the 4–7 dB range. This derivation assumed that there are no coupling losses between the amplifier and the input and output fibers. Having an input coupling loss degrades the noise figure of the amplifier (see Problem 4.5).

4.4.6 Bit Error Rates

Earlier, we calculated the bit error rate of an ideal direct detection receiver. Next, we will calculate the bit error rate of the practical receivers already considered, which must deal with a variety of different noise impairments.

The receiver makes decisions as to which bit (0 or 1) was transmitted in each bit interval by sampling the photocurrent. Because of the presence of noise currents, the receiver could make a wrong decision resulting in an erroneous bit. In order to compute this bit error rate, we must understand the process by which the receiver makes a decision regarding the transmitted bit.

First, consider a *pin* receiver without an optical preamplifier. For a transmitted 1 bit, let the received optical power $P = P_1$, and let the mean photocurrent $\bar{I} = I_1$. Then $I_1 = \mathcal{R}P_1$, and the variance of the photocurrent is

$$\sigma_1^2 = 2eI_1B_e + 4k_BTB_e/R_L.$$

If P_0 and I_0 are the corresponding quantities for a 0 bit, $I_0 = \mathcal{R}P_0$, and the variance of the photocurrent is

$$\sigma_0^2 = 2eI_0B_e + 4k_BTB_e/R_L.$$

For ideal OOK, P_0 and I_0 are zero, but we will see later (Section 5.3) that this is not always the case in practice.

Let I_1 and I_0 denote the photocurrent sampled by the receiver during a 1 bit and a 0 bit, respectively, and let σ_1^2 and σ_0^2 represent the corresponding noise variances. The noise signals are assumed to be Gaussian. The actual variances will depend on the type of receiver, as we saw earlier. So the bit decision problem faced by the receiver has the following mathematical formulation. The photocurrent for a 1 bit is a sample of a Gaussian random variable with mean I_1 and variance σ_1 (and similarly for the 0 bit as well). The receiver must look at this sample and decide whether the transmitted bit is a 0 or a 1. The possible probability density functions of the sampled photocurrent are sketched in Figure 4.8. There are many possible *decision rules* that the receiver can use; the receiver's objective is to choose the one that minimizes the bit error rate. This *optimum decision rule* can be shown to be the one that, given the observed photocurrent I, chooses the bit (0 or 1) that was *most likely* to have been transmitted. Furthermore, this optimum decision rule can be implemented as follows. Compare the observed photocurrent to a decision threshold I_{th}. If $I \geq I_{th}$, decide that a 1 bit was transmitted; otherwise, decide that a 0 bit was transmitted.

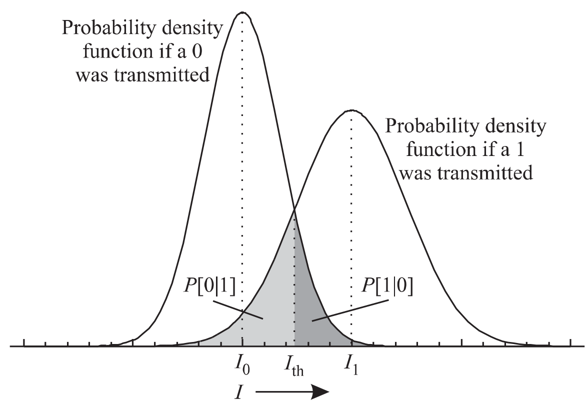

Figure 4.8 Probability density functions for the observed photocurrent.

For the case when 1 and 0 bits are equally likely (which is the only case we consider in this book), the threshold photocurrent is given approximately by

$$I_{\text{th}} = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1}. \tag{4.12}$$

This value is very close but not exactly equal to the optimal value of the threshold. The proof of this result is left as an exercise (Problem 4.7). Geometrically, I_{th} is the value of I for which the two densities sketched in Figure 4.8 cross. The probability of error when a 1 was transmitted is the probability that $I < I_{\text{th}}$ and is denoted by $P[0|1]$. Similarly, $P[1|0]$ is the probability of deciding that a 1 was transmitted when actually a 0 was transmitted and is the probability that $I \geq I_{\text{th}}$. Both probabilities are indicated in Figure 4.8.

Let $Q(x)$ denote the probability that a zero mean, unit variance Gaussian random variable exceeds the value x. Thus

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2}\, dy. \tag{4.13}$$

It now follows that

$$P[0|1] = Q\left(\frac{I_1 - I_{\text{th}}}{\sigma_1}\right)$$

and

$$P[1|0] = Q\left(\frac{I_{\text{th}} - I_0}{\sigma_0}\right).$$

Using (4.12), it can then be shown that the BER (see Problem 4.6) is given by

$$\text{BER} = Q\left(\frac{I_1 - I_0}{\sigma_0 + \sigma_1}\right). \tag{4.14}$$

The Q function can be numerically evaluated. Let $\gamma = Q^{-1}(\text{BER})$. For a BER rate of 10^{-12}, we need $\gamma \approx 7$. For a BER rate of 10^{-9}, $\gamma \approx 6$.

Note that it is particularly important to have a variable threshold setting in receivers if they must operate in systems with signal-dependent noise, such as optical amplifier noise. Many high-speed receivers do incorporate such a feature. However, many of the simpler receivers do not have a variable threshold adjustment and set their threshold corresponding to the average received current level, namely, $(I_1 + I_0)/2$. This threshold setting yields a higher bit error rate given by

$$\text{BER} = \frac{1}{2}\left[Q\left(\frac{(I_1 - I_0)}{2\sigma_1}\right) + Q\left(\frac{(I_1 - I_0)}{2\sigma_0}\right)\right].$$

We can use (4.14) to evaluate the BER when the received signal powers for a 0 bit and a 1 bit and the noise statistics are known. Often, we are interested in the inverse problem, namely, determining what it takes to achieve a specified BER. This leads us to the notion of *receiver sensitivity*. The receiver sensitivity $\bar{P}_{\text{sens}}$ is defined as the minimum average optical power necessary to achieve a specified BER, usually 10^{-12} or better. Sometimes the receiver sensitivity is also expressed as the number of photons required per 1 bit, M, which is given by

$$M = \frac{2\bar{P}_{\text{sens}}}{hf_c B},$$

where B is the bit rate.

In the notation introduced earlier, the receiver sensitivity is obtained by solving (4.14) for the average power per bit $(P_0 + P_1)/2$ for the specified BER, say, 10^{-12}. Assuming $P_0 = 0$, this can be obtained as

$$\bar{P}_{\text{sens}} = \frac{(\sigma_0 + \sigma_1)\gamma}{2G_m\mathcal{R}}. \tag{4.15}$$

Here, G_m is the multiplicative gain for APD receivers and is unity for *pin* photodiodes.

First consider an APD or a *pin* receiver, with no optical amplifier in the system. The thermal noise current is independent of the received optical power. However, the shot noise variance is a function of $\bar{P}_{\text{sens}}$. Assume that no power is transmitted for a 0 bit. Then $\sigma_0^2 = \sigma_{\text{thermal}}^2$ and $\sigma_1^2 = \sigma_{\text{thermal}}^2 + \sigma_{\text{shot}}^2$, where the shot noise variance

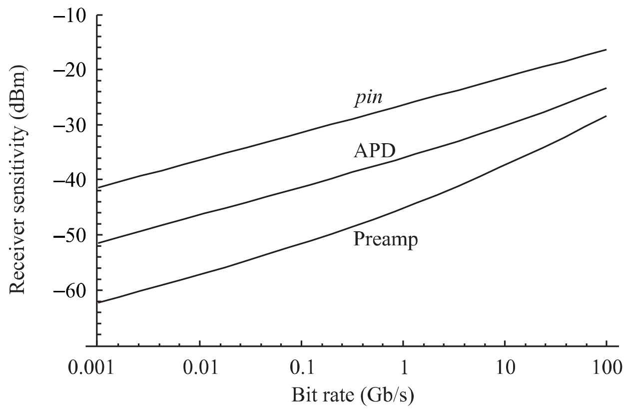

Figure 4.9 Sensitivity plotted as a function of bit rate for typical *pin*, APD, and optically preamplified receivers. The parameters used for the receivers are described in the text.

σ^2_{shot} must be evaluated for the received optical power $P_1 = 2\bar{P}_{\text{sens}}$ that corresponds to a 1 bit. From (4.4),

$$\sigma^2_{\text{shot}} = 4eG_m^2 F_A(G_m)\mathcal{R}\bar{P}_{\text{sens}}B_e.$$

Using this and solving (4.15) for the receiver sensitivity $\bar{P}_{\text{sens}}$, we get

$$\bar{P}_{\text{sens}} = \frac{\gamma}{\mathcal{R}}\left(eB_eF_A(G_m)\gamma + \frac{\sigma_{\text{thermal}}}{G_m}\right). \tag{4.16}$$

Assume that for a bit rate of B b/s, a receiver bandwidth $B_e = B/2$ Hz is required. Let the front-end amplifier noise figure $F_n = 3$ dB and the load resistor $R_L = 100\ \Omega$. Then, assuming the temperature $T = 300°$K, the thermal noise current variance, from (4.3), is

$$\sigma^2_{\text{thermal}} = \frac{4k_BT}{R_L}F_nB_e = 1.656 \times 10^{-22}B\ \text{A}^2. \tag{4.17}$$

Assuming the receiver operates in the 1.55 μm band, the quantum efficiency $\eta = 1$, $\mathcal{R} = 1.55/1.24 = 1.25$ A/W. Using these values, we can compute the sensitivity of a *pin* receiver from (4.16) by setting $G_m = 1$. For BER $= 10^{-12}$ and thus $\gamma \approx 7$, the receiver sensitivity of a *pin* diode is plotted as a function of the bit rate in Figure 4.9. In the same figure, the sensitivity of an APD receiver with $k_A = 0.7$ and an avalanche multiplicative gain $G_m = 10$ is also plotted. It can be seen that the APD receiver has a sensitivity advantage of about 8–10 dB over a *pin* receiver.

We now derive the sensitivity of the optically preamplified receiver shown in Figure 4.7. In amplified systems, the signal-spontaneous beat noise component usually dominates over all the other noise components, unless the optical bandwidth B_o is large, in which case the spontaneous-spontaneous beat noise can also be significant. Making this assumption, the bit error rate can be calculated, using (4.6), (4.9), and (4.14), as

$$\text{BER} = Q\left(\frac{\sqrt{GP}}{2\sqrt{(G-1)P_n B_e}}\right). \tag{4.18}$$

Let us see what receiver sensitivity can be obtained for an ideal preamplified receiver. The receiver sensitivity is measured either in terms of the required power at a particular bit rate or in terms of the number of photons per bit required. As before, we can assume that $B_e = B/2$. Assuming that the amplifier gain G is large and that the spontaneous emission factor $n_{\text{sp}} = 1$, we get

$$\text{BER} = Q\left(\sqrt{\frac{M}{2}}\right).$$

To obtain a BER of 10^{-12}, the argument to the $Q(.)$ function γ must be 7. This yields a receiver sensitivity of $M = 98$ photons per 1 bit. In practice, an optical filter is used between the amplifier and the receiver to limit the optical bandwidth B_o and thus reduce the spontaneous-spontaneous and shot noise components in the receiver. For practical preamplified receivers, receiver sensitivities of a few hundred photons per 1 bit are achievable. In contrast, a direct detection pinFET receiver without a preamplifier has a sensitivity of the order of a few thousand photons per 1 bit.

Figure 4.9 also plots the receiver sensitivity for an optically preamplified receiver, assuming a noise figure of 6 dB for the amplifier and an optical bandwidth $B_o =$ 50 GHz that is limited by a filter in front of the amplifier. From Figure 4.9, we see that the sensitivity of a *pin* receiver at a bit rate of 10 Gb/s is −21 dBm and that of an APD receiver is −30 dBm. For 10 Gb/s operation, commercial *pin* receivers with sensitivities of −18 dBm and APD receivers with sensitivities of −24 dBm are available today. From the same figure, at 2.5 Gb/s, the sensitivities of *pin* and APD receivers are −24 dBm and −34 dBm, respectively. Commercial *pin* and APD receivers with nearly these sensitivities at 2.5 Gb/s are available today.

In systems with cascades of optical amplifiers, the notion of sensitivity is not very useful because the signal reaching the receiver already has a lot of added amplifier noise. In this case, the two parameters that are measured are the average received signal power, $\bar{P}_{\text{rec}}$, and the received optical noise power, P_{ASE}. The *optical signal-to-noise ratio* (OSNR) is defined as $\bar{P}_{\text{rec}}/P_{\text{ASE}}$. In the case of an optically preamplified receiver, $P_{\text{ASE}} = 2P_n(G-1)B_o$. A system designer needs to relate the measured

OSNR with the bit error rate. Neglecting the receiver thermal noise and shot noise, it can be shown using (4.6), (4.9), (4.10), and (4.14) that the argument to the $Q(.)$ function, γ, is related to the OSNR as follows:

$$\gamma = \frac{2\sqrt{\frac{B_o}{B_e}}\text{OSNR}}{1 + \sqrt{1 + 4\text{OSNR}}}. \tag{4.19}$$

Consider a typical 2.5 Gb/s system with $B_e = 2$ GHz, with an optical filter with bandwidth $B_o = 36$ GHz placed between the amplifier cascade and the receiver. For $\gamma = 7$, this system requires an OSNR $= 4.37$, or 6.4 dB. However, this is usually not sufficient because the system must deal with a variety of impairments, such as dispersion and nonlinearities. We will study these in Chapter 5. A rough rule of thumb used by system designers is to design the amplifier cascade to obtain an OSNR of at least 20 dB at the receiver, so as to allow sufficient margin to deal with the other impairments.

4.4.7 Coherent Detection

We saw earlier that simple direct detection receivers are limited by thermal noise and do not achieve the shot noise limited sensitivities of ideal receivers. We saw that the sensitivity could be improved significantly by using an optical preamplifier. Another way to improve the receiver sensitivity is to use a technique called *coherent detection*.

The key idea behind coherent detection is to provide gain to the signal by mixing it with another local light signal from a so-called local-oscillator laser. At the same time, the dominant noise in the receiver becomes the shot noise due to the local oscillator, allowing the receiver to achieve the shot noise limited sensitivity. (In fact, a radio receiver works very much in this fashion except that it operates at radio, rather than light, frequencies.)

A simple coherent receiver is shown in Figure 4.10. The incoming light signal is mixed with a local-oscillator signal via a 3 dB coupler and sent to the photodetector. (We will ignore the 3 dB splitting loss induced by the coupler since it can be eliminated by a slightly different receiver design—see Problem 4.15.) Assume that the phase and polarization of the two waves are perfectly matched. The power seen by the photodetector is then

$$\begin{aligned} P_r(t) &= \left[\sqrt{2aP}\cos(2\pi f_c t) + \sqrt{2P_{\text{LO}}}\cos(2\pi f_{\text{LO}} t)\right]^2 \\ &= aP + P_{\text{LO}} + 2\sqrt{aPP_{\text{LO}}}\cos[2\pi(f_c - f_{\text{LO}})t]. \end{aligned} \tag{4.20}$$

Here, P denotes the input signal power, P_{LO} the local-oscillator power, $a = 1$ or 0 depending on whether a 1 or 0 bit is transmitted (for an OOK signal), and f_c and

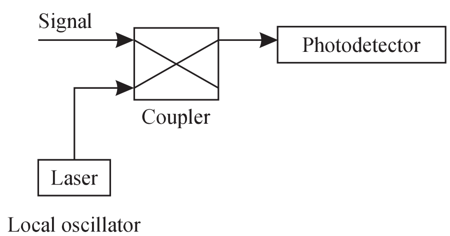

Figure 4.10 A simple coherent receiver.

f_{LO} represent the carrier frequencies of the signal and local-oscillator waves. We have neglected the $2f_c$, $2f_{LO}$, and $f_c + f_{LO}$ components since they will be filtered out by the receiver. In a *homodyne* receiver, $f_c = f_{LO}$, and in a *heterodyne* receiver, $f_c - f_{LO} = f_{IF} \neq 0$. Here, f_{IF} is called the intermediate frequency (IF), typically a few gigahertz.

To illustrate why coherent detection yields improved receiver sensitivities, consider the case of a homodyne receiver. For a 1 bit, we have

$$I_1 = \mathcal{R}(P + P_{LO} + 2\sqrt{PP_{LO}}),$$

and for a 0 bit,

$$I_0 = \mathcal{R}P_{LO}.$$

The key thing to note here is that by making the local-oscillator power P_{LO} sufficiently large, we can make the shot noise dominate over all the other noise components in the receiver. Thus the noise variances are

$$\sigma_1^2 = 2eI_1B_e$$

and

$$\sigma_0^2 = 2eI_0B_e.$$

Usually, P_{LO} is around 0 dBm and P is less than −20 dBm. So we can also neglect P compared to P_{LO} when computing the signal power, and both P and $\sqrt{PP_{LO}}$ compared to P_{LO} when computing the noise variance σ_1^2. With this assumption, using (4.14), the bit error rate is given by

$$\text{BER} = Q\left(\sqrt{\frac{\mathcal{R}P}{2eB_e}}\right).$$

As before, assuming $B_e = B/2$, this expression can be rewritten as

$$\text{BER} = Q(\sqrt{M}),$$

where M is the number of photons per 1 bit as before. For a BER of 10^{-12}, we need the argument of the $Q(.)$ function γ to be 7. This yields a receiver sensitivity of 49 photons per 1 bit, which is significantly better than the sensitivity of a simple direct detection receiver.

However, coherent receivers are generally quite complex to implement and must deal with a variety of impairments. Note that in our derivation we assumed that the phase and polarization of the two waves match perfectly. In practice, this is not the case. If the polarizations are orthogonal, the mixing produces no output. Thus coherent receivers are highly sensitive to variations in the polarizations of the signal and local-oscillator waves as well as any phase noise present in the two signals. There are ways to get around these obstacles by designing more complicated receiver structures [KBW96, Gre93]. However, direct detection receivers with optical preamplifiers, which yield comparable receiver sensitivities, provide a simpler alternative and are widely used today.

Yet another advantage is to be gained by using coherent receivers in a multichannel WDM system. Instead of using a demultiplexer or filter to select the desired signal optically, with coherent receivers, this selection can be done in the IF domain using electronic filters, which can be designed to have very sharp skirts. This allows very tight channel spacings to be achieved. In addition, in a WDM system, the receiver can be tuned between channels in the IF domain, allowing for rapid tunability between channels, a desirable feature to support fast packet switching. However, we will require highly wavelength-stable and controllable lasers and components to make use of this benefit. Such improvements may result in the resurrection of coherent receivers when WDM systems with large numbers of channels are designed in the future.

4.4.8 Timing Recovery

The process of determining the bit boundaries is called *timing recovery.* The first step is to extract the clock from the received signal. Recall that the clock is a periodic waveform whose period is the bit interval (Section 4.4). This clock is sometimes sent separately by the transmitter, for example, in a different frequency band. Usually, however, the clock must be extracted from the received signal. Even if the extracted clock has a period equal to the bit interval, it may still be out of phase with the received signal; that is, the clock may be offset from the bit boundaries. Usually, both the clock frequency (periodicity) and its phase are recovered simultaneously by a single circuit, as shown in Figure 4.11.

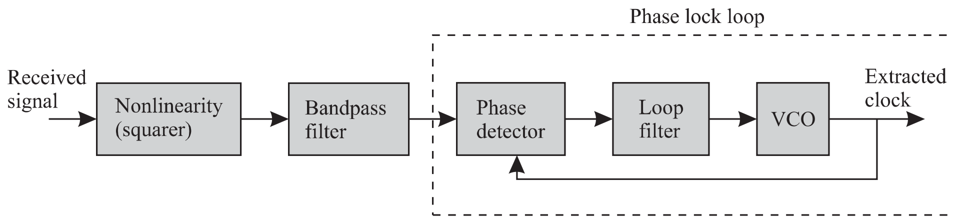

Figure 4.11 Block diagram illustrating timing, or clock, recovery at the receiver.

If we pass the received signal through a nonlinearity, typically some circuit that calculates the square of the received signal, it can be shown that the result contains a spectral component at $1/T$, where T is the bit period. Thus, we can filter the result using a bandpass filter as shown in Figure 4.11 to get a waveform that is approximately periodic with period T and that we call a *timing signal*. However, this waveform will still have considerable *jitter;* that is, successive "periods" will have slightly different durations. A "clean" clock with low jitter can be obtained by using the *phase lock loop* (PLL) circuit shown in Figure 4.11.

A PLL consists of a voltage-controlled oscillator (VCO), a phase detector, and a loop filter. A VCO is an oscillator whose output frequency can be controlled by an input voltage. A *phase detector* produces an error signal that depends on the difference in phase between its two inputs. Thus, if the timing signal and the output of the VCO are input to the phase detector, it produces an error signal that is used to adjust the output of the VCO to match the (average) frequency and phase of the timing signal. When this adjustment is complete, the output of the VCO serves as the clock that is used to sample the filtered signal in order to decide upon the values of the transmitted bits. The *loop filter* shown in Figure 4.11 is a critical element of a PLL and determines the residual jitter in the output of the VCO, as well as the ability of the PLL to track changes in the frequency and phase of the timing signal.

4.4.9 Equalization

We remarked in Section 4.4 with reference to Figure 4.5 that the receive filter that is used just prior to sampling the signal can incorporate an *equalization filter* to cancel the effects of intersymbol interference due to pulse spreading. From the viewpoint of the electrical signal that has been received, the entire optical system (including the laser, the fiber, and the photodetector) constitutes the *channel* over which the signal has been transmitted. If nonlinearities are ignored, the main distortion caused by this channel is the dispersion-induced broadening of the (electrical) pulse. Dispersion is

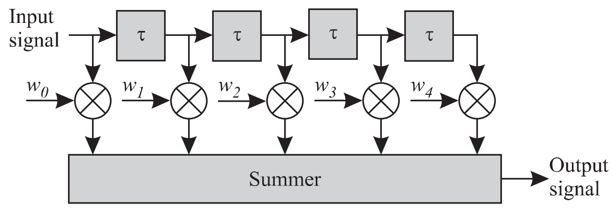

Figure 4.12 A transversal filter, a commonly used structure for equalization. The output (equalized) signal is obtained by adding together suitably delayed versions of the input signal, with appropriate weights.

a linear effect, and hence the effect of the channel on the pulse, due to dispersion, can be modeled by the response of a filter with transfer function $H_D(f)$. Hence, in principle, by using the inverse of this filter, say, $H_D^{-1}(f)$, as the equalization filter, this effect can be canceled completely at the receiver. This is what an equalization filter attempts to accomplish.

The effect of an equalization filter is very similar to the effect of dispersion compensating fiber (DCF). The only difference is that in the case of DCF, the equalization is in the optical domain, whereas equalization is done electrically when using an equalization filter. As in the case of DCF, the equalization filter depends not only on the type of fiber used but also on the fiber length.

A commonly used filter structure for equalization is shown in Figure 4.12. This filter structure is called a *transversal filter.* It is essentially a tapped delay line: the signal is delayed by various amounts and added together with individual weights. The choice of the weights, together with the delays, determines the transfer function of the equalization filter. The weights of the tapped delay line have to be adjusted to provide the best possible cancellation of the dispersion-induced pulse broadening.

Electronic equalization involves a significant amount of processing that is difficult to do at higher bit rates, such as 10 Gb/s. Thus optical techniques for dispersion compensation, such as the use of DCF for chromatic dispersion compensation, are currently much more widely used compared to electronic equalization.

4.5 Error Detection and Correction

An error-correcting code is a technique for reducing the bit error rate on a communicaton channel. It involves transmitting additional bits, called *redundancy,* along with the data bits. These additional bits carry redundant information and are used by the receiver to correct most of the errors in the data bits. This method of reducing the