Performance Analysis of PAM and QAM over AWGN Channels

Erik G. Ström

February 17, 2015

1 Introduction

This document should be viewed as a complement to Chapter 5 in [1]. Also included in this document is a short introduction to random variables in Appendix A.

2 White Gaussian Noise

The channel noise $n(t)$ is very often modeled as additive white Gaussian noise (AWGN). Hence, we think of $n(t)$ for every time instance t as a Gaussian random variable — this is what makes $n(t)$ a Gaussian random process[1]. White Gaussian noise is zero-mean, $\mathrm{E}[n(t)] = 0$, and a stationary random process, the latter implying that the autocorrelation function for $n(t)$,

$$\mathrm{E}[n(t)n(t+\tau)]$$

does not depend on t. In fact, the autocorrelation is

$$R_n(\tau) = \mathrm{E}[n(t)n(t+\tau)] = \frac{N_0}{2}\delta(\tau) \tag{1}$$

where $\delta(\tau)$ is the Dirac delta function. Due to the delta function, we see that the variance of the random variable $n(t)$, i.e., $R_n(0)$, is infinite for any time instance t, and that any two random variables $n(t_1)$ and $n(t_2)$ for $t_1 \neq t_2$ are uncorrelated,

[1]We are here not so careful about the exact defintion of a Gaussian random process, and we will not make a notational difference between the random process and its realization. To be more notationally strict, we could, e.g., denote the random process by $\{N(t)\}$ and its realization by $n(t)$, as is done in [1, Chapter 3.2].

since the crosscorrelation between $n(t_1)$ and $n(t_2)$ is

$$\begin{aligned}
\mathrm{E}[n(t_1)n(t_2)] &= R_n(t_2 - t_1) \\
&= \frac{N_0}{2}\delta(t_2 - t_1) \\
&= 0, \qquad t_1 \neq t_2.
\end{aligned} \tag{2}$$

These properties make white Gaussian noise a very bad noise process. In fact, one can show that additive white Gaussian noise is the worst-case noise among all noise distributions.

Interestingly enough, even though the variance of $n(t)$ is infinite, the variance of filtered noise $w(t) = n(t) * h(t)$ is finite, as long as the energy for the filter impulse response, $h(t)$, is finite,

$$E_h = \int_{-\infty}^{\infty} h^2(t)\, dt < \infty$$

It is easily shown that

$$\begin{aligned}
\mathrm{E}[w(t)] &= \mathrm{E}[n(t) * h(t)] \\
&= \mathrm{E}\left[\int_{-\infty}^{\infty} n(u)h(t-u)\, du\right] \\
&= \int_{-\infty}^{\infty} \mathrm{E}[n(u)]h(t-u)\, du = 0
\end{aligned} \tag{3}$$

and

$$\begin{aligned}
\mathrm{E}[w^2(t)] &= \mathrm{E}\left[\int_{-\infty}^{\infty} n(u)h(t-u)\, du \int_{-\infty}^{\infty} n(v)h(t-v)\, dv\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{E}[n(u)n(v)]h(t-u)h(t-v)\, dudv \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{N_0}{2}\delta(v-u)h(t-u)h(t-v)\, dudv && (4) \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty} h^2(t-u)\, du && (5) \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty} h^2(v)\, dv \\
&= \frac{N_0}{2}E_h, && (6)
\end{aligned}$$

where (4) follows from (2), and (5) follows after integration over v and using properties of the delta function.

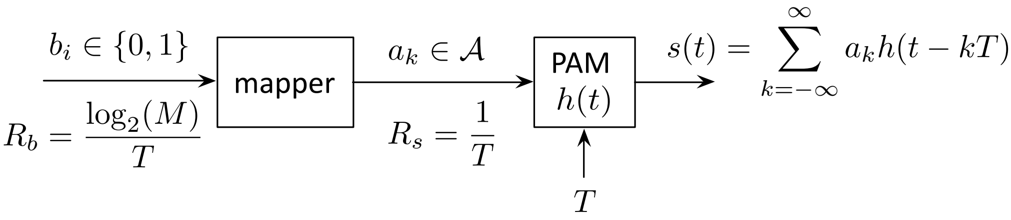

Figure 1: PAM transmitter.

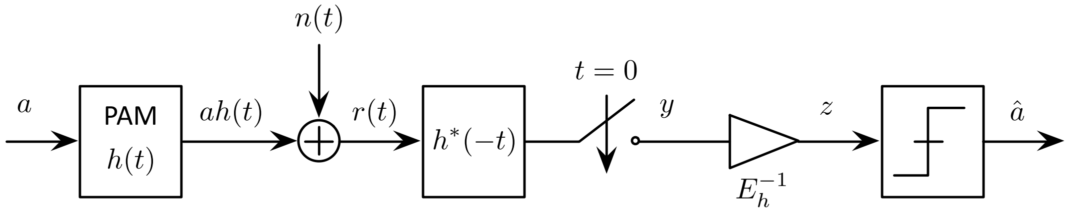

Figure 2: One-shot PAM transmitter, AWGN channel, and a one-shot minimum-distance receiver

Since filtering is a linear operation on the input, we have that $w(t)$ is a linear combination of Gaussian random variables

$$w(t) = \underbrace{\int_{-\infty}^{\infty}}_{\text{"sum"}} \underbrace{h(u)}_{\text{of scaled}} \underbrace{n(t-u)}_{\text{Gaussian r.v.'s}} \, du.$$

Hence, $w(t)$ is a Gaussian random variable, whose pdf is determined by its mean $\mathrm{E}[w(t)] = 0$ and variance $\mathrm{E}[w^2(t)] = E_h N_0/2$.

3 Error Probability for Baseband M-PAM

A general pulse-amplitude modulation (PAM) transmitter is depicted in Fig. 1. The mapper converts blocks of $\log_2 M$ consecutive bits $b_i \in \{0,1\}$ to the kth amplitude $a_k \in \mathcal{A}$, where i is the bit time index and k is the symbol time index. The mapping is one-to-one. Hence, since $\log_2 M$ bits can take on M different patterns, the alphabet $\mathcal{A}$ must consist of M amplitude values, i.e., $\mathcal{A} = \{A_1, A_2, \ldots, A_M\}$.

In the following, we will assume that the transmitted pulses $h(t-kT)$ are orthogonal, i.e., $\langle h(t-kT), h(t-lT \rangle = 0$ when $k \neq l$ (see Appendix B). In this case, we can detect each transmitted symbol independent of the others. We will therefore consider the one-shot, baseband PAM system in Figure 2. Since we are dealing with baseband PAM, we can assume that all signals and variables are real-valued. Hence, the pulse shape $h(t)$ is real-valued and $h^*(-t) = h(-t)$. We see that the received signal is

$$r(t) = ah(t) + n(t)$$

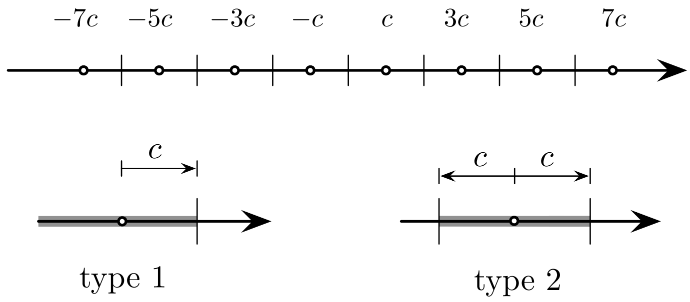

Figure 3: An M-PAM constellation and decision boundaries for $M = 8$ (top) and the two types of decision regions (bottom), Type 1 for the end constellation points and Type 2 for the inner constellation points.

where a is an M-ary PAM symbol, i.e., a member of the alphabet

$$\mathcal{A} = \{\pm c, \pm 3c, \dots, \pm(M-1)c\}.$$

From Appendix C, we know that c^2 is related to the energy for the alphabet E_a as

$$c^2 = \frac{3E_a}{M^2 - 1}.$$

From Figure 2, we see that

$$\begin{aligned} z &= E_h^{-1} y \\ &= E_h^{-1} [r(t) * h^*(-t)]|_{t=0} \\ &= E_h^{-1} \int_{-\infty}^{\infty} r(u) h^*(u)\, du \\ &= E_h^{-1} \int_{-\infty}^{\infty} a h(u) h^*(u) + n(u) h^*(u)\, du \\ &= a + n, \end{aligned}$$

where we can express n as $n = E_h^{-1} w(0)$, where $w(t) = n(t) * h^*(-t)$. From (3) and (6), we know that $w(0) \sim \mathcal{N}(0, E_h N_0/2)$, where E_h is the energy of the impulse response $h^*(-t)$, which is the same as the energy for $h(t)$. Hence, n is Gaussian with mean

$$\mathrm{E}[n] = \mathrm{E}[w(0)] E_h^{-1} = 0$$

and variance

$$\sigma^2 = \mathrm{E}[n^2] = \mathrm{E}[w^2(0)] E_h^{-2} = \frac{N_0}{2E_h}.$$

The M-PAM constellation diagram is depicted in Figure 3. As seen there are two types of decision regions: Type 1 and Type 2. Type 1 is applicable for the two end points and Type 2 for the $M-2$ interior points. The probability of

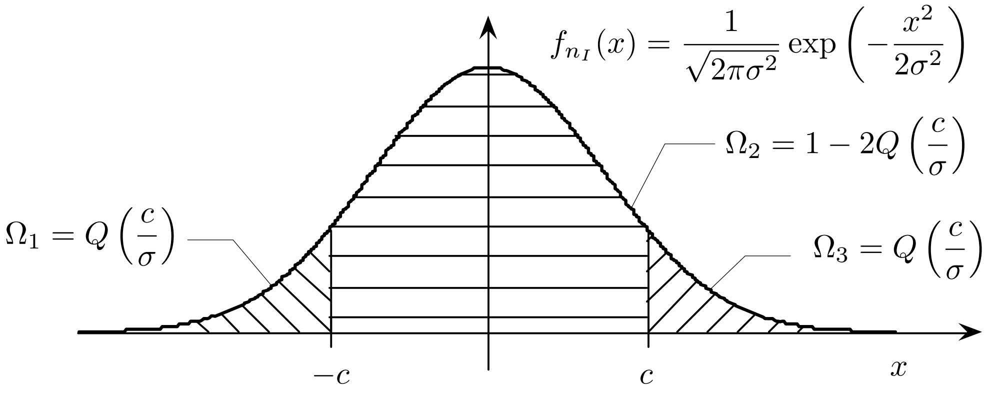

Figure 4: The probability density function for n is a zero mean Gaussian distribution with variance σ^2. The areas Ω_1, Ω_2, and Ω_3 represent the probabilities $\Omega_1 = \Pr\{n_I < -c\}$, $\Omega_2 = \Pr\{-c < n_I < c\}$, and $\Omega_3 = \Pr\{n_I > c\}$.

error, conditioned on that a Type 1 signal (end point) is sent is the probability that the noise carries the transmitted point outside the decision region,

$$P_{e|1} = \Pr\{n > c\} = \Pr\left\{\frac{n}{\sigma} > \frac{c}{\sigma}\right\} = Q\left(\frac{c}{\sigma}\right) = Q\left(\sqrt{\frac{c^2}{\sigma^2}}\right).$$

The probability $\Pr\{n > c\} = \Omega_3$ is illustrated in Figure 4.

To proceed, we recall that the energy per PAM symbol is $E_s = E_a E_h$, which yields

$$\frac{c^2}{\sigma^2} = \frac{3E_a}{M^2-1}\frac{2E_h}{N_0} = \frac{6}{M^2-1}\frac{E_s}{N_0},$$

and

$$P_{e|1} = Q\left(\sqrt{\frac{6}{M^2-1}\frac{E_s}{N_0}}\right).$$

The probability of error conditioned on that a Type 2 signal (interior point) was sent is (see Figure 4)

$$P_{e|2} = \Pr\{n < -c\} + \Pr\{n > c\} = 2\Pr\{n > c\} = 2P_{e|1}.$$

If we assume that all signals are sent with equal probability, i.e., with probability $1/M$, then the average error probability is

$$P_e = \frac{1}{M}[2P_{e|1} + (M-2)P_{e|2}] = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{6}{M^2-1}\frac{E}{N_0}}\right). \quad (7)$$

It is easily verified that (7) for $M = 4$ is the same as Eq. (5.104) in [1], which is the error probability for passband 4-PAM. This is no coincidence, since the error

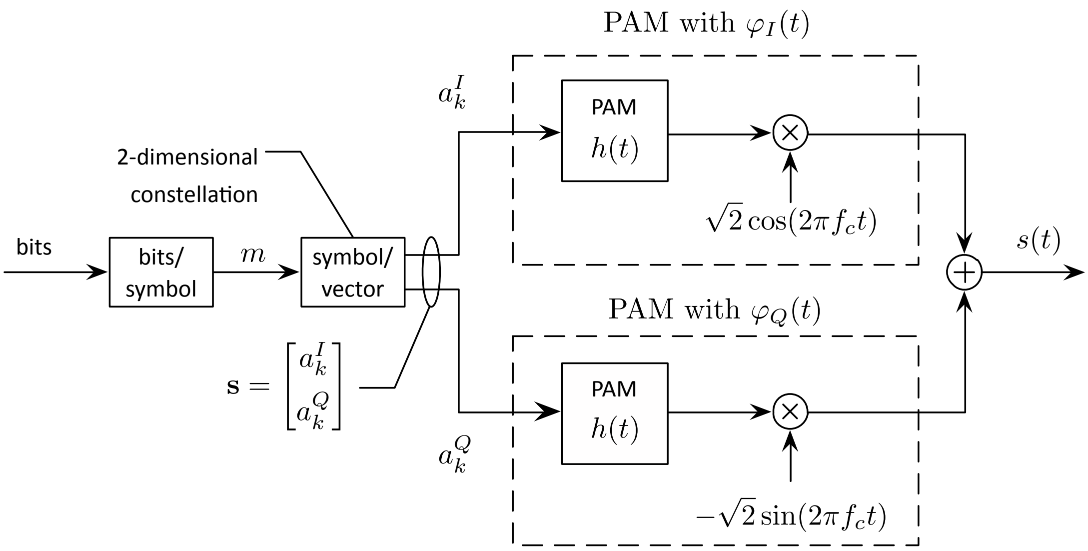

Figure 5: Generic IQ-modulator

probability for baseband PAM is exactly the same as for passband PAM. The latter conclusion follows since the decision variable z has the same statistics in both cases.

4 Quadrature Modulation

A block diagram for a generic inphase quadrature phase modulator (IQ-modulator) is found in Figure 5. We see that the transmitted signal can be viewed as the sum of two PAM processes with different pulse shapes: $\varphi_I(t)$ in the top (in-phase) branch and $\varphi_Q(t)$ in the lower (quadrature) branch, where

$$\begin{aligned}\varphi_I(t) &= h(t)\sqrt{2}\cos(2\pi f_c t)\\ \varphi_Q(t) &= -h(t)\sqrt{2}\sin(2\pi f_c t).\end{aligned}$$

The pulse shape and the carrier frequency f_c is chosen such that the power spectral density of the transmitted signal will fit the frequency response of the channel.[2]

As implied by the IQ-modulator block diagram, the signals $\varphi_I(t)$ and $\varphi_Q(t)$ span a two-dimensional signal space. In fact, it can be shown that signals are

[2]The astute reader will notice that the in-phase signal $\sqrt{2}\cos(2\pi f_c t)\sum_k a_k^I h(t-kT)$ is a PAM process with the pulse shape $\varphi_I(t)$ only if $\sqrt{2}\cos(2\pi f_c t)h(t-kT) = \varphi_I(t-kT)$, i.e., if $f_c = n/T$ for some integer n. The same argument can be made for the quadrature-phase signal $-\sqrt{2}\sin(2\pi f_c t)\sum_k a_k^Q h(t-kT)$. We will tacitly assume that $f_c = n/T$, which is not restrictive, since the main conclusions in this section can be shown to be valid also when $f_c \neq n/T$.

approximately orthogonal if the carrier frequency f_c is much larger than the bandwidth of $h(t)$. Let us show this for the special case when

$$h(t) = \begin{cases} A, & 0 \le t < T \\ 0, & \text{otherwise} \end{cases}.$$

The energy for $h(t)$ is $E_h = A^2T$ and the bandwidth is approximately $W = 1/T$. Now,

$$\begin{aligned} E_{\varphi_Q} &= \int_{-\infty}^{\infty} \varphi_Q^2(t)\, dt = \int_{-\infty}^{\infty} h^2(t) 2\sin^2(2\pi f_c t) dt = A^2 \int_0^T 2\sin^2(2\pi f_c t) dt \\ &= A^2 T - \frac{A^2 \sin(4\pi f_c T)}{4\pi f_c} = E_h \left(1 - \frac{\sin(4\pi f_c T)}{4\pi f_c T}\right). \end{aligned}$$

We see that the second term is is exactly zero when $2f_cT$ is an integer and small for large f_cT. We note that the condition, $f_cT \gg 1$ is the same as $f_c \gg 1/T = W$. We conclude that $E_{\varphi_Q} \approx E_h$ and, by using a very similar argument, we can show that $E_{\varphi_I} \approx E_h$. Finally,

$$\begin{aligned} \langle \varphi_Q(t), \varphi_I(t) \rangle &= \int_{-\infty}^{\infty} \varphi_Q(t) \varphi_I(t)\, dt \\ &= -\int_{-\infty}^{\infty} h^2(t) 2 \sin(2\pi f_c t) \cos(2\pi f_c t)\, dt \\ &= -A^2 \int_0^T \sin(4\pi f_c t)\, dt \\ &= A^2 \frac{\cos(4\pi f_c T) - 1}{4\pi f_c} \\ &= E_h \frac{\sin^2(2\pi f_c T)}{4\pi f_c T}. \end{aligned}$$

Again, we see that the inner product is exactly zero when $2f_cT$ is an nonzero integer and very small compared to E_h when $f_cT \gg 1$, i.e., when $f_c \gg W$.

The assumption that $f_c \gg W$ is valid for most practical systems, and we will therefore in the following assume that $\varphi_I(t)$ and $\varphi_Q(t)$ are orthogonal and that $E_{\varphi_I} = E_{\varphi_Q} = E_h$.

A generic demodulator is shown in Figure 6. We will assume that the channel noise is white and Gaussian with power spectral density $N_0/2$. We will now show that the noise vector elements, n_k^I and n_k^Q, are independent, identically distributed (iid) Gaussian random variables with zero mean and variance $\sigma^2 = N_0/(2E_h)$. The key observation is that n_k^I and n_k^Q are samples of filtered white Gaussian noise. This implies that n_k^I and n_k^Q are Gaussian and, since the white noise is

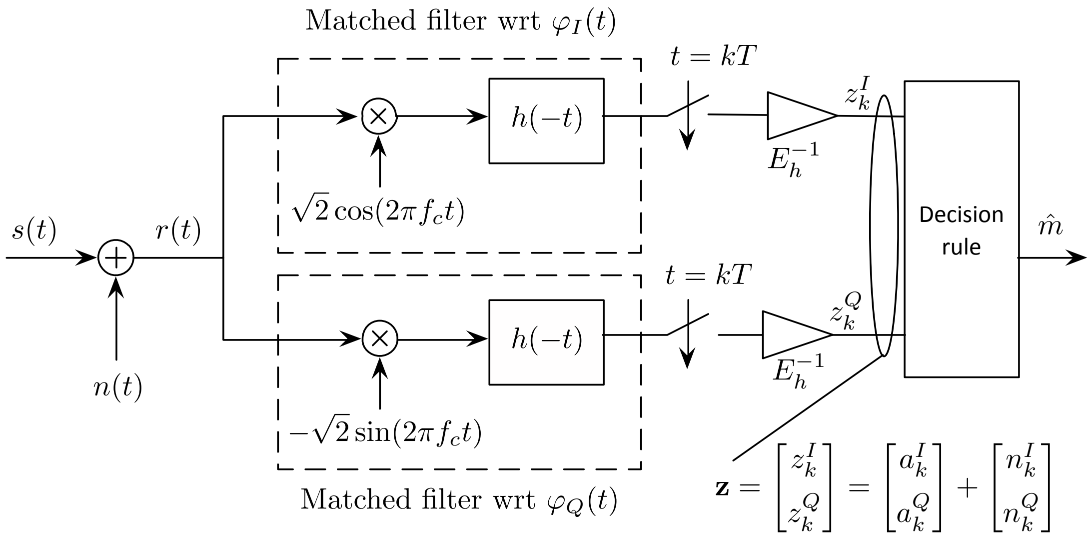

Figure 6: Generic IQ-demodulator

zero-mean, that n_k^I and n_k^Q, have zero means. To show independence, it is enough to show that n_k^I and n_k^Q are uncorrelated. Since n_k^I and n_k^Q have zero means, we need to show that $\mathrm{E}[n_k^Q n_k^I] = 0$,

$$
\begin{aligned}
\mathrm{E}[n_k^Q n_k^I] &= \mathrm{E}\left[\frac{1}{E_h}\int_{-\infty}^{\infty} n(u)\varphi_I(u)\,du \frac{1}{E_h}\int_{-\infty}^{\infty} n(v)\varphi_Q(v)\,dv\right] \\
&= \frac{1}{E_h^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{E}[n(u)n(v)]\varphi_I(u)\varphi_Q(v)\,dudv \\
&= \frac{1}{E_h^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{N_0}{2}\delta(u-v)\varphi_I(u)\varphi_Q(v)\,dudv \\
&= \frac{1}{E_h^2}\frac{N_0}{2}\int_{-\infty}^{\infty} \varphi_I(u)\varphi_Q(u)\,du \\
&= \frac{1}{E_h^2}\frac{N_0}{2}\langle\varphi_I(t), \varphi_Q(t)\rangle = 0.
\end{aligned}
$$

Hence, n_k^I and n_k^Q are independent. We can go through the same steps to derive the variance of n_k^I and n_k^Q by replacing $\varphi_Q(v)$ with $\varphi_I(v)$ or vice versa. Indeed,

$$
\begin{aligned}
\mathrm{E}[(n_k^Q)^2] &= \frac{1}{E_h^2}\frac{N_0}{2}\underbrace{\langle\varphi_Q(t), \varphi_Q(t)\rangle}_{=E_h} = \frac{N_0}{2E_h} \\
\mathrm{E}[(n_k^I)^2] &= \frac{1}{E_h^2}\frac{N_0}{2}\underbrace{\langle\varphi_I(t), \varphi_I(t)\rangle}_{=E_h} = \frac{N_0}{2E_h}.
\end{aligned}
$$

Let us consider the transmission of the signal vector $\mathbf{s}_k = [a_k^I \quad a_k^Q]^T$, which

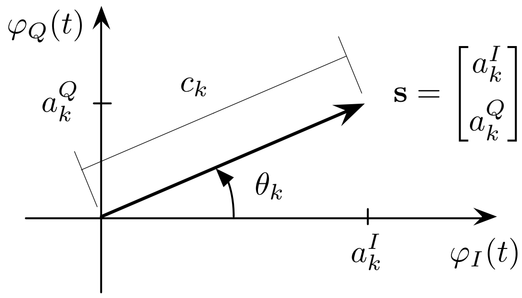

Figure 7: Geometric interpretation of c_k and θ_k.

implies that the transmitted signal is

$$\begin{aligned} s_k(t) &= a_k^I \varphi_I(t) + a_k^Q \varphi_Q(t) \\ &= a_k^I \sqrt{2} h(t) \cos(2\pi f_c t) - a_k^Q \sqrt{2} h(t) \sin(2\pi f_c t). \end{aligned}$$

If we define c_k and θ_k as (see Figure 7 for a geometrical interpretation)

$$\begin{aligned} c_k &= \|\mathbf{s}_k\| = \sqrt{(a_k^I)^2 + (a_k^Q)^2}, \\ \tan\theta_k &= \frac{a_k^Q}{a_k^I}, \end{aligned}$$

then we can write

$$\begin{aligned} a_k^I &= c_k \cos(\theta_k) \\ a_k^Q &= c_k \sin(\theta_k). \end{aligned}$$

Hence, the transmitted signal can written as

$$\begin{aligned} s_k(t) &= c_k \cos(\theta_k) h(t) \sqrt{2} \cos(2\pi f_c t) - c_k \sin(\theta_k) h(t) \sqrt{2} \sin(2\pi f_c t) \\ &= c_k h(t) \sqrt{2} \cos(2\pi f_c t + \theta_k), \end{aligned}$$

where we have used the trigonometric identity $\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) = \cos(\alpha + \beta)$.

We conclude that the transmitted signal can be seen as the pulse $h(t)$ multiplied with a cosine-carrier, where the amplitude and phase of the carrier are determined by a_k^I and a_k^Q. Hence, the transmitted information can affect both the amplitude and phase of the transmitted signal.

5 Common IQ Signal Constellations

We can choose the signal constellation such that the amplitude is the same for all signal alternatives by placing the signal vectors on a circle in the signal space.

This implies that $c_1 = c_2 = \cdots = c_M$, and the transmitted information is then carried by the phase of the carrier. Phase-shift keying (PSK) is an example of such a modulation scheme.

Conversely, if the signal vectors are placed on a straight line that crosses the origin in the signal space, then the carrier phase will be the same for all signal alternatives. (Which follows from the fact that $a_k^I / s_{Q,k}$ is the same for all k.) Bandpass PAM is an example of such a modulation scheme.

The general case, when both amplitude and phase are allowed to change between signal alternatives, is called quadrature amplitude modulation (QAM). There exist many QAM constellations, but we will limit our discussion to the case when the signal points are placed on a regular rectangular grid in the signal space, see Figure 8 for two examples.

A general rectangular M-ary QAM constellation has M_I amplitudes along the inphase axis and M_Q amplitudes along the quadrature axis. We will sometimes use the notation $(M_I \times M_Q)$-QAM to denote this type of constellation. Now, we have $M = M_I M_Q$, and

$$\begin{aligned} a_k^I &\in \mathcal{A}_I = \{\pm c, \pm 3c, \ldots, \pm(M_I - 1)c\} \\ a_k^Q &\in \mathcal{A}_Q = \{\pm c, \pm 3c, \ldots, \pm(M_Q - 1)c\}. \end{aligned}$$

When M is an even square, i.e., when $\sqrt{M}$ is an integer, we can set $M_I = M_Q = \sqrt{M}$. In Figure 8, we see a square 16-QAM constellation and a rectangular (4×2)-QAM constellation.

The energy for an $(M_I \times M_Q)$-QAM constellation, E_a, can be shown to be the sum for the energies for the PAM-constellations along the I and Q directions, see Appendix C. Hence,

$$E_a = E_{a,I} + E_{a,Q} = c^2 \frac{M_I^2 - 1}{3} + c^2 \frac{M_Q^2 - 1}{3}$$

and we can therefore relate c^2 to the constellation energy as

$$c^2 = E_a \frac{3}{M_1^2 + M_Q^2 - 2}$$

and for the special case when $M_I = M_Q = \sqrt{M}$, we have

$$c^2 = E_a \frac{3}{2(M-1)}.$$

As seen from Figure 8, the signal vectors are spaced with the distance $2c$ along the axes. Hence, the minimum distance of the constellation is $d_{\min} = 2c$.

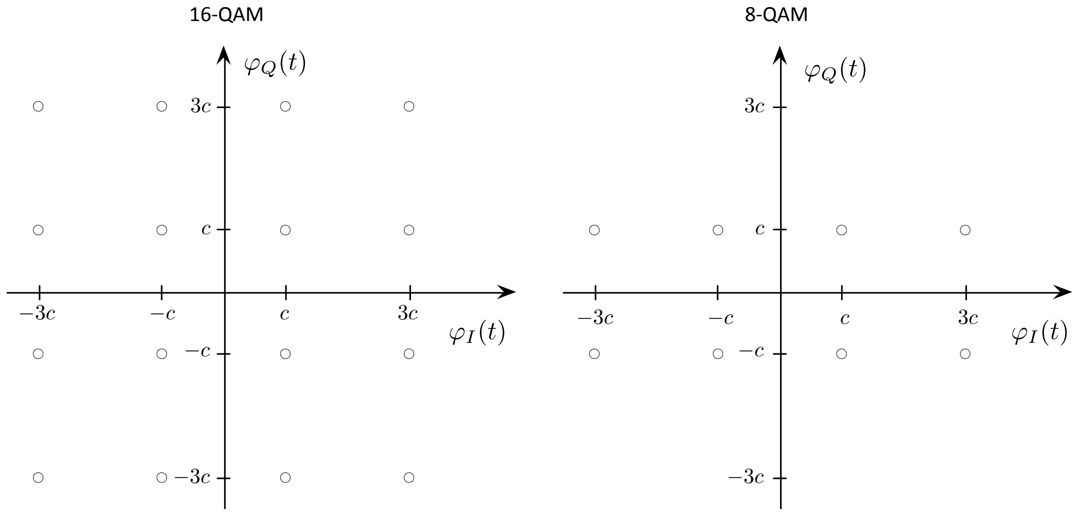

Figure 8: Constellations for square 16-QAM and rectangular (4×2)-QAM

6 Error Probability for M-ary QAM

If we assume that $h(t)$ is chosen such that the transmitted pulses are orthogonal to each other, i.e.,

$$\begin{aligned}
&\langle \varphi_I(t-kT), \varphi_I(t-lT)\rangle = \langle \varphi_Q(t-kT), \varphi_Q(t-lT)\rangle = 0, \qquad k \neq l \\
&\langle \varphi_I(t-kT), \varphi_Q(t-lT)\rangle = 0,
\end{aligned}$$

then we can detect the symbols for different k independent of each other. Hence, the symbol error probability is independent of k and we will remove k from the notation. For example, n_k^I will be denoted by n_I.

Regardless if M is an even square or not, the minimum-distance (maximum likelihood) decision regions will be rectangular-shaped. The regions can be squares (type 1), squares with one open side (type 2) or squares with two open sides (type 3), see Figure 9.

To compute the symbol error probability, we therefore need only to compute the conditional error probabilities for the three types of decision regions. Conditioned on that we send a symbol that has a decision region of type x, the probability of wrong decision is denoted $P_{e|x}$ and the probability of correct decision is denoted $P_{c|x}$. Since the elements of the noise vector, n_I and n_Q, are independent Gaussian random variables with zero mean and variance σ^2, we can write

$$\begin{aligned}
P_{c|1} &= \Pr\{-c < n_I < c, -c < n_Q < c\} \\
&= \Pr\{-c < n_I < c\} \Pr\{-c < n_Q < c\} \\
&= [\Pr\{-c < n_I < c\}]^2.
\end{aligned}$$

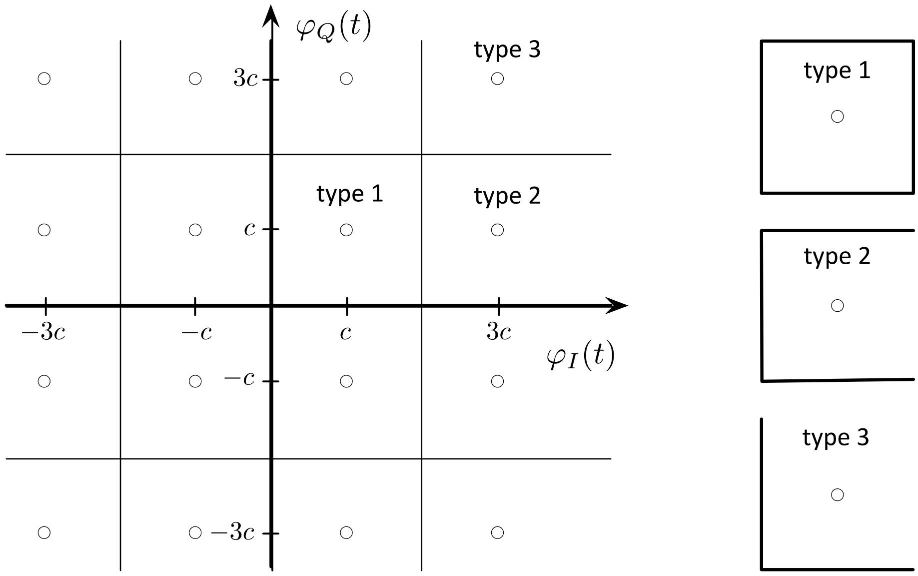

Figure 9: Minimum-distance (ML) receiver decision regions for 16-QAM

From Figure 4, we see that $\Pr\{-c < n_I < c\}$ can be written in terms of the Q-function as

$$\Pr\{-c < n_I < c\} = \Omega_2 = 1 - 2Q\left(\frac{c}{\sigma}\right).$$

Hence,

$$P_{e|1} = 1 - \left[1 - 2Q\left(\frac{c}{\sigma}\right)\right]^2 = 4Q\left(\frac{c}{\sigma}\right) - 4Q^2\left(\frac{c}{\sigma}\right). \tag{8}$$

Proceeding with the type 2 region, we note that

$$\begin{aligned}
P_{c|2} &= \Pr\{n_I > -c, -c < n_Q < c\} \\
&= \Pr\{n_I > -c\}\Pr\{-c < n_Q < c\} \\
&= (1-\Omega_1)\Omega_2 \\
&= \left[1 - Q\left(\frac{c}{\sigma}\right)\right]\left[1 - 2Q\left(\frac{c}{\sigma}\right)\right] \\
&= 1 - 3Q\left(\frac{c}{\sigma}\right) + 2Q^2\left(\frac{c}{\sigma}\right),
\end{aligned}$$

and

$$P_{e|2} = 3Q\left(\frac{c}{\sigma}\right) - 2Q^2\left(\frac{c}{\sigma}\right). \tag{9}$$

Finally,

$$\begin{aligned}
P_{c|3} &= \Pr\{-c < n_I, -c < n_Q\} \\
&= \Pr\{n_I > -c\} \Pr\{n_Q > -c\} \\
&= (1 - \Omega_1)(1 - \Omega_1) \\
&= \left[1 - Q\left(\frac{c}{\sigma}\right)\right]^2 \\
&= 1 - 2Q\left(\frac{c}{\sigma}\right) + Q^2\left(\frac{c}{\sigma}\right),
\end{aligned}$$

and

$$P_{e|3} = 2Q\left(\frac{c}{\sigma}\right) - Q^2\left(\frac{c}{\sigma}\right). \tag{10}$$

Suppose that all symbols are transmitted with the same probability $1/M$. If n_x denotes the number of type x regions in the constellation then n_x/M is the probability that a symbol with a type x decision region is transmitted. Hence, we can compute the (average) symbol error probability as

$$P_e = \frac{1}{M}(n_1 P_{e|1} + n_2 P_{e|2} + n_3 P_{e|3}). \tag{11}$$

To complete the derivation, we note that for a general $(M_I \times M_Q)$-QAM constellation with $M_I > 1$ and $M_Q > 1$, we have

$$\begin{aligned}
n_1 &= (M_I - 2)(M_Q - 2) \\
n_2 &= 2(M_I - 2) + 2(M_Q - 2) \\
n_3 &= 4
\end{aligned} \tag{12}$$

A general expression for the symbol error probability can now be formed by combining (8)–(12) and recalling that

$$\frac{c}{\sigma} = \sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_o}}. \tag{13}$$

As shown in Appendix D, we can now combine (8)–(13) to form the symbol error probability for $(M_I \times M_Q)$-QAM as

$$\begin{aligned}
P_e = {} & \frac{2(2M - M_I - M_Q)}{M} Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right) \\
& - \frac{4(M - M_I - M_Q + 1)}{M} Q^2\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right).
\end{aligned} \tag{14}$$

Now, since the second term in (14) is negative, we can form a bound on P_e as

$$\begin{aligned} P_e &< \frac{2(2M - M_I - M_Q)}{M} Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right) \\ &< 4Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right). \end{aligned} \tag{15}$$

For square constellation, $\sqrt{M} = M_I = M_Q$, the symbol error probability expression in (14) simplifies to

$$\begin{aligned} P_e = &\frac{4}{M}(M - \sqrt{M})Q\left(\sqrt{\frac{3E_s}{(M-1)N_0}}\right) \\ &- \frac{4}{M}(M - 2\sqrt{M} + 1)Q^2\left(\sqrt{\frac{3E_s}{(M-1)N_0}}\right), \end{aligned} \tag{16}$$

and the bound in (15) becomes

$$\begin{aligned} P_e &< \frac{4(M - \sqrt{M})}{M} Q\left(\sqrt{\frac{3}{M-1}\frac{E_s}{N_0}}\right) \\ &< 4Q\left(\sqrt{\frac{3}{M-1}\frac{E_s}{N_0}}\right). \end{aligned} \tag{17}$$

The bound in (15) is valid for any rectangular constellation; however, the bound (17) is only valid for square QAM .

Finally, we recall that the minimum distance of a QAM constellation is $d_{\min} = 2c$, and a standard union bound therefore yields

$$\begin{aligned} P_e &\le (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right) \\ &= (M-1)Q\left(\sqrt{\frac{2c^2}{N_0}}\right) \\ &= (M-1)Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right), \end{aligned} \tag{18}$$

which is not as tight as (15) or (17).

As an example, consider the (4×2) constellation in Figure 8. Plots of the exact error probability (14) and the bounds (15) and (18) are found in Figure 10.

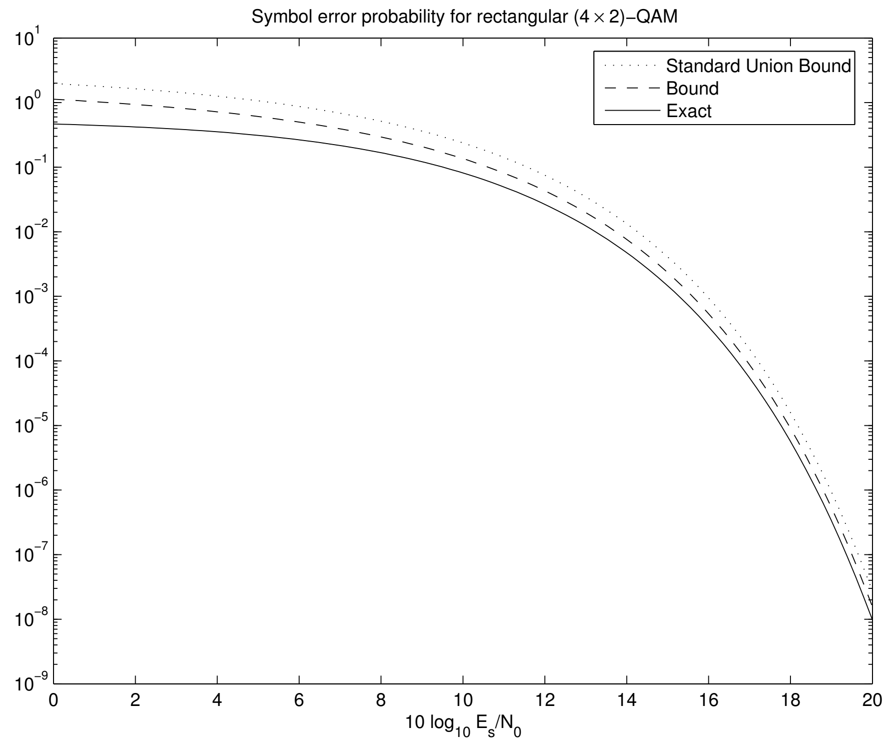

Figure 10: Plots of the exact expression and some upper bounds on the symbol error probability for rectangular (4×2)-QAM (see Figure 8). The curve marked 'Bound' is defined by (17), and the 'Standard Union Bound' is defined by (18).

7 Error Probability for M-ary QAM Revisited

An elegant derivation of the symbol error probability for M-ary QAM can be found by recognizing that a QAM constellation is essentially the combination of two PAM constellations. That is, an $(M_I \times M_Q)$ constellation consists of an M_I-PAM constellation along the $\varphi_I(t)$-direction and an M_Q-PAM constellation along the $\varphi_Q(t)$-direction.

To decode a QAM symbol correctly, we need to decode both PAM constellations correctly. Hence, the probability for correct decision is

$$\begin{aligned} P_c &= \Pr\{\text{correct decision for } M_I\text{-PAM and } M_Q\text{-PAM}\} \\ &= \Pr\{\text{correct decision for } M_I\text{-PAM}\} \Pr\{\text{correct decision for } M_Q\text{-PAM}\}. \end{aligned}$$

Now, the error probabilities for the PAM constellations are

$$\begin{aligned} P_{e,I} &= \frac{2(M_I - 1)}{M_I} Q\left(\frac{c}{\sigma}\right) \\ P_{e,Q} &= \frac{2(M_Q - 1)}{M_Q} Q\left(\frac{c}{\sigma}\right) \end{aligned}$$

and the symbol error probability for the QAM constellation is

$$\begin{aligned} P_e &= 1 - P_c = 1 - (1 - P_{e,I})(1 - P_{e,Q}) = P_{e,I} + P_{e,Q} - P_{e,I}P_{e,Q} \\ &= \frac{2(2M - M_I - M_Q)}{M} Q\left(\frac{c}{\sigma}\right) - \frac{4(M - M_I - M_Q + 1)}{M} Q^2\left(\frac{c}{\sigma}\right). \end{aligned}$$

The relation between c/σ and E_s/N_0 for a QAM-constellation is in (13). Using (13) in the above equation, we arrive at the final expression

$$\begin{aligned} P_e = &\frac{2(2M - M_I - M_Q)}{M} Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2} \frac{E_s}{N_0}}\right) \\ &- \frac{4(M - M_I - M_Q + 1)}{M} Q^2\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2} \frac{E_s}{N_0}}\right), \end{aligned}$$

which is the same equation as (14).

8 Summary

In these notes, we have under the assumption of minimum-distance (ML) detection, AWGN channel, and equally likely transmitted symbols, derived the symbol error probability for baseband and passband PAM (which is the same) and developed two methods for computing the exact symbol error probability for rectangular QAM. The general expression, found in (14), can be further simplified for the case when M is an even square, see (16). A number of upper bounds on the symbol error probability has also been presented in (15), (17), and (18).

A The Concept of a Random Variable

A.1 Introduction

The concept of a random variable is often found confusing. Perhaps that is part due to the fact that a random variable is more similar to a function than a variable.

Informally, a random variable is something that takes on values randomly. Before actually measuring or observing the random variable, we do not know its value. Let us take the outdoor temperature at noon as an example of a random variable. Before noon, we are not quite sure what the temperature will be. It depends on so many different mechanisms that we simply cannot predict it without error. In such situations, it is useful to model quantaties as random variables. Let us denote the temperature at noon by the random variable X. We cannot say what value X takes on (since we cannot predict the temperature).

However, we may be able to say something about the *probability* that the value of X is in a certain range. For instance, we may know what the probability of the event that the temperature will be less than, say, 20 degrees. This probability is denoted

$$\Pr\{X \leq 20\} = \text{Probability that } X \text{ is less or equal to 20 degrees.}$$

Here, $\Pr\{E\}$ denotes the probability of the event E. An *event* is some condition on the outcome of X. An example of an event is that the temperature will be between 10 and 30 degrees, which is formally written as $10 \leq X \leq 30$, and the probability of this event is $\Pr\{10 \leq X \leq 30\}$.

It is important to distinguish between the random variable and its outcome. In our example, the outcome is what the exact temperature will be at noon (which could be 20.1 degrees). The random variable is not a number; it is a description of the probabilities of the possible outcomes. To distinguish betweeen random variables and their outcomes, we usually use upper-case letters, $X, Y, Z, \ldots$ to denote random variables, and the outcomes with lower-case letters $x, y, z, \ldots$.

In many books, the random variable is introduced as a mapping between the sample space (the set of all possible outcomes of a random experiment) and the real numbers. This is a very useful view if we want to get a deeper understanding of random variables. However, we will skip that interpretation here to keep the presentation short.

A.2 Cumulative Distribution Function and Probability Density Function

We can completely describe a random variable X if we know $\Pr\{X \leq x\}$ for all values of x. For instance, we can compute the probability for the temperature to be between 10 and 30 degrees as

$$\Pr\{10 \leq X \leq 30\} = \Pr\{X \leq 30\} - \Pr\{X \leq 10\}.$$

We will use the short-hand notation $F_X(x) = \Pr\{X \leq x\}$. The function $F_X(x)$ is known as the *cumulative distribution function*, (cdf) of X. The cdf for the temperature at noon is plotted in Figure 11. As seen the probability that $X \leq 20$ is 50%. We also note that $F_X(-\infty) = 0$, $F_X(\infty) = 1$, and that $F_X(x)$ is monotonically increasing with x. These properties should be completely clear if we remember the definition of $F_X(x)$.

Another important description of a random variable is the *probability density function* (pdf), which is defined as

$$f_X(x) = \frac{d}{dx} F_X(x).$$

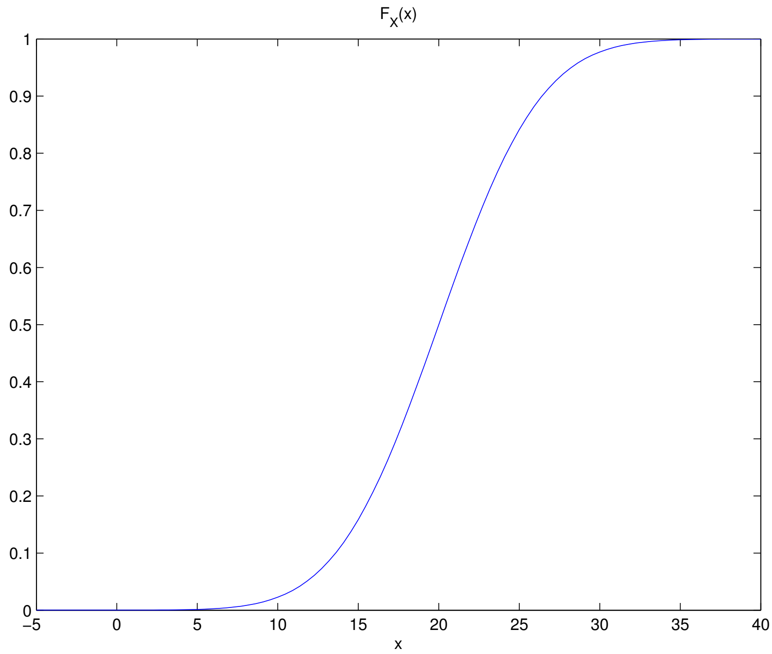

Figure 11: Cumulative distribution function (cdf).

From the definition, it is clear that

$$F_X(a) = \int_{-\infty}^{a} f(x)\,dx.$$

The pdf for X describes how the probability is distributed. The pdf for our example random variable is depicted in Figure 12. From the plot we see that $f_X(x)$ peaks around $x = 20$, and this indicates that there is a fairly large probability that the temperature is close to 20 degrees. To show this reasoning with mathematics, we note that

$$\begin{aligned}\Pr\{15 \leq X \leq 25\} &= F_X(25) - F_X(15)\\ &= \int_{-\infty}^{25} f_X(x)\,dx - \int_{-\infty}^{15} f_X(x)\,dx\\ &= \int_{15}^{25} f_X(x)\,dx.\end{aligned}$$

Now, if $f_X(x)$ is large in the integration interval, the probability that X will have an outcome in the interval is also large.

A.3 Expectations

From the pdf we can compute the *expected value* (also called *average* or *mean*) of the random variable. The expected value is pretty much what it sounds like. It is

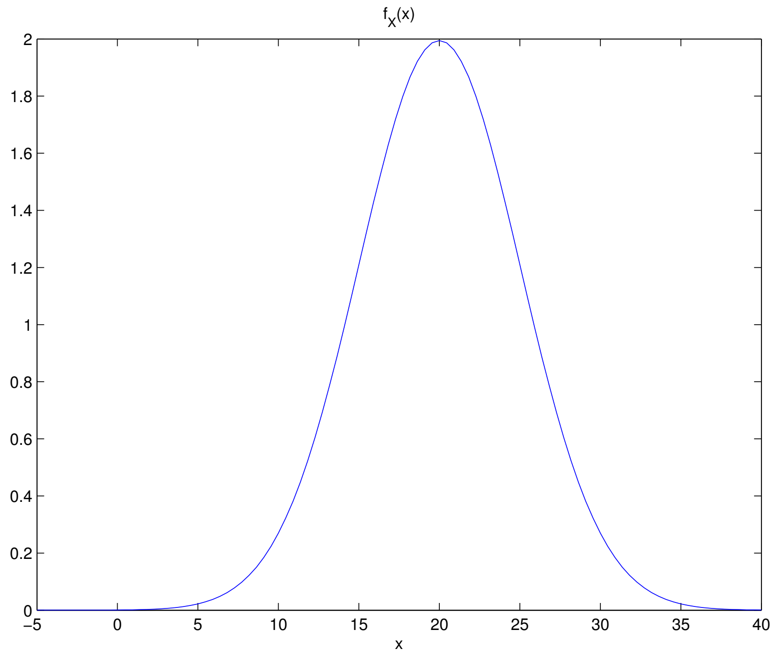

Figure 12: Probability density function (pdf).

our best guess what the outcome of the random variable will be. The definition of the expected value is

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx.$$

For symmetrical pdf's, like the one in Figure 12, the expected value is equal to the symmetry point on the x-axis. Hence, the expected temperature is 20 degrees in our example.

A measure on how much the outcome of the random variable deviates from the expected value is the *variance*. The variance of X is defined as

$$\mathrm{var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2].$$

In plain English, this says that "the variance is the average of the square of the difference between X and the expected value of X." Read that slowly a couple times. A large variance implies that the outcome is likely to be far from the expected value, while a small variance implies that the outcome is likely to be close to the expected value.

B Inner products and norms

Signals can be viewed as vectors in a normed vector space. Consider a complex-valued signall $x(t)$. The energy of $x(t)$ is defined as

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt.$$

Suppose we have two signals $x(t)$ and $y(t)$ with finite energies. We can then define the inner product between $x(t)$ and $y(t)$ as

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) y^*(t) \, dt \tag{19}$$

The inner product is sometimes called the scalar product or dot product. We say that the signals are orthogonal (or perpendicular) if the inner product is zero, just as two vectors in, say, 3-dimensional space, are orthogonal when their scalar product is zero.

The inner product has several useful properties, which are easily proved directly from the definition (19),

$$\begin{aligned}
\langle x(t), y(t) \rangle &= (\langle x(t), y(t) \rangle)^* \\
\langle a x(t), y(t) \rangle &= a \langle x(t), y(t) \rangle \\
\langle x(t), a y(t) \rangle &= a^* \langle x(t), y(t) \rangle \\
\langle x_1(t) + x_2(t), y(t) \rangle &= \langle x_1(t), y(t) \rangle + \langle x_2(t), y(t) \rangle \\
\langle x(t), y_1(t) + y_2(t) \rangle &= \langle x(t), y_1(t) \rangle + \langle x(t), y_2(t) \rangle.
\end{aligned}$$

The norm of the signal $x(t)$ is defined as

$$\|x(t)\| = \sqrt{\langle x(t), x(t) \rangle} = \sqrt{\int_{-\infty}^{\infty} |x(t)|^2 \, dt} = \sqrt{E_x}.$$

We can interpret the norm of $x(t)$ as its length and $\|x(t) - y(t)\|$ as the distance from $y(t)$ to $x(t)$. The Pythagorean theorem holds: for two orthogonal signals $x(t)$ and $y(t)$,

$$\|x(t) + y(t)\|^2 = \|x(t)\|^2 + \|y(t)\|^2.$$

We can easily prove this from the properties of the inner product,

$$\begin{aligned}
\|x(t) + y(t)\|^2 &= \langle x(t) + y(t), x(t) + y(t) \rangle \\
&= \langle x(t), x(t) \rangle + \underbrace{\langle x(t), y(t) \rangle}_{=0} + \underbrace{\langle y(t), x(t) \rangle}_{=0} + \langle y(t), y(t) \rangle \\
&= \|x(t)\|^2 + \|y(t)\|^2
\end{aligned}$$

Hence, if $x(t)$ are orthogonal, the energy for the signal $x(t) + y(t)$ is the sum of the signal energies E_x and E_y.

C PAM and QAM constellation energy

Recall that the alphabet for M-ary PAM is $\mathcal{A} = \{\pm(2i-1)c : i = 1, 2, \dots, M/2\}$. Suppose all symbols are equally likely, i.e., that the probability of $a = A_m$ is $1/M$ for all $A_m \in \mathcal{A}$. Then the average energy for the constellation is

$$E_a = \frac{1}{M}\sum_{i=1}^{M/2}[-(2i-1)c]^2 + [(2i-1)c]^2 = \frac{2}{M}\sum_{i=1}^{M/2}[(2i-1)c]^2 = c^2\frac{2}{M}\sum_{i=1}^{M/2}4i^2 - 4i + 1$$

We recall the following formulas (see, e.g., [2, p. 189])

$$\sum_{i=1}^{M/2} i = \frac{M(M+1)}{8}, \qquad \sum_{i=1}^{M/2} i^2 = \frac{M(M+1)(M+2)}{24}.$$

Hence, by combining the above three equations, we conclude that

$$\sum_{i=1}^{M/2}(2i-1)^2 = \frac{M}{2}\frac{M^2-1}{3}$$

and arrive at

$$E_a = c^2\frac{M^2-1}{3}.$$

A rectangular QAM constellation can be viewed as composed of two PAM constellations, one along the I-direction and one along the Q-direction. Hence, a QAM signal point is of the form $[\pm(2i-1)c \quad \pm(2j-1)c]^T$, where $i = 1, 2, \dots, M_I/2$ and $j = 1, 2, \dots, M_Q/2$, and the corresponding energy is $(2i-1)^2c^2 + (2j-1)^2c^2$. Due to symmetry, we can compute the average constellation energy as the average energy for the symbols in the first quadrant, i.e., for the $M/4$ points of the form $[(2i-1)c \quad (2j-1)c]^T$, where $i = 1, 2, \dots, M_I/2$ and $j = 1, 2, \dots, M_Q/2$,

$$\begin{aligned}
E_a &= \frac{4}{M}\sum_{i=1}^{M_I/2}\sum_{j=1}^{M_Q/2}[(2i-1)c]^2 + [(2j-1)c]^2 \\
&= c^2\frac{4}{M}\sum_{i=1}^{M_I/2}\left[\sum_{j=1}^{M_Q/2}(2i-1)^2 + (2j-1)^2\right] \\
&= c^2\frac{4}{M}\sum_{i=1}^{M_I/2}\left[\frac{M_Q}{2}(2i-1)^2 + \frac{M_Q}{2}\frac{M_Q^2-1}{3}\right] \\
&= c^2\frac{4}{M}\left[\frac{M_Q}{2}\frac{M_I}{2}\frac{M_I^2-1}{3} + \frac{M_I}{2}\frac{M_Q}{2}\frac{M_Q^2-1}{3}\right] \\
&= c^2\frac{M_I^2-1}{3} + c^2\frac{M_Q^2-1}{3}
\end{aligned}$$

Hence, the average energy for the QAM constellation is the sum of the average energies for the PAM constellations in the I and Q-directions.

D Derivation of (14)

We define

$$q = Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right).$$

From (8)–(13) we have

$$\begin{aligned}
MP_e &= n_1 P_{e|1} + n_2 P_{e|2} + n_3 P_{e|3} \\
&= (M_I - 2)(M_Q - 2)(4q - 4q^2) + [2(M_I - 2) + 2(M_Q - 2)](3q - 2q^2) + 4(2q - q^2) \\
&= q[4(M_I - 2)(M_Q - 2) + 6(M_I - 2) + 6(M_Q - 2) + 8] \\
&\quad + q^2[-4(M_I - 2)(M_Q - 2) - 4(M_I - 2) - 4(M_Q - 2) - 4] \\
&= q[4M - 8M_I - 8M_Q + 16 + 6M_I - 12 + 6M_Q - 12 + 8] \\
&\quad + q^2[-4M + 8M_I + 8M_Q - 16 - 4M_I + 8 - 4M_Q + 8 - 4] \\
&= 2(2M - M_I - M_Q)Q\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right) \\
&\quad - 4(M - M_I - M_Q + 1)Q^2\left(\sqrt{\frac{6}{M_I^2 + M_Q^2 - 2}\frac{E_s}{N_0}}\right).
\end{aligned}$$

which proves (14).

E References

[1] John R. Barry, Edward A. Lee, and David G. Messerschmidt. *Digital Communication.* Springer, third edition, 2003.

[2] Lennart Råde and Bertil Westergren. *Beta, Mathematics Handbook.* Studentlitteratur, 1998.