Chapter 3

COHERENT OPTICAL RECEIVERS AND IDEAL PERFORMANCE

Coherent detection of optical signal is first used for its superior receiver sensitivity compared to on-off keying. Equivalently speaking, the mixing of received signal with the local oscillator (LO) laser functions as an optical amplifier without noise enhancement. Even with the advances of Erbium-doped fiber amplifiers (EDFA), coherent detection can still provide better receiver sensitivity than amplified on-off keying. In this chapter, various structures and architectures of optical receiver are studied for phase-modulated or coherent optical communications. This chapter focuses on performance analysis to validate the receiver sensitivity improvement of phase-modulated optical communications, mainly for binary signals. However, for binary signals limited by amplifier noises, the improvement is only up to about 3 dB.

In a coherent receiver, phase-locked loop (PLL) may be required to track the phase of the received signal. In coherent optical communications, receiver with phase tracking is called synchronous receiver. Equivalent to the matched filter receiver in digital communications (Proakis, 2000), receivers with phase tracking always provide the optimal performance, at least at the linear regime.

Without phase tracking, the received signal has a random phase and can be detected based on the power or the envelope of the signal (Proakis, 2000). This type of noncoherent receiver is called asynchronous receiver in coherent optical communications. While the performance is typically inferior to synchronous receiver, asynchronous receiver has simple structure and provides low-cost implementation.

Coherent optical signal can also directly be detected without mixing with LO signal. While on-off keying or, equivalently, amplitude-shift keying (ASK) signal is directly detected by a photodiode, both phase-
or frequency-modulated signals have direct-detection receiver based on interferometer or optical filter. Direct-detection receiver may be the simplest receiver with low-cost implementation. However, direct-detection receiver for both phase- or frequency-modulated signal is still more complicated than a single photodiode to detect the presence or absence of light for on-off keying signal.

Another types of noncoherent receiver use phase-diversity techniques that combines two signals with a phase difference of 90°. Phase-diversity receiver is asynchronous receiver without the requirement of phase tracking. Equivalently speaking, phase-diversity receiver implements the envelope or power detection by combining optical and electrical techniques.

Other than polarization-diversity receivers, the mixing of two optical signals requires the alignment of their polarizations. In general, polarization alignment is provided by polarization controller. With optimal combining of the signal from both polarizations, polarization-diversity receivers have the same performance as receiver with polarization tracking.

Various types of coherent optical receiver will be discussed in this chapter.

1. Basic Coherent Receiver Structures

The basic structures of phase-shift keying (PSK) and differential phase-shift keying (DPSK) receivers have been shown in Figs. 1.3 and 1.4, respectively. This section studies each basic type of coherent optical receivers that mix the received signal with the LO laser. The signal-to-noise ratio (SNR) of each receiver type is derived, especially for systems limited by amplifier noises. The SNR is equal to the ratio of optical signal to the amplifier noise per polarization over an optical bandwidth equal to the data rate. For system without optical amplifiers, the SNR is equal to the number of photons per bit for heterodyne receiver.

1.1 Single-Branch Receiver

Figure 3.1 shows a typical structure of a single-branch coherent optical receiver. To enable optical signal mixing, the polarization of the received signal must be aligned to that of the LO laser. In Fig. 3.1, automatic polarization control (APC) is used to align the polarization of the received signal to that of the LO laser for optimal signal mixing. In general, the LO laser is phase or frequency locked to the received signal. Phase locking is used for homodyne receiver and frequency locking is required for heterodyne receiver for a fixed intermediate frequency (IF). Phase locking is facilitated by an optical PLL and frequency locking
is provided by an automatic frequency control (AFC) loop. The single-branch receiver is the simplest receiver structure. Most early heterodyne receivers used single-branch receiver for its simplicity (Goodwin, 1967, Nussmeier et al., 1974, Oliver, 1961, Peyton et al., 1972, Saito et al., 1981). Instead of 3-dB coupler, those early works used a power beam splitter or a power combiner, functioning as an 180° optical hybrid, to mix the received signal with the LO laser.

A 3-dB coupler is used in Fig. 3.1 to mix the received signal with the LO laser. The input and output relationship of a 3-dB coupler is\(^1\)

\[
S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.
\]  

(3.1)

The received signal is assumed to be

\[
E_r(t) = [A_s(t)e^{j\phi_s(t)} + n_x(t)]e^{j\omega_c t}x + n_y(t)e^{j\omega_c t}y,
\]

(3.2)

where \(A_s(t)\) and \(\phi_s(t)\) are the modulated amplitude and phase of the transmitted signal, respectively, \(\omega_c\) is carrier frequency of the signal, \(x\) is the polarization of the signal, and \(y\) is the polarization orthogonal to \(x\), \(n_x(t)\) and \(n_y(t)\) are the amplifier noise in the polarizations of \(x\) and \(y\), respectively. The complex signal of \(A_s(t)e^{j\phi_s(t)}\) is the low-pass representation of the signal. \(A_s(t)e^{j\phi_s(t)} = \pm A\) for both PSK and DPSK signals

\(^1\)The input and output relationship can also be

\[
S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}.
\]

without changing the phase relationship. Different input and output relationship for an 180° optical hybrid may be used in this book for convenience.

---

**Figure 3.1.** A single-branch coherent optical receiver.
and $A_s(t)e^{j\phi_s(t)} = \{0, A\}$ for ASK signals, where $A$ is the positive signal amplitude. Both $n_x(t)$ and $n_y(t)$ are the low-pass representation of the amplifier noise. Here, in this chapter, we assume that both the transmitted and LO lasers are a pure coherent source without phase noise. In next chapter, the impact of laser phase noise on phase-modulated signal is analyzed for further detail. The received signal may have a random phase of $\theta_0$ of $A_s(t)e^{j\phi_s(t)+j\theta_0}$ due to propagation delay. For system with phase tracking using PLL, we assume that $\theta_0 = 0$ for simplicity when the phase of LO laser tracks out the phase of $\theta_0$. For system without phase tracking, the phase of $\theta_0$ is included when necessary.

The optical SNR of the received signal is defined as

$$\text{SNR}_{o,s} = \frac{E\{|A_s(t)|^2\}}{E\{|n_x(t)|^2\} + E\{|n_y(t)|^2\}} = \frac{P_r}{2S_{ns}\Delta f_{o,s}}, \quad (3.3)$$

where $P_r$ is the received power, $S_{ns}$ is the spectral density of the received spontaneous emission in each polarization, and $\Delta f_{o,s}$ is the optical bandwidth of the received optical filter. In the complex representation of $n_x(t) = n_{x1}(t) + jn_{x2}(t)$, the spectral densities of $n_{x1}(t)$ and $n_{x2}(t)$ are same and equal to $S_{ns}/2$.

The LO signal is

$$E_{\text{LO}}(t) = [A_L + n_L(t)]e^{j\omega_{\text{LO}}t}, \quad (3.4)$$

where $A_L$ is the continuous-wave amplitude of LO laser, $n_L(t)$ is the noise of LO laser in the same polarization as the signal, and $\omega_{\text{LO}}$ is the angular frequency of LO laser. The polarization of the received signal of Eq. (3.2) and the LO laser of Eq. (3.4) is assumed to be the same using APC. The noise of $n_L(t)$ may originate from the optical amplifier used to boost up the LO power or from the relative intensity noise (RIN) of the LO laser. In the LO electric field of Eq. (3.4), the noise at the polarization orthogonal to the signal is ignored here for simplicity. In practice, the noise from polarization orthogonal to the signal can be filtered by a polarizer, especially at a receiver having a well-controlled LO laser.

In the single-branch receiver of Fig. 3.1, the electric field at the input of the photodiode is $[E_r(t) + E_{\text{LO}}(t)]/\sqrt{2}$, the photocurrent is

$$i(t) = \frac{R}{2} [E_r(t) + E_{\text{LO}}(t)]^2 + i_{\text{sh}} + i_{\text{th}}, \quad (3.5)$$

where $R$ is the responsivity of the photodiode, $i_{\text{sh}}$ is the photocurrent shot-noise, and $i_{\text{th}}$ is the thermal noise of the receiver. In later part of this chapter, other than specific, we ignored thermal noise for simplicity.
As an additive noise, the effect of thermal noise can be added to the signal afterward. In coherent optical communication systems with LO laser, thermal noise has less impact than both shot and amplifier noises.

The photodiode responsivity is equal to

\[ R = \eta \frac{e}{\hbar \omega_c}, \]  

where \( e = 1.6 \times 10^{-19} \) C is the charge per electron, \( \hbar \omega_c \) is the energy per photon, where \( \hbar = h/(2\pi) \) with \( h = 6.63 \times 10^{-34} \) J s as the Planck constant, \( \eta \) is the quantum efficiency of photodiode that is the average number of electrons generated per a photon by the photodiode.

The photocurrent of Eq. (3.5) is equal to

\[
i(t) = \frac{R}{2} \left\{ \left| A_L + n_L(t) \right|^2 + \left| A_s(t)e^{j\phi_s(t)} + n_x(t) \right|^2 + \left| n_y(t) \right|^2 \right\} 
+ RA_L A_s(t) \cos[\omega_{IF} t + \phi_s(t)] 
+ RRe \left\{ \left[ A_L n_x(t) + A_s(t)e^{j\phi_s(t)} n_L(t) \right] e^{j\omega_{IF}t} \right\} + i_{sh}. \]  

In the photocurrent of Eq. (3.7), the intermediate frequency (IF) is \( \omega_{IF} = \omega_c - \omega_{LO} \). Homodyne system has \( \omega_{IF} = 0 \) but heterodyne system has \( \omega_{IF} \neq 0 \). The photocurrent of Eq. (3.7) includes the intensity of LO laser with noise of \( |A_L + n_L(t)|^2 \), the intensity of received signal with noise of \( |A_s(t)e^{j\phi_s(t)} + n_x(t)|^2 \), the intensity of spontaneous emission from orthogonal polarization of \( |n_y(t)|^2 \), the beating of received signal with LO laser of \( RA_L A_s(t) \cos[\omega_{IF} t + \phi_s(t)] \), LO and spontaneous beating of \( A_L n_x(t) \), signal and LO-spontaneous beating of \( A_s(t)n_L(t) \), and together with shot noise \( i_{sh} \). In the photocurrent of Eq. (3.7), the small effect of the beating of spontaneous noise with shot noise is ignored. If necessary, thermal noise can add to the photocurrent of Eq. (3.7).

In the photocurrent of Eq. (3.7), the signal component is

\[ s(t) = RA_L A_s(t) \cos[\omega_{IF} t + \phi_s(t)]. \]  

For heterodyne system with \( \omega_{IF} = \omega_c - \omega_{LO} \neq 0 \), the signal power is

\[ P_s = \frac{R^2}{2} P_{LO} P_r, \]  

where \( P_{LO} = A_L^2 \) is the LO power, \( P_r = E\{|A_s(t)|^2\} \) is the received power. In PSK homodyne system with \( \omega_c = \omega_{LO} \), we obtain

\[ P_s = R^2 P_{LO} P_r. \]
has twice the signal power of heterodyne system. Conventionally, it was generally believed that homodyne system is 3-dB better than the corresponding heterodyne system (Betti et al., 1995, Hooijmans, 1994, Okoshi and Kikuchi, 1988, Ryu, 1995). However, for system dominated by optical amplifier noise, heterodyne system generally has the same performance as similar homodyne system.

The dominant noise source for the single-branch receiver depends on the system configuration. In the usual case that the LO power is significantly larger then the received power, i.e., $P_{LO} \gg P_r$, the dominant noise is usually LO-spontaneous beating noise, given by $\mathcal{R}\mathcal{R}\{A_Ln_x(t)e^{j\omega_1Ft}\}$ in the photocurrent of Eq. (3.7).

If the spontaneous emissions of LO and received signal are not negligible with $n_L(t) \neq 0$ and $n_x(t) \neq 0$, the LO-spontaneous beating including two components of $\mathcal{R}\mathcal{R}\{A_Ln_L(t)\}$ from $\frac{1}{2}R|A_L + n_L(t)|^2$ and $\mathcal{R}\mathcal{R}\{A_Ln_x(t)e^{j\omega_1Ft}\}$. If the optical bandwidth of $n_L(t)$ is $\Delta f_{o,L}$ centered around $\omega_L$, the bandwidth of $\mathcal{R}\mathcal{R}\{A_Ln_L(t)\}$ is about $\Delta f_{o,L}/2$. For heterodyne systems with $f_{IF} = \omega_{IF}/(2\pi) \neq 0$ and $f_{IF} > \frac{1}{2}\Delta f_{o,L} + B_d$, an electric filter can be designed such that the beating of $A_L$ with $n_L(t)$ does not affect the performance of the system, where $B_d$ is the symbol rate of the data channel. In homodyne system with $\omega_{IF} = f_{IF} = 0$, the beating of $A_L$ with $n_L(t)$ always contributes to the noise of the system.

In the LO and spontaneous noise $n_x(t)$ beating of $\mathcal{R}\mathcal{R}\{A_Ln_x(t)e^{j\omega_1Ft}\}$, if the optical bandwidth of $n_x(t)$ is $\Delta f_{o,s}$ centered around $\omega_c$, the upper and lower frequency of LO-spontaneous beating noise is $f_{IF} \pm \frac{1}{2}\Delta f_{o,s}$. Figure 3.2 illustrates the effect of the optical filter bandwidth of $\Delta f_{o,s}$ on the beating of LO laser and spontaneous emission of $n_x(t)$. The optical filter should have a center frequency align with the optical signal.

Without loss of generality and assume that $f_{IF} > 0$ as in Fig. 3.2, the lower frequency of $f_{IF} - \frac{1}{2}\Delta f_{o,s}$ may be a negative frequency for large $\Delta f_{o,s}$. The negative frequency noise affects the system if $f_{IF} - \frac{1}{2}\Delta f_{o,s}$ falls into the data bandwidth of $-f_{IF} + B_d$.

The upper two traces of Fig. 3.2 show the case when the optical bandwidth of $\Delta f_{o,s}$ is small and comparable to twice the data bandwidth of $B_d$. The beating noise is band-pass noise centered at $f_{IF}$. The lower two traces of Fig. 3.2 show the case when the optical bandwidth of $\Delta f_{o,s}$ is significantly larger than the data bandwidth of $B_d$. With $f_{IF} - \frac{1}{2}\Delta f_{o,s} < 0$, the beating noise may extend to $-f_{IF} + B_d$. In the worst case of having a wide-bandwidth receiver optical filter, the beating noise centered at $f_{IF}$ with a bandwidth of $2B_d$ is doubled compared with case of narrow-bandwidth optical filter. In the diagram of Fig. 3.2, the data bandwidth is assumed to be $2B_d$. In practical system, depending on linecode or spectral filtering, the data bandwidth may vary.
Figure 3.2. An illustration of the effects of receiver filter bandwidth of $\Delta f_{o,s}$ on LO-spontaneous beat noise. The optical bandwidth of $\Delta f_{o,s}$ is comparable (upper two traces) and larger than (lower two traces) than twice the data bandwidth.

The beating noise is not doubled if

$$ -f_{IF} + \frac{1}{2} \Delta f_{o,s} < f_{IF} - B_d $$

or

$$ f_{IF} > \frac{1}{4} (\Delta f_{o,s} + 2B_d). $$

In homodyne system, the LO and spontaneous noise $n_x(t)$ beating is within the frequencies of $\pm \Delta f_{o,s}/2$ and contributes to the positive and negative frequency noise twice. Because the power of homodyne receiver is twice that of heterodyne receiver by comparing Eq. (3.9) and (3.10), with the condition of Eq. (3.12), the output SNR of a single-branch homodyne and heterodyne receivers is the same. However, without the condition of Eq. (3.12), the noise of homodyne and heterodyne system is the same, homodyne system is 3 dB better than heterodyne system due to its higher power.

Under the condition of Eq. (3.12), heterodyne and homodyne systems have the same performance. With proper design, a coherent system with optical amplifiers is generally limited by the beating of LO-spontaneous emission beating between $\Delta f_{o,s}$ and spontaneous emission of $n_x(t)$. When the optical filter has a bandwidth of $\Delta f_{o,s} = 2B_d$, the IF must be larger than $f_{IF} > B_d$. In the limit, the narrowest optical bandwidth is actually
\( \Delta f_{o,s} = B_d \) for an optical match filter. However, for return-to-zero (RZ) signal with small duty cycle, the optical bandwidth of an optical match filter is larger than the data-rate of \( B_d \).

In general, the condition of Eq. (3.12) is a good approximation for most practical cases. While the bandwidth of \( f_{IF} \pm B_d \) is a good approximation, ideal sinc pulse has the bandwidth of \( f_{IF} \pm B_d/2 \) but RZ pulse has a bandwidth wider than \( 2B_d \).

Assuming the condition of Eq. (3.12), the LO-spontaneous beating noise is
\[
RA_Ln_{x1}(t)\cos(\omega_{IF}t) - RA_Ln_{x2}(t)\sin(\omega_{IF}t), \quad (3.13)
\]
where \( n_s(t) = n_{x1}(t) + jn_{x2}(t) \) with \( E\{n_{x1}^2(t)\} = E\{n_{x2}^2(t)\} = \frac{1}{2}S_{ns}B_d \).

The LO laser source of \( A_L \) beating with \( n_x(t) \) induces an electrical noise with spectrum density of
\[
N_{A_L-n_s} = \frac{1}{2}R^2S_{ns}P_{LO}, \quad (3.14)
\]
The LO laser source of \( A_L \) beating with \( n_L(t) \) induces electrical noise with spectrum density of
\[
N_{A_L-n_L} = \frac{1}{2}R^2S_{n_L}P_{LO}, \quad (3.15)
\]
where \( S_{n_L} \) is the spectrum density of the spontaneous of \( n_L(t) \) at optical domain.

The optical SNR of LO laser source is
\[
\text{SNR}_{o,LO} = \frac{P_{LO}}{2S_{n_L}\Delta f_{o,L}}, \quad (3.16)
\]
where the amplifier noise is in the same polarization as the signal and the orthogonal polarization, in contrast to the LO electric field of Eq. (3.4).

Because the receiver by itself should not induce noise into the system, the system must be designed for \( N_{A_L-n_s} > 10 \times N_{A_L-n_L} \), or \( S_{ns} > 10 \times S_{n_L} \), where the factor of 10 is for the condition that the additional penalty due to \( n_L(t) \) is less than 0.5 dB. For optical SNR defined for the same bandwidth of \( \Delta f_{o,s} = \Delta f_{o,L} \), the system requirement is
\[
\frac{\text{SNR}_{o,LO}}{\text{SNR}_{o,s}} > 10 \times \frac{P_{LO}}{P_r}, \quad (3.17)
\]
with the condition of \( P_{LO} \gg P_r \), the required optical SNR of the LO source must be far larger than that of the receive signal.

If the optical SNR of LO source is not substantially larger than that of the signal source, the system is dominated by noise from the LO source.
The noise of the LO source may be induced by an optical amplifier that boosts up the power of the LO laser or the RIN of the LO laser. In order to eliminate the effects of the LO laser noise, a balanced receiver can be used instead. Because balanced receiver has better performance than single-branch receiver, the analysis of this book usually assumes a balanced receiver.

In the receiver of Fig. 3.1, many methods can control the polarization though APC (Aarts and Khoe, 1989, Martinelli and Chipman, 2003, Noé et al., 1988a, 1991, Okoshi, 1985, Walker and Walker, 1990). The polarization control algorithm should be able to provide endless polarization control without reset. While early homodyne or heterodyne coherent optical communication systems used single-branch receiver, balanced receiver is more popular for its superior performance.

1.2 Balanced Receiver

Figure 3.3 shows a dual-photodiode balanced receiver that increases the signal power and eliminates the noise from the LO laser source. Similar to single-branch receiver of Fig. 3.1, balanced receiver of Fig. 3.3 requires both polarization alignment using APC and phase or frequency locking. The electric field at the input of the upper photodiode is \( \frac{[E_r(t) + E_{LO}(t)]}{\sqrt{2}} \), the photocurrent is the same as Eq. (3.5) and is equal to

\[
    i_1(t) = \frac{R}{2} |E_r(t) + E_{LO}(t)|^2 + i_{sh1}, \tag{3.18}
\]

where \( i_{sh1} \) is the shot noise. The electric field at the input of the lower photodiode is \( \frac{[E_r(t) - E_{LO}(t)]}{\sqrt{2}} \), the photocurrent is

\[
    i_2(t) = \frac{R}{2} |E_r(t) - E_{LO}(t)|^2 + i_{sh2}, \tag{3.19}
\]
where $i_{\text{sh}2}$ is the shot noise.

In both photocurrents of Eqs. (3.18) and (3.19), we assume that the two photodiodes are identical with the same responsivity and the 3-dB coupler or 180° optical hybrid is balanced 3-dB without excesses loss. The overall photocurrent is

\[
i(t) = i_1(t) - i_2(t) = 2R\Re\{E_r(t) \cdot E_{\text{LO}}^*(t)\} + i_{\text{sh}} \tag{3.20}
\]

or

\[
i(t) = 2RA_L A_s(t) \cos[\omega_{\text{IF}} t + \phi_s(t)]
+ 2R\Re\left\{\left[A_L n_x(t) + A_s(t) e^{j\phi_s(t)} n_L(t)\right] e^{j\omega_{\text{IF}} t}\right\} + i_{\text{sh}}, \tag{3.21}
\]

where $i_{\text{sh}} = i_{\text{sh}1} - i_{\text{sh}2}$ is the overall shot noise.

In the photocurrent of Eq. (3.21), the noise from the LO laser of $n_L(t)$ contributes to the system noise because of the beating term of $2R\Re\{A_s(t) n_L(t) e^{j\omega_{\text{IF}} t + j\phi_s(t)}\}$ with a spectral density of

\[
N_{A_s-n_L} = 2R^2 S_{n_L} P_r. \tag{3.22}
\]

Comparing with the LO-spontaneous beating noise of

\[
N_{A_L-n_s} = 2R^2 S_{n_s} P_{\text{LO}}, \tag{3.23}
\]

as long as the optical SNR has the relationship of $\text{SNR}_{o,\text{LO}} > 10 \times \text{SNR}_{o,s}$ or $N_{A_L-n_s} > 10 \times N_{A_s-n_L}$, the signal and LO noise beating does not provide a penalty more than 0.5 dB. Even when a booster amplifier is used, the LO laser required low-gain optical amplifier but the signal usually passes through a chain of many optical amplifiers. The optical SNR of the LO signal is usually much larger than that of the received signal. The signal and spontaneous emission beating of $N_{A_s-n_L}$ is usually much smaller than the LO-spontaneous emission beating of $N_{A_L-n_s}$.

The signal component of the photocurrent of Eq. (3.21) is

\[
s(t) = 2RA_L A_s(t) \cos[\omega_{\text{IF}} t + \phi_s(t)] \tag{3.24}
\]

with a signal amplitude twice of that in single-branch receiver of Eq. (3.8), giving four times larger received power. Since the noise is also increased by the same factor, balanced receiver does not increase the SNR for receiver limited by the beating of LO-spontaneous emission.

When the LO-spontaneous beating of $2R\Re\{A_L n_x(t) e^{j\omega_{\text{IF}} t}\}$ is the dominant noise, similar to the illustration of Fig. 3.2 for single-branch receiver, a heterodyne balanced receiver also requires an optical filter.
with the condition of Eq. (3.12) without doubling the LO-spontaneous
beating noise. For systems limited by optical amplifier noise, the perfor-
ance of heterodyne and homodyne receiver is the same with the con-
tion of Eq. (3.12).

Here, the SNR of a heterodyne system is evaluated with amplifier
noise and the condition of Eq. (3.12) for a balanced receiver, the signal
power is
\[ P_s = 2R^2P_{LO}P_r. \]  

The noise variance at a balanced receiver includes LO-spontaneous
beating noise of
\[ \sigma_{LO-sp}^2 = 2R^2P_{LO}S_{ns}B_d, \]  
and the signal-LO spontaneous beating noise variance of
\[ \sigma_{sig.-LOsp}^2 = 2R^2P_rS_{nL}B_d. \]

The shot noise has a variance of
\[ \sigma_{sh}^2 = 2eR(P_{LO} + P_r + 2S_{ns}\Delta f_{os} + 2S_{nL}\Delta f_{o,L})B_d \]  
for LO, signal, and spontaneous-emission induced shot noise.

For both heterodyne or homodyne systems, the SNR is
\[ \rho_s = \frac{P_r}{S_{ns}B_d + \frac{P_r}{P_{LO}}S_{nL}B_d + \frac{h\omega_c}{\eta} \left( 1 + \frac{P_r}{P_{LO}} \right) B_d}, \]
when the shot noise of spontaneous emission and the beating of sponta-
neous emission with shot noise are ignored.

The LO spontaneous emission noise usually comes from a single am-
plifier with spectral density of
\[ S_{nL} = (G_L - 1)n_{spL}\hbar\omega_{LO}, \]
where \( G_L \) and \( n_{spL} \) are the gain and spontaneous emission factor of the
LO amplifier.

The spontaneous emission noise of \( S_{ns} \) together with the received
signal is induced by a chain of optical amplifiers in the fiber link. Assume
\( N_A \) identical fiber spans with loss of \( 1/G_s \), equal to the gain of the \( N_A \)
identical optical amplifiers, the first optical amplifier has an amplified
spontaneous emission noise of
\[ (G_s - 1)n_{spa}\hbar\omega_c, \]
where \( n_{spa} \) is the spontaneous emission factor of the optical amplifiers
at each fiber span. The above spontaneous emission losses by the factor
of $1/G_s$ at the fiber span, and is amplified by $G_s$ by the second optical amplifier. The second optical amplifier also adds the same spontaneous emission as Eq. (3.31) to the optical signal. After $N_A$ fiber spans, we obtain the overall amplifier noise spectral density of

$$S_{n_s} = N_A(G_s - 1) n_{sp,S} h \omega_c. \quad (3.32)$$

The noise figure of optical amplifier is approximately equal to $2n_{sp,s}$. The spontaneous emission of Eq. (3.32) can be expressed using the noise figure of the optical amplifier as $S_{ns} \approx \frac{1}{2} G_s F_n h \omega_c$, where $F_n$ is the noise figure of each optical amplifier.

### Shot-noise limited systems

A heterodyne system without optical amplifiers is limited by the shot noise with $n_s(t) = 0$ and $n_L(t) = 0$, the signal power is that of Eq. (3.25) and the noise variance is that of Eq. (3.28). If the LO laser has significantly larger power than the receive power of $P_{LO} \gg P_r$, the SNR of a heterodyne system becomes

$$\rho_s = \frac{RP_r}{eB_d}. \quad (3.33)$$

For binary signal, if the average number of photons per bit is $N_s, P_r = N_s h \omega_c B_d$. For multilevel signal, $N_s$ is the average number of photons per symbol. With the photodiode responsivity defined by Eq. (3.6), we obtain

$$\rho_s = \eta N_s. \quad (3.34)$$

In homodyne system, the power is twice that of Eq. (3.25) but the shot noise remains the same as that of Eq. (3.28), the SNR is

$$\rho_s = 2\eta N_s. \quad (3.35)$$

In the quantum limit with $\eta = 1$, the quantum-limited SNR is the photon number per bit of $N_s$ and twice the photon number per bit of $2N_s$ for heterodyne and homodyne systems, respectively. The SNR of Eqs. (3.34) and (3.35) is usually used in the analysis of traditional coherent optical communication systems limited by shot noise (Betti et al., 1995, Okoshi and Kikuchi, 1988).

### Amplifier-noise limited systems

If the amplifier noise is the dominant noise source, with the condition of Eq. (3.12), the SNR after the balanced receiver for heterodyne system
with $\omega_{IF} \neq 0$ is

$$\rho_s = \frac{P_r}{S_{ns} B_d}.$$  \hspace{1cm} (3.36)

The SNR of Eq. (3.36) has a very simple and clear physical meaning of the received signal of $P_r$ over the optical noise within a bandwidth of $B_d$ in a single polarization. As the spontaneous emission has a spectral density of $S_{n_s}$, the noise power per polarization is $S_{n_s} B_d$. In a RF system, with proper filtering, mixing by upconversion or downconversion does not change the SNR. From Eq. (3.36), the SNR in optical domain, before the downconversion by balanced receiver, is also the SNR in the receiver. The definition of the SNR of Eq. (3.36) ignores the amplifier noise from orthogonal polarization that does not beat with the signal. In the photocurrent of Eq. (3.21), the optical amplifier noise from orthogonal polarization does not affect the system.

For a received power of $P_r = G_s N_s h\omega_c B_d$, we obtain

$$\rho_s = \frac{N_s}{N_A n_{sps}} \frac{G_s}{G_s - 1} \approx \frac{N_s}{N_A n_{sps}}.$$  \hspace{1cm} (3.37)

Comparing with Eq. (3.34), the SNR is degraded by the factor of $N_A n_{sps}$. The SNR of Eq. (3.37) is valid for both homodyne and heterodyne systems. For heterodyne systems, the condition of Eq. (3.12) must be satisfied such that the LO-spontaneous emission does not double.

The equivalent spontaneous emission factor can be defined as

$$n_{eq} = N_A n_{sps} (G_s - 1)/G_s \approx \frac{1}{2} N_A F_n.$$  \hspace{1cm} (3.38)

for $N_A$ identical optical amplifiers where $F_n$ is the noise figure of each optical amplifier. With the equivalent $n_{eq}$ of Eq. (3.38), the spectrum density of Eq. (3.32) becomes

$$S_{n_s} = (G_s - 1)n_{eq} h\omega_c,$$  \hspace{1cm} (3.39)

and the SNR is

$$\rho_s = \frac{N_s}{n_{eq}}.$$  \hspace{1cm} (3.40)

If the $N_A$ optical amplifiers are not the same with different gain and noise figure, the equivalent spontaneous emission factor is

$$n_{eq} = \sum_{k=1}^{N_A} \frac{P_{ink} G_k}{G_k} \frac{G_k - 1}{G_k} n_{sps_k},$$  \hspace{1cm} (3.41)

where $G_k, k = 1, \ldots, N_A$, are the gain of each optical amplifier, $P_{ink}, k = 1, \ldots, N_A$, are the input power of each optical amplifier, and $n_{sps_k}$ are the
spontaneous emission factor of each optical amplifier. If all amplifiers are identical with \( G_k = G_s \) and \( n_{sps_k} = n_{sps} \), we obtain \( n_{eq} = N_A n_{sps} (G_s - 1)/G_s \) of Eq. (3.38) again.

The SNR of Eq. (3.40) is the same as the optical SNR defined over a bandwidth of \( B_d \) for a single polarization of

\[
\frac{P_r}{S_{n_s}B_d} = \frac{N_s}{n_{eq}}. \tag{3.42}
\]

In the quantum limit of a single optical amplifier of \( N_A = 1 \) and \( F_n = 2 \) (or 3 dB), \( n_{eq} = 1 \), the quantum limited SNR is equal to the average number of photons per bit of \( N_s \), the same as that of shot-noise limited heterodyne system.

In practice, optical SNR of \( \text{SNR}_{o,s} \) is measured over an optical bandwidth of \( \Delta f_{o,s} \) using an optical spectrum analyzer. The SNR of Eq. (3.40) is equal to

\[
\rho_s = \text{SNR}_{o,s} \frac{2\Delta f_{o,s}}{B_d} \tag{3.43}
\]

where the factor of 2 is for two polarizations in most optical SNR measurement. The optical bandwidth in typical measurements is \( \Delta f_{o,s} = 12.5 \text{ GHz} \), corresponding to 0.1 nm or 1 Å in the wavelength around 1.55 μm. The difference between \( \rho_s \) and \( \text{SNR}_{o,s} \) is about 4.0 and -2.0 dB for 10 and 40-Gb/s signals, respectively.

If no optical filter is used or the optical filter has a wide bandwidth that does not conform to the condition of Eq. (3.12), the amplifier-noise limited SNR for heterodyne receiver is

\[
\rho_s = \frac{N_s}{2n_{eq}}. \tag{3.44}
\]

In later section, we always assume a SNR of Eq. (3.40) for heterodyne receiver. With a balanced receiver, the performance of a system depends solely on the SNR. Both heterodyne and homodyne systems have the same performance if the SNR of the system is the same.

Balanced receiver for coherent optical communications was first analyzed in detail by Abbas et al. (1985), Yuen and Chan (1983), and Alexander (1987) to suppress the LO noise and to obtain signal power gain. Without LO noise and ignored thermal noise, the performance of balanced receiver should be the same as that of a single-branch receiver.

### 1.3 Quadrature Receiver

Figure 3.4 shows a quadrature receiver to recover both the in- and quadrature-phase components of the optical signal. The quadrature re-
A quadrature homodyne receiver.

A receiver based on a 90° optical hybrid and two balanced receivers. Although a single-branch receiver of Fig. 3.1 can be used in Fig. 3.4, a balanced receiver has better performance, especially in the presence of LO laser noise.

The 90° optical hybrid composites of a 3-dB coupler and two polarization beam splitters (PBS). Using the two PBS as the reference polarization, the received signal must be controlled to be linearly polarized with a direction 45° from the PBS reference polarization. The received signal excluding noise is

$$E_r = \frac{1}{\sqrt{2}}(x + y)A_s(t)e^{j\phi_s(t) + j\omega_c t}.$$  \hspace{1cm} (3.45)

The LO laser must be circular polarized with an electric field of

$$E_{LO} = \frac{1}{\sqrt{2}}(x + e^{j\pi/2}y)A_L e^{j\omega_{LO} t}.$$  \hspace{1cm} (3.46)

At the output of the 3-dB coupler, the two electric fields before the PBS are

$$E_1 = \frac{1}{\sqrt{2}}(E_r + E_{LO})$$
$$= \frac{x}{2} \left[ A_s(t)e^{j\phi_s(t) + j\omega_c t} + A_L e^{j\omega_{LO} t} \right]_{0^\circ}$$
$$+ \frac{y}{2} \left[ A_s(t)e^{j\phi_s(t) + j\omega_c t} + A_L e^{j\omega_{LO} t + j\pi/2} \right]_{90^\circ},$$  \hspace{1cm} (3.47)
and

\[ E_2 = \frac{1}{\sqrt{2}} (E_r - E_{LO}) = \frac{x}{2} A_s(t) e^{j\phi_s(t) + j\omega_{c} t} - A_L e^{j\omega_{LO} t} \]

\[ + \frac{y}{2} A_s(t) e^{j\phi_s(t) + j\omega_{c} t} - A_L e^{j\omega_{LO} t + j\pi/2}. \tag{3.48} \]

The two PBS separate the above two electric fields to the polarization directions of \( x \) and \( y \). The upper balanced receiver combines the photocurrents corresponding to both 0° and 180°, the received photocurrent is

\[ i_I(t) = \frac{R}{4} \left| A_s(t) e^{j\phi_s(t) + j\omega_{c} t} + A_L e^{j\omega_{LO} t} \right|^2 \]

\[ - \frac{R}{4} \left| A_s(t) e^{j\phi_s(t) + j\omega_{c} t} - A_L e^{j\omega_{LO} t} \right|^2 = R A_s(t) A_L \cos[\omega_{IF} t + \phi_s(t)]. \tag{3.49} \]

The lower balanced receiver combines the photocurrents corresponding to both 90° and 270°, the received photocurrent is

\[ i_Q(t) = \frac{R}{4} \left| A_s(t) e^{j\phi_s(t) + j\omega_{c} t} + A_L e^{j\omega_{LO} t + j\pi/2} \right|^2 \]

\[ - \frac{R}{4} \left| A_s(t) e^{j\phi_s(t) + j\omega_{c} t} - A_L e^{j\omega_{LO} t + j\pi/2} \right|^2 = R A_s(t) A_L \sin[\omega_{IF} t + \phi_s(t)]. \tag{3.50} \]

In both photocurrents of Eqs. (3.49) and (3.50), an nonzero IF frequency of \( \omega_{IF} \) is assumed for a general quadrature receiver. Homodyne quadrature receiver has \( \omega_{IF} = 0 \).

Mathematically, a \( 2 \times 2 \) 90° optical hybrid has an input and output relationship of

\[ S = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & j \end{bmatrix}. \tag{3.51} \]

and a \( 2 \times 4 \) 90° optical hybrid has an input and output relationship of

\[ S = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & j & -1 & -j \end{bmatrix}. \tag{3.52} \]
For a homodyne receiver with a PLL of Fig. 3.4, $\omega_{IF} = 0$ and $i_I(t)$ gives the in-phase component of $A_s(t) \cos \phi_s(t)$, and $i_Q(t)$ provides the quadrature-phase component of $A_s(t) \sin \phi_s(t)$. The quadrature homodyne receiver can be used as the receiver of both $M$-ary PSK and quadrature-amplitude modulation (QAM) signals.

In the homodyne quadrature receiver of Fig. 3.4, the shot-noise limited SNR is the same as that in Eq. (3.35) and the amplifier-noise limited SNR is that of Eq. (3.40). The noises in the in- and quadrature-phase components are independent of each other. The SNR for the photocurrents of Eqs. (3.49) and (3.50) is identical.

For a heterodyne quadrature receiver with $\omega_{IF} \neq 0$, $i_I(t)$ and $i_Q(t)$ are quadrature components with a phase difference of $90^\circ$. With the condition of Eq. (3.12), the amplifier-noise limited SNR is that of Eq. (3.40). The shot-noise limited SNR is that of Eq. (3.34).

In heterodyne receiver, it may be more convenient to use electrical PLL to obtain both in- and quadrature-phase components. In another application as shown later, a heterodyne quadrature receiver similar to Fig. 3.4 can be used to provide image-rejection.

In- and quadrature-phase detection was first used to simultaneously measure phase and amplitude of an optical electric field (Walker and Carroll, 1984). The first application to coherent optical communications was by Hodgkinson et al. (1985). The $90^\circ$ optical hybrid of Fig. 3.4 is designed according to Kazovsky et al. (1987). Other implementations of $90^\circ$ optical hybrid were proposed by Delavaux and Riggs (1990), Delavaux et al. (1990), Hoffman et al. (1989), and Langenhorst et al. (1991). Recently, Cho et al. (2004c) shown an integrated LiNbO$_3$ optical $90^\circ$ hybrid. As shown later, both image-rejection and phase-diversity receivers are similar to the quadrature receiver of Fig. 3.4. Recently, without phase-locking, quadrature receivers similar to Fig. 3.4 were used for measurement purpose (Dorrer et al., 2003, 2005).

### 1.4 Image-Rejection Heterodyne Receiver

In a densely space coherent wavelength-division-multiplexed (WDM) systems, the most important issue is to improve the spectral efficiency by allotting more channels within the amplifier passband of the system. In the heterodyne receiver with $\omega_{IF} \neq 0$ in both Figs. 3.1 and 3.3, the signals at both $\omega_{c1}$ and $\omega_{c2}$ fall to the same IF band if $\omega_{IF} = \omega_{c1} - \omega_{LO} = \omega_{LO} - \omega_{c2}$. In a regular balanced heterodyne receiver of Fig. 3.3, a large guard-band of about $2\omega_{IF}$ is required such that the real- and image-band signals do not interfere with each other. The requirement of large guard-band limits the spectral efficiency of the system.
Figure 3.5 shows an image-rejection receiver with an optical front-end similar to the quadrature receiver of Fig. 3.4 with an 90° optical hybrid and two balanced receivers. The image-rejection receiver of Fig. 3.5 is for heterodyne receiver with $\omega_{IF} > 0$. Assume that there are two signals at $\omega_{c1}$ and $\omega_{c2}$ with a relationship of

$$\omega_{IF} = \omega_{c1} - \omega_{LO} = \omega_{LO} - \omega_{c2}.$$  \hfill (3.53)

If the two signals are

$$E_{c1}(t) = A_{s1}(t)e^{j\omega_{c1}t + j\phi_{s1}(t)}, \quad \text{and} \quad E_{c2}(t) = A_{s2}(t)e^{j\omega_{c2}t + j\phi_{s2}(t)},$$  \hfill (3.54)

similar to Eqs. (3.49) and (3.50) and without going into details, we obtain two photocurrents of

$$i_I(t) = RA_{s1}(t)A_{LO} \cos[(\omega_{c1} - \omega_{LO})t + \phi_{s1}(t)] + RA_{s2}(t)A_{LO} \cos[(\omega_{LO} - \omega_{c2})t - \phi_{s2}(t)],$$  \hfill (3.55)

and

$$i_Q(t) = RA_{s1}(t)A_{LO} \sin[(\omega_{c1} - \omega_{LO})t + \phi_{s1}(t)]$$
$$-RA_{s2}(t)A_{LO} \sin[(\omega_{LO} - \omega_{c2})t - \phi_{s2}(t)].$$  \hfill (3.56)

After a 90° microwave hybrid\(^2\), we obtain

$$i_{Re}(t) = \sqrt{2}RA_{s1}(t)A_{LO} \cos[(\omega_{c1} - \omega_{LO})t + \phi_{s1}(t)],$$  \hfill (3.57)

$$i_{Im}(t) = \sqrt{2}RA_{s2}(t)A_{LO} \cos[(\omega_{LO} - \omega_{c2})t - \phi_{s2}(t)].$$  \hfill (3.58)

\(^2\)The transfer matrix must be

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}$$

with a 90° shift from Eq. (3.1). The transfer matrix is called 90° microwave hybrid but similar matrix is for 180° optical hybrid.
for the real and image frequency band, respectively.

In an image-rejection heterodyne receiver of Fig. 3.5, even without an optical filter to filter-out the image-band amplified spontaneous emission, the amplifier-noise limited SNR is $\rho_s = N_s/n_{eq}$ from Eq. (3.40). Not only reject the signal from the image band, the image-rejection receiver also rejects the noise from the image band.

In another application of the image-rejection receiver of Fig. 3.5, the LO laser can have a frequency between two adjacent WDM channels. The real and image bands can be processed to receive two channels having larger or smaller frequency than the LO laser.

Used by Chikama et al. (1990a) and Lachs et al. (1990) for WDM systems, heterodyne image-rejection receiver was proposed by Darcie and Glance (1986), Glance (1986b), and Westphah and Strebel (1988). Even without optical filter, as indicated in Glance et al. (1988), Walker et al. (1990) and Jørgensen et al. (1992) showed that the SNR is improved by 3 dB using image-rejection receiver. The output SNR of image-rejection receiver is the same as that of homodyne receiver.

1.5 SNR of Basic Coherent Receivers

Table 3.1 summaries the SNR of system with different receiver structures. Single-branch receiver is more likely to be limited by the noise from LO laser.

For system limited by amplifier noise, without optical filter, heterodyne receiver is 3-dB worse than homodyne receiver. Image-rejection receiver eliminates the effects of the amplifier noise from the image frequency band even for the system without optical filtering. Homodyne and heterodyne receivers have the same SNR for both cases of having image-rejection or optical filtering. The optical filter of a heterodyne receiver must have a bandwidth conformed to the relationship of Eq. (3.12).

For system limited by shot-noise, the performance of heterodyne receiver is always 3-dB worse than homodyne receiver. This 3-dB difference was given in all standard textbooks (Agrawal, 2002, Betti et al., 1995, Okoshi and Kikuchi, 1988).

In later parts of this book, system performance is analyzed based on the representation of a received signal by

$$r(t) = A_s(t) \cos[\omega_{IF}t + \phi_s(t)] + n(t)$$

(3.59)

with the SNR from Table 3.1. The noise in the receiver is considered to be within the narrow receiver bandwidth of the receiver and with a band-pass representation of

$$n(t) = n_1(t) \cos \omega_{IF}t - n_2(t) \sin \omega_{IF}t$$

(3.60)
Table 3.1. Comparison of the SNR of Different Receiver Structures.

<table>
<thead>
<tr>
<th>Receiver Types</th>
<th>Limited Noise Sources</th>
<th>Shot Noise</th>
<th>Amplifier Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-branch heterodyne w/ optical filter†</td>
<td>ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
<tr>
<td>Single-branch heterodyne w/o optical filter‡</td>
<td>ηNs</td>
<td>N_s</td>
<td>2n_eq</td>
</tr>
<tr>
<td>Single-branch homodyne†</td>
<td>2ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
<tr>
<td>Balanced received heterodyne w/ optical filter</td>
<td>ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
<tr>
<td>Balanced received heterodyne w/o optical filter</td>
<td>ηNs</td>
<td>N_s</td>
<td>2n_eq</td>
</tr>
<tr>
<td>Balanced received homodyne</td>
<td>2ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
<tr>
<td>Quadrature homodyne</td>
<td>2ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
<tr>
<td>Image-rejection heterodyne receiver</td>
<td>ηNs</td>
<td>N_s</td>
<td>n_eq</td>
</tr>
</tbody>
</table>

†Single-branch receiver is more likely to be limited by LO noise.

with

\[ E\{n^2(t)\} = E\{n_{1}^2(t)\} = E\{n_{2}^2(t)\} = \sigma_{n}^2. \quad (3.61) \]

As discussed earlier, a quantum-limited system has the limit of a quantum efficiency of η = 1 or equivalent spontaneous emission of n_eq = 1. The quantum-limited SNR is equal to the average number of photons per bit (or per symbol) for heterodyne system limited by either shot or amplifier noise.

2. Performance of Synchronous Receivers

When an optical PLL is used to track the phase of the LO laser for homodyne systems or an electrical PLL is used to track the phase of an IF oscillator for heterodyne systems, the system is called a coherent system according to the terminology of conventional digital communications (Proakis, 2000). In coherent optical communications, system with phase tracking is called synchronous detection system. This section calculates the error probability of synchronous detection systems as a function of SNR. With the same SNR of \( \rho_s \), homodyne and heterodyne systems have
the same performance. This section considers only heterodyne systems with balanced receiver.

**2.1 Amplitude-Shift Keying**

When the optical carrier is ASK modulated, the signal current of Eq. (3.24) in a heterodyne receiver can be expressed as

\[
\begin{align*}
    s_1(t) &= A \cos \omega_{IF} t, \\
    s_2(t) &= 0.
\end{align*}
\]  

(3.62) (3.63)

The above binary ASK signal can be received by the heterodyne receiver of Fig. 1.3(b).

Including noise, the overall received signal is

\[
r(t) = s(t) + n(t) = \begin{cases} [A + n_1(t)] \cos \omega_{IF} t - n_2(t) \sin \omega_{IF} t & \text{for } s_1(t) \\ n_1(t) \cos \omega_{IF} t - n_2(t) \sin \omega_{IF} t & \text{for } s_2(t) \end{cases}.
\]

(3.64)

At the output of an ASK receiver with PLL, the decision variable is \( r_d(t) = A + n_1(t) \) and \( n_1(t) \) for \( s_1(t) \) and \( s_2(t) \), respectively. With decision threshold of \( \frac{1}{2} A \), the error probability is

\[
p_e = \frac{1}{2} \int_{-\infty}^{A/2} p_1(x) dx + \frac{1}{2} \int_{A/2}^{+\infty} p_2(x) dx,
\]

(3.65)

where

\[
\begin{align*}
p_1(x) &= \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-A)^2}{2\sigma_n^2}}, \\
p_2(x) &= \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{x^2}{2\sigma_n^2}}
\end{align*}
\]

(3.66) (3.67)

are the Gaussian probability density function (p.d.f.) of the decision random variables.

We obtain

\[
p_e = \frac{1}{2} \text{erfc} \left( \frac{A}{2\sqrt{2}\sigma_n} \right),
\]

(3.68)

\[
= \frac{1}{2} \text{erfc} \left( \frac{\rho_s}{2} \right),
\]

(3.69)

where the SNR is \( \rho_s = \frac{A^2 / 4}{\sigma_n^2} \) is the complementary error function, and

\[
\rho_s = \frac{A^2 / 4}{\sigma_n^2}.
\]

(3.70)
with signal power equal to $\frac{1}{2} \frac{A^2}{2} + 0 = \frac{A^2}{4}$.

To achieve an error probability of $10^{-9}$, a SNR of $\rho_s = 36$ (15.6 dB) is required. The quantum-limited binary ASK signal requires 36 photons/bit.

### 2.2 Phase-Shift Keying

When the light is PSK modulated, the signal of Eq. (3.24) in a heterodyne receiver can be expressed as

$$s_1(t) = A \cos \omega_{IF} t,$$

$$s_2(t) = -A \cos \omega_{IF} t. \quad (3.71)$$

The above binary PSK signal can be received by the heterodyne receiver of Fig. 1.3(b). Including noise, the overall received signal is

$$r(t) = s(t) + n(t) = \left\{ \begin{array}{ll}
[A + n_1(t)] \cos \omega_{IF} t - n_2(t) \sin \omega_{IF} t & \text{for } s_1(t) \\
[-A + n_1(t)] \cos \omega_{IF} t - n_2(t) \sin \omega_{IF} t & \text{for } s_2(t)
\end{array} \right. \quad (3.73)$$

At the output of the PLL, the decision random variable is $r_d(t) = \pm A + n_1(t)$. With a decision threshold of zero, the error probability is

$$p_e = \frac{1}{2} \int_{-\infty}^{0} p_1(x) dx + \frac{1}{2} \int_{0}^{+\infty} p_2(x) dx, \quad (3.74)$$

where

$$p_1(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-A)^2}{2\sigma_n^2}}, \quad (3.75)$$

$$p_2(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x+A)^2}{2\sigma_n^2}} \quad (3.76)$$

are the Gaussian p.d.f. of the decision random random variables.

We obtain

$$p_e = \frac{1}{2} \text{erfc} \left( \frac{A}{\sqrt{2}\sigma_n} \right) \quad (3.77)$$

$$= \frac{1}{2} \text{erfc} (\sqrt{\rho_s}), \quad (3.78)$$

where the SNR is

$$\rho_s = \frac{A^2/2}{\sigma_n^2}. \quad (3.79)$$

To achieve an error probability of $10^{-9}$, a SNR of $\rho_s = 18$ (12.5 dB) is required for an improvement of 3-dB over ASK signal. The quantum-limited binary PSK signal requires 18 photons/bit.
Synchronous receivers had been demonstrated mostly for PSK signal for its superior performance (Kahn et al., 1990, Kazovsky and Atlas, 1990, Kazovsky et al., 1990, Norimatsu et al., 1990, Schöpflin et al., 1990, Watanabe et al., 1989). There were other experiments to transmit some reference carriers without phase locking (Cheng and Okoshi, 1989, Wandernoth, 1992, Watanabe et al., 1992). There is recently interest to conduct PSK experiment (Cho et al., 2004c, Taylor, 2004).

2.3 Frequency-Shift Keying

When the optical carrier is frequency-shift keying (FSK) modulated, the signal of a heterodyne receiver can be expressed as

\[ s_1(t) = A \cos \omega_1 t, \]
\[ s_2(t) = A \cos \omega_2 t, \]

where \( \omega_1 \) and \( \omega_2 \) are two angular frequencies with orthogonal condition of

\[ \int_0^T s_1(t)s_2(t)dt = 0, \]

where \( T \) is the symbol interval. The two frequencies should be separated by \( \omega_1 - \omega_2 = k\pi/T, \) \( k = \pm 1, \pm 2, \ldots \), for the orthogonal condition. In a synchronous heterodyne FSK receiver of Fig. 3.6, two electrical PLL are required to phase lock the two oscillators with frequencies of either \( \omega_1 \) or \( \omega_2 \) for the two FSK signals. Equivalently, two matched filters are used with filter response matching to \( s_1(t) \) and \( s_2(t) \), respectively. The difference of the two outputs of Fig. 3.6 decides whether \( s_1(t) \) or \( s_2(t) \) is transmitted.
With the orthogonal condition, the overall received signal including noise is

\[
r(t) = s(t) + n(t) = \begin{cases} 
[A + n_{11}(t)] \cos \omega_1 t - n_{12}(t) \sin \omega_1 t & \text{for } s_1(t) \\
[A + n_{21}(t)] \cos \omega_2 t - n_{22}(t) \sin \omega_2 t & \text{for } s_2(t)
\end{cases}
\]

(3.83)

where \(n_{11}(t), n_{21}(t), n_{12}(t), \) and \(n_{22}(t)\) are independent of each other with the same variance of \(\sigma_n^2\). When \(s_1(t)\) is transmitted, the correct decision is \(A + n_{11}(t) > n_{21}(t)\) or \(A + n_{11}(t) - n_{21}(t) > 0\). The noise of \(n_{11}(t) - n_{21}(t)\) has a variance of \(2\sigma_n^2\). The error probability is the same as that of PSK signal but with a noise variance of \(2\sigma_n^2\), we obtain

\[
p_e = \frac{1}{2} \text{erfc} \left( \frac{A}{2\sigma_n} \right) \quad \text{(3.84)}
\]

\[
= \frac{1}{2} \text{erfc} \left( \sqrt{\frac{\rho_s}{2}} \right), \quad \text{(3.85)}
\]

where the SNR is

\[
\rho_s = \frac{A^2/2}{\sigma_n^2}. \quad \text{(3.86)}
\]

A FSK signal has the same performance as ASK signal. The same as ASK signal, to achieve an error probability of \(10^{-9}\), a SNR of \(\rho_s = 36\) (15.6 dB) is required. The quantum-limited binary FSK signal requires 36 photons/bit.

The performance of synchronous receiver, the same as digital signal with a matched filter provided by phase tracking, is analyzed in Proakis (2000) or the early paper of Yamamoto (1980).

3. **Performance of Asynchronous Receivers**

All of ASK, DPSK, and FSK signals can be detected without phase tracking. The detection is based on the comparison of the power or envelope of the signal. This type of detection is called noncoherent detection in conventional digital communications (Proakis, 2000) and asynchronous receiver in coherent optical communications. In this section, we consider asynchronous heterodyne receiver with signal processing by electrical circuits in the IF.

### 3.1 Envelope Detection of Heterodyne ASK Signal

Figure 3.7 shows the receiver for envelope detection of heterodyne binary ASK signal. The signal first passes through a band-pass filter (BPF) to limit the amount of noise. After the BPF, the signal is the
same as that of Eq. (3.64). The signal of Eq. (3.64) is squared, low-pass filtered (LPF), square-rooted, to obtain the envelope of

\[ r_d(t) = \begin{cases} \frac{[A + n_1(t)]^2 + n_2(t)}{\sqrt{n_1^2(t) + n_2^2(t)}} & \text{for } s_1(t) \\ \frac{[A + n_1(t)]^2 + n_2(t)}{\sqrt{n_1^2(t) + n_2^2(t)}} & \text{for } s_2(t) \end{cases} \]

(3.87)

with p.d.f. of

\[
p_1(r) = \frac{r}{\sigma_n^2} I_0 \left( \frac{r A}{\sigma_n^2} \right) \exp \left( -\frac{r^2 + A^2}{2\sigma_n^2} \right),
\]

(3.88)

\[
p_2(r) = \frac{r}{\sigma_n^2} \exp \left( -\frac{r^2}{2\sigma_n^2} \right),
\]

(3.89)
as Rice and Rayleigh distribution, respectively, where \( I_0(\cdot) \) is the zeroth-order modified Bessel function of the first kind. In the decision random variable of Eq. (3.87), some constant factors related to gain and loss of the squarer, LPF, and the square-root components are ignored. The envelope of Eq. (3.87) is the same as the amplitude of the signal. The error probability of the signal is

\[
p_e = \frac{1}{2} \left[ 1 - \int_{r_{\text{th}}}^{\infty} p_2(r) \, dr \right] + \frac{1}{2} \int_{r_{\text{th}}}^{\infty} p_1(r) \, dr,
\]

(3.90)

where \( r_{\text{th}} \) is the threshold. Using the Marcum’s Q-function defined by Eq. (3.A.3) from Appendix 3.A, we obtain

\[
p_e = \frac{1}{2} \left[ 1 - Q \left( \frac{A}{\sigma_n}, \frac{r_{\text{th}}}{\sigma_n} \right) \right] + \frac{1}{2} \exp \left( -\frac{r_{\text{th}}^2}{2\sigma_n^2} \right).
\]

(3.91)

The optimal threshold can be found by \( dp_e/dr_{\text{th}} = 0 \) as

\[
I_0 \left( \frac{Ar_{\text{th}}}{\sigma_n^2} \right) \exp \left( -\frac{A^2}{2\sigma_n^2} \right) = 1.
\]

(3.92)
The optimal decision threshold is difficult to find analytically, we may use the approximation of \( r_{\text{th}} = A/2 \). With this threshold, the second
term of Eq. (3.91) is larger than the first term and the error probability is approximately equal to

\[ p_e \approx \frac{1}{2} \exp \left( -\frac{A^2}{8\sigma_n^2} \right) \]
\[ = \frac{1}{2} \exp \left( -\frac{\rho_s}{2} \right). \]  
(3.93)

Based on the approximation of Eq. (3.93), the required SNR for an error probability of $10^{-9}$ is $\rho_s = 40$ (16 dB). The quantum limit is 40 photons/bit. The asynchronous receiver is about 0.4 dB worse than the synchronous receiver in Sec. 3.2.1.

### 3.2 Dual-Filter Detection of FSK Signal

Figure 3.8 shows an asynchronous heterodyne receiver for FSK signal based on two BPF matched to the FSK signals of $s_1(t)$ and $s_2(t)$ with center angular frequencies of $\omega_1$ and $\omega_2$, respectively. Based on Eq. (3.83), when $s_1(t)$ is transmitted, the received signal at the first filter centered at $\omega_1$ is the same as that of Eq. (3.87). Although not necessary, a square root is assumed for the signal of Fig. 3.8 for convenience. When $s_1(t)$ is transmitted, we obtain

\[ r_1(t) = \sqrt{(A + n_{11}(t))^2 + n_{12}(t)^2}, \]

and

\[ r_2(t) = \sqrt{n_{21}(t)^2 + n_{22}(t)^2} \]

(3.95)

with p.d.f. of

\[ p_1(r_1) = \frac{r_1}{\sigma_n^2} J_0 \left( \frac{r_1 A}{\sigma_n^2} \right) \exp \left( -\frac{r_1^2 + A^2}{2\sigma_n^2} \right), \]

(3.96)

\[ p_2(r_2) = \frac{r_2}{\sigma_n^2} \exp \left( -\frac{r_2^2}{2\sigma_n^2} \right) \]

(3.97)

as Ricc and Rayleigh distribution, respectively.
Bit error occurs when $r_1 < r_2$, or

$$p_e = \Pr \{ r_1 < r_2 \} = \Pr \{ r_1^2 < r_2^2 \}, \quad (3.98)$$

because both $r_1 \geq 0$ and $r_2 \geq 0$. From Eq. (3.98), the square-root component in the receiver of Fig. 3.8 is optional. The error probability is

$$p_e = \int_0^\infty p_1(r_1) \int_r^\infty p_2(r_2) dr_2 dr_1$$

$$= \int_0^\infty \frac{r_1}{\sigma_n^2} I_0 \left( \frac{r_1 A}{\sigma_n^2} \right) \exp \left( -\frac{2r_1^2 + A^2}{2\sigma_n^2} \right) dr_1$$

$$= \frac{1}{2} \exp \left( -\frac{A^2}{4\sigma_n^2} \right) \int_0^\infty \frac{2r_1}{\sigma_n^2} I_0 \left( \frac{2r_1 A}{2\sigma_n^2} \right) \exp \left( -\frac{4r_1^2 + A^2}{4\sigma_n^2} \right) dr_1$$

$$= \frac{1}{2} \exp \left( -\frac{\rho_s}{2} \right). \quad (3.99)$$

The above error probability is valid only if the two signals are orthogonal. The performance of FSK signal is similar to that of ASK signal. The required SNR for an error probability of $10^{-9}$ is $\rho_s = 40 \,(16 \,\text{dB})$. The quantum limit is 40 photons/bit.

A FSK signal can also be detected based on a single filter. The signal after the filter is the same as an ASK signal. The performance of FSK signal with a single filter is the same as that of ASK signal with an error probability of $p_e = \frac{1}{2} \exp \left( -\frac{\rho_s}{4} \right)$. While an ASK signal has no power at the “0” level, the power of FSK signal is the same at both “0” and “1” levels. The performance of single-filter detected FSK signal is 3-dB worse than the equivalent ASK signal. Intuitively, using a single filter, the FSK signal is also 3-dB worse than a dual-filter receiver.

Single filter FSK experiment was conducted by Emura et al. (1984) and Park et al. (1990), and dual-filter FSK experiment was conducted by Emura et al. (1990b).

### 3.3 Heterodyne Differential Detection of DPSK Signal

In another representation for a heterodyne DPSK system, the received signal is

$$r(t) = \Re \left\{ \left[ A e^{j\phi_t} + n(t) \right] e^{j\omega_{dp} t} \right\}, \quad (3.100)$$

or

$$r(t) = \Re \left\{ \tilde{r}(t) e^{j\omega_{dp} t} \right\}, \quad (3.101)$$
where \( \tilde{r}(t) = A e^{i\phi_s(t)} + n(t) \) is the baseband representation of the IF signal. The DPSK signal is demodulated by delay-and-multiplier circuits of Fig. 1.4(b). If \( \tilde{r}(t) \) is expressed by the polar representation of \( \tilde{r}(t) = \tilde{A}_1(t)e^{i\tilde{\phi}_1(t)} \) with \( \tilde{\phi}_1(t) \) include the noisy phase from the contribution of \( n(t) \), the demodulated signal after a low-pass filter is

\[
\begin{align*}
  r_d(t) &= \tilde{A}_1(t)\tilde{A}_1(t - T) \cos \left[ \omega_{IF}\tau + \tilde{\phi}_1(t) - \tilde{\phi}_1(t - T) \right] \\
       &= \Re \{ \tilde{r}(t)\tilde{r}^*(t - T) \}. \tag{3.102}
\end{align*}
\]

In Eq. (3.102), two symbols of \( \tau \) and \( T \) are used to represent the delay of the delay-and-multiplier circuits. The delay of \( \tau \) must be approximated equal to \( T \) and \( \omega_{IF}\tau \) must be an integer multiple of \( 2\pi \) for a noiseless decision variable of \( r_d(t) = \pm A^2 \) with \( \phi_s(t) - \phi_s(t - T) = 0, \pi \), respectively. For a large frequency of \( \omega_{IF} \), a small variation of \( \tau \) around the time interval \( T \) may give \( \omega_{IF}\tau \) equal to an integer multiple of \( 2\pi \). Later in this book assumes that \( \tau = T \) and \( \omega_{IF}T \) is equal to an integer multiple of \( 2\pi \) at the same time.

Some simple algebra gives

\[
r_d(t) = \left| \frac{\tilde{r}(t) + \tilde{r}(t - T)}{2} \right|^2 - \left| \frac{\tilde{r}(t) - \tilde{r}(t - T)}{2} \right|^2. \tag{3.103}
\]

The signals of \( \tilde{r}(t) + \tilde{r}(t - T) \) and \( \tilde{r}(t) - \tilde{r}(t - T) \) are independent of each other. With \( r_1 = |\tilde{r}(t) + \tilde{r}(t - T)|/2 \) and \( r_2 = |\tilde{r}(t) + \tilde{r}(t - T)|/2 \), assume that \( \phi_s(t) = \phi_s(t - T) \), the error probability is similar to Eq. (3.98) with

\[
p_e = \Pr\{r_1 < r_2\} = \Pr\{r_1^2 < r_2^2\} = \Pr\{r_d < 0\}. \tag{3.104}
\]

If \( \phi_s(t) = \phi_s(t - T) \), \( r_1^2 = |A + n(t)/2 + n(t - T)/2|^2 \) as the first term of Eq. (3.103) is the square of a Gaussian random variable with mean of \( A \) and variance of \( \sigma_n^2 \) and \( r_2^2 = |n(t) - n(t - T)|^2/4 \) as the second term of Eq. (3.103) is the square of a Gaussian random variable with zero mean and variance of \( \sigma_n^2 \). The error probability is the same as that for FSK signal of Eq. (3.85) with equivalent SNR of \( A^2/\sigma_n^2 \), we obtain

\[
p_e = \frac{1}{2} \exp \left( -\frac{A^2}{2\sigma_n^2} \right) = \frac{1}{2} \exp (-\rho_s). \tag{3.105}
\]

The required SNR for an error probability of \( 10^{-9} \) is \( \rho_s = 20 \) (13 dB) for DPSK signal. The quantum limit is 20 photons/bit. DPSK signal is about 3-dB better than ASK signal.
The error probability of Eq. (3.105) was derived by Cahn (1959) using approximation. DPSK signal was analyzed the same as orthogonal or FSK signal in Stein (1964). Using two time intervals, the performance of DPSK signal is 3-dB better than that of FSK signal. In Eq. (3.103), although the mean of the signal is the same as that of the FSK signal, the noise variance is reduced by half because the noise is the average over two time intervals.

Heterodyne DPSK signal was demonstrated in Chikama et al. (1988), Creancr et al. (1988), Meissner (1989), Naito et al. (1990), and Gnauck et al. (1990).

3.4 Heterodyne Receiver for CPFSK Signal

In binary continuous-phase frequency-shift keying (CPFSK) transmission, the signal in each time interval is equal to

$$s(t) = A \cos(\omega_{IF}t \pm \pi \Delta f t + \phi_0),$$

(3.106)

where $\phi_0$ depends the phase of previous symbols to ensure continuous phase operation and $\Delta f$ is the frequency deviation between the “0” and “1” states. With a received signal of $r(t) = s(t) + n(t)$, we may define

$$\tilde{r}(t) = Ae^{\pm j \pi \Delta f t + j \phi_0} + n(t) = A_1(t)e^{j\tilde{\phi}_1(t)}$$

similar to that for DPSK signal. The CPFSK signal can be demodulated using the asynchronous receiver of Fig. 3.9 based on the delay-and-multiplier circuits similar to that of Fig. 1.4(b) for DPSK signals. In the CPFSK receiver of Fig. 3.9, the delay is $\tau$ instead of the one-bit delay of $T$ in Fig. 1.4(b). Similar to the case of DPSK signal of Eq. (3.103), the decision random variable is

$$r_d = A_1(t)A_1(t-\tau) \cos[\omega_{IF} \tau + \tilde{\phi}_1(t) - \tilde{\phi}_1(t-\tau)]$$

$$= \Re\{\tilde{r}(t)\tilde{r}^*(t-\tau)e^{j\omega_{IF}\tau}\},$$

(3.107)

where $\tilde{\phi}_1(t) = \pm \pi \Delta f t + \phi_0 + \phi_n(t)$ with $\phi_n(t)$ as the noisy phase from $n(t)$. The differential phase is $\phi_1(t) - \phi_1(t-\tau) = \pm \pi \Delta f \tau + \phi_n(t) - \phi_n(t-\tau)$. The receiver achieves its optimal performance for a design of $\omega_{IF} \tau =$
\[ k \pi + \frac{1}{2} \pi \] with \( k \) as an integer. Excluding noise, the decision variable is \( r_d(t) = \pm A^2 \sin(\Delta f \pi \tau) \). For the best system efficiency, \( r_d(t) = \pm A^2 \) if \( \Delta f \tau = 1/2 \).

With the design of \( \Delta f \tau = 1/2 \), assume that the phase of noise of \( \phi_n(t) \) and \( \phi_n(t - \tau) \) are independent of each other, the performance of CPFSK signal with differential detection is the same as DPSK signal with error probability of

\[ p_e = \frac{1}{2} \exp(-\rho_s). \] (3.108)

The most interesting case is \( \Delta f = 1/2T \) when \( \tau = T \). The frequency separation of \( \Delta f = 1/2T \) is the minimum frequency separation for two orthogonal frequencies of Eq. (3.82). The CPFSK signal with \( \Delta f = 1/2T \) is called minimum shift keying (MSK) modulation. Similar to DPSK signal, MSK signal may be demodulated using a one-bit delay and multiplier. The performance of MSK signal is also the same as DPSK signal. From Proakis (2000), MSK signal has more compact spectrum than DPSK signal.

Because CPFSK signal can be generated directly modulated a semiconductor laser from Sec. 2.4, it was the most popular scheme for coherent optical communications (Emura et al., 1990a, Iwashita and Matsumoto, 1987, Iwashita and Takachio, 1988, Park et al., 1991, Takachio et al., 1989).

### 3.5 Frequency Discriminator for FSK Signal

FSK signal can also be detected asynchronously by frequency discriminator of Fig. 3.10. The frequency discriminator may be the most popular receiver for analog frequency modulation (FM) signal for its simplicity. The system must have high SNR to ensure the correct operation of the frequency discriminator.

The frequency discriminator of Fig. 3.10 consists of two band-pass filters having frequency responses of

\[ H_1(f) = \begin{cases} 2\pi K(f - f_{IF}) & f_{IF} < f < f + \Delta f/2 \\ 0 & \text{otherwise} \end{cases} \] (3.109)

and

\[ H_2(f) = \begin{cases} 2\pi K(f_{IF} - f) & f - \Delta f/2 < f < f_{IF} \\ 0 & \text{otherwise} \end{cases} \] (3.110)

where \( K \) is the slope of the frequency discriminator and \( \Delta f \) is the bandwidth of the frequency discriminator and the frequency separation of
two FSK signals of $2\pi \Delta f = \omega_1 - \omega_2$. In time domain, the operation of the frequency discriminator is equivalent to a linear operation of $K \frac{d}{dt}$.

Assume that $s_1(t)$ is transmitted with a signal before the discriminator as

$$A \cos \omega_1 t + n_1(t) \cos \omega_{IF} t - n_2(t) \sin \omega_{IF} t,$$

the output of $H_1(f)$ is

$$r_1(t) = AK(\omega_1 - \omega_{IF}) \sin \omega_1 t + K \frac{dn_{11}(t)}{dt} \cos (\omega_{IF} t + \pi \Delta f t/2)$$
$$- K \frac{dn_{12}(t)}{dt} \sin (\omega_{IF} t + \pi \Delta f t/2),$$

where $\omega_1 - \omega_{IF} = \pi \Delta f$, $n_{11}(t)$ and $n_{12}(t)$ are the part of $n_1(t)$ and $n_2(t)$ in the frequency between $f_{IF}$ and $f_{IF} + \Delta f / 2$. The output of $H_2(f)$ is

$$r_2(t) = K \frac{dn_{21}(t)}{dt} \cos (\omega_{IF} t + \pi \Delta f t/2) - K \frac{dn_{22}(t)}{dt} \sin (\omega_{IF} t + \pi \Delta f t/2).$$

where $n_{21}(t)$ and $n_{22}(t)$ are the part of $n_1(t)$ in the frequency smaller than the IF of $f_{IF}$. The signal of $r_1(t)$ is similar to an amplitude-modulated signal with a power of $\pi^2 [AK \Delta f]^2 / 2$. With the derivative operation as a linear filter, the noise of $Kdn_{11}(t)/dt$ has a variance of

$$K^2 E \left\{ \left| \frac{dn_{11}}{dt} \right|^2 \right\} = (2\pi K)^2 \int_0^{B_d} N_n f^2 df$$
$$= (2\pi K)^2 \frac{1}{3} B_d^3 N_n = \frac{4}{3} \pi^2 K^2 B_d^2 \sigma_n^2.$$

where $N_n$ is the spectral density of $n_{11}(t)$ and $\sigma_n^2 = N_0 B_d$. At the output of the frequency discriminator, the signal is similar to that case.
of orthogonal signals demonstrated by two BPF followed by envelope detection.

The equivalent SNR of the discriminator output signal is

$$\rho_{eq} = \frac{3 \rho_s}{8} \left( \frac{\Delta f}{B_d} \right)^2.$$  \hspace{1cm} (3.115)

Similar to the error probability of Eq. (3.99), we obtain

$$p_e = \frac{1}{2} \exp \left[ -\frac{3 \rho_s}{16} \left( \frac{\Delta f}{B_d} \right)^2 \right]. \hspace{1cm} (3.116)$$

For the case that $\Delta f = 2B_d$, we obtain $p_e = \frac{1}{2} \exp (-3\rho_s/4)$, 1.76 dB improvement over the asynchronous receiver of Eq. (3.99) but twice the spectral bandwidth. Using the method based on frequency discriminator, the system performance is improved with the frequency expansion of $\Delta f / B_d$. Similar to analog FM, detection based on frequency discriminator expands the signal bandwidth to obtain improved system performance. However, the error probability of Eq. (3.116) should be considered as an approximation. While the noise outside $B_d$ in Eq. (3.114) is ignored in the deviation, those noise gives noise-to-noise beating thought the squarer of Fig. 3.10 and degrades the performance of the system, especially at low SNR.

### 3.6 Envelope Detection of Correlated Binary Signals

FSK, DPSK, and MSK signals are special cases of signal modulation formats based on two orthogonal signals. The dual-filter detection of FSK signal and differential detection of both DPSK and MSK signals are asynchronous or noncoherent detection of two orthogonal signals. For envelope detection of binary signals, the receiver is the same as that of Fig. 3.8 with the band-pass filters representing two matched filters. If the two signals are correlated with a correlation coefficient of $|\rho|$, the p.d.f. of the outputs of two matched filters are

$$p_1(r_1) = \frac{r_1}{\sigma_n^2} I_0 \left( \frac{r_1A}{\sigma_n^2} \right) \exp \left( -\frac{r_1^2 + A^2}{2\sigma_n^2} \right), \hspace{1cm} (3.117)$$

$$p_2(r_2) = \frac{r_2}{\sigma_n^2} I_0 \left( \frac{r_2|\rho|A}{\sigma_n^2} \right) \exp \left( -\frac{r_2^2 + |\rho|^2A^2}{2\sigma_n^2} \right). \hspace{1cm} (3.118)$$

The random variables of $R_1$ and $R_2$ are correlated with each other. In order to derive the error probability of $p_e = \Pr \{ R_2 > R_1 \}$, a transform is required such that $R_2^2 - R_1^2 = \Gamma_2^2 - \Gamma_1^2$ in which the random variables of
\( \Gamma_2 \) and \( \Gamma_1 \) are independent of each other and also have Rice distribution. While the amplitude parameter of \( R_1 \) and \( R_2 \) are \( A \) and \( |\rho|A \), respectively, the amplitude parameters of \( \Gamma_1 \) and \( \Gamma_2 \) are \( A \left[ 1 \pm \sqrt{1 - |\rho|^2} \right]^{1/2} \), respectively. From Appendix 3.A, the error probability of correlated binary signal with envelope detection is

\[
p_e = Q(a, b) - \frac{1}{2} e^{-(a^2 + b^2)/2} I_0(ab),
\]

where

\[
a = \sqrt{\frac{\rho_s}{2}} \left( 1 - \sqrt{1 - |\rho|^2} \right),
\]

\[
b = \sqrt{\frac{\rho_s}{2}} \left( 1 + \sqrt{1 - |\rho|^2} \right).
\]

For orthogonal signal, \( |\rho| = 0 \), \( a = 0 \), and \( b = \sqrt{\rho_s} \), \( p_e = Q(0, b) - \frac{1}{2} e^{-b^2/2} = \frac{1}{2} e^{-\rho_s/2} \), the same as the error probability of Eq. (3.85). The error probability of Eq. (3.119) was first derived by Helstrom (1955) based on direct integration. The method to find the error probability here is based on Stein (1964) and Schwartz et al. (1966). Proakis (2000, Appendix B) also derived the error probability of Eq. (3.119). While the error probability of Eq. (3.119) is not useful if the two binary signals are well-designed without correlation, further degradation that induces correlation can be analyzed based on Eq. (3.119).

4. **Performance of Direct-Detection Receivers**

Other than the intensity-modulation/direct-detection (IMDD) systems of Fig. 1.1, both DPSK and FSK signals can be directly detected without mixing with an LO laser. DPSK and FSK signals can be detected using interferometer or optical filter. Direct-detection receiver is simpler than both homodyne and heterodyne receivers that require the mixing with an LO laser. This section analyzes the performance of typical direct-detection receivers for ASK or on-off keying, DPSK, and FSK signals.

4.1 **Intensity-Modulation/Direct-Detection Receiver**

IMDD systems of Fig. 1.1 are the simplest optical communication schemes to converse information with the presence or absence of light. The receiver is just a photodiode that converts the optical intensity to
photocurrent. For system limited by amplifier noises, the performance is analyzed in, for example, Agrawal (2002), Desurvire (1994), and Becker et al. (1999).

Quantum Limit for Systems without Amplifiers

In a system without optical amplifiers, due to the quantum nature of photons, the number of photons has a Poisson distribution. At the on-state, the probability of having \( k \) photons is

\[
p_k = \frac{N_b^k}{k!} e^{-N_b},
\]

where \( N_b \) is the number of photons in the on state of a binary on-off keying signal. If the detection is based on the presence or absence of photons, the error probability is

\[
p_e = \frac{1}{2} p_0 = \frac{1}{2} e^{-N_b}
\]

with no photon, where the factor of \( 1/2 \) is the probability of the off state. With no photon to send, there is no error for the off state. The average number of photons is \( N_s = N_b/2 \). In order to achieve an error probability of \( 10^{-9} \), an average of 10 photons/bit are required. If the quantum efficiency of \( \eta \) is less than unity, the required number of photons increases by the factor of \( \eta^{-1} \).

Practical on-off keying receivers require thousands of photons per bit, mostly due to the contribution from the thermal noise at the receiver circuitry. As an example, assume a thermal noise density of \( i_{th} = 5 \) pA/\( \sqrt{\text{Hz}} \), corresponding to a receiver sensitivity of \( -25.2 \) and \( -28.2 \) dBm for 10- and 2.5-Gb/s systems, respectively, if shot noise is ignored and a photodiode responsivity of \( R = 1 \) is assumed. For systems operating in 1.55 \( \mu \text{m} \), the number of photons per bit is \( 4.7 \times 10^4 \) and \( 2.3 \times 10^3 \) for 10- and 2.5-Gb/s systems for an error probability of \( 10^{-9} \). Without the usage of optical amplification, practical receiver always requires the number of photons per bit many order larger than the quantum limit of 10 photons/bit, even for receiver with very good sensitivity.

Amplifier-Noise Limited System

If the on-off keying system is limited by amplifier noise, the received electric field is

\[
E_r(r) = [A_s(t) + n_x(t)]e^{j\omega t}x + n_y(t)e^{j\omega t}y,
\]

where the transmitted data are contained in the amplitude of \( A_s(t) \in \{0, A\} \) for on-off keying signal. In the direct-detection receiver, the above
electric field is converted by a photodiode to photocurrent of
\[ i(t) = R|A_s(t) + n_x(t)|^2 + R|n_y(t)|^2, \]  
where shot noise is ignored by assuming that the amplifier noise is the major degradation. If the noise from orthogonal polarization of \( n_y(t) \) is ignored, the system is the same as that using envelope detection for heterodyne ASK receiver in Sec. 3.3.1.

The common factor of photodiode responsivity of \( R = 1 \) is assumed in Eq. (3.124) without loss of generality. Further assumed an optical matched filter preceding the photodiode, in the on state with \( A_s(t) = A \), we obtain
\[ r_{on}(t) = [A + n_{x1}(t)]^2 + n_{x2}^2(t) + n_{y1}^2(t) + n_{y2}^2(t) \]  
with p.d.f. of
\[ p_1(y) = \frac{1}{2}\sigma_n^2 \sqrt{y} e^{-y(A^2+y)/2\sigma_n^2} I_1 \left( \sqrt{\frac{y-A^2}{\sigma_n^2}} \right), \quad y \geq 0 \]  
as the noncentral chi-square (\( \chi^2 \)) p.d.f. with four degrees of freedom.

In the off state with \( A_s(t) = 0 \), we obtain
\[ r_{off}(t) = n_{x1}^2(t) + n_{x2}^2(t) + n_{y1}^2(t) + n_{y2}^2(t) \]  
with p.d.f. of
\[ p_2(y) = \frac{1}{4\sigma_n^2} y e^{-y/2\sigma_n^2}, \quad y \geq 0 \]  
as the \( \chi^2 \) p.d.f. with four degrees of freedom, also called Gamma distribution.

With a threshold of \( y_{th} \), the error probability of the system is
\[ p_e = \frac{1}{2} \int_0^{y_{th}} p_1(y)dy + \frac{1}{2} \int_{y_{th}}^{\infty} p_2(y)dy \]
\[ = \frac{1}{2} \left[ 1 - Q_2 \left( \frac{A}{\sigma_n}, \sqrt{\frac{y_{th}}{\sigma_n}} \right) \right] + \frac{1}{2} \exp \left( \frac{-y_{th}}{2\sigma_n^2} \right) \left( 1 + \frac{y_{th}}{2\sigma_n^2} \right), \]  
where
\[ Q_2(a, b) = Q(a, b) + \frac{b}{a} e^{-(a^2+b^2)/2} I_1(ab) \]  
is the second-order generalized Marcum Q function.

The optimal threshold can be determined by
\[ \sqrt{y_{th}} \frac{A}{2\sigma_n^2} = e^{-A^2/2\sigma_n^2} I_1 \left( \sqrt{y_{th}} \frac{A}{\sigma_n^2} \right), \quad y \geq 0. \]
In a simplified analysis, we can approximate the decision threshold as $y_{th} = A^2/4$, the error probability is approximately equal to

$$p_e \approx \frac{1}{2} \int_{y_{th}}^{\infty} p_2(y) dy$$

$$= \frac{1}{2} \exp \left( -\frac{A^2}{8\sigma_n^2} \right) \left( 1 + \frac{A^2}{8\sigma_n^2} \right)$$

$$= \frac{1}{2} \exp \left( -\frac{\rho_s}{2} \right) \left( 1 + \frac{\rho_s}{2} \right).$$

With the inclusion of the amplified noise from orthogonal polarization of $n_y(t)$, the error probability is increased by a factor of $1 + \rho_s/2$. Using a direct-detection receiver, based on the error probability of Eq. (3.132), a SNR of $\rho_s = 46.4$ (16.7 dB) is required to achieve an error probability of $10^{-9}$. The quantum-limited receiver requires 46.4 photons/bit. The inclusion of the noise from orthogonal polarization degrades the receiver by about 0.7 dB.

**Error Probability Based on Gaussian Approximation**

The error probability can be analyzed based on Gaussian approximation with the assumption of two received signals of

$$r_{on}(t) = I_1 + n_1(t),$$

$$r_{off}(t) = I_0 + n_0(t),$$

where $n_0(t)$ and $n_1(t)$ are assumed zero-mean Gaussian noise with variances of $\mathcal{E}\{n_1^2(t)\} = \sigma_1^2$ and $\mathcal{E}\{n_0^2(t)\} = \sigma_0^2$, respectively, $I_1$ and $I_0$ are the mean photocurrents for the on and off states. In general, $I_1 = RP$ and $I_0 = 0$.

The p.d.f. of the on and off states are

$$p_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-I_1)^2}{2\sigma_1^2}},$$

$$p_0(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(x-I_0)^2}{2\sigma_0^2}}.$$

The optimal decision threshold can be determined by

$$p_1(x_{th}) = p_0(x_{th}), \quad I_0 < x_{th} < I_1$$

of

$$x_{th} = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1}.$$
of \( \log(\sigma_1/\sigma_0) = 0 \), the optimal threshold is a difficult to calculate but without providing further accuracy.

Defined a \( Q \)-factor of

\[
Q = \frac{I_1 - x_{th}}{\sigma_1} = \frac{x_{th} - I_0}{\sigma_0} = \frac{I_1 - I_0}{\sigma_1 + \sigma_0},
\]

(3.139)

the error probability is equal to

\[
p_e = \frac{1}{2} \text{erfc} \left( \frac{Q}{\sqrt{2}} \right).
\]

(3.140)

With the Gaussian approximation, the noise variances are equal to

\[
\sigma_1^2 = \sigma_{A_s-n_s}^2 + \sigma_{n_s-n_s}^2 + \sigma_{sh}^2 + \sigma_{th}^2,
\]

(3.141)

\[
\sigma_0^2 = \sigma_{n_s-n_s}^2 + \sigma_{sh}^2 + \sigma_{th}^2,
\]

(3.142)

where

\[
\sigma_{A_s-n_s}^2 = 4R^2 P_r S_{sp} B_d,
\]

(3.143)

\[
\sigma_{n_s-n_s}^2 = 4R^2 S_{sp}^2 \Delta \nu_{opt} B_d,
\]

(3.144)

\[
\sigma_{sh}^2 = 2eR(P_r + 2S_{sp} \Delta \nu_{opt}) B_d,
\]

(3.145)

\[
\sigma_{th}^2 = \frac{4kT}{R_L}.
\]

(3.146)

Thermal noise is included in the above equations for the case that the received signal is very small or for system without amplifier noises. A direct-detection on-off keying system is potentially limited by thermal instead of amplifier noise.

For the specific case of the signals of Eqs. (3.125) and (3.127) with amplifier noise from orthogonal polarization, we obtain

\[
I_0 = 0, \quad I_1 = A^2,
\]

(3.147)

\[
\sigma_0^2 = 8\sigma_n^4, \quad \sigma_1^2 = 4\sigma_n^2 A^2 + 8\sigma_n^4,
\]

(3.148)

and

\[
Q = \frac{A^2}{\sqrt{2} \sigma_n^2 + 2 \sigma_n \sqrt{A^2 + 2\sigma_n^2}}
\]

\[
= \frac{\sqrt{2} \rho_s}{1 + \sqrt{2} \rho_s + 1},
\]

(3.149)

Based on the Gaussian approximation, the required SNR to achieve an error probability of \( 10^{-9} \) is \( \rho_s = 36 + 6\sqrt{2} = 44.5 \) (16.5 dB).
When the optical filter preceding the receiver has a wide bandwidth, direct-detection on-off keying system can be approximately analyzed by adding identical and independent noise terms to both Eqs. (3.125) and (3.127) (Humblet and Azizoğlu, 1991, Marcuse, 1990, 1991). The number of noise terms is equal to twice the ratio of the optical bandwidth to data rates, the factor of two taking into account the two polarizations of amplifier noise. While Marcuse (1990, 1991) and Humblet and Azizoğlu (1991) sum independent and identical random variables to model amplifier noises, Lee and Shim (1994) sums independent random variables with variance depending on the combined effects of electrical and optical filter responses. The optical filter may also model further linear effects of chromatic and polarization-mode dispersion. These two methods are widely used for performance evaluation in direct-detection on-off keying systems (Bosco et al., 2001, Chan and Conradi, 1997, Forestieri, 2000, Holzlochner et al., 2002, Roudas et al., 2002).

Comparison of Different Models

Figure 3.11 shows the error probability of ASK signal detected using various types of receiver that are also analyzed based on different assumptions. The error probability of Eq. (3.69) with synchronous receiver has the lowest error probability for the same SNR. The error probability of Eq. (3.93) for envelope detection is also shown for comparison together as the error probability calculated with the optimal threshold of Eq. (3.92) as dashed-lines. The error probability of Eq. (3.132) for direct-detection is shown as solid line with the corresponding error probability calculated with the optimal threshold of Eq. (3.131) as dashed-lines. The error probability with optimal threshold of Eq. (3.131) almost overlaps with Eq. (3.93). The Gaussian approximation of Eq. (3.140) using Q factor is shown in Fig. 3.11 as dotted-line.

From Fig. 3.11, the performance of ASK signals with envelope detection can be evaluated using Eq. (3.93). Compared with the error probability with optimal threshold, the approximation of Eq. (3.93) is just about 0.1 dB worse at the error probability of 10^-9.

For direct-detection receiver, the Gaussian approximation overestimates the error probability and gives an approximated SNR penalty of about 0.45 dB comparing with the one with optimal threshold. Because the Gaussian approximation also uses an optimal threshold according to its own model, the performance with Gaussian approximation is actually better than the error probability with sub-optimal threshold of Eq. (3.132) by 0.2 dB. In practical system, for either direct- or envelope-detection, the error probability of Eq. (3.93) can be used.
If the optical filter before the receiver is an optical matched filter, direct-detection receiver for on-off keying signal has very good receiver sensitivity as shown in both Atia and Bondurant (1999) and Caplan and Atia (2001).

4.2 Direct-Detection DPSK Receiver

Figure 3.12 redraws the direct-detection receiver for DPSK signal of Fig. 1.4(c). The DPSK receiver uses an asymmetric Mach-Zehnder interferometer in which the signal is splitted into two paths and combined after a path difference of an one-bit delay of $T$. In practice, the path difference of $\tau \approx T$ must be chosen such that $\exp(j\omega_0\tau) = 1$, where $\omega_0$ is the angular frequency of the signal. Ideally, the optical filter before the interferometer is assumed to be an optical matched filter for the transmitted signal. A balanced receiver similar to that of Fig. 3.3 is used to obtain the photocurrent. A low-pass filter reduces the receiver noise. We assume that the low-pass filter has a wide bandwidth and does not distort the received signal. With the assumption of matched filter, the analysis is applicable to both non-return-to-zero (NRZ) and RZ signals.
Figure 3.12. Direct-detection DPSK receiver using an unpolarized asymmetric Mach-Zehnder interferometer.

At the output of the unpolarized asymmetric Mach-Zehnder interferometer, the two output signals are

\[
E_1(t) = \frac{x}{2} [A e^{-j\phi_s(t)} + n_x(t)] + \frac{y}{2} n_y(t) + \frac{x}{2} [A e^{-j\phi_s(t-T)} + n_x(t-T)] + \frac{y}{2} n_y(t-T),
\]

and

\[
E_2(t) = \frac{x}{2} [A e^{-j\phi_s(t)} + n_x(t)] + \frac{y}{2} n_y(t) - \frac{x}{2} [A e^{-j\phi_s(t-T)} + n_x(t-T)] - \frac{y}{2} n_y(t-T).
\]

In the electric fields of Eqs. (3.150) and (3.151), the path difference of the interferometer is assumed exactly as the symbol time \(T\) and \(\exp(j\omega_c T) = 1\). The amplifier noises of \(n_x(t), n_y(t), n_x(t-T),\) and \(n_y(t-T)\) are independent identically distributed complex zero-mean circular Gaussian random variables. The noise variance is \(E\{|n_x(t)|^2\} = E\{|n_y(t)|^2\} = E\{|n_x(t-T)|^2\} = E\{|n_y(t-T)|^2\} = 2\sigma_n^2\), where \(\sigma_n^2\) is the noise variance per dimension. In a polarized receiver, \(n_y(t) = n_y(t-T) = 0\) and the error probability is the same as that for heterodyne DPSK system in Sec. 3.3.3.

In Eqs. (3.150) and (3.151), the loss in the interferometer is ignored. If the amplifier noise is the dominant noise source, both the interferometer loss and the photodiode responsivity does not affect the system performance.

Without loss of generality, we assume that \(\phi_s(t) = \phi_s(t-T) = 0\) when the consecutive transmitted phases are the same. Assume an unity photodiode responsivity, similar to that of Fig. 3.3, the photocurrent at the output of the balanced receiver is

\[
i(t) = |E_1(t)|^2 - |E_2(t)|^2,
\]
where

\[
|E_1(t)|^2 = \left| A + \frac{1}{2} \left[ n_x(t) + n_x(t-T) \right] \right|^2 + \frac{1}{4} \left| n_y(t) + n_y(t-T) \right|^2 ,
\]

and

\[
|E_2(t)|^2 = \frac{1}{4} \left| n_x(t) - n_x(t-T) \right|^2 + \frac{1}{4} \left| n_y(t) - n_y(t-T) \right|^2 .
\]

The error probability is equal to

\[
p_e = \Pr \{ i(t) < 0 \} = \Pr \{ |E_2(t)|^2 > |E_1(t)|^2 \} .
\]

Because \( n_x(t) + n_x(t-T) \) is independent of \( n_x(t-T) \) and \( n_y(t) + n_y(t-T) \) is independent of \( n_y(t) \), \( |E_1(t)|^2 \) and \( |E_2(t)|^2 \) are independent of each other. From the error probability of Eq. (3.155), similar to heterodyne DPSK signal in Sec. 3.3.3, DPSK signal can be analyzed as noncoherent detection of an orthogonal binary signal.

The p.d.f. of \( |E_1(t)|^2 \) of Eq. (3.153) is

\[
p_{|E_1|^2}(y) = \frac{1}{2\sigma^2 A} \sqrt{y} \exp \left( -\frac{A^2 + y}{2\sigma^2} \right) I_1 \left( \sqrt{\frac{y}{\sigma^2}} A \right) , \quad y \geq 0 ,
\]

where \( I_1(\cdot) \) is the first-order modified Bessel function of the first kind and the variance parameter \( \sigma^2 = \sigma_n^2 / 2 \). The p.d.f. of \( p_{|E_1|^2}(y) \) is noncentral \( \chi^2 \) distribution with four degrees of freedom with a variance parameter of \( \sigma^2 = \sigma_n^2 / 2 \) and noncentrality parameter of \( A^2 \). The variance of \( \sigma^2 = \sigma_n^2 / 2 \) is the variance per dimension of the random variables of \( [n_x(t) \pm n_x(t-T)]/2 \) and \( [n_y(t) \pm n_y(t-T)]/2 \) in Eqs. (3.153) and (3.154).

The p.d.f. of \( |E_2(t)|^2 \) of Eq. (3.154) is

\[
p_{|E_2|^2}(y) = \frac{1}{4\sigma^2} y \exp \left( -\frac{y}{2\sigma^2} \right) , \quad y \geq 0 .
\]

The p.d.f. of \( p_{|E_2|^2}(y) \) is the \( \chi^2 \) distribution with four degrees of freedom.

First, we need to find the probability of (Gradshteyn and Ryzhik, 1980, §3.351)

\[
\Pr \{ |E_2(t)|^2 > y \} = \int_y^{+\infty} p_{|E_2|^2}(y) \, dy = \exp \left( -\frac{y}{2\sigma^2} \right) \left[ 1 + \frac{y}{2\sigma^2} \right] .
\]

The error probability of Eq. (3.155) is

\[
p_e = \int_0^{+\infty} \Pr \{ |E_2(t)|^2 > y \} \, p_{|E_1|^2}(y) \, dy .
\]
Using the p.d.f. of Eq. (3.156) and the probability of Eq. (3.158), after some simplifications, we get

\[ p_e = \frac{e^{-\rho_s}}{\sqrt{2\rho_s}} \int_0^{+\infty} (1 + x)\sqrt{x}e^{-2x}I_1 \left(2\sqrt{2\rho_s x}\right) dx, \quad (3.160) \]

where \( x = y/(2\sigma^2) \), and \( \rho_s = E_0^2/\sigma^2 = E_0^2/(2\sigma_n^2) \) is the SNR. As the special case of Gradshteyn and Ryzhik (1980, §6.631), we get

\[ \int_0^{\infty} x^{1/2}e^{-2x}I_1(a\sqrt{x})dx = (a/8)e^{a^2/8}, \quad (3.161) \]

\[ \int_0^{\infty} x^{3/2}e^{-2x}I_1(a\sqrt{x})dx = [a(16 + a^2)/128]e^{a^2/8}. \quad (3.162) \]

The integration of Eq. (3.160) gives the error probability of

\[ p_e = \frac{\exp(-\rho_s)}{2} \left(1 + \frac{\rho_s}{4}\right). \quad (3.163) \]

Comparing the heterodyne error probability in Sec. 3.3.3, with the amplifier noise from the orthogonal polarization, the error probability is increased by a factor of \( 1 + \rho_s/4 \). The increase of the error probability is the similar to that for direct-detection ASK signals of Eq. (3.132).

Figure 3.13 shows the error probability of phase-modulated signal as a function of SNR \( \rho_s \). The error probability of synchronous detection of \( \frac{1}{2}\text{erfc}\sqrt{\rho_s} \) from Eq. (3.78), the error probability of asynchronous heterodyne differential detection of \( \frac{1}{2}e^{-\rho_s} \) from Eq. (3.105), and the error probability of direct-detection of Eq. (3.163) are also shown for comparison. For an error probability of \( 10^{-9} \), asynchronous heterodyne detection is about 0.45 dB worse than synchronous detection, and direct-detection is about 0.40 dB worse than asynchronous differential detection. The quantum-limited receivers require 18.0, 20.0, and 21.9 photons/bit.

The degradation of direct-detection is due to the inclusion of amplifier noise from orthogonal polarization. If a lossless polarizer precedes the detector, an improvement of 0.4 dB can be expected. Tonguz and Wagner (1991) shown that direct-detection DPSK receiver performs the same as heterodyne differential detection if the amplifier noise from orthogonal direction is ignored. The error probability of Eq. (3.163) was first derived by Okoshi et al. (1988) for DPSK signals with similar noise characteristics.

If the direct-detection receiver has a noise bandwidth far larger than the signal bandwidth, the system was analyzed in Humblet and Azizoglu (1991), Jacobsen (1993), and Chinn et al. (1996). The DPSK error probability of Eq. (3.105) assumes that there are two noise sources of
n_1(t) and n_2(t). The error probability of Eq. (3.163) assumes that there are four noise sources of n_x1(t), n_x2(t), n_y1(t), and n_y2(t) for both real and imaginary parts noise. With the assumption of both Marcuse (1990) and Humblet and Azizoğlu (1991) that there are 2k independent noise sources affects the DPSK signals, the error probability are

\[ p_e = \frac{e^{-\rho_s}}{2} \sum_{m=1}^{k} h_k \frac{\rho_s^m}{m!}, \]  

(3.164)

where

\[ h_k = \frac{1}{2^{2(k-1)}} \sum_{l=0}^{k-1} \binom{2k-l-2}{k-1-m}. \]  

(3.165)

Direct-detection DPSK signal is unquestionable the most popular phase-modulated optical communication scheme as shown in Table 1.2. DPSK receivers with very good receiver sensitivity were developed by Atia and Bondurant (1999), Gnauck et al. (2003a), and Sinsky et al. (2003). Both Gnauck and Winzer (2005) and Xu et al. (2004) reviewed the activities of direct-detection DPSK systems. DPSK signal can also
be detected using an optical filter similar to frequency discriminator (Lyubomirsky and Chien, 2005).

Single-branch direct-detection DPSK receiver converts the DPSK signal to an equivalent on-off keying signal. Comparing the receiver sensitivity of on-off keying with DPSK signal, single-branch direct-detection DPSK receiver has a receiver sensitivity 3-dB worse than the balanced receiver and has a performance the same as on-off keying signal.

4.3 Dual-Filter Direct-Detection of FSK Receiver

Figure 3.14 shows a dual-filter direct-detection FSK receiver. Figure 3.14(a) is the schematic of the receiver in which a balanced receiver is used with one detector connected to the output of each optical filter. The optical filters can be implemented using fiber Bragg grating or multilayer dielectric filters as shown in Figs. 3.14(b) and (c), respectively. The optical filters center at the optical frequencies of $f_1$ and $f_2$, corresponding to the two angular frequencies of $\omega_1$ and $\omega_2$ for binary FSK signal, respectively.

If the two optical filters are matched filter and the two FSK signals are orthogonal with each other, for lossless optical filter without loss of generality,

$$E_1(t) = [A \cos \omega_1 t + n_{x1}(t)]x + n_{y1}(t)y$$

(3.166)

if $s_1(t)$ is transmitted, where $n_{x1}(t)$ and $n_{y1}(t)$ are the amplifier noises in the polarization parallel and orthogonal to the signal, respectively.
The noise variance is $E\{|n_{x1}(t)|^2\} = E\{|n_{y1}(t)|^2\} = \sigma_n^2$ and the SNR is $\rho_s = A^2/2\sigma_n^2$. If $s_1(t)$ is transmitted, the electric field at the output of the optical filter centered at $f_2$ is

$$E_2(t) = n_{x2}(t)x + n_{y2}(t)y.$$  \hspace{1cm} (3.167)

With a photocurrent of $i(t) = R|E_1(t)|^2 - R|E_2(t)|^2$ and an error probability of $p_e = Pr\{i(t) < 0\}$, the error probability is the same as that for DPSK signal of Eq. (3.163) but half the SNR. For dual-filter direct-detection FSK receiver, the error probability is

$$p_e = \frac{1}{2} \exp \left( -\frac{\rho_s}{2} \right) \left( 1 + \frac{\rho_s}{8} \right).$$  \hspace{1cm} (3.168)

Direct-detection FSK signal is 3-dB worse than direct-detection DPSK signal. However, using the same receiver of Fig. 3.12, direct-detection MSK receiver has the same performance as DPSK signal.

Figure 3.15 shows the error probability of FSK signal demodulated using a synchronous receiver, asynchronous heterodyne receiver, and direct-detection dual-filter receiver. Compared with Fig. 3.13, frequency-modulated signal is 3-dB worse than phase-modulated signal. For an
error probability of \(10^{-9}\), asynchronous heterodyne detection is about 
0.45 dB worse than synchronous detection, and direct-detection is about 
0.40 dB worse than asynchronous differential detection. The quantum-
limited receivers require 36.0, 40.0, and 43.8 photons/bit.

Practical FSK receiver may use a single filter with a performance 
similar to ASK signal with 3-dB worse receiver sensitivity. If the two 
optical filters have crosstalk, the outputs of \(E_1(t)\) and \(E_2(t)\) have corre-
lation and the error probability is given by Eq. (3.119) with a correlation 
coefficient depending on the filter crosstalk. In order to improve the per-
formance, FSK signal with large frequency deviation can be used with 
discriminator-based detector for better performance. The performance 
of FSK signal with frequency discriminator is the same as heterodyne 
system analyzed in Sec. 3.3.2. However, high frequency-deviation FSK 
system has very small spectral efficiency.

Direct-detection receiver for frequency-modulated signal was used for 
a long time by Saito and Kimura (1964), Saito et al. (1983), and Ols-
son and Tang (1979). Single-filter direct-detection FSK receiver can use 
Fabry-Perot resonator (Chraplyvy et al., 1989, Kaminow, 1990, Kaminow 
et al., 1988, Malyon and Stallard, 1990, Willner, 1990, Willner et al.,
1990) or ring resonator (Oda et al., 1991, 1994). Using the interferom-
eter of Fig. 3.12, CPFSK signal was directly detected by Idler et al.
an optical filter is used to demodulate the FSK signal, it can also func-
tion as a demultiplexer to select the corresponding WDM channel.

5. Phase-Diversity Receiver

Phase-diversity receiver is another type of asynchronous detector for 
homodyne receiver. The phase-diversity receiver is based on the quadra-
ture receiver of Fig. 3.4. From the photocurrent of Eqs. (3.49) and (3.50) 
with \(\omega_{1F} = 0\), including a random phase of \(\theta_0\) from either the received 
signal or the LO signal, we obtain

\[
\begin{align*}
    r_I(t) &= A_s(t) \cos[\phi(t) + \theta_0] + n_I(t), \\
    r_Q(t) &= A_s(t) \sin[\phi(t) + \theta_0] + n_Q(t),
\end{align*}
\]

where \(A_s(t)\) is due to amplitude modulation, \(\phi(t)\) from phase modu-
lation, and \(n_I(t)\) and \(n_Q(t)\) are the identical independently distributed 
additive Gaussian noise. The random phase of \(\theta_0\) in Eq. (3.169) is used 
to model a receiver without phase locking. In Eq. (3.169), the random 
phase of \(\theta_0\) is a constant over a bit interval of \(T\) but can be changed 
slowly from bit to bit.
5.1 Phase-Diversity ASK Receiver

If the signal is amplitude-modulated with \( \phi_s(t) = 0 \) in the signals of Eq. (3.169), amplitude modulated signal with \( A(t) = \{0, A\} \) may be demodulated by the received envelope of

\[
r_d(t) = \sqrt{r_I^2(t) + r_Q^2(t)}. \tag{3.170}
\]

Note that the phase-diversity ASK receiver is similar to heterodyne envelope-detection receiver of Sec. 3.3.1. The error probability of phase-diversity ASK receiver is the same as Eq. (3.93) of \( p_e = \frac{1}{2} \exp(-\rho_s/2) \) if the threshold is chosen as the \( A/2 \). As a homodyne phase-diversity ASK receiver is mathematically the same as a heterodyne ASK receiver based on envelope detection, other aspects of a homodyne phase-diversity ASK receiver can also be analyzed the same as the corresponding receivers in Sec. 3.3.1 or Fig. 3.11.

The linear optical sampling scheme of Dorrer et al. (2003) is functionally a phase-diversity ASK receiver using LO laser with short optical pulse train.

5.2 Phase-Diversity DPSK Receiver

If the data in encode in the phase difference of \( \phi_s(t) - \phi_s(t-T) \) using DPSK modulation, the amplitude of \( A(t) = A \) is a constant. The phase difference can be demodulated using

\[
r_d(t) = r_I(t)r_I(t-T) + r_Q(t)r_Q(t-T) = A^2 \cos[\phi_s(t) - \phi_s(t-T)] + \text{noise terms}. \tag{3.171}
\]

Without noise, \( r_d(t) \) is proportional to \( \cos[\phi_s(t) - \phi_s(t-T)] \) and \( r_d(t) = \pm A^2 \) when \( \phi_s(t) - \phi_s(t-T) = 0 \) or \( \pi \), respectively. The phase-diversity receiver for DPSK signal has the same performance as a DPSK heterodyne receiver using differential detection of Sec. 3.3.3 or Fig. 3.13 with an error probability of

\[
p_e = \frac{1}{2} \exp(-\rho_s). \tag{3.172}
\]

5.3 Phase-Diversity Receiver for Frequency-Modulated Signals

For FSK signal, the received signals of Eq. (3.169) at the output of the quadrature homodyne receiver of Fig. 3.4 are

\[
\begin{align*}
r_I(t) & = A \cos(\pm \pi \Delta f t) + n_I(t), \tag{3.173} \\
r_Q(t) & = A \sin(\pm \pi \Delta f t) + n_Q(t). \tag{3.174}
\end{align*}
\]
where $\Delta f = (\omega_1 - \omega_2)/(2\pi)$ is the frequency difference between the binary FSK signal. The demodulated signal can be

$$r_d(t) = r_I(t) \frac{dr_Q(t)}{dt} + r_Q(t) \frac{dr_I(t)}{dt} = \mp \pi A^2 + \text{noise terms.}$$

(3.175)

The receiver sensitivity increases with the frequency difference of $\Delta f$. The performance of the system is similar to that of Sec. 3.3.5 using frequency discriminator.

For CPFSK signal, the output from the receiver is similar to the two signals of Eq. (3.174)

$$r_I(t) = A \cos(\pm \pi \Delta f t + \phi_0 + \theta_0) + n_I(t),$$

(3.176)

$$r_Q(t) = A \sin(\pm \pi \Delta f t + \phi_0 + \theta_0) + n_Q(t).$$

(3.177)

The demodulated signal is

$$r_d(t) = r_I(t)r_Q(t - \tau) + r_Q(t)r_I(t - \tau) = \pm A^2 \sin(\pi \Delta f \tau) + \text{noise terms.}$$

(3.178)

For MSK signal, $\tau = T$ and $\Delta f = 1/2T$, the receiver sensitivity is the same as differential detection DPSK signals of Eq. (3.105) or Fig. 3.13.


When the system is limited by optical amplifier noise, phase-diversity receiver performs about 0.4 dB better than direct-detection receiver for an error probability of $10^{-9}$. The main advantage of phase-diversity receiver is to provide narrow channel spacing for WDM systems or reduce the bandwidth requirement of the receiver. The same as the image-rejection heterodyne receiver of Sec. 3.1.4, phase-diversity receiver also has the advantage to reduce the channel spacing of a WDM system. Both phase-diversity homodyne and image-rejection heterodyne receivers use the same optical front-end of Fig. 3.4 with a 90° optical hybrid.

6. Polarization-Diversity Receiver

The single-branch receiver of Fig. 3.1 and the balanced receiver of Fig. 3.3 all require polarization control to match the polarization of
the received signal with that of the LO laser. The image-rejection receiver of Fig. 3.5 requires polarization control such that the polarization of the received signal is 45° linearly polarized with respect to the receiver polarization. System without polarization control is possible using polarization-diversity techniques.

Figure 3.16 shows a polarization-diversity receiver with an optical front end similar to the 90° optical hybrid of Fig. 3.4. However, unlike Fig. 3.4 in which the 90° optical hybrid is operated with linearly polarized received signal, the received optical field of the polarization-diversity receiver is generally elliptically polarized and uncontrolled. The LO laser is linearly polarized at 45° with respect to the receiver polarization. The received signal is mixed with the LO signal using a 3-dB coupler and forwards to two separated PBS.

With random polarization without APC, the received signal is assumed to have an electric field of

$$E_r(t) = A_s(t) \left[ \cos \varphi x + \sin \varphi e^{j\theta} x \right] e^{j\omega_c t + j\phi_s(t)}, \quad (3.179)$$

where the angles of $\varphi$ and $\theta$ are relative to the receiver polarization. Linearly 45° polarized to the receiver polarization, the LO laser has an electric field of

$$E_{LO}(t) = \frac{A_L}{\sqrt{2}} (x + y) e^{j\omega_{LO} t}. \quad (3.180)$$

The electric fields at the outputs of the 3-dB coupler are

$$E_1(t) = \frac{1}{\sqrt{2}} [E_r(t) + E_{LO}(t)]$$

$$= \frac{x}{2} \left[ A_s(t) \cos \varphi e^{j\omega_c t + j\phi_s(t)} + A_L e^{j\omega_{LO} t} \right]$$

$$+ \frac{y}{2} \left[ A_s(t) \sin \varphi e^{j\omega_c t + j\phi_s(t) + j\theta} + A_L e^{j\omega_{LO} t} \right], \quad (3.181)$$
and

\[
E_2(t) = \frac{1}{\sqrt{2}} [E_r(t) - E_{LO}(t)] \\
= \frac{x}{2} \left[ A_s(t) \cos \varphi e^{j\omega_c t + j\varphi_s(t)} - A_L e^{j\omega_{LO} t} \right] \\
+ \frac{y}{2} \left[ A_s(t) \sin \varphi e^{j\omega_c t + j\varphi_s(t) + j\theta} - A_L e^{j\omega_{LO} t} \right]. \tag{3.182}
\]

After the two PBS, similar to the quadrature receiver of Fig. 3.4, the photocurrents at the output of the two balanced receivers are

\[
i_i(t) = R \cos \varphi A_s(t) A_L \cos [\omega_{IF} t + \phi_s(t)], \tag{3.183}
\]

and

\[
i_Q(t) = R \sin \varphi A_s(t) A_L \cos [\omega_{IF} t + \phi_s(t) + \theta]. \tag{3.184}
\]

with a phase difference of \(\theta\).

In both photocurrents of Eqs. (3.183) and (3.184), the additive noise has the same variance and independent of each other. Including noise, the received signal is \(i_i(t) + n_I(t)\) and \(i_Q(t) + n_Q(t)\) where \(E\{n_I^2(t)\} = E\{n_Q^2(t)\} = \sigma_n^2\).

### 6.1 Combination in Polarization-Diversity Receiver

The polarization-diversity scheme is applicable to most modulation formats. Data are demodulated by combining the information from two polarization branches. The photocurrents can be processed in either the IF or the baseband. If the signal are combined in the IF stage, the carrier phase must be matched to cancel the phase difference of \(\theta\) between Eqs. (3.183) and (3.184). When the phase difference of \(\theta\) changes due to external disturbance, the phases of two signals must be adjusted adaptively before the combination process. In baseband combination, the signal of either \(A_s(t)\) or \(\phi_s(t)\) is demodulated independently for each polarization component. If the demodulation process tracks out the phase fluctuation of the IF signals, phase matching is not necessary. In practice, baseband combining is more practical with simple implementation.

Without loss of generality with \(R = 1\), we assume that the signals after phase matching for \(\theta = 0\) are

\[
r_i(t) = \cos \varphi A_s(t) \cos \phi_s(t) + n_i(t), \tag{3.185}
\]

and

\[
r_q(t) = \sin \varphi A_s(t) \cos \phi_s(t) + n_q(t), \tag{3.186}
\]

corresponding to the in- and quadrature-phase components.
Among all methods to combine the two polarization components, maximum ratio is the best that maximizes the output SNR. The in-phase component of Eq. (3.183) has a gain of \( \cos \varphi \) and the quadrature-phase component of Eq. (3.184) has a gain of \( \sin \varphi \). With maximum-ratio combination, the combined signal is

\[
r_c(t) = \cos \varphi r_i(t) + \sin \varphi r_q(t),
\]

\[
= A_s(t) \cos \phi_s(t) + \cos \varphi n_i(t) + \sin \varphi n_q(t),
\]

(3.187)

where the combined signal is the same as that with polarization control and the noise is also Gaussian noise having the same variance as \( n_i(t) \) or \( n_q(t) \). With maximum-ratio combination, there is no penalty using polarization diversity.

The simplest combining scheme may be equal-gain combining with a combined signal of

\[
r_c(t) = r_i(t) + r_q(t)
\]

\[
= (\cos \varphi + \sin \varphi) A_s(t) \cos \phi_s(t) + n_i(t) + n_q(t).
\]

(3.188)

The SNR penalty due to equal-gain combining is

\[
\delta_p = \frac{(\cos \varphi + \sin \varphi)^2}{2} = \frac{1 + \sin(2\varphi)}{2}, \quad 0 < \varphi < \frac{\pi}{2}
\]

(3.189)

Selection-combining scheme chooses the polarization component with the largest power. The penalty due to selection combining is

\[
\delta_p = \max(\cos^2 \varphi, \sin^2 \varphi) = \frac{1}{2} (1 + |\cos(2\varphi)|), \quad 0 < \varphi < \frac{\pi}{2}
\]

(3.190)

where the factor of 2 for noise enhancement is due to the addition of two identical and independent noise sources.

Another combining scheme can be used for either heterodyne or homodyne ASK signal to square and combine the two signals. The combined signal is

\[
r_c(t) = r_i^2(t) + r_q^2(t)
\]

\[
= A_s^2(t) \cos^2 \phi_s(t) + 2A_s(t) \cos \phi_s(t)[n_i(t) + n_q(t)]n_i^2(t) + n_q^2(t).
\]

(3.191)

In homodyne detection, square-combining is only possible for ASK signal in which \( \phi_s(t) = 0 \) and \( A_s(t) \in \{A, 0\} \). In heterodyne scheme, square combination is also possible for both DPSK and FSK signals with envelope detection.
Figure 3.17 shows the SNR penalty for various signal combination schemes. The SNR penalty is shown as a function of the polarization angle of $\varphi$. Square combination has a penalty of $0.4 \text{ dB}$ and is usually better than either equal or selected combining in most of the cases. Equal combining is the best when the polarization angle is around $\varphi = \pi/4$ and selected combining is the best when the polarization angle is around $\phi = 0$ and $\pi/2$.

First suggested by Okoshi (1985, 1986) for combination in IF, the above polarization-diversity schemes were discussed in detail in Ryu (1995). Figure 3.17 is almost identical to similar figure in Ryu (1995).

The two independent channels in polarization-division multiplexing (PDM) can be separated with maximum-ratio combination. If the orthogonal signals are $A_{s1}(t)e^{j\phi s1(t)}$ and $A_{s2}(t)e^{j\phi s2(t)}$, after IF processing, the two output signals are

$$r_i(t) = \cos \varphi A_{s1}(t) \cos \phi_{s1}(t) + \sin \varphi A_{s2}(t) \cos \phi_{s2}(t) + n_i(t), \quad (3.192)$$

and

$$r_q(t) = \sin \varphi A_{s1}(t) \cos \phi_{s1}(t) - \cos \varphi A_{s2}(t) \cos \phi_{s2}(t) + n_q(t). \quad (3.193)$$

The two independent signals can be recovered based on the combiner of

$$\cos \varphi r_i(t) + \sin \varphi r_q(t) \quad (3.194)$$
and
\[ \sin \varphi r_i(t) - \cos \varphi r_q(t), \quad (3.195) \]
respectively.

### 6.2 Heterodyne Differential Detection with Polarization Diversity

Figure 3.18 shows a polarization-diversity receiver for both DPSK and MSK signals using electrical delay-and-multiplier circuits. Without loss of generality and assumes a DPSK signal, after the delay-and-multiplier circuitry, the upper branch of Fig. 3.18 has a signal of

\[ r_1(t) = \cos^2 \varphi A^2 \cos[\phi_s(t) - \phi_s(t - T)] + \text{noise terms}, \quad (3.196) \]

and the lower branch of Fig. 3.18 has a signal of

\[ r_2(t) = \sin^2 \varphi A^2 \cos[\phi_s(t) - \phi_s(t - T)] + \text{noise terms}, \quad (3.197) \]

where both factors \( \cos^2 \varphi \) and \( \sin^2 \varphi \) are given by the multiplexer of Fig. 3.18.

When the above two signals are combined, the decision variable becomes

\[ r_d(t) = A^2 \cos[\phi_s(t) - \phi_s(t - T)] + \text{noise terms}. \quad (3.198) \]

In the above two equations, the common factor of \( R^2/2 \) is ignored for simplicity. If all noise terms are written down, the noise statistics is the same as that of direct-detection DPSK signal in Sec. 3.4.2 with the error probability of Eq. (3.163). The error probability of Eq. (3.163) was first derived by Okoshi et al. (1988) for phase-diversity DPSK signals. MSK signal should also have the same performance. From Fig. 3.13, the degradation of phase-diversity DPSK or MSK signal is about 0.40 dB compared with heterodyne DPSK signal, similar to the degradation of the square combination of Fig. 3.17.
When FSK signal is detected through envelope detection of Fig. 3.8 based on two filters, the signal passes through a squarer. For envelope detection of FSK signal, the decision variable is not sensitive to the phase difference of $\theta$ in Eq. (3.184). When the detected signals from two branches of Fig. 3.18 are combined together, the combination is actually based on square combination and has a degradation of 0.40 dB as from Fig. 3.13.


7. Polarization-Shift Keying Modulation

A single-mode optical fiber can support two polarizations and the electric field in an optical fiber can generally be expressed as

$$E_r(t) = [E_x(t)x + E_y(t)y] e^{j\omega_c t}.$$

The above electric field of Eq. (3.199) is the same as the signal of Eq. (3.179) but using different notation. In the single-branch receiver of Fig. 3.1, we assume that the signal of Eq. (3.2) has a single polarization and aligned with the reference polarization of the receiver of $x$.

Comparing the electric fields of Eq. (3.199) with Eq. (3.1), both $E_x(t)$ and $E_y(t)$ can be used independently to transmit two data streams using polarization-division multiplexing (PDM). In PDM system, a PBS is used to separate $E_x(t)$ and $E_y(t)$. APC is required to align the signal to the PBS. After the PBS, the two data streams encoded in $E_x(t)$ and $E_y(t)$ are then demodulated independently.

The electric fields of $E_x(t)$ and $E_y(t)$ of Eq. (3.199) can be used together to converse information by polarization-shift keying (PolSK). The simplest PolSK scheme is to transmit

$$s_1(t) = A_x e^{j\omega_c t},$$
and

$$s_2(t) = Ay e^{j\omega t}$$

(3.201)

with $E_x(t) = E_y(t) = A$ but used alternatively to carry either “0” or “1”. The performance of this simple PolSK scheme is identical to that of FSK signal with two orthogonal binary signals. Similar to FSK signal, a PolSK signal can be directly demodulated using a PBS, followed by a balanced receiver. The error probability of direct-detection PolSK is that of Eq. (3.85) of $p_e = \frac{1}{2} \exp(-\rho_s/2)$. When a heterodyne receiver is used, the LO laser can have a polarization that is 45° to both x and y. The performance of heterodyne receiver is also the same as that of Eq. (3.85).

Direct-detection polarization modulation has a long history (Daino et al., 1974, Pratt, 1966). For coherent optical communications, PolSK was proposed mainly to overcome laser phase noise (Benedetto and Poggiolino, 1990, Betti et al., 1988, Calvani et al., 1988, Dietrich et al., 1987, Imai et al., 1990b). When PolSK is designed based on the Stokes parameters, Dietrich et al. (1987) used the $s_1$ parameter, Calvani et al. (1988) detected the $s_2$ parameter, Imai et al. (1990b) based on the differential $s_1$ parameter, and Betti et al. (1988) and Benedetto and Poggiolino (1990) detected all Stokes parameters without optical polarization control and used signal processing to find the correct polarization states of the received signal. While PolSK systems usually use for heterodyne receiver, homodyne PolSK systems were described in Betti et al. (1991). PolSK systems are analyzed in details by Benedetto and Poggiolini (1992) and Benedetto et al. (1995a,b).

8. Comparison of Optical Receivers

Table 3.2 shows the performance of a quantum-limited optical receiver for various types of signal. An optimal receiver is assumed in Table 3.2 with, for example, optimal threshold setting and optical matched filter. The SNR penalty is also calculated in Table 3.2 compared with a homodyne or heterodyne PSK receiver limited by amplifier noise. The performance of Table 3.2 includes the performance of shot- and amplifier-noise limited signals.

Synchronous receivers for phase-modulated signals have best receiver sensitivity. Asynchronous receivers for phase-modulated signals have a degradation less than 1 dB compared to the best synchronous receiver. The performance of phase-diversity receiver is the same as asynchronous heterodyne receiver.

Based on the square combination, polarization-diversity receiver has the performance the same as direct-detection receiver. If maximum-
Table 3.2. Comparison of Different Optical Receivers for an Error Probability of $10^{-9}$.

<table>
<thead>
<tr>
<th>Modulation Format</th>
<th>Sensitivity (photons/bit)</th>
<th>Penalty (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homodyne PSK</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Heterodyne PSK</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Heterodyne DPSK, MSK</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Direct-Detection DPSK</td>
<td>–</td>
<td>22</td>
</tr>
<tr>
<td>Homodyne ASK</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>Heterodyne ASK, FSK</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Envelope Detection Heterodyne ASK</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>Direct-Detection ASK, FSK</td>
<td>–</td>
<td>40</td>
</tr>
<tr>
<td>Dual-Filter FSK, PolSK</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Single-Filter FSK</td>
<td>80</td>
<td>80</td>
</tr>
</tbody>
</table>

ratio combination is used, polarization-diversity receiver has the same performance the corresponding homodyne or heterodyne receiver.

This chapter just analyzes the receiver with only the dominant amplifier noise. In next chapter, in linear regime, other degradations to the signal is studied.

**APPENDIX 3.A: Marcum Q Function**

For a Gaussian random variable of $A + x_1$ and $x_2(t)$, the amplitude of $R = \sqrt{[A + x_1]^2 + x_2^2}$ has a Rice distribution of

$$p(r) = \frac{r}{\sigma_n^2} I_0 \left( \frac{rA}{\sigma_n^2} \right) \exp \left( -\frac{r^2 + A^2}{2\sigma_n^2} \right).$$

The cumulative distribution function of Rice distribution is

$$\int_x^{+\infty} p(r)dr = Q \left( \frac{A}{\sigma_n}, \frac{x}{\sigma_n} \right),$$

where the Marcum Q function was first used for radar theory (Marcum, 1960) and is very useful in the analysis of noncoherent or asynchronous detection of binary signals. This appendix presented some important properties of Marcum Q functions.
The Marcum Q function is a real function of

\[ Q(a, b) = \int_b^\infty x I_0(ax) \exp\left(-\frac{x^2 + a^2}{2}\right) \, dx, \]  

(3.3.3)

or

\[ Q\left(\sqrt{2a}, \sqrt{2b}\right) = \int_b^\infty e^{-(a+x)} I_0(2\sqrt{ax}) \, dx, \]  

(3.3.4)

where \( I_0(x) \) is the zeroth-order modified Bessel function of the first kind. From the definition, we have

\[ Q(0, b) = e^{-b^2/2}, \]  

(3.3.5)

\[ Q(a, 0) = 1. \]  

(3.3.6)

The modified Bessel function of \( I_0(x) \) can be represented as inverse Laplace transform of

\[ -\frac{1}{p} \exp\left(\frac{ap + b}{p}\right) \]  

(3.3.7)

where \( c \) is a real positive number. Substitute Eq. (3.3.7) into the Marcum Q function of Eq. (3.3.3) and exchange the order of integration, we get

\[ Q\left(\sqrt{2a}, \sqrt{2b}\right) = e^{-(a+b)} \frac{1}{2\pi j} \int_{c-\infty}^{c+j\infty} \frac{1}{p} \int_b^\infty e^{-x} \exp\left(\frac{ap + b}{p}\right) \, dx \, dp, \quad c > 0 \]  

(3.3.8)

or

\[ Q\left(\sqrt{2a}, \sqrt{2b}\right) = e^{-(a+b)} \frac{1}{2\pi j} \int_{c-\infty}^{c+j\infty} \frac{\exp\left(\frac{bp + a}{p}\right)}{p(p-1)} \, dp, \quad 0 < c < 1 \]  

(3.3.9)

and

\[ Q(a, b) = -e^{-(a^2+b^2)/2} \frac{1}{2\pi j} \int_{c-\infty}^{c+j\infty} \frac{\exp\left(\frac{b^2}{2} + \frac{b^2}{p}\right)}{p(p-1)} \, dp, \quad 0 < c < 1 \]  

(3.3.10)

By straightforward residue calculation involving shifting of the path, one immediately also derives the useful symmetry relationships

\[ Q(a, b) + Q(b, a) = 1 + e^{-(a^2+b^2)/2} I_0(ab), \]  

(3.3.11)

\[ Q(a, a) = \frac{1}{2} + \frac{1}{2} e^{-a^2} I_0(a^2). \]  

(3.3.12)

The Marcum Q function can be calculated using a function series of

\[ Q(a, b) = e^{-(a^2+b^2)/2} \sum_{m=0}^{\infty} \left(\frac{a}{b}\right)^m I_m(ab), \quad b > a \]  

(3.3.13)

or using the method in Cantrell and Ojha (1987).
In the derivation of the error probability of orthogonal modulation with correlation, we need to evaluate the probability that a Rice distributed random variable exceeds another. If the envelope of two Gaussian processes, \( R_1 \) and \( R_2 \), are independently distributed, with the well-known Rice p.d.f. of

\[
p_{R_1}(r_1) = \frac{t^1}{\sigma_1^1} \exp \left( -\frac{A^1_1 + r_1^2}{2\sigma_1^2} \right) I_0 \left( \frac{A_r^1}{\sigma_1^2} \right),
\]

\[
p_{R_2}(r) = \frac{t^2}{\sigma_2^2} \exp \left( -\frac{A^2_2 + r_2^2}{2\sigma_2^2} \right) I_0 \left( \frac{A_r^2}{\sigma_2^2} \right),
\]

the error probability is

\[
p_e = \Pr \{ R_1^2 < R_2^2 \} = \Pr \{ R_1 < R_2 \}
\]

\[
= \int_0^{\infty} p_{R_1}(r_1) \int_{r_1}^{\infty} p_{R_2}(r_2) dr_2 dr_1
\]

\[
= \int_0^{\infty} p_{R_1}(r_1) Q \left( \frac{A_2}{\sigma_2}, \frac{r_1}{\sigma_2} \right) dr_1.
\]

By using Eq. (3.A.8) in (3.A.17), interchanging orders of integrations, we find that

\[
p_e = \exp \left[ -\frac{1}{2} \left( \frac{A_1^2}{\sigma_1^2} + \frac{A_2^2}{\sigma_2^2} \right) \right] \cdot \frac{1}{2\pi j} \int_{-j\infty}^{c+j\infty} \exp \left( \frac{A_2^2 p}{2\sigma_2^2} \right) \frac{1}{p - 1} dp
\]

\[
\times \frac{1}{1 + \frac{A_2^2}{\sigma_2^2} \left( 1 - \frac{1}{p} \right)} \exp \left[ \frac{A_1^2}{2\sigma_1^2} \frac{1}{1 + \frac{A_2^2}{\sigma_2^2} \left( 1 - \frac{1}{p} \right)} \right] dp, \quad c > 1.
\]

Using the obvious partial fraction expansion, after some algebra via Eqs. (3.A.7) and (3.A.8), the error probability is

\[
p_e = Q(a, b) - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} e^{-\left( a^2 + b^2 \right) / 2} I_0(\sigma_1 ab),
\]

where

\[
a^2 = \frac{A_2^2}{\sigma_1^2 + \sigma_2^2}, \quad b^2 = \frac{A_1^2}{\sigma_1^2 + \sigma_2^2}.
\]

The results of Eq. (3.A.19) can also be written in more symmetric forms of

\[
p_e = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} [1 - Q(b, a)] + \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} Q(a, b),
\]

and

\[
p_e = \frac{1}{2} [1 - Q(b, a) + Q(a, b)] - \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2} e^{-\left( a^2 + b^2 \right) / 2} I_0(\sigma_1 ab).
\]

This Appendix follows the approaches of Stein (1964) for Marcum Q function. Both Schwartz et al. (1966) and Betti et al. (1995) also had similar Appendix. Higher-order Marcum Q functions are considered in Proakis (2000, Appendix B). In this book, only the second-order Marcum Q function of Eq. (3.130) is used.