polarized signals are denoted as vectors and are written in bold face, such as $\mathbf{y}_k$. Unless stated otherwise, we employ a linear (x, y) basis for decomposing dually polarized signals. In Section 5 dealing with impairment compensation, matrix and vector quantities are denoted in bold face.

3. Optical detection methods

3.1 Noncoherent detection

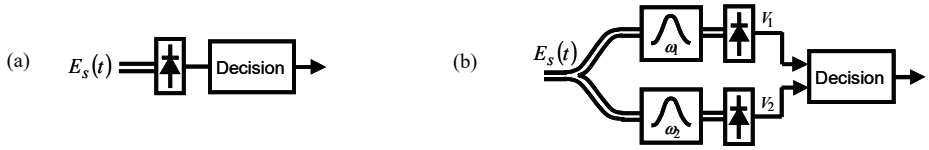

Fig. 1. Noncoherent receivers for (a) amplitude-shift modulation (ASK) and (b) binary frequency-shift keying (FSK).

In noncoherent detection, a receiver computes decision variables based on a *measurement of signal energy*. An example of noncoherent detection is direct detection of on-off-keying (OOK) using a simple photodiode (Fig. 1(a)). To encode more than one bit per symbol, multi-level amplitude-shift keying (ASK) – also known as pulse-amplitude modulation – can be used. Another example of noncoherent detection is frequency-shift keying (FSK) with wide frequency separation between the carriers. Fig. 1(b) shows a noncoherent receiver for binary FSK.

The limitations of noncoherent detection are: (a) detection based on energy measurement allows signals to encode only one degree of freedom (DOF) per polarization per carrier, reducing spectral efficiency and power efficiency, (b) the loss of phase information during detection is an irreversible transformation that prevents full equalization of linear channel impairments like CD and PMD by linear filters. Although maximum-likelihood sequence detection (MLSD) can be used to find the best estimate of the transmitted sequence given only a sequence of received intensities, the achievable performance is suboptimal compared with optical or electrical equalization making use of the full electric field [13].

3.2 Differentially coherent detection

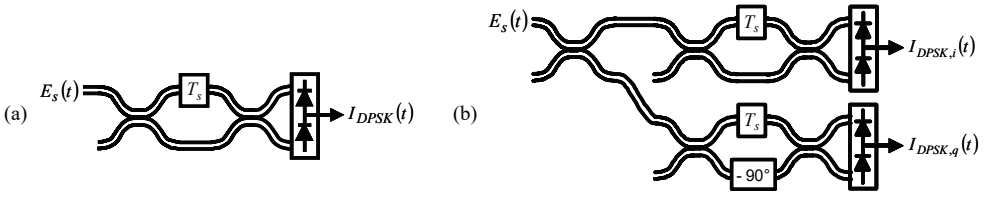

Fig. 2. Differentially coherent phase detection of (a) 2-DPSK (b) M-DPSK, $M > 2$.

In differentially coherent detection, a receiver computes decision variables based on *a measurement of differential phase between the symbol of interest and one or more reference symbol(s)*. In differential phase-shift keying (DPSK), the phase reference is provided by the previous symbol; in polarization-shift keying (PolSK), the phase reference is provided by the signal in the adjacent polarization. A binary DPSK receiver is shown in Fig. 2(a). Its output photocurrent is:

$$I_{DPSK}(t) = R\,\mathrm{Re}\left\{E_s(t)E_s^*(t-T_s)\right\}, \tag{1}$$

where $E_s(t)$ is the received signal, R is the responsivity of each photodiode, and T_s is the symbol period. This receiver can also be used to detect continuous-phase frequency-shift

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

keying (CPFSK), as the delay interferometer is equivalent to a delay-and-multiply demodulator. A receiver for *M*-ary DPSK, $M > 2$, can similarly be constructed as shown in Fig. 2(b). Its output photocurrents are:

$$I_{DPSK,i}(t) = \tfrac{1}{2} R \,\mathrm{Re}\left\{E_s(t)E_s^*(t-T_s)\right\}, \text{ and} \tag{2}$$

$$I_{DPSK,q}(t) = \tfrac{1}{2} R \,\mathrm{Im}\left\{E_s(t)E_s^*(t-T_s)\right\}. \tag{3}$$

A key motivation for employing differentially coherent detection is that binary DPSK has 2.8 dB higher sensitivity than noncoherent OOK at a BER of 10^{-9} [5]. However, the constraint that signal points can only differ in phase allows only one DOF per polarization per carrier, the same as noncoherent detection. As the photocurrents in Eq. (1) to (3) are not linear functions of the *E*-field, linear impairments, such as CD and PMD, also cannot be compensated fully in the electrical domain after photodetection.

A more advanced detector for *M*-ary DPSK is the multichip DPSK receiver, which has multiple DPSK receivers arranged in parallel, each with a different interferometer delay that is an integer multiple of T_s [14,15]. Since a multichip receiver compares the phase of the current symbol to a multiplicity of previous symbols, the extra information available to the detector enables higher sensitivity. In the limit that the number of parallel DPSK receivers is infinite, the performance of multi-chip DPSK approaches coherent PSK [15]. In practice, the number of parallel DPSK receivers required for good performance needs to be equal to the impulse duration of the channel divided by T_s. Although multi-chip DPSK does not require a local oscillator (LO) laser, carrier synchronization and polarization control, the hardware complexity can be a significant disadvantage.

3.3 Hybrid of noncoherent and differentially coherent detection

A hybrid of noncoherent and differentially coherent detection can be used to recover information from both amplitude and differential phase. One such format is polarization-shift keying (PolSK), which encodes information in the Stokes parameter. If we let $E_x(t) = a_x(t)e^{j\phi_x(t)}$ and $E_y(t) = a_y(t)e^{j\phi_y(t)}$ be the *E*-fields in the two polarizations, the Stokes parameters are $S_1 = a_x^2 - a_y^2$, $S_2 = 2a_x a_y \cos(\delta)$ and $S_3 = 2a_x a_y \sin(\delta)$, where $\delta(t) = \phi_x(t) - \phi_y(t)$ [16]. A PolSK receiver is shown in Fig. 3. The phase noise tolerance of PolSK is evident by examining S_1 to S_3. Firstly, S_1 is independent of phase. Secondly, S_2 and S_3 depend on the phase difference $\phi_x(t) - \phi_y(t)$. As $\phi_x(t)$ and $\phi_y(t)$ are both corrupted by the same phase noise of the transmitter (TX) laser, their arithmetic difference $\delta(t)$ is free of phase noise. In practice, the phase noise immunity of PolSK is limited by the bandwidth of the photodetectors [16]. It has been shown that 8-PolSK can tolerate laser linewidths as large as $\Delta\nu T_b \approx 0.01$ [17], which is about 100 times greater than the phase noise tolerance of coherent 8-QAM (Section 5.3.3). This was a significant advantage in the early 1990s, when symbol rates were only in the low GHz range. In modern systems, symbol rates of tens of GHz, in conjunction with tunable laser having linewidths <100 kHz, has diminished the advantages of PolSK. Recent results have shown that feedforward carrier synchronization enables coherent detection of 16-QAM at $\Delta\nu T_b \sim 10^{-5}$ [18], which is within the limits of current technology. As systems are increasingly driven by the need for high spectral efficiency, polarization-multiplexed QAM is likely to be more attractive because of its higher sensitivity (Section 4).

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

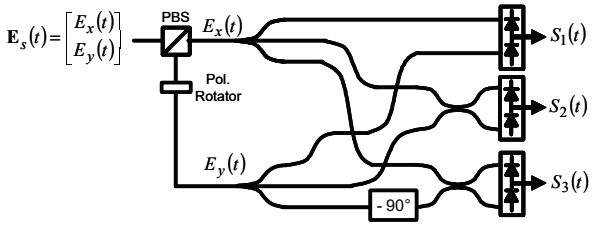

Fig. 3. Polarization-shift keying (PolSK) receiver.

3.4 Coherent detection

The most advanced detection method is coherent detection, where the receiver computes decision variables based on the *recovery of the full electric field*, which contains both amplitude and phase information. Coherent detection thus allows the greatest flexibility in modulation formats, as information can be encoded in amplitude and phase, or alternatively in both in-phase (I) and quadrature (Q) components of a carrier. Coherent detection requires the receiver to have knowledge of the carrier phase, as the received signal is demodulated by a LO that serves as an absolute phase reference. Traditionally, carrier synchronization has been performed by a phase-locked loop (PLL). Optical systems can use (i) an optical PLL (OPLL) that synchronizes the frequency and phase of the LO laser with the TX laser, or (ii) an electrical PLL where downconversion using a free-running LO laser is followed by a second-stage demodulation by an analog or digital electrical VCO whose frequency and phase are synchronized. Use of an electrical PLL can be advantageous in duplex systems, as the transceiver may use one laser as both TX and LO. PLLs are sensitive to propagation delay in the feedback path, and the delay requirement can be difficult to satisfy (Section 5.3.1). Feedforward (FF) carrier synchronization overcomes this problem. In addition, as a FF synchronizer uses both past and future symbols to estimate the carrier phase, it can achieve better performance than a PLL which, as a feedback system, can only employ past symbols. Recently, DSP has enabled polarization alignment and carrier synchronization to be performed in software.

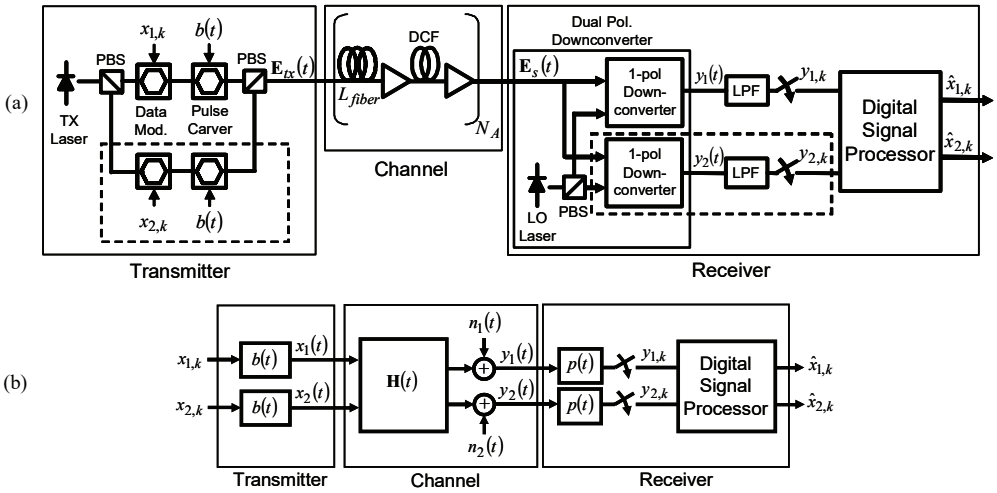

Fig. 4. Coherent transmission system (a) implementation, (b) system model.

A coherent transmission system and its canonical model are shown in Fig. 4. At the transmitter, Mach-Zehnder (MZ) modulators encode data symbols onto an optical carrier and perform pulse shaping. If polarization multiplexing is used, the TX laser output is split into

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

two orthogonal polarization components, which are modulated separately and combined in a polarization beam splitter (PBS). We can write the transmitted signal as:

$$\mathbf{E}_{tx}(t)=\begin{bmatrix}E_{tx,1}(t)\\E_{tx,2}(t)\end{bmatrix}=\sqrt{P_t}\sum_k \mathbf{x}_k\, b(t-kT_s)e^{j(\omega_s t+\phi_s(t))}\,, \tag{4}$$

where T_s is the symbol period, P_t is the average transmitted power, $b(t)$ is the pulse shape (e.g., non-return-to-zero (NRZ) or return-to-zero (RZ)) with the normalization $\int |b(t)|^2\, dt = T_s$, ω_s and $\phi_s(t)$ are the frequency and phase noise of the TX laser, and $\mathbf{x}_k=[x_{1,k}, x_{2,k}]^T$ is a 2×1 complex vector representing the *k*-th transmitted symbol. We assume that symbols have normalized energy: $E\left[|\mathbf{x}_k|^2\right]=1$. For single-polarization transmission, we can set the unused polarization component $x_{2,k}$ to zero.

The channel consists of N_A spans of fiber, with inline amplification and DCF after each span. In the absence of nonlinear effects, we can model the channel as a 2×2 matrix:

$$\mathbf{h}(t)=\begin{bmatrix}h_{11}(t) & h_{12}(t)\\ h_{21}(t) & h_{22}(t)\end{bmatrix} \overset{F}{\leftrightarrow} \begin{bmatrix}H_{11}(\omega) & H_{12}(\omega)\\ H_{21}(\omega) & H_{22}(\omega)\end{bmatrix}=\mathbf{H}(\omega), \tag{5}$$

where $h_{ij}(t)$ denote the response of the *i*-th output polarization due to an impulse applied at the *j*-th input polarization of the fiber. The choice of reference polarizations at the transmitter and receiver is arbitrary. Eq. (5) can describe CD, all orders of PMD, polarization-dependent loss (PDL), optical filtering effects and sampling time error [19]. In addition, a coherent optical system is corrupted by AWGN, which includes (i) amplified spontaneous emission (ASE) from inline amplifiers, (ii) receiver LO shot noise, and (iii) receiver thermal noise. In the canonical transmission model, we model the cumulative effect of these noises by an equivalent noise source $\mathbf{n}(t)=[n_1(t), n_2(t)]^T$ referred to the input of the receiver.

The *E*-field at the output of the fiber is $\mathbf{E}_s(t)=\left[E_{s,1}(t), E_{s,2}(t)\right]^T$, where:

$$E_{s,l}(t)=\sqrt{P_r}\sum_k\sum_{m=1}^{2} x_{m,k}c_{lm}(t-kT_s)e^{j(\omega_s t+\phi_s(t))}+E_{sp,l}(t)\,. \tag{6}$$

Under the assumption of Fig. 4 where inline amplification completely compensates propagation loss, $P_r = P_t$ is the average received power, $c_{lm}(t)=b(t)\otimes h_{lm}(t)$ is a normalized pulse shape, and $E_{sp,l}(t)$ is ASE noise in the *l*-th polarization. Assuming the N_A fiber spans are identical and all inline amplifiers have gain *G* and spontaneous emission factor n_{sp}, the two-sided power spectral density (psd) of $E_{sp,l}(t)$ is $S_{Esp}(f)=N_A n_{sp}\hbar\omega_s(G-1)/G$ W/Hz [20].

The first stage of a coherent receiver is a dual-polarization optoelectronic downconverter that recovers the baseband modulated signal. In a digital implementation, the analog outputs are lowpass filtered and sampled at $1/T = M/KT_s$, where M/K is a rational oversampling ratio. Channel impairments can then be compensated digitally before symbol detection.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

3.4.1 Single-polarization downconverter

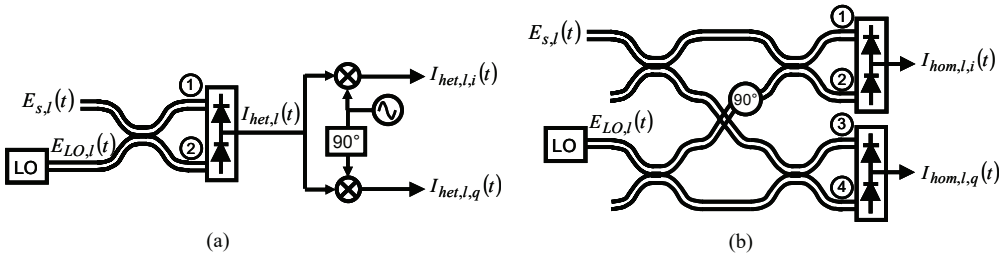

Fig. 5. Single-polarization downconverter employing a (a) heterodyne and (b) homodyne design.

We first consider a single-polarization downconverter, where the LO laser is aligned in the *l*-th polarization. Downconversion from optical passband to electrical baseband can be achieved in two ways: in a homodyne receiver, the frequency of the LO laser is tuned to that of the TX laser so the photoreceiver output is at baseband. In a heterodyne receiver, the LO and TX lasers differ by an intermediate frequency (IF), and an electrical LO is used to downconvert the IF signal to baseband. Both implementations are shown in Fig. 5. Although we show the optical hybrids as 3-dB fiber couplers, the same networks can be implemented in free-space optics using 50/50 beamsplitters; this was the approach taken by Tsukamoto [21]. Since a beamsplitter has the same transfer function as a fiber coupler, there is no difference in their performances.

In the heterodyne downconverter of Fig. 5(a), the optical frequency bands around $\omega_{LO}+\omega_{IF}$ and $\omega_{LO}-\omega_{IF}$ both map to the same IF. In order to avoid DWDM crosstalk and to avoid excess ASE from the unwanted image band, optical filtering is required before the downconverter. The output current of the balanced photodetector in Fig. 5(a) is:

$$I_{het,l}(t)=R\left(|E_1(t)|^2-|E_2(t)|^2\right)=2R\,\mathrm{Im}\left\{E_{s,l}(t)E_{LO,l}^*(t)\right\}+I_{sh,l}(t), \tag{7}$$

where $E_{LO,l}(t)=\sqrt{P_{LO,l}}\,e^{j(\omega_{LO}t+\phi_{LO}(t))}$ is the *E*-field of the LO laser, and $P_{LO,l}$, ω_{LO} and $\phi_{LO}(t)$ are its power, frequency and phase noise. $I_{sh,l}(t)$ is the LO shot noise. Assuming $P_{LO}>>P_s$, $I_{sh,l}(t)$ has a two-sided psd of $S_{Ish}(f)=qRP_{LO}$ A^2/Hz.. Substituting Eq. (6) into Eq. (7), we get:

$$I_{het,l}(t)=2R\sqrt{P_{LO,l}}\left(\sqrt{P_r}\left(y_{li}(t)\sin(\omega_{IF}t)+y_{lq}(t)\cos(\omega_{IF}t)\right)+E'_{sp,l}(t)\right)+I_{sh,l}(t), \tag{8}$$

where $\omega_{IF}=\omega_s-\omega_{LO}$ is the IF, $\phi(t)=\phi_s(t)-\phi_{LO}(t)$ is the carrier phase, and $y_{li}(t)$ and $y_{lq}(t)$ are the real and imaginary parts of:

$$y_{l0}(t)=\sum_k\sum_{m=1}^{2}x_{m,k}c_{lm}(t-kT_s)e^{j\phi(t)}. \tag{9}$$

The term $2R\sqrt{P_{LO,r}}E'_{sp,l}(t)$ in Eq. (8) is sometimes called *LO-spontaneous beat noise*, and $E'_{sp,l}(t)=\mathrm{Im}\left\{E_{sp,l}(t)e^{-j(\omega_{LO}t+\phi_{LO}(t))}\right\}$ has a two-sided psd of $\frac{1}{2}S_{Esp}(f)$.

It can similarly be shown that the currents at the outputs of the balanced photodetectors in the homodyne downconverter (Fig. 5(b)) are:

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

$$I_{hom,l,i}(t) = R\left(|E_1(t)|^2 - |E_2(t)|^2\right) = R\sqrt{P_{LO,l}}\left(\sqrt{P_r}\, y_{li}(t) + E'_{sp,li}(t)\right) + I_{sh,li}(t), \text{ and} \tag{10}$$

$$I_{hom,l,q}(t) = R\left(|E_3(t)|^2 - |E_4(t)|^2\right) = R\sqrt{P_{LO,l}}\left(\sqrt{P_r}\, y_{lq}(t) + E'_{sp,lq}(t)\right) + I_{sh,lq}(t), \tag{11}$$

where $E'_{sp,li}$ and $E'_{sp,lq}$ are white noises with two-sided psd $\frac{1}{2}S_{Esp}(f)$; and $I_{sh,li}$ and $I_{sh,lq}$ are white noises with two-sided psd $\frac{1}{2}S_{Ish}(f)$. Since it can be shown that thermal noise is always negligible compared to shot noise and ASE noise [22], we have neglected this term in Eq. (10) and (11). In long-haul systems, the psd of LO-spontaneous beat noise is typically much larger than that of LO shot noise; such systems are thus ASE-limited. Conversely, unamplified systems do not have ASE, and are therefore LO shot-noise-limited.

If one were to demodulate Eq. (8) by an electrical LO at ω_{IF}, as shown in Fig. 5(a), the resulting baseband signals $I_{het,l,i}(t)$ and $I_{het,l,q}(t)$ will be the same as Eqs. (10) and (11) for the homodyne downconverter in Fig. 5(b), with all noises having the same psd's. Hence, the heterodyne and the *two-quadrature* homodyne downconverters have the same performance [23]. A difference between heterodyne and homodyne downconversion only occurs when the transmitted signal occupies one quadrature (e.g. 2-PSK) and the system is LO shot-noise-limited, as this enables the use of a *single-quadrature* homodyne downconverter that has the optical front-end of Fig. 5(a), but has $\omega_s = \omega_{LO}$. Its output photocurrent is $2R\sqrt{P_{LO,l}}\left(\sqrt{P_r}\, y_{lq}(t) + E'_{sp,l}(t)\right) + I_{sh,l}(t)$. Compared to Eq. (11), the signal term is doubled (four times the power), while the shot noise power is only increased by two, thus yielding a sensitivity improvement of 3 dB compared to heterodyne or *two-quadrature* homodyne downconversion. This case is not of practical interest in this paper, however, as long haul systems are ASE-limited, not LO shot-noise-limited. Also, for good spectral and power efficiencies, modulation formats that encode information in both I and Q are preferred. Hence, there is no performance difference between a homodyne and a heterodyne downconverter provided optical filtering is used to reject image-band ASE for the heterodyne downconverter. Since the two downconverters in Fig. 5 ultimately recover the same baseband signals, we can combine Eqs. (10) and (11) and write a normalized, canonical equation for their outputs as:

$$y_l(t) = \sum_k \sum_{m=1}^{2} x_{m,k}\, c_{lm}(t - kT_s)\, e^{j\phi(t)} + n_l(t), \tag{12}$$

where $n_l(t)$ is complex white noise with a two-sided psd of:

$$S_{nn}(f) = N_0 = T_s/\gamma_s\ . \tag{13}$$

γ_s is the signal-to-noise ratio (SNR) per symbol. The values of γ_s for homodyne and heterodyne downconverters in different noise regimes are shown in Table 1. For the shot-noise limited regime using a heterodyne or two-quadrature homodyne downconverter, γ_s is simply the number of detected photons per symbol. We note that Eq. (12) is complex-valued, and its real and imaginary parts are the two baseband signals recovered in Fig. 5. For the remainder of this paper, it is understood that all complex arithmetic operations are ultimately implemented using these real and imaginary signals.

The advantages of heterodyne downconversion are that it uses only one balanced photodetector and has a simpler optical hybrid. However, the photocurrent in Eq. (8) has a bandwidth of $\omega_{IF} + BW$, where BW is the signal bandwidth (Fig. 6(a)). To avoid signal distortion caused by overlapping side lobes, ω_{IF} needs to be sufficiently large. Typically,

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

$\omega_{IF} \approx BW$, thus a heterodyne downconverter requires a balanced photodetector with at least twice the bandwidth of a homodyne downconverter, whose output photocurrents in Eqs. (10) and (11) only have bandwidths of BW (Fig. 6(b)). This extra bandwidth requirement is a major disadvantage. In addition, it is also difficult to obtain electrical mixers with baseband bandwidth as large as the IF. A summary of homodyne and heterodyne receivers is given in Table 2. A comparison of carrier synchronization options is given in Table 3.

Table 1. SNR per symbol for various receiver configurations. For the ASE-limited cases, $\overline{n}_s$ is the average number of photons received per symbol, N_A is the number of fiber spans, and n_{sp} is the spontaneous emission noise factor of the inline amplifiers. For the LO shot-noise-limited cases, $\overline{n}_r = \eta \overline{n}_s$ is the number of *detected* photons per symbol, where η is the quantum efficiency of the photodiodes.

Regime	Homodyne (Single Quadrature)	Homodyne (Two Quadratures)	Heterodyne
ASE-limited	$\frac{\overline{n}_s}{N_A n_{sp}}$	$\frac{\overline{n}_s}{N_A n_{sp}}$	$\frac{\overline{n}_s}{N_A n_{sp}}$
Shot-noise-limited	$\frac{1}{2}\overline{n}_r$	$\overline{n}_r$	$\overline{n}_r$

.

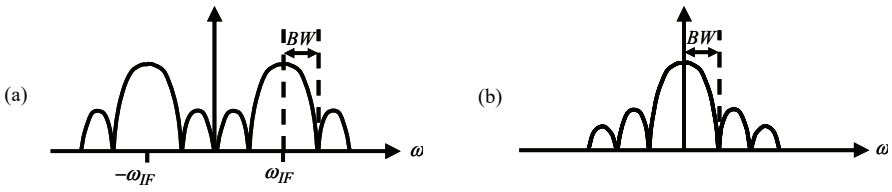

Fig. 6. Spectrum of a (a) heterodyne and (b) homodyne downconverter measured at the output of the balanced photodetector.

Table 2. Comparison between homodyne and heterodyne downconverters.

	Homodyne	Heterodyne
No. of balanced photodetectors per polarization required for QAM	2	1
Minimum photodetector bandwidth	BW	$2BW$

Table 3. Comparison of carrier synchronization options. All three can be used with either homodyne or heterodyne downconversion.

	Optical PLL	Electrical PLL	FF Carrier Recovery
Can the transceiver use same laser for TX and LO?	No	Yes	Yes
Does propagation delay degrade performance?	Yes	Yes	No
Carrier phase estimate depends on past or future symbols?	Past only	Past only	Past and future
Implementation	Analog	Analog or digital	Analog [24] or digital

3.4.2 Dual-polarization downconverter

In the single-polarization downconverter, the LO needed to be in the same polarization as the received signal. One way to align the LO polarization with the received signal is with a polarization controller (PC). There are several drawbacks with this: first, the received polarization can be time-varying due to random birefringence, so polarization tracking is required. Secondly, PMD causes the received Stokes vector to be frequency-dependent. When

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

a single-polarization receiver is used, frequency-selective fading occurs unless PMD is first compensated in the optical domain. Thirdly, a single-polarization receiver precludes the use of polarization multiplexing to double the spectral efficiency.

A dual-polarization downconverter is shown inside the receiver of Fig. 4(a). The LO laser is polarized at 45° relative to the PBS, and the received signal is separately demodulated by each LO component using two single-polarization downconverters in parallel, each of which can be heterodyne or homodyne. The four outputs are the I and Q of the two polarizations, which has the full information of $\mathbf{E}_S(t)$. CD and PMD are linear distortions that can be compensated quasi-exactly by a linear filter.

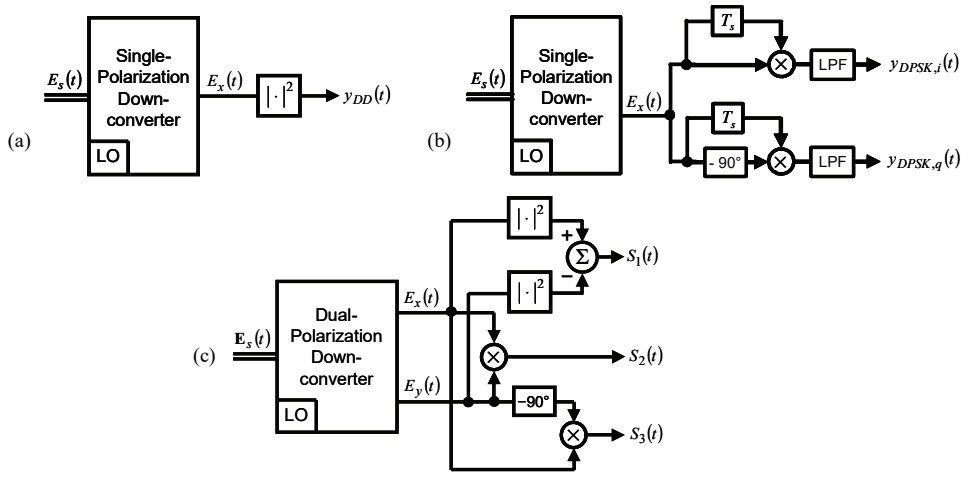

Fig. 7. Emulating (a) direct detection, (b) 4-DPSK detection and (c) PolSK detection with optoelectronic downconversion followed by non-linear signal processing in the electronic domain. The signals $E_x(t)$ and $E_y(t)$ are the complex-valued analog outputs described by Eq. (12) for each polarization. We note that in the case of the heterodyne downconverter, the non-linear operations shown can be performed at the IF output(s) of the balanced photoreceiver.

The lossless transformation from the optical to the electrical domain also allows the receiver to emulate noncoherent and differentially coherent detection by nonlinear signal processing in the electrical domain (Fig. 7). In long-haul transmission where ASE is the dominant noise source, these receivers have the same performance as those in Fig. 1−3. A summary of the detection methods discussed in this section is shown in Table 4.

Table 4. Comparison between noncoherent, differentially coherent and coherent detection. For the first two detection methods, direct detection refers to the receiver implementations shown in Figs. 1−3, while homodyne/heterodyne refers to the equivalent implementations shown in Fig. 7.

	Noncoherent Detection		Differentially Coherent Detection		Coherent Detection
	Direct	Hom./Het.	Direct	Hom./Het.	
Require LO?	No	Yes	No	Yes	Yes
Require Carrier Synchronization?	No	No	No	No	Yes
Can compensate CD and PMD by a linear filter after photodetection?	No	Yes	No	Yes	Yes
Degrees of freedom per polarization per carrier	1		1		2
Modulation formats supported	ASK, FSK, Binary PolSK		DPSK, CPFSK, Non-binary PolSK		PSK, QAM, PolSK, ASK, FSK, etc.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

4. Modulation formats

In this section, we compare the BER performance of different modulation methods for single-carrier transmission corrupted by AWGN. We assume that all channel impairments other than AWGN – including CD, PMD, laser phase noise and nonlinear phase noise – have been compensated using techniques discussed in Section 5. Since ASE and LO shot noise are Gaussian, the performance equations obtained are valid for both long haul and back-to-back systems, when the definition of SNR defined by Table 1 is used. Owing to fiber nonlinearity, it is desirable to use modulation formats that maximize power efficiency. Unless otherwise stated (PolSK being the only exception), the formulae provided assume transmission in one polarization, where noise in the unused polarization has been filtered. This condition is naturally satisfied when a homodyne or heterodyne downconverter is used. For noncoherent detection, differentially coherent detection and hybrid detection, the received optical signal needs to be passed through a polarization controller followed by a linear polarizer.

Since the two polarizations in fiber are orthogonal channels with statistically independent noises, there is no loss in performance by modulating and detecting them separately. The BER formulae provided are thus valid for polarization-multiplexed transmission provided there is no polarization crosstalk. Polarization multiplexing doubles the capacity while maintaining the same receiver sensitivity in SNR per bit. We write the received signal as:

$$y_k = x_k + n_k , \tag{14}$$

where x_k is the transmitted symbol and n_k is AWGN. For the remainder of this paper, our notation shall be as follows:

M is the number of signal points in the constellation.

$b = \log_2(M)$ is the number of bits encoded per symbol.

$T_b = T_s / b$ is the equivalent bit period.

$\gamma_s = E\left[|x_k|^2\right] \Big/ E\left[|n_k|^2\right]$ is the SNR per symbol in single-polarization transmission.

$\gamma_s = E\left[|\mathbf{x}_k|^2\right] \Big/ E\left[|\mathbf{n}_k|^2\right]$ is the SNR per symbol in dual-polarization transmission (e.g. polarization-multiplexed or PolSK).

$\gamma_b = \gamma_s / b$ is the SNR per bit

The maximum achievable spectral efficiency (bit/s/Hz) of a linear AWGN channel is governed by the Shannon capacity [1]:

$$b_{\max} = \log_2(1 + \gamma_s). \tag{15}$$

If N_D identical channels are available for transmission, and we utilize them all by dividing the available power equally amongst the channels, the total capacity is $b = N_D \log_2(1 + \gamma_s / N_D)$, which is an increasing function of N_D. Hence, the best transmission strategy is to use all the dimensions available. For example, suppose a target spectral efficiency of 4 bits per symbol is needed. Polarization-multiplexed 4-QAM, which uses the inphase and quadrature components of both polarizations, has better sensitivity than single-polarization 16-QAM.

ASK with noncoherent detection

Optical M-ary ASK with noncoherent detection has signal points evenly spaced in nonnegative amplitude [25]. The photocurrents for different signal levels thus form a

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

quadratic series. It can be shown that the optimal decision thresholds are approximately the geometric means of the intensities of neighboring symbols. Assuming the use of Gray coding, it can be shown that for large M and γ_b, the BER is approximated by [5]:

$$P_b^{ASK}(M) \approx \frac{1}{b}\left(\frac{M-1}{M}\right) erfc\left(\sqrt{\frac{3b\gamma_b}{2(M-1)(2M-1)}}\right). \tag{16}$$

DPSK with differentially coherent detection

Assuming the use of Gray coding, the BER for M-ary DPSK employing differentially coherent detection is [26]:

$$P_b^{DPSK}(M) \approx \frac{1}{b}\int_{\pi/M}^{\pi}\frac{1}{\pi}\int_{0}^{\pi/2}\sin\chi\left[1+b\gamma_b\left(1+\cos\eta\sin\chi\right)\right]\exp\left(-b\gamma_b\left(1-\cos\eta\sin\chi\right)\right)d\chi\, d\eta\,. \tag{17}$$

For binary DPSK, the above formula is exact, and can be simplified to [27]:

$$P_b^{DPSK}(2) = \frac{1}{2}\exp(-\gamma_b). \tag{18}$$

For quaternary DPSK, we have [27]:

$$P_b^{DPSK}(4) = Q_1(\alpha,\beta) - \frac{1}{2}I_0(\alpha\beta)\exp\left[-\tfrac{1}{2}\left(\alpha^2+\beta^2\right)\right], \tag{19}$$

where $\alpha = \sqrt{2\gamma_b\left(1-\sqrt{1/2}\right)}$ and $\beta = \sqrt{2\gamma_b\left(1+\sqrt{1/2}\right)}$. $Q_1(x,y)$ and $I_0(x)$ are the Marcum Q function and the modified Bessel function of the zeroth order, respectively.

Polarization-Shift Keying (PolSK)

PolSK is the special case in this section where the transmitted signal naturally occupies both polarizations. Thus, polarization multiplexing cannot be employed to double system capacity. The BER for binary PolSK is [16]:

$$P_b^{PolSK}(2) = \frac{1}{2}\exp(-\gamma_b). \tag{20}$$

For higher-order PolSK, the BER is well-approximated by [28],

$$P_b^{PolSK}(M) \approx \frac{1}{b}\left[1 - F_\theta(\theta_1) + \frac{n}{\pi}\int_{\theta_0}^{\theta_1}\cos^{-1}\left(\frac{\tan\theta_0}{\tan t}\right)f_\theta(t)dt\right], \tag{21}$$

where

$$F_\theta(t) = 1 - \frac{1}{2}\exp\left(-b\gamma_b\left(1-\cos t\right)\right)\left(1+\cos t\right), \text{ and} \tag{22}$$

$$f_\theta(t) = \frac{\sin t}{2}\exp\left(-b\gamma_b\left(1-\cos t\right)\right)\left(1+b\gamma_b\left(1+\cos t\right)\right). \tag{23}$$

n, θ_0 and θ_1 are related to the number of nearest neighbors and the shape of the decision boundaries on the Poincaré sphere. Table 5 shows their values for 4-PolSK and 8-PolSK. Square 4-PolSK denotes the constellation where the signal points lie at the vertices of a square

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

enclosed by the Poincaré square. In tetrahedral 4-PolSK and cubic 8-PolSK, the signal points lie at the vertices of a tetrahedron and a cube enclosed by the Poincaré square, respectively.

Table 5. Parameters for computing the BER in polarization-shift keying (PolSK).

	n	θ_0	θ_1
4-PolSK (square)	2	$\pi/4$	$\pi/4$
4-PolSK (tetrahedral)	3	$\frac{1}{2}\left(\pi-\tan^{-1}\sqrt{8}\right)$	$\pi-2\theta_0$
8-PolSK (cubic)	3	$\tan^{-1}\frac{1}{\sqrt{2}}$	$\pi/2-\theta_0$

PSK with coherent detection

Assuming the use of Gray coding, the BER for M-ary PSK employing coherent detection is given approximately by [29]:

$$P_b^{PSK}(M)\approx\frac{1}{b}erfc\left(\sqrt{b\gamma_b}\sin\left(\frac{\pi}{M}\right)\right). \tag{24}$$

For the special cases of BPSK and QPSK, we have the exact expressions:

$$P_b^{PSK}(2)=P_b^{PSK}(4)=\frac{1}{2}erfc\left(\sqrt{\gamma_b}\right). \tag{25}$$

QAM with coherent detection

Assuming the use of Gray coding, the BER for a square M-QAM constellation with coherent detection is approximated by [29]:

$$P_b^{QAM}(M)\approx\frac{2}{b}\left(\frac{\sqrt{M}-1}{\sqrt{M}}\right)erfc\left(\sqrt{\frac{3b\gamma_b}{2(M-1)}}\right). \tag{26}$$

The BER performance of 8-QAM with the signals points arranged as shown in Fig. 8 is [20]:

$$P_b^{QAM}(8)\approx\frac{11}{16}erfc\left(\sqrt{\frac{3\gamma_b}{3+\sqrt{3}}}\right). \tag{27}$$

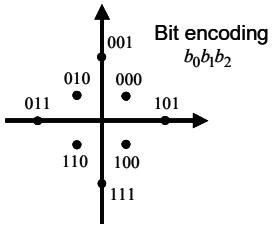

Fig. 8. 8-QAM constellation.

Using Eq. (16) to (27), we compute the SNR per bit required for each modulation format discussed to achieve a target BER of 10^{-3}, which is a typical threshold for receivers employing forward error-correction coding (FEC). The results are shown in Table 6. In Fig. 9, we plot spectral efficiency vs SNR per bit with polarization multiplexing to obtain fair comparison with PolSK (we also show results for 12-PolSK and 20-PolSK [28]). Since polarization-multiplexed ASK, DPSK and PSK all have two DOF (one per polarization), as is the case with PolSK, they all have similar slopes at high spectral efficiency. Because QAM

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

uses all four available DOF for encoding information, it has better SNR efficiency than the other formats, and exhibits a steeper slope at high spectral efficiency.

Table 6. SNR per bit (in dB) required to achieve BER=10^{-3}.

Bits per Symbol	Constellation Size M	Single-Polarization Transmission				PolSK
		ASK with Direct Detection	DPSK with Interferometric Detection	PSK with Coherent Detection	QAM with Coherent Detection	
1	2	9.8	7.9	6.8	6.8	7.9
2	4	15.0	9.9	6.8	6.8	8.0
3	8	20.0	13.1	10.0	9.0	9.4
4	16	25.0	17.4	14.3	10.5	

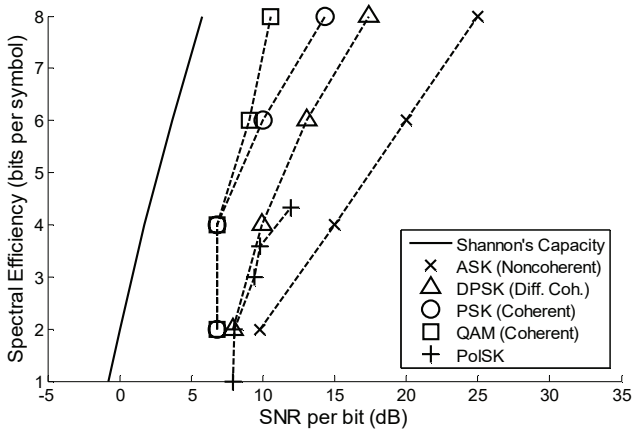

Fig. 9. Spectral efficiency vs. SNR per bit required for different modulation formats at a target BER of 10^{-3}. We assume polarization multiplexing for all schemes except PolSK. Also shown is the Shannon limit (15), which corresponds to zero BER.

5. Channel impairments and compensation techniques in single-carrier systems

In this section, we review the major channel impairments in fiber-optic transmission. We present the traditional methods of combating these, and show how compensation can be done electronically with coherent detection in single-carrier systems. Impairment compensation in multi-carrier systems is discussed in Section 6.

5.1 Linear impairments

5.1.1 Chromatic dispersion

CD is caused by a combination of waveguide and material dispersion [22]. In the frequency domain, CD can be represented by a scalar multiplication:

$$\mathbf{H}_{CD}(\omega) = e^{-j\left(\frac{1}{2}\beta_2 L_{fiber}(\omega-\omega_s)^2+\frac{1}{6}\beta_3 L_{fiber}(\omega-\omega_s)^3\right)}\mathbf{I}, \tag{28}$$

where L_{fiber} is the length of the fiber, β_2 is the dispersion parameter, β_3 is the dispersion slope, and ω_s is the signal carrier frequency. Uncompensated CD leads to pulse broadening, causing intersymbol interference (ISI). Long-haul systems use DCF to compensate CD optically [22]. However, inexact matching between the β_2 and β_3 of transmission fiber and DCF dictates the need for terminal dispersion compensation at high bit rates, typically 40

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

Gbit/s or higher [30]. In reconfigurable networks, data can be routed dynamically through different fibers, so the residual dispersion can be time-varying. This necessitates tunable dispersion compensators.

5.1.2 Polarization-mode dispersion

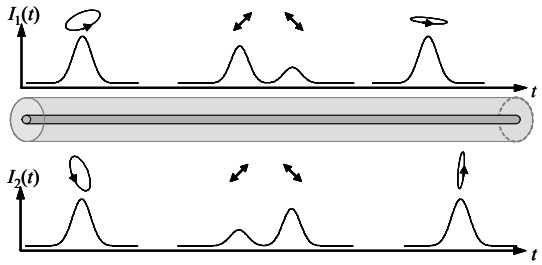

Fig. 10. First-order polarization-mode dispersion.

Polarization-mode dispersion (PMD) is caused by random birefringence in the fiber. In first-order PMD, a fiber possesses a "fast axis" along in polarization and a "slow axis" in the orthogonal polarization (Fig. 10). These states of polarization are known as the *principal states of polarization* (PSPs), and can be any vector in Stokes space in general. First-order PMD can be written as [31]:

$$\mathbf{H}_{PMD}(\omega) = \mathbf{R}_1^{-1}\mathbf{D}\mathbf{R}_2, \tag{29}$$

where $\mathbf{D} = diag\left(e^{j\omega\tau_{DGD}/2}, e^{-j\omega\tau_{DGD}/2}\right)$ is a diagonal matrix with τ_{DGD} being the differential group delay between the two PSPs, and $\mathbf{R}_1$ and $\mathbf{R}_2$ are unitary matrices that rotate the reference polarizations to the fiber's PSPs, which are elliptical in general. When a signal is launched in any polarization state other than a PSP, the receiver will detect two pulses at each reference polarization. Ignoring CD and other effects, the impulse response measured by a polarization-insensitive direct-detection receiver is $h(t) = a^2 \cdot \delta(t - \tau_{DGD}/2) + \left(1 - a^2\right) \cdot \delta(t + \tau_{DGD}/2)$, where a^2 is the proportion of transmitted energy falling in the slow PSP. In this simple two-path model, we see that PMD can lead to frequency-selective fading [32]. In contrast to CD, which is relatively static, PMD (both the PSPs and the DGD) can fluctuate on time scales on the order of a millisecond [33]. Thus PMD compensators thus need to be rapidly adjustable. The statistical properties of PMD have been studied in [34−36], and it has been shown that τ_{DGD} has a Maxwellian distribution, whose mean value $\bar{\tau}_{DGD}$ grows as the square-root of fiber length. In SMF, $\bar{\tau}_{DGD}$ is typically of order $0.1\,\mathrm{ps}/\sqrt{\mathrm{km}}$. PMD is a significant impact on systems at bit rates of 40 Gbit/s and higher, because τ_{DGD} can be a significant fraction of the symbol period. Uncompensated PMD can result in system outage [37].

One method of combating PMD is to use a tunable PC at the transmitter to ensure the input signal is launched into a PSP [38]. Receiver-based compensators for first-order PMD use a continuously tunable PC followed by a variable retarder, which inverts the DGD of the fiber [38,39]. By cascading multiple first-order PMD compensators, one can retrace the PMD vector of the transmission fiber. Such a device can compensate higher-order PMD [40]. Optical PMD compensators have been constructed using nonlinear chirped fiber Bragg gratings [41], planar lightwave circuits (PLC) [42] and polarization-maintaining fibers (PMF) twisted mechanically [40]. Compensation of τ_{DGD} as large as 1.7 symbols was demonstrated by Noé et al [40]. A major limitation of optical PMD compensation is that device performance

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

depends on the degree of tunability, and increasing the number degrees of freedom can require costly hardware. However, optical PMD compensators are transparent to the data rate and modulation format of the transmitted signal, and have been successfully employed for very high-data-rate systems, where digital compensation is currently impossible.

Electronic PMD equalization has gained considerable recent interest. Buchali and Bülow studied the use of a feedforward equalizer (FFE) with decision feedback equalizer (DFE) to combat PMD systems using direct detection of OOK [39]. As with electronic CD compensation in direct detection of OOK [43], the loss of phase during detection prevents quasi-exact compensation of PMD.

5.1.3 Other linear impairments

In addition to CD and PMD, a fiber optic link can also have polarization-dependent loss (PDL) due to anisotropy of network components such as couplers, isolators, filters, multiplexers, and amplifiers [44]. In DWDM transmission, arrayed waveguide gratings (AWG), interleavers and reconfigurable add-drop (de)multiplexers (ROADMs) cause attenuation at the band edges of a channel. When a signal has to pass through cascaded elements, bandwidth narrowing can be problematic. This is a major challenge in 40 Gbit/s RZ-DPSK at 50 GHz channel spacing [45]. Bandwidth narrowing can be equalized by tunable optical equalizers, but such devices are costly, introduce loss, and are difficult to make adaptive.

5.1.4 Compensation of linear impairments and computational complexity

Since a dual-polarization downconverter linearly recovers the full electric field, CD and PMD can be compensated quasi-exactly in the electronic domain after photodetection. One approach is to use a tunable analog filter. However, as in the case of optical compensators, it is difficult to implement the desired transfer function exactly, and analog filters are also difficult to make adaptive. In addition, parasitic effects like signal reflections can lead to signal degradation.

With improvements in DSP technology, digital equalization of CD and PMD is becoming feasible. When the outputs of a dual-polarization downconverter are sampled above the Nyquist rate, the digitized signal contains a full characterization of the received *E*-field, allowing compensation of linear distortions by a linear filter. CD compensation using a digital infinite impulse response (IIR) filter was studied by [46]. Although an IIR filter allows fewer taps, it is more difficult to analyze, and may require greater receiver complexity because of the need to implement time-reversal filters. In this paper, we concentrate on CD and PMD compensation using a finite impulse response (FIR) filter.

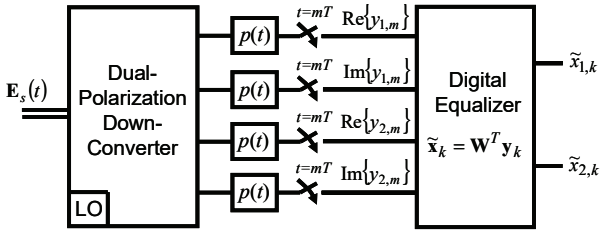

Fig. 11. Digital equalization for a dually polarized linear channel.

Linear equalization using an FIR filter for dually polarized coherent systems was studied in [47,48]. The canonical system model is shown in Fig. 11. The analog outputs of a dual-polarization downconverter are passed through anti-aliasing filters with impulse responses $p(t)$ and then sampled synchronously at a rate of $1/T = M/KT_S$, where M/K is a rational oversampling ratio. We assume that the sampling clock has been synchronized using well-known techniques [49]. In theory, the use of a matched filter in conjunction with symbol rate

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

sampling is optimal. In practice however, symbol-rate sampling is susceptible to sampling time errors. Fractionally spaced sampling has been shown to overcome this [50−52]. Let $q_{ij}(t) = b(t) \otimes h_{ij}(t) \otimes p(t)$ and $n_i'(t) = p(t) \otimes n_i(t)$. We can write the digital samples as:

$$y_i(kT) \overset{\Delta}{=} y_{i,k} = \sum_n \sum_{j=1}^{2} x_{j,n}\, q_{ij}(kT - nT_s) + n_i'(kT). \tag{30}$$

Suppose a linear equalizer takes the $N = 2L+1$ samples closest to symbol k to computes the minimum-mean-square error (MMSE) estimate of the k-th symbol $\tilde{\mathbf{x}}_k$. We have:

$$\tilde{\mathbf{x}}_k = \begin{bmatrix} \tilde{x}_{1,k} \\ \tilde{x}_{2,k} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{11}^T & \mathbf{W}_{21}^T \\ \mathbf{W}_{12}^T & \mathbf{W}_{22}^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_{1,k} \\ \mathbf{y}_{2,k} \end{bmatrix} = \mathbf{W}^T \mathbf{y}_k, \tag{31}$$

where $\mathbf{y}_{i,k} = \left[y_{i,\left\lfloor \frac{kM}{K} \right\rfloor + L}, y_{i,\left\lfloor \frac{kM}{K} \right\rfloor + L - 1}, \cdots, y_{i,\left\lfloor \frac{kM}{K} \right\rfloor - L} \right]^T$, $i = 1, 2$. It can be shown that the MMSE filter is a $2 \times 2N$ matrix given by [19]:

$$\mathbf{W}_{opt} = \mathbf{A}^{-1} \boldsymbol{\alpha}, \tag{32}$$

where $\mathbf{A} = E\left[\mathbf{y}_k^* \mathbf{y}_k^T\right]$ and $\boldsymbol{\alpha} = E\left[\mathbf{y}_k^* \mathbf{x}_k^T\right]$. When an non-integer oversampling ratio ($K > 1$) is used, there are K solutions to (32) depending on the value of $kM \bmod K$. In the example of 3/2 oversampling, there are separate Wiener solutions for odd and even symbols because of the difference in the sampling times relative to the centers of the symbols.

Defining $\boldsymbol{\varepsilon}_k = \mathbf{x}_k - \tilde{\mathbf{x}}_k$ to be the error (in the absence of other channel effects), it can be shown that the the mean-square error (MSE) matrix associated with equalizer $\mathbf{W}$ is a quadratic surface:

$$\mathcal{E} = E\left[\boldsymbol{\varepsilon}_k^* \boldsymbol{\varepsilon}_k^T\right] = \left(\mathbf{W} - \mathbf{W}_{opt}\right)^H \mathbf{A} \left(\mathbf{W} - \mathbf{W}_{opt}\right) + \left(P_x \mathbf{I} - \boldsymbol{\alpha}^H \mathbf{A}^{-1} \boldsymbol{\alpha} \right), \tag{33}$$

and that $tr\{E\}$ is minimized by choosing $\mathbf{W} = \mathbf{W}_{opt}$. Thus for a static channel $\mathbf{H}$, Eq. (32) gives the optimum equalizer of length $N = 2L+1$. It can be shown that as $M/K \to \infty$ and $N \to \infty$, and the anti-aliasing filter does not introduce amplitude distortion, the power penalty of a compensated CD/PMD channel approaches zero with respect to a pure AWGN channel at the same SNR. In practice, since $\mathbf{H}$ can be time-varying, an adaptive equalizer shown in Fig. 12 is desired. Owing to the quadratic nature of Eq. (33), we can use well-known algorithms such as least mean square (LMS) or recursive least squares (RLS) [53]. For LMS, we have the following update equation for the filter coefficients:

$$\mathbf{W}^{(m+1)} = \mathbf{W}^{(m)} + 2\mu \mathbf{y}_k^* \boldsymbol{\varepsilon}_k^T, \tag{34}$$

where $\mathbf{W}^{(m)}$ is the equalizer coefficients use to compute $\tilde{\mathbf{x}}_k$, and μ is the step size parameter that needs to satisfy $0 < \mu < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{A} = E\left[\mathbf{y}_k^* \mathbf{y}_k^T\right]$. When a non-integer oversampling ratio is used, K filters are required for all possible values of $kM \bmod K$. The index m indicates the m-th use of that particular filter.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

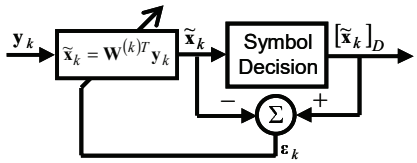

Fig. 12. Adaptive equalizer for polarization-multiplexed coherent detection.

Table 7. Equalizer length required (*N*) to compensate CD in a system using polarization-multiplexed 4-QAM at 100 Gbit/s. SMF (*D* = 17 ps/nm-km) with 2% under-compensation of CD is assumed. The oversampling ratio is 3/2.

Transmission Distance (km)	Polarization-Multiplexed Transmission		
	4-QAM (4 bits/symbol)	8-QAM (6 bits/symbol)	16-QAM (8 bits/symbol)
1,000	3	2	1
2,000	6	3	2
3,000	8	4	2
5,000	13	6	4

Ip and Kahn showed that when $p(t)$ is a fifth-order Butterworth filter with a 3-dB bandwidth of $0.4(M/K)R_s$, any arbitrary amount of CD and first-order PMD can be compensated with less than 2 dB power penalty provided the oversampling rate $M/K \geq 3/2$, and the filter length *N* satisfies:

$$N = 2\pi|\beta_2|LR_s^2(M/K), \text{ and} \tag{35}$$

$$N = \tau_{DGD}M/KT_s, \tag{36}$$

for mitigating CD and PMD, respectively [19]. It was found that for $M/K \geq 3/2$, the required value of *N* is insensitive to whether NRZ or RZ pulses are used, that system performance is insensitive to sampling time errors. Typically, the required value of *N* is dominated by CD considerations. In Table 7, we show the required value of *N* for different transmission distances, where inline DCF is used with 2% CD under-compensation. The target bit rate is 100 Gbit/s, and we consider polarization-multiplexed 4-QAM transmission. The complexity of directly implementing Eq. (32) is $4NR_s$ complex multiplications per second. For large *N*, linear equalization is more efficiently implemented using an FFT-based block processing [54]. Suppose a block length of *B* is chosen. An FFT-based implementation has a complexity of $R_s(N+B-1)(2\log_2(N+B-1)+4)/B$ complex multiplications per second. For a given *N*, there exists an optimum block length *B* that minimizes the number of operations required. It can be shown that the asymptotic complexity grows as $R_s \log_2 R_s$. We compare the complexity of single-carrier versus multi-carrier transmission (using OFDM) in Section 7.

5.2 Nonlinear impairments

5.2.1 Fiber nonlinearity

The dominant nonlinear impairments in fiber arise from the Kerr nonlinearity, which causes a refractive index change proportional to signal intensity. Signal propagation in the presence of fiber attenuation, CD and Kerr nonlinearity is described by the nonlinear Schrödinger equation (NLSE)

$$\frac{\partial E}{\partial z} + \frac{j\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + \frac{\alpha}{2}E = j\gamma|E|^2 E, \tag{37}$$

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

where $E(z,t)$ is the electric field, α is the attenuation coefficient, β_2 is the dispersion parameter, γ is the nonlinear coefficient, and z and t are the propagation direction and time, respectively.

Nonlinear effects include deterministic and statistical components. The nonlinearity experienced by a signal due to its own intensity called self-phase modulation (SPM). In WDM systems, a signal also suffers nonlinear effects due to the fields of neighboring channels. These are cross-phase modulation (XPM) and four-wave-mixing (FWM) [22], and their impact can be reduced by allowing non-zero local dispersion. In the absence of ASE, and given knowledge of the transmitted data, all these nonlinear effects are deterministic, and it is possible, in principle, to pre-compensate them at the transmitter. At the receiver, it would be also be possible to employ joint multi-channel detection techniques, though complexity precludes their implementation at this time. Here, we consider only receiver-based compensation of SPM-induced impairments in the presence of CD, which leads to intra-channel nonlinear effects.

In long-haul systems, interaction between ASE noise and signal through the Kerr nonlinearity leads to nonlinear phase noise (NLPN). When caused by the ASE and signal in the channel of interest, this is called SPM-induced NLPN. When caused by the ASE and signal in neighboring channels, it is called XPM-induced NLPN. Here, we consider only receiver-based compensation of SPM-induced NLPN.

In following two sections, we discuss (i) SPM in the presence of CD, and (ii) SPM-induced NLPN. We show how these can be compensated by exploiting their correlation properties.

5.2.2 Self-phase modulation with chromatic dispersion: intra-channel nonlinearity

We consider a noiseless transmitted signal $E(0,t) = \sum_k x_k b(t - kT_s) = \sum_k x_k b_k$, where T_s is the symbol period, b_k is the sampled pulse shape, and $x_k \in \left\{e^{j2\pi/M}, e^{j4\pi/M}, \cdots, e^{j2\pi}\right\}$ are the transmitted symbols chosen from an M-ary PSK constellation. We use first-order perturbation theory to gain insight into the effects of nonlinearity [55]. Let $E = E^{(lin)} + \Delta E = \sum_k x_k \left(b_k^{(lin)} + \Delta b_k\right)$, where $E^{(lin)}$ is the linear solution to the NLSE (obtained by setting the right hand side of (37) to zero) and ΔE is the perturbation due to Kerr nonlinearity. The NLSE can be re-written as:

$$\frac{\partial \Delta E}{\partial z} + j\beta_2 \frac{\partial^2 \Delta E}{\partial t^2} + \frac{\alpha}{2}\Delta E = j\gamma |E^{(lin)}|^2 E^{(lin)} = j\gamma \sum_{l,m,p} x_l x_m x_p^* b_l^{(lin)} b_m^{(lin)} b_p^{(lin)*} . \quad (38)$$

For typical terrestrial links, the accumulated dispersion is such that the only the terms on the right hand side of (38) that will induce a sizeable effect on the pulse at symbol k are those with indices satisfying $k = l + m - p$. Without loss of generality, we focus on symbol $k = 0$. The NLSE can then be simplified to:

$$\frac{\partial \Delta b_0}{\partial z} + j\beta_2 \frac{\partial^2 \Delta b_0}{\partial t^2} + \frac{\alpha}{2}\Delta b_0 = j\gamma \sum_{l,m} x_l x_m x_{l+m}^* b_l^{(lin)} b_m^{(lin)} b_{l+m}^{(lin)*} . \quad (39)$$

The term $l = m = 0$ is a deterministic distortion of a pulse by itself, and is referred to as *intra-channel self-phase modulation* (ISPM). The terms where $l = 0, m \neq l$ (and $m = 0, l \neq m$) are distortions to a pulse by neighboring pulses, and are known as *intra-channel cross-phase modulation* (IXPM), as they are analogous to signal distortion by neighboring channels in XPM. Finally, the remaining terms $l, m \neq 0$ are called *intra-channel four-wave mixing*

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

(IFWM), because they are analogous to the interacting frequencies in FWM. We emphasize that the "intra-channel" effects all originate from SPM.

It is well-known that in OOK with direct detection, IXPM causes timing jitter to the pulses, while IFWM causes amplitude jitter (or "ghost pulses") in the zero bits [56,57]. One can minimize these effects by careful dispersion map design [58,59], phase alternations [60] and coding [61]. When coherent detection is used in conjunction with constant-intensity modulation formats (e.g., PSK or DPSK), IXPM is deterministic, since $|x_l|^2$ is constant. Hence the randomness in $b_0^{(lin)}$ is only due to IFWM. Let $C'_{l,m}(t)$ be the solution to

$$\frac{\partial C'_{l,m}}{\partial z} + j\beta_2 \frac{\partial^2 C'_{l,m}}{\partial t^2} + \frac{\alpha}{2} C'_{l,m} = j\gamma b_l^{(lin)} b_m^{(lin)} b_{l+m}^{(lin)*}, \tag{40}$$

which represents signal propagation through one span of fiber. The IFWM-induced phase noise at symbol 0 is given by:

$$\phi_{IFWM}(t) = \mathrm{Im}\left\{ \sum_{l,m\neq 0} \frac{x_l x_m x_{l+m}^* C'_{l,m}(t)}{x_0 (b_0^{(lin)} + \Delta b^{(S+X)})} \right\} = \mathrm{Im}\left\{ \sum_{l,m\neq 0} x_l x_m x_{l+m}^* x_0^* C_{l,m}(t) \right\}, \tag{41}$$

where $\Delta b^{(S+X)}$ is the deterministic perturbation to $b_0^{(lin)}$ due to ISPM and IXPM. IFWM phase noise was first studied by Wei and Liu [62], who showed that ϕ_{IFWM} are correlated between symbols. In long-haul transmission over multiple identical spans of fiber, ϕ_{IFWM} add coherently for each span. Lau and Kahn have shown the autocorrelation function of IFWM phase noise: $R(k,t) = E[\phi_{IFWM}(t)\phi_{IFWM}(t-kT)]$ is given by [63]:

$$\begin{aligned} R(0) &= \frac{1}{2}\sum_{l,m} \mathrm{Im}\{C_{l,m}\}^2, \\ R(k) &= \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k-m} C_{-m,m-k}^*\right\} - \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k} C_{m,-k}\right\} \\ &\quad + \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k} C_{m,-k}^*\right\} - \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k-m} C_{-m,m-k}\right\}, \; k>0 \end{aligned} \tag{42}$$

for BPSK systems, and

$$R(0) = \frac{1}{2}\sum_{l,m} |C_{l,m}|^2, \quad R(k) = \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k-m} C_{-m,m-k}^*\right\} - \frac{1}{2}\sum_m \mathrm{Re}\left\{C_{m,k} C_{m,-k}\right\}, \; k>0 \tag{43}$$

for *M*-ary PSK with $M > 2$. Fig. 13 shows $R(k,t)$ for single-polarization RZ-QPSK at 80 Gbit/s, where we used Gaussian pulses with 33% duty cycle. We compare Eq. (43) with Monte Carlo simulations at different sampling times *t*. The simulations used 5000-trial propagations through a typical terrestrial system with a random 32-pulse sequence. We observe good agreement between theory and simulation results. It is thus possible to reduce effects of IFWM by implementing a linear noise predictor in DSP using knowledge of $R(k)$. Assuming the received symbols are sampled at the optimal instants, we have $\theta_k = \arg\{x_k\} + \phi_{IFWM}(kT)$. A 1.8-dB improvement in performance was obtained using a linear predictor for IFWM when IFWM is the dominant system impairment.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

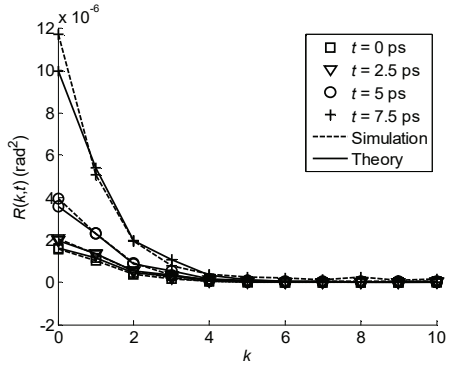

Fig. 13. Autocorrelation function for IFWM phase noise in 80 Gbit/s QPSK transmission with 33% Gaussian pulses. Each span consists of 80 km of SMF with α = 0.25 dB/km, D = 17 ps/nm-km, γ = 1.2 $W^{-1}km^{-1}$, followed by DCF with α = 0.6 dB/km, D = −85 ps/nm-km, γ = 5.3 $W^{-1}km^{-1}$. The mean nonlinear phase shift is Φ_{NL} = 0.0215 rad/span.

5.2.3 Nonlinear phase noise

SPM-induced NLPN is often called the Gordon-Mollenauer (G-M) effect [64]. Fig. 14(b) illustrates the G-M effect for QPSK. The received constellation is spiral-shaped, as signal points with larger amplitude experiences larger phase shifts.

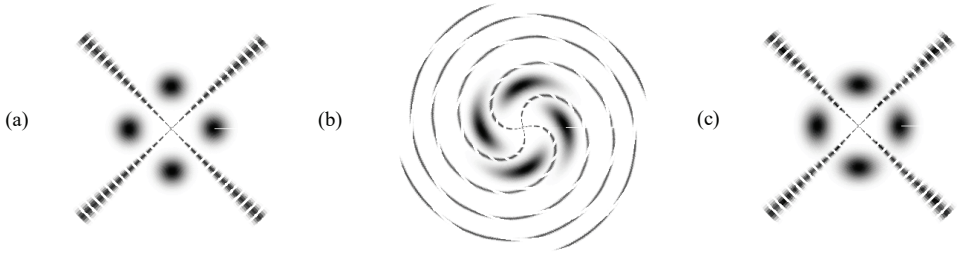

Fig. 14. Constellation diagrams of the received signal showing ML decision boundaries for (a) no NLPN, (b) with NLPN before compensation by θ_{opt}, and (c) with NLPN after compensation by θ_{opt}.

Let the ASE noise of the i-th amplifier be $n_i \sim N\left(0,\sigma^2 I\right)$. We assume the n_i are i.i.d. with $\sigma^2 \propto G-1$, where G is the amplifier gain. In the absence of CD and multi-channel nonlinear effects, the NLPN of a system with N uniformly spaced identical amplifiers is:

$$\phi_{NL} = \gamma L_{eff} \sum_{i=1}^{N} \left| E + \sum_{k=1}^{i} n_k \right|^2 . \tag{44}$$

It can be shown that the variance of Eq. (44) is:

$$\sigma_{NL}^2 = \frac{2}{3} N(N+1)\left(\gamma L_{eff}\sigma\right)^2 \left[(2N+1)|E|^2 + (N^2+N+1)\sigma^2 \right], \tag{45}$$

where L_{eff} is the effective nonlinear length of each span [65] . Ho studied the probability density function (pdf) of NLPN for distributed amplification and obtained analytical formulae

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

for the BER of PSK and DPSK systems [66]. Although ϕ_{NL} is not Gaussian, σ_{NL}^2 is a good measure of the impact of NLPN on system performance. Lau and Kahn studied joint minimization of the NLPN and linear phase noise variances by optimizing the gains and spacings of amplifiers in long-haul transmission [67,68]. In addition, as ϕ_{NL} is correlated with instantaneous received power P_{rec}, one can perform receiver-based compensation of NLPN by applying a phase rotation proportional to P_{rec}. Ho and Kahn [69], Ly-Gagnon and Kikuchi [70] and Liu et al [71] have shown that the optimal phase rotation θ_{opt}, which minimizes $\sigma_{NL}^2 = E\left[(\phi_{NL} - \theta_{opt})^2\right]$, is given by:

$$\theta_{opt} = -\gamma L_{eff} \frac{N+1}{2} P_{rec}. \tag{46}$$

Compared to the uncompensated case, σ_{NL}^2 is reduced by a factor of four (6 dB). This phase rotation can be performed by a phase modulator, or digitally by DSP. Various experiments have confirmed the performance improvement of this technique [72−74]. Ho [75] also studied mid-span phase rotation proportional to instantaneous signal power, and showed that a reduction of σ_{NL}^2 by a factor of nine (9.5 dB) can be obtained when the compensation is performed at 2/3 the length of the transmission link. Recently, Lau and Kahn showed that the ML detection boundaries for *M*-ary PSK in the presence of NLPN are of the form $\theta_{ML} = aP_{rec} + b\sqrt{P_{rec}} + c$ [76]. Hence, ML detection can be implemented by rotating the received signal by θ_{ML} and then using straight decision boundaries (Fig. 14(c)). The phase rotation techniques discussed can also be applied to 16-QAM where it has been shown that ML detection is well-approximated by straight-line decision boundaries when phase pre-compensation and/or post-compensation is implemented at the transmitter and/or receiver [76].

When CD is also considered, NLPN becomes more complicated. For signal propagation over a single span of fiber with perfect dispersion compensation and at high OSNR, the nonlinear phase noise is given by

$$\phi_{NL}(t) = \mathrm{Im}\left[\frac{2j\gamma\int_0^L \sum_k \left|b_k^{(lin)}(z,t)\right|^2 n(z,t) \otimes h_{-z}(t) e^{-\alpha z} dz}{E^{(lin)}(L,t)}\right], \tag{47}$$

where $h_{-z}(t)$ is the impulse response due to the CD of the fiber from *z* to *L*. The variance of Eq. (47) was studied by Green [77], Kumar [78] and Ho [79], and it was shown that the temporal profile of $\sigma_{\phi_{NL}}^2(t)$ is asymmetric with respect to $t = 0$ due to the dispersive effect of ASE noise. Kumar et al [80] and Boivin et al [81] proposed the use of optical phase conjugation to mitigate $\sigma_{\phi_{NL}}^2(t)$, while Serena et al [82] presented a method of characterizing BER analytically in the presence of NLPN based on a parametric gain approach. Further statistical properties of NLPN in the presence of CD, including its pdf and psd, have yet to be investigated.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

5.2.4 Comparison of IFWM phase noise and nonlinear phase noise

The relative impact of IFWM phase noise and NLPN depend on the system parameters. Eqs. (41) and (44) show that the statistical components of IFWM and NLPN varies as $\phi_{IFWM} \propto |E|^2$ and $\phi_{NL} \propto |E \cdot n|$. Hence, the ratio $\sigma^2_{\phi_{IFWM}} / \sigma^2_{\phi_{NL}} \propto P/\sigma^2$ is a function of OSNR. For a system with N amplifiers, $\sigma^2_{\phi_{NL}} \propto N^3$ [64], while $\sigma^2_{\phi_{IFWM}} \propto N^2$, as ϕ_{IFWM} adds coherently between spans. Ho and Wang [83] and Zhang et al [84] investigated the relative impact of ϕ_{IFWM} and ϕ_{NL} for DPSK, and showed that ϕ_{IFWM} increases with the amount of local CD as ϕ_{IFWM} requires strong pulse overlap to occur, while ϕ_{NL} increases with decreasing CD. Zhang et al studied the effects of dispersion pre-compensation and having residual dispersion per span for 40 Gbit/s RZ-DPSK and showed that when ϕ_{IFWM} dominates, the optimal dispersion pre-compensation is similar to that for RZ-OOK [85]. Finally, Zhu et al [86] studied the BER performance of DPSK in the presence of NLPN, IFWM phase noise and linear phase noise through semi-analytical characterizations of the joint phase noise variance.

5.3 Laser phase noise

Laser phase noise is caused by spontaneous emission [87], and is modeled as a Wiener process [88]:

$$\phi(t) = \int_{-\infty}^{t} \delta\omega(\tau) d\tau , \tag{48}$$

where $\phi(t)$ is the instantaneous phase, $\delta\omega(t)$ is frequency noise with zero mean and autocorrelation $R_{\delta\omega\delta\omega}(\tau) = 2\pi\Delta\nu\,\delta(\tau)$. It can be shown that the laser output $E_0(t) = Ae^{j(\omega_c t + \phi(t))}$ has a Lorentzian spectrum with a 3-dB linewidth Δv. Schawlow and Townes showed that laser linewidth is inversely proportional to output power [89], so it is desirable to operate the TX and LO lasers at maximum power, attenuating their outputs as required.

Phase noise is an important impairment in coherent systems as it impacts carrier synchronization. In noncoherent detection, the carrier phase is unimportant because the receiver only measures energy. In DPSK, information is encoded by phase changes, and Δv only needs to be small enough such that the phase fluctuation over a symbol period is small.

In Eq. (12), we showed that the baseband signal $y(t)$ is modulated by $e^{j\phi(t)}$. In the absence of other impairments, this manifests as a rotation of the received constellation. Carrier synchronization is required to ensure $\phi(t)$ is small so the transmitted symbols can be detected with low power penalty. Since phase noise is a Wiener process with temporal correlation, it can be mitigated by signal processing. In the next two subsections, we consider carrier synchronization in a single polarization using a PLL and FF carrier synchronizer.

5.3.1 Phase-locked loop

The traditional method of carrier synchronization is to use a PLL. A system diagram of a PLL is shown in Fig. 15(a). The phase estimator removes the data modulation so that $\phi(t)$ can be measured. This can be done by a number of methods, including raising the signal to the M-th power in M-ary PSK [23]. The phase estimator output is an error signal that is then passed through a loop filter, producing a control signal for the LO frequency. In an OPLL, the control signal drives the LO laser, whereas in an electrical PLL, the control signal drives the electrical

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

VCO (Fig. 5(a)). Both types of PLL can be analyzed in the same manner, and the design of PLLs has been extensively studied in [29,90].

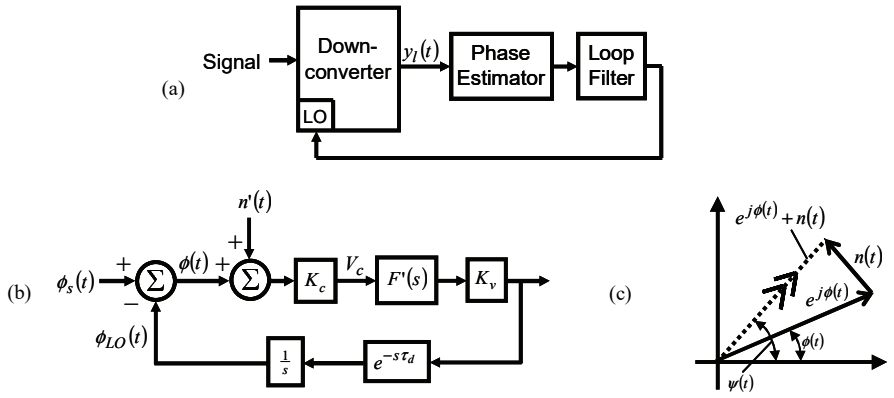

Fig. 15. Phase-locked loop: (a) System model, (b) analytical model, (c) phase estimation.

The performance of a PLL is usually analyzed with the linear model shown in Fig. 15(b). We assume the LO has no phase noise, while the TX laser has phase noise equal to the sum of the linewidths of the two lasers. The signal to be tracked is $\phi_s(t)$, which evolves as Eq. (48). Owing to AWGN, the phase estimator measures $\psi(t)=\phi(t)+n'(t)$ (Fig. 15(c)), and produces a voltage $K_c\psi(t)$ for the loop filter $F'(s)$. We assume the control port of the LO has a slope K_v Hz/V. Delay in the loop is modeled as $e^{-s\tau_d}$, where τ_d includes signal propagation plus the rise times of intermediate components. Delay degrades system performance. An electrical PLL is usually superior to an OPLL since according to Fig. 5(a), the optical path has to pass through the optical hybrid and balanced photodetectors. However, an LO laser is likely to have a larger tuning range than an electrical LO. If the frequency drift of the lasers is significant, an OPLL may be preferable.

The performance of the PLL is determined by the loop filter. For a given $F'(s)$, we have:

$$\phi(s)=\frac{s}{s+F(s)e^{-s\tau_d}}\phi_s(s)-\frac{F(s)e^{-s\tau_d}}{s+F(s)e^{-s\tau_d}}n'(s), \tag{49}$$

where $F(s)=K_cK_vF'(s)$. Ignoring delay, the denominator of Eq. (49) gives the loop order. For a first-order loop, $F'(s)=K_f$. The design parameter is the loop bandwidth $\omega_n=K_cK_fK_v$. For a second-order loop, $F(s)=2\zeta\omega_n+\omega_n^2/s$, where ζ and ω_n are the damping factor and natural frequency, respectively. The performance 4-QAM employing a second-order PLL was studied in [91], while the performance of 8- and 16-QAM employing a second-order PLL was studied in [20]. A damping factor of $1/\sqrt{2}$ is typically chosen as a compromise between a rapid response and low steady-state variance. For both first- and second-order PLLs, there is an optimal ω_n that minimizes the phase-error variance:

$$\sigma_\phi^2=\frac{\sigma_p^2}{2\zeta\omega_nT_s}\Gamma_{PN}(\omega_n\tau_d)+\frac{\left(1+4\zeta^2\right)\omega_nT_s}{4\zeta}\frac{\eta_c}{2\gamma_s}\Gamma_{AWGN}(\omega_n\tau_d), \tag{50}$$

where

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

$$\Gamma_{PN}(\omega_n\tau_d)=\frac{2\zeta\omega_n}{\pi}\int_{-\infty}^{\infty}\left|j\omega+e^{-j\omega\tau_d}F(\omega)\right|^{-2}d\omega\text{, and} \tag{51}$$

$$\Gamma_{AWGN}(\omega_n\tau_d)=\frac{2\zeta}{\pi\left(1+4\zeta^2\right)\omega_n}\int_{-\infty}^{\infty}\left|\frac{F(\omega)}{j\omega+e^{-j\omega\tau_d}F(\omega)}\right|^2 d\omega. \tag{52}$$

The first term in Eq. (50) arises from phase noise and is proportional to $\sigma_p^2=2\pi\Delta\nu T_s$, which is the ratio between the laser linewidth and signal bandwidth. The second term arises from AWGN, and is inversely proportional to the received OSNR. Assuming the use of a decision-directed PLL [23], $\eta_c=E\left[|x|^2\right]E\left[1/|x|^2\right]$ is a penalty factor associated with the transmitted constellation [20]. A larger ω_n allows the LO to adapt more quickly to phase fluctuations in the TX laser, but the loop becomes more susceptible to noise. These two conflicting requirements give rise to an optimum ω_n, which needs to evaluated numerically. A typical plot of σ_ϕ versus ω_n is shown in Fig. 16 for a second-order PLL. We have assumed single-polarization 16-QAM at 100 Gbit/s. The laser beat linewidth and received OSNR are 100 kHz and 11.5 dB, respectively, which is 1 dB above sensitivity for BER = 10^{-3} (Table 6). We observe that delay increases σ_ϕ. Above $\tau_d=125T_b$, there is no ω_n that can give 1-dB power penalty due to phase error ($\sigma_\phi=2.7^\circ$ [20]). At 100 Gbit/s, this maximum delay corresponds to 1.25 ns, which corresponds to a transmission line only ~ 25 cm long. Even with careful circuit design, this is probably not feasible. Hence for high-data-rate systems, FF carrier recovery techniques are likely to be required.

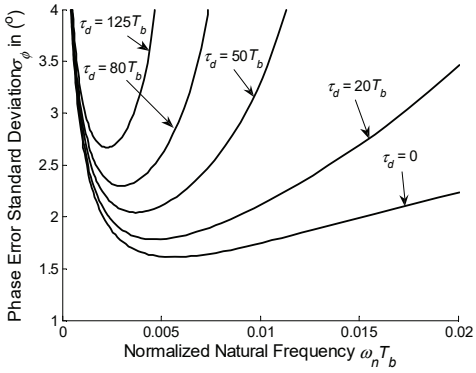

Fig. 16. σ_ϕ versus ω_n for 16-QAM at 100 Gbit/s in a single polarization employing a second-order PLL, with Δν = 100 kHz and an OSNR of 11.5 dB.

5.3.2 Feedforward carrier recovery

The PLL is a feedback system as the control phase at time *t* can only depend on past input phases up to $t-\tau_d$. However, the laser phase noise process described by (48) has a symmetric autocorrelation function $E[\phi(t)\phi(t-\tau)]=2\pi\Delta\nu|\tau|$, so its value at time *t* has the same correlation with phases before and after *t*. The PLL is suboptimal as it does not exploit possible knowledge of $\{\phi(t-\tau):\tau<\tau_d\}$.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

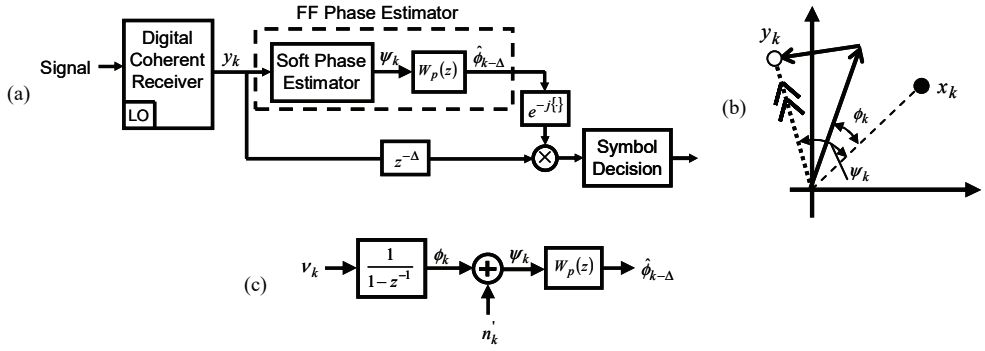

Fig. 17. Feedforward carrier phase estimation. (a) System model, (b) soft phase estimation, (c) analytical model.

FF carrier synchronization for an intradyne receiver using analog electronics was described in [24], while digital FF carrier synchronization was studied in [21,92−94]. In this section, we focus on digital FF carrier synchronization. Consider the system model shown in Fig. 17(a), where we assume all other channel impairments have been compensated by the digital coherent receiver, whose outputs are $y_k = x_k e^{j\phi_k} + n_k$, where x_k is the transmitted symbol, and ϕ_k and n_k are the carrier phase and AWGN, respectively (Fig. 17(b)). Instead of using a PLL to ensure that ϕ_k is small, a FF phase estimator directly estimates the carrier phase and then de-rotates the received signal by this estimate so symbol decisions can be made at low BER.

The FF phase estimator has a *soft phase estimator* that first computes a symbol-by-symbol estimate ψ_k of ϕ_k, followed by a MMSE filter $W_p(z)$. A number of algorithms exist for finding ψ_k. In the case of *M*-ary PSK, raising y_k to the *M*-th power which removes the data modulation. For more general signaling formats, decision-directed (DD) phase estimation can be used [18]. The symbol-by-symbol estimate is corrupted by AWGN (Fig. 17(b)) so that $\psi_k = \phi_k + n_k'$, where n_k' is the projection of n_k onto a vector orthogonal to $x_k e^{j\phi_k}$. Since ϕ_k is correlated by the Wiener process, we use a linear filter to compute an MMSE estimate $\hat{\phi}_{k-\Delta}$. Using the analytical model shown in Fig. 17(c), whose input is the discrete frequency noise process v_k with zero mean and variance $\sigma_p^2 = 2\pi\Delta\nu T_s$, it can be shown that the MMSE filter for $\Delta = 0$ has coefficients:

$$w_n = \begin{cases} \dfrac{\alpha r}{1-\alpha^2}\alpha^n & n \geq 0 \\ \dfrac{\alpha r}{1-\alpha^2}\alpha^{-n} & n < 0 \end{cases}, \tag{53}$$

where $r = \sigma_p^2 / \sigma_{n'}^2 > 0$ is the ratio between the magnitudes of frequency noise and AWGN, and $\alpha = (1 + r/2) - \sqrt{(1 + r/2)^2 - 1}$. The MMSE filter is non-causal, as it has two exponentially decaying tails toward the past and future. In practice, one can truncate Eq. (53) to L_p significant coefficients and implement it as an FIR filter with delay $\Delta = \left\lfloor \frac{L_p - 1}{2} \right\rfloor$. For any

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008

causal filter $W_p(z)$ with L_p coefficients and delay Δ, it can be shown that the error $\phi_k - \hat{\phi}_k$ is Gaussian distributed with zero mean and a variance of:

$$\sigma_\phi^2(W,\Delta) = \sigma_p^2 \cdot \left[\sum_{m'=0}^{\Delta-1} \left(\sum_{l'=0}^{m'} w_{l'} \right)^2 + \sum_{m'=\Delta+1}^{L_p-1} \left(\sum_{l'=m'}^{L_p-1} w_{l'} \right)^2 \right] + \frac{\eta_c}{2\gamma_s} \cdot \left[\sum_{l=0}^{L_p-1} w_l^2 \right]. \tag{54}$$

As with the PLL, the variance has two terms that arise from phase noise and AWGN.

5.3.3 Power Penalty from Phase Error

Regardless of whether a PLL or a FF carrier synchronizer is used, the coherent receiver makes symbol decisions on $y_k = x_k e^{j\varepsilon_k} + n_k$. For a PLL, $\varepsilon_k = \phi_k$ has a Tikhonov distribution [95], while for a FF carrier synchronizer, $\varepsilon_k = \phi_k - \hat{\phi}_k$ has Gaussian distribution. In the limit of high SNR, the Tikhonov is well-approximated by the Gaussian distribution. The method for computing BER for a given phase-error distribution can be found in [95]. With the phase error variances for the PLL and the FF carrier synchronizer given by Eqs. (50) and (54), the power penalty can be determined. In Table 8, we compare the linewidth requirements for receivers that use a PLL and a FF carrier synchronizer, assuming a 1-dB power penalty at a target BER of 10^{-3} [18]. We observe that FF carrier recovery can tolerate 50% to 100% wider laser linewidth than a PLL, and is also insensitive to propagation delay.

Table 8. Linewidth requirements for various single-polarization modulation formats using a PLL and a FF carrier synchronizer at a target BER of 10^{-3}.

Modulation Format	OSNR per bit (dB)	Max. Tolerable σ_ϕ for BER = 10^{-3}	Max. Linewidth using a PLL ($\Delta\nu T_b$)	Max. Linewidth using Feedforward ($\Delta\nu T_b$)
4-QAM	7.79	4.91°	6.9×10^{-5}	1.3×10^{-4}
8-QAM	10.03	5.01°	8.3×10^{-5}	1.3×10^{-4}
16-QAM	11.52	2.70°	7.9×10^{-6}	1.5×10^{-5}

5.3.4 Carrier synchronization in polarization-multiplexed systems

When polarization multiplexing is employed, the receiver has two independent signals from which to estimate carrier phase. Consider the system models shown in Fig. 18. In the PLL-based receiver, the baseband signals for each polarization are passed through phase error estimators that compute independent estimates $\{\psi_i(t), i=1,2\}$ of $\phi(t)$. In the digital FF carrier synchronizer, the soft-decision phase estimators give independent estimates $\{\psi_{i,k}, i=1,2\}$ of ϕ_k. Assuming both phase estimates are equally reliable, we take their average. Since AWGN is the only impairment preventing errorless measurement of carrier phase, and since noises in the two polarizations are independent, $\psi(t)$ and ψ_k have half as much AWGN as $\psi_i(t)$ and $\psi_{i,k}$. For a given noise psd, the AWGN contribution to Eqs. (50) and (54) is halved. However, to obtain the same symbol-error rate for each polarization, the SNR per polarization must be preserved. This requires the transmit energy per symbol (now a 2×1 complex vector) be doubled. Hence, the dependence of phase error variance on *SNR per symbol* is preserved, and Eqs. (50) and (54) hold for polarization-multiplexed transmission.

#86543 - $15.00 USD Received 20 Aug 2007; revised 9 Nov 2007; accepted 12 Nov 2007; published 9 Jan 2008