Bit Error Rate Performance for Digital Communications

Version 1.3

Prof. Rui Dinis

This section considers the performance evaluation of digital communication systems employing signals in the PAM format where each symbol can take one of *M* values, which are selected from the data bits to be transmitted according to a given mapping rule. We focus on the receiver design and the computation of the corresponding BER (Bit Error Rate) performance for base-band signals. The analysis can be extended for bandpass signals by taking advantage of the base-band equivalent of a given digital modulation. This extension will be done in the future.

Let us consider the PAM signal

$$x(t) = \sum_k a_k r(t - kT).$$

where $r(t)$ is the pulse shape and a_k is the amplitude related to the value of the symbol *k*. For an M-PAM signal, each a_k can take one of *M* possible values and corresponds to the transmission of $\log_2(M)$ bits. For instance, for polar binary signals, a_k can take the values –A or A for bits 0 or 1, respectively, and for polar quaternary signals a_k can take the values -3A, -A, A and 3A for the dibits 01, 00, 10, 11, respectively.

The pulse shape associated to the *k*th symbol is

$$a_k r(t - kT)$$

and the corresponding energy of a modulated pulse shape is

$$E_k = \int_{-\infty}^{+\infty} |a_k r(t-kT)|^2 \, dt = |a_k|^2 \int_{-\infty}^{+\infty} |r(t)|^2 \, dt = |a_k|^2 E_r$$

where

$$E_r = \int_{-\infty}^{+\infty} |r(t)|^2 \, dt$$

denotes the energy of the basic pulse shape $r(t)$. This means that the average symbol energy is

$$E_s = \overline{|a_k|^2}$$

The average bit energy is

$$E_b = \frac{E_s}{\log_2(M)} = \frac{\overline{|a_k|^2} E_r}{\log_2(M)}$$

It can be shown that

$$E_b = P_x T_b = \frac{P_x}{R_b}$$

where P_x denotes the average power of the signal $x(t)$, R_b denotes the bit rate and

$$T_b = \frac{T}{\log_2(M)} = \frac{1}{R_b}$$

denotes the bit duration, It should be noted that the bit duration, T_b, is not the pulse duration, T, not even in the binary case (for RZ signals the pulse duration is smaller than T and for Nyquist, raised cosine pulses the pulse duration is infinity, not T_b).

If we assume an ideal channel, the received signal will be

$$z(t) = x(t) + w(t) = \sum_k a_k r(t - nT) + w(t)\,,$$

where w(t) denotes the noise term, which is assumed to be AWGN (Additive White Gaussian Noise). An AWGN noise has the following PSD (Power Spectral Density)

$$S_w(f) = \frac{N_0}{2}$$

whose average value is 0 (there is no spectral line at frequency 0) and its power is

$$P_w = \int_{-\infty}^{+\infty} S_w(f)df = +\infty\,.$$

Clearly, the transmission of a finite-power signal $x(t)$ with infinite power noise is not possible, since the corresponding SNR (Signal to Noise Ratio) is

$$SNR_{Channel} = \frac{P_x}{P_w} = 0\,.$$

However, most of the power of the useful signal is concentrated in a limited band, while the power of the AWGN noise is evenly spread over all frequencies. This means that we can reduce the power of the noise by filtering it. If this filtering is done appropriately, the power of the signal term is not severely affected.

Let us consider a detection filter with impulse response $h(t)$ and frequency response $H(f)$. The signal at the output of the detection filter will be

$$q(t) = z(t) * h(t) = x(t) * h(t) + w(t) * h(t) = y(t) + n(t)$$

Clearly, the useful signal at the filter's output still has the PAM format, and will be given by

$$y(t) = \sum_k a_k p(t - nT)$$

where the new pulse shape will be given by

$$p(t) = r(t) * h(t)$$

Regarding the noise term at the filter's output, it has the PSD

$$S_n(f) = S_w(f)\,|\,H(f)\,|^2 = \frac{N_0}{2}\,|\,H(f)\,|^2$$

and the corresponding power is

$$P_n = \int_{-\infty}^{+\infty} S_n(f)df = \frac{N_0}{2}\int_{-\infty}^{+\infty} \left|H(f)\right|^2 df = \frac{N_0}{2}\int_{-\infty}^{+\infty} \left|h(t)\right|^2 dt \, .$$

If the filter has a very narrow band then P_n can be made arbitrarily small, but the signal term will also vanish. On the other hand, if the filter's band is large we can maintain the useful signal unchanged, but we will end up with a scenario similar to the original, unfiltered case. Clearly, there will be an optimum filter somewhere in-between. It can be shown that the optimum filter that maximizes the SNR at its output is the so-called matched filter, which has impulse response

$$h(t) = K\, r^*(t_0 - t)\,,$$

where K is a scaling factor that does not affect the SNR (it affects equally the signal and noise terms) and t_0 is a delay, that is only required to make the filter causal. If the pulse shape $r(t)$ is real, we can assume without loss of generally that

$$h(t) = r(-t) \leftrightarrow H(f) = R^*(f)\,,$$

i.e., the matched filter is simply a reflected version of the pulse shape $r(t)$. In this case, the pulse shape at the detection filter output becomes

$$p(t) = r(t) * r(-t) \leftrightarrow P(f) = \left|R(f)\right|^2 = \left|H(f)\right|^2 \, .$$

Regarding the noise, we have

$$S_n(f) = S_w(f)\,|\,H(f)\,|^2 = \frac{N_0}{2}P(f)\,,$$

the autocorrelation of the noise samples is

$$R_n(\tau) = \frac{N_0}{2}\, p(\tau) \qquad (1)$$

And its power is

$$P_n = \int_{-\infty}^{+\infty} S_n(f)df = \frac{N_0}{2}\int_{-\infty}^{+\infty} P(f)df = \frac{N_0}{2}\, p(0)\, .$$

To obtain the estimates of the transmitted symbols (and bits), the signal at the output of the detection filter will be sampled at the sampling instants

$$t_k = t_0 + kT = kT$$

(assuming that t_0=0), leading to the samples q_k

$$q_k = q(kT) = y(kT) + n(kT) = y_k + n_k$$

The sampled signal is

$$y_k = y(kT) = \sum_m a_m p(kT - mT) = \sum_m a_{k-m} p(mT) = \underbrace{a_k p(0)}_{\text{Useful term}} + \underbrace{\sum_{m\neq 0} a_{k-m} p(mT)}_{\text{ISI term}}$$

For the case where we do not have ISI (i.e., when we employ Nyquist pulses like the raised cosine pulses), we have

$$p(kT) = 0, k \neq 0,$$

which means that

$$y_k = a_k p(0)$$

The sampled noise will still be Gaussian (a linear filter does not change the Gaussian nature of the input signal), with mean 0 (we still do not have a spectral line at frequency 0 in $S_n(f)$) and with variance

$$\sigma^2 = E\left[\left|n_k\right|^2\right] = E\left[\left|n(t)\right|^2\right] = P_n = \int_{-\infty}^{+\infty} S_n(f)df = \frac{N_0}{2}\int_{-\infty}^{+\infty} P(f)df = \frac{N_0}{2} p(0).$$

Moreover, from (1), the correlation between noise samples n_k and $n_{k'}$ is

$$R_n((k-k')T) = \frac{N_0}{2} p((k-k')T).$$

For the case where we do not have ISI this leads to

$$R_n((k-k')T) = \begin{cases} N_0 p(0)/2, & k = k' \\ 0, & k \neq k' \end{cases}$$

which means different noise samples are uncorrelated. Therefore, since the data samples y_k are also uncorrelated, we can perform the optimum detection of the kth symbol using only the sample q_k.

From the Gaussian nature of the noise, we can define the noise's probability density function

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

and the probability that the noise is above a given value is

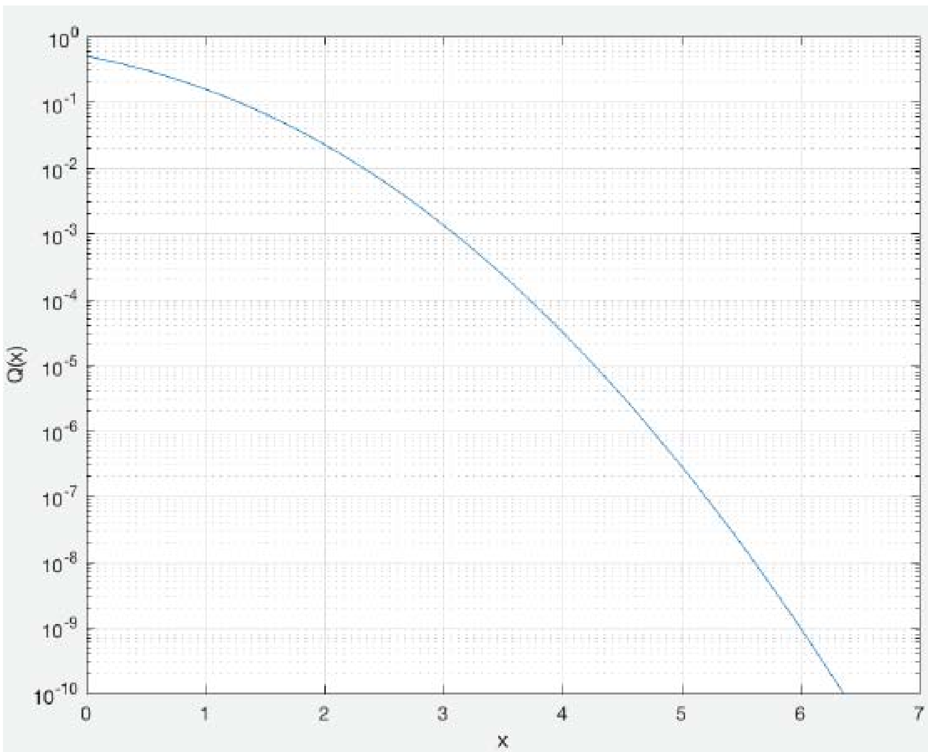
$$\text{Prob.}(n_k > D) = \int_D^\infty p(\alpha)d\alpha = \int_D^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\alpha^2}{2\sigma^2}\right)d\alpha =$$

$$\overset{(a)}{=} \int_{D/\sigma}^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\alpha^2}{2}\right)d\alpha \overset{(b)}{=} Q\left(\frac{D}{\sigma}\right) = \text{Prob.}\left(\frac{n_k}{\sigma} > \frac{D}{\sigma}\right)$$

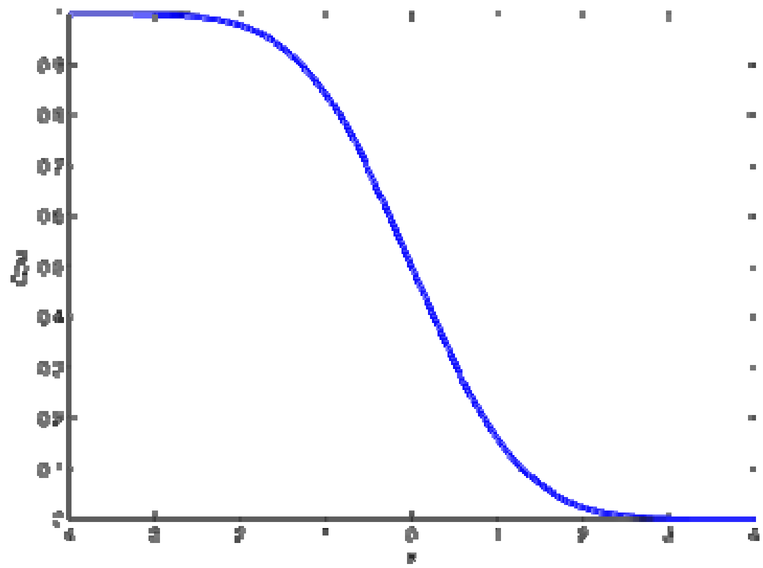
where (a) results from a change of variables in the integral and (b) takes advantage of the definition of the Q() function

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\alpha^2}{2}\right) d\alpha \ .$$

In statistics, the *Q*-function is the tail distribution function of the standard normal distribution (i.e., a Gaussian distribution with zero mean and variance one). In other words, *Q*(*x*) is the probability that a normal (Gaussian) random variable will obtain a value larger than *x* standard deviations. Equivalently, *Q*(*x*) is the probability that a standard normal random variable takes a value larger than *x*. This function does not have a closed form and usually is obtained from tables or graphs like the one in the figure.



If we use a linear scale (instead of the semi-logarithmic one as in the figure above) the graph of the Q(x) function would look like the one represented in the following figure (not so relevant to obtain the values we want).



Example: Binary Polar NRZ

In this case,

$$a_k = \begin{cases} -A, & \text{bit } 0 \\ A, & \text{bit } 1 \end{cases}$$

and

$$r(t) = \text{rect}(t/T)\ .$$

We have

$$\begin{aligned} p(0) &= T \\ E_b = E_s &= A^2 T \\ P_x &= A^2 \\ \sigma^2 &= \frac{N_0}{2} T \end{aligned}$$

The useful term of the signal at the sampling instants takes the values

$$y_k = a_k p(0) = \begin{cases} -AT, & \text{bit } 0 \\ AT, & \text{bit } 1 \end{cases}$$

and we can perform the decision:

$$\hat{a}_k = \begin{cases} -A, & \text{if } q_k < 0 \\ A, & \text{if } q_k > 0 \end{cases}$$

We will have an error in the following cases:

$$\begin{aligned} \hat{a}_k = A \text{ and } a_k = -A &\Leftrightarrow n_k > AT \\ \hat{a}_k = -A \text{ and } a_k = A &\Leftrightarrow n_k < -AT \end{aligned}$$

Both probabilities are given by

$$\text{Prob}(n_k > AT) = \text{Prob}(n_k < -AT) = Q\left(\frac{AT}{\sigma}\right) = Q\left(\sqrt{\frac{A^2T^2}{\sigma^2}}\right) = Q\left(\sqrt{\frac{2A^2T}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Which means that the average BER (Bit Error Rate) will be

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

Example: Binary Polar RZ (Duty Cycle 50%)

In this case,

$$a_k = \begin{cases} -A, & \text{bit 0} \\ A, & \text{bit 1} \end{cases}$$

and

$$r(t) = \text{rect}\left(\frac{t}{T/2}\right).$$

We have

$$p(0) = T/2$$
$$E_b = E_s = A^2T/2$$
$$P_x = A^2/2$$
$$\sigma^2 = \frac{N_0}{4}T$$

The useful term of the signal at the sampling instants takes the values

$$y_k = a_k\,p(0) = \begin{cases} -AT/2, & \text{bit 0} \\ AT/2, & \text{bit 1} \end{cases}$$

and we can perform the decision:

$$\hat{a}_k = \begin{cases} -A, & \text{if } q_k < 0 \\ A, & \text{if } q_k > 0 \end{cases}$$

We will have an error in the following cases:

$$\hat{a}_k = A \text{ and } a_k = -A \Leftrightarrow n_k > AT/2$$
$$\hat{a}_k = -A \text{ and } a_k = A \Leftrightarrow n_k < -AT/2$$

Both probabilities are given by

$$\text{Prob}(n_k > AT/2) = \text{Prob}(n_k < -AT/2) = Q\left(\frac{AT/2}{\sigma}\right) =$$
$$= Q\left(\sqrt{\frac{A^2T^2}{4\sigma^2}}\right) = Q\left(\sqrt{\frac{A^2T}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

This means that the average BER will be

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

which is the same of the binary polar NRZ case.

Example: Binary Polar Without ISI (General Case)

In this case,

$$a_k = \begin{cases} -A, & \text{bit 0} \\ A, & \text{bit 1} \end{cases}$$

and the adopted pulse shape is such that we have a Nyquist pulse at the matched filter output. This means that

$$p(t) = r(t) * h(t) = r(t) * r(-t) \leftrightarrow P(f) = R(f)H(f) = |R(f)|^2 = |H(f)|^2$$
$$p(kT) = 0, \quad k \neq 0$$
$$S_x(f) = \frac{A^2}{T}|R(f)|^2 = \frac{A^2}{T}P(f) \to P_x = \int_{-\infty}^{+\infty} S_x(f)df = \frac{A^2}{T}\int_{-\infty}^{+\infty} P(f)df = \frac{A^2}{T}p(0).$$
$$E_b = E_s = P_x T = A^2 p(0)$$
$$\sigma^2 = \frac{N_0}{2}\int_{-\infty}^{+\infty} |H(f)|^2 df = \frac{N_0}{2}\int_{-\infty}^{+\infty} P(f)df = \frac{N_0}{2}p(0)$$

The useful term of the signal at the sampling instants takes the values

$$y_k = a_k p(0) = \begin{cases} -Ap(0), & \text{bit 0} \\ Ap(0), & \text{bit 1} \end{cases}$$

and we can perform the decision:

$$\hat{a}_k = \begin{cases} -A, & \text{if } q_k < 0 \\ A, & \text{if } q_k > 0 \end{cases}$$

We will have an error in the following cases:

$$\hat{a}_k = A \text{ and } a_k = -A \Leftrightarrow n_k > Ap(0)$$
$$\hat{a}_k = -A \text{ and } a_k = A \Leftrightarrow n_k < -Ap(0)$$

Both probabilities are given by

$$\text{Prob}(n_k > Ap(0)) = \text{Prob}(n_k < -Ap(0)) = Q\left(\frac{Ap(0)}{\sigma}\right) =$$
$$= Q\left(\sqrt{\frac{A^2 p^2(0)}{\sigma^2}}\right) = Q\left(\sqrt{\frac{2A^2 p(0)}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

This means that the average BER will be

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

which is independent of the adopted pulse shape, provided there is no ISI at the output of the matched filter.

Example: Binary Unipolar NRZ

In this case,

$$a_k = \begin{cases} 0, & \text{bit } 0 \\ A, & \text{bit } 1 \end{cases}$$

and

$$r(t) = \text{rect}(t/T)\,.$$

We have

$$\begin{aligned} &p(0) = T \\ &E_b = E_s = A^2 T/2 \\ &P_x = A^2/2 \\ &\sigma^2 = \frac{N_0}{2} T \end{aligned}$$

The useful term of the signal at the sampling instants takes the values

$$y_k = a_k\, p(0) = \begin{cases} 0, & \text{bit } 0 \\ AT, & \text{bit } 1 \end{cases}$$

and we can perform the decision:

$$\hat{a}_k = \begin{cases} 0, & \text{if } q_k < AT/2 \\ A, & \text{if } q_k > AT/2 \end{cases}$$

We will have an error in the following cases:

$$\begin{aligned} &\hat{a}_k = A \text{ and } a_k = 0 \Leftrightarrow n_k > AT/2 \\ &\hat{a}_k = 0 \text{ and } a_k = A \Leftrightarrow n_k < -AT/2 \end{aligned}$$

Both probabilities are given by

$$\text{Prob}(n_k > AT/2) = \text{Prob}(n_k < -AT/2) = Q\left(\frac{AT/2}{\sigma}\right) =$$

$$= Q\left(\sqrt{\frac{A^2T^2}{4\sigma^2}}\right) = Q\left(\sqrt{\frac{A^2T/2}{N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Which means that the average BER (Bit Error Rate) will be

$$P_b = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

When compared with the binary polar case, we have a degradation of a factor of 2 (or 3dB). This degradation is due to the fact that the average value of the transmitted signal $x(t)$ (i.e., its DC component) is $A/2$ and has power $A^2/4$, which corresponds to half of the power of $x(t)$. This DC component is not used for transmitting any information and, therefore, is "wasting" half of the transmitted power, hence the 3dB degradation when compared with the polar binary case (where there is no DC component and no power is wasted).

Example: Quaternary Polar NRZ

In this case,

$$a_k = \begin{cases} 3A, & \text{dibit } 11 \\ A, & \text{dibit } 10 \\ -A, & \text{dibit } 00 \\ -3A, & \text{dibit } 01 \end{cases}$$

and

$$r(t) = \mathrm{rect}(t/T)\ .$$

We have

$$\begin{aligned} &p(0) = T \\ &E_s = 5A^2 T \\ &E_b = \frac{E_s}{2} = \frac{5}{2} A^2 T \\ &P_x = 5A^2 \\ &\sigma^2 = \frac{N_0}{2} T \end{aligned}$$

The useful term of the signal at the sampling instants takes the values

$$y_k = a_k p(0) = \begin{cases} 3AT, & \text{dibit } 11 \\ AT, & \text{dibit } 10 \\ -AT, & \text{dibit } 00 \\ -3AT, & \text{dibit } 01 \end{cases}$$

and we can perform the decision:

$$\hat{a}_k = \begin{cases} 3A, & \text{if } q_k > 2AT \\ A, & \text{if } 0 < q_k < 2AT \\ -A, & \text{if } 0 > q_k > -2AT \\ -3A, & \text{if } q_k < -2AT \end{cases}$$

The different thresholds are separated by a distance 2AT and the noise needs to be above *AT* or below *–AT* to lead to an error at the symbol level (we transmit a symbol and detect a different symbol). These probabilities are given by

$$p = Q\left(\frac{AT}{\sigma}\right) = Q\left(\sqrt{\frac{A^2T^2}{\sigma^2}}\right) = Q\left(\sqrt{\frac{2A^2T}{N_0}}\right) = Q\left(\sqrt{\frac{4E_b}{5N_0}}\right).$$

The probability of having a symbol error depends on the transmitted symbol:

- Transmission of a_k=3A. We have an error if $n_k < -AT$, which has the probability *p*.
- Transmission of a_k=A. We have n error if $n_k < -AT$ or $n_k > AT$. Both have probability *p*, and the joint probability is 2*p*.
- Transmission of a_k=-A. We have n error if $n_k < -AT$ or $n_k > AT$. Both have probability *p*, and the joint probability is 2*p*.
- Transmission of a_k=-3A. We have an error if $n_k > AT$, which has the probability *p*.

By averaging the different cases (which have all probability ¼), we obtain the average symbol probability

$$P_s = \frac{p+2p+2p+p}{4} = \frac{3}{2}Q\left(\sqrt{\frac{4E_b}{5N_0}}\right)$$

If we employ a Gray mapping between the dibits and the symbols, then a symbol error is very likely to lead to a single bit error (i.e., we move to an adjacent symbol, that differs only in one bit). Therefore, the average BER will be approximately given by

$$P_b \approx \frac{P_s}{2} = \frac{3}{4}Q\left(\sqrt{\frac{4E_b}{5N_0}}\right).$$

If we ignore the factor ¾ (which has a minor impact for low BER values), there is a degradation relatively to the binary polar case of a factor 5/2=2.5 (about 4dB). This is the price that we pay for transmitting with twice the bit rate for the same bandwidth.