Chapter 2 Modulation and Coding for Ideal Gaussian channels

This chapter is devoted to the discussion on orthonormal PAM/QAM systems, capacity calculations for ideal AWGN channels, and the performance evaluation of small signal constellations.

Note: Some of material presented in this chapter is based on Forney's lecture notes.

2.1 Data Modulation and Demodulation

In data modulation we convert information bits into waveforms that are suitable for transmission over a communication channel. This is done, in general, by splitting data modulation (or channel encoding) into separate layers: mapping from a sequence of binary digits into a sequence of real or complex numbers (numerical signals), followed by the conversion of signals to a waveform. Demodulation is the corresponding process at the receiver. The demodulator maps the received waveform to a sequence of received signals, which is then mapped to a bit sequence.

In a passband system, modulators usually include a third layer, which maps a baseband-encoded waveform $x_b(t)$ into a passband waveform $x(t)=\sqrt{2}\Re\left[x_b(t)e^{j2\pi f_c t}\right]$ with carrier frequency f_c. At the receiver this passband waveform is mapped back to baseband before performing the other components of decoding. This frequency conversion operation at encoder and decoder is often referred to as modulation and demodulation, but it is more common today to use the word modulation for the entire process of mapping bits to waveforms. Figure 2.0 illustrates these three layers.

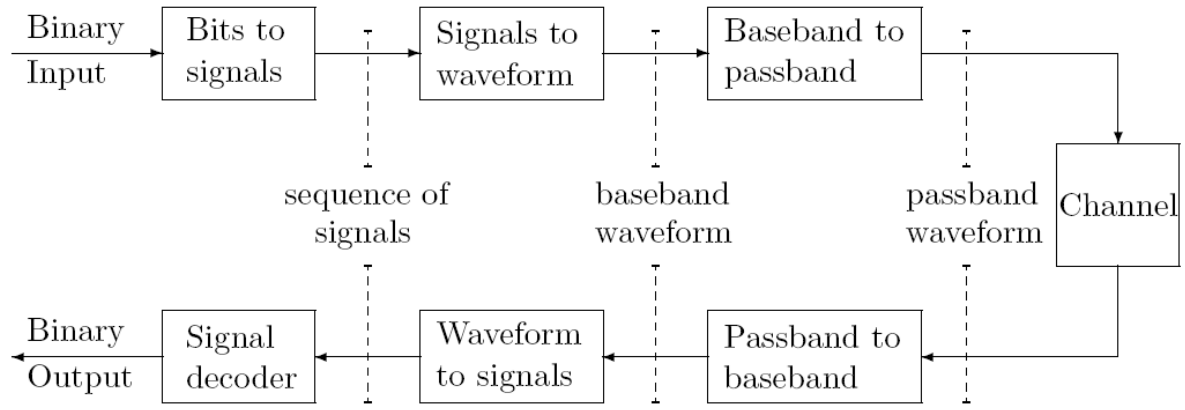


Figure 2.0 The layers of a modulator (channel encoder) and demodulator (channel decoder)

2.1.1 Mapping of vectors to waveform

It is well-known that any continuous-time function $x(t)\in\mathcal{L}_2$ defined on the interval [0, T] can be decomposed according to some set of N orthonormal basis function $\{\phi_k(t), 1\le k\le$

$N\}$ as

$$x(t) = \sum_{k=1}^{N} x_k \phi_k(t), \quad 0 \le t \le T \tag{2.1}$$

Relationship in (2.1) indicates a correspondence between an N-dim real-valued vector $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix}^T \in \mathbb{R}^N$ and the waveform $x(t) \in \mathcal{L}_2$. Let $\mathbf{x}$ be a data symbol. Then the modulated signal (waveform) corresponding to $\mathbf{x}$ is given by (2.1).

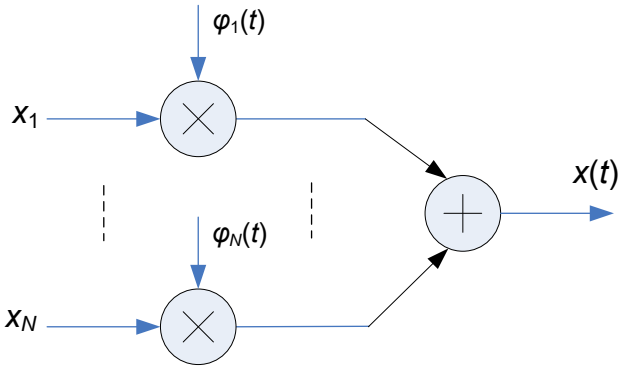


Figure 2.1 Modulator implements the basis expansion.

Similarly, a set of continuous-time function $\{x_i(t)\}$ corresponds to a set of discrete N-dim points $\{\mathbf{x}_i\}$ known as a *signal constellation*.

This decomposition of continuous-time functions can be extended to random processes using what is known as *Karhunen-Loeve expansion*. Decomposition of random processes is fundamental to demodulation and detection in the presence of noise. Modulation constructively assembles random signals for the communication system from a set of basis functions $\{\phi_k(t)\}$ and a set of signal points $\{\mathbf{x}_i\}$.

2.1.2 Demodulation

The demodulation takes the continuous-time waveforms and extracts the received version. Given the basis expansion of (2.1), the demodulation extracts the coefficients of the expansion by projecting the (received) signal onto its basis as shown below.

$$\int_0^T x(t)\phi_n(t)dt = \int_0^T \sum_{k=1}^{N} x_k \phi_k(t)\phi_n(t)dt$$

$$= \sum_{k=1}^{N} x_k \int_0^T \phi_k(t)\phi_n(t)dt$$

$$= \sum_{k=1}^{N} x_k \delta_{kn} = x_n \tag{2.2}$$

Therefore, in the noiseless case, demodulation is just recovering the coefficients of the basis function, and it can be implemented by a correlator. See Fig. 2.2.

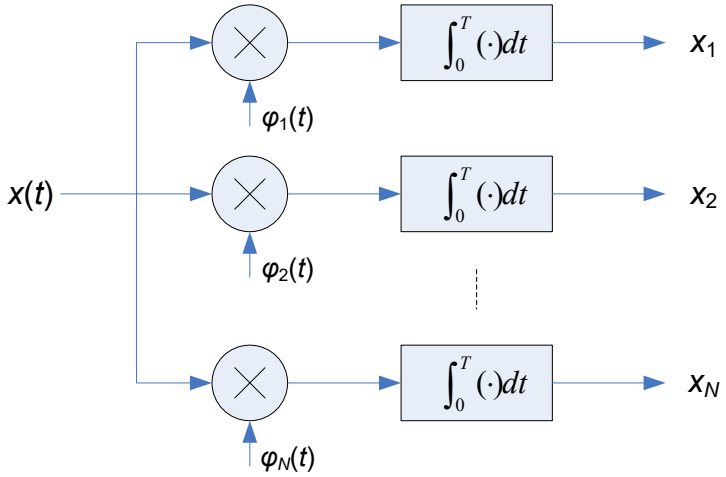


Figure 2.2 The correlative demodulator

The computation in (2.2) can be more easily implemented by a matched filter since it can be written as

$$\int_0^T x(t)\phi_n(t)dt = x(t) * \phi_n(T-t)\big|_{t=T}$$

where "*" denotes convolution operation. Matched-filter demodulator is illustrated in Fig. 2.3.

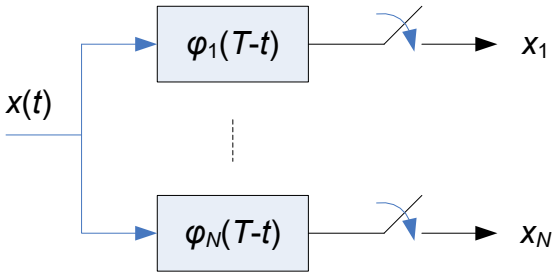


Figure 2.3 The matched-filter demodulator

2.2 The Ideal AWGN Channel

The continuous-time ideal band-limited AWGN channel is modeled by

$$y(t) = x(t) + n(t)$$

where $x(t)$ is the channel input signal with an average power of P and a positive-frequency bandwidth of W Hz, and $n(t)$ is AWGN with one-sided p.s.d. N_0.

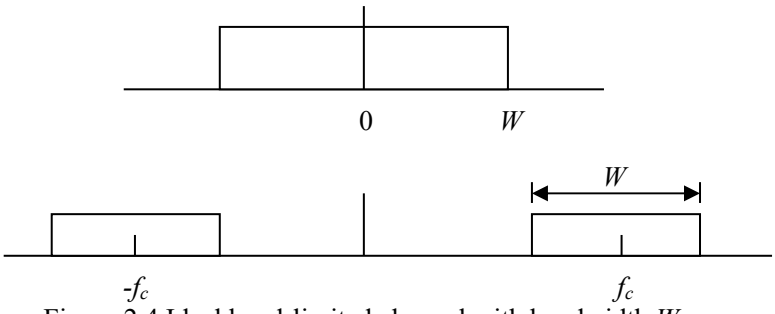


Figure 2.4 Ideal band-limited channel with bandwidth W

- The channel band B is a positive-frequency interval with bandwidth W Hz. The channel is said to be baseband if $B = [0, W]$, and passband otherwise.

2.1.1 Conversion from the continuous-time AWGN channel into a vector channel

Let $\{\phi_k(t), 1 \le k \le N\}$ denote a set of real orthonormal waveforms. Then the transmitted signal can be expressed as

$$x(t) = \sum_{k=1}^{N} x_k \phi_k(t), \quad 0 \le t \le T \tag{2.1}$$

With optimal receiver, the sampled output of the matched-filter (MF) demodulator is given by

$$y_k = \langle y(t), \phi_k(t) \rangle = \int_0^T y(t)\phi_k(t)dt = x_k + n_k \tag{2.2}$$

where

- $\{y_k\}$ is a set of sufficient statistics about $\left\{x_k = \int_0^T x(t)\phi_k(t)dt\right\}$;
- $\left\{n_k = \int_0^T n(t)\phi_k(t)dt\right\}$ is a set of i.i.d. real Gaussian random variables with zero mean and variance $\sigma_n^2 = N_0/2$.

Equation (2.2) is the discrete-time ideal AWGN channel model. Notice that there is no intersymbol interference (ISI).

2.3 Signal Constellations

An N-dimensional *signal constellation* (signal set, alphabet) of size M is a set $\mathcal{A} = \{\mathbf{a}_j, 1 \le j \le M\}$ of M real N-tuples $\mathbf{a}_j \in \mathbb{R}^N$. Each distinct point (called signal point) in the signal constellation corresponds to a different modulated waveform, but all the waveforms share the same set of basis functions. Such a set is able to transmit $\log_2 M$ bits per N dimensions. Some examples of signal constellations are shown in Fig. 2.5.

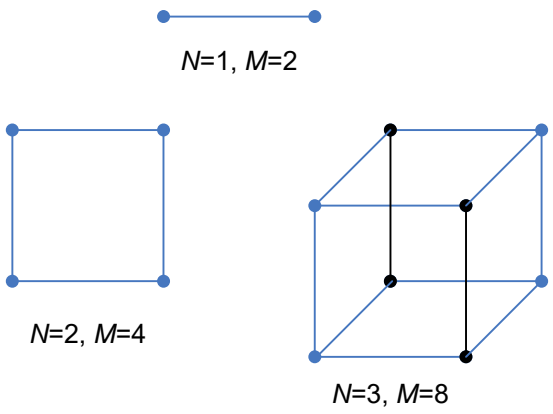


Figure 2.5 (a) Cubic constellations for N=1, 2 and 3.

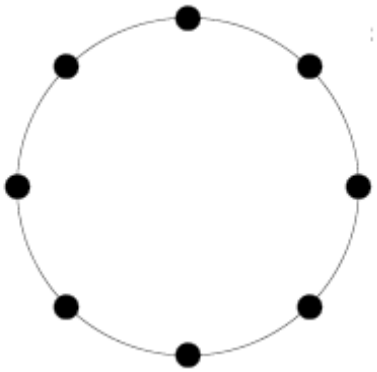


Figure 2.5 (b) A circular constellation

The average energy of $\mathcal{A}$ (per N dimensions) is defined by

$$E_N = \mathsf{E}\left[\| \mathbf{a}_j \|^2 \right] = \sum_{j=1}^{M} \| \mathbf{a}_j \|^2 \Pr(j) \tag{2.3}$$

where Pr(j) is the probability of choosing $\mathbf{a}_j$. E_N is closely related to the average power P by $P = E_N / T$ with T standing for the signal duration (symbol interval).

The *minimum distance* between constellation points is denoted by

$$d_{\min}(\mathcal{A}) = \min_{i \neq j} \| \mathbf{a}_i - \mathbf{a}_j \|, \tag{2.4}$$

which is related to the probability of error of detector.

The above power constraints and minimum distance constitutes two key parameters of a signal constellation. Other parameters are as follows.

- A system with symbol interval T and bandwidth W has a number of dimensions available for constellation construction
$$N = 2WT$$
- The average number of bits per dimension is
$$b = \frac{\log_2 M}{N} \qquad \text{bits / dim}$$
- The data rate is $R = \log_2 M / T$ bits/s.
- Then $2b = R/W \equiv \rho$, which is the *spectral efficiency* of a modulation scheme (bits / s / Hz), or
$$\rho = \frac{2}{N} \log_2 M \qquad \text{bits / 2D}$$
- The average energy of $\mathcal{A}$ per dimension is given by $E_d = E_N / N$

  If the M signal points in $\mathcal{A}$ are assumed to be equiprobable, then

$$E_N = \frac{1}{M}\sum_{j=1}^{M} \| \mathbf{a}_j \|^2 \quad \text{and} \quad E_d = \frac{1}{MN}\sum_{j=1}^{M} \| \mathbf{a}_j \|^2$$

- To simplify performance assessment, we define a *figure of merit* (CFM) for a constellation as

$$\zeta \equiv \frac{d^2_{\min}(\mathcal{A})}{4E_d} \tag{2.5}$$

As we have stated in Chapter 1, from the signal space viewpoint, the data modulation can be viewed as a two-step process: a signal mapper followed by a simple waveform converter. A *signal mapper* performs the mapping of a binary vector of length $\log_2 M$ into one of the signal points in a constellation of size M. Usually, the binary vector is referred to as the *label* of the signal point. Mapping is not arbitrary, clever choices lead to better performance over noisy channels.

In some channels it is suitable to label points that are close in Euclidean distance to map to being close in Hamming distance. An example for such mapping is the Gray-mapping. Fig. 2.6 illustrates an example.

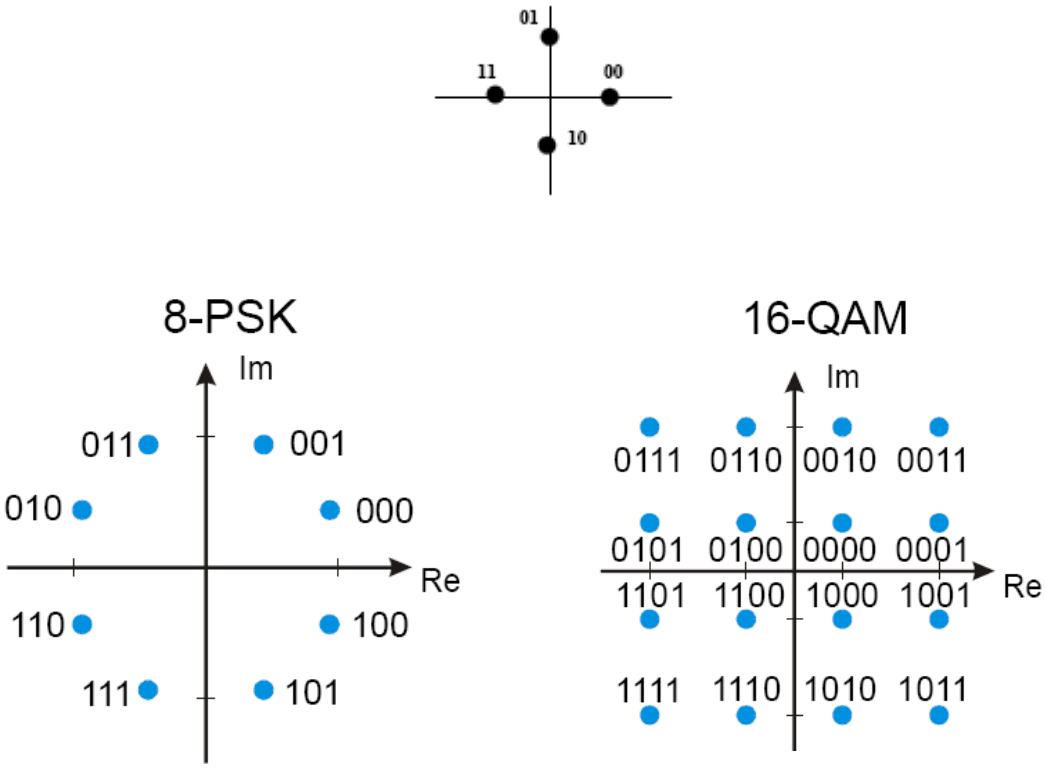


Figure 2.6 Gray-coded mapper for QPSK, 8-PSK, and 16-QAM constellations. Neighboring symbols differ in only one bit.

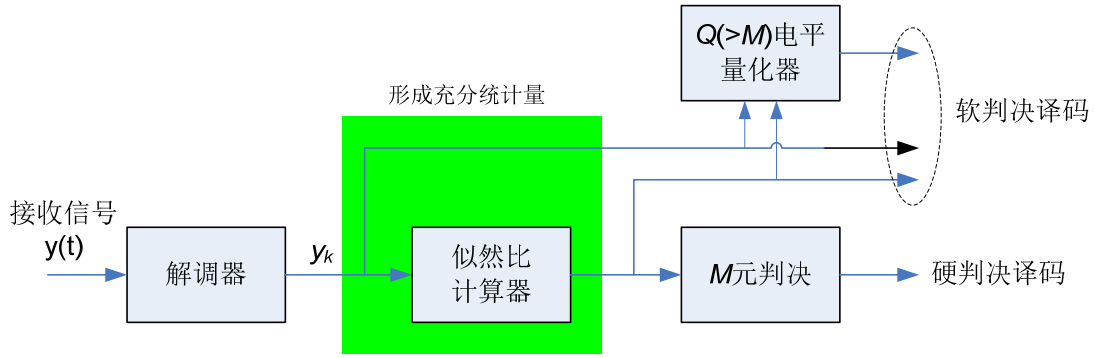
✧ **解调器（解映射器）与译码器之间的接口：Hard-decisions v.s. Soft decisions**

（1） Hard-decisions：When coding is used, it is possible to make ML decisions on

each signal separately (by the 解调器), and then to use the coding constraints to correct errors in the detected sequence (by 纠错译码器). These individual signal decisions are called *hard decisions*.

（2） Soft decisions：For each k, 解调器 pass a sufficient statistics (such as y_k or LLR(y_k)) for the decision on x_k or information on the reliability of the decisions (taken by the demodulator) to the 译码器. This is called a *soft decision* since it save all the relevant information needed for an ML decision (made by 译码器) between the set of possible codewords.

Hard-decision decoding neglects some of the information available at the receiver, 而这些信息 could be used to improve 译码效率。Soft decisions yield smaller error probabilities than hard decisions. 下图给出了解调器与译码器之间的接口描述。



对于采用二元 FEC 码的 BPSK 系统，关于 x_k 的充分统计量通常简单地选为接收值 y_k 或对数似然比（LLR）：

$$L(x_k)=\log\left(\frac{p(y_k\mid x_k=+1)}{p(y_k\mid x_k=-1)}\right)=\frac{4y_k}{N_0},\qquad \text{（假定 AWGN 信道）}$$

基于此未量化信息或由此进一步产生的判决可靠性信息（量化信息），FEC 译码器可进行软判决译码。

注：使用 LLR 的优点是：

1） 在一定程度上压缩了似然比的范围；

2） 可以将联合事件的概率的乘积运算转化为简单的 LLR 值求和运算。

对于采用二元 FEC 码的高阶调制系统，关于 x_k 的充分统计量通常选为接收值 y_k 或符号概率（似然值）、LLR 值等，这是符号级上的软解调。在此基础上，我们可以做硬判决，并将判决符号的 label 值作为比特向量向译码器输出硬判决结果；也可以进一步做 bit-level soft demodulation.

- **星座点转换为信号波形：**

Denote the jth signal vector in $\mathcal{A}$ by $\mathbf{a}_j=(a_{j,1},a_{j,2},...,a_{j,N})$. Then the selected signal vector $\mathbf{a}_j$ is modulated into a signal waveform $b_j(t)=\sum_{k=1}^{N}a_{j,k}\phi_k(t)$, where $\{\phi_1(t),...,\phi_N(t)\}$

is a set of N orthonormal waveforms. This provides a general scenario for mapping the symbols 1 to M into a set of signal waveforms $b_1(t)$ to $b_M(t)$. If we transmit a sequence of such M-ary symbols at T-spaced symbol-intervals, the most straightforward way of accomplishing this is to choose the orthonormal waveforms $\phi_1(t),...,\phi_N(t)$ in such a way that $\phi_k(t-lT)$ and $\phi_m(t-l'T)$ are orthonormal for all $k,m,1\le k,m\le N$, and all integers l,l'.

Note that PAM is a special case of this scenario where the dimension $N = 1$, and the function $\phi_1(t)$ is the real modulation pulse $p(t)$ for baseband transmission; QAM is another special case where $N = 2$ at passband. In this case, the complex signals a_j are viewed as 2D real signals. The orthonormal waveforms (assuming real $p(t)$) are $\phi_1(t)=\sqrt{2}p(t)\cos(2\pi f_c t)$ and $\phi_2(t)=\sqrt{2}p(t)\sin(2\pi f_c t)$. See the following for a brief discussion. A more detailed discussion can be found in [Gallgar, 2008].

2.4 Orthonormal PAM / QAM Systems

We now proceed to consider the modulation system with PAM/QAM signaling. The system block model is depicted in Fig. 2.7, together with the notation to be used in the following discussions.

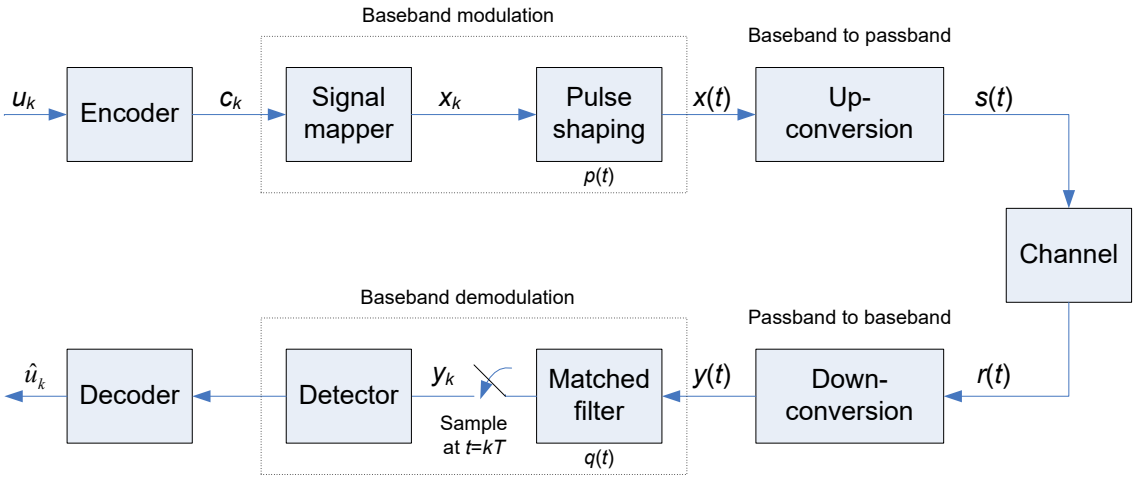


Figure 2.7 Block model of modulation and demodulation

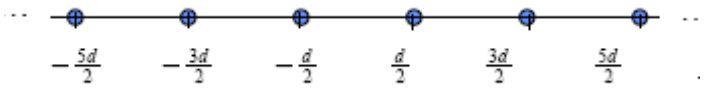
2.4.1 Orthonormal PAM

- PAM Signal Set

A standard M-PAM signal set $\mathcal{A}$ consists of $M = |\mathcal{A}| = 2^b$ d-spaced real numbers located symmetrically about the origin; i.e.,

$$\mathcal{A} = \{-d(M-1)/2, ..., -d/2, d/2, ..., d(M-1)/2\}$$

where b is the number of bits transmitted per symbol. See the figure below.



Assume that the signal points are equiprobable. Then the *average energy per symbol* is

$$\overline{A^2} = \int_{-dM/2}^{dM/2} a^2 \cdot \frac{1}{dM} da - \int_{-d/2}^{d/2} \frac{a^2}{d} da = \frac{d^2(M^2-1)}{12} \tag{2.6}$$

For example, for $M = 2$, 4 and 8, $\overline{A^2} = d^2/4, 5d^2/4$ and $21d^2/4$, respectively.

An alternative derivation:

$$\overline{A^2} = \frac{2}{M} \sum_{i=1}^{M/2} \left(\frac{2i-1}{2} \right)^2 d^2 = \frac{d^2}{2M} \left(\frac{M^3}{6} - \frac{M}{6} \right)$$

思考題： Let U be a random variable uniformly distributed in $[-d, d]$ and let S be a discrete random variable independent of U and uniformly distributed over the PAM constellation {±d, ±3d, …, ±(m−1)d}, where m is an even integer. Let $V = S + U$.

(a) Find the distribution of V.

(b) Find the variance of U and that of V.

(c) Use part (b) to determine the variance of S. (i.e., the average energy of the PAM with uniform distribution.)

■ PAM Modulator

A PAM modulator converts a discrete-time sequence $\{x_k, k \in \mathbb{Z}\}$ of real symbols into a continuous-time real waveform

$$x(t) = \sum_{k \in \mathbb{Z}} x_k p(t - kT)$$

where T is the symbol interval, $p(t)$ is the basic modulation pulse, and $x_k \in \mathcal{A}$.

A PAM system is said to be *orthonormal* if

$$\langle p(t-kT), p(t-jT) \rangle = \delta_{k-j}$$

that is $\{\phi_k(t)\} = \{p(t\text{-}kT)\}$.

With the assumption that the underlying channel model is an ideal AWGN channel, the received baseband waveform is given by

$$y(t) = x(t) + n(t)$$

In an orthonormal PAM system, the sequence $\{y_k\}$ of matched-filter outputs is given by

$$y_k = \langle y(t), p(t-kT) \rangle = x_k + n_k$$

Thus, we obtain without loss of optimality an equivalent discrete-time AWGN channel model

$$\mathbf{y} = \mathbf{x} + \mathbf{n}$$

2.4.2 Orthonormal QAM

- QAM Signal Set

A standard ($M\times M$)-QAM signal set consists of the Cartesian product of two standard M-PAM d-spaced signal set; i.e.,

$$\mathcal{A} = \{a = (a' + ja'') \mid a' \in \mathcal{A}', a'' \in \mathcal{A}'\}$$

where $\mathcal{A}' = \{-d(M-1)/2, ..., -d/2, d/2, ..., d(M-1)/2\}$. Thus the size of $\mathcal{A}$ is $|\mathcal{A}| = M^2$, and the number of bits per symbol is $2\log_2 M = 2b$. See Fig. 2.8 for examples. Assuming that the symbols in $\mathcal{A}$ are equiprobable, the *average energy per complex symbol* is

$$\overline{A^2} = 2 \times \frac{d^2(M^2-1)}{12} = \frac{d^2(M^2-1)}{6} \tag{2.7}$$

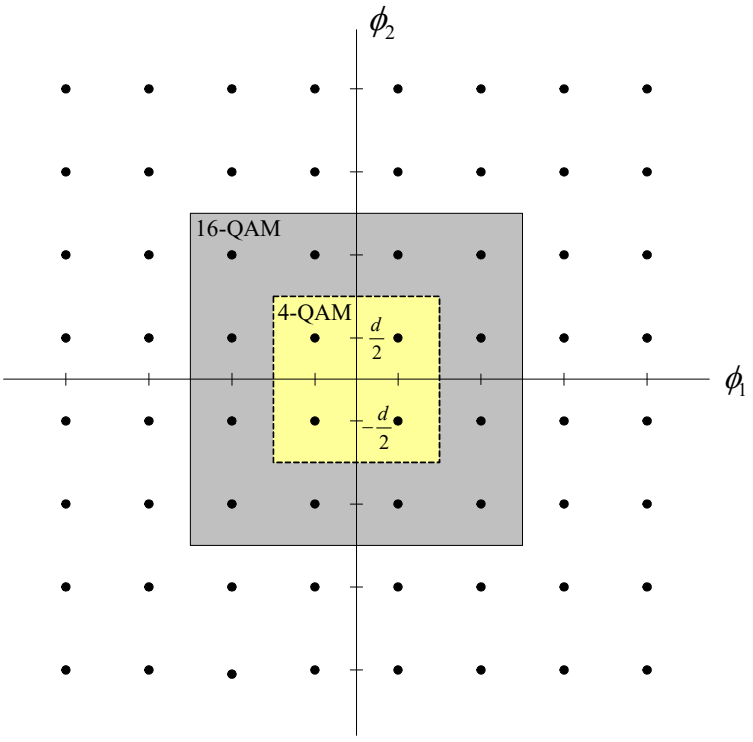


Figure 2.8 QAM constellations

- QAM Modulator

A QAM baseband modulator converts a discrete-time sequence $\{x_k,\ k\in\mathbb{Z}\}$ of complex symbols into a continuous-time complex waveform

$$x(t) = \sum_{k \in \mathbb{Z}} x_k p(t - kT)$$

where T is the symbol interval, $p(t)$ is the complex modulation pulse, and $x_k \in \mathcal{A} \subset \mathbb{C}$.

From the implementation viewpoint, it is helpful to view the modulation in terms of filtering, see Fig. 2.9. If $p(t)$ is real, this corresponds to two real filtering operations.

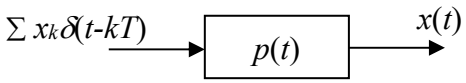


Figure 2.9

A QAM system is said to be *orthonormal* if

$$\langle p(t - kT), p(t - jT) \rangle = \delta_{k-j}$$

where the inner product is defined as $\langle f(t), g(t) \rangle = \int f(t) g^*(t) dt$.

In a passband QAM system, $x(t)$ is then up-modulated to passband. After transmission over an AWGN channel and down-modulation, the received signal is

$$y(t) = x(t) + n(t)$$

where $n(t)$ is independent of $x(t)$ and can be represented as

$$n(t) = \sum_k n_k p(t - kT) + n^{\perp}(t)$$

Here $\{n_k\}$ is a sequence of i.i.d. zero-mean complex Gaussian variables, each with variance N_0, and $n^{\perp}(t)$ is a Gaussian process with $< n^{\perp}(t), p(t\text{-}kT) > = 0$ for all k.

The output sequence of a QAM demodulator is given by

$$y_k = \langle y(t), p(t - kT) \rangle = x_k + n_k$$

which represents an equivalent discrete-time AWGN channel model obtained with an orthonormal QAM system.

2.4.3 Nyquist Criterion and Minimum Bandwidth

由信号检测理论可知，在 AWGN 下，接收滤波器采用匹配滤波器，即 $q(t) = p^*(-t)$，能够使得系统在采样时刻的输出信噪比达到最大，因此匹配滤波器（或等效地，相关器）通常被用作最佳接收机的前端。这是从抗噪声性能的角度来考虑的。从抗 ISI（无失真传输）的角度来讲，我们需要选择发送滤波器 $p(t)$、接收滤波器 $q(t)$使得接收波形

$$y'(t) = \int_{-\infty}^{\infty} x(\tau) q(t - \tau) d\tau = \sum_k x_k g(t - kT)$$

在 $t=kT$ 时的样值 $y'(kT) = x_k$ for each k, 其中 $g(t) = p(t) * q(t) = \int_{-\infty}^{\infty} p(\tau) q(t - \tau) d\tau$.

考虑上述两个因素，我们要求 $g(t) = p(t) * p^*(-t)$ 满足 $g(0) = 1$ and $g(kT) = 0$ for

each nonzero integer k. Waveforms with this property are said to be *idea Nyquist*.

We now consider the design of the modulation pulse $p(t)$ such that the PAM/QAM system has no intersymbol interference. (That is, find the condition under which $\{p(t\text{-}kT)\}$ forms a orthonormal set.)

■ Nyquist Pulses

In 1928, Nyquist recognized that the *sinc* pulse in the Sampling theorem was significant in two ways:

- the shifts of *sinc* pulse are orthogonal;
- the *sinc* pulse has the value 0 at all sampling points other than at time 0.

Nyquist realized that, to avoid ISI, the crucial property was to *find a pulse whose shifts are orthogonal*.

We now find the conditions under which a real waveform $p(t)$ has the property:

$$\int_{-\infty}^{\infty} p(t-kT)\,p^*(t)dt = \begin{cases} 1, & \text{for } k=0 \\ 0, & \text{for } k\neq 0 \end{cases} \tag{2.8}$$

Note that the Fourier transform $\mathcal{F}[p(t-kT)] = e^{-j2\pi kTf}P(f)$, we have

$$\int_{-\infty}^{\infty} p(t-kT)\,p^*(t)dt = \int_{-\infty}^{\infty} e^{-j2\pi kTf}P(f)P^*(f)df \quad \text{(by Parseval's theorem)}$$

$$= \sum_m \int_{\frac{m}{T}-\frac{1}{2T}}^{\frac{m}{T}+\frac{1}{2T}} e^{-j2\pi kTf}P(f)P^*(f)df$$

Making the substitution $f_m = f - m/T$ in each interval and since $e^{j2\pi km} = 1$ for all integers m and k, we have

$$\int_{-\infty}^{\infty} p(t-kT)\,p^*(t)dt = \sum_m \int_{-\frac{1}{2T}}^{\frac{1}{2T}} e^{-j2\pi kT(f_m+\frac{m}{T})}P(f_m+\frac{m}{T})P^*(f_m+\frac{m}{T})df_m$$

$$= \int_{-\frac{1}{2T}}^{\frac{1}{2T}} e^{-j2\pi kTf}\sum_m P(f+\frac{m}{T})P^*(f+\frac{m}{T})df$$

$$= \int_{-\frac{1}{2T}}^{\frac{1}{2T}} e^{-j2\pi kTf}\sum_m G(f+\frac{m}{T})df \tag{2.9}$$

where we have defined

$$G(f) \triangleq P(f)P^*(f) = |P(f)|^2$$

Combining (2.9) with (2.8) yields

$$\sum_m G\left(f+\frac{m}{T}\right) = \sum_m \left|P\left(f+\frac{m}{T}\right)\right|^2 = T$$

Theorem 2.1 (*Orthonormality condition*): The time shifts $\{p(t\text{-}kT)\}$ of a real or complex $\mathcal{L}_2$ waveform $p(t)$ with Fourier transform $P(f)$ are orthonormal if and only if

$$\frac{1}{T}\sum_m |P(f+m/T)|^2 = 1 \qquad \text{for } -\frac{1}{2T} \le f \le \frac{1}{2T} \tag{2.10}$$

This is often called *Nyquist criterion* for zero ISI.

■ An Alternative Derivation of (2.10)

Let $g(t) = p(t) * p^*(-t)$. Noting that $p(t) * p^*(-t) = \int_{-\infty}^{\infty} p(\tau) p^*(\tau - t) d\tau$, equation (2.8) is equivalent to

$$p(t) * p^*(-t)\,|_{t=kT} = g(kT) = \delta_k$$

From the sampling theorem, the discrete-time Fourier transform (DTFT) of the sequence $\{g(kT)\}$ is

$$\tilde{G}(f) = G(e^{j2\pi fT}) = \frac{1}{T}\sum_{m=-\infty}^{\infty} G(f - m/T)$$

$$= \frac{1}{T}\sum_m |P(f - m/T)|^2 = 1 \quad \text{for all } f$$

where $G(f) = \mathcal{F}[g(t)]$. The $g(t)$ can be viewed as the output of a filter matched to $p(t)$, see Fig. 2.10.

Figure 2.10

■ Minimum bandwidth of Orthonormal Pulses

The orthonormality condition implies a bandwidth of at least $W = 1/(2T)$ is required for ISI-free transmission.

There are many pulse shapes that satisfy the Nyquist criterion. The minimum bandwidth pulse satisfying (2.10) is the *sinc* function; i.e., $g(t) = p(t) * p^*(-t) = \dfrac{\sin(\pi t/T)}{\pi t/T}$ if $f_{\max} = 1/(2T)$. However, such a signal is not practically realizable.

Commonly used pulses $p(t)$ that satisfy the Nyquist criterion are the raised-cosine pulses, which satisfy the "Nyquist band-edge symmetric condition":

$$T - G(W - \Delta) = G^*(W + \Delta)$$

where $\Delta = f - W$. See Fig. 2.11.

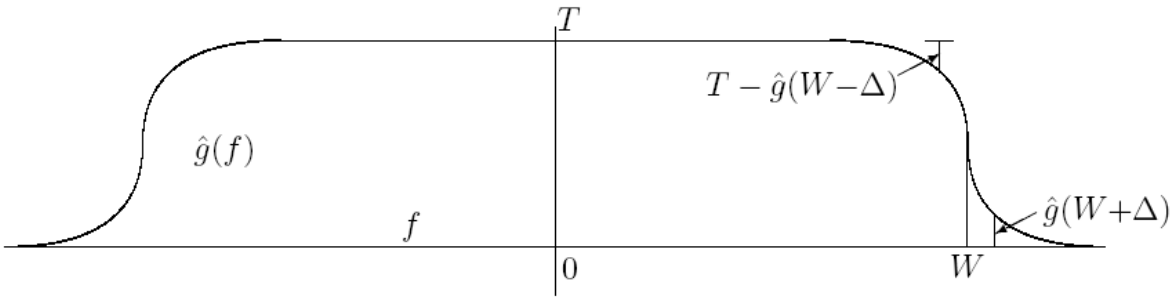


Figure 2.11 Illustration of the spectrum $G(f) = |P(f)|^2$ of a Nyquist pulse.

Therefore we will say that the *nominal* (or *Nyquist*) *bandwidth* of a discrete-time PAM (real) symbol sequence with symbol interval T is $W = 1/2T$. For a QAM (complex) symbol sequence with symbol interval T, the nominal bandwidth is $W = 1/T$, because the baseband interval $[-1/2T, 1/2T]$ is translated to a positive-frequency passband interval $[f_c-1/2T, f_c+1/2T]$ of width $W = 1/T$ on the actual channel.

Remark: *We can send 2W = 1/T real dimensions per second with a PAM system, or W = 1/T complex dimensions per second with a QAM system*.

2.4.4 More discussions on PAM and QAM transmission

Let

$$x(t) = \sum_{k\in\mathbb{Z}} x_k p(t-kT)$$

be the baseband transmitted waveform for a PAM modulator, which is defined by the discrete-time symbol sequence $\{x_k\}$, the symbol interval T and the real waveform $p(t)$ ("pulse"). If $p(t)$ is a Nyquist pulse with a one-sided bandwidth $W=1/(2T)$, then $x(t)$ has a nominal bandwidth W. The actual baseband bandwidth B is defined as the smallest number B such that $X(f) = 0$ for $|f| > B$. To avoid ISI, we must have $B \geq W$. Usually, $W \leq B \leq 2W$. See Fig. 2.11.

Suppose now that the channel is a passband channel. We need to modulate a baseband PAM waveform to some high-frequency passband around some carrier f_c.

- Using double sideband amplitude modulation (AM-DSB), i.e., letting $s_b(t)=x(t)$, the passband PAM signal is given by

  $$s(t) = \sqrt{2}\Re\left[x(t)e^{j2\pi f_c t}\right] = \sqrt{2}x(t)\cos(2\pi f_c t)$$

  which has a passband bandwidth $2W$, going from f_c-W to f_c+W.

- Using AM-SSB, with $s_b(t) = \frac{1}{\sqrt{2}}[x(t) + j\hat{x}(t)]$, the resulting passband signal has only upper-sideband frequency components above f_c. Its bandwidth is $W=1/(2T)$.

For QAM transmission, let $x'(t)$ and $x''(t)$ be two real-valued information-bearing signal. We simutanously modulate them by letting $x(t) = x'(t) + jx''(t)$. In this view, we have the

baseband complex waveform

$$s_b(t) = x(t) = \sum_{k\in\mathbb{Z}} (x_k' + jx_k'')\,p(t-kT) = \sum_k x_k\,p(t-kT)$$

where the complex coefficient $x_k = x_k' + jx_k''$. As in the PAM case, we choose the basic pulse $p(t)$ to be a Nyquist pulse with a one-sided bandwidth W_b=1/(2T). (W_b is called the nominal baseband bandwidth of the QAM modulator.)

The baseband waveform $x(t)$ is then up-converted to a passband around the carrier frequency f_c, producing the real passband waveform (by adding a negative frequency component)

$$s(t) = \frac{1}{\sqrt{2}}\left[x(t)e^{j2\pi f_c t} + x^*(t)e^{-j2\pi f_c t}\right]$$

This can also be written equivalently as

$$s(t) = \sqrt{2}\Re\left[x(t)e^{j2\pi f_c t}\right]$$

$$= \sqrt{2}\Re[x(t)]\cos(2\pi f_c t) - \sqrt{2}\Im[x(t)]\sin(2\pi f_c t)$$

The passband waveform $s(t)$ has a nominal passband bandwidth $W = 2W_b = 1/T$. See Fig. 2.12.

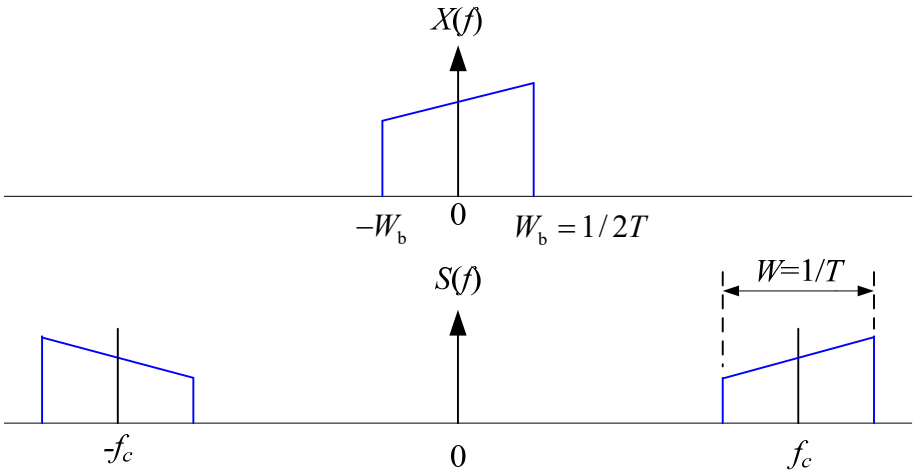


Figure 2.12

- Thus, when talking about PAM, we defined the nominal bandwidth W=1/(2T), where T is the symbol interval. On the other hand, with QAM and passband communication, W=1/T.
- Summary [Degrees of freedom]:
  With PAM, we could choose real symbols separated by T and transmit them in a bandwidth W=1/(2T). Thus, over a long interval T_0, we can transmit $T_0/T = 2WT_0$ real symbols.
  Using QAM, with symbols separated by T we need a passband bandwidth W=1/T.

However, each symbol is complex, and the real and imaginary parts can be selected independently. Thus, again, over a long interval T_0, we can transmit $2WT_0$ real symbols.

Therefore, in principle, QAM and baseband PAM with the same overall bandwidth are equivalent in terms of the number of degrees of freedom that can be used to transmit real symbols.

2.5 Orthogonal (Multipulse) Modulation and Generalized Nyquist Criterion

In baseband PAM, a single pulse shape $p(t)$ is used in one symbol interval, and amplitude modulated by the data symbol x_k. We can generalize this model by allowing the pulse shape in any symbol interval to be chosen from a set of N possibilities, $\{p_n(t), 0 \le n \le N-1\}$, to represent $\log_2 N$ bits of information. The transmitted signal can then be written as

$$x(t) = \sum_{k \in \mathbb{Z}} p_{x_k}(t - kT)$$

where $x_k \in [0, N\text{-}1]$. The data symbol thus indexes which pulse is transmitted in the kth symbol interval, rather than the amplitude of the pulse that is transmitted. If the $\{p_n(t)\}$ is a set of orthogonal pulses, i.e.,

$$\int_{-\infty}^{+\infty} p_i(t) p_j^*(t) dt = \sigma_p^2 \delta_{i-j}$$

then we call this *orthogonal multipulse modulation*.

Consider the transmission of a single isolated pulse $p_n(t)$ for some $0 \le n \le N-1$. Assume that the received signal $y(t)$ is corrupted by AWGN,

$$y(t) = p_n(t) + n(t)$$

An optimum demodulator can be implemented as a bank of matched filters. Define

$$q_n(t) = p_n^*(-t)$$

Then the sampled output of the MF with impulse response $q_n(t)$ is

$$y_n = y(t) * q_n(t)\big|_{t=0} = \int y(\tau) p_n^*(\tau) d\tau$$

A matched-filter receiver is shown in Fig. 2.12.

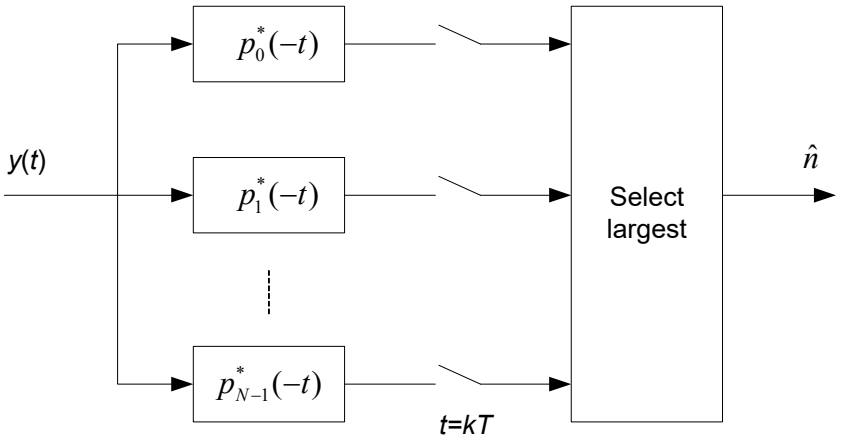


Figure 2.12 MF receiver for multipulse modulation

■ The Generalized Nyquist Criterion

There is a fundamental lower bound on the bandwidth required by an orthogonal multipulse signal, assuming that we wish to avoid ISI. We now generalize the Nyquist criterion aforementioned to the case of orthogonal signal.

Assuming the receiver structure shown in Fig. 2.12. To avoid ISI, if the signal input is pulse $p_n(t)$, then the samples at the output of the filter matched to $p_n(t)$ must satisfy the ordinary Nyquist criterion,

$$p_n(t) * p_n^*(-t)\big|_{t=kT} = \delta_k, 0 \le n \le N-1$$

In addition, to avoid crosstalk between pulses, if $p_n(t)$ is the input to a filter matched to pulse $p_l(t)$ for $l \neq n$, then the output sampled at $t = kT$ must be zero for all k,

$$p_n(t) * p_l^*(-t)\big|_{t=kT} = 0, \quad l \neq n, k \in \mathbb{Z}$$

These conditions can be written together in the form

$$p_n(t) * p_l^*(-t)\big|_{t=kT} = \delta_k \delta_{l-n} \tag{2.11}$$

Let $P_n(f)$ be the Fourier transform of $p_n(t)$. Then (2.11) can be expressed in terms of an equivalent frequency-domain condition as

$$\frac{1}{T}\sum_{m\in\mathbb{Z}} P_n(f+m/T)P_l^*(f+m/T) = \delta_{l-n} \tag{2.12}$$

Equation (2.12) is called the *generalized Nyquist criterion*. Using this, we can show that in order to avoid ISI, the minimum bandwidth of orthogonal multipulse is $N/(2T)$ Hz, or N times the minimum bandwidth of ordinary PAM. A detailed proof is provided in [Lee94, pp. 266].

■ Spectral Efficiency

The best spectral efficiency of orthogonal multipulse is

$$\rho = \frac{\log_2 N}{T \cdot N/(2T)} = \frac{2\log_2 N}{N} \tag{2.13}$$

From $\frac{d\rho}{dN} = 0$, we have the maximum spectral efficiency is at $N = e$. Of course, N must be an integer, so the best we can do is N=3, and the resulting $\rho_{\max} = 1.056$ bits/s/Hz.

PAM and orthogonal multipulse modulation can be combined, resulting in the so-called *multicarrier modulation*. Two currently popular multicarrier modulation schemes are discrete multitone (DMT) and orthogonal frequency-division multiplexing (OFDM).

2.4.1 Combined PAM and Orthogonal (Multipulse) Modulation

Consider general serial/parallel PAM modulation in which real N-vectors $\mathbf{x}_k = [x_{k,0}, x_{k,1}, \cdots, x_{k,N-1}]$ are transmitted at a rate of $1/T$ vectors per second over an ideal AWGN channel. Let $\mathbf{p}(t) = [p_0(t), p_1(t), \cdots, p_{N-1}(t)]$ be an N-vector of real pulses. A general PAM modulator transmits

$$x(t) = \sum_{k\in\mathbb{Z}} \langle \mathbf{x}_k, \mathbf{p}(t-kT) \rangle = \sum_{k\in\mathbb{Z}} \sum_{n=0}^{N-1} x_{k,n} p_n(t-kT) \tag{2.14}$$

(i.e., amplitude-modulating each pulse shape with a different symbol from an alphabet of M symbols). If $\{p_n(t-kT)\}$ is a set of orthonormal functions, the resulting combination is called orthogonal PAM (OPAM). In each symbol interval of length T, N symbols are simultaneously transmitted using N distinct pulses. The spectral efficiency is $\rho = \frac{\log_2 M^N}{T \cdot N/(2T)} = 2\log_2 M$. An optimum demodulator may recover noisy estimates $\mathbf{z}_k$ of $\mathbf{x}_k$ from $y(t) = x(t) + n(t)$ by matched filtering

$$\mathbf{y}_k = \langle y(t), \mathbf{p}(t-kT) \rangle = \int y(t)\mathbf{p}(t-kT)dt = \mathbf{x}_k + \mathbf{n}_k \tag{2.15}$$

The sequence $\{\mathbf{n}_k\}$ is a sequence of i.i.d. Gaussian N-vectors, whose elements have variance σ_n^2. Thus the AWGN waveform channel is converted to a discrete-time vector AWGN channel, which is equivalent to N real discrete-time AWGN channels operating in parallel.

Let $\mathbf{P}(f) = [P_0(f), P_1(f), \cdots, P_{N-1}(f)]$, where $P_n(f)$ is the Fourier transform of $p_n(t)$. Then the frequency-domain condition for orthonormality is the generalized Nyquist criterion

$$\tilde{\mathbf{G}}(f) = \frac{1}{T}\sum_{m\in\mathbb{Z}} \mathbf{G}(f + m/T)$$

$$= \frac{1}{T}\sum_{m\in\mathbb{Z}} \mathbf{P}^{\dagger}(f+m/T)\mathbf{P}(f+m/T)$$

$$= \mathbf{I}_N, \qquad |f| \le 1/(2T) \tag{2.16}$$

where "†" denotes conjugate transpose, $\mathbf{G}(f) = \left\{P_n^*(f)P_{n'}(f), 0 \le n, n' \le N-1\right\}$ is the $N\times N$ matrix, and $\mathbf{I}_N$ is the $N\times N$ identity matrix. It can be shown that the total one-sided bandwidth for which some $P_n(f), f \ge 0$ is nonzero must be at least $N/(2T)$ Hz.

Multicarrier modulation is the generic term used for any OPAM scheme in which each of the orthogonal pulses is roughly localized in the frequency domain. It includes as special cases OFDM and DMT. In OFDM, the pulse set $\{p_n(t)\}$ is

$$p_n(t) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{j2\pi nk/N} g(t-kT/N), \quad \text{for } n = 0, \ldots, N\text{-}1$$

2.5 Discrete-time and Continuous-time Parameters

We now examine the relationship between the parameters of continuous-time and discrete-time AWGN channel models. The parameters involved are, namely, W, SNR, ρ and C.

Continuous-time band-limited AWGN channel model: $y(t) = x(t) + n(t)$

Discrete-time (real or complex) AWGN channel model: $\mathbf{y} = \mathbf{x} + \mathbf{n}$

- The nominal bandwidth W of the continuous-time channel is equal to the inverse $1/T$ of the symbol interval T of the discrete-time channel when the latter is complex, or to $1/2T$ when it is real.

- The nominal spectral efficiency of the continuous-time channel is defined as

  $$\rho = R/W \quad \text{bits/s/Hz}$$

  If the discrete-time channel sends b bits per *real dimension*, then the date rate $R = b/T$ [b/s] for PAM or $R = 2b/T$ [b/s] for QAM. In either case $R = 2bW$, so

  $$\rho = 2b \quad \text{bits/s/Hz} \tag{2.17}$$

  which is the number of bits sent per two real dimensions (bits/2D).

- The average energy per symbol is $E_s = \mathbb{E}\left[|x_k|^2\right]$; the average energy per second, i.e., the average power of $x(t)$, is $P = E_s/T$. Thus,

  $$P = 2E_sW \quad \text{for PAM} \tag{2.17b}$$
  $$P = E_sW \quad \text{for QAM}$$

- For the continuous-time channel, the signal-to-noise ratio (SNR) is defined as

  $$SNR = \frac{\text{signal power}}{\text{noise power}} = \frac{P}{N_0W}$$

For the discrete-time channel, the SNR is defined as

$$SNR = \frac{E_d}{\sigma^2} = \frac{E_N}{N\sigma^2} = \begin{cases} \dfrac{2E_s}{N_0}, & \text{for PAM} \\ \dfrac{E_s}{N_0}, & \text{for QAM} \end{cases}$$

where N is the dimension of the signal constellation under consideration.

The relation $P = E_s/T$ shows that *the SNR is unchanged*. $SNR = \dfrac{E_s / T}{N_0 W}$

- The channel capacity of a continuous-time band-limited AWGN channel is given by

$$C_{[b/s]} = W \log_2(1 + SNR) = W \log_2\left(1 + \frac{P}{N_0 W}\right) \quad \text{b/s} \tag{2.18}$$

(which is obtained with a Gaussian distribution over $x(t)$.)
The capacity of a real or complex discrete-time AWGN channel is given by

$$C_{[b/2D]} = \log_2(1 + SNR) = \log_2\left(1 + \frac{E_s}{N\sigma^2}\right) \quad \text{b/2D} \tag{2.19}$$

regardless of whether it is real or complex. (The derivation of (2.19) is provided in Section 2.6.3)
Since the channel supports W two-dimensional real symbols per second, the capacity is unchanged.

下面研究 PAM/QAM 系统的性能，包括极限性能与未编码系统的错误概率。

2.6 Capacity of Discrete-Time Channels

In this section, we will discuss the capacity of the discrete-time AWGN channel under the constraint that a certain constellation is employed; this is the benchmark for the discussion of performance of coded communication systems.

Consider a discrete-time memoryless channel with input random process $\{X_k\}$ and output $\{Y_k\}$, for which the current output Y_k is independent of all inputs except X_k. Such a channel is fully characterized by the conditional probabilities $P_{Y|X}(y|x)$ for all $x \in \mathcal{A}_X$ and $y \in \mathcal{A}_Y$; i.e.

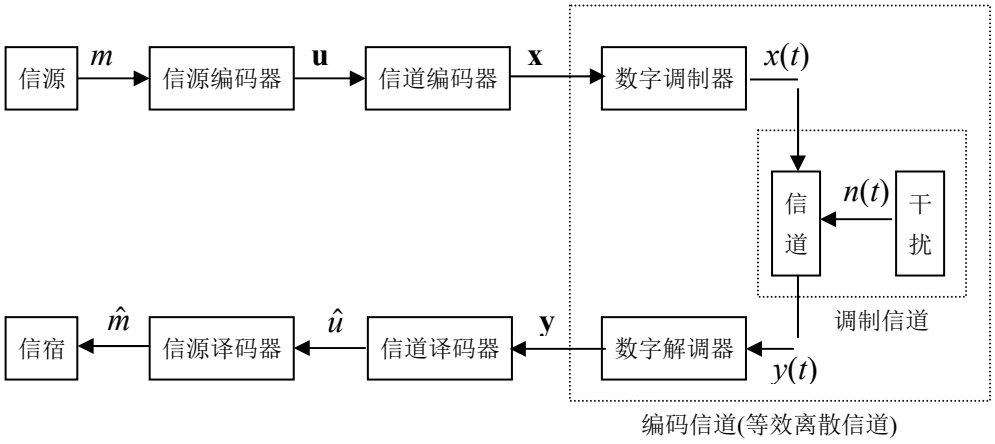
$$P(y_k \mid x_1, ..., x_{k-1}, x_k, y_1, ..., y_{k-1}) = P_{Y|X}(y_k \mid x_k), \quad \text{for } k=1,2,\ldots \tag{2.20}$$

It should be noted that (2.20) implies that the channel is *time-invariant* in the sense that the probability distribution $P_{Y_k|X_k}$ does not depend on k.

For such channels, the ***channel capacity*** is defined as the maximum average mutual information $I(X; Y)$ that can be obtained by choice of $P(x)$, i.e.,

$$C \equiv \max_{P_X} I(X;Y) = \max_{P_X} \; [H(Y) - H(Y \mid X)] \tag{2.21}$$

Note: When $P(X)$ is fixed (e.g., a uniform distribution), the "capacity" would then be $I(X; Y)$, and the max in (2.21) will be removed. In Shannon theory, this type of capacity is more properly called an ***information rate***.



2.6.1 Discrete-Valued Inputs and Outputs

Consider a binary symmetric channel (BSC) with $\mathcal{A}_X = \mathcal{A}_Y = \{0, 1\}$, and $P_{Y|X}(0 \mid 1) = P_{Y|X}(1 \mid 0) = p$. See Fig. 2.13.

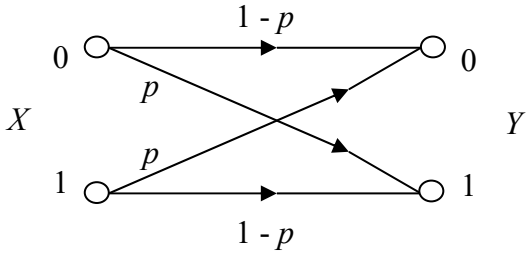


Figure. 2.13 The equivalent discrete channel: BSC

Note that for BSC, $I(X;Y) = H(Y) - H(Y \mid X)$

$$\begin{aligned} &= H(Y) + \sum_x P(x) \sum_y P(y \mid x) \log P(y \mid x) \\ &= H(Y) - H(p) \\ &\le 1 - H(p) \end{aligned} \quad ,$$

where $H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ is the binary entropy function. The equality is achieved when $P(Y=0) = P(Y=1) = 1/2$. This can only occur when $P(X=0) = P(X=1) = 1/2$, so the optimum input distribution is uniform. Hence the capacity of a BSC with crossover p is

$$C_{BSC} = 1 + p \log_2 p + (1-p) \log_2 (1-p) \tag{2.22}$$

If we use the 2-PAM signaling over the AWGN channel, then $p = Q(\sqrt{2E_bR_c / N_0})$ (see Section 2.72), where R_c is the code rate. Let $R_c = C_{BSC}$. Substituting p into (2.22) we can obtain the performance curves of the hard-decision decoding as shown in Fig. 2.15.

For a general discrete memoryless channel (DMC) with M input letters and L output letters, i.e., $\mathcal{A}_X = \{a_1, a_2, ..., a_M\}$ and $\mathcal{A}_Y = \{b_1, b_2, ..., b_L\}$, its capacity is

$$C = \max_{P_X} \sum_{j=1}^{M} \sum_{l=1}^{L} P_X(a_j) P(b_l \mid a_j) \log_2 \frac{P(b_l \mid a_j)}{\sum_{i=1}^{M} P_X(a_i) P(b_l \mid a_i)} \quad \text{bits/symbol}$$

2.6.2 Discrete Inputs and Continuous Outputs

■ One-Dimensional Constellations

Consider the following AWGN channel

$$Y = X + N, \qquad N \sim \mathcal{N}(0, \sigma^2) \tag{2.23}$$

where the channel input X is restricted to take on values in an alphabet $\mathcal{A}_X = \{a_j \in R, 1 \le j \le M\}$, but no restriction is imposed on the channel output, i.e., $\mathcal{A}_Y$ = (-∞, +∞). The channel transition probability density function (p.d.f.) is

$$p(y \mid x) \equiv p_{Y|X}(Y = y \mid X = x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - x)^2}{2\sigma^2} \right) \tag{2.24}$$

The capacity of this channel is given by

$$C = \max_{P_X} \sum_{j=1}^{M} P_X(a_j) \int_{-\infty}^{+\infty} p(y \mid a_j) \log \frac{p(y \mid a_j)}{p(y)} dy \tag{2.25}$$

and the average signal energy per symbol is $E_s = \sum_{j=1}^{M} P_X(a_j) a_j^2$.

If all signal points a_j are used equally likely, the capacity will depend only on the a_j and the maximization will be over all possible choices of the signaling alphabet subject to the power constraint. Then the capacity becomes

$$C = \max_{A_X} \sum_{j=1}^{M} \frac{1}{M} \int_{-\infty}^{+\infty} p(y \mid a_j) \log \frac{p(y \mid a_j)}{M^{-1} \sum_{i=1}^{M} p(y \mid a_i)} dy \tag{2.26}$$

Furthermore, we take the signaling alphabet as given, and consisting of equally spaced real numbers. With these constraints, the capacity (more precisely, the information rate) reduces to

$$C = \sum_{j=1}^{M} \frac{1}{M} \int_{-\infty}^{+\infty} p(y \mid a_j) \log \frac{p(y \mid a_j)}{M^{-1} \sum_{i} p(y \mid a_i)} dy$$

$$= \sum_{j=1}^{M} \frac{1}{M} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-a_j)^2}{2\sigma^2} \right) \log \frac{\exp\{-(y-a_j)^2/2\sigma^2\}}{M^{-1}\sum_i \exp\{-(y-a_i)^2/2\sigma^2\}} dy$$

Let $z = (y-a_j)/\sigma$. The formula becomes

$$C = \log M + \sum_{j=1}^{M} \frac{1}{M} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) \log \frac{\exp(-z^2/2)}{\sum_i \exp\{-(z\sigma + a_j - a_i)^2/2\sigma^2\}} dz$$

Replace a_i with $a'_i\sigma$ (i.e., a_i is normalized by the noise variance), we obtain (the prime on a'_i is droped)

$$C = \log_2 M - \frac{1}{M} \sum_{j=1}^{M} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) \log_2 \sum_{i=1}^{M} \exp\left(-\frac{(a_j - a_i)^2}{2} \right) \exp\left(-z(a_j - a_i) \right) dz \quad (2.27)$$

and $SNR = \sum_j a_j^2 / M$.

Then we employ numerical integration to evaluate C as a function of SNR for each choice of signal constellation. The results are shown in Fig. 2.14. In particular, we compare the C_{BSC} with C for BPSK in Fig. 2.15, which represent the results of hard-decision and soft-decision, respectively. It is seen from Fig. 2.15 that, *in terms of capacity, soft decisions are about 2dB more efficient than hard decisions*.

Note: Equation (2.27) can be expressed as

$$C = \log_2 M - \frac{1}{M} \sum_{j=1}^{M} \mathsf{E}_{p(z)} \left\{ \log_2 \sum_{i=1}^{M} \exp\left(-\frac{(a_j - a_i)^2}{2} \right) \exp\left(-z(a_j - a_i) \right) \right\}$$

$$= \log_2 M - \frac{1}{M} \sum_{j=1}^{M} \mathsf{E}_{p(z)} \left\{ \log_2 \sum_{i=1}^{M} \exp\left(-\frac{(a_j - a_i + z)^2 - z^2}{2} \right) \right\}$$

where the expectation is taken with respect to the random variable z.

- Two-Dimensional Constellations

Now the channel is assumed to be complex Gaussian. The channel input alphabet is a set of M complex numbers. The additive noise is a zero-mean circularly symmetric white Gaussian process, i.e., $N \sim \mathcal{CN}(0, N_0)$. Then the channel output is a complex number with p.d.f.

$$p(y \mid x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|y-x|^2}{2\sigma^2} \right)$$

Let $z \equiv u + jv = (y-a_j)/\sigma$. Using similar methods, the capacity for two-dimensional constellations can be calculated as

$$C = \log M - \sum_{j=1}^{M} \frac{1}{M} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|z|^2}{2} \right) \log \sum_{i=1}^{M} \exp\left(-\frac{|a_j - a_i|^2}{2} \right) \exp\left\{ -\mathrm{Re}[z^*(a_j - a_i)] \right\} du dv$$

By replacing integration by expectation over the noise variable N and noting that $y=a_j+n$ (so $z = n/\sigma$), this equation (or (2.26)) can be written as

$$C^* = \log_2 M - \frac{1}{M}\sum_{j=1}^{M} \mathsf{E}\left\{ \log_2 \sum_{i=1}^{M} \exp\left(-\frac{|\, a_j + n - a_i \,|^2 - |\, n \,|^2}{N_0} \right) \right\} \tag{2.28}$$

(Note: with a_i normalized (so that σ=1), these two equations are equivalent.)
Thus we can evaluate C^* by Monte Carlo average of (2.28) as in (2.100). See Fig. 2.16.

$$C^* = \log_2 M - \frac{1}{MK}\sum_{j=1}^{M}\sum_{k=1}^{K} \log_2 \sum_{i=1}^{M} \exp\left(-\frac{|\, a_j + n_k - a_i \,|^2 - |\, n_k \,|^2}{N_0} \right) \tag{2.100}$$

■ Numerical Results

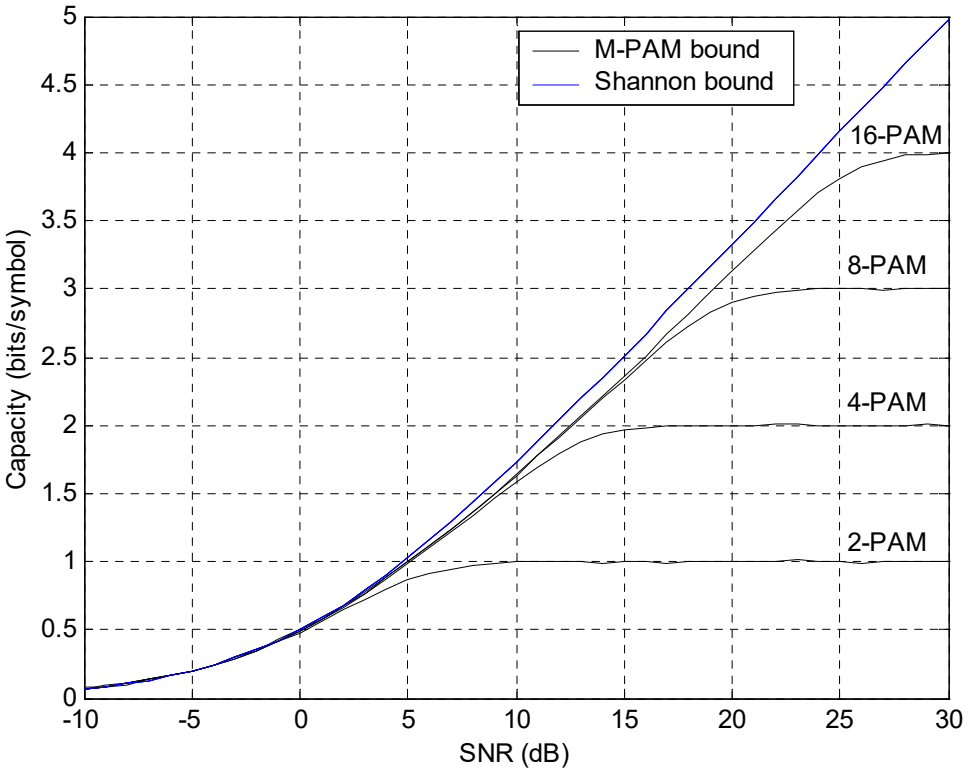


Figure 2.14 Capacity of some one-dimension constellations

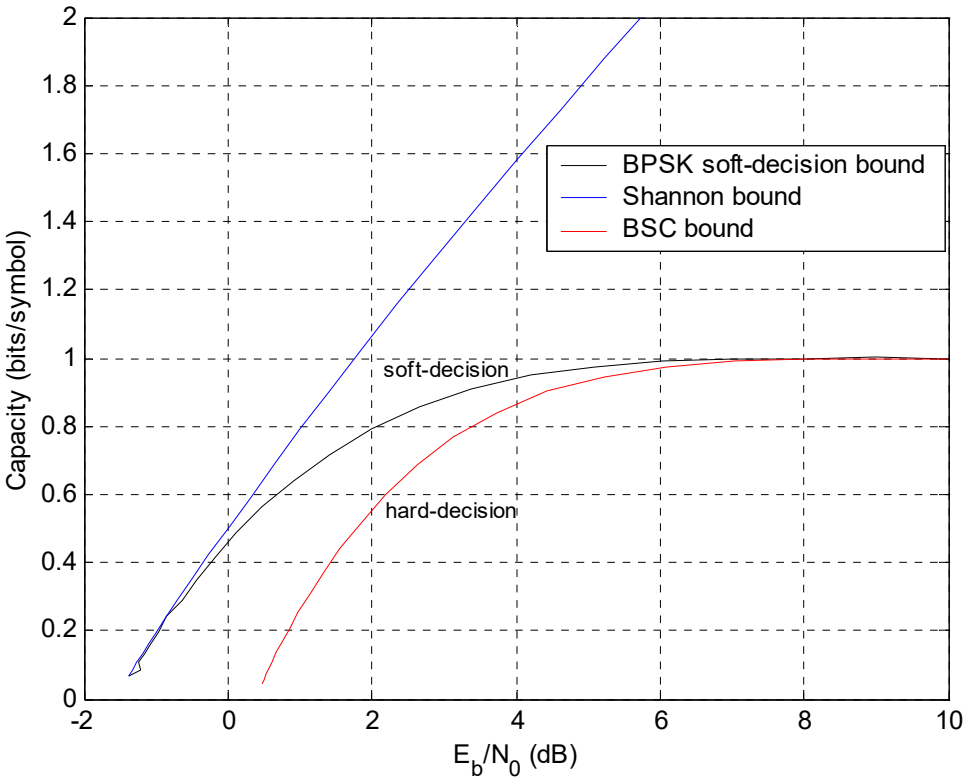
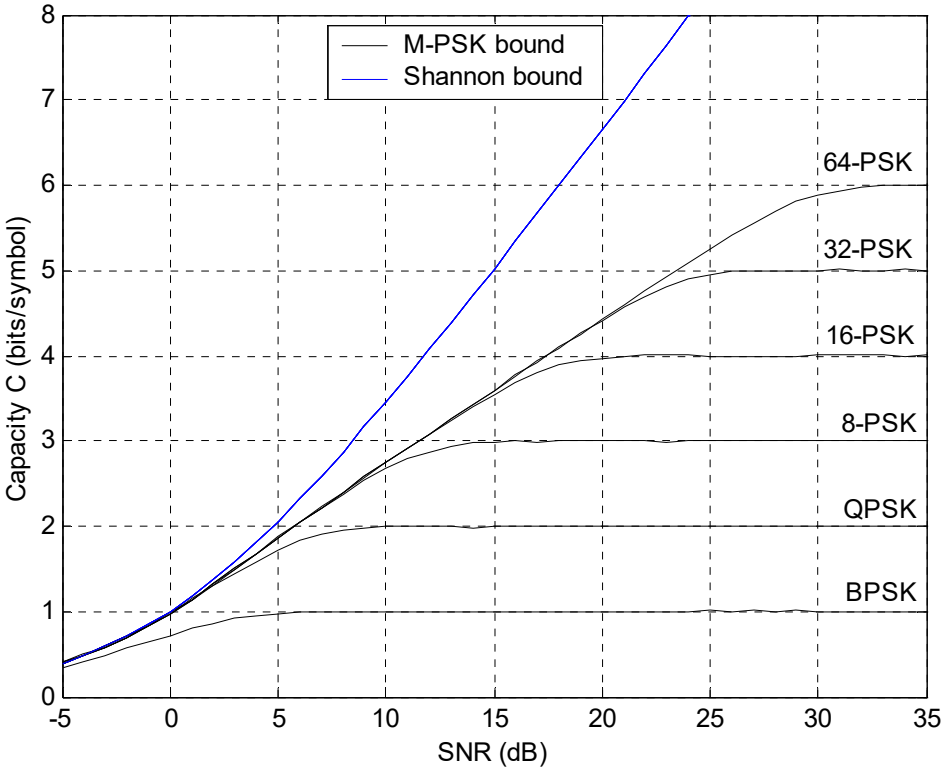
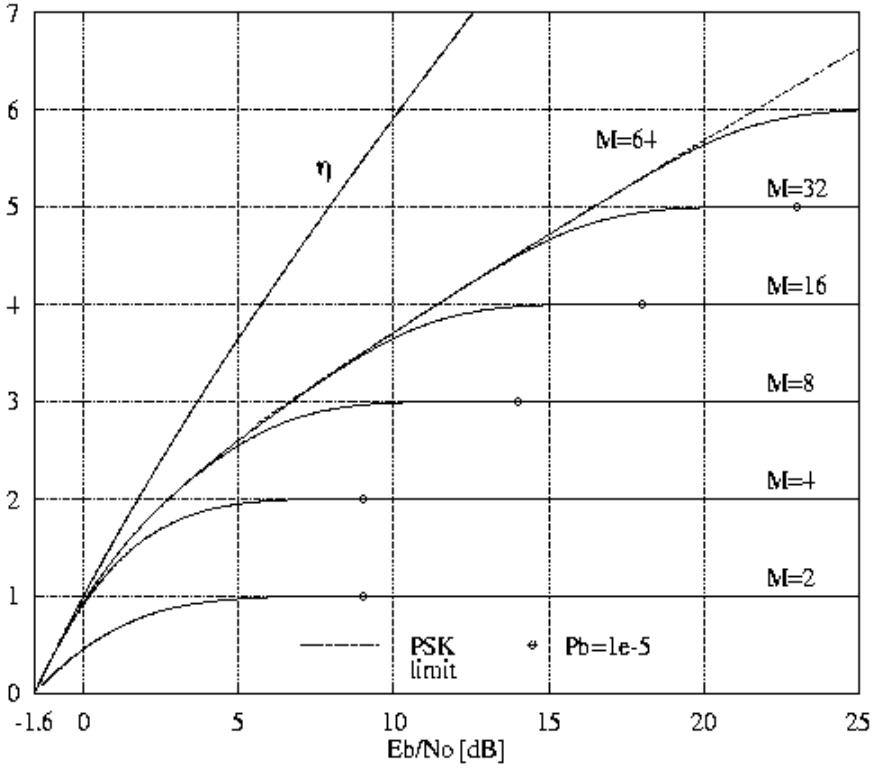
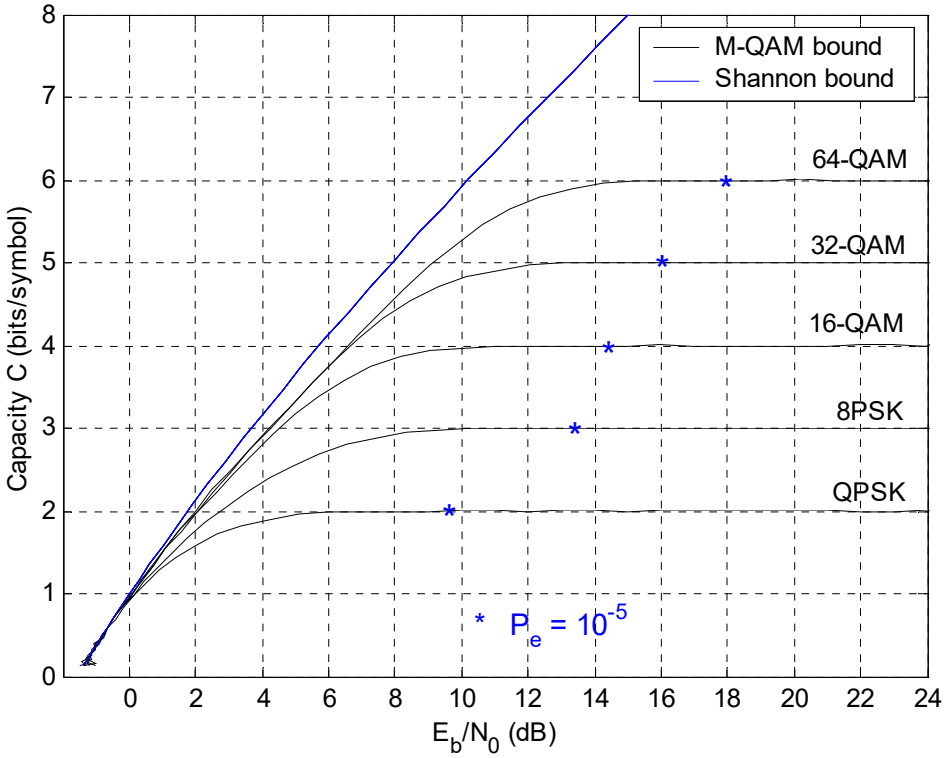


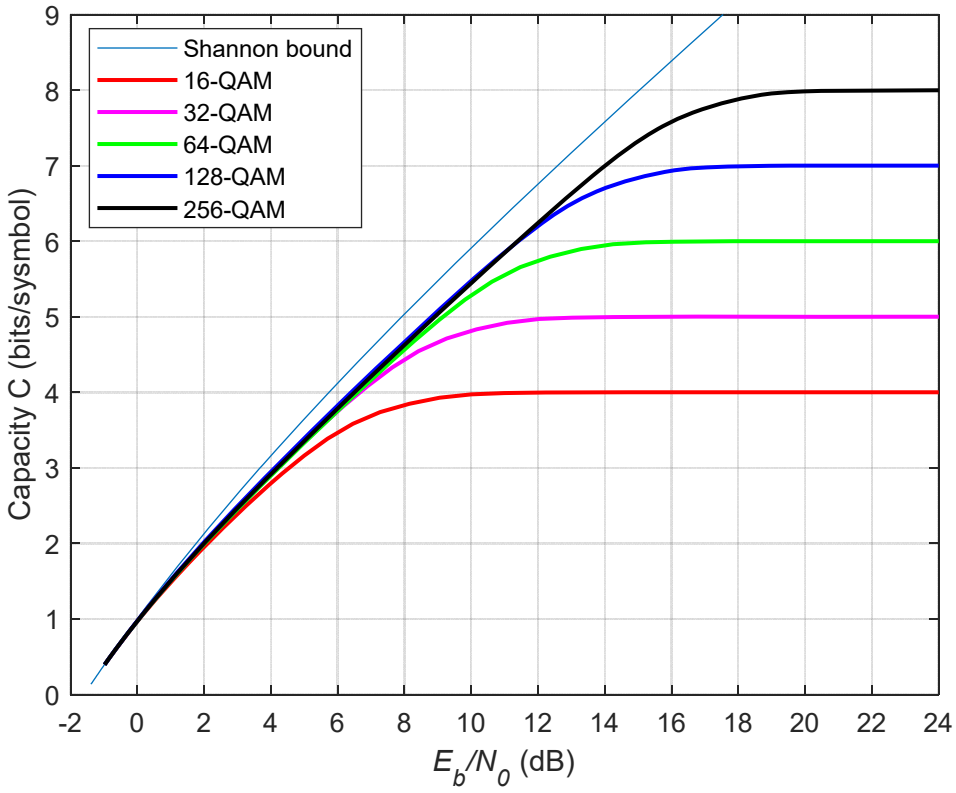
图 2.15　软判决与硬判决译码的信道容量





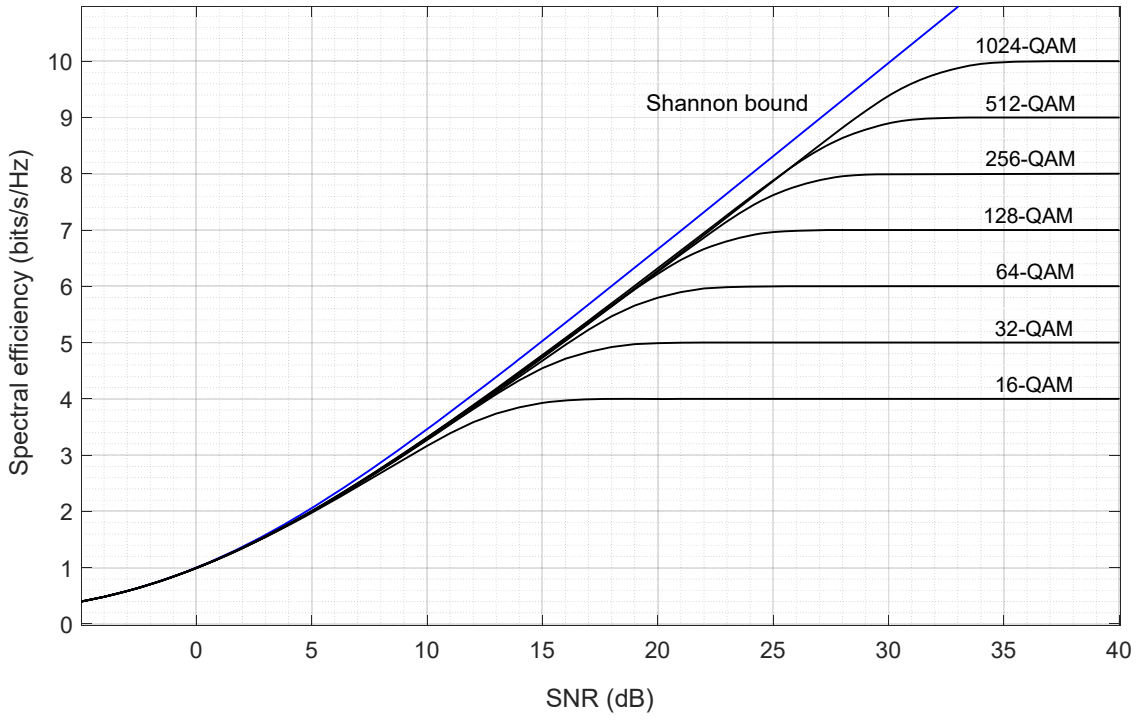
(a) for *M*-PSK constellations





(b) for QAM constellations

Figure 2.16 Capacity of some two-dimension constellations



2.6.3 Continuous-Valued Inputs and Outputs

■ Scalar AWGN Channel

Question: *What is lost by choosing a specific discrete alphabet at the channel input?*

Consider an AWGN channel with a continuous-valued input X,

$$Y = X + N, \qquad N \sim \mathcal{N}(0, \sigma^2)$$

Assuming an average power constraint; that is, $Var[X] = \sigma_X^2$. Since N is independent of X, from (2.21), we have

$$C = \max_{P_X} [H(Y) - H(X + N \mid X)] = \max_{P_X} [H(Y) - H(N)]$$

Note that $H(N) = \frac{1}{2}\log_2 2\pi e\sigma^2$ and $\mathbb{E}[Y^2] = \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[N] + \mathbb{E}[N^2] = \sigma_X^2 + \sigma^2$.

Since the Gaussian distribution maximizes the entropy for a given variance, it follows that

$$\textit{Maximize H(Y)} \rightarrow \textit{Y is Gaussian} \rightarrow \textit{X is Gaussian}$$

That is, $H(Y) \le \frac{1}{2}\log_2[2\pi e(\sigma_X^2 + \sigma^2)]$ with equality iff Y is Gaussian. Thus,

$$\begin{aligned} C &= \frac{1}{2}\log_2[2\pi e(\sigma_X^2 + \sigma^2)] - \frac{1}{2}\log_2 2\pi e\sigma^2 \\ &= \frac{1}{2}\log_2(1 + \sigma_X^2/\sigma^2) \qquad \text{bits per dimension} \end{aligned} \tag{2.29}$$

and the maximum is attained when $X \sim \mathcal{N}(0, \sigma_X^2)$. Other terms for the unit of capacity are also often used, such as "bits per channel use, bits per transmission, and bits per symbol."

■ Vector AWGN Channel

Consider a channel modeled by

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}$$

where **X**, **Y** and **N** are N-vectors, **X** and **N** are independent, and the components of **N** are i.i.d. Gaussian variables each with variance σ^2. The input variance is constrained to $\mathbb{E}[\| \mathbf{X} \|^2] = \sigma_{\mathbf{X}}^2$. The capacity of this channel is given by

$$C = \frac{N}{2}\log_2\left(1 + \frac{\sigma_{\mathbf{X}}^2}{N\sigma^2}\right) \tag{2.30}$$

and the optimal input distribution is a zero-mean Gaussian vector with independent components, with variance $\sigma_{\mathbf{X}}^2 / N$. By noting that $SNR = \frac{\sigma_X^2}{\sigma^2} = \frac{\sigma_{\mathbf{X}}^2}{N\sigma^2}$, we have

$$C_{[b/2D]} = \log_2(1 + SNR) \qquad \text{bits per complex-dimension}$$

for a complex discrete-time AWGN channel.

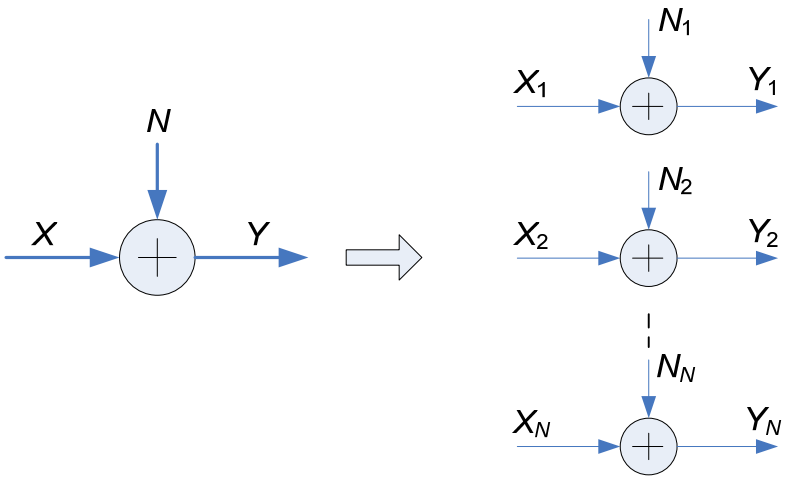


Figure 2.100 Conversion of a vector AWGN channel into multiple scalar channels

2.6.4 Extension to Channels with a Defined Error Rate

In the above derivations for capacity limits, we assumed a required probability of decoding error, P_b, arbitrarily close to zero. When the data is binary and we are willing to tolerate a certain given value of P_b, we can obtain an alternative Shannon capacity.

接收端译码后误比特率为 p_b 的情形等效于对信源做平均 Hamming 失真为 p_b 的信源编码而接收端信道译码无误的情形。According to the Shannon's rate distortion theory, if we are transmitting the output of a binary symmetric source and can tolerate an average distortion of KP_b for a block of K information symbols, then we can represent R bits of information per second with only $R\left(1-h(P_b)\right)$ bits per second, where

$$h(p) = -p\log_2 p - (1-p)\log_2(1-p) \tag{2.31}$$

is the binary entropy function. These $R\left(1-h(P_b)\right)$ bits per second should now be transmitted over the channel with an error probability arbitrarily close to zero. Hence, we have the inequality

$$R\left(1-h(P_b)\right) < C_{[b/s]} \tag{2.32}$$

Substituting (2.18) into (2.32) yields

$$\frac{R\left(1-h(P_b)\right)}{W} = \rho\left(1-h(P_b)\right) < \log_2\left(1+SNR\right) = \log_2\left(1+\frac{\rho E_b}{N_0}\right)$$

or, equivalently,

$$\left.\frac{E_b}{N_0}\right|_{P_b} > \frac{2^{\rho\left(1-h(P_b)\right)}-1}{\rho} \tag{2.33}$$

With one-dimensional signaling, it becomes

$$\left.\frac{E_b}{N_0}\right|_{P_b} > \frac{2^{2R_c(1-h(P_b))}-1}{2R_c}$$

This is the ultimate minimum energy needed by a rate R_c data bits/channel use code that attains P_b. As an example, Fig. 2.17 shows the P_b versus E_b/N_0 for R_c = 1/2 and R_c = 1/3, respectively.

Note:

1）对于一个离散信源 X，其 Hamming 失真定义为

$$d(x,\hat{x}) = \begin{cases} 0, & \text{if } x = \hat{x} \\ 1, & \text{otherwise} \end{cases}$$

显然，平均失真即给出错误概率：$P(X \neq \hat{X}) = \mathsf{E}d(x,\hat{x})$.

2）序列 x^n 和 $\hat{x}^n$ 之间的失真定义为

$$d(x^n,\hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n} d(x_i,\hat{x}_i)$$

令 $\hat{x}^n = g(f(x^n))$ 表示 x^n 经信源编译码之后的估计值，一个 $(2^{nR},n)$ 率失真码的失真为

$$D = \sum_{x^n} P(x^n) d(x^n,\hat{x}^n)$$

3）在 Hamming 失真度量下，一个二进制信源 $X \sim \text{Bernoulli}(p)$ 的率失真函数为

$$R(D) = H(p) - H(D),\ 0 \le D \le \min\{p, 1-p\}$$

它给出了失真为 D 时信源编码的最小可达速率。

■ An Alternative Derivation of (2.33)

The inverse to the coding theorem states that

$$P_b \log(M-1) + h(P_b) \ge H(U) - \frac{\tau_s}{\tau_c} C \tag{2.34}$$

For equiprobable binary source, $M = 2$, $H(U) = 1$. Noting that $\frac{\tau_s}{\tau_c} = \frac{N}{K} = \frac{1}{R_c}$, we have

$$h(P_b) \ge 1 - \frac{1}{R_c} C \implies C \ge R_c\left(1 - h(P_b)\right) \quad \text{bits/symbol} \tag{2.35}$$

Error free capacity (bits/symbol)

Transmission rate with BER>0 i.e., C_{ber} (bits/symbol)

From (2.19),

$$C = \frac{1}{2}\log_2\left(1+\frac{2R_cE_b}{N_0}\right) \quad \text{bits/dim} \tag{2.36}$$

Combining (2.35) and (2.36) yields

$$\left.\frac{E_b}{N_0}\right|_{P_b} > \frac{2^{2R_c(1-h(P_b))}-1}{2R_c} = \frac{2^{\rho(1-h(P_b))}-1}{\rho} \tag{2.37}$$

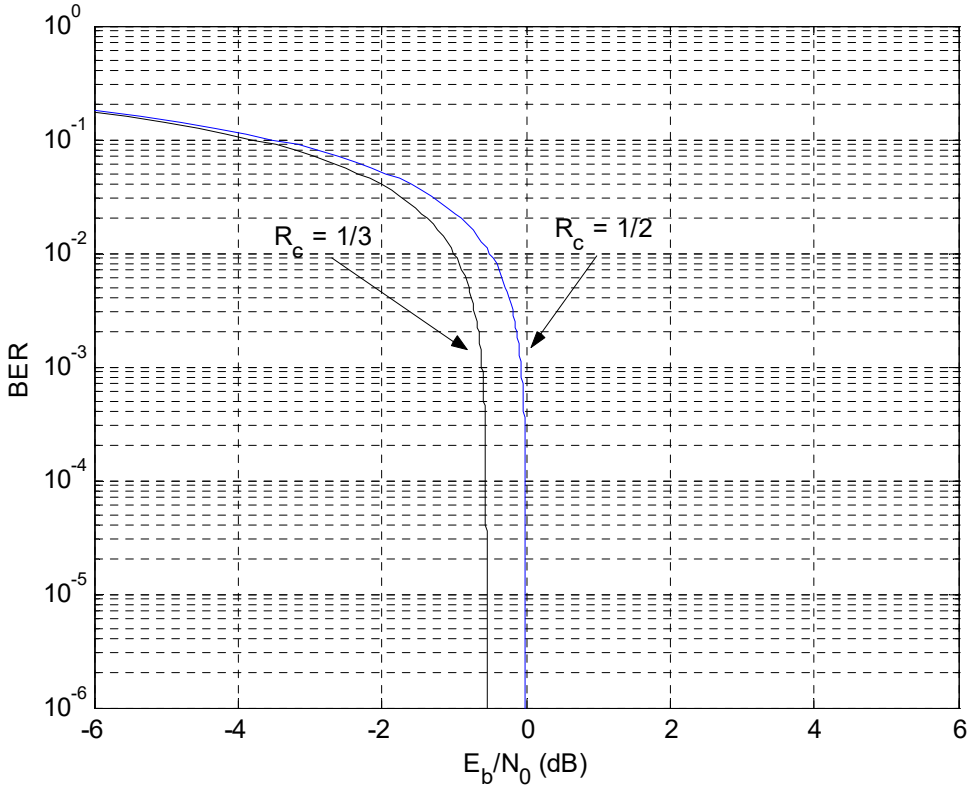


Figure 2.17 Shannon capacity for nonzero bit error rates

对于有限字符集，在某一调制方式下（例如 *M*-QAM），对于给定的信噪比 SNR，我们可计算其对应的容量 $C(SNR)$。由式（2.32），对于给定的谱效率 ρ 和错误概率 P_b，达到该性能要求需要满足

$$\rho\left(1-h(P_b)\right) < C\left(SNR\right)$$

因此，所需的信噪比要求为

$$SNR > SNR_{\min}, \text{ where } SNR_{\min} \text{满足} C_{\text{QAM}}(SNR_{\min}) = \rho\left(1-h(P_b)\right)$$

相应地，有

$$\left[\frac{E_b}{N_0}\right]_{\min} = \frac{SNR_{\min}}{\rho}$$

图 2.18 给出了 16-QAM 调制下，不同谱效率在不同错误概率要求下所对应的容量限。

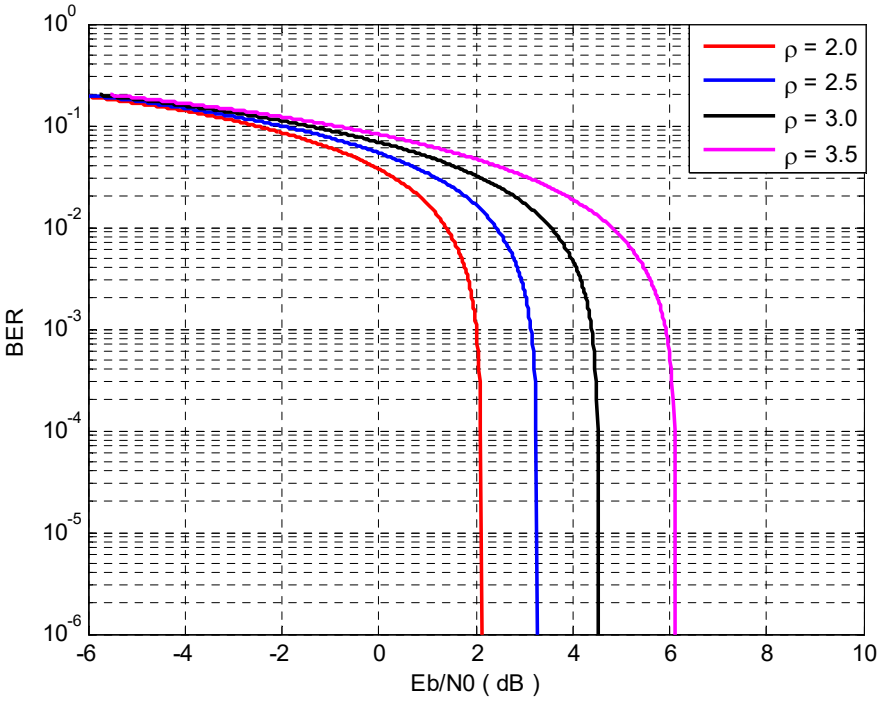


图 2.18　16-QAM 调制下，给定错误的容量限曲线

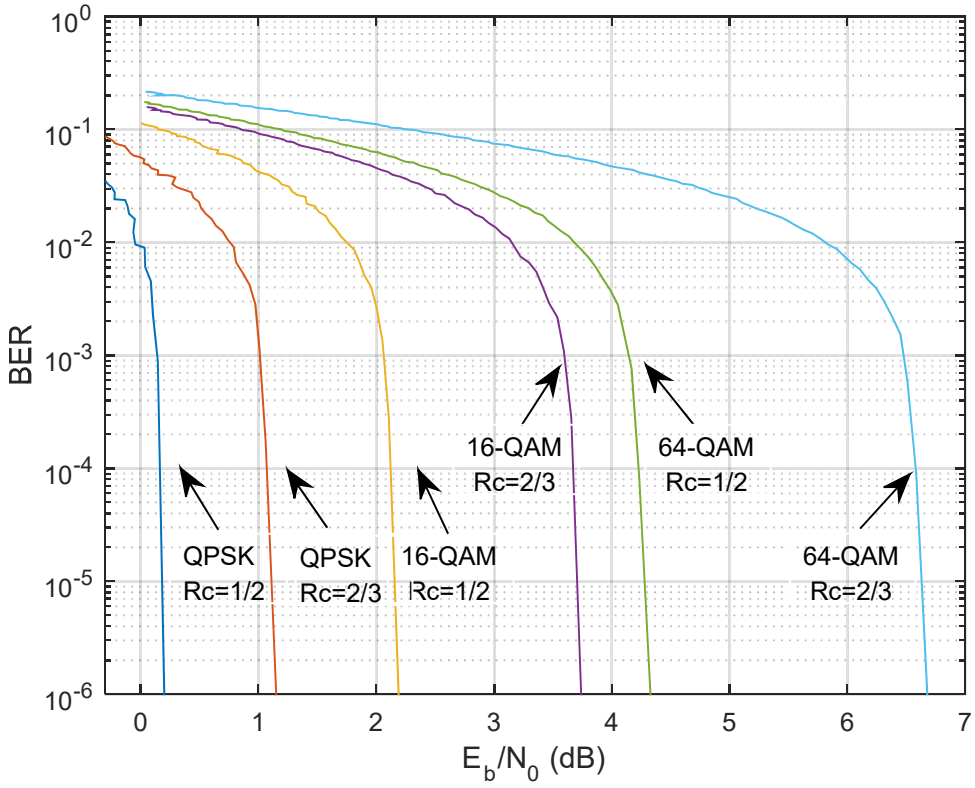


图 2.19 不同码率与调制模式下的给定错误的容量限曲线

2.7 Normalized SNR and Potential Coding Gain

2.7.1 Shannon Limit and Normalized SNR

From (2.19), in order to be capable of transmitting at rate R b/s at an arbitrarily small error probability, the spectral efficiency must satisfy

$$\rho = R/W < \log_2(1+SNR) = C_{[b/2D]}$$

This result is called the *Shannon limit*. An equivalent statement is $SNR > 2^{\rho} - 1$.
This motivates defining the *normalized SNR* as

$$SNR_{\text{norm}} = \frac{SNR}{2^{\rho}-1} \tag{2.38}$$

(It summarizes SNR and ρ in a single parameter so that the Shannon limit is very simply expressed in terms of SNR_{norm}.)
For a capacity-approaching modulation and coding scheme, $\rho = C_{[b/2D]}$ and $SNR_{\text{norm}} = 1$ (0dB). If $\rho < C_{[b/2D]}$, as will always be the case in practice, then

$$SNR_{\text{norm}} > 1$$

The values of SNR_{norm} in dB thus measures how far a given coding and modulation scheme is operating from the Shannon limit (the "*gap to capacity*").

2.7.2 E_b/N_0

Another commonly used normalized measure of SNR is E_b/N_0, where E_b denotes the *average (received) energy per information bit*. From

$$SNR = \frac{P}{N_0 W} = \frac{E_b}{N_0} \cdot \frac{R}{W} = \frac{E_b}{N_0} \times \rho$$

we have

$$\frac{E_b}{N_0} = \frac{SNR}{\rho} = \frac{2^\rho - 1}{\rho} SNR_{\text{norm}} \tag{2.39}$$

and (using (2.17b))

$$E_b = \begin{cases} 2E_s / \rho, & \text{for PAM} \\ E_s / \rho, & \text{for QAM} \end{cases} \tag{2.40}$$

The Shannon limit in terms of E_b/N_0, may be expressed as

$$\frac{E_b}{N_0} > \frac{2^\rho - 1}{\rho} \tag{2.41}$$

This lower bound is a monotonic function of ρ.

- For $\rho = 1$, $E_b/N_0 > 1$ (0dB).
- For $\rho = 2$, $E_b/N_0 > 1.76$ dB.
- As $\rho \to 0$, it approaches the *ultimate Shannon limit:* $E_b/N_0 > \ln 2$ (-1.59dB).

Note 1: Consider a coded communication system in which a one- or two-dimensional signaling constellation, $\mathcal{A} = \{a_j, 1 \le j \le M\}$, of size M is used in conjunction with a rate-R_c FEC code.

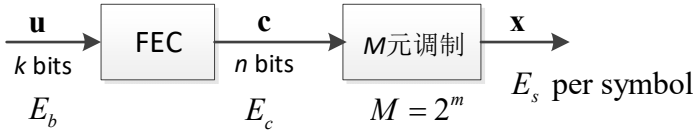


图 2.20

The amount of energy for the n coded symbols has to be equal to the amount of energy for k data symbols, i.e. $kE_b = nE_c$ or $E_c = \frac{k}{n} E_b = R_c E_b$. And for M-ary modulation, the average energy per modulated symbol is $E_s = (\log_2 M) E_c = mE_c$.

From the discussion above, we have

$$E_b = \frac{E_s}{R_c \log_2 |\mathcal{A}|}, \text{ and}$$

$$\frac{E_b}{N_0} = \frac{E_s}{N_0} \frac{1}{R_c \log_2 |\mathcal{A}|}$$

Note 2:

1) The symbol rate for a FEC code is $R_s = K / N$.

2) If the source emits binary digits with the probability $P(U=1) = p$, then the information rate $R_c = H_2(p)\frac{K}{N}$ [bits per coded symbol]

2.7.3 Power-Limited and Bandwidth-Limited Channels

In general, we take ρ = 1 (b/s/Hz) as the boundary to divide the ideal band-limited AWGN channels into two categories: power-limited channels and bandwidth-limited channels. The behavior of capacity is different in the two regimes. In the *power-limited regime* (small SNR), we have

$$\rho \le C_{[b/2D]} \approx SNR \log_2 e$$
$$SNR_{\text{norm}} \approx (E_b / N_0) \log_2 e$$

Clearly, the capacity goes *linearly* with SNR.

On the other hand, in the *bandwidth-limited regime* (large SNR), we have

$$\rho \le C_{[b/2D]} \approx \log_2 SNR$$
$$SNR_{\text{norm}} \approx SNR / 2^{\rho}$$

The capacity goes *logarithmically* with SNR. Every additional 3dB in SNR yields an increase in achievable spectral efficiency of only 1 (b/s/Hz).

2.7.4 Coding Gain

To compare the coded and uncoded schemes it is common to use the coding gain parameter. The *effective coding gain* of a coded modulation scheme is the difference between the signal-to-noise ratios required to achieve a certain target error probability with a coded modulation scheme and to obtain the same error probability with an uncoded scheme of the same throughput. That is,

$$\gamma_{\text{eff}} = \left(\frac{E_b}{N_0}\right)_{\text{uncoded}} - \left(\frac{E_b}{N_0}\right)_{\text{coded}}, \text{ or } \gamma_{\text{eff}} = [SNR_{\text{norm}}]_{\text{uncoded}} - [SNR_{\text{norm}}]_{\text{coded}}$$

- In the power-limited regime, we usually use 2-PAM (or equivalent 4-QAM) as a baseline uncoded modulation scheme;
- In the bandwidth-limited regime, we will use M-PAM (or equivalent $(M\times M)$-QAM) as a baseline uncoded modulation scheme.

2.7.5 Performance of Baseline Uncoded Scheme

- 2-PAM

With a 2-PAM signal set $\mathcal{A}=\{\pm d/2\}$, the probability of error of a maximum likelihood (ML) detector is given by

$$\Pr(e) = Q\left(\frac{d}{2\sigma}\right)$$

where $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} e^{-x^2/(2\sin^2\phi)} d\phi$. Substituting $E_b = d^2 / 4$ and $\sigma^2 = N_0 / 2$,

we have the *error probability per bit*

$$P_b(e) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{2.42}$$

To achieve $P_b(e) \approx 10^{-5}$, we will need to achieve $E_b/N_0 \approx 9.6$ dB. The nominal spectral efficiency of 2-PAM is $\rho = 2$ (b/2D), and the corresponding Shannon limit is

$$E_b / N_0 > \frac{2^\rho - 1}{\rho} = 3/2 \text{ (1.76 dB)}.$$

Thus, at $\rho = 2$ and $P_b(e) \approx 10^{-5}$, we can achieve a coding gain of up to about 8dB with powerful codes. As $\rho \to 0$, a coding gain of over 11dB can be achieved.

- *M*-PAM and (*M*×*M*)-QAM

  For a *d*-spaced *M*-PAM, the probability of error per dimension is

$$\Pr(e) = \frac{2(M-1)}{M} Q\left(\frac{d}{2\sigma}\right) \approx 2Q\left(\frac{d}{2\sigma}\right) \quad (M \text{ large})$$

The probability of error per two dimensions is, therefore,

$$P_s(e) = 1 - \left(1 - \Pr^{M-\text{PAM}}(e)\right)^2 = 2\Pr^{M-\text{PAM}}(e) - \left(\Pr^{M-\text{PAM}}(e)\right)^2 \approx 2\Pr(e)$$

$$P_s(e) = 2\Pr(e) \approx 4Q\left(\frac{d}{2\sigma}\right)$$

Substituting $E_s = d^2(M^2-1)/6, \sigma^2 = N_0/2$, and

$$SNR_{\text{norm}} = \frac{SNR}{2^\rho - 1} = \frac{E_s / N_0}{(M^2 - 1)}$$

we obtain

$$P_s(e) \approx 4Q\left(\sqrt{\frac{3E_s}{(M^2-1)N_0}}\right) = 4Q\left(\sqrt{3SNR_{\text{norm}}}\right) \tag{2.43}$$

If we want to achieve $P_s(e) \approx 10^{-5}$ with uncoded (*M*×*M*)-QAM, then we need to achieve $SNR_{\text{norm}} \approx 8.4$ dB.

Fig. 2.20 shows the $P_b(e)$ versus E_b/N_0 for uncoded 2-PAM, and $P_s(e)$ versus SNR_{norm} for uncoded *M*-PAM.

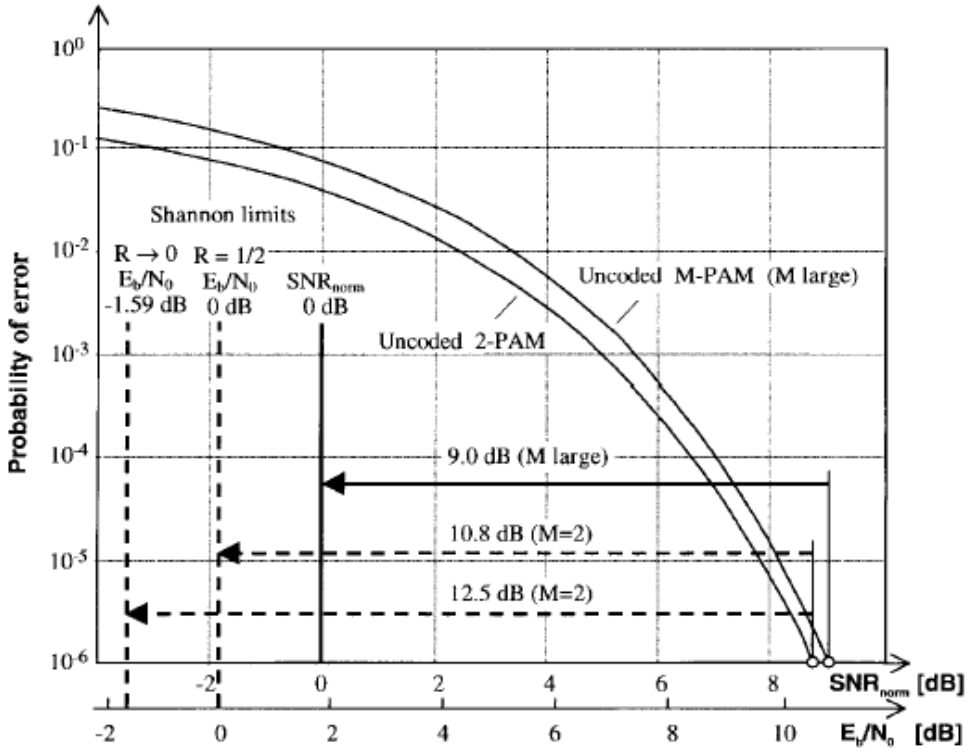
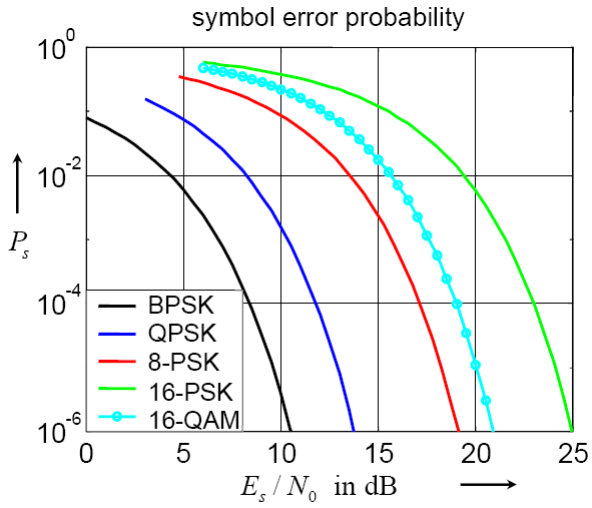
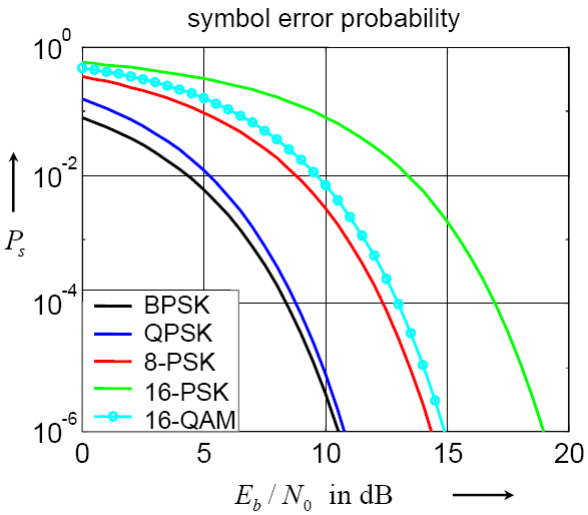
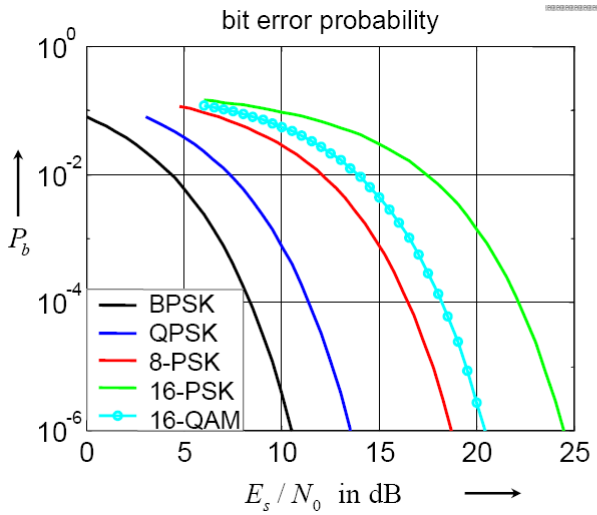
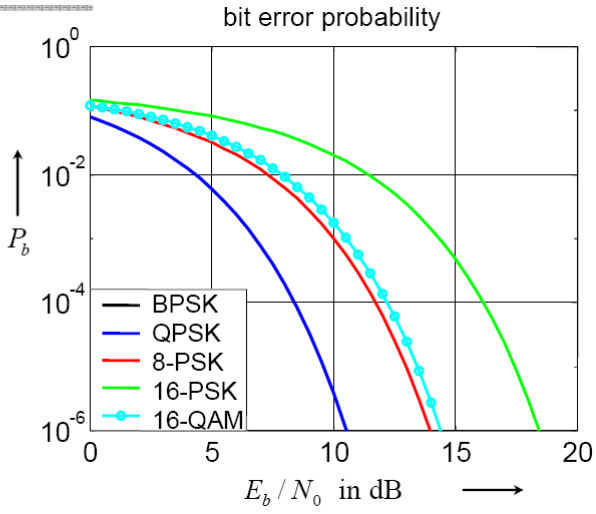


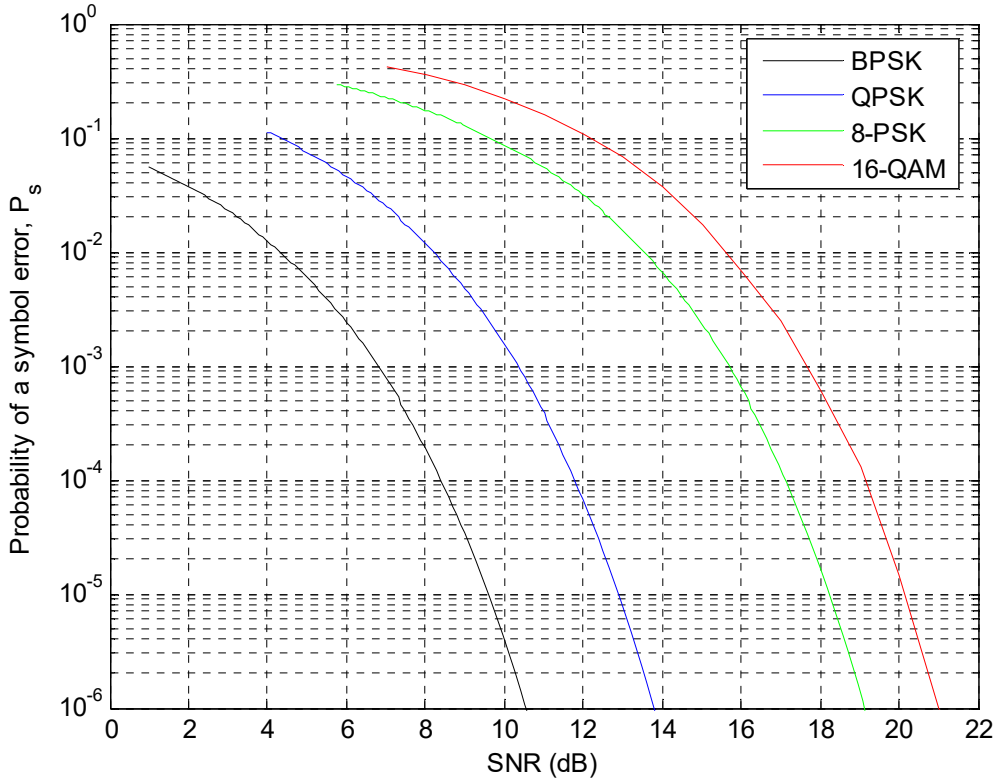
Figure 2.20 (From Forney-TIT1998)

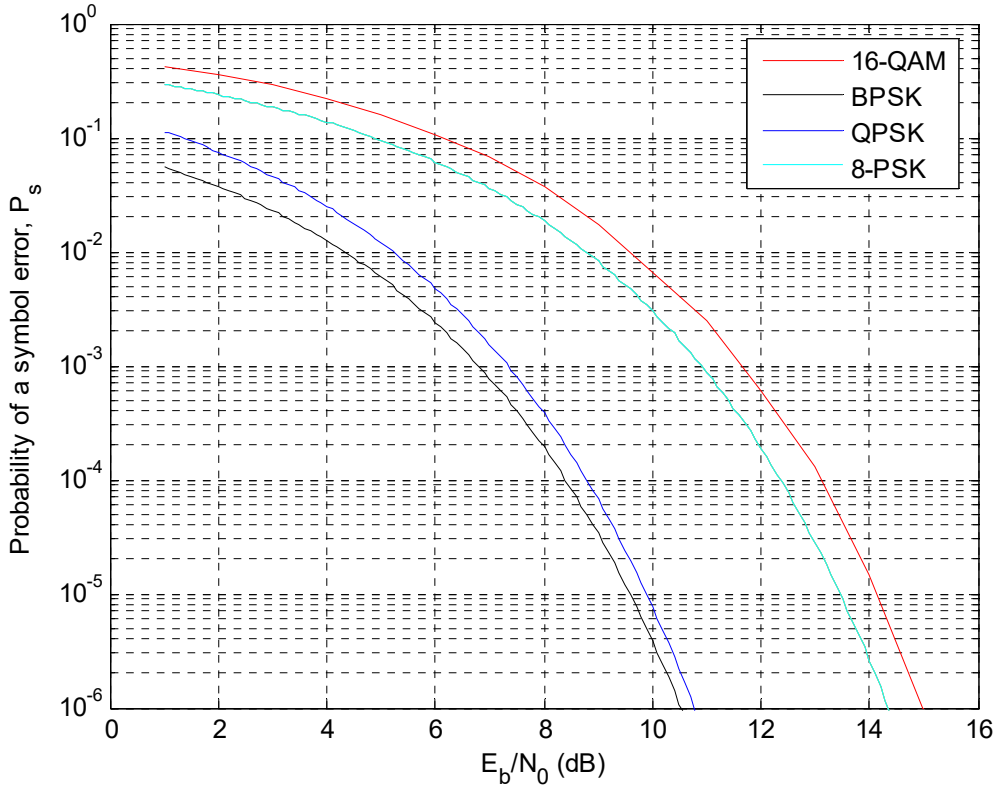

symbol error probability
10^0
10^{-2}
10^{-4}
10^{-6}
P_s
BPSK
QPSK
8-PSK
16-PSK
16-QAM
0
5
10
15
20
25
E_s / N_0 in dB

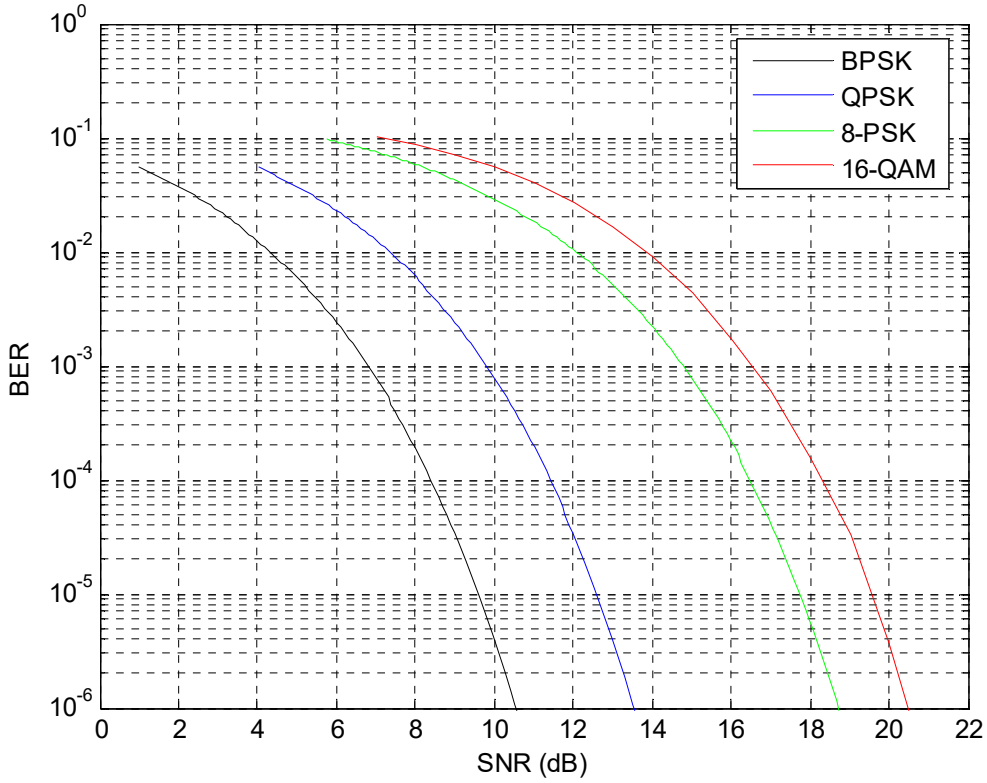

symbol error probability
10^0
10^{-2}
10^{-4}
10^{-6}
P_s
BPSK
QPSK
8-PSK
16-PSK
16-QAM
0
5
10
15
20
E_b / N_0 in dB


bit error probability
10^0
10^{-2}
10^{-4}
10^{-6}
P_b
BPSK
QPSK
8-PSK
16-PSK
16-QAM
0
5
10
15
20
25
E_s / N_0 in dB


bit error probability
10^0
10^{-2}
10^{-4}
10^{-6}
P_b
BPSK
QPSK
8-PSK
16-PSK
16-QAM
0
5
10
15
20
E_b / N_0 in dB


BPSK
QPSK
8-PSK
16-QAM
Probability of a symbol error, P_s
SNR (dB)


16-QAM
BPSK
QPSK
8-PSK
Probability of a symbol error, P_s
E_b/N_0 (dB)



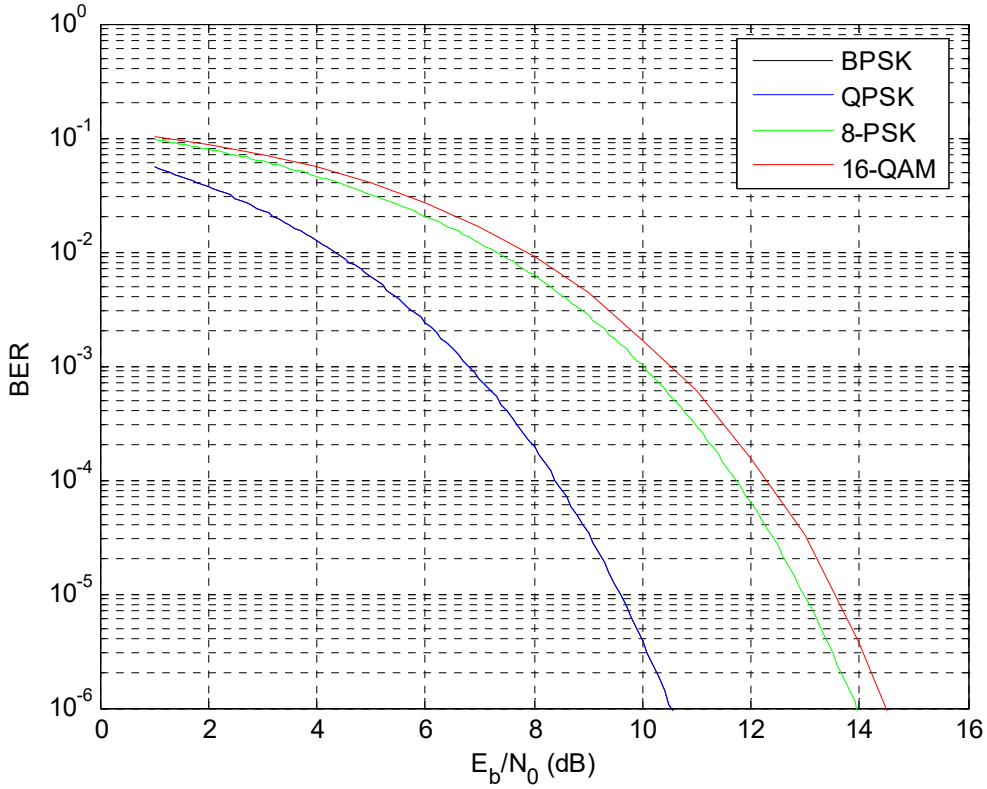


Figure 2.21 Error probability on the AWGN channel. $P_b \approx \frac{1}{m} P_s$

2.8 Performance of Small Signal Constellations

In this section, we consider the design and performance characteristics of optimum receivers for small signal constellations. It is convenient to subdivided the receiver into two parts: the *signal demodulator* and the *detector*, as shown in Fig. 2.22. Usually, we use the correlator or matched-filter as the signal demodulator, which coverts the received waveform into an N-dimensional vector. (Refer to Fig. 2.7.)

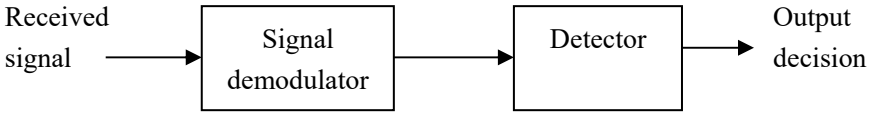


Figure 2.22 Receiver structure

2.8.1 Discrete-Time Channel Model

We have seen that with orthonormal PAM or QAM, the channel model reduces to an analogous real or complex discrete-time AWGN channel model

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}$$

where $\mathbf{X}$ represents the (random) input signal point sequence, and $\mathbf{N}$ is an i.i.d. Gaussian noise sequence with zero mean and variance $\sigma^2 = N_0 / 2$ per real dimension.

In this section, we consider N-dimensional signal constellations for this channel. Assume that the transmitted random vector $\mathbf{X}$ is chosen equiprobably from the M N-vectors in $\mathcal{A} = \{\mathbf{a}_j, 1 \le j \le M\}$, where $\mathbf{a}_j \in \mathbb{R}^N$. Both $\mathbf{Y}$ and $\mathbf{N}$ are also N-vectors. The noise p.d.f. is

$$p_N(\mathbf{n}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\|\mathbf{n}\|^2}{2\sigma^2}\right)$$

2.8.2 Optimum Data Detection

Let $\mathbf{a}$ be the transmitted vector, and $\hat{\mathbf{a}}$ be the output decision of the detector. The probability of decision error is defined as $\Pr(e|\mathbf{a} \text{ sent}) = \Pr(e) = P(\hat{\mathbf{a}} \neq \mathbf{a})$. The probability of being correct is $\Pr(c) = 1 - \Pr(e) = P(\hat{\mathbf{a}} = \mathbf{a})$. In digital communications, we are usually interested in the *minimum-probability-of-error* (MPE) decision rule: given a received vector $\mathbf{Y} = \mathbf{y}$, choose the signal point $\hat{\mathbf{a}} \in \mathcal{A}$ to minimize $\Pr(e)$.

■ <u>The MAP detector</u>

Since

$$\begin{aligned}\Pr(e) &= P(\hat{\mathbf{a}} \neq \mathbf{a}) \\ &= \int_{\mathcal{Y}} p_{\mathbf{Y}}(\mathbf{y}) P(\hat{\mathbf{a}} \neq \mathbf{a} \mid \mathbf{Y} = \mathbf{y}) d\mathbf{y} \\ &= \int_{\mathcal{Y}} p_{\mathbf{Y}}(\mathbf{y}) \left[1 - P(\hat{\mathbf{a}} = \mathbf{a} \mid \mathbf{Y} = \mathbf{y})\right] d\mathbf{y}\end{aligned}$$

$$= 1 - \int_{\mathcal{Y}} p_{\mathbf{Y}}(\mathbf{y}) P(\hat{\mathbf{a}} = \mathbf{a} \mid \mathbf{Y} = \mathbf{y}) d\mathbf{y},$$

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 MAP decision rule can be expressed as

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}_j} \{p(\mathbf{y} \mid \mathbf{a}_j) P(\mathbf{a}_j)\}$$

■ The ML detector

If the transmitted signal vectors $\mathbf{a}_j$ are equiprobable, so $P(\mathbf{a}_j) = 1/M$ for all j, then the MAP decision rule becomes the *maximum likelihood* (ML) decision rule: choose the $\hat{\mathbf{a}} \in \mathcal{A}$ such that $P(\mathbf{y} \mid \hat{\mathbf{a}})$ is maximum among all $P(\mathbf{y} \mid \mathbf{a}_j), \mathbf{a}_j \in \mathcal{A}$. That is,

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}_j} \{p(\mathbf{y} \mid \mathbf{a}_j)\}$$

■ The MD detector

For AWGN channels, using the noise p.d.f., we have

$$p(\mathbf{y} \mid \mathbf{a}_j) = p_{\mathbf{N}}(\mathbf{y} - \mathbf{a}_j) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\| \mathbf{y} - \mathbf{a}_j \|^2}{2\sigma^2} \right)$$

Clearly, the ML rule is equivalent to the *minimum distance* (MD) rule: choose the $\hat{\mathbf{a}} \in \mathcal{A}$ such that $\| \mathbf{y} - \hat{\mathbf{a}} \|^2$ is minimum among all $\| \mathbf{y} - \mathbf{a}_j \|^2, \mathbf{a}_j \in \mathcal{A}$. It can be expressed as

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}_j} \| \mathbf{y} - \mathbf{a}_j \|^2$$

Define $d^2(\mathbf{y}, \mathbf{a}_j) = \| \mathbf{y} - \mathbf{a}_j \|^2 = \sum_{k=1}^{N} (y_k - a_{jk})^2$. We will call $d^2(\mathbf{y}, \mathbf{a}_j)$, $\forall j$, the squared distance metrics.

Note that $\| \mathbf{y} - \mathbf{a}_j \|^2 = \| \mathbf{y} \|^2 - 2\,\mathrm{Re}\{< \mathbf{y}, \mathbf{a}_j >\} + \| \mathbf{a}_j \|^2$, the MD rule reduces to choosing the signal that maximizes the following metric

$$c(\mathbf{y}, \mathbf{a}_j) = -2\,\mathrm{Re}\{< \mathbf{y}, \mathbf{a}_j >\} + \| \mathbf{a}_j \|^2$$

In summary, under the assumption of equiprobable inputs and i.i.d. Gaussian noise, the MD rule is the MPE rule. Therefore we will consider only MD detection.

2.8.3 Decision Regions

Under a MD decision rule, real N-space $\mathbb{R}^N$ is partitioned into M disjoint decision regions $\mathcal{R}_j$, $1 \le j \le M$, where $\mathcal{R}_j$ consists of the received vectors $\mathbf{y} \in \mathbb{R}^N$ that are at least as close to $\mathbf{a}_j$ as to any other point in $\mathcal{A}$:

$$\mathcal{R}_j = \{\mathbf{y} \in \mathbb{R}^N : \| \mathbf{y} - \mathbf{a}_j \|^2 \le \| \mathbf{y} - \mathbf{a}_i \|^2, \forall i \ne j\}$$

$$\begin{cases} \mathbb{R}^N = \bigcup_j \mathcal{R}_j \\ \mathcal{R}_j \cap \mathcal{R}_i = \phi, \quad i \ne j \end{cases}$$

Under the MD rule, given a received vector $\mathbf{y}$, the decision is $\mathbf{a}_j$ if and only if $\mathbf{y} \in \mathcal{R}_j$. The decision regions $\mathcal{R}_j$ are disjoint except on their boundaries. The boundary between two regions $\mathcal{R}_j$ and $\mathcal{R}_i$, if any, consists of the points in $\mathbb{R}^N$ that satisfy the following equivalent equations:

$$\| \mathbf{y} - \mathbf{a}_j \|^2 = \| \mathbf{y} - \mathbf{a}_i \|^2$$

$$< \mathbf{y}, \mathbf{a}_i - \mathbf{a}_j > = \left\langle \frac{\mathbf{a}_j + \mathbf{a}_i}{2}, \mathbf{a}_i - \mathbf{a}_j \right\rangle \tag{2.44}$$

Since the noise vector N is a continuous random vector, the probability that $\mathbf{y}$ will actually fall precisely on the boundary hyperplane H is zero, so it does not matter the decision is $\mathbf{a}_j$ or $\mathbf{a}_i$ in that case.

Probability of decision error: The probability of decision error given that $\mathbf{a}_j$ is transmitted is

$$P(e \mid \mathbf{a}_j) = P(\mathbf{y} \notin \mathcal{R}_j \mid \mathbf{a}_j \text{ sent}) = 1 - \int_{\mathcal{R}_j} p(\mathbf{y} \mid \mathbf{a}_j) d\mathbf{y} = 1 - \int_{\mathcal{R}_j} p_{\mathbf{n}}(\mathbf{y} - \mathbf{a}_j) d\mathbf{y}$$

The average probability of decision error is given by

$$\begin{aligned} \Pr(e) &= \sum_{\mathbf{a} \in \mathcal{A}} P(e \mid \mathbf{a}) P(\mathbf{a}) \\ &= \sum_{\mathbf{a} \in \mathcal{A}} \left[1 - \int_{\mathcal{R}(\mathbf{a})} p(\mathbf{y} \mid \mathbf{a}) d\mathbf{y} \right] P(\mathbf{a}) \\ &= 1 - \sum_{\mathbf{a}} \int_{\mathcal{R}(\mathbf{a})} p(\mathbf{y} \mid \mathbf{a}) P(\mathbf{a}) d\mathbf{y} \\ &= 1 - \sum_{\mathbf{a}} \int_{\mathcal{R}(\mathbf{a})} P(\mathbf{a} \mid \mathbf{y}) p(\mathbf{y}) d\mathbf{y} \end{aligned}$$

where $\mathcal{R}(\mathbf{a})$ denote the decision region associated with the transmitted signal $\mathbf{a}$. Pr(e) is minimized by independently maximizing each term in the sum, which is obtained by

choosing $\mathcal{R}(\mathbf{a})$ as the region where $P(\mathbf{a}|\mathbf{y})$ is maximum over $\mathbf{a}$. Thus, the MAP rule yields the minimum Pr(e).

2.8.4 Pairwise Error Probabilities

The pairwise error probability $P_2(\mathbf{a}_j \to \mathbf{a}_i)$ is defined as the probability that, if $\mathbf{a}_j$ is transmitted, the received point $\mathbf{Y} = \mathbf{a}_j + \mathbf{N}$ is closer to $\mathbf{a}_i$ than to $\mathbf{a}_j$; that is,

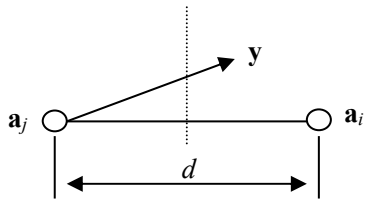
$$P_2(\mathbf{a}_j \to \mathbf{a}_i) \triangleq P\left(\| \mathbf{y} - \mathbf{a}_i \|^2 < \| \mathbf{y} - \mathbf{a}_j \|^2 \,\middle|\, \mathbf{a}_j \text{ sent}\right)$$

It can be shown that an error may occur if and only if the one-dimensional noise component $n_1 = \mathbf{n}_{|\mathbf{a}_i - \mathbf{a}_j}$, the projection of $\mathbf{n}$ onto the difference vector $\mathbf{a}_i - \mathbf{a}_j$, exceeds half the distance $d(\mathbf{a}_i, \mathbf{a}_j) = \| \mathbf{a}_i - \mathbf{a}_j \|$ between $\mathbf{a}_j$ and $\mathbf{a}_i$. Due to fact that the distribution $p_{\mathbf{N}}(\mathbf{n})$ of noise vector $\mathbf{N}$ is spherically symmetric, the p.d.f. of n_1 is given by

$$p_N(n_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_1^2}{2\sigma^2}\right)$$

Therefore, the pairwise error probability $P_2(\mathbf{a}_j \to \mathbf{a}_i)$ can be calculated as

$$P_2(\mathbf{a}_j \to \mathbf{a}_i) = \frac{1}{\sqrt{2\pi}\sigma} \int_{d(\mathbf{a}_i, \mathbf{a}_j)/2}^{\infty} e^{-x^2/2\sigma^2} dx = Q\left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sigma}\right) \tag{2.45}$$



2.8.5 The Union Bound

The union bound on error probability is based on the elementary union bound of probability theory: if A and B are any two events, then $\Pr(A \cup B) \le \Pr(A) + \Pr(B)$. We define the error event E_{ji} as the event that $\mathbf{a}_i \in \mathcal{A}$ is closer than $\mathbf{a}_j$ to $\mathbf{y}$, while $\mathbf{a}_j$ is the transmitted signal vector. Thus the probability of detection error $\Pr(e \mid \mathbf{a}_j)$ with MD detection is given and upper-bounded by

$$\Pr(e \mid \mathbf{a}_j) = P\{\bigcup_{\substack{i=1 \\ i \neq j}}^{M} E_{ji}\}$$

$$\leq \sum_{\substack{i=1 \\ i\neq j}}^{M} P_2(\mathbf{a}_j \rightarrow \mathbf{a}_i) = \sum_{\substack{i=1 \\ i\neq j}}^{M} Q\left(\frac{d(\mathbf{a}_j, \mathbf{a}_i)}{2\sigma}\right)$$

or,

$$\begin{aligned}
\Pr(e \mid \mathbf{a}_j) &= P\left(\mathbf{y} \notin \mathcal{R}_j \mid \mathbf{a}_j \text{ sent}\right) \\
&= P\left(\bigcup_{\substack{i=1 \\ i\neq j}}^{M} \text{Event}(\mathbf{y} \in R_i \mid \mathbf{a}_j)\right) \\
&= \sum_{\substack{i=1 \\ i\neq j}}^{M} P\left(\mathbf{y} \in \mathcal{R}_i \mid \mathbf{a}_j\right) \\
&\leq \sum_{\substack{i=1 \\ i\neq j}}^{M} P_2(\mathbf{a}_j \rightarrow \mathbf{a}_i) = \sum_{\substack{i=1 \\ i\neq j}}^{M} Q\left(\frac{d(\mathbf{a}_j, \mathbf{a}_i)}{2\sigma}\right)
\end{aligned}$$

where the inequality follows from the fact that each $\mathcal{R}_i$ cannot be larger with M-ary signals than it would be if there were just two signals $\mathbf{a}_i$ and $\mathbf{a}_j$ in the set.

Let D denote the set of distances between elements of $\mathcal{A}$; then we have

$$\Pr(e \mid \mathbf{a}_j) \leq \sum_{d \in D} K_d(\mathbf{a}_j) Q(d / 2\sigma) \tag{2.46}$$

where $K_d(\mathbf{a}_j)$ is the number of signals $\mathbf{a}_i \neq \mathbf{a}_j \in \mathcal{A}$ at distance d from $\mathbf{a}_j$.

From (2.4), the minimum distance between signal points is $d_{\min}(\mathcal{A}) = \min_{i \neq j} \| \mathbf{a}_i - \mathbf{a}_j \|$. The *union bound estimate* (UBE) of $\Pr(e \mid \mathbf{a}_j)$ is based on the idea that the nearest neighbors to $\mathbf{a}_j$ at distance $d_{\min}(\mathcal{A})$ will dominate the sum in (2.46). If there are $K_{\min}(\mathbf{a}_j)$ neighbors at distance $d_{\min}(\mathcal{A})$ from $\mathbf{a}_j$, then

$$\Pr(e \mid \mathbf{a}_j) \approx K_{\min}(\mathbf{a}_j) Q\left(\frac{d_{\min}(\mathcal{A})}{2\sigma}\right) \tag{2.47}$$

Notice that this estimate is valid only if the next nearest neighbors are at a significantly greater distance and there are not too many of them.

If there is at least one neighbors $\mathbf{a}_i$ at distance $d_{\min}(\mathcal{A})$ from $\mathbf{a}_j$, then we have the *pairewise lower bound*

$$\Pr(e \mid \mathbf{a}_j) \geq Q\left(\frac{d_{\min}(\mathcal{A})}{2\sigma}\right)$$

Similarly, the total error probability is given by

$$\begin{aligned}\Pr(e) &= \sum_{\mathbf{a}_j \in \mathcal{A}} \Pr(e \mid \mathbf{a}_j) P(\mathbf{a}_j) \\ &\le \sum_{\mathbf{a}_j \in \mathcal{A}} P(\mathbf{a}_j) \sum_{d \in D} K_d(\mathbf{a}_j) Q(d/2\sigma) \\ &= \sum_{d \in D} \sum_{\mathbf{a}_j \in \mathcal{A}} P(\mathbf{a}_j) K_d(\mathbf{a}_j) Q(d/2\sigma) = \sum_{d \in D} K_d(\mathcal{A}) Q(d/2\sigma) \end{aligned} \tag{2.48}$$

where $K_d(\mathcal{A})$ is the average number of signals $\mathbf{a}_i \neq \mathbf{a}_j \in \mathcal{A}$ at distance d from $\mathbf{a}_j$, or simply the average number of neighbors for the constellation $\mathcal{A}$.

If $K_{\min}(\mathcal{A}) = \overline{K_{\min}(\mathbf{a}_j)}$ is the average number of nearest neighbors at distance $d_{\min}(\mathcal{A})$, then the UBE of $\Pr(e)$ is

$$\Pr(e) \approx K_{\min}(\mathcal{A}) Q\left(\frac{d_{\min}(\mathcal{A})}{2\sigma}\right) \tag{2.49}$$

which is also called the *nearest neighbor union bound* (NNUB).

2.8.6 Coding Gain in the Power-Limited Regime

In this regime we take 2-PAM as a baseline constellation. The *effective coding gain* $\gamma_{\text{eff}}(\mathcal{A})$ of a signal set $\mathcal{A}$ at a given target error probability $P_b(e)$ will be defined as the difference in dB between the E_b/N_0 required to achieve the target $P_b(e)$ with $\mathcal{A}$ and the E_b/N_0 required to achieve the target $P_b(e)$ with (uncoded) 2-PAM. In this definition, the effective coding gain includes any gain that result from using a lower spectral efficiency $\rho < 2$.

For an N-dimensional constellation $\mathcal{A}$ that supports $b=\log_2|\mathcal{A}|$ bits per symbol, the UBE allows us to estimate the effective coding gain as follows. We have *error probability per bit* (not in general the same as the bit error probability!)

$$P_b(e) = \frac{\Pr(e)}{b} \approx \frac{K_{\min}(\mathcal{A})}{b} Q\left(\sqrt{\frac{d^2_{\min}(\mathcal{A})}{2N_0}}\right)$$

We define the *nominal (or asymptotic) coding gain* $\gamma_c(\mathcal{A})$ as

$$\gamma_c(\mathcal{A}) = \frac{d^2_{\min}(\mathcal{A})}{4E_b}$$

This definition is normalized so that for 2-PAM, $\gamma_c(\mathcal{A}) = 1$. The UBE then becomes

$$P_b(e) \approx K_b(\mathcal{A}) Q\left(\sqrt{2\gamma_c(\mathcal{A}) E_b / N_0}\right)$$

where $K_b(\mathcal{A}) = K_{\min}(\mathcal{A})/b$ is the average nearest neighbors per transmitted bit. Thus if

$K_b(\mathcal{A}) = 1$, then $\gamma_{\text{eff}}(\mathcal{A}) = \gamma_c(\mathcal{A})$ for all $P_b(e)$. However, if $K_b(\mathcal{A}) > 1$, then $\gamma_{\text{eff}}(\mathcal{A}) < \gamma_c(\mathcal{A})$.

2.8.7 Coding Gain in the Bandwidth-Limited Regime

In the bandwidth-limited regime, we take the (M×M)-QAM signal set as a baseline and normalize all quantities per two dimensions. Let $P_s(e)$ be error probability per two dimensions. The *effective coding gain* $\gamma_{\text{eff}}(\mathcal{A})$ is defined as the difference in dB between the SNR_{norm} required to achieve the target $P_s(e)$ with $\mathcal{A}$ and the SNR_{norm} required to achieve the target $P_s(e)$ with M-PAM or (M×M)-QAM.

The effective coding gain is estimated by the UBE, as follows. We have

$$P_s(e) = \frac{2\Pr(e)}{N} \approx \frac{2K_{\min}(\mathcal{A})}{N} Q\left(\sqrt{\frac{d^2_{\min}(\mathcal{A})}{2N_0}}\right)$$

In the bandwidth-limited regime, we define the *nominal (or asymptotic) coding gain* $\gamma_c(\mathcal{A})$ as

$$\gamma_c(\mathcal{A}) = \frac{(2^\rho - 1)d^2_{\min}(\mathcal{A})}{6E_s}$$

This definition is normalized so that for M-PAM or (M×M)-QAM, $\gamma_c(\mathcal{A}) = 1$. Then the UBE becomes

$$P_s(e) \approx K_s(\mathcal{A}) Q\left(\sqrt{3\gamma_c(\mathcal{A}) SNR_{norm}}\right)$$

where $K_s(\mathcal{A}) = 2K_{\min}(\mathcal{A}) / N$.

2.9 Design of Signal Constellations

From the above discussion, the probability of detection error depends primarily on the ratio $d_{\min}(\mathcal{A})/2\sigma$.

- Therefore, the objective of constellation design is to maximize $d_{\min}(\mathcal{A})$ while not exceeding the power constraint, or equivalently to minimize E_s or E_b for a given $d_{\min}(\mathcal{A})$; or equivalently to maximize a figure of merit such as

$$\zeta = \frac{d^2_{\min}(\mathcal{A})}{4E_s} \quad \text{or} \quad \zeta = \frac{d^2_{\min}(\mathcal{A})}{4E_b}$$

- Secondarily, one would also like to minimize the "error coefficient" $K_{\min}(\mathcal{A})$.

The geometric problem of optimally arraying points in a vector space with minimum E_N while maintaining a minimum distance between points is the well-stuided

sphere-packing problem.

A signal constellation usually has the following properties:

- The performance of a constellation is invariant under rotation and translation.
- The power will be minimized if it has zero mean.

2.9.1 Invariance of probability of detection error to Translation

Consider an arbitrary signal set $\mathcal{A} = \{\mathbf{a}_j, 1 \le j \le M\}$. Let $\mathcal{A}$' be $\mathcal{A}$ translated by a constant vector amount **m**, i.e.,

$$\mathcal{A}' \equiv \{\mathbf{a}_{j,translte}, 1 \le j \le m\} = \mathcal{A} - \mathbf{m} = \{\mathbf{a}_j - \mathbf{m}, 1 \le j \le m\}$$

The received vector **y** is correspondingly translated by the same vector amount, as shown by

$$\mathbf{y}_{translate} = \mathbf{y} - \mathbf{m}$$

We therefore immediately deduce that

$$\| \mathbf{y}_{translate} - \mathbf{a}_{j,translate} \| = \| \mathbf{y} - \mathbf{a}_j \|, \forall j$$

Thus we have the following remark:

Remark 2.9.1: If a signal constellation is translated by a constant vector, then the error probability of an optimum detector over an AWGN channel is completely unchanged.

- Optimality of zero-mean constellations

Let $E_{\mathcal{A}}$ and $E_{\mathcal{A}'}$ denote the average energies of $\mathcal{A}$ and $\mathcal{A}$'. Then

$$E_{\mathcal{A}'} = \sum_{j=1}^{M} \| \mathbf{a}_j - \mathbf{m} \|^2 \, P(j)$$

where $P(j)$ is the probability that $\mathbf{a}_j$ is transmitted. The above equation can be written as

$$E_{\mathcal{A}'} = \sum_{j=1}^{M} \| \mathbf{a}_j \|^2 \, P(j) - 2\sum_{j=1}^{M} \mathbf{m}^T \mathbf{a}_j P(j) + \| \mathbf{m} \|^2 \sum_{j=1}^{M} P(j)$$

$$= E_{\mathcal{A}} - 2\mathbf{m}^T \mathsf{E}[\mathbf{A}] + \| \mathbf{m} \|^2$$

where $\mathsf{E}[\mathbf{A}] = \sum_{j=1}^{M} \mathbf{a}_j P(j)$. Let $\frac{dE_{\mathcal{A}'}}{d\mathbf{m}} = 0$. We find that the minimum translate is

$$\mathbf{m}_{\min} = \mathsf{E}[\mathbf{A}]$$

The minimum average energy of the constellation $\mathcal{A}$' is given by

$$E_{\mathcal{A}',\min} = E_{\mathcal{A}} - \| \mathbf{m}_{\min} \|^2$$

Thus we have the following remark:

Remark 2.9.2: The signal constellation with zero mean has minimum average energy.

2.9.2 Invariance of probability of detection error to Rotation

Again consider an arbitrary signal set $\mathcal{A} = \{\mathbf{a}_j, 1 \le j \le M\}$. The effect of a rotation applied to all signal points in $\mathcal{A}$ is equivalent to multiplying the N-vector $\mathbf{a}_j$ by an $N \times N$ orthonormal matrix $\mathbf{Q}$ for all j. The matrix $\mathbf{Q}$ satisfies the condition $\mathbf{QQ}^T = \mathbf{I}$. Thus the rotated signal is

$$\mathbf{a}_{j,rotate} = \mathbf{Q}\mathbf{a}_j, \qquad \forall j$$

The received vector for the rotated signal constellations can be expressed as

$$\mathbf{y}_{rotate} = \mathbf{Q}\mathbf{a}_j + \mathbf{n} \qquad \forall j$$

Thus we have

$$\| \mathbf{y}_{rotate} - \mathbf{a}_{j,rotate} \|^2 = (\mathbf{y}_{rotate} - \mathbf{a}_{j,rotate})^T (\mathbf{y}_{rotate} - \mathbf{a}_{j,rotate})$$

$$= (\mathbf{y} - \mathbf{a}_j)^T \mathbf{Q}^T \mathbf{Q} (\mathbf{y} - \mathbf{a}_j)$$

$$= \| \mathbf{y} - \mathbf{a}_j \|^2$$

Therefore, *the error probability of an optimum detector remains unchanged on an AWGN channel.*

2.9.3 Examples of Signal Constellations

- The QAM cross constellations

The cross constellation allows for odd numbers of bits per symbol in QAM data transmission. Its can be constructed as follows.

Augments a square-QAM constellation from b-1 bits per symbol by adding 2^{b-1} data symbols that extend the sides of the QAM square. See Fig. 2.23. As an example, a 32-QAM cross constellation is shown in Fig. 2.24.

It can be shown that a M-ary cross constellation has the average energy

$$E = \frac{d^2}{6}\left(\frac{31}{32}M - 1\right).$$

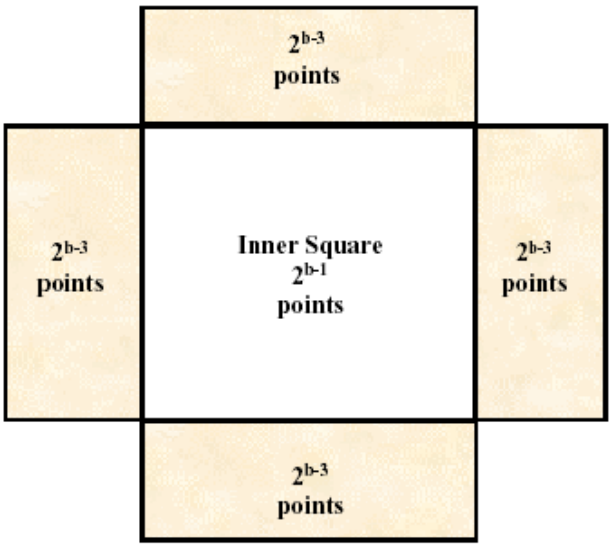


Figure 2.23 Cross constellations

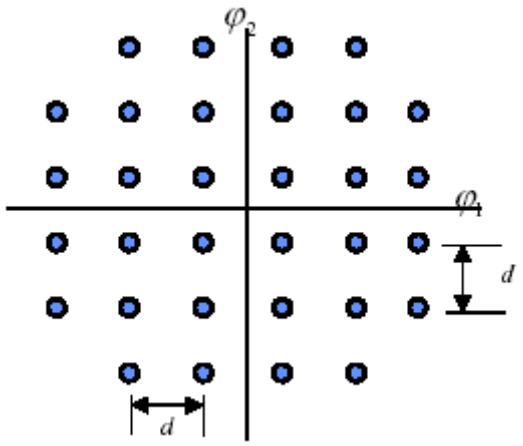


Figure 2.24 A 32-cross constellation

■ Optimal Hexagonal constellations in two dimensions

A performance improvement can be achieved with hexagonal constellations. The term hexagonal refers to the shape of the decision regions. See Fig. 2.25.

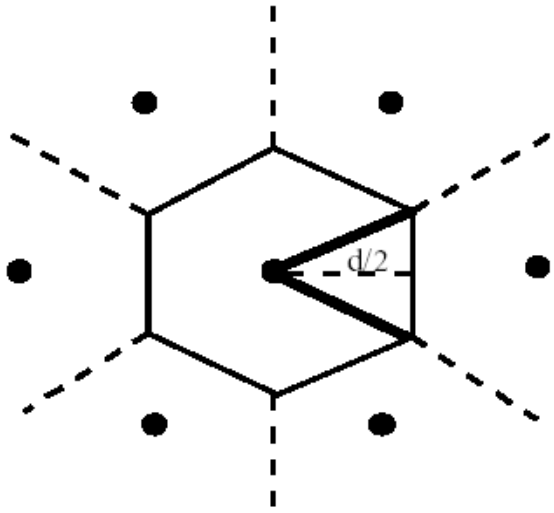


Figure 2.25 Hexagonal constellations.

■ Cubic constellations

Cubic constellations are commonly used on simple data communication channels. The construction of a cubic constellation directly maps a sequence of $N = b$ bits into the components of the basis vectors in a corresponding N-dimensional signal constellation. Following is an example. The bit stream . . . 010010 . . . may be grouped as the sequence of two-dimensional vectors . . . (01)(00)(10) . . .

The n-fold *Cartesian product* of a 2-PAM signal set $\mathcal{A}=\{\pm d/2\}$ is the n-dimension signal set

$$\mathcal{A}' = \mathcal{A}^n = \{(a_1, a_2, ..., a_n) \mid a_k \in \mathcal{A}, 1 \le k \le n\}$$

which corresponds to independent transmission of n bits using 2-PAM. Geometrically, $\mathcal{A}$' is the vertex set of an n-cube of side d. It has dimension $N' = n$, size $M' = |\mathcal{A}|^n = 2^n$, nominal spectral efficiency $\rho = 2$ (b/D), average energy $E_{\mathcal{A}'} = nd^2/4$, average energy per bit $E_b = d^2 / 4$, minimum squared distance $d^2_{\min}(\mathcal{A}') = d^2$, the average number of nearest neighbors $K'_{\min}(\mathcal{A}') = n$. Its probability of error is then

$$\Pr(e)' = nQ(\sqrt{2E_b / N_0})$$

The error probability per dimension (or per bit) is $P_b(e) = \Pr(e)' / n$.

A general *Cartesian-product constellation* $\mathcal{A}^K$ can be expressed as

$$\mathcal{A}^K = \{(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K) \mid \mathbf{x}_k \in \mathcal{A}\}$$

which is the set of all sequences of K points from an elementary constellation $\mathcal{A}$.

In general, the concatenation of two successively transmitted N-dimensional signals taken from the same N-dimensional signal constellation as a single $2N$-dimensional signal causes the resulting $2N$-dimensional constellation, formed as a Cartesian product of the constituent N-dimensional constellations, to have the same average energy per dimension as the N-dimensional constellation. Thus, *simple concatenation of a signal set with itself does not improve the design*. However, careful packing of signals in increasingly larger dimensional signals sets can lead to a reduction in the energy per dimension required to transmit a given set of messages.

■ Constellation design based on mutual information

See [S. Y. Le Goff, "Signal constellations for bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol.49, no.1, pp.307-313, Jan. 2003.]

Appendix

The MATLAB program used to compute the channel capacity for *M*-PAM constellations in Section 2.5 is listed below. It is based on Monte Carlo method.

```
% ---------------------------------------- begin ----------------------------------------------
snr_indB = -10:1:30;
snr = 10 .^ (0.1 * snr_indB);
EbN0 = [];
C_mpam = [];
logpy = [];
alphabet = [];
trials = 50000;
b = 3;
M = 2 ^ b;
for k = 1 : 1 : M;
   alphabet(k) = 2*k - 1 - M;
end
average = sum(alphbet .^ 2) / M;    % average energy Es

for i = 1 : 1 : length(snr_indB);
   sigma2 = average / snr(i);
   for j = 1 : 1 : trials;
      x(j) = alphabet(floor(rand * M) + 1);
   end
   y = x + sqrt(sigma2) * randn(1, trials);
   for j = 1 : 1 : trials;
      pyx = 0;                      % compute p(y) from p(y|x)
      for k = 1 : 1 : M;
         pyx = pyx + exp(-(y(j) - alphabet(k)).^2 / (2*sigma2));
      end
      logpy(j) = log2(1/M / sqrt(2*pi*sigma2) * pyx(1));
   end
   % I(X; Y) = H(Y) - H(Y|X) = H(Y) - H(N)
   C_mpam(i) = -mean(logpy) - 0.5*log2(2*pi * exp(1) * sigma2);
end

EbN0 = 10 * log10(snr ./ C_mpam / 2);
plot(snr_indB, C_mpam, 'k-');
%plot(EbN0, C_mpam, 'k-');
hold on;
```

```
C = [];                                 % Shannon capacity
for i = 1 : 1 : length(snr_indB);
   C(i) = 0.5 * log2(1 + snr(i));
   EbN0(i) = 10 * log10(snr(i) / C(i) / 2);
end
plot(snr_indB, C, 'b-');
%plot(EbN0, C, `b-`);

legend('M-PAM bound', 'Shannon bound');
xlabel('SNR (dB)');
%xlabel('Eb/N0 (dB)');
ylabel('Capacity (bits/symbol)');
grid on;
% ---------------------------------------- end ----------------------------------------------------
```

Suggested Reading

[1] R. G. Gallager, *Principles of Digital Communication*. Cambridge University Press, 2008. (影印版：人民邮电出版社，2010 年)

[2] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inform. Theory*, vol.44, no.6, pp. 2384-2415, Oct. 1998.

[3] G. D. Forney, *Principles of Digital Communication - Part II*, Lecture notes. MIT, 2005.