Spectral Efficiency Limits and Modulation/Detection Techniques for DWDM Systems

Joseph M. Kahn, *Fellow, IEEE,* and Keang-Po Ho, *Senior Member, IEEE*

Invited Paper

***Abstract*—Information-theoretic limits to spectral efficiency in dense wavelength-division-multiplexed (DWDM) transmission systems are reviewed, considering various modulation techniques (unconstrained, constant-intensity, binary), detection techniques (coherent, direct), and propagation regimes (linear, nonlinear). Spontaneous emission from inline optical amplifiers is assumed to be the dominant noise source in all cases. Coherent detection allows use of two degrees of freedom per polarization, and its spectral efficiency limits are several b/s/Hz in typical terrestrial systems, even considering nonlinear effects. Using either constant-intensity modulation or direct detection, only one degree of freedom per polarization can be used, significantly reducing spectral efficiency. Using binary modulation, regardless of detection technique, spectral efficiency cannot exceed 1 b/s/Hz per polarization. When the number of signal and/or noise photons is small, the particle nature of photons must be considered. The quantum-limited spectral efficiency for coherent detection is slightly smaller than the classical capacity, but that for direct detection is 0.3 b/s/Hz higher than its classical counterpart. Various binary and nonbinary modulation techniques, in conjunction with appropriate detection techniques, are compared in terms of their spectral efficiencies and signal-to-noise ratio requirements, assuming amplified spontaneous emission is the dominant noise source. These include a) pulse-amplitude modulation with direct detection, b) differential phase-shift keying with interferometric detection, c) phase-shift keying with coherent detection, and d) quadrature-amplitude modulation with coherent detection.**

***Index Terms*—Differential phase-shift keying, heterodyning, homodyne detection, information rates, optical fiber communication, optical modulation, optical signal detection, phase-shift keying, pulse-amplitude modulation.**

I. Introduction

THE throughput of a dense wavelength-division-multiplexed (DWDM) transmission system can be increased by using a wider optical bandwidth, by increasing spectral efficiency, or by some combination of the two. Utilizing a wider bandwidth typically requires additional amplifiers and other optical components, so raising spectral efficiency is often the more economical alternative. Ultimate limits to spectral efficiency are determined by the information-theoretic capacity per unit bandwidth [1], [2]. While closely approaching these limits may require high complexity and delay [3], [4], establishing accurate estimates of the limits can yield insights useful in practical system design. In this paper, we review spectral efficiency limits, considering various modulation techniques (unconstrained, constant-intensity [5]–[7] or binary), various detection techniques (coherent or direct [8]), and various propagation regimes (linear or nonlinear [9], [10]). In all cases, amplified spontaneous emission (ASE) from inline optical amplifiers is assumed to be the dominant noise source. We show that coherent detection allows information to be encoded in two degrees of freedom per polarization, and its spectral efficiency limits are several b/s/Hz in typical terrestrial systems, even considering nonlinear effects. Using either constant-intensity modulation or direct detection, only one degree of freedom per polarization can be exploited, reducing spectral efficiency. Using binary modulation, regardless of detection technique, spectral efficiency cannot exceed 1 b/s/Hz per polarization.

When the number of signal and/or noise photons is small, the information-theoretic capacity of optical communication systems is also limited by the particle nature of photons. Coherent communication is equivalent to detecting the real and imaginary parts of the coherent states [11], [12]. Direct detection is equivalent to counting the number of photons in the number states [11], [12]. In the coherent states, quantum effects add one photon to the noise variance [12] (if both signal and noise are expressed in terms of photon number), yielding a channel capacity slightly smaller than the classical limit of [1] and [2]. The quantum limit of direct detection is determined by photon statistics and yields a slightly higher channel capacity than the classical limit of [8].

While currently deployed DWDM systems use binary on–off keying (OOK) with direct detection, in an effort to improve spectral efficiency and robustness against transmission impairments, researchers have investigated a variety of binary and nonbinary modulation techniques, in conjunction with various detection techniques. In this paper, we compare the spectral efficiencies and power efficiencies of several techniques, assuming in all cases that ASE is the dominant noise source. Techniques considered include pulse-amplitude modulation (PAM) with direct detection [13], [14], differential phase-shift keying (DPSK)

Manuscript received November 3, 2003; revised December 17, 2003. This work was supported in part by the National Science Foundation under Grant ECS-0335013 and in part by the National Science Council of R.O.C. under Grant NSC-92-2218-E-002-034.

J. M. Kahn is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: jmk@ee.stanford.edu).

K.-P. Ho is with the Institute of Communication Engineering and Department of Electrical Engineering, National Taiwan University, 106 Taipei, Taiwan, R.O.C. (e-mail: kpho@cc.ee.ntu.edu.tw).

Digital Object Identifier 10.1109/JSTQE.2004.826575

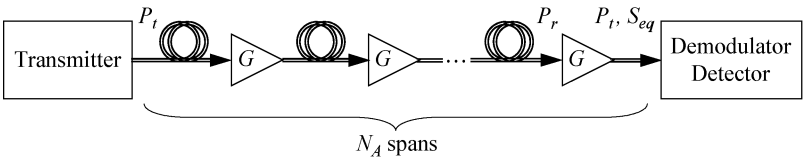

Fig. 1. Equivalent block diagram of multispan system. After each of the N_A fiber spans, an amplifier of gain G exactly compensates the span loss. The average power level per channel at key locations is indicated. At the output of the final amplifier, the total amplified spontaneous emission has a power spectral density per polarization given by S_{eq}.

with interferometric detection[1] [15]–[17], phase-shift keying (PSK) with coherent detection [18], [19], and (d) quadrature-amplitude modulation (QAM) with coherent detection [20]. We argue that at spectral efficiencies below 1 b/s/Hz per polarization, binary PAM (OOK) and binary DPSK are attractive options. At spectral efficiencies between 1 and 2 b/s/Hz, quaternary DPSK and PSK are perhaps the most attractive techniques. Techniques such as 8-PSK or 8- and 16-QAM are necessary to achieve spectral efficiencies above 2 b/s/Hz per polarization. Historically, the main advantages of coherent optical detection were considered to be high receiver sensitivity and the ability to perform channel demultiplexing and chromatic dispersion compensation in the electrical domain [21]. In this paper, we argue that from a more modern perspective, the principal advantage of coherent detection is superior spectral efficiency. We remind the reader that in ASE-limited coherent systems, there is no fundamental sensitivity difference between 2- and 4-PSK, nor is there one between synchronous homodyne and heterodyne detection.

The remainder of this paper is organized as follows. In Section II, we review information-theoretic spectral efficiency limits for various modulation and detection techniques, in linear and nonlinear propagation regimes. We also calculate the quantum-limited capacity due to the particle nature of photons. In Section III, we compare the power efficiencies and spectral efficiencies of various uncoded modulation and detection techniques, and comment on nonlinear phase noise and on coding techniques for nonbinary modulation. Conclusions and suggestions for future work are given in Section IV.

II. Spectral Efficiency Limits

A. Preliminaries

According to Shannon [1], the capacity of a communication channel is the maximum bit rate that can be transmitted without error, taking into account of noise, available bandwidth, and constrained power. Computing the capacity of a channel requires knowledge of the conditional probability of the output given the input, i.e., the characteristics of the channel "noise." Capacity calculations for a variety of discrete and continuous channels are described in [2]. It is remarkable that the capacity can be computed without explicitly considering any specific modulation, coding, or decoding scheme. Likewise, computation of the capacity does not generally tell us which specific modulation, coding, or decoding schemes we should use in order to achieve the capacity. The theory does tell us the optimal probability density of the transmitted signal. It indicates that we must use strong error-correcting codes [3], [4], and that the decoding complexity and delay must increase exponentially as we approach the capacity.

In a DWDM system, the spectral efficiency limit is simply the capacity per channel divided by the channel spacing. In the linear regime, the spectral efficiency limit is independent of chromatic dispersion, because it is possible, in principle, to fully compensate dispersion at the receiver. Alternatively, dispersion effects may be avoided by using narrow-band channels with small channel spacing. Initially, we consider propagation in a single polarization, neglecting the impact of polarization-mode dispersion. Polarization effects are addressed briefly in Section II-B4.

Throughout Section II, B denotes the occupied bandwidth per channel, Δf denotes the channel spacing, C denotes the capacity per channel, and $S = C/\Delta f$ denotes the spectral efficiency limit. Note that C and S have units of bits per second (b/s) and b/s/Hz, respectively. For concreteness, we consider a multispan system as shown in Fig. 1. The system comprises N_A fiber spans, each with gain $1/G$. After each span, an amplifier of gain G compensates the span loss. The average transmitted power per channel is P_t, while the average power at the input of each amplifier is $P_r = P_t/G$. We assume that for all detection schemes, ASE noise dominates over other noise sources, including local-oscillator shot noise or receiver thermal noise, thereby maximizing the receiver signal-to-noise ratio (SNR) [22]. At the output of the final amplifier, the ASE in one polarization has a power spectral density (PSD) given by

$$S_{eq} = N_A(G-1)n_{sp}h\nu = (G-1)n_{eq}h\nu \tag{1}$$

where n_{sp} is the spontaneous emission noise factor of one amplifier and we have defined the equivalent noise factor of the multispan system by $n_{eq} = N_A n_{sp}$. At the output of the final amplifier, the ASE in one polarization in the channel bandwidth B has a power

$$P_n = S_{eq}B. \tag{2}$$

Hence, at the output of the final amplifier, the optical SNR in one polarization in the channel bandwidth B is given by

$$\text{SNR} = \frac{P_t}{P_n}. \tag{3}$$

In optical communications, "coherent detection" has often been used to denote any detection process involving photoelectric mixing between an optical signal and local oscillator [21]. Information encoded in both in-phase and quadrature field components can be detected only using synchronous coherent detection. Synchronous homodyne or heterodyne detection requires an optical or electrical phase-locked loop (PLL) or some other carrier-recovery technique. It has been pointed out that in ASE-

[1]This DPSK detection technique uses an interferometer to convert phase modulation to intensity modulation, followed by a direct-detection receiver. While it is often called "direct detection," here we call it "interferometric detection" to distinguish it from simple direct detection.

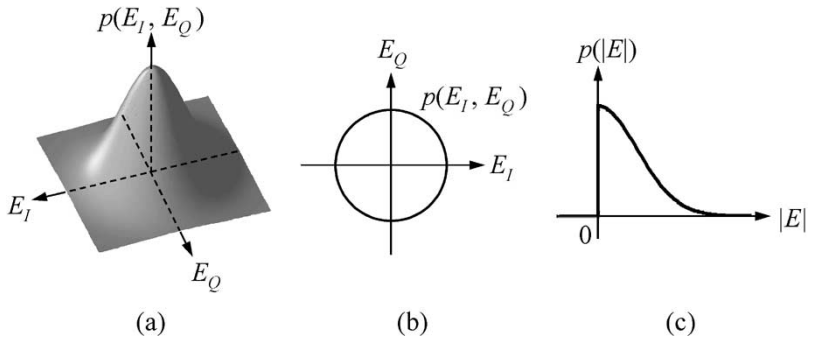

Fig. 2. Optimal probability density functions of the transmitted signal. (a) Unconstrained modulation, coherent detection: field is complex circular Gaussian-distributed. (b) Constant-intensity modulation, coherent detection: complex field is uniformly distributed on a circle. (c) Unconstrained modulation, direct detection: field magnitude is half-Gaussian-distributed.

limited systems, the sensitivity of a synchronous heterodyne receiver is equivalent to a synchronous homodyne receiver provided that the ASE is narrow-band-filtered or that image rejection is employed [23]. Throughout this paper, our use of the term "coherent detection" assumes that all of the above conditions have been satisfied and is fully consistent with its use in nonoptical communications [20]. These points are further elaborated upon in Section III-B.

B. Linear Regime

1) Unconstrained Modulation With Coherent Detection: In the case of coherent detection, information can be encoded in two degrees of freedom: the in-phase and quadrature field components E_I and E_Q. Alternatively, we can think of the two degrees of freedom as intensity and phase. ASE noise is modeled as additive, signal-independent complex circular Gaussian noise. When no constraints are placed on the modulation format, the optimal transmitted electric field is also complex circular Gaussian-distributed, as shown in Fig. 2(a). The capacity is given by the well-known Shannon formula [1]

$$C = B\log_2(1+\text{SNR}) \tag{4}$$

and the spectral efficiency limit is

$$S = \frac{B}{\Delta f}\log_2(1+\text{SNR}). \tag{5}$$

Note that at high SNR, the spectral efficiency (5) is given asymptotically by $S \sim (B/\Delta f)\log_2(\text{SNR})$.

2) Constant-Intensity Modulation With Coherent Detection: Various modulation techniques, such as DPSK and continuous-phase frequency-shift keying (CPFSK), encode information in optical signals having nominally constant intensity. These techniques are often demodulated using differentially coherent detection [20], to which interferometric detection is mathematically equivalent [24]. The capacity of constant-intensity modulation with coherent detection upper bounds the capacity with differentially coherent or interferometric detection. The capacity with coherent detection was first presented in [5] and was later derived independently in [6] and [7]. The optimal transmitted electric field is uniformly distributed on a circle, as shown in Fig. 2(b). At arbitrary SNR, the capacity is given by

$$C = B\left[-2\pi\int_0^\infty rf(r)\log_2 f(r)dr - \log_2 2\pi eP_n\right] \tag{6}$$

where

$$f(r) = \frac{1}{2\pi P_n}\exp\left(-\frac{P_s+r^2}{2P_n}\right)I_0\left(\frac{r\sqrt{P_s}}{P_n}\right). \tag{7}$$

At high SNR, using the asymptotic expression $I_0(u) \sim e^u/\sqrt{2\pi u}$, (7) is approximated as $f(r) \sim (1/((2\pi)^{3/2}(P_sP_n)^{1/2}))\exp(-((r-\sqrt{P_s})^2/2P_n))$, yielding the asymptotic capacity

$$C \sim B\left[\frac{1}{2}\log_2(\text{SNR}) + 1.10\right]. \tag{8}$$

The asymptotic spectral efficiency is $S \sim (B/\Delta f)[(1/2)\log_2(\text{SNR}) + 1.10]$. Observe that this limit is 1.1 b/s/Hz more than half the Shannon limit (5), assuming $B/\Delta f = 1$. Intuitively, the factor of two reduction results from discarding one degree of freedom, i.e., the field intensity.

3) Unconstrained Modulation With Direct Detection: In the case of direct detection, the transmitted optical signal is modeled as a nonnegative, real electric field magnitude.[2] In a well- designed system, the dominant noise is signal-spontaneous beat noise, which is additive and signal-dependent. To date, spectral efficiency limits have not been derived for arbitrary SNR. Mecozzi and Shtaif [8] have addressed the high-SNR limit. When no constraint is placed on the modulation format, the optimal transmitted field magnitude follows a half-Gaussian distribution, as shown in Fig. 2(c). Asymptotically, the capacity is given by

$$C \sim B\left[\frac{1}{2}\log_2(\text{SNR}) - 1.00\right] \tag{9}$$

and the spectral efficiency limit is $S \sim (B/\Delta f)[(1/2)\log_2(\text{SNR}) - 1.00]$. Note that this limit is 1.0 b/s/Hz less than half the Shannon limit (5), assuming $B/\Delta f = 1$. Intuitively, the factor-of-two reduction results from discarding one degree of freedom, i.e., one of the two field quadratures. The additional 1.0-b/s/Hz loss is caused by discarding the sign of the field.

4) Discussion of Linear Regime: Spectral efficiency limits in the linear regime are compared in Fig. 3 for a system using

[2]In practice, one may not only modulate the magnitude but also modulate the phase, either intentionally (e.g., duobinary encoding) or unintentionally (e.g., chirp). While phase modulation affects the spectrum, it does not affect the detected photocurrent, provided that dispersion is well compensated and the receiver does not employ an interferometer to convert phase modulation to intensity modulation.

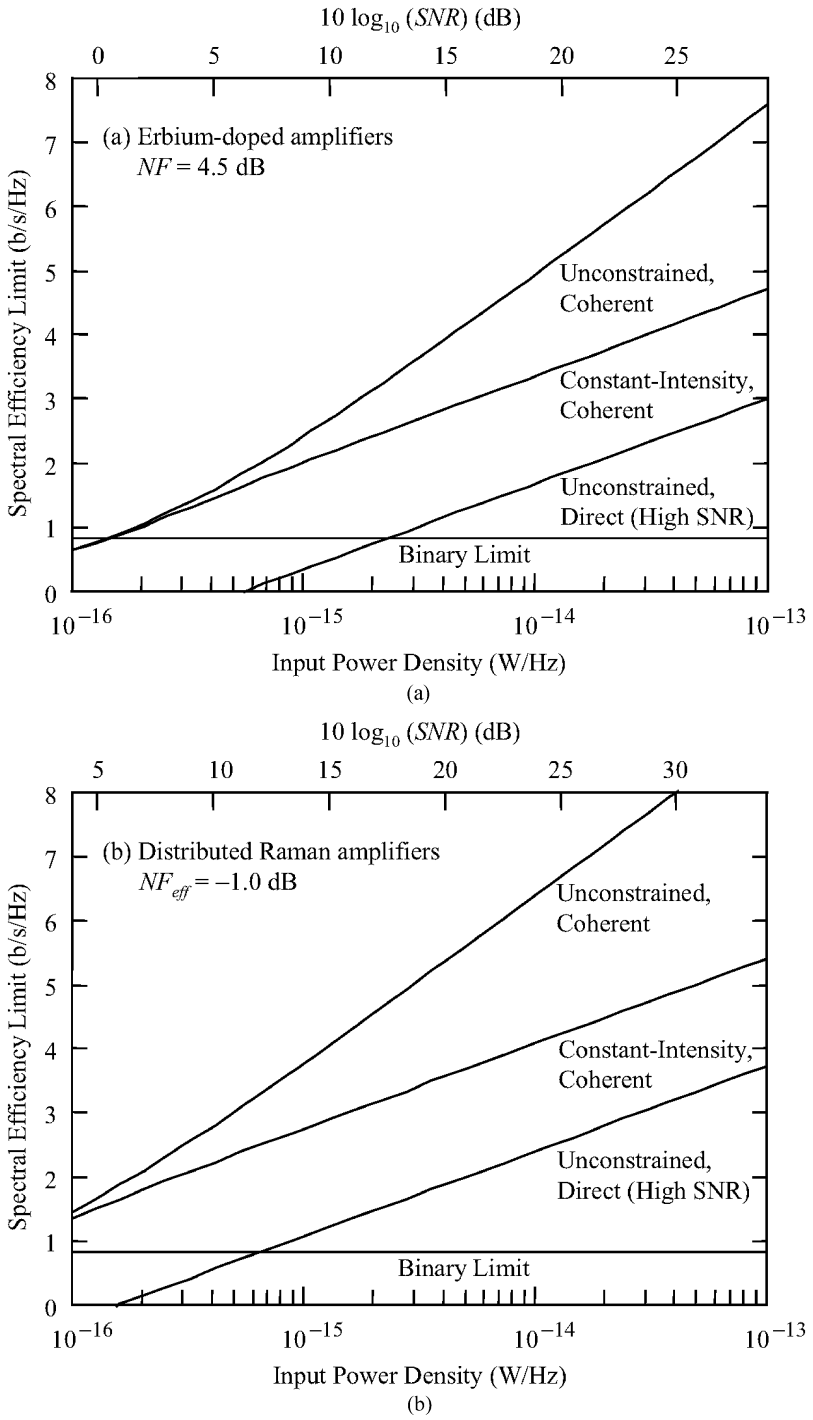

Fig. 3. Spectral efficiency limits in DWDM systems in linear regime. The x-axes correspond to input power density $P_t/\Delta f$ and SNR= $P_t/S_{eq}B$. The y-axis plots $S = C/\Delta f$. (a) Erbium-doped fiber amplifiers, NF= 4.5 dB; (b) distributed Raman amplifiers, $\text{NF}_{eff} = -1.0$ dB. Four cases are plotted: 1) unconstrained modulation with coherent detection, 2) constant-intensity modulation with coherent detection, 3) unconstrained modulation with direct detection (high-SNR asymptote), and 4) the binary modulation limit $B/\Delta f$ (which is the same for both amplifier types). Parameters assumed: bandwidth per channel $B = 40$ GHz, channel spacing $\Delta f = 50$ GHz, 25 fiber spans, 80 km span length, $\alpha = 0.20$ dB/km, $G = 16$ dB, $\lambda = 1.55$ μm, single polarization.

25 spans of 80 km length, assuming a fractional bandwidth utilization $B/\Delta f = 0.8$. Fig. 3(a) considers erbium-doped fiber amplifiers with noise figure NF= 4.5 dB, while Fig. 3(b) considers distributed Raman amplifiers with counterpropagating pump and effective noise figure $\text{NF}_{eff} = -1.0$ dB. In each case, the spontaneous emission noise factor n_{sp} is related to the noise figure by $n_{sp} = \text{NF}/2$. It is instructive to compare the high-SNR slopes of the spectral efficiency limits in Fig. 3(a) and (b). Unconstrained modulation with coherent detection (which offers two degrees of freedom per polarization) has a slope twice as large as constant-intensity modulation with coherent detection or unconstrained modulation with direct detection (each offers one degree of freedom per polarization). At realistic values of the SNR, unconstrained modulation with coherent detection offers a spectral efficiency limit more than twice as high as unconstrained modulation with direct

detection, and all three schemes offer spectral efficiencies well above the binary limit of $B/\Delta f = 0.8$ b/s/Hz.

Polarization-mode dispersion (PMD) has no fundamental impact on spectral efficiency limits. By making the bandwidth per channel B sufficiently small, we can neglect PMD-induced distortion.[3] In principle, the spectral efficiency limits described above can be doubled using polarization-division multiplexing (PDM), i.e., launching pairs of signals in orthogonal polarizations and employing polarization-resolved detection. When PMD is nonnegligible, pairs of signals must be launched into the two principal states of polarizations (PSPs). Because the PSPs change with time, adaptive polarization control is essential for PDM to operate without crosstalk between paired signals.

C. Nonlinear Regime

1) Spectral Efficiency Limits With Coherent Detection: Fiber nonlinearities limit the transmission distance and the overall capacity of DWDM systems [25]–[28]. The major fiber nonlinearities are the Kerr effect, stimulated Raman scattering, and stimulated Brillouin scattering [29]. The Kerr effect leads to self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). In the Kerr effect, the intensity of the aggregate optical signal perturbs the fiber refractive index, thereby modulating the phase of signals. In DWDM systems, SPM and XPM arise when the phase of a channel is modulated by its own intensity and by the intensity of other channels, respectively. FWM arises when two channels beat with each other, causing intensity modulation at the difference frequency, thereby phase-modulating all the channels and generating new frequency components. Fiber propagation with the Kerr effect is modeled using a nonlinear Schrödinger equation for single channel systems and coupled nonlinear Schrödinger equations for DWDM systems. Among the various nonlinearities, the Kerr effect has the greatest impact on channel capacity.

Early studies focused on the effect of fiber nonlinearity on specific modulation and detection techniques, including OOK with direct detection [25], [28] or simple modulations with coherent detection [26], [27]. Recently, the combined effect of ASE and Kerr nonlinearity on the Shannon capacity [1] has been studied [9], [10], [30]–[37].

Mitra and Stark [10] argued that the capacity of DWDM systems is limited most fundamentally by XPM. As a signal propagates, chromatic dispersion converts XPM-induced phase modulation to intensity noise. Capacity limitations caused by XPM are further studied in [30] and [31]. In fibers with nonzero dispersion, XPM has a much greater impact than FWM. DWDM systems with many channels are likely to be limited by XPM, perhaps allowing SPM to be ignored to first order. However, the methods of [10], [30], and [31] do not quantify the importance of SPM relative to XPM and cannot be applied single-channel systems, which are limited primarily by SPM.

With constant-intensity modulation, such as phase or frequency modulation [7], ideally, both SPM and XPM cause only time-invariant phase shifts, eliminating both phase and intensity distortion. If one could further neglect FWM, propagation would be linear; the capacity would be given by the expressions in [5]–[7], and increasing the launched power would lead to a monotonic increase in spectral efficiency. In reality, chromatic dispersion converts phase or frequency modulation to intensity modulation [32], and laser intensity noise and imperfect modulation cause additional intensity fluctuations. Hence, it is difficult to maintain constant intensity along an optical fiber. Furthermore, in reality, constant-intensity modulation is subject to FWM.

Tang [33]–[35] solved the nonlinear Schrödinger equation using a series expansion, similar to the Volterra series in [36]. The linear term is considered to be signal and all higher order terms are considered to be noise. If a sufficient number of terms is included, methods based on series expansion are very accurate. Tang [33]–[35] has included many terms, yielding a quasi-exact closed-form treatment.

In a single-channel system, the channel capacity is limited by SPM. In a DWDM system, when all channels are detected together, the impact of XPM can be reduced using a multiuser detection or interference-cancellation scheme. Using perturbation methods, Narimanov and Mitra [9] found the channel capacity of single-channel systems. For a single-channel system with zero average dispersion, the nonlinear Schrödinger equation has been solved analytically in [37] to find the channel capacity.

To quantify the SNR in the presence of XPM noise, Mitra and Stark [10] introduced a nonlinear intensity scale I_0. For transmitted power per channel well below I_0, increasing the power raises the SNR, increasing capacity. As the transmitted power approaches I_0, however, XPM noise increases rapidly, causing capacity to decrease precipitously. This nonlinear intensity scale I_0 is also applicable to the models of Tang [33]–[35], Narimanov and Mitra [9], constant-intensity modulation [7], and even the quantum limit [37]. In each of the models of [7], [9], [10], [30], [31], [33]–[35], the launched power that maximizes capacity increases with I_0. In DWDM systems, the nonlinear intensity scale I_0, and thus the capacity, increases with fiber dispersion, channel spacing, and signal bandwidth, and decreases with the total number of spans and the total number of channels.

In the pioneering paper [10], the nonlinear intensity scale for XPM was found to be

$$I_0 = \sqrt{\frac{BD\Delta\lambda}{2\gamma^2 \ln\left(\frac{N_c}{2}\right) L_{\text{eff}}}} \tag{10}$$

and the maximum spectral efficiency is lower bounded by

$$S = \frac{B}{\Delta f}\log_2\left(1 + \frac{e^{-\left(\frac{P_t}{I_0}\right)^2} P_t}{P_n + \left(1 - e^{-\left(\frac{P_t}{I_0}\right)^2}\right) P_t}\right) \tag{11}$$

where D is the fiber dispersion coefficient, γ is the fiber nonlinear coefficient, N_c is the number of DWDM channels, $L_{\text{eff}} \approx N_A/\alpha$ is the overall nonlinear length of the fiber for N_A fiber spans, and α is the fiber attenuation coefficient. Using the spectral efficiency lower bound (11), the power per channel that maximizes spectral efficiency is approximately equal to $(I_0^2 P_n/2)^{1/3}$, and the maximum spectral efficiency is approximately equal to $(B/\Delta f)(2/3)\log_2(2I_0/P_n)$ [10].

[3] At somewhat larger values of B, we can restrict attention to first-order PMD, and by launching a signal into one of the principal states of polarization (which requires adaptive polarization control at the transmitter), we can avoid PMD-induced distortion.

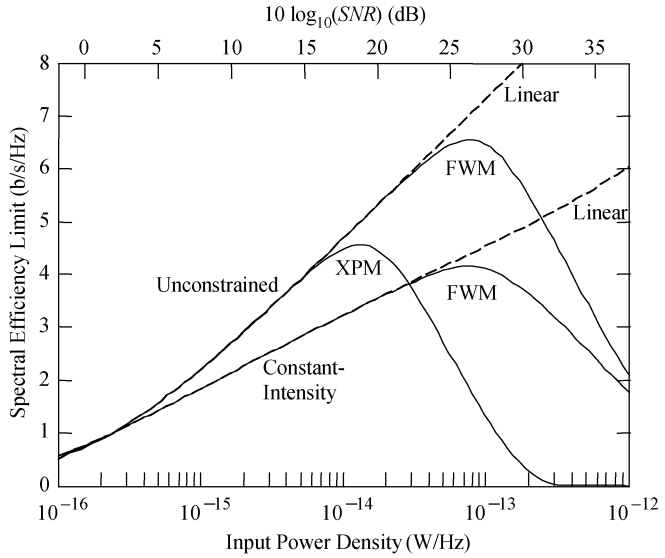

Fig. 4. Spectral efficiency limits in DWDM systems with coherent detection. Dashed lines describe the linear regime and are identical to those in Fig. 3(a). Solid lines correspond to the nonlinear regime as limited by XPM or FWM. All system parameters are the same as in Fig. 3(a), with the following additional parameters assumed: $\gamma = 1.24/\text{W/km}$, $N_c = 101$, and $D = 17$ ps/km/nm.

In a DWDM system limited by FWM instead of XPM, the spectral efficiency bound is given by (11), and the nonlinear intensity scale is given by

$$I_0^{-1} = N_A \sum_{\substack{p,q,p\neq 0,q\neq 0 \\ |p+q|\leq \frac{(n_c-1)}{2}}} \frac{\gamma^2 \left(\frac{D_{pq}}{3}\right)^2}{\alpha^2 + \Delta k_{pq}^2} \tag{12}$$

where $-(N_c-1)/2 \leq p, q \leq (N_c-1)/2$, $D_{pq}/3 = 1$ if $p = q$ and $D_{pq}/3 = 2$ if $p \neq q$, $\Delta k_{pq} = 2\pi\lambda^2 D \Delta f^2 qp/c$, λ is the optical wavelength, and c is the speed of light. Expression (12) has been derived for the center (worst case) channel.

Fig. 4 shows the nonlinear spectral efficiency limits of a 101-channel DWDM system using coherent detection and either unconstrained or constant-intensity modulation, assuming the same parameters as in Fig. 3(a). The limits in the linear regime, taken from Fig. 3(a), are shown as dashed lines. When either unconstrained or constant-intensity modulation is limited by FWM, an optimal power density of about 10^{-13}W/Hz maximizes spectral efficiency. Fig. 4 also shows the XPM-limited spectral efficiency for unconstrained modulation from [10], which is maximized for an input power density of about 10^{-14}W/Hz. For the specific parameters assumed in Fig. 4, the FWM-limited spectral efficiency of constant-intensity modulation is comparable to, but less than, the XPM-limited spectral efficiency of unconstrained modulation.

2) Discussion of Nonlinear Regime: While the nonlinear Schrödinger equation with noise provides a very accurate model for nonlinear propagation in fiber, the equation does not have an analytical solution except in some special cases [37]. While [7], [9], [10], [30], [31], [33]–[35], and [37] are based on this accurate formulation, they make different assumptions, leading to different estimates of channel capacity.

In order to illustrate the major qualitative differences between the various models, we consider a simplified memoryless monotonic transfer characteristic $y = f(x) = x + \varepsilon x^3$, where x and y are the input and output, respectively, and ε is a small number. While there is no nonlinear fiber channel having transfer characteristic $f(x)$, this simple function with linear term x and nonlinear term εx^3 yields insight into fiber systems. For a monotonic, one-to-one function such as $f(x)$, if we interpret both terms x and εx^3 as signal, then in the absence of any noise, the entropy of the output given the input $H(Y|X)$ is equal to zero. The mutual information between the input and the output, and thus the channel capacity, equals the entropy of the input x.

We draw an equivalence to the models [7], [9], [10], [30], [31], [33]–[35] by considering the linear term x to be signal and the nonlinear term εx^3 to be noise. In DWDM systems with many channels, the nonlinear term εx^3 includes "intermodulation products" corresponding to the XPM and FWM caused by other channels. As all channels are typically independent from one another, the models [10], [30], [31] concerning XPM-induced distortion can indeed model XPM as noise independent from the signal. Likewise, the model [7] concerning FWM components from other channels for constant-intensity modulation can model FWM to be noise independent from the signal. In a single-channel system [9], the nonlinear distortion caused by SPM depends on the signal and cannot be modeled as signal-independent noise. While only contributions from XPM are modeled as signal-independent noise in [10], [30], and [31], the series expansion model of [33]–[35] regards all higher order terms as noise independent of the signal. In fact, the two main groups of models are not complete because Tang [33]–[35] cannot account for the dependence of higher order terms on the signal and Mitra *et al.* [10], [30], [31] cannot include the higher order terms caused by SPM. In DWDM systems with many channels, the methods of [10], [30], [31], and [33]–[35] can be considered to be equivalent if the effect of SPM is negligible compared to XPM. In single-channel systems, where SPM must be considered, only the method of Turitsyn *et al.* [37] yields the probability density of the channel output including nonlinearity and uses it to calculate the channel capacity. In all cases, if the nonlinear term of εx^3 is considered as noise, the channel capacity decreases at high launched power, and a nonlinear intensity scale similar to (10) and (12) can be evaluated. In the single-channel system of Narimanov and Mitra [9], the channel capacity curve behaves like the curves in Fig. 4 for FWM.

The single-channel capacity of Turitsyn *et al.* [37] has been evaluated for fiber links with zero average dispersion. Only nonlinear phase noise, or the Gordon–Mollenauer effect [39], modulates the phase of the signal. The channel capacity is calculated using the probability density of the signal with nonlinear phase noise [37], [40]. In the limit of very high nonlinear phase noise, the capacity degenerates to that of direct detection (discussed in Section II-B3), which increases logarithmically with launched power. The nonlinear phase noise causes no amplitude noise.

The impact of Kerr nonlinearity can be reduced or canceled using phase conjugation [41], [42]. In DWDM systems with many channels, Kerr effect compensation reduces or cancels the nonlinear terms originating from other channels. In such a case, Kerr effect compensation yields the obvious benefit of reducing the "noise." In single-channel systems, Kerr effect compensation changes the statistics of the signal with noise. While midspan or distributed Kerr effect compensation can improve the capacity, Kerr effect compensation just before the receiver

does not improve capacity, and may actually reduce capacity by adding more noise.

D. Quantum Regime

When the number of signal and/or noise photons is small, quantum effects must be considered to compute the capacity of an optical channel [11], [12]. While a classical continuous-time channel can be converted to a discrete-time channel using the sampling theorem [1], the particle-based quantum channel does not have a corresponding sampling theorem. However, one may assume that a measurement is made within a time interval limited by the channel bandwidth. Hence, most studies of quantum-limited capacity assume a discrete-time channel [11], [12], [43]–[48].

When the coherent states of a radiation are used for communication, coherent detection can be used to detect the real and imaginary parts of the quantum states. The spectral efficiency with Gaussian noise is given by [12], [44], [46]

$$S = \log_2\left(1 + \frac{\overline{n}_s}{1+\overline{n}_{\rm ASE}}\right) \tag{13}$$

where $\overline{n}_s$ is the mean number of signal photons and $\overline{n}_{\rm ASE}$ is the mean number of ASE photons. The spectral efficiency (13) is the same as (4) with $B/\Delta f = 1$ and a quantum SNR given by $\overline{n}_s/(1+\overline{n}_{\rm ASE})$. Comparing this with the classical SNR $\overline{n}_s/\overline{n}_{\rm ASE}$, it is intuitively clear that there is a minimum of one noise photon in the coherent states [11], [12], [44]. In a system with fiber nonlinearity, the noise introduced by intermodulation products is equivalent to an increase in the number of ASE photons [38].

The intensity-modulation direct-modulation channel considered in Section II-B3 is equivalent to the classical limit of a photon-counting channel. The continuous-time photon-counting channel is modeled by information theorists as a Poisson channel with unlimited bandwidth but peak and average power constraints [49]–[52]. A discrete-time Poisson channel can be used to model a quantum-limited bandlimited photon-counting channel [11], [12], [43]–[46].

Optical amplification alters the photon statistics of amplified light. While an amplified signal has Poisson statistics, ASE noise obeys Bose–Einstein statistics. For a signal having n_S signal photons and $\overline{n}_{\rm ASE}$ ASE photons, the ASE noise is equivalent to Poisson-distributed light where the mean number of photons has the distribution

$$p_{\rm ASE}(\overline{n}) = \frac{1}{\overline{n}_{\rm ASE}} \exp\left(-\frac{\overline{n}}{\overline{n}_{\rm ASE}}\right), \quad \overline{n} \geq 0 \tag{14}$$

and the output photon-number distribution is given by [22], [53]

$$p_{n_S}(n) = \frac{\overline{n}_{\rm ASE}^n}{(1+\overline{n}_{\rm ASE})^{n+1}} \exp\left(-\frac{n_S}{1+\overline{n}_{\rm ASE}}\right) \times L_n\left(-\frac{\frac{n_S}{\overline{n}_{\rm ASE}}}{1+\overline{n}_{\rm ASE}}\right) \tag{15}$$

where L_n is the Laguerre polynomial

$$L_n(x) = \sum_{k=0}^{n} \binom{n}{k} \frac{(-x)^k}{k!}. \tag{16}$$

The distribution (15) includes only ASE noise having the same polarization as the signal.

If the input distribution is $p_S(n_S)$, the output distribution is $p_N(n) = \int_0^\infty p_S(n_S) p_{n_S}(n) dn_S$. The maximum spectral efficiency is

$$S = \max_{p_S(n_S)} \left\{ \int_0^\infty p_S(n_S) \sum_{n=0}^{\infty} p_{n_S}(n) \log_2 p_{n_S}(n) dn_S - \sum_{n=0}^{\infty} p_N(n) \log_2 p_N(n) \right\}. \tag{17}$$

As in [11], [43], and [44], the maximum spectral efficiency can be derived by maximizing the output entropy

$$H_N = -\sum_{n=0}^{\infty} p_N(n) \log_2 p_N(n). \tag{18}$$

If the input electric field is Gaussian distributed with $p_S(n_S) = \exp(-n_S/\overline{n}_S)/\overline{n}_S$, the output photon distribution is $p_N(n) = \overline{n}^n)/(1+\overline{n})^{(n+1)}$ and

$$H_N = \log_2(1+\overline{n}) + \overline{n}\log_2\left(1+\frac{1}{\overline{n}}\right) \tag{19}$$

where $\overline{n} = \overline{n}_s + \overline{n}_{\rm ASE}$. The spectral efficiency limit can be evaluated numerically using (17) and (19).

For a large number of photons and high SNR, the summation $\sum_{n=0}^{\infty} p_{n_S}(n) \log_2 p_{n_S}(n)$ can be approximated by integration, and the conditional distribution $p_{n_S}(n)$ can be assumed to be Gaussian distributed with variance $\sigma_n^2(n_s) = n_s + \overline{n}_{\rm ASE} + 2n_s\overline{n}_{\rm ASE} + \overline{n}_{\rm ASE}^2$ describing signal shot noise, ASE shot noise, signal-spontaneous beat noise, and spontaneous–spontaneous beat noise, respectively [22]. Based on the Gaussian approximation

$$\sum_{n=0}^{\infty} p_{n_S}(n) \log_2 p_{n_S}(n) \approx -\frac{1}{2}\log_2\left[2\pi e \sigma_s^2(n_s)\right]. \tag{20}$$

Similar to [12], we obtain

$$S \approx \frac{1}{2}\log_2\left(\frac{\overline{n}_s e^{1+\gamma_e}}{\overline{n}_{\rm ASE}\pi}\right) - 1 = \frac{1}{2}\log_2\left(\frac{\overline{n}_s}{\overline{n}_{\rm ASE}}\right) - 0.688 \tag{21}$$

where $\gamma_e \approx 0.5772$ is the Euler constant.

The asymptotic limit (21) is 0.312 bit/s/Hz higher than the classical intensity-modulation/direct-detection limit (9) for the same classical SNR $\overline{n}_s/\overline{n}_{\rm ASE}$. The limit (21) is also independent of the number of modes (or polarizations) in the ASE noise. For example, ASE in single and dual polarizations gives the same asymptotic limit (21). When the particle nature of photon is considered, spectral efficiency can be improved by 0.312 bit/s/Hz compared to the classical assumption of a continuous-amplitude channel. Based on photon statistics, the asymptotic limits in Fig. 3(a) and (b) for intensity modulation/direct-detection can be moved up by 0.312 bit/s/Hz.

Fig. 5(a) and (b) presents spectral efficiency limits given by (17) for a photon-counting channel with ASE noise. Fig. 5(a) shows the spectral efficiency (17) as a function of the classical SNR $\overline{n}_s/\overline{n}_{\rm ASE}$ for various values of the mean number of signal photons $\overline{n}_s$. Fig. 5(a) also shows the asymptotic limit (21) for large number of signal photons and high SNR. Fig. 5(b) shows the spectral efficiency (17) as a function of the mean signal photon number $\overline{n}_s$ for various values of the SNR. The case with infinite SNR corresponds to the spectral efficiency for Poisson-distributed photons [43], [44].

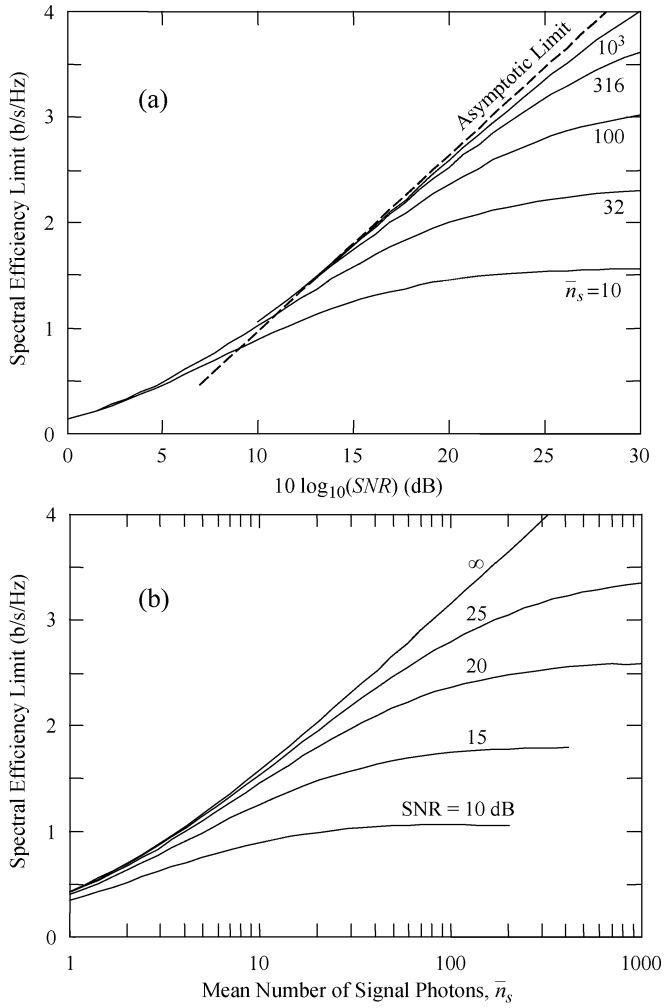

Fig. 5. Spectral efficiency limits for photon-counting channel with ASE noise (a) versus SNR for various mean numbers of signal photons and (b) versus mean number of signal photons for various SNRs.

Unlike the classical case in Fig. 3(a) and (b), in which the spectral efficiency depends only on the SNR, the quantum-limited spectral efficiency depends on both the SNR and the number of signal photons. Even at high SNR, high spectral efficiency cannot be achieved with a small number of signal photons. As shown in [38], the effect of fiber nonlinearity upon quantum-limited spectral efficiency is equivalent to an increase in the number of ASE photons.

III. Modulation, Detection, and Coding Techniques

A. Preliminaries

In this section, we compare the spectral efficiencies and SNR efficiencies of various modulation and detection techniques. We mainly consider uncoded modulation, deferring discussion of coding to Section III. The effects of laser phase noise and fiber nonlinearity (including nonlinear phase noise) are also deferred to Section III. Unless stated otherwise, we consider transmission in a single polarization.

Binary modulation encodes one bit per symbol, while nonbinary modulation encodes more than one bit per symbol, leading to higher spectral efficiency. While not considered in detail in this paper, nonbinary modulation also offers improved resistance to transmission impairments, including chromatic and polarization-mode dispersions [13], [54]. There are two reasons for this improved robustness. For a fixed bit rate, nonbinary modulation can employ a lower symbol rate than binary modulation, reducing signal bandwidth and thus reducing pulse spreading caused by chromatic dispersion. Also, because nonbinary modulation employs a longer symbol duration than binary modulation, it can often tolerate greater pulse spreading caused by chromatic and polarization-mode dispersions.

The information bit rate per channel in one polarization is given by

$$R_b = R_s R_c \log_2 M. \tag{22}$$

The bit rate R_b has units of b/s, and the symbol rate R_s has units of baud (symbols/s). For an occupied bandwidth per channel B, avoidance of intersymbol interference requires $R_s \leq B$, with $R_s = B$ being achieved by ideal pulses of the form $\mathrm{sinc}(R_s t)$ [20].[4] The unitless parameter $R_c \leq 1$ is the rate of an error-correction encoder that is used to add redundancy to the signal in order to improve the SNR efficiency. The uncoded case corresponds to $R_c = 1$. The parameter M is the number of points in the signal constellation. If the channel spacing is Δf, the spectral efficiency per polarization is

$$S = \frac{R_b}{\Delta f} = \frac{R_s R_c \log_2 M}{\Delta f} \leq \frac{B R_c \log_2 M}{\Delta f}. \tag{23}$$

Our figure of merit for spectral efficiency is $\log_2 M$, the number of coded bits per symbol, which determines spectral efficiency for fixed $R_s/\Delta f$ and fixed R_c.

Referring to Fig. 1, we note that at the input of the final amplifier, the average energy per information bit is $E_b = P_r/R_b$. At the output of the final amplifier, the average energy per information bit is

$$GE_b = \frac{GP_r}{R_b} = \frac{P_t}{R_b} \tag{24}$$

i.e., it is identical to the average transmitted energy per information bit. Recall that S_{eq} denotes the PSD of ASE at the output of the final amplifier (in one polarization). Our figure of merit for SNR efficiency is the value of the received SNR per information bit GE_b/S_{eq} required to achieve an information bit error ratio (BER) $P_b = 10^{-9}$. This figure of merit indicates the average energy that must be transmitted per information bit for fixed ASE noise, making it appropriate for systems in which the transmitted energy is constrained by fiber nonlinearities. Defining the average number of photons per information bit at the input of the final amplifier $n_b = E_b/h\nu$, and using (1), the figure of merit is equal to

$$\frac{GE_b}{S_{eq}} = \left(\frac{G}{G-1}\right)\frac{n_b}{n_{eq}} \approx \frac{n_b}{n_{eq}} \tag{25}$$

[4]This assumes double-sideband transmission. Using single-sideband transmission with real symbols and direct detection can reduce the bandwidth required per channel, but only in the limit of high SNR when a strong carrier is transmitted [55]. With coherent detection, use of quadrature-amplitude modulation is typically preferred over single-sideband transmission [20].

TABLE I
COMPARISON OF MODULATION AND DETECTION SCHEMES. M DENOTES THE NUMBER OF POINTS IN THE SIGNAL CONSTELLATION. NUMBERS GIVEN REPRESENT THE VALUES OF $GE_b/S_{eq} = n_b/n_{eq}$ (PHOTONS/BIT) REQUIRED FOR AN INFORMATION BER $P_b = 10^{-9}$. NUMBERS IN PARENTHESES ARE THE CORRESPONDING VALUES OF $10\log_{10}(GE_b/S_{eq}) = 10\log_{10}(n_b/n_{eq})$(dB). ONE AND TWO POLARIZATIONS ARE CONSIDERED. VALUES FOR PAM, $M \geq 4$, ARE ESTIMATED. QUESTION MARKS DENOTE UNKNOWN VALUES

M	$\log_2 M$	PSK/Coherent	QAM/Coherent	DPSK/Interferometric		PAM/Direct	
		One Polarization	One Polarization	One Polarization	Two Polarizations	One Polarization	Two Polarizations
2	1	18 (12.6)	Not applicable	20 (13.0)	22 (13.4)	38 (15.8)	41 (16.1)
4	2	18 (12.6)	18 (12.6)	31 (14.9)	?	134 (21.3)	?
8	3	41 (16.2)	29 (14.6)	83 (19.2)	?	443 (26.5)	?
16	4	119 (20.8)	45 (16.6)	240 (23.8)	?	1472 (31.7)	?

where the approximate equality is valid for high gain $G \gg 1$. Equation (25) indicates that the figure of merit can be viewed as the receiver sensitivity at the final amplifier input n_b (photons/bit), divided by the equivalent noise factor of the multispan system n_{eq}. The definition of n_b/n_{eq} is similar to the classical SNR $\overline{n}_s/\overline{n}_{\mathrm{ASE}}$ used in the quantum-limited spectral efficiencies (13) and (21).

B. PSK and QAM With Coherent Detection

As stated in Section II-A, we use "coherent detection" to denote synchronous homodyne or heterodyne detection; heterodyne detection is assumed to use narrow-band ASE filtering or image rejection so that it achieves the same performance as homodyne detection [23]. These forms of coherent detection achieve optimal SNR efficiency and, with QAM, can approximately double spectral efficiency by enabling two degrees of freedom per polarization.

It is not necessary for us to consider heterodyne or phase-diversity homodyne detection with differentially coherent (delay) demodulation of DPSK or CPFSK [20]. The interferometric detection scheme described in Section III-C is mathematically equivalent [24] and is easier to implement. Note also that CPFSK, particularly in nonbinary form, does not offer high spectral efficiency [20]. Likewise, it is not necessary for us to consider heterodyne or phase-diversity homodyne detection with noncoherent demodulation (envelope detection) of PAM, since the direct detection scheme described in Section III-D is mathematically equivalent [24] and is more easily implemented.

Homodyne detection requires a receiver electrical bandwidth approximately equal to the symbol rate R_s, a potential advantage at very high symbol rates. It requires a pair of balanced receivers for detection of M-ary PSK and QAM, $M \geq 4$, and usually employs an optical PLL. Heterodyne detection requires a receiver electrical bandwidth of approximately $2R_s$, which is potentially problematic at very high symbol rates. For detection of all modulation techniques, it requires only one balanced receiver, and it can use either an optical or electrical PLL. Most DWDM systems use optical demultiplexers that provide narrow-band filtering of the received signal and ASE, in which case, image rejection is not required for heterodyne to achieve the same performance as homodyne detection. Both homodyne and heterodyne detection require polarization tracking or polarization diversity. Our analysis assumes tracking, as it requires fewer photodetectors. Coherent system performance is optimized by using high amplifier gain G and a strong local-oscillator laser, so that local-oscillator-ASE beat noise dominates over receiver thermal noise and other noise sources [22]. This corresponds to the standard case of additive white Gaussian noise [20].

M-ary PSK uses a constellation consisting of M points equally spaced on a circle. In the case of uncoded 2- or 4-PSK, the BER is given by [20]

$$P_b = Q\left(\sqrt{\frac{2GE_b}{S_{eq}}}\right) = Q\left(\sqrt{\frac{2n_b}{n_{eq}}}\right) \tag{26}$$

where the Q function is defined in [20]. Achieving a BER $P_b = 10^{-9}$ requires $n_b/n_{eq} = 18$. The BER performance of M-PSK, $M > 4$, is computed in [20]. The spectral and SNR efficiencies of M-PSK for $M = 2, 4, 8, 16$ are compiled in Table I and are plotted as circles in Fig. 6.

M-ary QAM uses a set of constellation points that are roughly uniformly distributed within a two-dimensional region. In the cases $M = 2^{2m}$ ($M = 4, 16, \ldots$), the points are evenly arrayed in a $2^m \times 2^m$ square, while in the cases $M = 2^{2m+1}$ ($M = 8, 32, \ldots$), the points are often arranged in a cross. The BER performance of M-QAM is computed in [20]. The spectral and SNR efficiencies of M-QAM for $M = 4, 8, 16$ are compiled in Table I and are plotted as squares in Fig. 4.

C. DPSK With Interferometric Detection

M-ary DPSK uses a signal constellation consisting of M points equally spaced on a circle, like M-PSK. While M-PSK encodes each block of $\log_2 M$ bits in the *phase* of the transmitted symbol, M-DPSK encodes each block of $\log_2 M$ bits

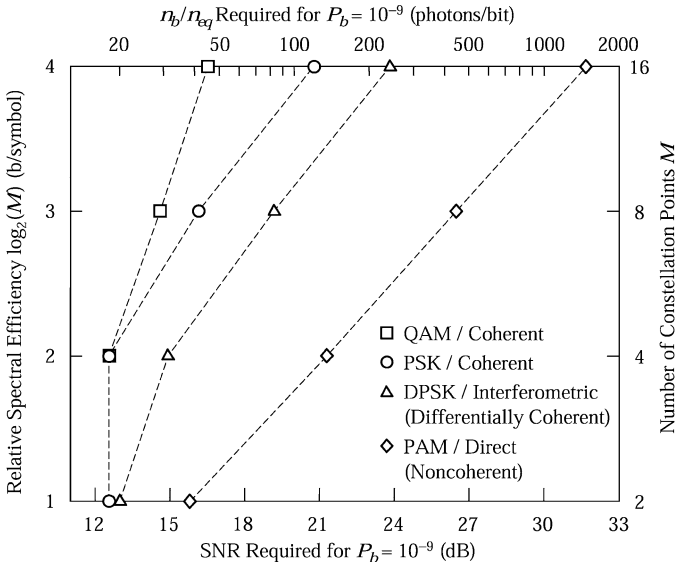

Fig. 6. Spectral efficiency versus SNR requirement for various modulation and detection techniques. The SNR requirement is the value of $10\log_{10}(GE_b/S_{eq}) = 10\log_{10}(n_b/n_{eq})$ required to achieve a BER $P_b = 10^{-9}$. One polarization is assumed.

in the *phase change* between successively transmitted symbols [20]. DPSK has recently been employed for fiber transmission with various formats, including nonreturn-to-zero (NRZ), return-to-zero (RZ), and chirped RZ (CRZ). In the absence of fiber nonlinearity, with proper dispersion compensation and matched filtering, all these formats fundamentally achieve the same SNR and spectral efficiency figures of merit considered here.

Under interferometric detection of two-DPSK, a Mach–Zehnder interferometer with a path difference of one-symbol duration compares the phases transmitted in successive symbols, yielding an intensity-modulated output that is detected using a balanced optical receiver. In the case of M-DPSK, $M \geq 4$, a pair of Mach–Zehnder interferometers (with phase shifts of 0 and $\pi/2$) and a pair of balanced receivers are used to determine the in-phase and quadrature components of the phase change between successive symbols.

Tonguz and Wagner [24] showed that for DPSK, the performance with optical preamplification and interferometric detection is equivalent to that with standard differentially coherent detection. Their work indicates that 2-DPSK requires $n_b/n_{eq} = 20$ with single-polarization filtering[5] and $n_b/n_{eq} = 22$ with polarization diversity in order to achieve a BER $P_b = 10^{-9}$ [24]. The performance of M-DPSK for $M \geq 4$ with single-polarization polarization filtering is described by the analysis in [20]. At low BERs, the SNR requirement for 4-DPSK is 2.3 dB higher than for 2- or 4-PSK, while for 8- and 16-DPSK, the SNR requirements are very nearly 3.0 dB higher than for 8- and 16-PSK, respectively. The spectral and SNR efficiencies of M-DPSK for $M = 2, 4, 8, 16$ with single-polarization filtering are compiled in Table I and are plotted as triangles in Fig. 4. The authors are not aware of a performance analysis of M-DPSK for $M \geq 4$ with polarization diversity.

[5]Single-polarization filtering uses a polarization-tracking receiver with a polarization analyzer that blocks the ASE polarized orthogonal to the signal.

D. PAM With Direct Detection

When used in conjunction with direct detection, M-ary PAM encodes a block of $\log_2 M$ bits by transmitting one of M intensity levels. Hence, the signal constellation consists of M points on the nonnegative real line.[6] Binary PAM, corresponding to OOK, is the most commonly used modulation technique. While OOK and nonbinary PAM have been employed with NRZ, RZ, and CRZ formats, all these formats fundamentally achieve the same SNR and spectral efficiency figures of merit considered here. A finite extinction ratio can strongly impact the SNR efficiency of M-PAM, particularly for $M \geq 4$. For simplicity, we consider an infinite extinction ratio.

Henry [58] and Humblet and Azizoglu [59] analyzed the performance of 2-PAM (OOK) with optical preamplification and direct detection. Including signal-spontaneous beat noise and spontaneous–spontaneous beat noise, the photocurrents are squares of Rayleigh- and Rician-distributed random variables (i.e., central and noncentral chi-square random variables) for intensity levels zero and one, respectively. In order to achieve a BER $P_b = 10^{-9}$, 2-PAM requires $n_b/n_{eq} = 38$ with single-polarization filtering and $n_b/n_{eq} = 41$ with polarization diversity.

The authors are not aware of an exact performance analysis of M-PAM for $M \geq 4$. Neglecting all noises except the dominant signal-spontaneous beat noise, at each intensity level, the photocurrent is Gaussian-distributed, with a variance proportional to the intensity. Setting the $M-1$ decision thresholds at the geometric means of pairs of adjacent levels approximately equalizes the downward and upward error probabilities at each threshold. In order to equalize the error probabilities at the $M-1$ different thresholds, the M intensity levels should form a quadratic series [14], i.e., 0, 1, 4, 9, etc. Under these assumptions, and assuming Gray coding, it is easily shown that the BER is approximately given by

$$P_b \approx \frac{1}{\log_2 M} Q\left(\sqrt{\frac{3\log_2 M}{(2M-1)(M-1)}\frac{GE_b}{S_{eq}}}\right) = \frac{1}{\log_2 M} Q\left(\sqrt{\frac{3\log_2 M}{(2M-1)(M-1)}\frac{n_b}{n_{eq}}}\right). \quad (27)$$

For $M = 2$, (27) indicates that $n_b/n_{eq} = 36$ is required for $P_b = 10^{-9}$, which is lower by 0.2 dB than the requirement $n_b/n_{eq} = 38$. For $M \geq 4$, (27) indicates that the SNR requirement increases by a factor $(3\log_2 M)/[(2M-1)(M-1)]$, corresponding to penalties of 5.5, 10.7, and 15.9 dB for $M = 4, 8, 16$, respectively. In order to estimate the SNR requirements of M-PAM for $M \geq 4$ with single-polarization filtering, we assume the exact requirement $n_b/n_{eq} = 38$ for $M = 2$ and then add the respective penalties for $M = 4, 8, 16$. The spectral and SNR efficiencies of M-PAM for $M = 2, 4, 8, 16$ with single-polarization filtering are compiled in Table I and are plotted as diamonds in Fig. 6.

[6]The constellation is sometimes expanded to $2M-1$ points by including the negatives of the $M-1$ nonzero points. The optical spectrum can be controlled by appropriately choosing between the positive and negatives points, e.g., using duobinary encoding [56], [57] or other line coding schemes. These techniques do not affect the SNR or spectral efficiency figures of merit considered here.

TABLE II
COMPARISON OF DETECTION TECHNIQUES. SHADING DENOTES AN ADVANTAGE

Attribute	Direct	Interferometric	Coherent
Maximum degrees of freedom per polarization	1	1	2
Signal-to-noise requirement for binary modulation (relative to 2-PAM with direct detection)	0 dB (2-PAM)	−2.8 dB (2-DPSK)	−3.2 dB (2-PSK)
Signal-to-noise requirement for quaternary modulation (relative to 2-PAM with direct detection)	+5.5 dB (4-PAM)	−0.9 dB (4-DPSK)	−3.2 dB (4-PSK)
Electrical filtering can select wavelength-division-multiplexed channel	No	No	Yes
Chromatic dispersion is linear distortion, making electrical compensation more effective	No	No	Yes
Local oscillator laser required at receiver	No	No	Yes
Polarization control or diversity required	No	No	Yes

TABLE III
LASER LINEWIDTH REQUIREMENTS FOR VARIOUS MODULATION AND DETECTION TECHNIQUES, ASSUMING A 0.5-dB PENALTY. FOR INTERFEROMETRIC DETECTION, TRANSMITTER HAS LINEWIDTH $\Delta\nu$, WHILE FOR COHERENT DETECTION, EACH OF THE TRANSMITTER AND LOCAL OSCILLATOR HAS LINEWIDTH $\Delta\nu$. QUESTIONS MARKS DENOTE UNKNOWN VALUES

Modulation	Detection	$\Delta\nu/R_b$	$\Delta\nu$ for $R_b = 10$ Gb/s	Reference
2-DPSK	Interferometric	3.0×10^{-3}	30 MHz	[77]
4-DPSK	Interferometric	?	?	
2-PSK	Coherent	8.0×10^{-4}	8 MHz	[78]
4-PSK	Coherent	2.5×10^{-5}	250 kHz	[79]

E. Discussion

Fig. 6 compares the spectral efficiencies and SNR requirements of the various modulation and detection techniques described in the sections above. We observe that for $M > 2$, the SNR requirements for PAM increases very rapidly, while the SNR requirements of the other three techniques increase at a more moderate rate. Note that for large M, the SNR requirements increase with roughly equal slopes for PAM, DPSK, and PSK, while QAM exhibits a distinctly slower increase of SNR requirement. This behavior can be traced to the fact that PAM, DPSK, and PSK offer one degree of freedom per polarization, while QAM offers two degrees of freedom per polarization. Based on Fig. 6, at spectral efficiencies below 1 b/s/Hz per polarization, 2-PAM (OOK) and 2-DPSK are attractive techniques. Between 1 and 2 b/s/Hz, 4-DPSK and 4-PSK are perhaps the most attractive techniques. At spectral efficiencies above 2 b/s/Hz, 8-PSK and 8- and 16-QAM become the most attractive techniques.

Table II compares some of the key attributes of direct, interferometric and coherent detection. The key advantages of interferometric detection over direct detection lie in the superior SNR efficiency of 2- and 4-DPSK as compared to 2- and 4-PAM. Coherent detection is unique in offering two degrees of freedom per polarization, leading to outstanding SNR efficiency for 2- and 4-PSK, and still reasonable SNR efficiency for 8-PSK and for 8- and 16-QAM. Coherent detection also enables electrical channel demultiplexing and chromatic dispersion compensation. Coherent detection requires a local oscillator laser and polarization control, which are two significant drawbacks.

Laser phase noise has traditionally been a concern for optical systems using DPSK, PSK, or QAM. In DPSK, interferometric detection can be impaired by changes in laser phase between successive symbols. Laser linewidth requirements for 2-DPSK have been studied in [77] but have not been studied for 4-DPSK, to our knowledge. In PSK and QAM systems, a PLL (optical or electrical) attempts to make the local oscillator track the laser phase noise, but the PLL operation is corrupted by additive Gaussian noise. Laser linewidth requirements for coherent detection have been studied for 2-PSK [78] and 4-PSK [79].[7] The laser linewidth requirements for 2-DPSK, 2-PSK, and 4-PSK are summarized in Table III. At a bit rate $R_b = 10\ \mathrm{Gb/s}$, the linewidth requirements for 2-DPSK (30 MHz) and 2-PSK (8 MHz) can be accommodated by standard distributed-feedback lasers. Four-PSK requires a much narrower linewidth (250 kHz), which can be achieved by compact external cavity lasers, which offer linewidths below 200 kHz [60].

Nonlinear phase noise [39] is induced by the interaction of ASE and signal through the Kerr effect. SPM- and XPM-induced nonlinear phase noise are added to the signal phase, limiting the performance of systems using phase-modulated signals, such as DPSK and PSK [61]–[66]. The statistical properties of nonlinear phase noise depend only on the mean nonlinear phase shift and the SNR [64]. The variance of SPM-induced nonlinear phase noise is proportional to the square of launched power per channel and inversely propor-

[7]These works addressed the local oscillator shot noise-dominated regime, where the sensitivity limit for heterodyne detection of 2-PSK and homodyne or heterodyne detection of 4-PSK is $n_b = 18$ for $P_b = 10^{-9}$. The results are also applicable to homodyne or heterodyne detection of 2- and 4-PSK in the ASE-limited regime, where the limit is $n_b/n_{eq} = 18$.

tional to the SNR [39], [64]. The SNR is proportional to the launched power per channel, so there is an optimal launched power per channel or, equivalently, an optimal mean nonlinear phase shift. As shown in [64]–[66], the optimal mean nonlinear phase shifts for 2-PSK and 2-DPSK are about 1.25 and 1.00 rad, respectively. For 4-DPSK, the optimal mean nonlinear phase shift is 0.89 rad [67]. As shown in Fig. 6, higher level modulations like QAM require higher SNR. As a high SNR decreases the variance of nonlinear phase noise, the optimal mean nonlinear phase shifts are likely to be around 1 rad, close to the estimate given in [39].

SPM-induced nonlinear phase noise is correlated with the received intensity, making it possible to partially compensate the nonlinear phase noise using the received intensity [64], [66]–[70]. When such compensation is performed, the variance of nonlinear phase noise can be reduced by a factor of two, doubling the optimal mean nonlinear phase shift. If nonlinear phase noise is the dominant impairment, the transmission distance can also be doubled. While SPM-induced nonlinear phase noise has received the most attention recently [62], [64]–[70], there have been few studies on XPM-induced nonlinear phase noise [63]. While SPM-induced nonlinear phase noise requires complicated models [64], XPM-induced nonlinear phase noise should be Gaussian-distributed and may be modeled as mathematically equivalent to laser phase noise.

Current fiber systems use binary modulation with error-control coding schemes that are well matched to binary modulation [71], such as Reed–Solomon codes, BCH codes, turbo codes [72], or low-density parity-check codes [73]. Nonbinary modulation will be needed to achieve spectral efficiencies above 1 b/s/Hz. Following lessons learned in nonoptical systems, making good use of nonbinary modulation requires appropriate coding schemes, including trellis-coded modulation (TCM) [74], turbo TCM [75], and shaping codes [76]. Systems using coherent detection can use coding schemes devised for nonoptical systems with Gaussian noise. In systems using direct detection or interferometric detection, the dominant noise at the receiver is not Gaussian-distributed; as a result, coding schemes developed for Gaussian-noise channels are not optimal. There is a need to study TCM, turbo TCM, and shaping codes for these systems. Accurate analytical bounds on error probability must be developed, which can lead to design criteria for good codes. Once good codes are identified, their coding gains must be evaluated.

IV. Conclusion and Future Work

Increasing spectral efficiency is often the most economical means to increase DWDM system capacity. In this paper, we have reviewed information-theoretical spectral efficiency limits for various modulation and detection techniques in both classical and quantum regimes, considering both linear and nonlinear fiber propagation regimes. Spectral efficiency limits for unconstrained modulation with coherent detection are several b/s/Hz in terrestrial DWDM systems, even considering nonlinear effects. Spectral efficiency limits are reduced significantly using either constant-intensity modulation or direct detection. Using binary modulation, regardless of detection technique, spectral efficiency cannot exceed 1 b/s/Hz per polarization.

We have compared the spectral efficiency and SNR requirements of various modulation and detection techniques in the ASE-limited regime. At spectral efficiencies below 1 b/s/Hz, binary PAM (OOK) and DPSK are attractive options. Between 1 and 2 b/s/Hz, quaternary DPSK and PSK are perhaps the most attractive techniques. Techniques such as 8-PSK or 8- and 16-QAM are necessary to achieve spectral efficiencies above 2 b/s/Hz per polarization.

Many interesting challenges remain. Nonlinear spectral efficiency limits for DWDM systems that simultaneously consider SPM and XPM are needed. Further investigations of nonlinear effects, including nonlinear phase noise, on DPSK, PSK, and QAM systems are warranted. Compensation of nonlinear effects using such diverse methods as optical phase conjugation and multiuser detection should be explored further. The impact of chromatic dispersion and polarization-mode dispersion on nonbinary modulation methods, including PAM, DPSK, PSK, and QAM, needs to be more fully quantified. Optimal coding techniques for nonbinary DPSK and PAM should be developed.

Optical signals propagating in fibers offer several degrees of freedom, including time, frequency, and polarization. The combined coding over these degrees of freedom has been seldom explored as a means to increase transmission capacity in fibers, especially as a way to combat or benefit from fiber nonlinearity and polarization-mode dispersion.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, July and Oct. 1948.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[3] C. Berrou, "The ten-year-old turbo codes are entering into service," *IEEE Commun. Mag.*, vol. 41, pp. 110–116, Aug. 2003.

[4] S.-Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, pp. 58–60, 2001.

[5] J. M. Geist, "Capacity and cutoff rate for dense M-ary PSK constellations," in *Proc. IEEE Mil. Commun. Conf.*, Monterey, CA, Sept. 30–Oct. 3 1990, pp. 168–770.

[6] J. P. Aldis and A. G. Burr, "The channel capacity of discrete time phase modulation in AWGN," *IEEE Trans. Inform. Theory*, vol. 39, pp. 184–185, 1993.

[7] K.-P. Ho and J. M. Kahn, "Channel capacity of WDM systems using constant-intensity modulation formats," in *Opt. Fiber Comm. Conf. OFC '02*, 2002, paper ThGG85.

[8] A. Mecozzi and M. Shtaif, "On the capacity of intensity modulated systems using optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 1029–1031, 2001.

[9] E. E. Narimanov and P. Mitra, "The channel capacity of a fiber optics communication system: perturbation theory," *J. Lightwave Technol.*, vol. 20, pp. 530–537, 2002.

[10] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fiber communications," *Nature*, vol. 411, pp. 1027–1030, 2001.

[11] C. M. Caves and P. D. Drummond, "Quantum limits on Bosonic communication rates," *Rev. Mod. Phys.*, vol. 66, pp. 481–537, 1994.

[12] M. J. W. Hall, "Gaussian-noise and quantum-optical communication," *Phys. Rev. A.*, vol. 50, pp. 3295–3303, 1994.

[13] S. Walklin and J. Conradi, "Multilevel signaling for increasing the reach of 10 Gb/s lightwave systems," *J. Lightwave Technol.*, vol. 17, pp. 2235–2248, 1999.

[14] J. Rebola and A. Cartaxo, "Optimization of level spacing in quaternary optical communication systems," in *Proc. SPIE*, vol. 4087, Quebec, Canada, June 12–16, 2000, pp. 49–59.

[15] M. Rohde, C. Caspar, N. Heimes, M. Konitzer, E.-J. Bachus, and N. Hanik, "Robustness of DPSK direct detection transmission format in standard fiber WDM systems," *Electron. Lett.*, vol. 36, pp. 1483–1484, Aug. 1999.

[16] A. H. Gnauck, G. Raybon, S. Chandrasekhar, J. Leuthold, C. Doerr, L. Stulz, A. Agarwal, S. Banerjee, D. Grosz, S. Hunsche, A. Kung, A. Marhelyuk, D. Maywar, M. Movagassaghi, X. Liu, C. Xu, X. Wei, and D. M. Gill, "2.5 Tb/s (64 × 42.7 Gb/s) transmission over 40 × 100 km NZDSF using RZ-DPSK format and all-Raman-amplified spans," in *Tech. Dig. Postdeadline Papers Optical Fiber Commun. Conf. 2002*, pp. FC2.1–FC2.3.

[17] R. A. Griffin and A. C. Carter, "Optical differential quadrature phase-shift key (oDQPSK) for high capacity optical transmission," in *Proc. Optical Fiber Commun. Conf. 2002*, 2002, paper WX6.

[18] J. M. Kahn, A. H. Gnauck, J. J. Veselka, S. K. Korotky, and B. L. Kasper, "4 Gbit/s PSK homodyne transmission system using phase-locked semiconductor lasers," *IEEE Photon. Technol. Lett.*, vol. 2, pp. 285–287, Apr. 1990.

[19] S. Norimatsu, K. Iwashita, and K. Noguchi, "An 8 Gb/s QPSK optical homodyne detection experiment using external-cavity laser diodes," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 765–767, 1992.

[20] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[21] G. P. Agrawal, *Fiber Optic Communication Systems*, 3rd ed. New York: Wiley, 2002.

[22] E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications*. New York: Wiley, 1994.

[23] B. F. Jorgensen, B. Mikkelsen, and C. J. Mahon, "Analysis of optical amplifier noise in coherent optical communication systems with optical image rejection receivers," *J. Lightwave Technol.*, vol. 10, pp. 660–671, 1992.

[24] O. K. Tonguz and R. E. Wagner, "Equivalence between preamplified direct detection and heterodyne receivers," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 835–837, 1991.

[25] A. R. Chraplyvy, "Limitation on lightwave communications imposed by optical-fiber nonlinearities," *J. Lightwave Technol.*, vol. 8, pp. 1548–1557, 1990.

[26] N. Shibata, K. Nosu, K. Iwashita, and Y. Azuma, "Transmission limitations due to fiber nonlinearities in optical FDM systems," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1068–1077, 1990.

[27] R. G. Waarts, A. A. Friesem, E. Lichtman, H. H. Yaffe, and R.-P. Baurn, "Nonlinear effects in coherent multichannel transmission through optical fibers," *Proc. IEEE*, vol. 78, pp. 1344–1368, 1990.

[28] A. R. Chraplyvy and R. W. Tkach, "What is the actual capacity of single-mode fibers in amplified lightwave systems?," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 666–668, 1993.

[29] G. P. Agrawal, *Nonlinear Fiber Optics*, 2nd ed. New York: Academic, 1995.

[30] J. B. Stark, P. Mitra, and A. Sengupta, "Information capacity of nonlinear wavelength division multiplexing fiber optic transmission line," *Opt. Fiber Technol.*, vol. 7, pp. 275–288, 2001.

[31] A. G. Green, P. B. Littlewood, P. P. Mitra, and L. G. L. Wegener, "Schrödinger equation with a spatially and temporally random potential: Effects of cross-phase modulation in optical communication," *Phys. Rev. E*, vol. 66, 2002.

[32] S. Norimatsu and K. Iwashita, "The influence of cross-phase modulation on optical FDM PSK homodyne transmission systems," *J. Lightwave Technol.*, vol. 11, pp. 795–804, 1993.

[33] J. Tang, "The Shannon channel capacity of dispersion-free nonlinear optical fiber transmission," *J. Lightwave Technol.*, vol. 19, pp. 1104–1109, 2001.

[34] ——, "The multispan effects of Kerr nonlinearity and amplifier noises on Shannon channel capacity for a dispersion-free nonlinear optical fiber," *J. Lightwave Technol.*, vol. 19, pp. 1110–1115, 2001.

[35] ——, "The channel capacity of a multispan DWDM system employing dispersive nonlinear optical fibers and an ideal coherent optical receiver," *J. Lightwave Technol.*, vol. 20, pp. 1095–1101, 2002.

[36] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single mode fibers," *J. Lightwave Technol*, vol. 15, pp. 2232–2241, 1997.

[37] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, p. 203 901, 2003.

[38] E. Desurvire, "A quantum model for optically amplified nonlinear transmission systems," *Opt. Fiber Technol.*, vol. 8, pp. 210–230, 2002.

[39] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, 1990.

[40] J.-A. Huang and K.-P. Ho, "Probability density of signal with nonlinear phase noise," in *Pacific Rim Conf. Laser Electro-Optic CLEO/PR '03*, Taipei, Taiwan, 2003.

[41] D. M. Pepper and A. Yariv, "Compensation for phase distortions in nonlinear media by phase conjugation," *Opt. Lett.*, vol. 5, pp. 59–60, 1980.

[42] I. Brener, B. Mikkelsen, K. Rottwitt, W. Burkett, G. Raybon, J. B. Stark, K. Parameswaran, M. H. Chou, M. M. Fejer, E. E. Chaban, R. Harel, D. L. Philen, and A. Kosinski, "Cancellation of all Kerr nonlinearities in long fiber spans using a $\mathrm{LiNbO_3}$ phase conjugator and Raman amplification," in *Opt. Fiber Comm. Conf., OFC '00*, vol. 4, 2000, pp. 266–268.

[43] T. Stern, "Some quantum effects in information channels," *IEEE Trans. Inform. Theory*, vol. IT-6, pp. 435–440, 1960.

[44] J. P. Gordon, "Quantum effects in communications systems," *Proc. IRE*, vol. 50, pp. 1898–1908, 1962.

[45] Y. Yamamoto and H. A. Haus, "Preparation, measurement and information capacity of optical quantum states," *Rev. Mod. Phys.*, vol. 58, pp. 1001–1020, 1986.

[46] M. J. W. Hall and M. J. O'Routke, "Realistic performance of the maximum information channel," *Quantum Opt.*, vol. 5, pp. 161–180, 1993.

[47] M. Ban, M. Sasaki, and M. Takeoka, "Continuous variable teleportation as a generalized thermalizing quantum channel," *Phys. Rev. A*, vol. 35, pp. 401–405, 2002.

[48] M. Ban, "Transmission rate of classical information through the thermalizing quantum channel," *J. Phys. A.*, vol. 36, pp. 7397–7409, 2003.

[49] Y. M. Kabanov, "The capacity of a channel of Poisson type," *Theory Prob. Appl.*, vol. 23, pp. 143–147, 1978.

[50] M. H. A. Davis, "Capacity and cutoff rate for Poisson-type channels," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 710–715, 1980.

[51] J. L. Massey, "Capacity, cutoff-rate and coding for a direct-detection optical channel," *IEEE Trans. Commun.*, vol. COM-29, pp. 1616–1621, 1981.

[52] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel—Part I and II," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1449–1471, 1988.

[53] P. Diament and M. C. Teich, "Evolution of the statistical properties of photons passed through a traveling-wave laser amplifier," *IEEE J. Quantum Electron.*, vol. 28, pp. 1325–1334, 1992.

[54] J. Wang and J. M. Kahn, "Impact of chromatic and polarization-mode dispersions on DPSK systems using interferometric demodulation and direct detection," *J. Lightwave Technol.*, vol. 22, no. 2, pp. 362–371, 2004.

[55] A. Mecozzi and M. Shtaif, "Noiseless amplification and signal-to-noise ratio in single-sideband transmission," *Opt. Lett.*, vol. 28, pp. 203–205, 2003.

[56] T. Ono, Y. Yano, K. Fukuchi, T. Ito, H. Yamazaki, M. Yamaguchi, and K. Emura, "Characteristics of optical duobinary signals in terabit/s capacity, high-spectral-efficiency WDM systems," *IEEE J. Lightwave Technol.*, vol. 16, pp. 188–797, 1998.

[57] J. M. Kahn and K.-P. Ho, "Transmission and reception of duobinary multilevel pulse-amplitude-modulated optical signals using finite-state machine-based encoder," U.S. Patent 6 424 444, July 23, 2002.

[58] P. S. Henry, "Error-rate performance of optical amplifiers," in *Proc. Conf. Optical Fiber Commun.*, Washington, DC, 1989, p. 170.

[59] P. A. Humblet and M. Azizoglu, "On the bit error rate of lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 9, pp. 1576–1582, 1991.

[60] J. D. Berger, Y. Zhang, J. D. Grade, H. Lee, S. Hrinya, H. Jerman, A. Fennema, A. Tselikov, and D. Anthon, "Widely tunable external cavity diode laser using a MEMS electrostatic rotary actuator," in *Proc. 27th Euro. Conf. Optical Commun.*, Amsterdam, the Netherlands, Sept. 30–Oct. 4 2001.

[61] A. Mecozzi, "Limits to long-haul coherent transmission set by the Kerr nonlinearity and noise of the in-line amplifiers," *J. Lightwave Technol.*, vol. 12, no. 11, pp. 1993–2000, 1994.

[62] H. Kim and A. H. Gnauck, "Experimental investigation of the performance limitation of DPSK systems due to nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 320–322, 2003.

[63] H. Kim, "Cross-phase-modulation-induced nonlinear phase noise in WDM direct-detection DPSK systems," *J. Lightwave Technol.*, vol. 21, pp. 1770–1774, 2003.

[64] K.-P. Ho, "Statistical properties of nonlinear phase noise," in *Advances in Optics and Laser Research*, W. T. Arkin, Ed. Hauppauge, NY: Nova Science, 2003, vol. 3.

[65] ——, "Performance degradation of phase-modulated systems with nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 1213–1215, 2003.

[66] ——, "Compensation improvement of DPSK signal with nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 1216–1218, 2003.

[67] J.-A. Huang and K.-P. Ho, "Exact error probability of DQPSK signal with nonlinear phase noise," in *Pacific Rim Conf. Laser Electro-Optic CLEO/PR '03*, Taipei, Taiwan, 2003.
[68] X. Liu, X. Wei, R. E. Slusher, and C. J. McKinstrie, "Improving transmission performance in differential phase-shift-keyed systems by use of lumped nonlinear phase-shift compensation," *Opt. Lett.*, vol. 27, pp. 1616–1618, 2002.
[69] C. Xu and X. Liu, "Postnonlinearity compensation with data-driven phase modulators in phase-shift keying transmission," *Opt. Lett.*, vol. 27, pp. 1619–1621, 2002.
[70] K.-P. Ho and M. Kahn, "Compensation technique to mitigate nonlinear phase noise," *J. Lightwave Technol.*, vol. 22, no. 3, pp. 779–783, 2004.
[71] P. V. Kumar, M. Z. Win, H.-F. Lu, and C. N. Georghiades, "Error-control coding techniques and applications," in *Optical Fiber Telecommunications IV B*, I. P. Kaminow and T. Li, Eds. San Diego, CA: Academic, 2002, pp. 902–964.
[72] T. Mizuochi, K. Ouchi, T. Kobayashi, Y. Miyata, K. Kuno, H. Tagami, K. Kubo, H. Yoshida, M. Akita, and K. Motoshima, "Experimental demonstration of net coding gain of 10.1 dB using 12.4 Gb/s block turbo code with 3-bit soft decision," in *Proc. OFC 2003*, postdeadline paper 21.
[73] B. Vasic, I. B. Djordjevic, and R. K. Kostuk, "Low-density parity check codes and iterative decoding for long-haul optical communication systems," *J. Lightwave Technol.*, vol. 21, pp. 438–446, Feb. 2003.
[74] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets—Part I: Introduction," *IEEE Commun. Mag.*, vol. 25, pp. 5–11, Feb. 1987.
[75] P. Robertson and T. Worz, "A novel bandwidth-efficient coding scheme employing turbo codes," in *Proc. Int. Conf. Commun.*, New York, 1996, pp. 962–967.
[76] G. D. Forney and L. Wei, "Multidimensional constellations—Part I: Introduction, figures of merit and generalized cross constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 941–958, Aug. 1989.
[77] C. P. Kaiser, P. J. Smith, and M. Shafi, "An improved optical heterodyne DPSK receiver to combat laser phase noise," *J. Lightwave Technol.*, vol. 13, pp. 525–533, Mar. 1995.
[78] S. Norimatsu and K. Iwashita, "Linewidth requirements for optical synchronous detection systems with nonnegligible loop delay time," *J. Lightwave Technol.*, vol. 10, pp. 341–349, Mar. 1992.
[79] J. R. Barry and J. M. Kahn, "Carrier synchronization for homodyne and heterodyne detection of optical quadriphase-shift keying," *J. Lightwave Technol.*, vol. 10, pp. 1939–1951, Dec. 1992.

Joseph M. Kahn (M'90–SM'98–F'00) received the A.B., M.A., and Ph.D. degrees in physics from the University of California (UC) at Berkeley, in 1981, 1983, and 1986, respectively.

From 1987 to 1990, he was with AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990 to 2003, he was on the Faculty of the Department of Electrical Engineering and Computer Sciences at UC Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering at Stanford University. His current research interests include single- and multimode optical fiber communications, free-space optical communications, and MEMS for optical communications.

Prof. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993–2000, he was a Technical Editor of IEEE PERSONAL COMMUNICATIONS MAGAZINE.

Keang-Po Ho (S'91–M'95–SM'03) received the B.S. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1991 and the M.S. and Ph.D. degrees from the University of California at Berkeley in 1993 and 1995, respectively, all in electrical engineering.

He performed research at the IBM T. J. Watson Research Center, Hawthorne, NY, on all-optical networks in the summer of 1994. He was a Research Scientist with Bellcore (currently Telcordia Technologies), Red Bank, NJ, from 1995 to 1997, where he conducted research on optical networking, high-speed lightwave systems, and broadband access. He was with the Department of Information Engineering, Chinese University of Hong Kong, from 1997 to 2001. He was Cofounder and Chief Technology Officer of StrataLight Communications, Campbell, CA, from 2000 to 2003, where he developed high-spectral efficiency 40-Gb/s systems. He has been with the Institute of Communication Engineering, National Taiwan University, since 2003. His research interests include optical communication systems, high-speed transmission, combined source-channel coding, and communication theory. He has published more than 120 journal and conference papers in those fields.