As shown in the helix shape scattergram of Figs. 5.1, nonlinear phase noise is correlated with the received intensity. The received intensity can be used to compensate for the nonlinear phase noise. Ideally, the optimal compensator should minimize the error probability of the system after compensation. However, the optimal compensator or detector given by the maximum a posteriori probability (MAP) criterion to minimize the error probability may be difficult to find and design. The minimum mean-square error (MMSE) compensator is always analytically simple to find and leads to practical implementation.

In this chapter, the linear compensator is optimized in term of the variance, or MMSE criterion, of the residual nonlinear phase noise. The linear MMSE compensator is derived when a single compensator preceding the receiver or in the middle of the fiber span. The MMSE compensator is generalized to the cases when many compensators are located in the fiber link for mid-span or distributed compensation.

The exact error probability with a single linear compensator preceding the receiver is derived analytically. With a simple expression to calculate the error probability, numerical optimization is used to find the linear MAP compensator to minimize the error probability. While linear MAP compensator gives improvement over MMSE compensator for phase-shift keying (PSK) signals, linear MMSE compensator performs close to linear MAP compensator for differential phase-shift keying (DPSK) signals.

When the joint distribution of received amplitude and phase is derived analytically for signal with nonlinear phase noise. The optimal MAP detector can be derived to minimize the error probability of the signal after the detector. Based on the joint distribution of received amplitude
and phase, the error probability of nonlinear MMSE or MAP detector can also be evaluated numerically for PSK signals.

1. Electronic Compensator for Nonlinear Phase Noise

The received intensity $P_N = R_r^2$ may be used to compensate the nonlinear phase noise of Eq. (5.5), where the received amplitude is

$$R_r = |E_0 + n_1 + \cdots + n_N|.$$  \hfill (6.1)

The received amplitude of Eq. (6.1) ignores the fiber loss of the last fiber span and the required optical amplifier to compensate for it. The actual received electric field is

$$E_r = (E_0 + n_1 + \cdots + n_N) \exp(-j\Phi_{NL})$$ \hfill (6.2)

with nonlinear phase rotation of $\Phi_{NL}$ as shown in Figs. 5.1. In the received electric field of Eq. (6.2), the nonlinear phase noise of Eq. (5.5) does not change the received amplitude $R_r$ of Eq. (6.1). For binary PSK and DPSK signals, only the real part of the received electric field or the corresponding differential component is required to detect the transmitted data. Both real and imaginary parts of the electric field of $E_r$ can be measured by an optical phase-locked loop (PLL) for PSK signal using the quadrature receiver of Fig. 3.4. The corresponding differential components can be measured by two interferometers for DPSK signal. In order to compensate the nonlinear phase noise, either the received phase or the complex representation of the received electric field of $E_r$ must be measured.

In the simplest nontrivial example of two fiber spans, the received electric field in linear regime is $E_2 = E_0 + n_1 + n_2$. The nonlinear phase shift is $\Phi_{NL} = \gamma L_{eff} \left( |E_0 + n_1|^2 + |E_0 + n_1 + n_2|^2 \right)$, and the received electric field with nonlinear phase shift is $E_2 \exp(-j\Phi_{NL})$. Hence, the received phase is $\Phi_2 = \theta_2 - \Phi_{NL} = \theta_2 - \gamma L_{eff} \left( |E_0 + n_1|^2 + |E_0 + n_1 + n_2|^2 \right)$, where $\theta_2$ is the phase of $E_2 = |E_2|e^{j\theta_2}$. Because the received signal intensity $P_2 = |E_2|^2 = |E_0 + n_1 + n_2|^2$ is correlated with $\Phi_{NL}$, the impact of the nonlinear phase shift can be reduced by adding a correction term of $\gamma L_{eff} P_2$ to the received phase. The corrected estimation of the received phase is $\Phi_2 + \gamma L_{eff} P_2 = \theta_2 - \gamma L_{eff} |E_0 + n_1|^2$. In fact, the impact of the nonlinear phase noise can be reduced further, because the remaining nonlinear phase shift of $\gamma L_{eff} |E_0 + n_1|^2$ is still correlated with the received intensity $P_2$. As shown below, the optimal correction term is about $1.5 \gamma L_{eff} P_2$.

Figure 6.1(a) shows an example of a homodyne optical receiver using an optical PLL to detect both in-phase and quadrature components.
Figure 6.1. (a) Typical coherent receiver detecting both in-phase and quadrature components of the received electric field $E_r$ (see also Fig. 3.4). (b) Nonlinear phase noise is compensated by using a spiral-boundary decision device. (c) Nonlinear phase noise is compensated by using a compensator, followed by a straight-boundary decision device. [From Ho and Kahn (2004a), © 2005 IEEE]

of the received electric field $E_r$, similar to the quadrature receiver of Fig. 3.4. A $90^\circ$ optical hybrid is used to combine the signal with a phase-locked local oscillator (LO) laser, yielding four combinations with relative phase shifts of $0^\circ$, $180^\circ$, $90^\circ$ and $270^\circ$. A pair of balanced photodetectors provides in-phase and quadrature photocurrents, $i_I$ and $i_Q$, equivalently the two components of $\cos \Phi_r$ and $\sin \Phi_r$ or the whole complex number of $E_r = |E_r| e^{j\Phi_r}$. In systems where the dominant noise source is amplified spontaneous emission from optical amplifiers, a synchronous heterodyne optical receiver can also be used without loss of
An alternative method, shown in Fig. 6.1(c), employs a compensator that subtracts from the received phase a correction as a function of the received intensity. The compensated phase, depending on the received intensity, is the central phase of the decision region of either zero or one of Fig. 6.2(a). The compensator is followed by a straight-boundary decision device. Figure 6.2(b) shows the distribution of the residual phase noise after the compensator, together with the straight-boundary deci-
sion boundary as the $y$-axis. The compensator phase can be a linear function of the received intensity, as suggested by Liu et al. (2002b), Xu et al. (2003), and Ho and Kahn (2004a) or a nonlinear function of the received intensity, as suggested by Ho (2003d) and Appendix 6.A. Theoretically, the optimal compensator should be nonlinear compensator without linearity constraint.

Figure 6.3 shows a schematic diagram to compensate nonlinear phase noise using a phase modulator, first suggested by Xu and Liu (2002) and implemented in both Xu et al. (2002) and Hansryd et al. (2004, 2005). A passive fiber tap is used to split out part of the optical signal, amplified using both a TIA and driver amplifier to drive a phase modulator. The modulated phase of the phase modulator cancels part of the nonlinear phase noise in the fiber link. Both linear and nonlinear compensators can be implemented using a phase modulator of Fig. 6.3 using either a linear or nonlinear driver amplifier. While most external modulators are polarization sensitive, polarization independent modulator is also available (Chen et al., 1992, 1997a, Sugiyama, 2003). If polarization dependent modulator is used in Fig. 6.3, automatic polarization control (APC) is required preceding the phase modulator.

While the electronic compensator of Fig. 6.1(c) or the Yin-Yang detector of Fig. 6.1(b) is used in the receiver, the electro-optical compensator of Fig. 6.3 can be used in the middle of a fiber link for distributed compensation of nonlinear phase noise. Both electronic and electro-optical implementations of Figs. 6.1 and 6.3 can be used as both linear and nonlinear compensator, some nonlinear optical devices can be used solely as a linear compensator (Liu et al., 2002b).
2. Linear MMSE Compensator for Finite Number of Fiber Spans

The simplest compensator is a linear compensator to minimize the variance of the residual nonlinear phase noise. Although the probability density function (p.d.f.) of nonlinear phase noise was derived in Chapter 5 for finite and infinite number of fiber spans, this type of MMSE compensator can be derived without using the p.d.f. or joint characteristic between nonlinear phase noise and received intensity but the correlation between them. In this section, we find either the decision boundary in Fig. 6.2(a) or the corresponding compensated curve to minimize the variance of the residual nonlinear phase noise. In terms of the variance of residual nonlinear phase noise, linear and nonlinear MMSE compensators do not have significant difference.

2.1 Minimum Mean-Square Error Compensation

In an $N_A$-span system, the linear MMSE compensator can be derived by finding a scale factor $\alpha$ to minimize the variance of the residual nonlinear phase noise of $\Phi_{NL} - \alpha P_N$. The corrected phase estimation is $\Phi_r + \alpha P_N$, where $\Phi_r$ is the phase of the received electric field $E_r$ of Eq. (6.2).

The intensity of signal plus noise has a noncentral chi-square ($\chi^2$) distribution. The MMSE compensator minimizes the variance that depends on the covariance between its compositions. First, we consider a simple mathematical problem to find the covariance between two $\chi^2$ distributed random variables. For a real amplitude of $A = |E_0|$ and two independent complex circular Gaussian distributed random variables $\zeta_1$ and $\zeta_2$, both $|A + \zeta_1|^2$ and $|A + \zeta_1 + \zeta_2|^2$ are noncentral $\chi^2$ distributed random variables with two degrees of freedom. The mean and variance of $|A + \zeta_1|^2$ are

$$E\{|A + \zeta_1|^2\} = A^2 + 2\sigma^2_{\zeta_1},$$

$$\sigma^2_{|A+\zeta_1|^2} = f(\sigma^2_{\zeta_1}) = 4A^2\sigma^4_{\zeta_1} + 4\sigma^4_{\zeta_1},$$

where $E\{\zeta_1^2\} = 2\sigma^2_{\zeta_1}$ is the variance of $\zeta_1$. In Eq. (6.4), the variance of the random variable $|A + \zeta_1|^2$ is defined as a function of $f(\sigma^2_{\zeta_1})$. Without going into detail, the covariance between $|A + \zeta_1|^2$ and $|A + \zeta_1 + \zeta_2|^2$ is

$$E\left\{(|A + \zeta_1|^2 - E\{|A + \zeta_1|^2\}) (|A + \zeta_1 + \zeta_2|^2 - E\{|A + \zeta_1 + \zeta_2|^2\})\right\} = 4A^2\sigma^2_{\zeta_1} + 4\sigma^4_{\zeta_1} = f(\sigma^2_{\zeta_1}).$$

The covariance relationship of Eq. (6.5) is obvious, the random variable of $|A + \zeta_1|^2$ does not depend on, and is not correlated with, the
random variable of $\zeta_2$. Note that $\sigma_{\zeta_1}^2$ is the covariance between $\zeta_1$ and $\zeta_1 + \zeta_2$.

Using the expressions of Eqs. (6.4) and (6.5), the variance of the nonlinear phase noise of Eq. (5.5) is found to be

$$\sigma_{\Phi_{NL}}^2 = (\gamma L_{\text{eff}})^2 \left[ \sum_{k=1}^{N_A} f(k \sigma^2) + 2 \sum_{k=1}^{N_A} (N_A - k) f(k \sigma^2) \right].$$

(6.6)

The first summation of Eq. (6.6) corresponds to $\zeta_1 = n_1 + \ldots + n_k$ with $\sigma_{\zeta_1}^2 = k \sigma^2$ for all terms of $|E_0 + n_1 + \ldots + n_k|^2$. The second summation of Eq. (6.6) corresponds to all the covariance terms between $|E_0 + n_1 + \ldots + n_k|^2$ and $|E_0 + n_1 + \ldots + n_{k+1}|^2 + \ldots + |E_0 + n_1 + \ldots + n_{N_A}|^2$ with $\zeta_1 = n_1 + \ldots + n_k$ and $\zeta_2 = n_{k+1} + \ldots + n_{N_A}$. Substituting Eq. (6.4) to (6.6), we obtain

$$\sigma_{\Phi_{NL}}^2 = \frac{4}{3} N_A (N_A + 1) (\gamma L_{\text{eff}} \sigma_0)^2 \left[ \left( N_A + \frac{1}{2} \right) |E_0|^2 + (N_A^2 + N_A + 1) \sigma_0^2 \right].$$

(6.7)

Similar to Eq. (6.6), the variance of the residual nonlinear phase noise of $\Phi_{NL} - \alpha P_N$, is found using Eqs. (6.4) and (6.6) to be

$$\sigma_{\Phi_{NL} - \alpha P_N}^2 = (\gamma L_{\text{eff}})^2 \left[ \sigma_{\Phi_{NL}}^2 + (\alpha - 1)^2 f(N_A \sigma^2) - 2(\alpha - 1) \sum_{k=1}^{N_A-1} f(k \sigma^2) \right].$$

(6.8)

The optimal scale factor can be found by solving $d\sigma_{\Phi_{NL} - \alpha P_N}^2 / d\alpha = 0$ to obtain

$$\alpha_{\text{mse}} = 1 + \frac{\sum_{k=1}^{N_A-1} f(k \sigma^2)}{f(N_A \sigma^2)}.$$

(6.9)

Using Eq. (6.4), the optimal scale factor is found to be

$$\alpha_{\text{mse}} = \gamma L_{\text{eff}} \frac{N_A + 1}{2} \cdot \frac{|E_0|^2 + (2N_A + 1) \sigma_0^2}{|E_0|^2 + N_A \sigma_0^2}/3.$$

(6.10)

At high signal-to-noise ratio (SNR), we obtain

$$\alpha_{\text{mse}} \approx \gamma L_{\text{eff}} \frac{N_A + 1}{2}.$$

(6.11)

The variance of the residual nonlinear phase noise is reduced to

$$\sigma_{\Phi_{NL} - \alpha_{\text{mse}} P_N}^2 = (N_A - 1) N_A (N_A + 1) (\gamma L_{\text{eff}} \sigma_0)^2 \times \frac{|E_0|^4 + 2N_A \sigma_0^2 |E_0|^2 + (2N_A^2 + 1) \sigma_0^4}{3(|E_0|^2 + N_A \sigma_0^2)}. $$

(6.12)
The mean residual nonlinear phase shift is
\[ \langle \Phi_{\text{RES}} \rangle = \langle \Phi_{\text{NL}} \rangle - \alpha_{\text{mse}}(|E_0|^2 + 2N_A\sigma_0^2). \] (6.13)

In high SNR, similar to Eq. (5.80), the variance of nonlinear phase noise of Eq. (6.7) can be approximated as
\[ \sigma_{\Phi_{\text{NL}}}^2 \approx \frac{4}{3}(\gamma L_{\text{eff}}\sigma_0|E_0|)^2 \approx \frac{2}{3}\langle \Phi_{\text{NL}} \rangle^2. \] (6.14)

and the variance of residual nonlinear phase noise of Eq. (6.12) can be approximated as
\[ \sigma_{\Phi_{\text{NL}}-\alpha_{\text{PN}}}^2 \approx \frac{1}{3}(\gamma L_{\text{eff}}\sigma_0|E_0|)^2 \approx \frac{\langle \Phi_{\text{NL}} \rangle^2}{6\rho_s}, \] (6.15)

where the mean nonlinear phase shift is
\[ \langle \Phi_{\text{NL}} \rangle = N_A\gamma L_{\text{eff}}[|E_0|^2 + (N_A + 1)\sigma_0^2] \approx N_A\gamma L_{\text{eff}}|E_0|^2. \] (6.16)

To derive Eqs. (6.10), (6.7), and (6.12), the following relationships are used:
\[ \sum_{k=1}^{N} k^2 = \frac{1}{6}N(2N + 1)(N + 1), \] (6.17)
\[ \sum_{k=1}^{N} (N - k)^2 = \frac{1}{6}N(2N - 1)(N - 1), \] (6.18)
\[ \sum_{k=1}^{N} k^2(N - k) = \sum_{k=1}^{N} k(N - k)^2 = \frac{1}{12}N^2(N^2 - 1), \] (6.19)
\[ \sum_{k=1}^{N} k(N - k) = \frac{1}{6}N(N^2 - 1). \] (6.20)

Comparing the approximated variances of Eqs. (6.14) and (6.15), using linear compensation, the variance of nonlinear phase noise is reduced by a factor of 4.

At high SNR, the variance of the phase of amplifier noise is \( \sigma_{\Phi_{\text{a}}}^2 \approx 1/2\rho_s \) from Eq. (4.A.15). Gordon and Mollenauer (1990) gave an insight that the optimal operating point can be found by choosing \( \sigma_{\Phi_{\text{a}}}^2 = \sigma_{\Phi_{\text{NL}}}^2 \) or \( \langle \Phi_{\text{NL}} \rangle = \sqrt{0.75} = 0.866 \) rad for both PSK and DPSK signals. Gordon and Mollenauer (1990) actually gave an estimation of \( \langle \Phi_{\text{NL}} \rangle = 1 \) rad as a simple approximated result. Proportional to the mean nonlinear phase...
shift of \(\langle \Phi_{NL} \rangle\), the standard deviation (STD) of the residual nonlinear phase noise is reduced by a factor of 2 using linear compensator and enable the maximum \(\langle \Phi_{NL} \rangle\) increases by a factor of 2. Because the mean nonlinear phase shift of Eq. (6.16) is proportional to the number of fiber spans, as shown in Eq. (6.16), doubling the mean nonlinear phase shift doubles the number of fiber spans, and thus doubles the transmission distance, assuming that nonlinear phase noise is the primary limitation.

Linear compensator for nonlinear phase noise was first independently proposed by Liu et al. (2002b) using special nonlinear optical devices, Xu and Liu (2002) using a phase modulator, and Ho and Kahn (2004a) using electronic circuits. To use a scale version of the received intensity to compensate for nonlinear phase noise may be the simplest implementation.

While the foregoing discussion has focused on binary PSK, DPSK signals have generated much more interest recently from Table 1.2. In a DPSK system, information is encoded in phase differences between successive symbols, and is decoded using equivalently the differential phase of \(\Delta \Phi_r = \Phi_r(t) - \Phi_r(t-T)\), where \(T\) is the symbol interval. When the differential phase is corrupted by the nonlinear phase noise difference of \(\Phi_{NL}(t+T) - \Phi_{NL}(t)\), the impact of nonlinear phase noise can be compensated by the linear combination of \(\Delta \Phi_r + \alpha \Delta P_N\), where \(\Delta P_N = P_N(t) - P_N(t-T)\) is the power difference between consecutive symbols. The optimal scale factor for DPSK systems is precisely analogous to that for PSK systems, and optimal compensation also approximately doubles the transmission distance.

Using electronic circuits of Fig. 6.1, there are two implementations of the linear compensator (or any compensators). Using the received electric field shown in Figs. 5.1, a Yin-Yang detector can be implemented based on the spiral-boundary decision device. Alternatively, when both received phase \(\Phi_r\) and intensity \(P_N\) are given, a simple linear combiner of \(\Phi_r + \alpha P_N\) can be used as a compensator.

A practical quadrature receiver of Fig. 3.4 may yield the in-phase and quadrature components of \(\cos \Phi_r\) and \(\sin \Phi_r\). Instead of the corrected phase of \(\Phi_r + \alpha P_N\), the corrected quadrature components can be calculated using electronic signal processing techniques, as \(\cos(\Phi_r + \alpha P_N) = \sin \Phi_r \sin(\alpha P_N) - \cos \Phi_r \cos(\alpha P_N)\) and \(\sin(\Phi_r + \alpha P_N) = \sin \Phi_r \cos(\alpha P_N) + \cos \Phi_r \sin(\alpha P_N)\). Compensator for DPSK signal can also be implemented similarly.

To compensate for nonlinear phase noise, the electro-optic implementation of Fig. 6.3 and all-electronic implementation of Figs. 6.1 can use electronic signal processing to provide arbitrary linear and nonlinear functions. To correct the nonlinear phase noise, a fundamental prob-
The problem is to find the optimal MMSE compensator without the constraint of linearity. It is also important to verify whether the linear MMSE compensator is close to the optimal MMSE compensator. Appendix 6.A derives the optimal nonlinear compensator to minimize the variance of nonlinear phase noise after the compensator.

Figure 6.4 shows the simulated phase noise STD as a function of mean nonlinear phase shift with and without compensation. For all 5000 points in Figs. 5.1, the STD of the nonlinear phase noise $\Phi_{NL}$ of Eq. (5.5), denoted as $\sigma_{\Phi_{NL}}$, is calculated. When the nonlinear phase noise is compensated using a linear compensator of $\alpha_{\text{inse}}P_N$ proportional to the received intensity of $P_N$, the STD of the residual nonlinear phase noise of $\Phi_{NL} - \alpha_{\text{inse}}P_N$ is calculated and denoted as $\sigma_{\Phi_{NL}-\alpha P_N}$, where $\alpha_{\text{inse}}$ is the scale factor of Eq. (6.10). When the nonlinear phase noise is compensated using the nonlinear MMSE compensator $E\{\Phi_{NL}|R_r\}$ of Eq. (6.A.15) of Appendix 6.A, the STD of the residual nonlinear phase noise of $\Phi_{NL} - E\{\Phi_{NL}|R_r\}$ is calculated and denoted as $\sigma_{\Phi_{NL}-E\{\Phi_{NL}|R_r\}}$.

Figure 6.4 also plots the approximation of

$$\sigma_{\Phi_{NL}} \approx 0.1925 \langle \Phi_{NL} \rangle \quad \text{and} \quad \sigma_{\Phi_{NL}-\alpha R_r^2} \approx 0.0962 \langle \Phi_{NL} \rangle. \quad (6.21)$$

given by Eqs. (6.14) and (6.15), respectively, as solid lines for $\rho_s = 18$ (12.6 dB).
Figure 6.4 shows that there is almost no difference between the STD of the residual nonlinear phase noise of $\sigma_{\Phi_{NL} - \alpha P_N}$ using the scale factor given by Eq. (6.10) and the STD of the residual nonlinear phase noise of $\sigma_{\Phi_{NL} - E\{\Phi_{NL}|R_r\}}$ using the nonlinear MMSE compensator of Eq. (6.A.15). In general, $\sigma_{\Phi_{NL} - E\{\Phi_{NL}|R_r\}}$ is smaller than $\sigma_{\Phi_{NL} - \alpha R^2}$ by about only 0.2%. Figure 6.4 also shows that the approximation of Eq. (6.21) is also very accurate. The linear compensator, if optimized, is very close to the optimal nonlinear MMSE compensator.

As shown in Fig. 6.4, the variances of the residual nonlinear phase noise of the linear and optimal MMSE compensators have little difference. When linear and optimal MMSE compensators are used, the p.d.f. of the residual nonlinear phase noise may not be the same, especially the tail probabilities. With the approximation of Eq. (6.A.11) from the Appendix 6.A, the p.d.f. of the residual nonlinear phase noise is Gaussian distributed. In practice, the residual nonlinear phase noise has a larger tail probability than that of the Gaussian distribution. While it is possible to derive the p.d.f. of the residual nonlinear phase noise for the linear compensator, that for nonlinear compensator is difficult to derive.

The usage of linear compensation to reduce nonlinear phase noise was proposed by Liu et al. (2002b), Xu and Liu (2002), and Ho and Kahn (2004a). The reduction of the variance of nonlinear phase noise by a factor of 4 was first given by the observation of Liu et al. (2002b) and Xu and Liu (2002) by simulation and Ho and Kahn (2004a) by analysis.

### 2.2 Probability Density of Residual Nonlinear Phase Noise

The probability density of residual nonlinear phase noise can also be derived using method similar to that in Sec. 5.1.2. Using a notation similar to Sec. 5.1.2 by ignoring the constant fact of $\gamma L_{eff}$, with the linear correction of the phase noise using received intensity, the residual nonlinear phase noise is

$$
\Phi_{\text{RES}} = |E_0 + n_1|^2 + |E_0 + n_1 + n_2|^2 + \ldots + |E_0 + n_1 + \ldots + n_{N_A-1}|^2 - (\alpha_{\text{mse}} - 1)|E_0 + n_1 + \ldots + n_{N_A}|^2, \quad (6.22)
$$

As from Eq. (6.10), $\alpha_{\text{mse}} \approx (N_A + 1)/2$ is the optimal scale factor to correct the nonlinear phase noise of Eq. (5.5) using the received intensity of $P_N$. The random variable corresponding to $\varphi_1$ of Eq. (5.13) becomes

$$
(N_A - \alpha_{\text{mse}}) A^2 + 2Aw_r^T \bar{x} + \bar{x}^T C_r \bar{x}, \quad (6.23)
$$
where \( \tilde{w}_r = \tilde{w} - \alpha_{\text{mse}} \tilde{w}_I \) with \( \tilde{w}_I = (1, 1, \ldots, 1) \) and

\[
C_r = (M - L)^T(M - L) - (\alpha_{\text{mse}} - 1)C_r^T L,
\]

where

\[
L = \begin{bmatrix}
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 \\
1 & 1 & \cdots & 1 & 1
\end{bmatrix}.
\]

(6.25)

Following the procedure from Eqs. (5.13) to (5.22), the characteristic function of \( \Phi_{\text{RES}} \) is

\[
\Psi_{\Phi_{\text{RES}}} (\nu) = \frac{\exp \left[ -j\nu(N_A - \alpha_{\text{mse}})A^2 - 2\sigma_0^2\nu^2A^2\tilde{w}_r^T(I - 2j\nu\sigma_0^2C_r)^{-1}\tilde{w}_r \right]}{\det [I - 2j\nu\sigma_0^2C_r]}.
\]

(6.26)

The characteristic functions of \( \Phi_{\text{RES}} \) in the form of eigenvalues and eigenvectors are similar to that of Eqs. (5.23) and (5.24). The characteristic functions of \( \Phi_{\text{RES}} \) has the same expression as Eq. (5.24) using a new set of eigenvalues and eigenvectors of the covariance matrix \( C_r \) and the vector of \( \tilde{w}_r \). The characteristic function of Eq. (6.26) is a special case of Eq. (5.24).

Except the first and last rows, the matrix \( C_r^{-1} \) is also approximately a Toeplitz matrix for the series of \( 2, -1, 0, \ldots \) For large number of fiber spans of \( N_A \), the eigenvalues of \( C_r \) are asymptotically equal to

\[
\frac{1}{\lambda_k} \approx 4 \sin^2 \left[ \frac{(k - 1.25)\pi}{2(N_A - 1)} \right], \quad k = 2, \ldots, N_A, \quad \lambda_1 \approx -\sum_{k=2}^{N_A} \lambda_k.
\]

(6.27)

Other than the largest one in absolute value, the eigenvalues of \( C_r \) are all positive. All eigenvalues of the covariance matrix \( C_r \) sum to approximately zero and multiple to \( \alpha_{\text{mse}} - 1 \approx (N_A - 1)/2 \).

The p.d.f. of the residual nonlinear phase noise of \( \Phi_{\text{RES}} \) of Eq. (6.22) can be calculated by taking the inverse Fourier transform of the corresponding characteristic functions of \( \Psi_{\Phi_{\text{RES}}} (\nu) \) [Eq. (6.26)]. Figure 6.5 shows the p.d.f. of \( \Phi_{\text{RES}} \). The same as Fig. 5.2, Fig. 6.5 is plotted for the case that the SNR of \( \rho_s = A^2/(2N_A\sigma_0^2) \) = 18 (12.6 dB), corresponding to an error probability of \( 10^{-9} \) for PSK signals if the amplifier noise is the only impairment as shown in Eq. (3.78) and Fig. 3.13. The number of spans is \( N_A = 32 \). The x-axis is normalized with respect to \( N_A A^2 \), approximately equal to the mean nonlinear phase shift.
Similar to $\Phi_{NL}$, the random variables of $\Phi_{RES}$ can be modeled as the combination of $N_A = 32$ independently distributed noncentral $\chi^2$-random variables with two degrees of freedom. Some studies implicitly assume a Gaussian distribution by using the $Q$-factor to characterize the random variables. While the Gaussian assumption for $\Phi_{RES}$ may not be valid, other than the two noncentral $\chi^2$-random variables corresponds to the two largest eigenvalues, the other random variables should sum to Gaussian distribution. By modeling the summation of random variables with smaller eigenvalues as Gaussian distribution, the nonlinear phase noise of Eq. (6.26) can be modeled as a summation of three instead of $N_A = 32$ independently distributed random variables.

From Fig. 6.5, the p.d.f. of $\Phi_{RES}$ has significant difference with that of a Gaussian distribution. Figure 6.6 divides the p.d.f. of $\Phi_{RES}$ into the convolution of three parts. The first part has no observable difference with a Gaussian p.d.f. and corresponds to the third largest to the smallest eigenvalues, $\lambda_k, k = 3, \ldots, N_A$, of the characteristic function in the format of Eq. (5.24). The second and third parts are noncentral $\chi^2$-p.d.f. with two degrees of freedom and corresponds to the largest and second largest eigenvalues $\lambda_1$ and $\lambda_2$, respectively. The p.d.f. of $\Phi_{RES}$ in Fig. 6.5 is also plotted in Fig. 6.6 for comparison.

Figure 6.5. The p.d.f. of residual nonlinear phase noise of $\Phi_{RES}$. 

![Figure 6.5. The p.d.f. of residual nonlinear phase noise of $\Phi_{RES}$](image-url)
3. Linear Compensator for Infinite Number of Fiber Spans

With linear compensator, the decision variable is a linear combination of nonlinear phase noise $\Phi_{\text{NL}}$ [either Eqs. (5.5) or (5.32)], the phase of amplifier noise $\Theta_n$, and the received intensity $Y$. This section first finds the received phase with linear compensation. Like Sec. 5.3, the p.d.f. of decision variable is then expanded using as a Fourier series and the error probability can be found analytically as a series summation. This section is based on the distributed model of Sec. 5.2 but the results and methods are also applicable to a system with finite number of fiber spans using the model of Appendix 5.B.

3.1 Minimum Mean-Square Error Compensation

The simplest method to compensate the nonlinear phase noise is to add a scaled received intensity into the received phase. The optimal linear MMSE compensator minimizes the variance of the normalized residual phase noise of $\Phi_\alpha = \Phi - \alpha R^2 = \Phi - \alpha Y$. Using the joint characteristic function of Eq. (6.B.7) from Appendix 6.B, the characteristic function for the normalized residual nonlinear phase noise is

$$\Psi_{\Phi_\alpha}(\nu) = \Psi_{\Phi,Y}(\nu, -\alpha \nu).$$  (6.28)
The mean of the normalized residual nonlinear phase noise is

\[ \langle \Phi_\alpha \rangle = -j \frac{d}{d\nu} \Psi_{\Phi_\alpha} (\nu) \bigg|_{\nu=0} = \rho_s + \frac{1}{2} - \alpha (\rho_s + 1). \]  

(6.29)

The variance of the normalized residual nonlinear phase noise is

\[ \sigma_{\Phi_\alpha}^2 = - \frac{d^2}{d\nu^2} \Psi_{\Phi_\alpha} (\nu) \bigg|_{\nu=0} - \langle \Phi_\alpha \rangle^2 \]

\[ = \frac{2}{3} \rho_s + \frac{1}{6} - 2 \left( \rho_s + \frac{1}{3} \right) \alpha + (2\rho_s + 1) \alpha^2. \]  

(6.30)

Solving \( d\sigma_{\Phi_\alpha}^2 / d\alpha = 0 \), the optimal scale factor for linear compensator is

\[ \alpha_{\text{mse}} = \frac{1}{2} \frac{\rho_s + \frac{1}{3}}{\rho_s + \frac{1}{2}}. \]  

(6.31)

In high SNR, \( \alpha_{\text{mse}} \to 1/2 \). Other than the normalization, the optimal scale factor of Eq. (6.31) is the same as that from Sec. 6.2. Here, the optimal scale factor is derived by the joint characteristic function between nonlinear phase noise and the received intensity. Similar to Sec. 6.2, the results here can also be derived using the covariance properties of the received intensity and nonlinear phase noise.

With the optimal scale factor of Eq. (6.31), the mean and variance of the normalized residual nonlinear phase noise are

\[ \langle \Phi_{\alpha_{\text{mse}}} \rangle = \frac{1}{2} \frac{\rho_s^2 + \frac{2}{3} \rho_s + \frac{1}{6}}{\rho_s + \frac{1}{2}}, \]  

(6.32)

\[ \sigma_{\Phi_{\alpha_{\text{mse}}}}^2 = \frac{1}{6} \frac{\rho_s^2 + \rho_s + \frac{1}{6}}{\rho_s + \frac{1}{2}}. \]  

(6.33)

The mean of the residual nonlinear phase noise is about half the mean of the nonlinear phase noise of \( \langle \Phi \rangle \) from Eq. (5.49). The variance of the residual nonlinear phase noise is about a quarter of that of the variance of the nonlinear phase noise of \( \sigma_{\Phi}^2 \) from Eq. (5.52).

After the linear compensation, the characteristic function of the residual normalized nonlinear phase noise is

\[ \Psi_{\Phi_{\alpha_{\text{mse}}}} (\nu) = \Psi_{\Phi \nu} \left( \nu, -\nu \frac{\rho_s + \frac{1}{3}}{2 \rho_s + \frac{1}{2}} \right). \]  

(6.34)
The p.d.f. of the residual normalized nonlinear phase noise is the inverse Fourier transform of $\Psi_{\Phi_{\text{non}}}(\nu)$.

The p.d.f. of the residual nonlinear phase noise for finite number of fiber spans was derived by modeling the residual nonlinear phase noise $\Phi_\alpha$ as the summation of $N_A$ independently distributed random variables in Sec. 6.2.2. Figure 6.7 shows a comparison of the p.d.f. for $N_A = 4, 8, 16, 32$ of fiber spans with the distributed case of Eq. (6.34). The residual nonlinear phase noise is scaled by the mean normalized phase shift of $\langle \Phi \rangle = \rho_s + \frac{1}{2}$ from Eq. (5.49). Using a SNR of $\rho_s = 18$, Figure 6.7 is plotted in logarithmic scale to show the difference in the tail. Figure 6.7 also provides an inset in linear scale of the same p.d.f. to show the difference around the mean. Like that of Fig. 5.5, the asymptotic p.d.f. for residual nonlinear phase noise of Fig. 6.7 is also very accurate for $N_A \geq 32$ fiber spans. Unlike that of Fig. 5.5, the asymptotic p.d.f. for residual nonlinear phase noise of Fig. 6.7 has slightly larger spreading then that of the finite cases. The mean of the residual nonlinear phase noise is about $0.5(\rho_s + \frac{1}{2})$, the same as that of Eq. (6.32). Comparing Figs. 5.5 and 6.7, with linear compensation, both the mean and STD of the residual nonlinear phase noise is about half of that of the nonlinear phase noise before compensation.

Figure 6.7. The asymptotic p.d.f. of the residual nonlinear phase noise $\Phi_\alpha$ as compared with the p.d.f. of $N_A = 4, 8, 16, 32$ fiber spans. The p.d.f. in linear scale is shown in the inset.
3.2 Distribution of the Linearly Compensated Received Phase

With nonlinear phase noise, received by a quadrature receiver of Fig. 3.4 with an optical PLL, before compensation, the overall received phase is that of Eq. (5.63). With linear compensation, the compensated received phase is

$$\Phi_{cm} = \Theta_n - \frac{\langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} (\Phi - \alpha Y). \quad (6.35)$$

The compensated received phase is confined to the range of $[-\pi, +\pi)$. The p.d.f. of the compensated received phase is a periodic function with a period of $2\pi$. If the characteristic function of the compensated received phase is $\Psi_{\Phi_{cm}}(\nu)$, the p.d.f. of the compensated received phase has a Fourier series expansion of

$$p_{\Phi_{cm}}(\theta) = \frac{1}{2\pi} \sum_{m=-\infty}^{+\infty} \Psi_{\Phi_{cm}}(m) \exp(jm\theta), \quad (6.36)$$

or

$$p_{\Phi_{cm}}(\theta) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{m=1}^{+\infty} \Re \{\Psi_{\Phi_{cm}}(m) \exp(jm\theta)\}, \quad (6.37)$$

where $\Re\{\cdot\}$ denotes the real part of a complex number.

Using the joint characteristic function of $\Psi_{\Phi,Y,\Theta_n}(\nu, \omega, m)$ with $\Theta_n$ at integer “angular” frequency from Eq. (6.B.10) of Appendix 6.B, similar to that of Eq. (6.34), from the compensated received phase of Eq. (6.35), the Fourier series coefficients are

$$\Psi_{\Phi_{cm}}(m) = \Psi_{\Phi,Y,\Theta_n} \left( \frac{-m \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}}, \alpha \frac{m \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}}, m \right). \quad (6.38)$$

Using the scale factor of $\alpha_{mse}$ from Eq. (6.31), Figure 6.8 shows the p.d.f. of the compensated received phase of Eq. (6.37) with mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 0, 1, 2, \text{ and } 3 \text{ rad}$. Shifted by the mean residual nonlinear phase shift of

$$\langle \Phi_{RES} \rangle = \frac{\langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \langle \Phi_{\alpha_{mse}} \rangle, \quad (6.39)$$

the p.d.f. is plotted in logarithmic scale to show the difference in the tail. Not shifted by $\langle \Phi_{RES} \rangle$, the same p.d.f. is plotted in linear scale in the inset. Figure 6.8 is plotted for the case that the SNR is equal to $\rho_s = 18$ (12.6 dB), the same as that of Fig. 5.6 without compensation.
From Fig. 6.8, when the p.d.f. is broadened by the nonlinear phase noise, similar to the case without compensation of Fig. 5.6, the broadening is not symmetrical with respect to the mean residual nonlinear phase shift. With small mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 1$ rad, the received phase spreads further in the positive than the negative phase. With large mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 3$ rad, the received phase spreads further in the negative than the positive phase. The difference in the spreading for small and large mean nonlinear phase shift is due to the dependence between the residual nonlinear phase noise and the phase of amplifier noise.

### 3.3 PSK Signals

For PSK signal with two antipodal phases, assume that the decision angles are $\pm \pi/2 - \theta_c$ with the center phase of $\theta_c$, the error probability is

$$p_e = 1 - \int_{-\frac{3}{2} \pi - \theta_c}^{\frac{3}{2} \pi - \theta_c} p_{\Phi_{cm}}(\theta) d\theta,$$

or, similar to the error probability of Eq. (5.69),

$$p_e = \frac{1}{2} \left[ 2 - \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1} 2 e^{-j(2k+1)\theta_c} \right].$$
From both Eqs. (6.38) and (6.B.10), the coefficients for the error probability Eq. (6.41) are

\[
\Psi_{\Phi_{cm}}(2k + 1) = \frac{\sqrt{\pi r_{k,\omega}^3}}{2r_k} e^{-r_k + \frac{r_{k,\omega}}{2}} \Psi_{\Phi} \left[ \frac{(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right] \\
\times \left[ I_k \left( \frac{T_{k,\omega}}{2} \right) + I_{k+1} \left( \frac{T_{k,\omega}}{2} \right) \right], \quad k \geq 0,
\]  

(6.42)

where, from Eq. (5.A.12) of Appendix 5.A and Eq. (6.B.3) of Appendix 6.B,

\[
2 \left[ \frac{j(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right]^{\frac{1}{2}} \\
\sin \left\{ 2 \left[ \frac{j(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right]^{\frac{1}{2}} \rho_s \right\}
\]

(6.43)

and

\[
r_{k,\omega} = \frac{r_k}{1 + \alpha \left[ \frac{j(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right]^{\frac{1}{2}} \tan \left\{ \left[ \frac{j(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right]^{\frac{1}{2}} \right\}},
\]

(6.44)

and \(\Psi(\nu)\) is the marginal characteristic function of nonlinear phase noise of Eq. (5.48) that depends solely on SNR. Note that the angular SNR of Eq. (6.43) is the same as Eq. (5.71).

Based on the MMSE criterion, the center phase in Eq. (6.41) is \(\theta_c = \langle \Phi_{RES} \rangle\) of Eq. (6.39) and the scale factor in Eq. (6.44) is \(\alpha = \alpha_{\text{mse}}\) of Eq. (6.31). Because the series summation of Eq. (6.41) is a simple expression, numerical optimization can be used to find the linear MAP compensator to minimize the error probability.

If the residual nonlinear phase noise is assumed to be independent to the phase of amplifier noise, similar to Eq. (5.72), the error probability based on MMSE criterion can be approximated as

\[
p_c \approx \frac{1}{2} - e^{-\rho_s^2} \sqrt{\frac{\rho_s}{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k + 1} \left[ I_k \left( \frac{\rho_s}{2} \right) + I_{k+1} \left( \frac{\rho_s}{2} \right) \right] \\
\times \Re \left\{ \Psi_{\Phi_{mse}} \left[ \frac{(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + \frac{1}{2}} \right] e^{-j(2k+1)\langle \Phi_{RES} \rangle} \right\},
\]

(6.45)

where the characteristic function \(\Psi_{\Phi_{mse}}(\nu)\) is the characteristic function of Eq. (6.34) with \(\alpha = \alpha_{\text{mse}}\) of Eq. (6.31).
Figure 6.9 shows the exact [Eq. (6.41)] and approximated [Eq. (6.45)] error probabilities as a function of SNR $\rho_s$ when nonlinear phase noise is compensated using the linear MMSE compensator. The error probability of Eq. (6.41) is also minimized based on the MAP criterion and shown in Fig. 6.9. Figure 6.9 also plots the error probability without nonlinear phase noise of Fig. 3.13.

Figure 6.10 shows the SNR penalty of PSK signal for an error probability of $10^{-9}$ calculated by the exact Eq. (6.41) and approximated Eq. (6.45) error probability formulae with the linear MMSE compensator. The SNR penalty with the linear MAP compensator is also shown in Fig. 6.10. The corresponding optimal center phase and scale factor of Fig. 6.10 are shown in Fig. 6.11.

Figure 6.10 also shows the SNR penalty of PSK signal without compensation from Fig. 5.8 using the exact error probability with optimal center phase there. For the same SNR penalty, the mean nonlinear phase shift with compensation is slightly larger than twice of that without compensation from Fig. 5.8.

The discrepancy between the exact and approximated error probability is smaller for small and large mean nonlinear phase shift of $\langle \Phi_{NL} \rangle$. 
Figure 6.10. With linear compensation, the SNR penalty of PSK signal as a function of mean nonlinear phase shift $\langle \Phi_{NL} \rangle$.

Figure 6.11. The optimal center phase corresponding to the operating point of Fig. 6.10 as a function of mean nonlinear phase shift $\langle \Phi_{NL} \rangle$. 
With the MMSE criterion, the largest discrepancy between the exact and approximated SNR penalty is about 0.37 dB at a mean nonlinear phase shift around $\langle \Phi_{NL} \rangle = 2.53$ rad. Like the conclusion of Sec. 5.3 without compensation, with the MMSE criterion, the approximated error probability Eq. (6.45) is not accurate enough for linearly compensated PSK signals.

Figure 6.10 also shows that the linear MMSE compensator using the optimal scale factor of Eq. (6.31) does not perform well as compared with the linear MAP compensator. The largest discrepancy is about 0.34 dB at mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 1.68$ rad. The major reason of this large discrepancy is due to the non-symmetrical p.d.f. of Fig. 6.8.

Using the linear MAP compensator, the mean nonlinear phase shift $\langle \Phi_{NL} \rangle$ must be less than 2.30 rad for a SNR penalty less than 1 dB, slightly more than twice that of Fig. 5.8 of 1 rad without compensation. The optimal operating level is for a mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 2.15$ rad, slightly less than twice that of Fig. 5.8 of 1.25 rad without compensation.

From the optimal center phase from Fig. 6.11 for the linear MAP compensator, the optimal center phase is less than the mean nonlinear phase shift when the mean nonlinear phase shift is less than about 2.68 rad. The changing of the optimal center phase with the mean nonlinear phase shift is consistent with Figs. 5.9 and 6.8. Similar to Fig. 5.9, when the residual nonlinear phase noise is assumed to be independent of the phase of amplifier noise using Eq. (6.45), the optimal center phase is always larger than the mean residual nonlinear phase shift. The approximated error probability of Eq. (6.45) is not useful to find the optimal center phase.

The optimal scale factor from Fig. 6.11 is also not equal to the optimal MMSE scale factor from Eq. (6.31) except when the mean nonlinear phase shift is very large. From Figs. 6.10 and 6.11, the MMSE criterion is not close to the performance of the optimal linearly compensated PSK signals based on MAP criterion.

### 3.4 DPSK Signals

A compensated DPSK signal is demodulated using interferometer to find the compensated differential phase of

$$
\Delta \Phi_{cm} = \Phi_{cm}(t) - \Phi_{cm}(t - T) = \Theta_n(t) - \Phi_{RES}(t) - \Theta_n(t - T) + \Phi_{RES}(t - T),
$$

where $\Phi_{cm}(\cdot)$, $\Theta_n(\cdot)$, and $\Phi_{RES}(\cdot)$ are the compensated received phase, the phase of amplifier noise, and the residual nonlinear phase noise as a
function of time, and $T$ is the symbol interval. The phases at $t$ and $t-T$ are independent of each other but are identically distributed random variables similar to that of Eqs. (5.63) and (6.35). The differential phase of Eq. (6.46) assumes that the transmitted phases at $t$ and $t-T$ are the same.

From the p.d.f. of Eq. (6.37), similar to Eq. (5.75), the p.d.f. of the differential compensated phase Eq. (6.46) is

$$p_{\Delta \Phi_{cm}}(\theta) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{m=1}^{+\infty} |\Psi_{\Phi_{cm}}(m)|^2 \cos(m\theta), \quad (6.47)$$

that is symmetrical with respect to the zero phase.

Figure 6.12 shows the p.d.f. of the differential received phase with linear compensation of Eq. (6.47) for the mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 0, 1, 2,$ and 3 rad using the linear MMSE compensator with $\alpha = \alpha_{\text{mse}}$ of Eq. (6.31). The p.d.f. is plotted in logarithmic scale to show the difference in the tail. The same p.d.f. is plotted in linear scale in the inset. Figure 6.12 is plotted for the case that the SNR is equal to $\rho_s = 20$ (13 dB), corresponding to an error probability of $10^{-9}$ with only amplifier noise from Fig. 3.13. From Fig. 6.12, when the p.d.f. of differential phase is broadened by the residual nonlinear phase noise, the broadening is symmetrical with respect to the zero phase.
Similar to Eqs. (5.77) and (6.41), the error probability for DPSK signal is

\[ p_e = \frac{1}{2} - \frac{2}{\pi} \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1} |\Psi_{\phi_{cm}}(2k+1)|^2, \quad (6.48) \]

where the coefficients of \( \Psi_{\phi_{cm}}(2k+1) \) are given by Eq. (6.42) with parameters from Eqs. (6.43) and (6.44).

Similar to the approximation for PSK signal Eq. (6.45), if the nonlinear phase noise is assumed to be independent to the phase of amplifier noise, the error probability of Eq. (6.48) can be approximated as

\[ p_e \approx \frac{1}{2} - \frac{\rho_s e^{-\rho_s}}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \left[ I_k\left(\frac{\rho_s}{2}\right) + I_{k+1}\left(\frac{\rho_s}{2}\right) \right]^2 \]

\[ \times \left| \Psi_{\phi_{\text{mmse}}} \left[ \frac{(2k+1) (\Phi_{NL})}{\rho_s + \frac{1}{2}} \right] \right|^2. \quad (6.49) \]

Figure 6.13 shows the exact [Eq. (6.48)] and approximated [Eq. (6.49)] error probabilities as a function of SNR \( \rho_s \) for DPSK signal with linear MMSE compensator. The scale factor in Eq. (6.44) is also numerically
optimized to find the linear MAP compensator to minimize the exact error probability of Eq. (6.48). From Fig. 6.13, the linear MMSE and MAP compensators do not have big difference for linearly compensated DPSK signal. Figure 6.13 also plots the error probability without nonlinear phase noise. From Fig. 6.13, the approximated error probability Eq. (6.49) always slightly overestimates the error probability.

Figure 6.14 shows the SNR penalty of DPSK signal for an error probability of $10^{-9}$ calculated by the exact and approximated error probability formulae for DPSK signal with linear MMSE and MAP compensation. The discrepancy between the exact and approximated error probability with either MMSE or MAP criteria is very small. The largest discrepancy between the exact and approximated SNR penalty with MMSE criterion is about 0.1 dB at a mean nonlinear phase shift of $\langle \Phi_{NL} \rangle = 1.74$ rad. Both with the exact error probability Eq. (6.48), the linear MMSE and MAP compensators have the largest discrepancy of 0.06 dB at a mean nonlinear phase shift $\langle \Phi_{NL} \rangle$ of 0.95 rad.

Figure 6.14 also shows the SNR penalty of DPSK signal without compensation from Eq. (5.13) using the exact error probability formula. For a power penalty less than 1 dB, the mean nonlinear phase shift of $\langle \Phi_{NL} \rangle$ must be less than 1.31 rad, slightly larger than twice that of Fig. 5.13 of 0.57 rad without compensation. The optimal operating level of the mean nonlinear phase shift is about 1.81 rad such that the increase of
power penalty is always less than the increase of mean nonlinear phase shift, slightly smaller than twice that of Fig. 5.13 of 1 rad without compensation.

To verify the accuracy of the error probability in Fig. 6.13, Figure 6.15 compares the theoretical and simulated error probability as a function of SNR for a typical linearly compensated DPSK system having mean nonlinear phase shift of \( \langle \Phi_{NL} \rangle = \sqrt{2} \) rad. The Monte-Carlo simulation is similar to that of Figs. 5.10 and 5.14 for \( N_A = 32 \) fiber spans using the optimal scale factor \( \alpha_{\text{mse}} \) of Eq. (6.31). Including both exact [Eq. (6.48)] and approximated [Eq. (6.49)] error probability, the theoretical results are the same as that in Fig. 6.13 but extend to high error probability. Figure 6.15 shows that the approximated, exact and simulated results have insignificant difference. From Fig. 6.15, we may conclude that the exact error probability of Eq. (6.48) is very accurate to evaluate the error probability of DPSK signal with linearly compensated nonlinear phase noise.

The SNR penalty given by the approximated error probability of Eq. (6.49) is the same as that in Ho (2003b) but for large number of
fiber spans. With the assumption of independent additive phase noise, the approximated error probability of Eq. (6.49) is the same as that of Eqs. (5.78), (5.82), and (4.40) with nonlinear and linear phase noise. Similar to Ho (2003b), for the same SNR penalty, DPSK signal with linear compensator can tolerate slightly larger than twice the mean nonlinear phase noise without linear compensation.

In this section, the exact error probability is derived analytically for PSK and DPSK signals with linearly compensated nonlinear phase noise. The p.d.f. of the compensated received phase is first expressed as a Fourier series. The Fourier coefficients are given by the joint characteristic function of nonlinear phase noise, received intensity, and the phase of amplifier noise.

For PSK signal, the linear MMSE compensator is up to 0.34 dB worse than the linear MAP compensator that minimizes the error probability after compensation. For PSK signal with linear MAP compensator, the mean nonlinear phase shift must be less than 2.30 rad for a SNR penalty less than 1 dB. The optimal mean nonlinear phase shift is about 2.15 rad such that the increase of required SNR is smaller than the increase of launched optical power.

For DPSK signal with linear MAP compensator, the phase of amplifier noise can be assumed to be independent of the residual nonlinear phase noise. The mean nonlinear phase shift must be less than 1.31 rad for a SNR penalty less than 1 dB. The optimal mean nonlinear phase shift is about 1.81 rad.

The formulae to evaluate the error probability of DPSK signal with linearly compensated nonlinear phase noise are verified by Monte-Carlo simulation. Theoretical and simulation results have no significant difference.

4. Mid-Span Linear MMSE Compensation

Unlike the electronic compensator of Fig. 6.1, the phase-modulator based compensator of Fig. 6.3 can be used in the middle of the fiber link. The system performance can be greatly improved using mid-span compensation even with only a single compensator. When more than one compensator are used, the performance of the system can be further improved. With a compensator preceding the receiver, the variance of the residual nonlinear phase noise can be reduced to about a quarter of the variance of nonlinear phase noise without compensation as shown earlier. With a compensator located optimally in the middle of fiber span, the variance of the residual nonlinear phase noise can be reduced to about 1/9 of the variance of nonlinear phase noise without compensation.
The compensator is independent of signal format and applicable to NRZ and RZ PSK and DPSK signals.

The optimal locations for mid-span compensation are derived analytically here. For simplicity, the model of nonlinear phase noise is according to the distributed model with infinite number of fiber spans. The optimal linear compensator is derived using the covariance between optical intensity and nonlinear phase noise.

4.1 Single Compensator

If only one compensator is used, the compensator can be located anywhere in the fiber link to optimize the system performance. Using the compensation factor of \( \alpha \), based on the distributed model, the optimal compensator minimizes the variance of the residual nonlinear phase noise of

\[
\Phi_\alpha = \Phi - \alpha |\sqrt{\rho_s} + b(s)|^2,
\]

where \( s \) is the location of the compensator. In order to derive the optimal factor of \( \alpha \), we need to find the covariance of

\[
\kappa(s, t) = E \left\{ \left( |\sqrt{\rho_s} + b(s)|^2 - m_1|\sqrt{\rho_s} + b(s)|^2 \right) \times \left( |\sqrt{\rho_s} + b(t)|^2 - m_1|\sqrt{\rho_s} + b(t)|^2 \right) \right\}
\]

for Wiener process of \( b(s) \), where \( m_1|\sqrt{\rho_s} + b(s)|^2 \) and \( m_1|\sqrt{\rho_s} + b(t)|^2 \) are the mean values of \( |\sqrt{\rho_s} + b(s)|^2 \) and \( |\sqrt{\rho_s} + b(t)|^2 \), respectively. Because \( |\sqrt{\rho_s} + b(t)|^2 \) is noncentral \( \chi^2 \) random variable with two degrees of freedom, the noncentrality parameter is \( \sqrt{\rho_s} \) and the variance is \( t/2 \). We have \( m_1|\sqrt{\rho_s} + b(s)|^2 = \rho_s + s \) and \( m_1|\sqrt{\rho_s} + b(t)|^2 = \rho_s + t \).

The covariance between \( b(s) \) and \( b(t) \) is \( \min(s, t) \) from the properties of Wiener process. Similar to Eq. (6.5), the covariance of Eq. (6.51) is

\[
\kappa(s, t) = 2\rho_s \min(s, t) + \min(s, t)^2.
\]

For the normalized nonlinear phase noise \( \Phi \) defined in Eq. (5.32), the covariance between \( \Phi \) and the intensity of \( |\sqrt{\rho_s} + b(s)|^2 \) is

\[
\xi(s) = E \left\{ (|\sqrt{\rho_s} + b(s)|^2 - \rho_s - s) \left( \int_s^1 |\sqrt{\rho_s} + b(t)|^2 dt - \rho_s - \frac{1}{2} \right) \right\}
\]

\[
= \int_s^1 \kappa(s, t) dt
\]

\[
= \rho_s s(2 - s) + s^2 \left( 1 - \frac{2s}{3} \right).
\]
Using both Eqs. (6.52) and (6.53), the variance of the residual nonlinear phase noise is

$$\sigma_{\Phi,\alpha}^2 = \sigma_{\Phi}^2 + \alpha^2 \kappa(s, s) - 2\alpha \xi(s),$$  \hspace{1cm} (6.54)

The variance of Eq. (6.54) is a quadratic function of $\alpha$, the optimal compensation factor is given by

$$\alpha_{\text{opt}} = \frac{\xi(s)}{\kappa(s, s)} = \frac{\rho_s s^2(2 - s)}{2\rho_s s + s^2}. $$  \hspace{1cm} (6.55)

If $s = 1$, the optical compensation factor of $\alpha_{\text{opt}}$ is the same as Eq. (6.31) when the compensator precedes the receiver. If $s = 0$, there is no compensation. For high SNR of $\rho_s$, the optimal compensation factor is approximately equal to $\alpha_{\text{opt}} \rightarrow 1 - s/2$.

When the optimal compensation factor of Eq. (6.55) is substituted to the variance of Eq. (6.54), we obtain

$$\sigma_{\Phi,\alpha,\text{min}}^2 = \sigma_{\Phi}^2 - \frac{\xi^2(s)}{\kappa(s, s)}. $$  \hspace{1cm} (6.56)

The location of the compensator is optimized by solving $d\sigma_{\Phi,\alpha,\text{min}}^2 / ds = 0$ to obtain

$$s_{\text{opt}} = \frac{3}{8} + \frac{9\rho_s}{8} \left(1 + \frac{9 + 42\rho_s}{81\rho_s^2} - 1\right) \approx \frac{2}{3} + \frac{2}{91\rho_s}. $$  \hspace{1cm} (6.57)

In most value of SNR, the optimal location is about $2/3$ from the beginning of the fiber link.

At high SNR of $\rho_s$, the ratio of the variance of residual nonlinear phase noise to that of nonlinear phase noise of Eq. (5.52) is equal to

$$\frac{\sigma_{\Phi,\alpha,\text{min}}^2}{\sigma_{\Phi}^2} \rightarrow 1 - 3s + 3s^2 - \frac{3}{4}s^3 $$  \hspace{1cm} (6.58)

with a minimum value of $1/9$ when $s_{\text{opt}} \rightarrow 2/3$.

Figure 6.16 shows the ratio of standard deviation (STD) of residual nonlinear phase noise of $\sigma_{\text{RES}}$ to the mean nonlinear phase shift $\langle \Phi_{\text{NL}} \rangle$. The STD of residual nonlinear phase noise is given by $\sigma_{\text{RES}} = \langle \Phi_{\text{NL}} \rangle \sigma_{\Phi,\alpha}/(\rho_s + 1/2)$ without normalization. In Fig. 6.16, the compensator is located at $s = 1/4, 1/2$ and $3/4$ from the beginning of fiber link, or preceding the receiver of $s = 1$. Figure 6.16 also shows $\sigma_{\text{RES}}/\langle \Phi_{\text{NL}} \rangle$ without compensation and with optimal location of $s = s_{\text{opt}}$ of Eq. (6.57).
From Fig. 6.16, the same as Sec. 6.2, a compensator preceding the receiver about halves the STD of the nonlinear phase noise. From Fig. 6.16, a compensator at the optimal location of $s = s_{opt}$ of Eq. (6.57) reduces the STD of nonlinear phase noise to about $1/3$ of the case without compensation.

To illustrate the effects of compensator location to the performance of the nonlinear phase noise compensation scheme, Figure 6.17 shows the ratio of $\sigma_{RES}/\langle \Phi_{NL} \rangle$ as a function of compensator location for several values of SNR $\rho_s$. The residual nonlinear phase noise reduces when a compensator is used in the fiber link. For a location before about $2/3$ of the fiber link, the STD of residual nonlinear phase noise decreases with the location of the compensator and reaches a value of about $1/3$ of the value of STD without compensation when $s = 0$. For a location after about $2/3$ of the fiber link, the STD of residual nonlinear phase noise increases with the location of the compensator and reaches a value of about $1/2$ of the value of STD without compensation when the compensator precedes the receiver.

Mid-span compensation is commonly better than compensation preceding the receiver. The nonlinear phase noise of Eq. (5.32) is an integration of the noisy power of $|\sqrt{\rho_s} + b(t)|^2$ over distance of $t$. As an example, the mid-span intensity of $|\sqrt{\rho_s} + b(1/2)|^2$ has less noise than the received intensity of $|\sqrt{\rho_s} + b(1)|^2$. However, the nonlinear phase noise
Figure 6.17. The ratio of STD to mean of nonlinear phase noise as a function of compensator location.

[Reprinted from Ho (2005a), copyright (2005), with permission from Elsevier]

after the compensator of \( \int_{1/2}^{1} |\sqrt{\rho_s} + b(t)|^2 dt \) can still be compensated using \( |\sqrt{\rho_s} + b(1/2)|^2 \). Because the intensity of \( |\sqrt{\rho_s} + b(1/2)|^2 \) is not suitable to compensate the nonlinear phase noise at the end of the fiber link, i.e., \( \int_{0.8}^{1} |\sqrt{\rho_s} + b(t)|^2 dt \), the optimal location needs to balance the usage of less noisy optical intensity and the ability to compensate the nonlinear phase noise at the end of the fiber link. The optimal location of about 2/3 is the balance between these two factors. The optimal compensation factor is \( \alpha_{\text{opt}} \to 2/3 \) at high SNR when \( s = 2/3 \). From the same length of fiber interval, the nonlinear phase noise of \( \int_{0.8}^{1} |\sqrt{\rho_s} + b(t)|^2 dt \) is significantly larger than \( \int_{0.6}^{0.2} |\sqrt{\rho_s} + b(t)|^2 dt \), the optimal compensator location is not at the middle of \( s = 1/2 \) but \( s = 2/3 \) toward the end of the fiber link.

It is also interesting to find the variance of nonlinear phase noise as a function of distance. The compensation factor of Eq. (6.55) minimizes the variance of residual nonlinear phase noise at the end of the fiber link. The compensator may increase the phase noise immediately at the output of the compensator. However, at the end of the fiber link, the residual nonlinear phase noise is definitely minimized. For a compensator located at \( s \) with a compensator factor of \( \alpha \), the variance of the residual nonlinear phase noise as a function of distance is equal to

\[
\sigma_{\phi}^2(t) + \alpha^2 \kappa(s,s) - 2\alpha \xi_t(s), \quad t \geq s,
\]

(6.59)
Figure 6.18. The ratio of STD to mean of nonlinear phase noise as a function of normalized distance when a compensator is located at $s = 1/2, 2/3$ and $3/4$. [Reprinted from Ho (2005a), copyright (2005), with permission from Elsevier]

where

$$\xi_t(s) = \int_0^t \kappa(s, \tau) d\tau$$

(6.60)

is the correlation of the nonlinear phase noise at $t$ with the intensity at $s$.

Figure 6.18 shows the ratio of $\sigma_{RES}/\langle \Phi_{NL} \rangle$ as a function of normalized distance when a compensator is located at $s = 1/2, 2/3$ and $3/4$. Without compensation, the STD of nonlinear phase noise as a function of distance is given by Eq. (6.59). Without compensation, the STD of nonlinear phase noise as a function of distance is given by

$$\sigma^2_{\Phi}(t) = \int_0^t \xi_t(s) ds = \frac{2\rho_s t^3}{3} + \frac{t^4}{6}.$$  (6.61)

When the amplifier noise is accumulated with distance, the nonlinear phase noise is increased faster at the end of the fiber link.

From Fig. 6.18, the STD of nonlinear phase noise is the largest at the compensator output for both $s = 1/2$ and $2/3$. When the compensator is located at $s = 1/2$, the compensator actually increases its output phase noise. When the compensator is located at $s = 3/4$, the compensator reduces its output phase noise. However, in all cases of Fig. 6.18, the compensator minimizes the nonlinear phase noise at the end of the fiber link.
4.2 Multiple Compensators

The system can further be improved when more than one compensator are used for mid-span compensation. If there are $N_c$ compensators located at $s_1, s_2, \ldots, s_{N_c}$, the optimal compensation factors of $\alpha_1, \alpha_2, \ldots, \alpha_{N_c}$ minimize the variance of the residual nonlinear phase noise of

$$
\Phi_0 = \Phi - \sum_{k=1}^{N_c} \alpha_k |\sqrt{\rho_s} + b(s_k)|^2.
$$

(6.62)

The variance of the residual nonlinear phase noise is equal to

$$
\sigma_{\Phi_0}^2 = \sigma_{\Phi_0}^2 - 2 \sum_{k=1}^{N_c} \alpha_k \xi(s_k) + \sum_{k=1}^{N_c} \sum_{l=1}^{N_c} \alpha_k \alpha_l \kappa(s_k, s_l).
$$

(6.63)

where the second term is the covariance of the intensity at $s_k$ of $|\sqrt{\rho_s} + b(s_k)|^2$ with the nonlinear phase noise of Eq. (5.32) and the third term is the covariance of the intensity at $s_k$ and $s_l$.

The optimal compensation factors can be found by solving all equations of $\partial \sigma_{\Phi_0}^2 / \partial \alpha_k = 0$, or

$$
\sum_{l=1}^{N_c} \alpha_l \kappa(s_k, s_l) = \xi(s_k), \quad k = 1, \ldots, N_c.
$$

(6.64)

In the case of $N_c = 1$, the solution is the same as that of Eq. (6.55). The $N_c$ linear equations in Eq. (6.64) determine the $N_c$ optimal compensation factors of $\alpha_k, k = 1, \ldots, N_c$. The optimal compensation factors are equal to

$$
\tilde{\alpha}_{\text{opt}} = K^{-1} \tilde{Z},
$$

(6.65)

where the covariance matrix has elements of $K_{kl} = \kappa(s_k, s_l)$ and the correlation vector is $\tilde{Z} = [\xi(s_1), \xi(s_2), \ldots, \xi(s_{N_c})]^T$. The solution of Eq. (6.65) is the same as that for most MMSE algorithms (Proakis, 2000, p. 626).

Figure 6.19 shows the ratio of $\sigma_{\text{RES}} / \langle \Phi_{\text{NL}} \rangle$ as a function of SNR $\rho_s$ when multiple compensators are used. In Fig. 6.19(a), the $N_c$ compensators are evenly located at $s_k = k/N_c, k = 1, 2, \ldots, N_c$, including one compensator preceding the receiver of $s_{N_c} = 1$. In Fig. 6.19(b), the $N_c$ compensators are evenly located at $s_k = 2k/(2N_c + 1), k = 1, 2, \ldots, N_c$. The compensator spacings in Fig. 6.19(a) and (b) are $1/N_c$ and $2/(2N_c + 1)$, respectively. The two curves in Fig. 6.16 without compensation and with one compensator of $s = 1$ are also shown in Fig. 6.19 for comparison.
From Fig. 6.19(a), the STD of nonlinear phase noise is reduced by the factor of about $2N_c$. From Fig. 6.19(b), the STD of nonlinear phase noise is reduced by the factor of about $2N_c + 1$. Because the intensity at $s_{N_c} = 1$ has larger noise, better performance is achieved when $s_{N_c} =$
1 - 1/(2Nc + 1). In Fig. 6.17, the STD of the residual nonlinear phase noise is very close to its minimum value for a large range of locations from s = 0.6 to s = 0.75. Because the compensation scheme is optimized over a large range of locations, the locations of sk = 2k/(2Nc + 1) are very close to the optimal locations.

The optimal compensation factor of αopt can also be evaluated numerically. With the compensator location of sk = k/Nc, the optimal compensation factors are αopt,k → 1/Nc, k = 1, . . . , Nc - 1, and αopt,Nc → 1/2Nc. With the compensator locations of sk = 2k/(2Nc + 1), the optimal compensation factors all approach the normalized values of αopt,k → 2/(2Nc + 1), k = 1, . . . , Nc, that is the same as the compensator spacing. With the compensator locations of sk = 2k/(2Nc + 1), all compensators have approximately the same compensation factor.

From Fig. 6.19, when sufficient number of compensators is used, nonlinear phase noise can be completely eliminated. We may infer that if one compensator is used per fiber span in the system of Sec. 6.2, nonlinear phase noise can be completely eliminated.

When Nc compensators are used in optimal locations, the STD of nonlinear phase noise is reduced by a factor of about 2Nc+1. If nonlinear phase noise is the dominant noise source, the transmission distance can be increased by a factor of 2Nc+1. However, the amplifier noises modeled by the Wiener process of b(t) also limits the transmission distance by adding noise into the signal for system with more fiber spans. With higher tolerance of nonlinear phase noise, the launched power to the system can be increased accordingly. Assumed the same amplifier noises, the system SNR is increased with the launched power. If the operation point is estimated when the variances of linear and nonlinear phase noise are equal, the overall transmission distance can be increased by a factor of √2Nc+1.

Another method to reduce nonlinear phase noise is the usage of phase conjugation (McKinstrie et al., 2003). With a single compensator, the optimal location is also at 2/3 the fiber link and the STD is also reduced by a factor of 1/3. However, the techniques and methods for analysis are different. Instead of phase conjugator in McKinstrie et al. (2003), a phase modulator is used here instead. While the analysis of McKinstrie et al. (2003) is only for soliton, the method here is applicable to more general signals.

The inclusion of compensators in the fiber link requires additional amplification to compensate for the compensator loss. If the compensator is properly located, for example, in the mid-stage of an amplifier, compensator loss gives minimal increase to the amplifier noise figure and overall amplifier noise to the system.
In this section for mid-span compensation, with a single compensator, the optimal placement of the compensator is not preceding the receiver but at about 2/3 of the fiber link. With this optimal location, a single compensator reduces the STD of nonlinear phase noise by a factor of 3 instead of a factor of 2 for a compensator preceding the receiver. The optimal compensation factor is approximately equal to 2/3.

If \( N_c \) compensators are located at \( 2k/(2N_c + 1) \) of the fiber link, the STD of nonlinear phase noise is reduced by a factor of \( 2N_c + 1 \). The optimal compensation factors for all compensators are all approximately equal to \( 2/(2N_c + 1) \).

The analysis of this section follows Ho (2005a).

5. Nonlinear Compensation

In previous sections, linear compensator is investigated in detail. From Sec. 6.2, in term of variance, the linear MMSE compensator performs very close to the optimal nonlinear MMSE compensator. Usually, in digital communications, most types of compensator or equalizer are designed based on the MMSE criterion for simplicity (Proakis, 2000, chs. 10, 11).

However, as from Sec. 6.3, it is obvious that the MMSE compensator does not necessary perform the same as MAP compensator that minimizes the error probability. The linear MAP compensator was found to perform up to 0.35 dB better than the linear MMSE compensator for binary phase-shift keying (PSK) signals.

In order to find the optimal nonlinear MAP compensator, this section first derives the joint distribution of the received amplitude and phase. The optimal nonlinear MAP compensator is then derived for PSK signal with nonlinear phase noise. The error probability of both nonlinear MAP and MMSE compensators is calculated using the joint distribution of received amplitude and phase.

As shown in Sec. 6.1, a nonlinear compensator can be implemented as a Yin-Yang detector with a spiral decision boundary or equivalently to add a nonlinear compensated phase to the received phase.

The MAP detector provides the ultimate performance that is possible for a PSK signal with nonlinear phase noise. Nonlinear compensator can be implemented by phase modulator with nonlinear driver circuits or electronic signal processing techniques. The optimal nonlinear detector derived here is for a compensator preceding or combined with a receiver, distributed nonlinear compensator in the middle of fiber link is not considered here.
5.1 Joint Distribution of the Received Amplitude and Phase

The received phase of Eq. (5.63) is confined to the range of \([-\pi, +\pi]\). The joint p.d.f. of received amplitude and phase \(p_{R,\Phi|\theta_0}(r, \theta)\) can be modeled as a periodic function of \(\theta\) with a period of \(2\pi\) and expanded as a Fourier series as

\[
p_{R,\Phi|\theta_0}(r, \theta) = \frac{p_R(r)}{2\pi} + \frac{1}{\pi} \sum_{m=1}^{+\infty} \Re \left\{ C_m(r)e^{im(\theta-\theta_0)} \right\}, \quad r \geq 0, \quad (6.66)
\]

where

\[
p_R(r) = 2r \exp \left[ -(r^2 + \rho_s) \right] I_0(2r\sqrt{\rho_s}) \quad (6.67)
\]

is the Rice p.d.f. of the received amplitude, \(C_m(r)\) is the \(m\)th Fourier coefficient as a function of the received amplitude \(r\), and \(\Re \{ \cdot \} \) denotes the real part of a complex number. The p.d.f. of Eq. (6.66) has been simplified using the relationship of \(C_{-m}(r) = C^*_m(r)\). It is also obvious that the marginal p.d.f. of amplitude is the Rice distribution of \(\int_{-\pi}^{+\pi} p_{R,\Phi|\theta_0}(r, \theta) d\theta = p_R(r)\).

As the Fourier series of the p.d.f. of \(p_{R,\Phi|\theta_0}(r, \theta)\), the Fourier coefficients of \(C_m(r)\) are the Fourier transform of the p.d.f. of Eq. (6.66) with respect to \(\theta\) at integer "angular frequency" of \(m\). The Fourier transform of a p.d.f. is its characteristic function and \(C_m(r)\) can come from the "partial" Fourier transform of the joint p.d.f. of nonlinear phase noise, received intensity, and the phase of amplifier noise.

Using \(q_{\Phi,Y,\Theta_n}(\nu, y, m) = \mathcal{F}_\omega^{-1} \{ \Psi_{\Phi,Y,\Theta_n}(\nu, y, m) \} \) to denote the partial p.d.f. and characteristic function as the inverse Fourier transform of the characteristic function of \(\Psi_{\Phi,Y,\Theta_n}(\nu, y, m)\) from Eq. (6.B.2), we have \(q_{\Phi,R,\Theta_n}(\phi, r, m) = 2r q_{\Phi,Y,\Theta_n}(\phi, r^2, m)\) for the amplitude of \(R\) instead of intensity of \(Y\). The Fourier coefficients of \(C_m(r)\) are

\[
C_m(r) = 2r q_{\Phi,Y,\Theta_n}^*(\nu, y, m) \frac{\langle \Phi_{NL} \rangle}{\rho_s + 1/2, r^2, m}. \quad (6.68)
\]

Using Eq. (6.B.10) and \(\theta_0 = 0\), we obtain (Gradshteyn and Ryzhik, 1980, §8.431)

\[
q_{\Phi,Y,\Theta_n}(\nu, y, m) = \frac{\Psi_{\Phi}(\nu)}{2\sigma_\nu^2} \exp \left( \frac{y + |\xi_\nu|^2}{2\sigma_\nu^2} \right) I_m \left( \frac{\sqrt{y} |\xi_\nu|}{\sigma_\nu^2} \right). \quad (6.69)
\]

Based on Eqs. (6.68) and (6.69), we obtain

\[
C_m(r) = \frac{r \Psi_m}{s_m} \exp \left( -\frac{r^2 + \alpha_m^2}{2s_m} \right) I_m \left( \frac{\alpha_m r}{s_m} \right), \quad m \geq 1, \quad (6.70)
\]
Figure 6.20. The distribution of received electric field for binary PSK signal. The contour lines are in logarithmic scale and both x and y axes are normalized to unity mean amplitude. The dashed line is the decision boundary. The SNR is $\rho_s = 18$ (12.6 dB) and the mean nonlinear phase shift is $\langle \Phi_{NL} \rangle = 2$ rad.

with

$$\Psi_m = \Psi_{\Phi} \left( m \frac{\langle \Phi_{NL} \rangle}{\rho_s + 1/2} \right),$$

(6.71)

$$\alpha_m = \sqrt{\rho_s \sec \left( \frac{jm}{2} \frac{\langle \Phi_{NL} \rangle}{\rho_s + 1/2} \right)},$$

(6.72)

and

$$s_m = \frac{1}{2} \left( jm \frac{\langle \Phi_{NL} \rangle}{\rho_s + 1/2} \right)^{-\frac{1}{2}} \tan \left( \frac{jm}{2} \frac{\langle \Phi_{NL} \rangle}{\rho_s + 1/2} \right),$$

(6.73)

where $\Psi_{\Phi}(\nu)$ is the characteristic function of the normalized nonlinear phase noise of Eq. (5.48).

For a PSK signal with $\theta_0 \in \{0, \pi\}$, Figure 6.20 shows the distribution of the received electric field similar to Fig. 5.1(b). The contour lines of Fig. 6.20 are the logarithmic of the p.d.f. of $\frac{1}{2} \left[ p_{R,\Phi_0}(r, \theta) + p_{R,\Phi_\pi}(r, \theta) \right]$ after the conversion to rectangular coordinate. The SNR of $\rho_s = 18$ (12.6
dB) is chosen for an error probability of $10^{-9}$ for binary PSK signal without nonlinear phase noise. The mean nonlinear phase shift of $\langle \Phi_{\text{NL}} \rangle = 2$ rad is close to the optimal operating point for linearly compensated PSK signals from Fig. 6.10.

Similar to those of Fig. 5.1, the helix shape distribution of Fig. 6.20 clearly shows that the rotation due to nonlinear phase noise is correlated with the amplitude (the distance to the origin). The optimal MAP detector is derived here.

The derivation of joint p.d.f. here follows Huang and Ho (2003). Using method in quantum field theory, a p.d.f. similar to Eq. (6.66) was derived in Turitsyn et al. (2003). Similar p.d.f. was also derived in Mecozzi (1994a,b).

### 5.2 Optimal MAP Detector

For a PSK signal with $\theta_0 \in \{0, \pi\}$, using the Neyman-Pearson lemma with unity likelihood ratio (McDonough and Whalen, 1995, §5.3), the optimal decision regions of the MAP detector are given by

$$
R_0 = \left\{ r, \theta \mid p_{R, \Phi_r \mid \theta}(r, \theta) \geq p_{R, \Phi_r \mid \pi}(r, \theta) \right\},
$$

(6.74)

$$
R_1 = \left\{ r, \theta \mid p_{R, \Phi_r \mid \pi}(r, \theta) > p_{R, \Phi_r \mid 0}(r, \theta) \right\}
$$

(6.75)

for the transmitted phases of $\theta_0 = 0$ and $\theta_0 = \pi$, respectively. The error probability is

$$
p_e = \frac{1}{2} \int_{R_0} p_{R, \Phi_r \mid \pi}(r, \theta) d\theta d\theta + \frac{1}{2} \int_{R_1} p_{R, \Phi_r \mid 0}(r, \theta) d\theta d\theta
$$

$$
= \int_{R_0} p_{R, \Phi_r \mid \pi}(r, \theta) d\theta d\theta
$$

$$
= \int_{R_1} p_{R, \Phi_r \mid 0}(r, \theta) d\theta d\theta.
$$

(6.76)

The decision regions of $R_0$ and $R_1$ have a boundary of $\theta_c(r) \pm \pi/2$, $r \geq 0$, where $\theta_c(r)$ is the center phase of the decision regions of $R_0$. With the joint p.d.f. of Eq. (6.66), the error probability for MAP detector of Eq. (6.76) becomes

$$
p_e = \frac{1}{2} - \frac{2}{\pi} \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k + 1} \int_0^\infty \Re \left\{ C_{2k+1}^*(r) e^{j(2k+1)\theta_c(r)} \right\} dr.
$$

(6.77)

Using only the p.d.f. of $p_{R, \Phi_r \mid \theta}(r, \theta)$, because of the relationship of $p_{R, \Phi_r \mid \pi}(r, \theta) = p_{R, \Phi_r \mid 0}(r, \theta - \pi)$, the center phase of $\theta_c(r)$ can be deter-
mined by the relationship of

$$p_{R,\Phi_c|0}(r,\theta_c(r) + \frac{\pi}{2}) = p_{R,\Phi_c|0}(r,\theta_c(r) - \frac{\pi}{2}). \quad (6.78)$$

Figure 6.20 also shows the decision boundary of $\theta_c(r) \pm \pi/2$ based on Eqs. (6.74) and (6.75) with center phase given by Eq. (6.78). The decision boundary is the “valley” between two peaks of the p.d.f. of $p_{R,\Phi_c|0}(r,\theta)$ and $p_{R,\Phi_c|\pi}(r,\theta)$, respectively.

The optimal MAP detector can be implemented based on the decision boundary of Fig. 6.20 of the Yin-Yang detector. The same MAP detector can also be implemented as a nonlinear compensator by adding an angle of $\theta_c(r)$ from Eq. (6.78) to the received phase. The Yin-Yang detector is equivalent to the nonlinear compensator.

In this section, the optimal MAP detector is derived based on $\theta_c(r)$ as a function of the received amplitude. Because the relationship between intensity and amplitude is a monotonic function of $Y = R^2$, $R \geq 0$, the nonlinear phase noise is considered to be compensated by the received intensity instead of the received amplitude.

The more popular differential phase-shift keying (DPSK) signals can also be compensated using the differential phase of $\theta_c(r_1) - \theta_c(r_2)$, where $r_1$ and $r_2$ are the received amplitudes in two consecutive symbols. While the compensator for DPSK signals can be constructed using the center phase of Eq. (6.78), the evaluation of the error probability of nonlinearly compensated DPSK signals is difficult.

### 5.3 Optimal MMSE Detector

The optimal MMSE compensator estimates the normalized nonlinear phase using the received intensity by the conditional mean of $E\{\Phi|R\} = E\{\Phi|Y\} = E\{\Phi|R^2\}$ from Appendix 6.A. The conditional mean is the Bayes estimator that minimizes the variance of the residual nonlinear phase noise without the constraint of linearity (McDonough and Whalen, 1995, §10.2). We call the estimation of $E\{\Phi|R\}$ an estimation based on received intensity although it is actually based on received amplitude.

To find the conditional mean $E\{\Phi|R\}$, either the conditional p.d.f. of $p_{\Phi|R}(\theta)$ or the conditional characteristic function of $\Psi_{\Phi|R}(\nu)$, or equivalently $\Psi_{\Phi|Y}(\nu)$, is required. The conditional characteristic function of $\Psi_{\Phi|Y}(\nu)$ can be found using the relationship of

$$\Psi_{\Phi|Y}(\nu) = \frac{q_{\Phi,Y}(\nu, y)}{p_Y(y)}, \quad (6.79)$$

where $q_{\Phi,Y}(\nu, y)$ of Eq. (6.B.8) is the partial characteristic function of normalized nonlinear phase noise and p.d.f. of received intensity and
\(p_Y(y)\) of Eq. (6.8B.9) is the p.d.f. of the received intensity. Using both Eqs. (6.B.8) and (6.B.9), we obtain

\[
\Psi_{\Phi|Y}(\nu) = \frac{\Psi_{\Phi}(\nu) \exp \left[ -\frac{y + |\xi_\nu|^2}{2\sigma_\nu^2} \right] I_0 \left( \sqrt{\frac{y |\xi_\nu|}{\sigma_\nu^2}} \right)}{2\sigma_\nu^2 \exp \left[-(y + \rho_s)\right] I_0 \left( 2\sqrt{y\rho_s} \right)}.
\]

(6.80)

The optimal MMSE compensator is the conditional mean of the normalized nonlinear phase noise given the received intensity of \(Y = R^2\), we obtain

\[
E\{\Phi|R\} = -j \frac{d}{d\nu} \Psi_{\Phi|Y}(\nu|y) \Bigg|_{\nu=0, y=r^2} = \frac{1}{6} + \frac{1}{3} \rho_s + \frac{1}{3} r^2 + \frac{\sqrt{\rho_s r} I_1(2r\sqrt{\rho_s})}{3 I_0(2r\sqrt{\rho_s})}.
\]

(6.81)

Other than the normalization, the MMSE compensator of Eq. (6.81) is similar to that of Eq. (6.A.15) for finite number of fiber spans from Appendix 6.A and can be derived by the method there. The optimal MMSE compensator Eq. (6.81) depends solely on the system SNR.

The normalized residual nonlinear phase noise is \(\phi_e = \phi - E\{\Phi|R\}\), its conditional variance is

\[
\sigma^2_{\phi_e|R}(r) = -\frac{d^2}{d\nu^2} \Psi_{\Phi|Y}(\nu|y) \Bigg|_{\nu=0, y=r^2} - E\{\Phi|R\}^2
\]

\[
= \frac{1 + 4(\rho_s + r^2)}{90} + \frac{r^2 \rho_s}{9} + \frac{\sqrt{\rho_s r} I_1(2r\sqrt{\rho_s})}{45 I_0(2r\sqrt{\rho_s})} - \frac{r^2 \rho_s I_1^2(2r\sqrt{\rho_s})}{9 I_0^2(2r\sqrt{\rho_s})},
\]

(6.82)

that also depends solely on SNR.

The variance of the normalized residual nonlinear phase noise can be numerically integrated as

\[
\sigma^2_{\phi_e} = \int_0^\infty \sigma^2_{\phi_e|R}(r)p_R(r)dr,
\]

(6.83)

where \(p_R(r)\) is the p.d.f. of the received amplitude of Eq. (6.67).

Similar to the error probability of Eq. (6.77), the error probability for MMSE detector is

\[
p_e = \frac{1}{2} - \frac{2}{\pi} \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k + 1} \int_0^\infty \left\{ C_{2k+1}^* \exp \left[ \frac{j(2k + 1) \langle \Phi_{NL} \rangle}{\rho_s + 1/2} E\{\Phi|R\} \right] \right\} dr.
\]

(6.84)
Figure 6.21. The STD of normalized nonlinear phase noise with and without compensation. The STD with linear and nonlinear compensator are showed as dotted- and solid lines, respectively.

The factor of \( \langle \Phi_{NL} \rangle / (\rho_s + 1/2) \) scales the normalized nonlinear phase noise \( \Phi \) in Eq. (6.81) to the actual nonlinear phase noise \( \Phi_{NL} \).

Figure 6.21 plots the STD of the normalized nonlinear phase noise with and without compensation as a function of SNR \( \rho_s \). The STD of residual nonlinear phase noise with linear and nonlinear compensator are shown as almost overlapped solid and dotted lines, respectively. The STD of nonlinear compensator is about 0.2% less than that of linear compensator. Figure 6.21 confirms the results of Fig. 6.4 that linear and nonlinear MMSE compensator performs the same in term of the variance of residual nonlinear phase noise. However, as shown later, the error probability of nonlinear MMSE compensator is far better that for linear MMSE compensator.

5.4 Numerical Results

Figure 6.22 plots the center phase of \( \theta_c(r) \) from Eq. (6.78) as a function of the normalized received intensity of \( r^2/\rho_s \). The system parameters of Fig. 6.22 are the same as that of Fig. 6.20. As discussed earlier, the center phase of Yin-Yang detector is the same as the compensated phase of a compensator. The center (or compensated) phase of the nonlinear MMSE compensator of Eq. (6.81) is also plotted in Fig. 6.22 for comparison. The phase of Eq. (6.81) is scaled by \( \langle \Phi_{NL} \rangle / (\rho_s + 1/2) \). The
Figure 6.22. The compensated phase as a function of received intensity. System parameters are the same as that in Fig. 6.20.

compensated phases of the linear compensator designed by MMSE or MAP criteria are also plotted in Fig. 6.22 as dashed-lines for comparison. In Fig. 6.22, the received intensity is normalized with respect to the SNR of $\rho_s$.

From Fig. 6.22, nonlinear MMSE or MAP compensated curves are very close to the linear compensated curves, especially when the received intensity is near its mean value of about $r^2/\rho_s = 1$. When the STD of Fig. 6.21 is evaluated, the STD is mainly contributed from the region where the random variable is close to its mean value. Compared the linear and nonlinear compensated phases of Fig. 6.22, the STD of Fig. 6.21 for linear and nonlinear MMSE compensator should not have significant difference. From Fig. 6.21, the linear and nonlinear MAP compensated phases are also very close to each other. In the region close to $r^2/\rho_s = 1$, because the nonlinear MAP compensated phase is closer to the linear MAP compensated phase than the nonlinear MMSE compensated phase, the linear MAP compensated phase has better performance than the nonlinear MMSE compensated phase.

Figure 6.23 shows the error probability given by both Eqs. (6.77) and (6.84) for optimal MAP and MMSE detectors, respectively, as a function of SNR $\rho_s$. Figure 6.23 also plots the error probability of a PSK signal without nonlinear phase noise of Eq. (3.78). From Fig. 6.23, the optimal MMSE detector does not minimize the error probability. The optimal
compensated phase of Eq. (6.78) always gives a smaller error probability that the compensated phase of Eq. (6.81).

Figure 6.24 shows the SNR penalty of PSK signal for an error probability of $10^{-9}$ calculated by the MAP of Eq. (6.77) and MMSE of Eq. (6.84) error probability formulae. The SNR penalty with linear compensator is also plotted as dashed lines for comparison. The SNR penalty without compensation is also shown for comparison.

From Fig. 6.24, the nonlinear MMSE compensator performs up to 0.23 dB better than the linear MMSE compensator. Although Fig. 6.21 shows that linear and nonlinear MMSE compensators perform almost the same in term of the STD of residual nonlinear phase noise, the error probability has significant difference. The comparison between Figs. 6.21 and 6.24 shows that the variance does not correlate well with the error probability. From Fig. 6.24, the optimal linear MAP detector performs very close to the optimal nonlinear MAP detector. The optimal nonlinear MAP detector performs only up to 0.14 dB better than the linear MAP detector.

For binary PSK signal, Table 6.1 shows the mean nonlinear phase shift corresponding to a system with 1-dB SNR penalty and the optimal operating point. The optimal operating point is found by the condition

\[ \text{Equation}(6.78) \]

\[ \text{Equation}(6.81) \]

\[ \text{Equation}(6.77) \]

\[ \text{Equation}(6.84) \]
Figure 6.24. The SNR penalty of PSK signal as a function of mean nonlinear phase shift $\langle \Phi_{NL} \rangle$.

Table 6.1. PSK Signals with Nonlinear Phase Noise Compensation.

<table>
<thead>
<tr>
<th></th>
<th>Mean Nonlinear Phase Shift $\langle \Phi_{NL} \rangle$</th>
<th>Max. Diff. to Nonlin. MAP (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-dB Penalty (rad)</td>
<td>Optimal Point (rad)</td>
</tr>
<tr>
<td>Without Compensation</td>
<td>1.00</td>
<td>1.25</td>
</tr>
<tr>
<td>MMSE</td>
<td>Linear</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td>Nonlinear</td>
<td>2.31</td>
</tr>
<tr>
<td>MAP</td>
<td>Linear</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>Nonlinear</td>
<td>2.35</td>
</tr>
</tbody>
</table>

that the increase of SNR penalty is less than the increase of SNR that is proportional to the mean nonlinear phase shift. Because of the steepness of the slope, the optimal MAP detector actually gives smaller optimal operating point than other compensation schemes.

On the basis of the distribution of a received signal with nonlinear phase noise, the optimal MAP detector is derived for a phase-modulated signal to minimize the error probability. The error probability of nonlinear MAP and MMSE detector is also evaluated using the distribution of the received signal. Having the same variance of residual nonlinear phase noise, nonlinear MMSE compensator performs up to 0.23 dB better than
linear MMSE compensator. The optimal nonlinear MAP compensator performs very close to the optimal linear MAP compensator with a difference less than 0.14 dB.

This section discussed the optimal nonlinear compensator for PSK systems with infinite number of fiber spans based on the model of Appendix 5.A and 6.B. For system with finite number of fiber span, similar nonlinear compensator can be derived using the model of Appendix 5.B.

6. Summary

The compensation of nonlinear phase noise is studied in detail. Based on the correlation of received intensity with nonlinear phase noise, the simplest compensator is linear compensator to use a scale version of the received intensity to reduce nonlinear phase noise. The linear compensator can be optimized using the MMSE criterion to minimize the variance of the residual nonlinear phase noise. The variance of the residual nonlinear phase noise is about a quarter of that of nonlinear phase noise, doubling the transmission distance if nonlinear phase noise is the major impairment.

While the MMSE criterion does not provide a minimum error probability, the linear MAP compensator optimized by numerical methods is able to well approximate the optimal nonlinear detector. In practice, the optimal detector can be implemented as a Yin-Yang detector or a compensator.

The optimal compensation scheme is distributed or mid-span compensator. Even with a single compensator, if located 2/3 from the beginning of the fiber link, the STD of residual nonlinear phase noise is reduced to 1/3 that without compensation, far better than a compensator co-located with the receiver.

Ideally, the nonlinear phase noise compensators should minimize the error probability of a phase-modulated system with nonlinear phase noise. The nonlinear MAP compensator is found numerically as the optimal detector to minimize the error probability. In term of variance, nonlinear MMSE compensator performs the same as linear MMSE compensator. Nonlinear MAP compensator performs up to about 0.21 dB better than nonlinear MMSE compensator but just 0.14 dB better than linear MAP compensator.
APPENDIX 6.A: Nonlinear MMSE Compensation

Using the received intensity of $P_N = r_r^2$, the optimal compensator is the MMSE estimator based on the conditional mean of

$$E\{\Phi_{NL}|r_r\}, \quad (6.A.1)$$

where $E\{\cdot\}$ denotes expectation.

The variance of the difference of $\Phi_{NL} - E\{\Phi_{NL}|r_r\}$ is minimized using the nonlinear MMSE compensator of Eq. (6.A.1). The optimal compensator can be implemented using a phase modulator driven by the waveform of Eq. (6.A.1) or an electrical compensator to calculate $\hat{E}_r \exp[j E\{\Phi_{NL}|r_r\}]$ using electronic signal processing. With the received electric field $E_r$, the detector can also be implemented as a Yin-Yang detector. If the received phase is demodulated as $\Phi_r = \arg(E_r)$, the compensator is a combiner of $\Phi_r + E\{\Phi_{NL}|r_r\}$. There is no difference between $E\{\Phi_{NL}|r_r^2\}$ and the expectation of Eq. (6.A.1) because the received amplitude is a positive number $r_r \geq 0$. The estimator of Eq. (6.A.1) is called a correction by received intensity although the expression of Eq. (6.A.1) is based on received amplitude.

Without loss of generality, assume that the transmitted electric field is $E_0 = A$, we need to calculate the estimation of term by term of

$$E\left\{|A + \tilde{n}_1 + \cdots + \tilde{n}_{N_A}|^2\left|\tilde{E}_0 + \tilde{n}_1 + \cdots + \tilde{n}_{N_A}\right|\right\}, \quad (6.A.2)$$

The estimation of Eq. (6.A.2) is a special case of the estimation of

$$\varphi(\sigma_{\zeta_1}, \sigma_{\zeta_2}) = E\left\{|\tilde{E}_0 + \tilde{\zeta}_1|^2\right\}, \quad (6.A.3)$$

for a real value $A$ and two zero-mean complex circular Gaussian variables $\zeta_1$ and $\zeta_2$, where $\rho_2 = |E_0 + \zeta_1 + \zeta_2|$, $E\{\zeta_1\} = E\{\zeta_2\} = 0$, $E\{|\zeta_1|^2\} = 2\sigma_{\zeta_1}^2$, and $E\{|\zeta_2|^2\} = 2\sigma_{\zeta_2}^2$. The estimation of Eq. (6.A.3) is used to derive the MMSE estimator of Eq. (6.A.1) and the value of Eq. (6.A.2). The real value of $A$ can be used to represent the amplitude of $|E_0|$ that is a constant for both PSK and DPSK signals.

Define $x_1 + jy_1 = A + \zeta_1$ and $x_2 + jy_2 = A + \zeta_1 + \zeta_2$, the joint p.d.f. of $x_1, y_1, x_2, y_2$ is

$$f_{xy}(x_1, y_1, x_2, y_2) = \frac{1}{(2\pi\sigma_{\zeta_1}\sigma_{\zeta_2})^2} \exp\left[-\frac{(x_1 - A)^2 + y_1^2}{2\sigma_{\zeta_1}^2}\right]\times \exp\left[-\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2\sigma_{\zeta_2}^2}\right]. \quad (6.A.4)$$

Using the joint p.d.f. of Eq. (6.A.4), changing the variable to $x_2 = \rho_2 \cos \theta_2$ and $y_2 = \rho_2 \sin \theta_2$, the marginal p.d.f. of $\rho_2$ is a Rice distribution of

$$f_{\rho_2}(\rho_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{xy}(x_1, y_1, \rho_2 \cos \theta_2, \rho_2 \sin \theta_2) \rho_2 d\theta_2 dx_1 dy_1
\quad = \frac{\rho_2}{(\sigma_1^2 + \rho_2^2)^{1/2}} \exp\left(-\frac{A^2 + \rho_2^2}{\sigma_1^2 + \rho_2^2}\right) I_0\left(\frac{A\rho_2}{\sigma_1^2 + \rho_2^2}\right). \quad (6.A.5)$$

The estimation of Eq. (6.A.3) is

$$\varphi(\sigma_{\zeta_1}, \sigma_{\zeta_2}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1^2 + y_1^2) f_{xy|\rho_2}(x_1, y_1|\rho_2) dx_1 dy_1, \quad (6.A.6)$$
where the conditional p.d.f. is
\[ f_{x,y|\rho_2}(x_1,y_1|\rho_2) = \frac{1}{f_{\rho_2}(\rho_2)} \int_{-\infty}^{+\infty} f_{x,y}(x_1,y_1,\rho_2 \cos \theta_2, \rho_2 \sin \theta_2) \, d\rho_2 \, d\theta_2. \tag{6.6.7} \]

The integration of Eq. (6.6.6) becomes
\[ \frac{\rho_2}{f_{\rho_2}(\rho_2)} \int_{-\pi}^{+\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1^2 + y_1^2) f_{x,y}(x_1,y_1,\rho_2 \cos \theta_2, \rho_2 \sin \theta_2) \, dx_1 \, dy_1 \, d\theta_2. \tag{6.6.8} \]

Integrated first over \(x_1\) and \(x_2\), we obtain
\[ \frac{1}{2\pi(\sigma_{c_1}^2 + \sigma_{c_2}^2)^2} \int_{-\pi}^{+\pi} e^{-A\rho_2 \cos \theta_2/(\sigma_{c_1}^2 + \sigma_{c_2}^2)} \left[ \sigma_{c_1}^4 (\rho_2^2 + 2\sigma_{c_2}^2) + \rho_2^2 \sigma_{c_1}^2 \sigma_{c_2}^2 \cos \theta_2 \right] \, d\theta_2. \tag{6.A.9} \]

The integration of Eq. (6.A.6) is
\[ \varphi(\sigma_{c_1}, \sigma_{c_2}) = \frac{\sigma_{c_1}^4 (\rho_2^2 + 2\sigma_{c_2}^2) + \sigma_{c_2}^4 (A^2 + 2\sigma_{c_1}^2)}{(\sigma_{c_1}^2 + \sigma_{c_2}^2)^2} + \frac{2A\rho_2 \sigma_{c_1}^2 \sigma_{c_2}^2 I_0 (\rho_2/(\sigma_{c_1}^2 + \sigma_{c_2}^2))}{(\sigma_{c_1}^2 + \sigma_{c_2}^2)^2} \frac{I_0 (A/(\sigma_{c_1}^2 + \sigma_{c_2}^2))}{I_0 (A/(\sigma_{c_1}^2 + \sigma_{c_2}^2))}. \tag{6.A.10} \]

For the case of high SNR, the expectation of Eq. (6.A.3) can be approximated by
\[ \varphi(\sigma_{c_1}, \sigma_{c_2}) = E\left\{ |E_0 + \zeta_1|^2 \right| E_0 + \zeta_1 + \zeta_2 \right\} \approx E\left\{ A^2 + 2A(x_1 - A) \right\} \tag{6.A.11} \]

Using the approximation of Eq. (6.A.11), the expectation of Eq. (6.A.3) becomes
\[ \varphi(\sigma_{c_1}, \sigma_{c_2}) \approx \frac{\sigma_{c_2}^2}{\sigma_{c_1}^2 + \sigma_{c_2}^2} A^2 + \frac{\sigma_{c_1}^2}{\sigma_{c_1}^2 + \sigma_{c_2}^2} \rho_2^2. \tag{6.A.12} \]

Compared the estimation of Eq. (6.6.2) with (6.6.3), we have \( \zeta_1 = n_1 + \cdots + n_k \) and \( \zeta_2 = n_{k+1} + \cdots + n_{N_n} \) that are zero-mean complex circular Gaussian random variables with variance of \( E(\zeta_1^2) = 2k\sigma_0^2 \) and \( E(\zeta_2^2) = 2(N_n - k)\sigma_0^2 \). Using the function of \( \varphi(\sigma_{c_1}, \sigma_{c_2}) \) defines in Eq. (6.A.3), we obtain
\[ E\left\{ |A + n_1 + \cdots + n_k|^2 \right| \rho_r \right\} = \varphi \left( \sqrt{k} \sigma_0, \sqrt{N_n - k} \sigma_0 \right). \tag{6.A.13} \]

For the estimator of Eq. (6.6.1), we obtain
\[ E\left\{ \Phi_{NL} | \rho_r \right\} = \gamma L_{\text{eff}} \sum_{k=1}^{N_n} \varphi \left( \sqrt{k} \sigma_0, \sqrt{N_n - k} \sigma_0 \right). \tag{6.A.14} \]

Substitute the function of \( \varphi(\sigma_{c_1}, \sigma_{c_2}) \) of Eq. (6.A.10) to Eq. (6.A.14), we get
\[ E\left\{ \Phi_{NL} | \rho_r \right\} = \gamma L_{\text{eff}} \left\{ \frac{(2N_n - 1)(N_n - 1)}{6N_n} \rho_r^2 + \frac{2N_n^2 - 1}{3} \sigma_0^2 \right\} + \frac{2N_n + 1}{6N_n} \rho_r^2 + \frac{(N_n^2 - 1)A \rho_r}{3N_n} \right\} \tag{6.A.15} \]
To derive Eq. (6.A.15), the relationships of Eqs. (6.18) to (6.20) are used. Because $I_1(x)/I_0(x) \sim x/2$ for small $x < 0.5$, the phase estimator of Eq. (6.A.15) is a linear function of the received intensity of $\rho_r^2$ when the received signal is small. Because $I_1(x)/I_0(x) \sim 1$ for large $x > 10$, the phase estimator of Eq. (6.A.15) is a quadratic function of the received amplitude of $\rho_r$ when the received signal is large. The interested systems usually have high SNR and the received signal has an amplitude around $A$, i.e., $\rho_r \sim A$, the phase estimator of Eq. (6.A.15) mostly functions as a nonlinear compensator.

For large number of fiber span of $N_A \gg 1$ and high SNR, the phase estimator of Eq. (6.A.15) is approximately

$$E\{\Phi_{NL}\rho_r\} \approx \frac{\langle \Phi_{NL} \rangle}{3} \left[ 1 + \frac{\rho_r^2}{A^2} + \frac{\rho_r I_1(\rho_A/\rho_r)}{A I_0(\rho_A/\rho_r)} \right], \quad \text{(6.A.16)}$$

where the mean nonlinear phase noise $\langle \Phi_{NL} \rangle$ is given by Eq. (6.16). The estimation of Eq. (6.A.16) divides the estimation into three parts, the first term is one-third the mean nonlinear phase shift, and the second term proportional to received intensity, and third term is the nonlinear term.

While most of the above discussion is related to PSK systems, the optimal compensator for DPSK systems follows the same principle. For DPSK signals, the differential phase of $\Phi_r(t) - \Phi_r(t - T)$ is detected, where $T$ is the symbol period. The optimal MMSE compensator should be

$$E\{\Phi_{NL}(t)\rho_r(t)\} - E\{\Phi_{NL}(t - T)\rho_r(t - T)\}, \quad \text{(6.A.17)}$$

similar to Eq. (6.A.15).

**APPENDIX 6.B: Joint Characteristic Function**

Here, the joint characteristic function of nonlinear phase noise, received intensity, and the phase of amplifier noise is derived. Corresponding to the Fourier coefficients, only the characteristic function at integer “frequency” of the phase of amplifier noise of $\Theta_n$ is derived. Using the partial p.d.f. and characteristic function of Eq. (5.A.10), with $m$ as a non-negative integer, we obtain (Gradshteyn and Ryzhik, 1980, §8.431)

$$\Psi_{\Phi,\Theta_m}(\nu,\omega, m) = \frac{\Psi_{\Phi}(\nu)}{2\sigma_{\Phi}^2} \int_0^\infty \exp \left[ -\frac{y + \xi_\nu^2}{2\sigma_{\Phi}^2} \right] I_m \left( \frac{\sqrt{y}}{\sigma_{\Phi}} \right) e^{iy\omega} dy, \quad m \geq 0, \quad \text{(6.B.1)}$$

and $\Psi_{\Phi,\Theta_m}(\nu,\omega, -m) = \Psi_{\Phi,\Theta_m}(\nu,\omega, m)$, where $I_m(\cdot)$ is the $m$th-order modified Bessel function of the first kind.


$$\Psi_{\Phi,\Theta_m}(\nu,\omega, m) = \frac{\Psi_{\Phi}(\nu) \exp \left[ \frac{-\xi_\nu^2}{2\sigma_{\Phi}^2} \right]}{2\sigma_{\Phi}^2} \frac{\Gamma(m/2 + 1)}{m!} \times \frac{\sqrt{2\sigma_{\nu}^{(m+1)/2}}}{(1 - 2j\omega\sigma_{\nu}^{2})^{1/2} |\xi_\nu|^{1/2}} \text{F}_1 \left( \frac{m}{2} + 1; m + 1; \gamma_{\nu,\omega} \right), \quad m \geq 0, \quad \text{(6.B.2)}$$
where
\[ \gamma_{\nu, \omega} = \frac{1}{1 - 2j\omega \sigma_0^2} \frac{|\xi_{\omega}|^2}{2\sigma_x^2} \]
\[ = \frac{2j\nu}{(\sqrt{j\nu} - j\omega \tan \sqrt{j\nu})} \sin (2\sqrt{j\nu}) \rho_s, \] (6.3)
\[ \Gamma(\cdot) \] is the Gamma function, and \( \Gamma_1(a; b; \cdot) \) is the confluent hypergeometric function of the first kind.

Using \( \gamma_{\nu} \) defined by Eq. (5.12), we can rewrite Eq. (6.2) as
\[ \Psi_{\Phi, \gamma_{\omega}}(\nu, \omega, m) = \Psi_{\Phi}(\nu) e^{-\gamma_{\nu}} \Gamma \left( \frac{1}{2} m + 1 \right) \frac{\gamma_{\omega}^{\frac{m}{2} + 1}}{\gamma_{\nu}} \Gamma_1 \left( \frac{m}{2} + 1; m + 1; \gamma_{\nu, \omega} \right). \] (6.4)

With \( \omega = 0 \) in Eq. (6.4), using Gradshteyn and Ryzhik (1980, §9.212), the joint Fourier coefficients of nonlinear phase noise and the phase of amplifier noise are
\[ \Psi_{\Phi, \gamma_{\omega}}(\nu, m) = \Psi_{\Phi}(\nu) e^{-\gamma_{\nu}} \gamma_{\nu} \frac{\gamma_{\omega}^m}{m!} \Gamma \left( \frac{m}{2} + 1 \right) \Gamma_1 \left( \frac{m}{2}; m + 1; \gamma_{\omega} \right), \] (6.5)
the same as Eq. (5.62). We may also derive Eq. (5.62) using the method here.

If \( m = 0 \) in Eq. (6.4), using the relationship of \( \Gamma_1(1; 1; z) = e^z \) (Gradshteyn and Ryzhik, 1980, §9.215), the joint characteristic function of nonlinear phase noise and the received intensity is
\[ \Psi_{\Phi, \gamma}(\nu, \omega) = \Psi_{\Phi}(\nu) \frac{\gamma_{\omega}^{\nu, \omega}}{\gamma_{\nu}} e^{-\gamma_{\omega, 0} + \gamma_{\nu, \omega}} = \frac{\Psi_{\Phi}(\nu)}{1 - 2j\omega \sigma_0^2} \exp \left[ \frac{j\omega |\xi_{\omega}|^2}{1 - 2j\omega \sigma_0^2} \right], \] (6.6) or
\[ \Psi_{\Phi, \gamma}(\nu, \omega) = \frac{1}{\cos \sqrt{j\nu} - j\omega \sin \sqrt{j\nu} \sqrt{j\nu}} \exp \left[ \frac{j\omega \rho_s}{\cos^2 \sqrt{j\nu} - j\omega \sin(2\sqrt{j\nu})} \right]. \] (6.7)

The characteristic function of Eq. (6.6) is the same as the characteristic function of noncentral \( \chi^2 \) distribution with two degrees of freedom. Using the joint characteristic function of Eq. (6.6), take the inverse Fourier transforms, we obtain
\[ \mathcal{F}_{\omega}^{-1} \{ \Psi_{\Phi, \gamma}(\nu, \omega) \} = \Psi_{\Phi}(\nu) \frac{2\sigma_0^2}{\sigma_0^2 + |\xi_{\omega}|^2} \exp \left[ -\frac{|\xi_{\omega}|^2}{2\sigma_0^2} \right] I_0 \left( 2\sqrt{j\nu} \rho_s \sigma_0^2 \right), \] (6.8)
similar to a noncentral \( \chi^2 \) distribution with two degrees of freedom. With \( \nu = 0 \), the p.d.f. of received intensity is
\[ p_Y(y) = \exp[-(y + \rho_s)] I_0(2\sqrt{j\nu} \rho_s). \] (6.9)

To simplify the characteristic function of Eq. (6.4) using the modified Bessel functions, from Gradshteyn and Ryzhik (1980, §9.212, §9.238) and similar to Jain
(1974), Jain and Blachman (1973) and Blachman (1981, 1988), we get

\[
\psi_{\Phi, Y, \Theta_n}(\nu, \omega, m) = \Psi_{\Phi}(\nu) \frac{\sqrt{\pi} \gamma_{\nu, \omega}^{3/2}}{2 \gamma_{\nu}} \exp \left( -\gamma_{\nu} + \frac{\gamma_{\nu, \omega}}{2} \right) \left[ I_{m-\frac{1}{2}} \left( \frac{\gamma_{\nu, \omega}}{2} \right) + I_{m+\frac{1}{2}} \left( \frac{\gamma_{\nu, \omega}}{2} \right) \right].
\]

(6.B.10)

In this chapter, the joint characteristic function of Eq. (6.B.10) is used to calculate the error probability of phase-modulated signal with linearly compensated nonlinear phase noise. The received phase is the linear combination of \( \Phi, Y, \) and \( \Theta_n \).