H_{2^{m-1}} & -H_{2^{m-1}} \end{bmatrix}.$$

(a) Show by induction that:

(i) $(H_{2^m})^T = H_{2^m}$, *where* ^{T} *denotes the transpose; i.e.,* H_{2^m} *is symmetric;*

(ii) The rows or columns of H_{2^m} *form a set of mutually orthogonal vectors of length* 2^m*;*

(iii) The first row and the first column of H_{2^m} *consist of all* $+1$*s;*

(iv) There are an equal number of $+1$s and -1s in all other rows and columns of H_{2^m};

(v) $H_{2^m}H_{2^m} = 2^m I_{2^m}$; i.e., $(H_{2^m})^{-1} = 2^{-m}H_{2^m}$, where $^{-1}$ denotes the inverse.

We first verify that (i)-(v) hold for H_2. We then suppose that (i)-(v) hold for $H_{2^{m-1}}$. We can then conclude by induction that:

(i)

$$(H_{2^m})^T = \begin{bmatrix} +(H_{2^{m-1}})^T & +(H_{2^{m-1}})^T \\ +(H_{2^{m-1}})^T & -(H_{2^{m-1}})^T \end{bmatrix} = \begin{bmatrix} +H_{2^{m-1}} & +H_{2^{m-1}} \\ +H_{2^{m-1}} & -H_{2^{m-1}} \end{bmatrix} = H_{2^m}.$$

(ii) The first 2^{m-1} rows of H_{2^m} are of the form $\mathbf{h}_j = (\mathbf{g}_j, \mathbf{g}_j)$, where $\mathbf{g}_j$ is the corresponding row of $H_{2^{m-1}}$, and the second 2^{m-1} rows of H_{2^m} are of the form $\mathbf{h}_{j+2^{m-1}} = (\mathbf{g}_j, -\mathbf{g}_j)$. Suppose that the rows $\mathbf{g}_j$ of $H_{2^{m-1}}$ are mutually orthogonal. Then the inner product $\langle \mathbf{h}_j, \mathbf{h}_{j'} \rangle$ is 0 whenever $j \neq j'$ and $j \neq j' \pm 2^{m-1}$, because the inner product is the sum of the inner products of the first half-rows and the second half-rows, which are both zero. If $j = j' - 2^{m-1}$, then the inner product is

$$\langle \mathbf{h}_j, \mathbf{h}_{j'} \rangle = \langle \mathbf{g}_j, \mathbf{g}_j \rangle + \langle \mathbf{g}_j, -\mathbf{g}_j \rangle = \langle \mathbf{g}_j, \mathbf{g}_j \rangle - \langle \mathbf{g}_j, \mathbf{g}_j \rangle = 0,$$

and similarly $\langle \mathbf{h}_j, \mathbf{h}_{j'} \rangle = 0$ when $j = j' + 2^{m-1}$. Thus $\langle \mathbf{h}_j, \mathbf{h}_{j'} \rangle = 0$ whenever $j \neq j'$, so the rows of H_{2^m} form a set of mutually orthogonal vectors. Since $(H_{2^m})^T = H_{2^m}$, so do the columns.

(iii) The first row of H_{2^m} is $\mathbf{h}_0 = (\mathbf{g}_0, \mathbf{g}_0)$, where $\mathbf{g}_0$ is the first row of $H_{2^{m-1}}$. By the inductive hypothesis, $\mathbf{g}_0 = (+1, +1, \ldots, +1)$, so $\mathbf{h}_0 = (+1, +1, \ldots, +1)$; *i.e.,* all columns of H_{2^m} have a $+1$ as their first component. Since $(H_{2^m})^T = H_{2^m}$, so do all the rows.

(iv) The remaining rows of H_{2^m} are orthogonal to $\mathbf{h}_0$ by (ii), and thus must have an equal number of $+1$s and -1s. Since $(H_{2^m})^T = H_{2^m}$, so must the remaining columns.

(v) Since $(H_{2^m})^T = H_{2^m}$, the matrix $H_{2^m}H_{2^m} = H_{2^m}(H_{2^m})^T$ is the matrix of inner products of rows of H_{2^m}. By (ii), all off-diagonal elements of this matrix are zero. The diagonal elements are the squared norms $||\mathbf{h}_j||^2 = 2^m$, since $\mathbf{h}_j$ is a vector of length 2^m in which each component has squared norm 1. Thus $H_{2^m}H_{2^m} = 2^m I_{2^m}$.

(b) A biorthogonal signal set is a set of real equal-energy orthogonal vectors and their negatives. Show how to construct a biorthogonal signal set of size 64 as a set of $\{\pm 1\}$-valued sequences of length 32.

By (a)(ii), the rows of H_{32} form a set O of 32 orthogonal $\{\pm 1\}$-valued sequences of length 32, each with energy 32. It follows that the rows of H_{32} and their negatives form a biorthogonal set $B = \pm O$ of 64 $\{\pm 1\}$-valued sequences of length 32.

(c) A simplex signal set S is a set of real equal-energy vectors that are equidistant and that have zero mean $\mathbf{m}(S)$ under an equiprobable distribution. Show how to construct a simplex signal set of size 32 as a set of 32 $\{\pm 1\}$-valued sequences of length 31. [Hint: The fluctuation $O - \mathbf{m}(O)$ of a set O of orthogonal real vectors is a simplex signal set.]

As in (b), the rows of H_{32} form a set O of 32 orthogonal equal-energy (and therefore equidistant) $\{\pm 1\}$-valued sequences of length 32. By (a)(iii)-(iv), the mean of O is $\mathbf{m}(O) = (+1, 0, 0, \dots, 0)$, since all rows have $+1$ in the first column and an equal number of $+1$s and -1s in the remaining columns. Thus $S = O - \mathbf{m}(O)$ is a zero-mean, equal-energy and equidistant set of 32 row vectors of length 32 which have 0 in the first coordinate and the elements of the rows of H_{32} in the remaining coordinates. Since the first coordinate always has value 0, it may be deleted without affecting any norms, distances or inner products; in particular, the vectors remain zero-mean, equal-energy and equidistant. Thus we obtain a simplex signal set S' consisting of a set of 32 $\{\pm 1\}$-valued sequences of length 31.

(d) Let $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ be the received sequence on a discrete-time AWGN channel, where the input sequence $\mathbf{X}$ is chosen equiprobably from a biorthogonal signal set B of size 2^{m+1} constructed as in part (b). Show that the following algorithm implements a minimum-distance decoder for B (i.e., given a real 2^m-vector $\mathbf{y}$, it finds the closest $\mathbf{x} \in B$ to $\mathbf{y}$):

(i) Compute $\mathbf{z} = H_{2^m}\mathbf{y}$, where $\mathbf{y}$ is regarded as a column vector;

(ii) Find the component z_j of $\mathbf{z}$ with largest magnitude $|z_j|$;

(iii) Decode to $\mathrm{sgn}(z_j)\mathbf{x}_j$, where $\mathrm{sgn}(z_j)$ is the sign of the largest-magnitude component z_j and $\mathbf{x}_j$ is the corresponding column of H_{2^m}.

Given $\mathbf{y}$, minimizing the squared distance $||\mathbf{y}-\mathbf{x}||^2$ over $\mathbf{x} \in B$ is equivalent to maximizing the inner product $\langle \mathbf{y}, \mathbf{x} \rangle$, since

$$||\mathbf{y} - \mathbf{x}||^2 = ||\mathbf{y}||^2 - 2\langle \mathbf{y}, \mathbf{x} \rangle + ||\mathbf{x}||^2,$$

and $||\mathbf{x}||^2$ is equal to a constant (2^m) for all $\mathbf{x} \in B$. The vector $\mathbf{z} = H_{2^m}\mathbf{y}$ is the set of inner products $\langle \mathbf{y}, \mathbf{x} \rangle$ as $\mathbf{x}$ runs through the 2^m rows of H_{2^m}. The set of inner products $\langle \mathbf{y}, \mathbf{x} \rangle$ as $\mathbf{x}$ runs through the 2^{m+1} elements of B are therefore just the elements of $\mathbf{z}$ and their negatives. Maximizing $\langle \mathbf{y}, \mathbf{x} \rangle$ is therefore equivalent to finding the element z_j of $\mathbf{z}$ with largest magnitude $|z_j|$, and deciding on the corresponding row $\mathbf{x}_j$ of H_{2^m} if the sign of z_j is positive, or on $-\mathbf{x}_j$ if the sign is negative.

Remark. Note that the matrix multiplication $\mathbf{z} = H_{2^m}\mathbf{y}$ corresponds to implementing a bank of matched filters, one for each of the rows of H_{2^m}, which form the set of correlations of $\mathbf{y}$ with each of the rows of H_{2^m}. Since the rows of H_{2^m} span the signal space $\mathcal{S} \supset B$, by the theorem of irrelevance (see Chapter 2) the outputs $\mathbf{z}$ of this bank of matched filters form a set of sufficient statistics for detection of a signal $\mathbf{x} \in \mathcal{S}$ in the presence of AWGN. In this case we have been able to show directly that we can find an optimal decision rule based on $\mathbf{z}$ which is very simple.

(e) Show that a circuit similar to that shown in Figure 1 below for $m = 2$ can implement the $2^m \times 2^m$ matrix multiplication $\mathbf{z} = H_{2^m}\mathbf{y}$ with a total of only $m \times 2^m$ addition and subtraction operations. (This is called the "fast Hadamard transform," or "Walsh transform," or "Green machine.")

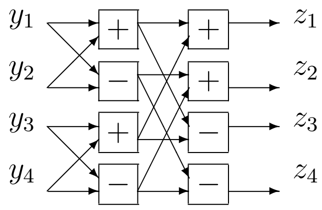


Figure 1. Fast $2^m \times 2^m$ Hadamard transform for $m = 2$.

This circuit is based on the following recursion for $\mathbf{z} = H_4\mathbf{y}$:

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} +H_2 & +H_2 \\ +H_2 & -H_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} +H_2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + H_2 \begin{bmatrix} y_3 \\ y_4 \end{bmatrix} \\ +H_2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - H_2 \begin{bmatrix} y_3 \\ y_4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} y'_1 \\ y'_2 \end{bmatrix} + \begin{bmatrix} y'_3 \\ y'_4 \end{bmatrix} \\ \begin{bmatrix} y'_1 \\ y'_2 \end{bmatrix} - \begin{bmatrix} y'_3 \\ y'_4 \end{bmatrix} \end{bmatrix}.$$

In other words, we first group the elements of $\mathbf{y}$ into pairs $(y_1, y_2), (y_3, y_4), \ldots$. The first set of arithmetic elements computes the 2×2 Walsh-Hadamard transform of each pair; *e.g.,*

$$\begin{bmatrix} y'_1 \\ y'_2 \end{bmatrix} = H_2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + y_2 \\ y_1 - y_2 \end{bmatrix}.$$

We then group the elements of $\mathbf{y}'$ into pairs $(y'_1, y'_3), (y'_2, y'_4), \ldots$, and again compute the 2×2 Walsh-Hadamard transform of each pair; *e.g.,*

$$\begin{bmatrix} z_1 \\ z_3 \end{bmatrix} = H_2 \begin{bmatrix} y'_1 \\ y'_3 \end{bmatrix} = \begin{bmatrix} y'_1 + y'_3 \\ y'_1 - y'_3 \end{bmatrix}.$$

Thus we can compute the 4×4 Walsh-Hadamard transform $\mathbf{z} = H_4\mathbf{y}$ by computing two stages of two 2×2 Walsh-Hadamard transforms, as illustrated in Figure 1.

Similarly, we can compute a $2^m \times 2^m$ Walsh-Hadamard transform $\mathbf{z} = H_{2^m}\mathbf{y}$ using the recursion

$$\begin{bmatrix} \mathbf{z}_0 \\ \mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} +H_{2^{m-1}} & +H_{2^{m-1}} \\ +H_{2^{m-1}} & -H_{2^{m-1}} \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} = \begin{bmatrix} H_{2^{m-1}}\mathbf{y}_0 + H_{2^{m-1}}\mathbf{y}_1 \\ H_{2^{m-1}}\mathbf{y}_0 - H_{2^{m-1}}\mathbf{y}_1 \end{bmatrix},$$

by computing two $2^{m-1} \times 2^{m-1}$ Walsh-Hadamard transforms, and then combining their outputs in one more stage involving 2^{m-1} 2×2 Walsh-Hadamard transforms. If each $2^{m-1} \times 2^{m-1}$ Walsh-Hadamard transform requires $m - 1$ stages of 2^{m-2} 2×2 Walsh-Hadamard transforms, then the $2^m \times 2^m$ Walsh-Hadamard transform requires m stages of 2^{m-1} 2×2 Walsh-Hadamard transforms. Each 2×2 transform requires one addition and one subtraction, so a total of only $m \times 2^{m-1} \times 2$ additions and subtractions is required.

Thus the complexity of an $M = 2^m$-point Walsh-Hadamard transform is only of the order of $M \log_2 M$, rather than M^2. This is why this algorithm is called "fast."

Problem 1.3 (16-QAM signal sets)

Three 16-point 2-dimensional quadrature amplitude modulation (16-QAM) signal sets are shown in Figure 2, below. The first is a standard 4×4 signal set; the second is the V.29 signal set; the third is based on a hexagonal grid and is the most power-efficient 16-QAM signal set known. The first two have 90° symmetry; the last, only 180°. All have a minimum squared distance between signal points of $d^2_{\min} = 4$.

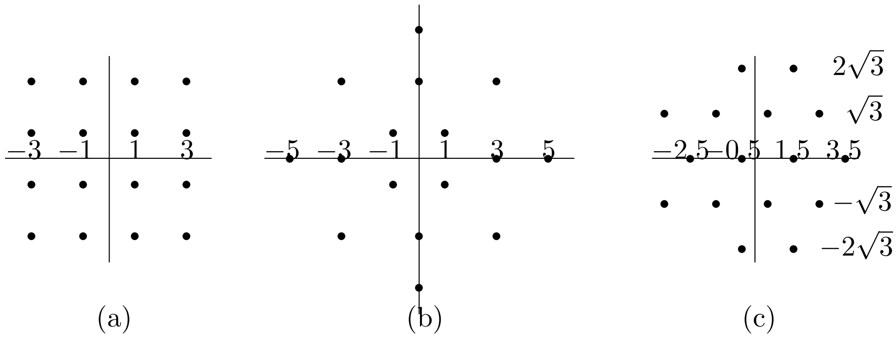


Figure 2. 16-QAM signal sets. (a) (4×4)-QAM; (b) V.29; (c) hexagonal.

(a) Compute the average energy (squared norm) of each signal set if all points are equiprobable. Compare the power efficiencies of the three signal sets in dB.

In Figure 2(a), there are 4 points with squared norm 2, 8 with squared norm 10, and 4 with squared norm 18, so

$$E_a = \frac{1}{4}2 + \frac{1}{2}10 + \frac{1}{4}18 = 10 \text{ (10.00 dB)}.$$

Alternatively, both coordinates have equal probability of having squared norm 1 or 9, so the average energy per coordinate is 5, and thus $E_a = 10$.

In Figure 2(b), there are 4 points with squared norm 2, 4 with squared norm 9, 4 with squared norm 18, and 4 with squared norm 25, so

$$E_b = \frac{1}{4}2 + \frac{1}{4}9 + \frac{1}{4}18 + \frac{1}{4}25 = 13.5 \text{ (11.30 dB)}.$$

Thus this signal set is 1.3 dB less power-efficient than that of Figure 2(a).

In Figure 2(c), there is 1 point with squared norm 1/4, 1 with 9/4, 2 with 13/4, 2 with 21/4, 1 with 25/4, 2 with 37/4, 3 with 49/4, 2 with 57/4 and 2 with 61/4, so

$$E_c = \frac{1 + 9 + 26 + 42 + 25 + 74 + 147 + 114 + 122}{4 \times 16} = 8.75 \text{ (9.42 dB)}.$$

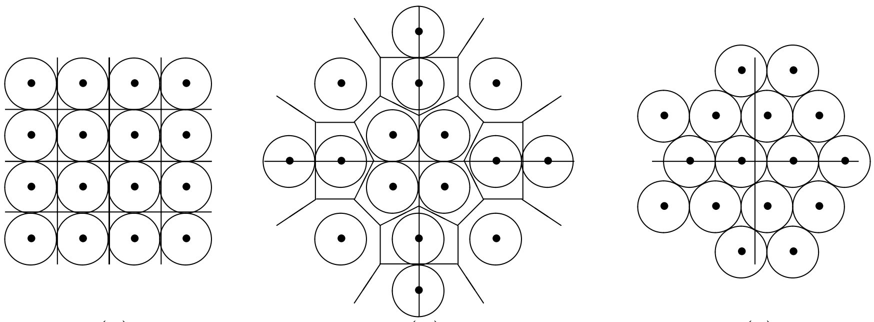
Thus this signal set is about 0.6 dB more power-efficient than that of Figure 2(a).

Remark. A more elegant way of doing this calculation is to shift the origin by $(1/2, 0)$ to the least-energy point. The resulting signal set has mean $(1/2, 0)$, and 1 point with squared norm 0, 6 with 4, 6 with 12, and 3 with 16, for an average of 9. Subtracting the squared norm 1/4 of its mean, we get $E_c = 8.75$ (9.42 dB) for the zero-mean signal set of Figure 2(c).

(b) Sketch the decision regions of a minimum-distance detector for each signal set.

The minimum-distance decision regions are sketched below for Figures 2(a) and 2(b). For Figure 2(c), the decision regions for signals in the interior of the constellation are hexagons (and are hard to draw in LaTeX).

Note that in Figure 2(a) the minimum-distance decision regions correspond to two independent minimum-distance decisions on each 4-PAM coordinate.



(a) (b) (c)

We also draw 2-spheres (circles) of radius 1 about each signal point. The fact that these 2-spheres are disjoint except where they kiss (at their points of tangency) shows that $d_{\min} = 2$. This makes it obvious that 2(a) is more densely packed (more power-efficient) than 2(b), and in turn that 2(c) is more densely packed than 2(a). In fact, 2(c) might well be what we would come up with if we took 16 pennies and tried to pack them together as densely as possible.

Since Gaussian noise is circularly symmetric and its probability density decreases exponentially with distance, the dominant types of errors in AWGN will be those that occur at the points of tangency of these 2-spheres, which lie on decision region boundaries.

Remark. Figure 2(b) could obviously be made somewhat more power-efficient without changing its essential character by somewhat decreasing the radii of its outer points until their associated 2-spheres become tangent to innermore 2-spheres.

(c) Show that with a phase rotation of $\pm 10°$ the minimum distance from any rotated signal point to any decision region boundary is substantially greatest for the V.29 signal set.

The question here is: if each signal point $\mathbf{x} = (x, y)$ is rotated by a small angle θ to a rotated point $\mathbf{x}' = (x \cos\theta - y \sin\theta, y \cos\theta + x \sin\theta)$, what is the worst-case reduction in distance to the nearest decision boundary?

For small phase rotations ($|\theta| \le 10°$), we may use the approximations $\cos\theta \approx 1$, $\sin\theta \approx \theta$, or $\mathbf{x}' \approx (x - \theta y, y + \theta x)$.

For Figure 2(a), a rotation of either a point of type $(3, 1)$ or of type $(3, 3)$ by a small angle θ therefore reduces the distance to the nearest decision boundary by approximately 3θ, or by approximately 0.5 when $|\theta| \approx 10°$. Thus the minimum distance to a decision boundary is cut approximately by a factor of 2, or the minimum squared distance by a factor of 4, which we will see later amounts to a reduction of about 6 dB in signal-to-noise margin.

For Figure 2(b), all of the outer points have enough distance from their nearest decision boundaries in the tangential direction so that the minimum distance is still at least 1 after a 10° rotation. For example, a point of type $(3,0)$ rotates approximately to $(3,0.5)$, which is still distance 1 from the nearest point $(3,1.5)$ on the decision boundary. The worst case is therefore an inner point of type $(1,1)$, whose minimum distance of 1 is reduced by about 0.17 by a 10° rotation. Since $(0.83)^2$ is about 1.6 dB, this amounts to about a 1.6 dB reduction in signal-to-noise margin. Thus even though the Figure 2(b) signal set has 1.3 worse signal-to-noise margin to start with, in the presence of uncompensated $\pm 10^\circ$ phase rotations ("phase jitter") it becomes more than 3 dB better than Figure 2(a).

For Figure 2(c), it is clear that the outer points are affected by phase rotations similarly to the outer points of 2(a). For example, a point of type $(2.5, \sqrt{3})$ has squared norm 9.25 and thus radius 3.04. A phase rotation of θ moves it directly by an amount 3.04θ toward its nearest decision boundary, so as in 2(a) a rotation of about $\theta = 10^\circ$ cuts the distance to the nearest decision boundary by a factor of about 2, for a reduction in SNR margin of about 6 dB.

Problem 1.4 (Shaping gain of spherical signal sets)

In this exercise we compare the power efficiency of n-cube and n-sphere signal sets for large n.

An n-cube signal set is the set of all odd-integer sequences of length n within an n-cube of side $2M$ centered on the origin. For example, the signal set of Figure 2(a) is a 2-cube signal set with $M = 4$.

An n-sphere signal set is the set of all odd-integer sequences of length n within an n-sphere of squared radius r^2 centered on the origin. For example, the signal set of Figure 3(a) is also a 2-sphere signal set for any squared radius r^2 in the range $18 \le r^2 < 25$. In particular, it is a 2-sphere signal set for $r^2 = 64/\pi = 20.37$, where the area πr^2 of the 2-sphere (circle) equals the area $(2M)^2 = 64$ of the 2-cube (square) of the previous paragraph.

Both n-cube and n-sphere signal sets therefore have minimum squared distance between signal points $d^2_{\min} = 4$ (if they are nontrivial), and n-cube decision regions of side 2 and thus volume 2^n associated with each signal point. The point of the following exercise is to compare their average energy using the following large-signal-set approximations:

- *The number of signal points is approximately equal to the volume $V(\mathcal{R})$ of the bounding n-cube or n-sphere region $\mathcal{R}$ divided by 2^n, the volume of the decision region associated with each signal point (an n-cube of side 2).*
- *The average energy of the signal points under an equiprobable distribution is approximately equal to the average energy $E(\mathcal{R})$ of the bounding n-cube or n-sphere region $\mathcal{R}$ under a uniform continuous distribution.*

(a) Show that if $\mathcal{R}$ is an n-cube of side $2M$ for some integer M, then under the two above approximations the approximate number of signal points is M^n and the approximate average energy is $nM^2/3$. Show that the first of these two approximations is exact.

The first approximation is that the number m_{cube} of signal points is approximately

$$m_{\text{cube}} \approx \frac{V(\mathcal{R})}{2^n} = \frac{(2M)^n}{2^n} = M^n.$$

This approximation is exact, because it can be seen that an n-cube constellation of side $2M$ is simply the n-fold Cartesian product $\mathcal{A}^n = \{(x_1, x_2, \ldots, x_n) \mid x_k \in \mathcal{A}\}$ of an M-PAM costellation $\mathcal{A} = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$, the set of all odd integers in the interval $[-M, M]$. (For example, Figure 2(a) is the 2-fold Cartesian product $\mathcal{A}^2$ of a 4-PAM constellation $\mathcal{A}$.)

The second approximation is that the average energy E_{cube} is approximately

$$E_{\text{cube}} \approx E(\mathcal{R}) = n(M^2/3),$$

where we observe that the average energy over an n-cube $\mathcal{R}$ of side $2M$ under an equiprobable distribution $p(\mathbf{x})$, whose marginals are the uniform distributions $p(x_k) = 1/2M$, is

$$E(\mathcal{R}) = \int_{\mathcal{R}} ||\mathbf{x}||^2 p(\mathbf{x}) d\mathbf{x} = \sum_{k=1}^{n} \int_{-M}^{M} x_k^2 \; p(x_k) dx_k = n\frac{2M^3}{3}\frac{1}{2M} = n\frac{M^2}{3}.$$

(In this case the exact expression is

$$E_{\text{cube}} = n\frac{M^2 - 1}{3},$$

since the constellation $\mathcal{A}^n$ is the n-fold Cartesian product of an M-PAM constellation $\mathcal{A}$ whose average energy is $E_{\mathcal{A}} = (M^2 - 1)/3$.)

(b) For n even, if $\mathcal{R}$ is an n-sphere of radius r, compute the approximate number of signal points and the approximate average energy of an n-sphere signal set, using the following known expressions for the volume $V_{\otimes}(n, r)$ and the average energy $E_{\otimes}(n, r)$ of an n-sphere of radius r:

$$\begin{aligned} V_{\otimes}(n, r) &= \frac{(\pi r^2)^{n/2}}{(n/2)!}; \\ E_{\otimes}(n, r) &= \frac{nr^2}{n+2}. \end{aligned}$$

The first approximation is that the number m_{sphere} of signal points is approximately

$$m_{\text{sphere}} \approx \frac{V(\mathcal{R})}{2^n} = \frac{(\pi r^2)^{n/2}}{2^n (n/2)!}.$$

The second approximation is that the average energy E_{sphere} is approximately

$$E_{\text{sphere}} \approx E(\mathcal{R}) = \frac{nr^2}{n+2}.$$

(c) For $n = 2$, show that a large 2-sphere signal set has about 0.2 dB smaller average energy than a 2-cube signal set with the same number of signal points.

In general, in n dimensions, to make $m_{\text{cube}} = m_{\text{sphere}}$ we choose M and r so that

$$M^n = \frac{(\pi r^2)^{n/2}}{2^n (n/2)!};$$

i.e.,

$$M^2 = \frac{\pi r^2}{2^2((n/2)!)^{2/n}}.$$

Then the ratio of the average energy of the n-cube to that of the n-sphere is

$$\frac{E_{\text{cube}}}{E_{\text{sphere}}} = \frac{n+2}{nr^2}\frac{nM^2}{3} = \frac{\pi(n+2)}{12((n/2)!)^{2/n}}.$$

For example, for $n = 2$, setting the volumes equal, $(2M)^2 = \pi r^2$, we have

$$\frac{E_{\text{cube}}}{E_{\text{sphere}}} = \frac{\pi}{3} \text{ (0.20 dB)}.$$

Thus in two dimensions using a large circular rather than square constellation saves only about 0.2 dB in power efficiency.

(d) For $n = 16$, show that a large 16-sphere signal set has about 1 dB smaller average energy than a 16-cube signal set with the same number of signal points. [Hint: $8! = 40320$ (46.06 dB).]

For $n = 16$, however, we have

$$\frac{E_{\text{cube}}}{E_{\text{sphere}}} = \frac{18\pi}{12(8!)^{1/8}} = \frac{3\pi}{2(40320)^{1/8}} = (4.77 + 4.97 - 3.01 - \frac{1}{8}46.06) = 0.97 \text{ dB};$$

i.e., using a 16-sphere rather than 16-cube constellation saves nearly 1 dB in power efficiency (signal-to-noise margin).

(e) Show that as $n \to \infty$ a large n-sphere signal set has a factor of $\pi e/6$ (1.53 dB) smaller average energy than an n-cube signal set with the same number of signal points. [Hint: Use Stirling's approximation, $m! \to (m/e)^m$ as $m \to \infty$.]

Using the hint, as $n \to \infty$ we have

$$((n/2)!)^{2/n} \to \frac{n}{2e}.$$

Therefore

$$\frac{E_{\text{cube}}}{E_{\text{sphere}}} = \frac{\pi(n+2)}{12((n/2)!)^{2/n}} \to \frac{\pi e(n+2)}{6n} \to \frac{\pi e}{6} \text{ (1.53 dB)}.$$

Since the ratio is monotonically increasing, we conclude that the greatest possible gain in any number of dimensions is 1.53 dB.