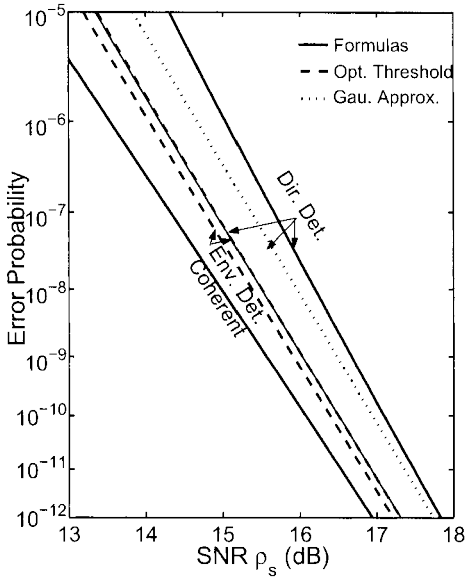

Figure 3.11. The error probability of amplitude-modulated signals as a function of SNR ρ_s.

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 ω_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.

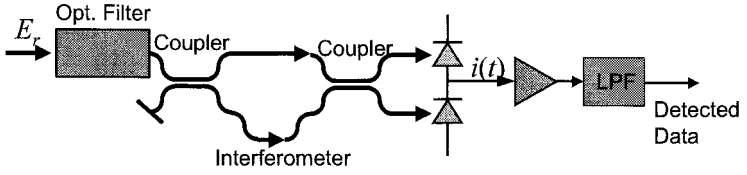

Figure 3.12. Direct-detection DPSK receiver using an unpolarized asymmetric Mach-Zehnder interferometer.

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

$$E_1(t) = \frac{\mathbf{x}}{2}[Ae^{-j\phi_s(t)} + n_x(t)] + \frac{\mathbf{y}}{2}n_y(t) + \frac{\mathbf{x}}{2}[Ae^{-j\phi_s(t-T)} + n_x(t-T)] + \frac{\mathbf{y}}{2}n_y(t-T), \tag{3.150}$$

and

$$E_2(t) = \frac{\mathbf{x}}{2}[Ae^{-j\phi_s(t)} + n_x(t)] + \frac{\mathbf{y}}{2}n_y(t) - \frac{\mathbf{x}}{2}[Ae^{-j\phi_s(t-T)} + n_x(t-T)] - \frac{\mathbf{y}}{2}n_y(t-T). \tag{3.151}$$

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 σ_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

$$i(t) = |E_1(t)|^2 - |E_2(t)|^2, \tag{3.152}$$

where

$$|E_1(t)|^2 = \left|A + \frac{1}{2}[n_x(t) + n_x(t-T)]\right|^2 + \frac{1}{4}\left|n_y(t) + n_y(t-T)\right|^2, \tag{3.153}$$

and

$$|E_2(t)|^2 = \frac{1}{4}\left|n_x(t) - n_x(t-T)\right|^2 + \frac{1}{4}\left|n_y(t) - n_y(t-T)\right|^2. \tag{3.154}$$

The error probability is equal to

$$p_e = \Pr\{i(t) < 0\} = \Pr\left\{|E_2(t)|^2 > |E_1(t)|^2\right\}. \tag{3.155}$$

Because $n_x(t) + n_x(t-T)$ is independent of $n_x(t) - n_x(t-T)$ and $n_y(t) + n_y(t-T)$ is independent of $n_y(t) - n_y(t-T)$, $|E_1(t)|^2$ 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

$$p_{|E_1|^2}(y) = \frac{1}{2\sigma^2 A}\sqrt{y}\exp\left(-\frac{A^2+y}{2\sigma^2}\right) I_1\left(\sqrt{y}\frac{A}{\sigma^2}\right), \quad y \geq 0, \tag{3.156}$$

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 χ^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

$$p_{|E_2|^2}(y) = \frac{1}{4\sigma^4} y \exp\left(-\frac{y}{2\sigma^2}\right), \quad y \geq 0. \tag{3.157}$$

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

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

$$\begin{aligned} \Pr\left\{|E_2(t)|^2 > y\right\} &= \int_y^{+\infty} p_{|E_2|^2}(y)\mathrm{d}y \\ &= \exp\left(-\frac{y}{2\sigma^2}\right)\left[1 + \frac{y}{2\sigma^2}\right]. \end{aligned} \tag{3.158}$$

The error probability of Eq. (3.155) is

$$p_e = \int_0^{+\infty} \Pr\left\{|E_2(t)|^2 > y\right\} p_{|E_1|^2}(y)\mathrm{d}y. \tag{3.159}$$

Using the p.d.f. of Eq. (3.156) and the probability of Eq. (3.158), after some simplifications, we get

$$p_e = \frac{e^{-2\rho_s}}{\sqrt{2\rho_s}} \int_0^{+\infty} (1+x)\sqrt{x} e^{-2x} I_1\left(2\sqrt{2\rho_s x}\right) \mathrm{d}x, \tag{3.160}$$

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

$$\int_0^{\infty} x^{1/2} e^{-2x} I_1(a\sqrt{x}) \mathrm{d}x = (a/8) e^{a^2/8}, \tag{3.161}$$

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

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

$$p_e = \frac{\exp(-\rho_s)}{2}\left(1 + \frac{\rho_s}{4}\right). \tag{3.163}$$

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 ρ_s. The error probability of synchronous detection of $\frac{1}{2}\mathrm{erfc}\sqrt{\rho_s}$ 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 Azizoğlu (1991), Jacobsen (1993), and Chinn et al. (1996). The DPSK error probability of Eq. (3.105) assumes that there are two noise sources of

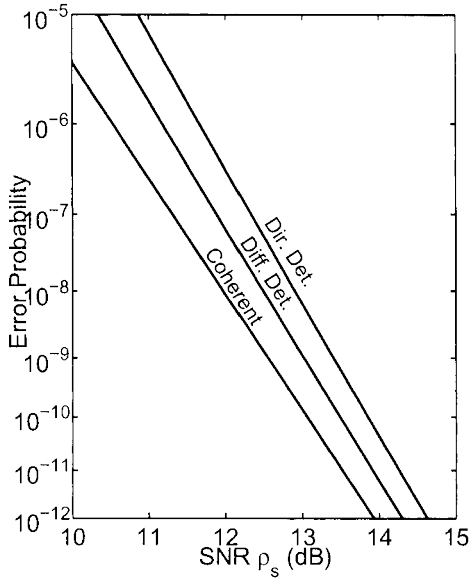

Figure 3.13. The error probability of phase-modulated signals as a function of SNR ρ_s.

$n_1(t)$ 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

$$p_e = \frac{e^{-\rho_s}}{2} \sum_{m=1}^{k} h_k \frac{\rho_s^m}{m!}, \tag{3.164}$$

where

$$h_k = \frac{1}{2^{2(k-1)}} \sum_{l=0}^{k-1} \begin{pmatrix} 2k-l-2 \\ k-1-m \end{pmatrix}. \tag{3.165}$$

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

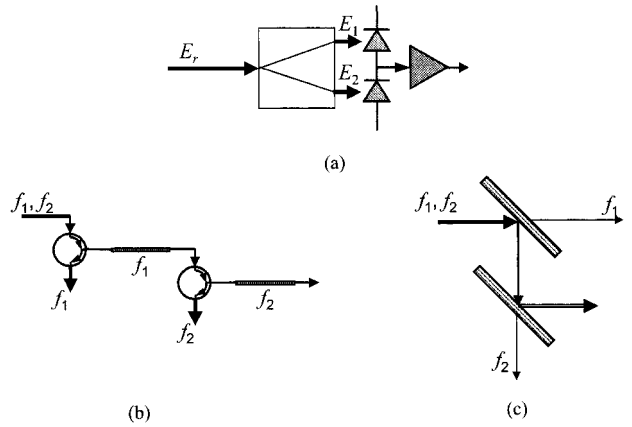

Figure 3.14. Direct-detection FSK receiver using two optical filters. (a) The schematic of the receivers. (b) Optical filter using fiber Bragg gratings. (c) Optical filter using multilayer dielectric filters

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 ω_1 and ω_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,

$$E_1(t) = [A\cos\omega_1 t + n_{x1}(t)]\,\mathbf{x} + n_{y1}(t)\mathbf{y} \tag{3.166}$$

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.

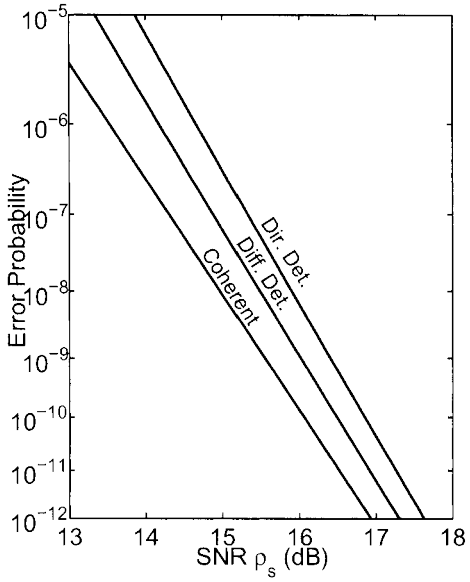

Figure 3.15. The error probability of frequency-modulated signals as a function of SNR ρ_s.

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

$$E_2(t) = n_{x2}(t)\mathbf{x} + n_{y2}(t)\mathbf{y}. \tag{3.167}$$

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

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

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 correlation and the error probability is given by Eq. (3.119) with a correlation coefficient depending on the filter crosstalk. In order to improve the performance, 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 Olsson 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 interferometer of Fig. 3.12, CPFSK signal was directly detected by Idler et al. (2004), Malyon and Stallard (1989) and Toba et al. (1990, 1991). When an optical filter is used to demodulate the FSK signal, it can also function 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 quadrature receiver of Fig. 3.4. From the photocurrent of Eqs. (3.49) and (3.50) with $\omega_{\rm IF} = 0$, including a random phase of θ_0 from either the received signal or the LO signal, we obtain

$$\begin{aligned} r_I(t) &= A_s(t)\cos[\phi_s(t)+\theta_0] + n_I(t), \\ r_Q(t) &= A_s(t)\sin[\phi_s(t)+\theta_0] + n_Q(t), \end{aligned} \tag{3.169}$$

where $A_s(t)$ is due to amplitude modulation, $\phi_s(t)$ from phase modulation, and $n_I(t)$ and $n_Q(t)$ are the identical independently distributed additive Gaussian noise. The random phase of θ_0 in Eq. (3.169) is used to model a receiver without phase locking. In Eq. (3.169), the random phase of θ_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

$$\begin{aligned} r_d(t) &= r_I(t)r_I(t-T) + r_Q(t)r_Q(t-T) \\ &= A^2\cos\left[\phi_s(t) - \phi_s(t-T)\right] + \text{noise terms}. \end{aligned} \tag{3.171}$$

Without noise, $r_d(t)$ is proportional to $\cos\left[\phi_s(t) - \phi_s(t-T)\right]$ and $r_d(t) = \pm A^2$ when $\phi_s(t) - \phi_s(t-T) = 0$ or π, 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

$$r_I(t) = A\cos(\pm\pi\Delta f t) + n_I(t), \tag{3.173}$$

$$r_Q(t) = A\sin(\pm\pi\Delta f t) + n_Q(t), \tag{3.174}$$

where $\Delta f = (\omega_1 - \omega_2)/(2\pi)$ is the frequency difference between the binary FSK signal. The demodulated signal can be

$$\begin{aligned} r_d(t) &= r_I(t)\frac{dr_Q(t)}{dt} + r_Q(t)\frac{dr_I(t)}{dt} \\ &= \mp\pi\Delta f A^2 + \text{noise terms.} \end{aligned} \tag{3.175}$$

The receiver sensitivity increases with the frequency difference of Δf. 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)

$$r_I(t) = A\cos(\pm\pi\Delta f t + \phi_0 + \theta_0) + n_I(t), \tag{3.176}$$

$$r_Q(t) = A\sin(\pm\pi\Delta f t + \phi_0 + \theta_0) + n_Q(t). \tag{3.177}$$

The demodulated signal is

$$\begin{aligned} r_d(t) &= r_I(t)r_Q(t-\tau) + r_Q(t)r_I(t-\tau) \\ &= \pm A^2\sin(\pi\Delta f\tau) + \text{noise terms.} \end{aligned} \tag{3.178}$$

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.

Homodyne phase-diversity receiver was mostly for ASK and DPSK signals (Cheng et al., 1989, Davis et al., 1987, Davis and Wright, 1986, Hodgkinson et al., 1985, 1988, Kazvosky et al., 1987, Okoshi and Cheng, 1987, Smith, 1987). Phase-diversity receiver for FSK and CPFSK modulation is not as popular (Davis et al., 1987, Noé et al., 1988b, Siuzdak and van Etten, 1991, Tsao et al., 1990, 1992). Reviewed by Kazovsky (1989), phase-diversity receiver was also analyzed by Hao and Wicker (1995), Nicholson and Stephens (1989), Siuzdak and van Etten (1989), and Ho and Wang (1995), especially those using 120° optical hybrid.

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

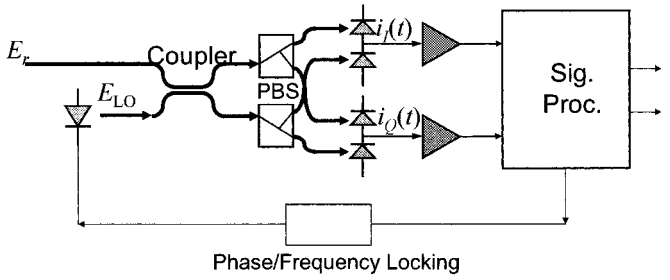

Figure 3.16. A polarization-diversity receiver.

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

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

where the angles of φ and θ are relative to the receiver polarization. Linearly 45° polarized to the receiver polarization, the LO laser has an electric field of

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

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

$$\begin{aligned} E_1(t) &= \frac{1}{\sqrt{2}}\left[E_r(t) + E_{\rm LO}(t)\right] \\ &= \frac{\mathbf{x}}{2}\left[A_s(t)\cos\varphi e^{j\omega_c t + j\phi_s(t)} + A_L e^{j\omega_{\rm LO}t}\right] \\ &\quad + \frac{\mathbf{y}}{2}\left[A_s(t)\sin\varphi e^{j\omega_c t + j\phi_s(t) + j\theta} + A_L e^{j\omega_{\rm LO}t}\right], \end{aligned} \tag{3.181}$$

and

$$\begin{aligned} E_2(t) &= \frac{1}{\sqrt{2}}\left[E_r(t) - E_{\rm LO}(t)\right] \\ &= \frac{\mathbf{x}}{2}\left[A_s(t)\cos\varphi e^{j\omega_c t + j\phi_s(t)} - A_L e^{j\omega_{\rm LO}t}\right] \\ &\quad + \frac{\mathbf{y}}{2}\left[A_s(t)\sin\varphi e^{j\omega_c t + j\phi_s(t) + j\theta} - A_L e^{j\omega_{\rm LO}t}\right]. \end{aligned} \quad (3.182)$$

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

$$i_I(t) = R\cos\varphi A_s(t) A_L \cos[\omega_{\rm IF}t + \phi_s(t)], \quad (3.183)$$

and

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

with a phase difference of θ.

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 θ between Eqs. (3.183) and (3.184). When the phase difference of θ 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

$$r_i(t) = \cos\varphi A_s(t)\cos\phi_s(t) + n_i(t), \quad (3.185)$$

and

$$r_q(t) = \sin\varphi A_s(t)\cos\phi_s(t) + n_q(t), \quad (3.186)$$

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

$$\begin{aligned} r_c(t) &= \cos\varphi r_i(t) + \sin\varphi r_q(t), \\ &= A_s(t)\cos\phi_s(t) + \cos\varphi n_i(t) + \sin\varphi n_q(t), \end{aligned} \tag{3.187}$$

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

$$\begin{aligned} r_c(t) &= r_i(t) + r_q(t) \\ &= (\cos\varphi + \sin\varphi)A_s(t)\cos\phi_s(t) + n_i(t) + n_q(t). \end{aligned} \tag{3.188}$$

The SNR penalty due to equal-gain combining is

$$\delta_p = \frac{(\cos\varphi + \sin\varphi)^2}{2} = \frac{1 + \sin(2\varphi)}{2}, \qquad 0 < \varphi < \frac{\pi}{2} \tag{3.189}$$

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

$$\delta_p = \max(\cos^2\varphi, \sin^2\varphi) = \frac{1}{2}\left(1 + |\cos(2\varphi)|\right), \qquad 0 < \varphi < \frac{\pi}{2} \tag{3.190}$$

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

$$\begin{aligned} r_c(t) &= r_i^2(t) + r_q^2(t) \\ &= A_s^2(t)\cos^2\phi_s(t) + 2A_s(t)\cos\phi_s(t)[n_i(t) + n_q(t)]n_i^2(t) + n_q^2(t). \end{aligned} \tag{3.191}$$

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.

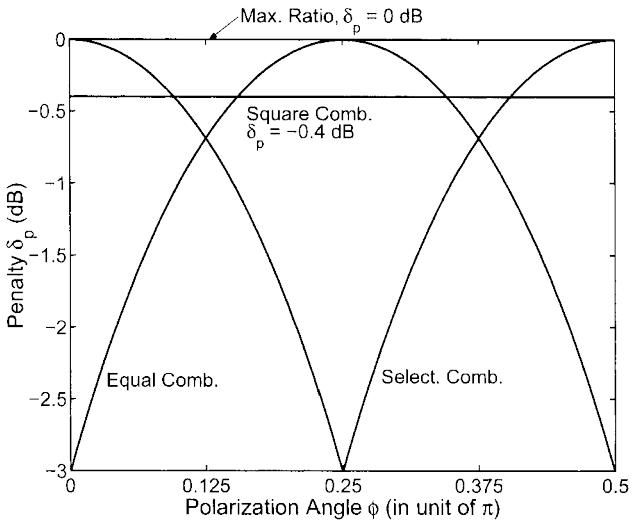

Figure 3.17. The penalty using various schemes to combine the signal from two polarization-diversified signals.

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 φ. Square combination has a penalty of 0.4 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

$$\cos\varphi r_i(t) + \sin\varphi r_q(t) \quad (3.194)$$

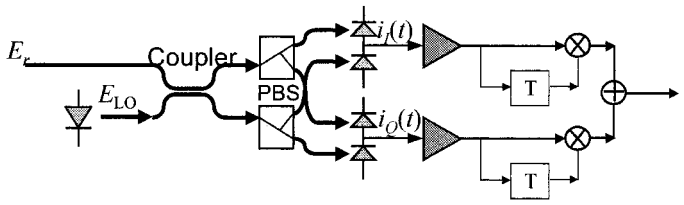

Figure 3.18. A polarization-diversity receiver for DPSK and MSK signals.

and

$$\sin\varphi r_i(t) - \cos\varphi r_q(t), \tag{3.195}$$

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

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

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

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

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

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

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 θ 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.

Polarization diversity has been considered mostly for DPSK and FSK signals using baseband combination (Cline et al., 1990, Glance, 1987, Imai et al., 1991, Kavehrad and Glance, 1988, Okoshi and Cheng, 1987, Ryu et al., 1987, 1991a, Shibutani and Yamazaki, 1989). Most early field trials of coherent optical communications used polarization diversity with frequency modulated signal (Bødker et al., 1991, Imai et al., 1990a, Ryu et al., 1988, 1992). In additional to polarization alignment using APC and polarization-diversity, polarization scrambling provides random polarization within the bit interval (Caponio et al., 1991, Cimini Jr. et al., 1988, Habbab and Cimini Jr., 1988, Hodgkinson et al., 1987, Meada and Smith, 1990). Not compatible with PDM, polarization scrambling has a power penalty of 3 dB due to the loss of signal to the polarization orthogonal to the receiver polarization.

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

$$E_r(t) = [E_x(t)\mathbf{x} + E_y(t)\mathbf{y}]\, e^{j\omega_c t}. \tag{3.199}$$

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 $\mathbf{x}$.

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

$$s_1(t) = A\mathbf{x}e^{j\omega_c t}, \tag{3.200}$$

and

$$s_2(t) = A\mathbf{y}e^{j\omega_c t} \tag{3.201}$$

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 $\mathbf{x}$ and $\mathbf{y}$. 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}.

Modulation Format	Sensitivity (photons/bit)		Penalty
	Shot-Noise	Amplifier-Noise	(dB)
Homodyne PSK	9	18	0
Heterodyne PSK	18	18	0
Heterodyne DPSK, MSK	20	20	0.45
Direct-Detection DPSK	–	22	0.85
Homodyne ASK	18	36	3
Heterodyne ASK, FSK	36	36	3
Envelope Detection Heterodyne ASK	39	39	3.4
Direct-Detection ASK, FSK	–	40	3.5
Dual-Filter FSK, PolSK	40	40	3.5
Single-Filter FSK	80	80	6.5

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

$$p(r) = \frac{r}{\sigma_n^2} I_0\left(\frac{rA}{\sigma_n^2}\right) \exp\left(-\frac{r^2 + A^2}{2\sigma_n^2}\right). \tag{3.A.1}$$

The cumulative distribution function of Rice distribution is

$$\int_x^{+\infty} p(r)\mathrm{d}r = Q\left(\frac{A}{\sigma_n}, \frac{x}{\sigma_n}\right), \tag{3.A.2}$$

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

$$Q(a,b) = \int_b^\infty x I_0(ax) \exp\left(-\frac{x^2+a^2}{2}\right) \mathrm{d}x, \tag{3.A.3}$$

or

$$Q\left(\sqrt{2a}, \sqrt{2b}\right) = \int_b^\infty e^{-(a+x)} I_0\left(2\sqrt{ax}\right) \mathrm{d}x, \tag{3.A.4}$$

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

$$Q(0,b) = e^{-b^2/2}, \tag{3.A.5}$$

$$Q(a,0) = 1. \tag{3.A.6}$$

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

$$I_0(2\sqrt{ab}) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{1}{p} \exp\left(ap + \frac{b}{p}\right) \mathrm{d}p, \qquad c > 0 \tag{3.A.7}$$

where c is a real positive number. Substitute Eq. (3.A.7) into the Marcum Q function of Eq. (3.A.3) and exchange the order of integration, we get

$$\begin{aligned} Q\left(\sqrt{2a}, \sqrt{2b}\right) &= \frac{e^{-a}}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{1}{p} \int_b^\infty e^{-x} \exp\left(ap + \frac{x}{p}\right) \mathrm{d}x\mathrm{d}p \\ &= e^{-(a+b)} \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{\exp\left(ap + \frac{b}{p}\right)}{p-1} \mathrm{d}p, \qquad c > 1 \end{aligned} \tag{3.A.8}$$

or

$$Q\left(\sqrt{2a}, \sqrt{2b}\right) = e^{-(a+b)} \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{\exp\left(bp + \frac{a}{p}\right)}{p(p-1)} \mathrm{d}p, \qquad 0 < c < 1 \tag{3.A.9}$$

and

$$Q(a,b) = -e^{-(a^2+b^2)/2} \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{\exp\left(\frac{b^2 p}{2} + \frac{a^2}{2p}\right)}{p(p-1)} \mathrm{d}p, \qquad 0 < c < 1 \tag{3.A.10}$$

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

$$Q(a,b) + Q(b,a) = 1 + e^{-(a^2+b^2)/2} I_0(ab), \tag{3.A.11}$$

$$Q(a,a) = \frac{1}{2} + \frac{1}{2} e^{-a^2} I_0(a^2). \tag{3.A.12}$$

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

$$Q(a,b) = e^{-(a^2+b^2)/2} \sum_{m=0}^{\infty} \left(\frac{a}{b}\right)^m I_m(ab), \qquad b > a \tag{3.A.13}$$

$$= 1 - e^{-(a^2+b^2)/2} \sum_{m=0}^{\infty} \left(\frac{b}{a}\right)^m I_m(ab), \qquad a > b. \tag{3.A.14}$$

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

$$p_{R_1}(r_1) = \frac{r_1}{\sigma_1^2} \exp\left(-\frac{A_1^2 + r_1^2}{2\sigma_1^2}\right) I_0\left(\frac{A_1 r_1}{\sigma_1^2}\right), \tag{3.A.15}$$

$$p_{R_2}(r) = \frac{r_2}{\sigma_2^2} \exp\left(-\frac{A_2^2 + r_2^2}{2\sigma_2^2}\right) I_0\left(\frac{A_2 r_2}{\sigma_2^2}\right), \tag{3.A.16}$$

the error probability is

$$\begin{aligned} p_e &= \Pr\{R_1^2 < R_2^2\} = \Pr\{R_1 < R_2\} \\ &= \int_0^\infty p_{R_1}(r_1) \int_{r_1}^\infty p_{R_2}(r_2) \mathrm{d}r_2 \mathrm{d}r_1 \\ &= \int_0^\infty p_{R_1}(r_1) Q\left(\frac{A_2}{\sigma_2}, \frac{r_1}{\sigma_2}\right) \mathrm{d}r_1. \end{aligned} \tag{3.A.17}$$

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

$$\begin{aligned} p_e = \exp\left[-\frac{1}{2}\left(\frac{A_1^2}{\sigma_1^2} + \frac{A_2^2}{\sigma_2^2}\right)\right] \cdot \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{\exp\left(\frac{A_1^2}{2\sigma_2^2} p\right)}{p-1} \\ \times \frac{1}{1 + \frac{\sigma_1^2}{\sigma_2^2}\left(1 - \frac{1}{p}\right)} \exp\left[\frac{A_1^2}{2\sigma_1^2} \frac{1}{1 + \frac{\sigma_1^2}{\sigma_2^2}\left(1 - \frac{1}{p}\right)}\right] \mathrm{d}p, \quad c > 1. \end{aligned} \tag{3.A.18}$$

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

$$p_e = Q(a,b) - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} e^{-(a^2+b^2)/2} I_0(ab), \tag{3.A.19}$$

where

$$a^2 = \frac{A_2^2}{\sigma_1^2 + \sigma_2^2}, \qquad b^2 = \frac{A_1^2}{\sigma_1^2 + \sigma_2^2}. \tag{3.A.20}$$

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

$$p_e = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} [1 - Q(b,a)] + \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} Q(a,b), \tag{3.A.21}$$

and

$$p_e = \frac{1}{2} [1 - Q(b,a) + Q(a,b)] - \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2} e^{-(a^2+b^2)/2} I_0(ab). \tag{3.A.22}$$

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.