Carrier Synchronization for 3- and 4-bit-per-Symbol Optical Transmission

Ezra Ip and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—We investigate carrier synchronization for coherent detection of optical signals encoding 3 and 4 bits/symbol. We consider the effects of laser phase noise and of additive white Gaussian noise (AWGN), which can arise from local oscillator (LO) shot noise or LO-spontaneous beat noise. We identify 8- and 16-ary quadrature amplitude modulation (QAM) schemes that perform well when the receiver phase-locked loop (PLL) tracks the instantaneous signal phase with moderate phase error. We propose implementations of 8- and 16-QAM transmitters using Mach–Zehnder (MZ) modulators. We outline a numerical method for computing the bit error rate (BER) of 8- and 16-QAM in the presence of AWGN and phase error. It is found that these schemes can tolerate phase-error standard deviations of 2.48° and 1.24°, respectively, for a power penalty of 0.5 dB at a BER of 10^{-9}. We propose a suitable PLL design and analyze its performance, taking account of laser phase noise, AWGN, and propagation delay within the PLL. Our analysis shows that the phase error depends on the constellation penalty, which is the mean power of constellation symbols times the mean inverse power. We establish a procedure for finding the optimal PLL natural frequency, and determine tolerable laser linewidths and PLL propagation delays. For zero propagation delay, 8- and 16-QAM can tolerate linewidth-to-bit-rate ratios of 1.8×10^{-5} and 1.4×10^{-6}, respectively, assuming a total penalty of 1.0 dB.**

I. Introduction

MOST DENSE wavelength-division-multiplexing (DWDM) systems currently employ binary modulation schemes, such as ON-OFF keying (OOK) or binary differential phase-shift keying (2-DPSK). Systems employing these binary schemes cannot achieve spectral efficiencies exceeding 1 bit/s/Hz per polarization [1]. Recent research has focused on increasing spectral efficiency by using nonbinary modulation formats, such as 4-DPSK [2], which encodes 2 bits/symbol, or even 8-DPSK [3], which encodes 3 bits/symbol. Information-theoretic studies suggest that even when optical nonlinearities are considered, spectral efficiencies of several bits per symbol are possible [1], [4].

Fig. 1 shows the spectral efficiencies and signal-to-noise ratio (SNR) per bit required to achieve a bit error rate (BER) of 10^{-9} for various modulation and detection techniques [1].[1] It is evident that the SNR requirements for M-DPSK with differentially coherent detection increase substantially as M increases to 4, 8, and 16, corresponding to spectral efficiencies of 2, 3, and 4 bits/symbol. At such high spectral efficiencies, substantially lower SNR requirements are obtained using 4-PSK, 8-quadrature amplitude modulation (QAM), and 16-QAM with coherent detection.[2]

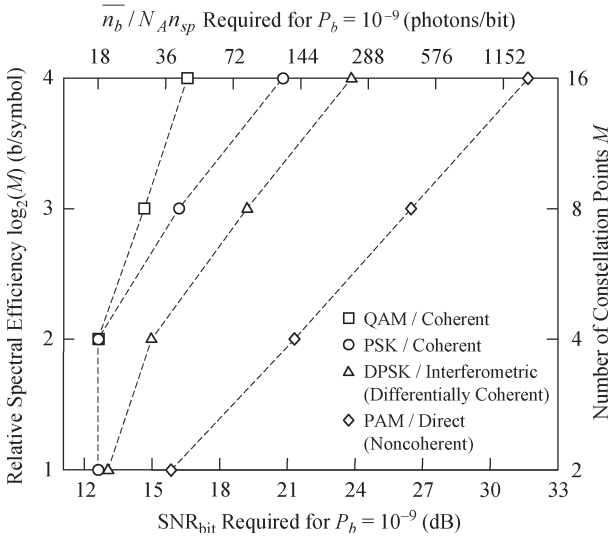

Fig. 1. Spectral efficiency versus SNR per bit required for a BER of 10^{-9} for various transmission schemes. One polarization is assumed (taken from [1]).

A major challenge in coherent detection lies in carrier synchronization between the local oscillator (LO) and the transmitter at optical frequencies. Phase noise is an inherent property of lasers, and can also arise from nonlinear processes [5] such as self- and cross-phase modulation and four-wave mixing. Phase noise increases the tracking errors in a coherent receiver's phase-locked loop (PLL). When the phase noise is large, it can dominate over additive white Gaussian noise (AWGN) arising from LO shot noise and inline amplifier noise, becoming the principal source of system degradation. The PLL tracking error depends on: 1) the statistics of laser phase noise, 2) the symbol rate, 3) the type of PLL employed, and 4) the delay in the PLL feedback path. To ensure that PLL phase error does not impose an excessive receiver-sensitivity penalty, we need to mathematically characterize the combined effects of phase noise and AWGN on PLL tracking performance. Carrier synchronization has been studied thoroughly for 4-PSK [6], [7]. In this

Manuscript received March 28, 2005; revised June 26, 2005. This research was supported by National Science Foundation Grant ECS-0335013.

The authors are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: wavelet@stanford.edu; jmk@ee.stanford.edu).

Digital Object Identifier 10.1109/JLT.2005.859428

[1]We shall define the SNR per bit in Sections III-A and IV-B.

[2]Throughout this paper, "coherent detection" denotes synchronous detection by a phase-locked LO, and should be distinguished from noncoherent detection (e.g., of OOK) and differentially coherent detection (e.g., of DPSK) [1].

paper, we study carrier synchronization for 8- and 16-QAM. We focus on the effect of laser phase noise, neglecting nonlinear phase noise. We assume perfect matching between the signal and LO states of polarization.

This paper is organized as follows. In Section II, we review various signaling schemes that are suitable candidates for 8- and 16-ary optical transmission. We examine the relative merits of these schemes and determine which constellations are most attractive at moderate phase errors. We propose transmitter implementations using Mach–Zehnder (MZ) modulators. In Section III, we compute the BER performance for the chosen 8- and 16-ary constellations at different phase-error standard deviations. We also determine the maximum phase errors that these constellations can tolerate for different power penalties at various target BERs. In Section IV, we analyze carrier synchronization in the presence of phase noise, AWGN, and PLL propagation delay, determining the maximum tolerable laser linewidths and PLL propagation delays. Our methods for computing BER and analyzing PLL tracking error are sufficiently general that they can be applied to any signaling scheme employing amplitude and phase modulation in two-dimensional (2-D) signal space.

II. System Considerations

A. Various 8- and 16-Point Constellations

Fig. 2 shows some well-known 8- and 16-point constellations that have been employed in nonoptical systems, and which are suitable candidates for consideration here. Owing to differences in their packing densities and the different angular separation between their signal points, these constellations perform differently with respect to AWGN and phase noise [8]. Also, implementation of transmitters and receivers is less complex for some constellations than for others.

We first consider the 16-point constellations shown in Figs. 2(d)–(f). The most common 16-point constellation is 16-QAM [Fig. 2(d)], in which points are arrayed on a 4×4 square grid. A major attraction of 16-QAM is its relatively low implementation complexity. A 16-QAM transmitter requires the fewest MZ modulators among the 16-point schemes considered here.[3] The in-phase (I) and quadrature (Q) components are separable, so a receiver for 16-QAM can make decisions on I and Q independently. Performance advantages are also offered by 16-QAM. It has the highest packing density under an average power constraint of 16-point constellations considered here, so it has the best performance with respect to AWGN. Of the 16-point constellations, 16-QAM is the only one for which Gray coding between bits and symbols is possible, minimizing the BER for a given symbol error rate. The main disadvantage of 16-QAM is that the signal points at the corners of the constellation are poorly separated in angle, making the scheme susceptible to phase error. Only when the phase-error standard deviation is under 1° does 16-QAM have the best performance among the schemes considered here [8]. Other 16-point options include the 1–5–10 and 5–11 constellations shown in Fig. 2(e) and (f), respectively. Their main advantage over 16-QAM is that the outer signal points are uniformly distributed on concentric circles, maximizing their angular separation. Hence, the 1–5–10 and 5–11 constellations are attractive options when the phase-error standard deviation is in the range of 1–1.5° [8]. Implementation of transmitters and receivers is more complex for 1–5–10 and 5–11 than for 16-QAM, and Gray coding is not possible.

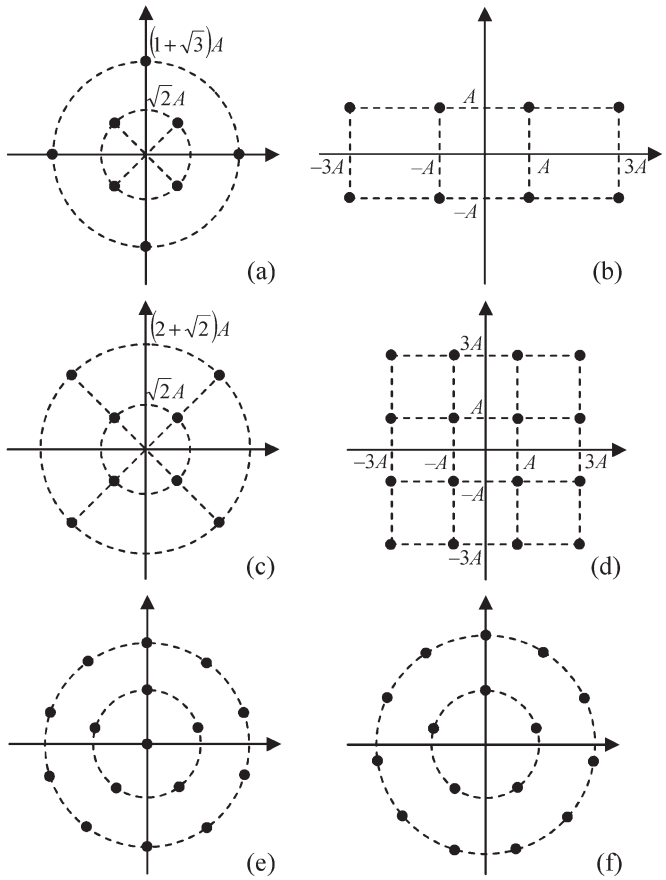

Fig. 2. Various constellations having [(a), (b), and (c)] 8 and [(d), (e), and (f)] 16 points.

Based on the above considerations, we identify the 16-QAM square constellation shown in Fig. 2(d) as the most attractive choice, so it is the only 16-point constellation considered in the remainder of this paper.

We now consider the 8-point constellations shown in Figs. 2(a)–(c). The constellation shown in Fig. 2(b), with points arranged on a 2×4 grid, has separable I and Q components, potentially simplifying receiver implementation. It does not have the highest packing density under an average power constraint among 8-point constellations, so it does not offer the best performance against AWGN. Most importantly, the signal points furthest from the origin are poorly separated in angle, leading to poor performance in the presence of phase error. The constellations shown in Fig. 2(a) and (c) each consist of two sets of four points uniformly distributed on concentric circles. Hence, they have identical performance against phase error in the limit that phase noise dominates over AWGN. Among the 8-point constellations considered here, the 8-QAM cross

[3]As explained in Section III-C, we assume that a transmitter uses MZ modulators that are driven into saturation in order to optimize transmitted signal quality [9].

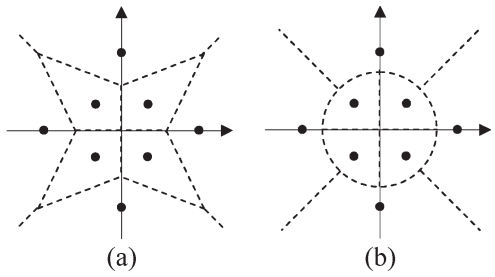

Fig. 3. ML decision regions for 8-QAM. (a) Optimal in the AWGN-limited case. (b) Optimal in the phase-noise-limited case.

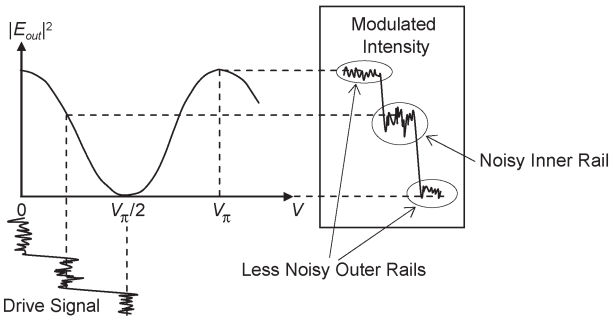

Fig. 4. MZ modulator output for a noisy drive signal.

shown in Fig. 2(a) has the highest packing density, and thus, the best performance against AWGN. Although this is the only one of the three 8-point constellations for which Gray coding is not possible, Gray coding is secondary to packing density at the high SNRs of interest. A transmitter for this constellation also can be implemented with reasonable complexity, as shown below.

Based on the above considerations, we identify the 8-QAM cross constellation shown in Fig. 2(a) as the most attractive choice, so it is the only 8-point constellation considered in the remainder of this paper.

B. Detection of 8- and 16-QAM

The optimal detector employs maximum-likelihood (ML) detection, which minimizes the probability of symbol error. The decision regions in an ML detector depend on the relative magnitudes of AWGN and phase error. When phase errors are absent and all signal points in the constellation are equally likely to be transmitted, an ML detector makes decisions in favor of the symbol located at the smallest Euclidean distance from the received signal. When a phase error is present however, the ML decision regions are determined by a non-Euclidean metric [8]. An example of this is illustrated in Fig. 3, where the optimal decision regions for AWGN-limited and phase-noise-limited cases are shown for 8-QAM. In a more general situation where neither AWGN nor phase error completely dominates, the ML decision regions are more complicated.

Practical considerations need to be taken into account when implementing a decision device, particularly at high symbol rates. Arbitrary decision regions can be implemented by digitizing I and Q components and using a digital lookup table. As analog-to-digital converters at rates of 10 GHz and above become increasingly practical, this approach will be the most flexible and will likely lead to the best performance.

Other means exist to implement decision devices. For 16-QAM in the AWGN-limited case, one can achieve rectangular decision regions by splitting the signal into I and Q components and applying three decision thresholds to each component. However, these rectangular decision regions are not optimal in the presence of significant phase error. For 8-QAM in the AWGN-limited case, the decision regions shown in Fig. 3(a) can be implemented by forming various linear combinations of I and Q components and applying decision thresholds to them. For 8-QAM in the phase-noise-limited case, the decision regions shown in Fig. 3(b) can be implemented by using an envelope detector and threshold to determine whether the signal lies in the inner or outer circle; and by comparing I and Q components, and their sum and difference, to determine the quadrant in which the received signal lies.

In this paper, we analyze the performance of 8- and 16-QAM under the assumption of optimal ML decisions for any particular combination of AWGN and phase noise that may be present. Although a practical receiver may need to implement some approximation of the ML decision regions, this assumption is necessary for mathematical tractability.

C. Transmitters for 8- and 16-QAM

Using MZ modulators and couplers arranged in various structures, 8- and 16-QAM may be generated. An MZ modulator has a transfer characteristic given by

$$\frac{E_{\rm out}}{E_{\rm in}} = \frac{1}{2}\left[\exp\left(j\pi\frac{V_1}{V_\pi}\right) + \exp\left(j\pi\frac{V_2}{V_\pi}\right)\right] \qquad (1)$$

where $E_{\rm in}$ and $E_{\rm out}$ are input and output electric fields, and V_π is the device-specific voltage that produces a π phase shift between the two arms of the MZ modulator. In a dual-drive modulator, V_1 and V_2 are two independent drive voltages, while in a single-drive modulator, the drive voltage is $V = V_1 = -V_2$. Ho showed that an arbitrary QAM signal set may be generated with just one dual-drive MZ modulator [9]. Although the hardware savings in this implementation appear attractive, the set of drive signals become arbitrarily large as the number of signal points increases. Furthermore, the nonlinear characteristic of the MZ modulator causes drive signal noise[4] to have an unequal effect on different signal levels. Consider M-pulse amplitude modulation (PAM) generated with a single-drive modulator. The output intensity versus drive voltage is shown in Fig. 4. Noise causes the outer eye rails to have smaller deviations than the inner rails owing to the MZ-modulator characteristic saturating at its maxima and minima. In order to minimize this effect, it is desirable to use drive signals that change by integer multiples of V_π between successive symbols periods.

[4]Although we refer to this as "noise" for brevity, it typically arises from signal distortion in practice.

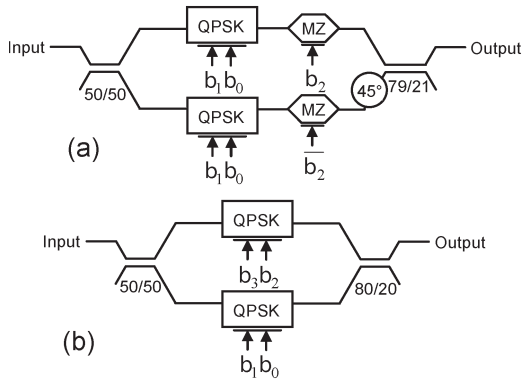

Fig. 5. Possible transmitter configurations for (a) 8-QAM and (b) 16-QAM.

We observe that 8- and 16-QAM may be decomposed into simpler quaternary phase-shift keying (QPSK) constellations. We can construct 8-QAM as a superposition of two QPSK constellations having different powers; the points in the outer QPSK circle are rotated 45° from those in the inner QPSK circle. This scheme may be implemented using the structure shown in Fig. 5(a). In each symbol period, a block of three signal bits control the transmitter. The least significant bits drive identical QPSK modulators in the lower and upper branches. The most significant bit drives complementary MZ modulators biased for binary OOK transmission. In either polarity of the most significant bit, one of the OOK transmitters is turned ON and the other is turned OFF. The outputs of the OOK transmitters are combined in a coupler having a splitting ratio of 78.9%/21.1% to produce QPSK rings with an average power ratio of $(1+\sqrt{3})^2 : (\sqrt{2})^2$.

We can decompose 16-QAM similarly into a Cartesian product of two QPSK constellations, as shown in Fig. 6. The structure that implements this is shown in Fig. 5(b). The two most significant bits control the QPSK modulator in the upper branch while the least significant bits control the QPSK modulator in the lower branch. The outputs are combined with a coupler having a splitting ratio of 80%/20%.

The transmitters shown in both Fig. 5(a) and (b) require QPSK modulators. Each QPSK modulator may be realized with either one dual-drive MZ modulator, or two MZ modulators that are single or dual drive. The latter configuration requires an extra modulator, but the I and Q eye patterns do not exhibit overshoot between successive symbols [9].

III. BER Performance of 8- and 16-QAM

The effect of phase error is usually evaluated in terms of a power penalty, which is defined as the additional SNR required at the receiver to achieve the same target BER as that required in the absence of phase error. In this section, we compute the BER performance of 8- and 16-QAM for different phase-noise standard deviations. We assume that AWGN and laser phase noise are the only sources of impairment. We assume that intersymbol interference and nonlinear phase noise are negligible. The latter requirement can be satisfied by using sufficiently low transmit power [5]. This enables us to model phase error using well-known statistical distributions for PLL tracking error.

A. Computing the Conditional Probability for the Received Signal

Let $\mathbf{s}_j$ be a vector in 2-D space representing the transmitted signal. $\mathbf{s}_j$ can be from an arbitrary constellation. When the signal at the detector of the receiver has phase error θ, its mean position is rotated from $\mathbf{s}_j$, as shown in Fig. 7. AWGN adds a random noise vector $\mathbf{n}$ to the rotated signal to produce a vector $\boldsymbol{z}$ of the form

$$\mathbf{z} = \mathbf{R}(\theta)\mathbf{s}_j + \mathbf{n}. \tag{2}$$

$\mathbf{R}(\theta)$ is a 2×2 rotation matrix given by $\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Let the probability density function (pdf) of phase error be $p(\theta)$, where θ can take on values between $-\pi$ and π. The probability that $\mathbf{z}$ is received given symbol $\mathbf{s}_j$ was sent is

$$p(\mathbf{z}|\mathbf{s}_j) = \int_{-\pi}^{\pi} p(\mathbf{z}|\mathbf{s}_j, \theta) \cdot p(\theta) d\theta. \tag{3}$$

Since phase error and AWGN are the only channel impairments, if we assume the noise in I and Q are independent, the conditional probability $p(\mathbf{z}|\mathbf{s}_j, \theta)$ is a circular Gaussian distribution, whose variance is N_0 in each dimension:

$$p(\mathbf{z}|\mathbf{s}_j, \theta) = \frac{1}{2\pi N_0} \exp\left\{ -\frac{\|\mathbf{z} - \mathbf{R}(\theta)\mathbf{s}_j\|^2}{2N_0} \right\}. \tag{4}$$

The SNR per symbol is

$$\text{SNR} = \frac{E\left[|s_k|^2\right]}{2N_0}. \tag{5}$$

To compare the average-power-limited performance of systems with different spectral efficiencies, the SNR per bit is a more useful parameter [1]. The SNR per bit is given by $\text{SNR}_{\text{bit}} = (1/\log_2 M)\text{SNR}$ for a constellation with M signal points encoding $\log_2(M)$ bits per symbol.

In order to compute $p(\mathbf{z}|\mathbf{s}_j)$ using (3), we need to find a suitable pdf for $p(\theta)$. If the receiver employs a first-order PLL whose input is a sine wave corrupted by Gaussian noise, the slowly varying component of phase error has a Tikhonov distribution, whose pdf is [8]

$$p(\theta) = \frac{1}{2\pi} \frac{\mathrm{e}^{\alpha \cos\theta}}{I_0(\alpha)}. \tag{6}$$

$I_0(\alpha)$ is the modified Bessel function of the zeroth order. The parameter α is the carrier-to-noise ratio (CNR) at the PLL input. Although this distribution is derived for a first-order PLL, it is a good approximation of the phase error in a second-order PLL [10], which is the type we analyze in

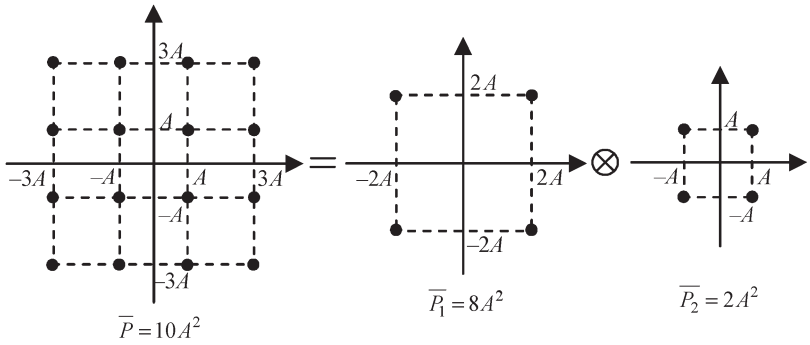

Fig. 6. Decomposed of 16-QAM into the Cartesian product of two QPSK systems can be done.

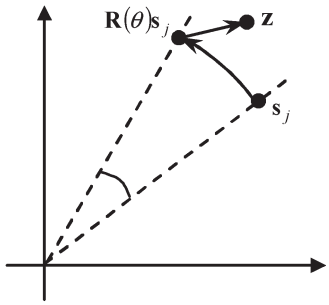

Fig. 7. Effect of phase error and AWGN on the received signal.

Section IV. Substituting (4) and (6) into (3), the conditional probability $p(\mathbf{z}|\mathbf{s}_j)$ is [8]

$$p(\mathbf{z}|\mathbf{s}_j) = \frac{1}{2\pi N_0} \frac{I_0\left(\frac{\beta_j}{N_0}\right)}{I_0(\alpha)} \cdot \exp\left\{-\frac{1}{2N_0}\left[\|\mathbf{z}-\mathbf{s}_j\|^2 + 2\langle \mathbf{z}, \mathbf{s}_j\rangle\right]\right\} \quad (7)$$

where

$$\beta_j = \sqrt{\|\mathbf{z}\|^2\|\mathbf{s}_j\|^2 + 2\alpha N_0\langle \mathbf{z}, \mathbf{s}_j\rangle + (\alpha N_0)^2}. \quad (8)$$

Equation (7) can be simplified using the following approximation for $I_0(x)$ at large x:

$$I_0(x) \approx 0.4\frac{\mathrm{e}^x}{\sqrt{x}}. \quad (9)$$

A plot of the ratio between the left- and right-hand sides of (9) is shown in Fig. 8. We observe that this approximation is accurate to within 5% for $x > 3$. In all practical systems, α will be much greater than 3 (4.8 dB). We can therefore evaluate (7) as follows:

1) If $\beta_j/N_0 > 3$

$$p(\mathbf{z}|\mathbf{s}_j) = \frac{1}{2\pi N_0}\sqrt{\frac{\alpha N_0}{\beta_j}} \cdot \exp\left\{-\frac{1}{2N_0}\left[\|\mathbf{z}-\mathbf{s}_j\|^2 + 2\langle \mathbf{z}, \mathbf{s}_j\rangle + 2\alpha N_0 - 2\beta_j\right]\right\}. \quad (10)$$

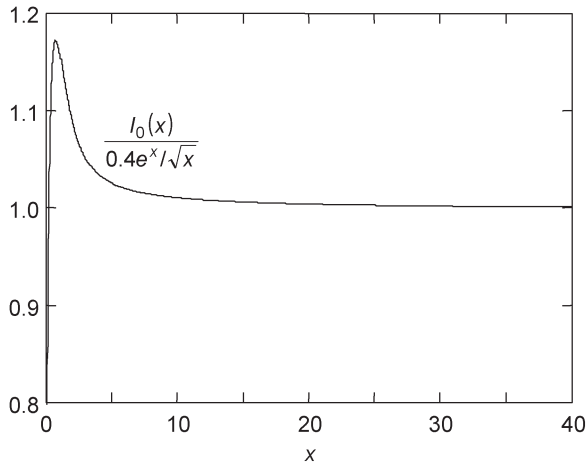

Fig. 8. Approximation to the Bessel function $I_0(x)$.

2) Else

$$p(\mathbf{z}|\mathbf{s}_j) = \frac{1}{2\pi N_0}\frac{\sqrt{\alpha}}{0.4} \cdot I_0\left(\frac{\beta_j}{N_0}\right) \cdot \exp\left\{-\frac{1}{2N_0}\left[\|\mathbf{z}-\mathbf{s}_j\|^2 + 2\langle \mathbf{z}, \mathbf{s}_j\rangle + 2\alpha N_0\right]\right\}. \quad (11)$$

B. Computing the Probability of Symbol Error

As discussed in Section II-B, the probability of symbol error depends on the decision regions employed by the receiver. In this paper, we assume ML decision regions, even though a practical receiver may have to implement some approximation to them. When the system is operating at low probability of error, the penalty associated with using suboptimal decision regions should be small.

It is generally not possible to obtain analytical formulas for the probability of symbol error when phase error and AWGN are present simultaneously. In these circumstances, numerical methods are used to compute system performance. We use a procedure developed by Foschini *et al.* [8] in which the 2-D space of the received signal is sampled at $N \times N$ points $\{\mathbf{z}_i = (x_i, y_i) : i = 1, 2, \ldots, N^2\}$ distributed evenly in a rectangular region $(x, y) : -\gamma A \le x \le \gamma A$, $-\gamma A \le y \le \gamma A$, as shown in Fig. 9. This method is not specific to the constellations we are considering in this paper. It may be used to compute the

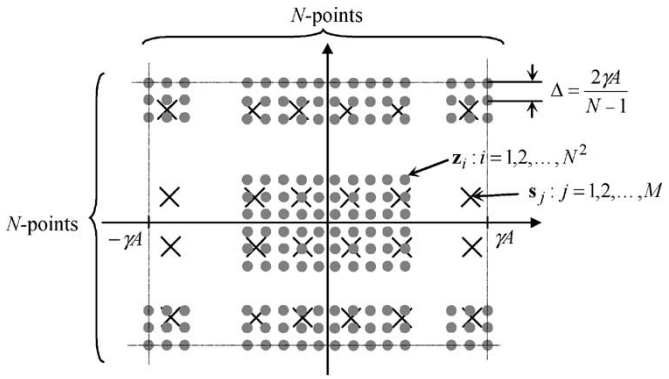

Fig. 9. Discretizing the received signal space to allow numerical calculation of the probability of symbol error.

BER of any general 2-D signaling scheme employing amplitude and/or phase modulation. γ is chosen to be sufficiently large to encompass all regions where the conditional probability (7) is significant. The larger the value of N, the more accurately we can evaluate the probability of symbol error.

The probability of the detector correctly decoding a symbol can be approximated by a discrete sum:

$$P(\text{correct}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(\text{correct}|\mathbf{z}) \cdot p(\mathbf{z})\mathrm{d}\mathbf{z} \cong \sum_{i=1}^{N^2} P(\text{correct}|\mathbf{z}_i) \cdot p(\mathbf{z}_i)\Delta x \Delta y \qquad (12)$$

where Δx and Δy are the horizontal and vertical separations between neighboring samples of $\mathbf{z}_i$. In Fig. 9, we observe that

$$\Delta x = \Delta y = \frac{2\gamma A}{N-1}. \qquad (13)$$

In ML detection, we declare symbol $\mathbf{s}_k$ was transmitted if

$$p(\mathbf{z}_i|\mathbf{s}_k) = \max_j \{p(\mathbf{z}_i|\mathbf{s}_j)\}. \qquad (14)$$

The probability of a correct decision is then equal to

$$P(\text{correct}|\mathbf{z}_i) = p(\mathbf{s}_k|\mathbf{z}_i). \qquad (15)$$

Using Baye's rule on the right-hand side of (15), we have

$$p(\mathbf{s}_k|\mathbf{z}_i) = \frac{p(\mathbf{z}_i|\mathbf{s}_k) \cdot p(\mathbf{s}_k)}{p(\mathbf{z}_i)}. \qquad (16)$$

Assuming all constellation points are equally likely to be transmitted, $p(\mathbf{s}k) = 1/M$, we may substitute (16) into (15) to compute the conditional probability of making a correct decision

$$P(\text{correct}|\mathbf{z}_i) \cdot p(\mathbf{z}_i) = \frac{1}{M} p(\mathbf{z}_i|\mathbf{s}_k). \qquad (17)$$

Thus, according to (12), we have

$$P(\text{correct}) \cong \frac{1}{M}\Delta x \Delta y \sum_{i=1}^{N^2} p(\mathbf{z}_i|\mathbf{s}_k) = \frac{1}{M}\Delta x \Delta y \sum_{i=1}^{N^2} \max_j \{p(\mathbf{z}_i|\mathbf{s}_j)\}. \qquad (18)$$

There are $N^2 \times M$ conditional probabilities $p(\mathbf{z}_i|\mathbf{s}_j)$ that need to be evaluated to compute (18). We can write a matrix equation for $p(\mathbf{z}_i|\mathbf{s}_j)$ as follows:

$$\begin{bmatrix} p(\mathbf{z}_1) \\ \vdots \\ p(\mathbf{z}_{N^2}) \end{bmatrix} = \begin{bmatrix} p(\mathbf{z}_1|\mathbf{s}_1) & \cdots & p(\mathbf{z}_1|\mathbf{s}_M) \\ \vdots & \ddots & \vdots \\ p(\mathbf{z}_{N^2}|\mathbf{s}_1) & \cdots & p(\mathbf{z}_{N^2}|\mathbf{s}_M) \end{bmatrix} \cdot \begin{bmatrix} p(\mathbf{s}_1) \\ \vdots \\ p(\mathbf{s}_M) \end{bmatrix}. \qquad (19)$$

The probability of making a correct decision on the received signal is the sum of the largest elements in each row of the matrix $p(\mathbf{z}|\mathbf{s})$. The larger we make N, the more accurately we can compute (18). In practice, we evaluate the probability of symbol error for different values of N to see at what point $P(\text{error}) = 1 - P(\text{correct})$ changes negligibly with an increase in N. When stable results are obtained, the computation terminates and the results are declared accurate.

C. Computing the Probability of Bit Error

Computation of the BER requires knowledge of the relationship between symbol errors and bit errors. When a receiver is operating at a low probability of error (symbol or bit), it is assumed that symbol errors are made in favor of the nearest neighbor(s) of the transmitted signal [11]. A constellation that has a Gray code mapping has the benefit that a symbol error results in only a single bit error. If the constellation encodes b bits per symbol, the probability of bit error is approximately $1/b$ times the probability of symbol error.

The computation of BER in the presence of phase error is complicated by the fact that ML detection uses a non-Euclidean metric for computing the conditional probability, as shown in (7). This changes the "distance" relationship between signal points as well as the number of nearest neighbors that each signal point possesses in the constellation. If we assume phase error to be small, such that AWGN is the dominant source of signal corruption, the distance metric will be roughly Euclidean. This allows us to find the nearest neighbors for each signal point using a Euclidean distance metric as is usually done when phase errors are absent. For 16-QAM, Gray code mapping leads to

$$P_{\mathrm{b}}^{16-\mathrm{QAM}} = \frac{1}{4} P_{\mathrm{s}}^{16-\mathrm{QAM}} \qquad (20)$$

where P_{b} and P_{s} refer to BER and symbol error ratio, respectively.

For 8-QAM, a Gray code mapping does not exist because the signal points in the inner circle have four nearest neighbors.

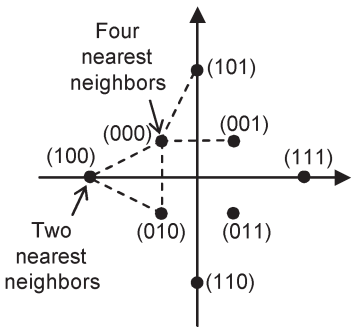

Fig. 10. Optimal bit-to-symbol mapping for 8-QAM.

These are shown by dashed lines in Fig. 10, where we have also labeled the symbols by an optimal bit-to-symbol mapping. We observe that the outer points have two nearest neighbors each, so when one of these points is transmitted and a symbol error is made, there is 50% chance of making one bit error and 50% chance of making two bit errors. Likewise, when an inner point is transmitted, a symbol error has 75% chance of making one bit error and 25% chance of making two bit errors. The probability of bit error for 8-QAM is then

$$P_{\rm b}^{8-{\rm QAM}} = \frac{1}{3}\left[\frac{1}{2}\left(\frac{1}{2}\times 1+\frac{1}{2}\times 2\right) + \frac{1}{2}\left(\frac{3}{4}\times 1+\frac{1}{4}\times 2\right)\right]\cdot P_{\rm s}^{8-{\rm QAM}} = \frac{1.375}{3}P_{\rm s}^{8-{\rm QAM}}. \quad (21)$$

Equations (20) and (21) are accurate when the effect of phase errors is small compared to AWGN and when the system is already operating at a low probability of error.

D. Results

We used the numerical integration technique in (18) to compute the probability of symbol error. Assuming that the probability of error is low and that the effects of phase errors are small compared to AWGN,[5] we invoked (20) and (21) to compute the probability of bit error. The BER results for 8- and 16-QAM are shown in Fig. 11(a) and (b) versus SNR per bit at the detector for various values of phase-error standard deviations. The analytical solution for the BER of 16-QAM in the absence of phase errors

$$P_{\rm b}^{16-{\rm QAM}} = \frac{3}{8}{\rm erfc}\left(\sqrt{\frac{2{\rm SNR}_{\rm bit}}{5}}\right) \quad (22)$$

is also shown in Fig. 11(b). We observe that the analytical curve lies on top of our numerical result for $\sigma_\varepsilon = 0^\circ$. Table I shows the maximum phase-error standard deviations that 8- and 16-QAM can tolerate for the power penalty to remain within 0.5 and 1.0 dB, respectively, at BERs of 10^{-8}, 10^{-9}, and 10^{-10}.

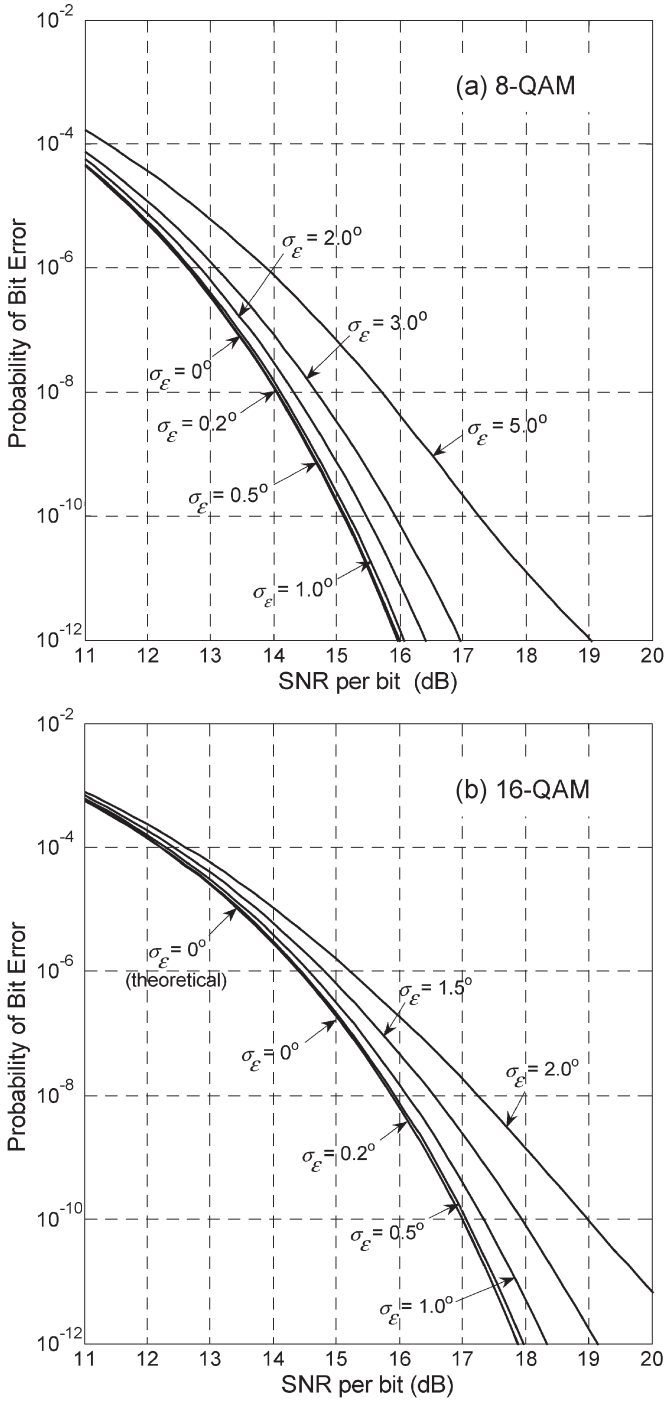

Fig. 11. BER versus SNR per bit for (a) 8-QAM and (b) 16-QAM.

While 8-QAM can tolerate between 2° and 4° phase-error standard deviation without significant loss of receiver sensitivity, 16-QAM can only tolerate 1–2°. Thus, in going from 3 to 4 bits per symbol, the requirement on the PLL becomes markedly more stringent. In the absence of phase errors, 8- and 16-QAM require 14.60 and 16.46 dB SNR per bit to achieve a BER of 10^{-9}.

[5]These conditions are required for the assumed relationship between symbol errors and bit errors to hold in (20) and (21). It can be shown that these conditions are satisfied inside the domains of Fig. 11 by separately considering the probability of a symbol error due to either AWGN or phase noise alone.

TABLE I
POWER PENALTIES FOR (A) 8-QAM AND (B) 16-QAM

Target BER	SNR per bit required with no phase error ($\sigma_\varepsilon = 0°$)	Maximum Tolerable Phase Error Standard Deviation (σ_ε)	
		0.5 dB power penalty	1.0 dB power penalty
10^{-8}	14.03 dB	2.64°	3.76°
10^{-9}	14.60 dB	2.48°	3.53°
10^{-10}	15.10 dB	2.33°	3.32°

Target BER	SNR per bit required with no phase error ($\sigma_\varepsilon = 0°$)	Maximum Tolerable Phase Error Standard Deviation (σ_ε)	
		0.5 dB power penalty	1.0 dB power penalty
10^{-8}	15.87 dB	1.34°	1.77°
10^{-9}	16.46 dB	1.24°	1.63°
10^{-10}	16.98 dB	1.15°	1.53°

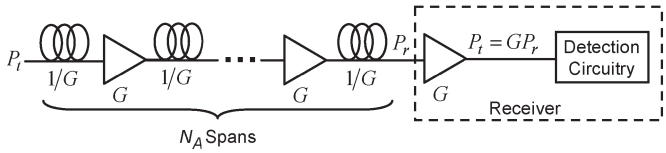

Fig. 12. Long-haul system with N_A spans of fiber employing inline OAs.

IV. CARRIER SYNCHRONIZATION FOR 8- AND 16-QAM

Now that the maximum allowable phase-error standard deviations have been computed for 8- and 16-QAM, we turn our attention to carrier synchronization. Our goal is to determine what PLL parameters are required to ensure that the standard deviation of PLL tracking error is smaller than the values computed in Table I. Our analysis shall focus on long-haul transmission where lumped optical amplifiers (OAs) are used periodically to reamplify the signal attenuated by fiber losses. A model long-haul system is shown in Fig. 12. Each of the N_A spans is of length L. The output power at the end of each span is $1/G$ times the input power. We insert an OA with gain G at the output of each span to compensate for transmission losses. The receiver consists of a preamplifier followed by the detection circuit that includes the PLL. The system's SNR is established at the output of the preamplifier. The SNR per symbol is the signal power divided by the LO-spontaneous noise occupying the band of the signal. In a long-haul system, LO-spontaneous beat noise is the dominant source of AWGN at the receiver if the LO laser is operated at sufficiently high power.

A. PLL Models and Analysis

Either heterodyne or homodyne detection may be used to coherently detect 8- and 16-QAM. Canonical receiver structures are shown in Fig. 13. The sensitivities of the heterodyne and homodyne receivers are identical when LO shot noise dominates [7]. In the LO-spontaneous beat-noise-limited case, a heterodyne receiver can achieve the same performance as a homodyne receiver, provided the heterodyne receiver employs image rejection or narrowband optical filtering to reject the spontaneous emission noise in the image band [12]. For concreteness, we analyze a heterodyne receiver that uses narrowband optical filtering of spontaneous emission. We note that a homodyne or heterodyne receiver can employ either an optical PLL or an electrical PLL [7] for carrier synchronization. Our analysis is valid whether an optical or electrical PLL is used. It is also possible to implement the PLL digitally by means of digital signal processing [13], [14]. Such an all-digital implementation is mathematically equivalent to the analog implementation analyzed here, provided that matched filtering is used and that the sampling rate (equal to the symbol rate) is much greater than the PLL natural frequency. Since the latter condition is easily satisfied, our analysis is applicable to all-digital PLLs. We assume perfect matching between the signal and LO states of polarization.

In a laser that has a Lorentzian lineshape, whose full-width at half-maximum (FWHM) linewidth is $\Delta\nu$, the instantaneous frequency is a white Gaussian random process. The phase noise $\varphi(\mathrm{t})$ is a Wiener process, whose power spectral density (PSD) is

$$S_{\varphi\varphi}(\omega) = \frac{\Delta\nu}{\omega^2}\ \mathrm{rad}^2/\mathrm{Hz}. \tag{23}$$

The analysis in this section is an extension of [7], which treated carrier synchronization for QPSK.

Consider the signals at the inputs of the 180° hybrid of the heterodyne receiver, as shown in Fig. 14. The fields of the received signal and LO laser are

$$E_{\mathrm{in}}(t) = T \cdot \sum_k \sqrt{P_k} b(t - kT) \cdot \mathrm{e}^{j(\omega_{\mathrm{c}} t + \theta_k + \varphi_{\mathrm{in}}(t))} + E_{\mathrm{sp}}(t) \tag{24}$$

$$E_{\mathrm{LO}}(t) = \sqrt{P_{\mathrm{LO}}} \cdot \mathrm{e}^{j(\omega_{\mathrm{LO}} t + \psi(t) + \varphi_{\mathrm{LO}}(t))}. \tag{25}$$

In (24) and (25), we have the following.

1) $s_k = \sqrt{P_k}\mathrm{e}^{j\theta_k}$ is a complex number denoting the kth transmitted symbol. The real and imaginary parts of s_k are its I and Q components. In this section, we use an amplitude-phase notation for the received symbols, as opposed to Section III where we used a vector notation $\mathbf{s}_k$. Since the inline OAs in the system are assumed to have completely compensated for fiber attenuation, P_k is equal to the transmitted power of the kth symbol.
2) $b(t)$ is the pulse shape. In this paper, we assume rectangular pulse shaping [nonreturn to zero (NRZ)], so $b(t)$ takes the form

$$b(t) = \begin{cases} \frac{1}{T}, & t \in (0, T] \\ 0, & t \notin (0, T] \end{cases} \tag{26}$$

where T is the symbol period. It should be noted that our analytical results to follow are also valid for any general pulse shape (including RZ), provided that the low-pass filters (LPFs) in Fig. 13 are matched to $b(t)$.
3) $E_{\mathrm{LO}} = \sqrt{P_{\mathrm{LO}}}\mathrm{e}^{j\psi(t)}$ is a phasor representing the amplitude and phase of the LO.
4) $E_{\mathrm{sp}}(t)$ is a phasor representing spontaneous emission.
5) $\varphi_{\mathrm{in}}(t)$ and $\varphi_{\mathrm{LO}}(t)$ are the phase noises of the source and LO lasers.

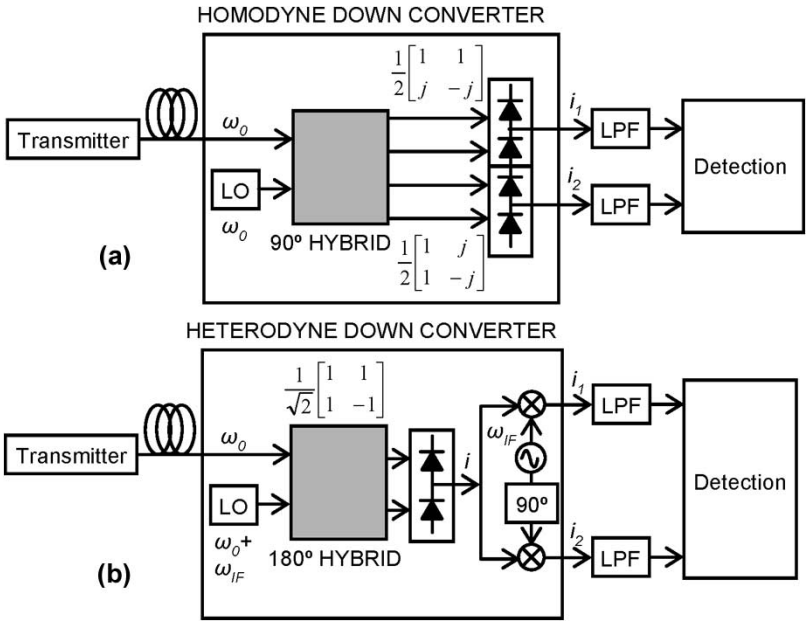

Fig. 13. Coherent detection using a (a) homodyne receiver and (b) heterodyne receiver.

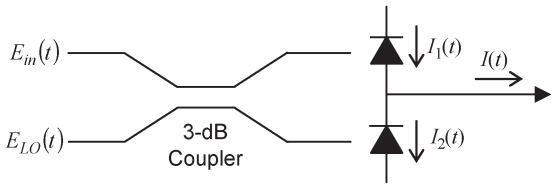

Fig. 14. 180° hybrid of the heterodyne receiver.

6) $\omega_{\rm c}$ and $\omega_{\rm LO}$ are the frequencies of the signal and the LO. In a heterodyne system, their algebraic difference $\omega_{\rm c} - \omega_{\rm LO}$ is the intermediate frequency (IF) $\omega_{\rm IF}$.
7) Let the responsivity of each photodiode in Fig. 14 be R. It can be shown that the balanced photodetector produces an output $I(t)$ of the form

$$I(t) = 2R\sum_k T\sqrt{P_k P_{\rm LO}}\,b(t-kT) \cdot \sin\left(\omega_{\rm IF}t + \theta_k + \varphi(t) - \psi(t)\right) - 2R\sqrt{P_{\rm LO}} \cdot {\rm Im}\left\{{\rm e}^{j\phi_{\rm LO}(t)}E^*_{\rm sp}(t)\right\} + n_{\rm sh}(t) \quad (27)$$

where the two phase noises have been combined into a single variable $\varphi(t) = \varphi_{\rm in}(t) - \varphi_{\rm LO}(t)$. We have thus attributed phase noise entirely to the signal laser; the LO laser is assumed to be phase-noise free. The spectrum of ${\rm e}^{j\varphi(t)}$ is Lorentzian, whose linewidth $\Delta\nu$ is equal to the sum of the linewidths of the transmitter and LO lasers, which we refer to as the beat linewidth. The first term of (27) is the desired signal modulated by the IF frequency; the second term is LO-spontaneous beat noise; and the third term $n_{\rm sh}(t) = n_{{\rm sh},1}(t) - n_{{\rm sh},2}(t)$ is shot noise produced by the two photodiodes. The power of the desired signal is

$$P_{\rm sig} = 2R^2 G P_{\rm r} P_{\rm LO}. \quad (28)$$

The one-sided PSDs of the noise terms are

$$S_{\rm LO-spont}(f) = 2R^2 P_{\rm LO} S_{\rm sp}(f) \quad (29)$$

$$S_{\rm shot}(f) = 2qRP_{\rm LO} \quad (30)$$

where $S_{\rm sp}(f)$ is the spontaneous noise PSD given by [15]

$$S_{\rm sp}(f) = N_A n_{\rm sp}(G-1)h\nu. \quad (31)$$

$n_{\rm sp}$ is the spontaneous emission factor and $h\nu$ is the photon energy. The signal power in (28) divided by the LO-spontaneous noise power in the same band as the one passing through a filter matched to the signal gives the SNR per symbol at the detector.

Let $n(t) = 2R\sqrt{P_{\rm LO}}{\rm Im}\{{\rm e}^{-j\phi_{\rm LO}(t)}E_{\rm sp}(t)\} + n_{\rm sh}(t)$ be the total noise in the system. Its PSD $S_{nn}(f)$ is the sum of (29) and (30). In the limit that one noise source dominates, as in a long-haul system (considered by this paper) or a back-to-back measurement, we may ignore one of the terms.

A suitable PLL that can be used to recover the phase of the signal is shown in Fig. 15. The photocurrent $I(t)$ is mixed with the I and Q phases of an IF electrical LO. I and Q mixer outputs are passed through baseband matched filters $B(s)$. The functions performed by F and G depend on the type of loop implemented, and are discussed in more detail below. The outputs of the F and G elements in the I and Q branches are cross multiplied and subtracted, and fed to a loop filter $F(s)$, which is proportional plus integral in a second-order PLL. The loop filter output is used to control the IF phase. In an optical PLL, the IF phase is controlled by frequency modulation of the LO laser, while in an electrical PLL, the IF phase is controlled by frequency modulation of the IF electrical LO.

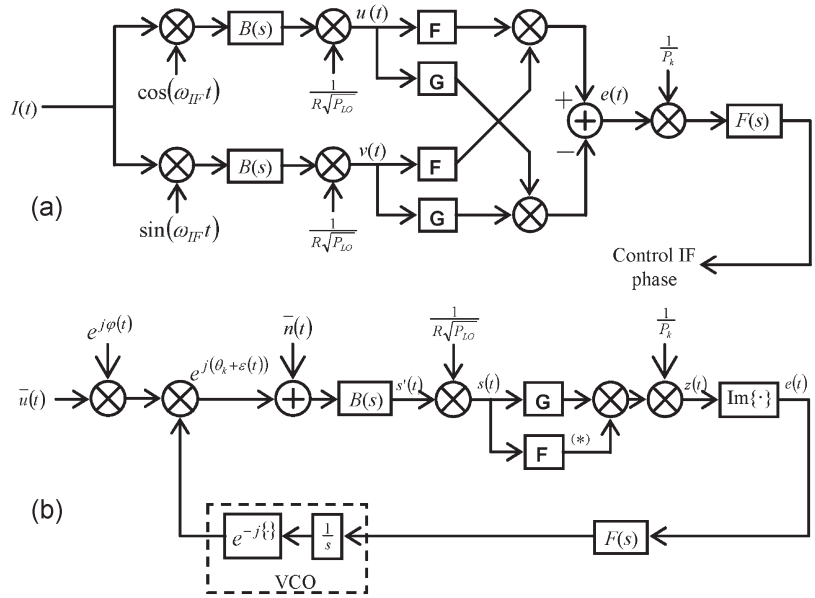

Fig. 15. Fig. 15. Decision-directed loop. (a) Implementation. (b) Complex model.

TABLE II
TRANSFER CHARACTERISTICS F AND G

Loop Type	F{x(t)}	G{x(t)}
Discrete-Time Decision-Directed	Decision: [x(kT)]	Sample & Hold: x(kT)
Analog Decision Directed	Decision: [x(kT)]	Delay: x(t-T/2)

In Table II, we show the functional forms of F and G in a digital (discrete-time) decision-directed loop and in an analog decision-directed loop. Either loop may be used to phase lock 8- and 16-QAM. Compared to carrier synchronization for QPSK [7], fewer loop options are available: The Costas loop cannot be used because the I and Q components have more than two levels; similarly, the fourth-power loop cannot be used because the constellations have more than four phases. Although an eighth-power loop is theoretically possible for phase-locking 8-QAM, the bandwidth requirement on its components makes it prohibitive to implement. As the analog decision-directed loop is an approximation of the digital decision-directed loop [7], we shall analyze the latter only. At low BERs, the performance of the analog decision-directed loop will approach that of the digital decision-directed loop.

We can simplify our analysis by writing a mathematically equivalent complex model of the PLL. This is shown in Fig. 15(b). The functions F and G in this model operate on complex-valued inputs, performing independent actions on the real and imaginary components:

$$\mathsf{F}\{u+jv\} = \mathsf{F}\{u\} + j\mathsf{F}\{v\}. \tag{32}$$

The choices for F and G are now apparent in this model, as illustrated in Fig. 16. Let s_k be the signal that appears at the output of the matched filter at the kth sampling instant. When phase error is the only signal corruption, and its value is small,

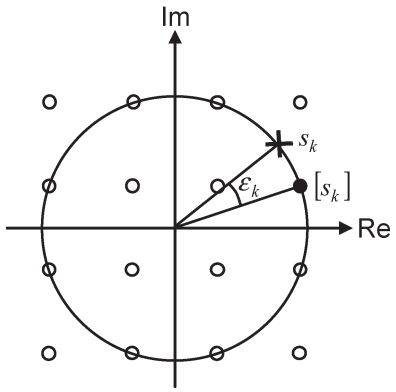

Fig. 16. Received signal and the decision.

the decision device correctly recovers the transmitted signal. The product of the sample-and-hold output and the complex conjugate of the decision $[s_k]$ gives

$$s_k \cdot [s_k]^* \approx |s_k|^2 \mathrm{e}^{j\varepsilon_k}. \tag{33}$$

If the value of the phase error ε_k is small, its value can be estimated by computing

$$\mathrm{Im}\left[\frac{s_k}{|s_k|^2} \cdot [s_k]^*\right] = \sin\varepsilon_k \approx \varepsilon_k. \tag{34}$$

Using the photocurrent $I(t)$ computed in (27), the signal $s'(t)$ that appears at the matched filter output in Fig. 15(b) is

$$s'(t) = b(t) \otimes \left\{ R\sum_k T\sqrt{P_k P_{\mathrm{LO}}}\, b(t-kT) \cdot \mathrm{e}^{j(\theta_k+\varphi(t)-\psi(t))} + \overline{n}(t) \right\} \tag{35}$$

where $\overline{n}(t) = jn(t)\mathrm{e}^{-j\omega_{\mathrm{IF}}t}$ has the same PSD as $S_{nn}(f)$ given by (29) or (30). Since $b(t)$ is a rectangular function, the signal $s'(t)$ sampled at $t = kT$ has a value

$$s'(kT) = R\sqrt{P_{k-1}P_{\mathrm{LO}}}\cdot \mathrm{e}^{j\theta_{k-1}} \cdot \left(\frac{1}{T}\int_{(k-1)T}^{kT} \mathrm{e}^{j\varepsilon(t')}dt'\right) + \frac{1}{T}\int_{(k-1)T}^{kT} \overline{n}(t')dt'. \quad (36)$$

In general, the phase error $\varepsilon(t) = \varphi(t) - \psi(t)$ changes very little over a symbol period.[6] We can make the approximation that $\varepsilon(t) \approx \varepsilon_{k-1}$ for the entire duration $(k-1)T < t \leq kT$. Defining a new variable

$$\overline{n_k} = \frac{1}{T}\int_{kT}^{(k+1)T} \overline{n}(t')dt' \quad (37)$$

we have the following after dividing (36) by $R\sqrt{P_{\mathrm{LO}}}$:

$$s(kT) = \mathrm{e}^{j\varepsilon_{k-1}}\cdot s_{k-1} + \frac{1}{R\sqrt{P_{\mathrm{LO}}}}\overline{n_{k-1}}. \quad (38)$$

The first term of (38) is the receiver's estimate of the transmitted symbol corrupted by phase error. The second term is corruption by AWGN. We observe that (38) is the same form as (2), which we assumed in deriving the results in Section III. If the phase error and AWGN are small, the decision device always makes a correct decision

$$\mathsf{F}\left[s(kT)\right] = s_{k-1} = \sqrt{P_{k-1}}\mathrm{e}^{j\theta_{k-1}} \quad \text{for } kT < t \leq (k+1)T. \quad (39)$$

The output of the sample-and-hold device is

$$\mathsf{G}\left[s(kT)\right] = s(kT), \qquad \text{for } kT < t \leq (k+1)T \quad (40)$$

and the product of (40) and the complex conjugate of (39) gives

$$\mathsf{F}\left[s(kT)\right]\mathsf{G}\left[s(kT)\right]^* = \mathrm{e}^{j\varepsilon_{k-1}}P_{k-1} + \frac{\sqrt{P_{k-1}}}{R\sqrt{P_{\mathrm{LO}}}}\overline{n_{k-1}}\mathrm{e}^{-j\theta_{k-1}} \quad \text{for } kT < t \leq (k+1)T. \quad (41)$$

Dividing (41) by the signal power P_{k-1} and taking its imaginary component yields the receiver's estimate of the phase error $e(t)$:

$$e(t) = \sin(\varepsilon_{k-1}) + \frac{1}{R\sqrt{P_{\mathrm{LO}}P_{k-1}}}\mathrm{Im}\left\{\mathrm{e}^{-j\theta_{k-1}}\cdot\overline{n_{k-1}}\right\} \approx \varepsilon_{k-1} + w(t-T), \qquad \text{for } kT < t \leq (k+1)T \quad (42)$$

[6]The fast component of the tracking error $\varepsilon(t)$ is due to phase noise. Lasers suitable for coherent detection have linewidths of the order of tens of kilohertz, while the symbol rate is of the order of tens of gigahertz.

TABLE III
CONSTELLATION PENALTIES FOR VARIOUS WELL-KNOWN SIGNALING SCHEMES

Signaling Scheme	Constellation Penalty η_c
2-PSK (BPSK)	1
4-PSK (QPSK)	1
8-PSK	1
8-QAM	1.5
16-QAM	1.889
64-QAM	2.685

where

$$w(t) = \frac{1}{R\sqrt{P_{\mathrm{LO}}P_k}}\mathrm{Im}\left\{\mathrm{e}^{-j\theta_k}\cdot\overline{n_k}\right\} \quad \text{for } kT < t \leq (k+1)T. \quad (43)$$

$w(t)$ is the effect of AWGN on our estimate of the true phase error, which is linear in the phase error of the first order.

The key difference between the result (43) and the one in [7] is that (43) is valid for an arbitrary signaling format employing amplitude and/or phase modulation. If constant-envelope modulation is used, our result reduces to [7]. The statistics of $w(t)$ over the duration $kT < t \leq (k+1)T$ depend on the power of the kth transmitted symbol. Provided the transmitted symbols are a stationary process, it can be shown that $w(t)$ is stationary and is well approximated by a white-noise process with spectrum given by (see Appendix)

$$S_{WW}(\omega) = \frac{\eta T_{\mathrm{b}}}{\overline{n_{\mathrm{b}}}} \quad (44)$$

where $T_{\mathrm{b}} = T/\log_2(M)$ is the bit period and $\overline{n_{\mathrm{b}}} = P_{\mathrm{r}}T_{\mathrm{b}}/h\nu$ is the mean number of photons per bit incident on the receiver shown in Fig. 12. The minimum value of $\overline{n_{\mathrm{b}}}$ that achieves satisfactory performance is the receiver sensitivity. η depends on which source of AWGN is dominant. For an LO-spontaneous beat-noise-limited system, we have

$$\eta = \frac{N_A n_{\mathrm{sp}}}{2}\left(\frac{G-1}{G}\right)\cdot E[P_k]E\left[\frac{1}{P_k}\right] \quad (45)$$

where $E[P_k]$ is the mean transmitted power, and $E[1/P_k]$ is the mean inverse of the transmitted power. These expectations are taken over the signal points in the signal constellation. We can define the product of these quantities as

$$\eta_{\mathrm{c}} = E[P_k]E\left[\frac{1}{P_k}\right]. \quad (46)$$

η_{c} is a unitless "constellation penalty," whose value depends only on the arrangement of signal points in the constellation. Table III shows the value of η_{c} for some commonly used constellations including 8- and 16-QAM. All constant-envelope (PSK) schemes have unit constellation penalty.

The key utility of our complex-model analysis is that it reduces the receiver's phase-error estimate to a linear sum of the actual phase error and $w(t)$. We can now analyze the PLL using the linear model shown in Fig. 17. In this model, the input signal $\dot{\varphi}(t)$ represents frequency noise. For a laser with

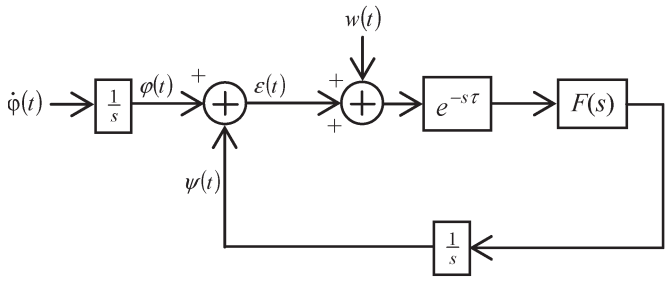

Fig. 17. Linearized model of PLL.

a Lorentzian lineshape, $\dot{\varphi}(t)$ is a white Gaussian process. Its integral represents phase noise, whose PSD is given by (23). Phase error is the difference between phase noise $\varphi(t)$, which we have attributed entirely to the signal laser, and the control phase $\psi(t)$ of the LO. The feedback path of the PLL includes a loop filter $F(s)$ and a delay element that lumps together all delays in the PLL arising from signal propagation and component group delays. The delay time is given by τ. For a loop filter of the form

$$F(s) = 2\zeta\omega_{\rm n} + \frac{\omega_{\rm n}^2}{s} \tag{47}$$

the PLL has a second-order transfer function, whose natural frequency and damping factor are given by $\omega_{\rm n}$ and ζ, respectively. Grant *et al.* [16] showed that the variance of phase error is given by

$$\sigma_\varepsilon^2 = \frac{\pi\Delta\nu}{2\zeta\omega_{\rm n}}\Gamma_{\rm PN}(\omega_{\rm n}\tau) + \frac{(1+4\zeta^2)\omega_{\rm n}}{4\zeta}\frac{\eta T_{\rm b}}{\overline{n_{\rm b}}}\Gamma_{\rm AWGN}(\omega_{\rm n}\tau) \tag{48}$$

where

$$\Gamma_{\rm PN}(\omega_{\rm n}\tau) = \frac{2\zeta\omega_{\rm n}}{\pi}\cdot\int_{-\infty}^{\infty}\left|j\omega + {\rm e}^{-j\omega\tau}F(\omega)\right|^{-2}d\omega \tag{49}$$

$$\Gamma_{\rm AWGN}(\omega_{\rm n}\tau) = \frac{2\zeta}{\pi(1+4\zeta^2)\omega_{\rm n}} \cdot\int_{-\infty}^{\infty}\left|\frac{F(\omega)}{j\omega + {\rm e}^{-j\omega\tau}F(\omega)}\right|^2 d\omega. \tag{50}$$

The subscripts PN and AWGN denote that these terms are contributions by phase noise and AWGN, respectively. We observe that they add independently and linearly. $\Delta\nu$ is the beat linewidth between the signal and LO lasers and η can be computed using (45). Equations (48)–(50) allow a system designer to calculate the phase-error standard deviation for any coherent system. The value of σ_ε computed should be compared with Table I to determine whether the system achieves satisfactory performance.

B. Example

In Section III-D, the values of SNR per bit required to achieve a BER of 10^{-9} for 8- and 16-QAM were computed to be 14.60 and 16.46 dB, respectively. Since matched filtering is used by the receiver, the LO-spontaneous beat-noise variance is

$$\sigma_{\rm LO-spont}^2 = \int_{-\infty}^{\infty} S_{\rm LO-spont}(f)\,|B(f)|^2\,df = \frac{2R^2P_{\rm LO}S_{\rm sp}(f)}{T}. \tag{51}$$

Noting that (28) divided by (51) gives the SNR per symbol at the detector, the SNR per bit is

$$\text{SNR}_{\rm bit} = \frac{1}{\log_2 M}\frac{GP_{\rm r}T}{(G-1)N_A n_{\rm sp}h\nu} = \left(\frac{G}{G-1}\right)\frac{\overline{n_{\rm b}}}{N_A n_{\rm sp}}. \tag{52}$$

Assume a system that has a single OA ($N_A = 1$) with high gain ($G \to \infty$), whose spontaneous emission factor is $n_{\rm sp} = 1$. We have $\text{SNR}_{\rm bit} \approx \overline{n_{\rm b}}$. The receiver sensitivities for 8- and 16-QAM are 28.8 and 44.3 photons per bit, respectively.

As in [7], we assume our system operates 1-dB away from the fundamental AWGN limit. Of this 1 dB, we allot 0.5 dB to system margin and the remaining 0.5 dB to phase-error penalty. Let the gain of the OA be 20 dB. For the PLL, we assume a second-order loop with a damping factor $\zeta = 1/\sqrt{2}$. This is a common choice as it balances the requirements of a quick transient response and a low steady-state variance. We assume a 10-Gb/s system so the bit period is $T_{\rm b} = 100$ ps. We assume that the transmitter and LO lasers each have linewidths of 15 kHz, so the beat linewidth is $\Delta\nu = 30$ kHz.[7]

Fig. 18 shows the phase-error standard deviation in such a system, as a function of the PLL's natural frequency at various feedback delays. To ensure that the phase-error penalty remains below 0.5 dB, we note in Table I that σ_ε should be less than 2.48° for 8-QAM and 1.24° for 16-QAM. Provided the loop delay is less than $34.5T_{\rm b}$, 8-QAM is able to satisfy this requirement, corresponding to 3.45 ns. For 16-QAM, there is no value τ for which the 0.5-dB margin is satisfied. Consequently, a system implementing 16-QAM must either tolerate a larger phase penalty or use lasers having smaller linewidths.

We observe in Fig. 18 that for every value of τ, there exists an optimal PLL natural frequency that minimizes the phase-error standard deviation. A plot of the optimal PLL natural frequency $\omega_{\rm n}$ versus delay is shown in Fig. 19. As the delay increases, the optimal loop frequency decreases like the magnitude response of a first-order LPF [7].

[7]This number was chosen based on measured linewidths of microelectromechanical system (MEMS)-based external-cavity tunable lasers [17].

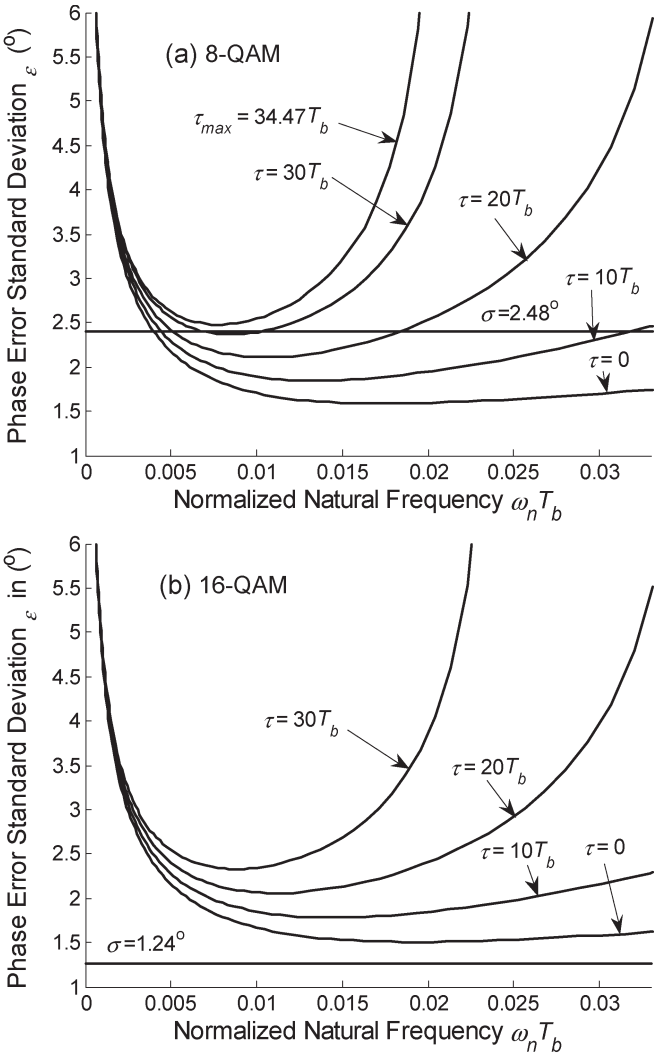

Fig. 18. Phase-error standard deviation versus PLL natural frequency at various feedback delays for (a) 8-QAM and (b) 16-QAM.

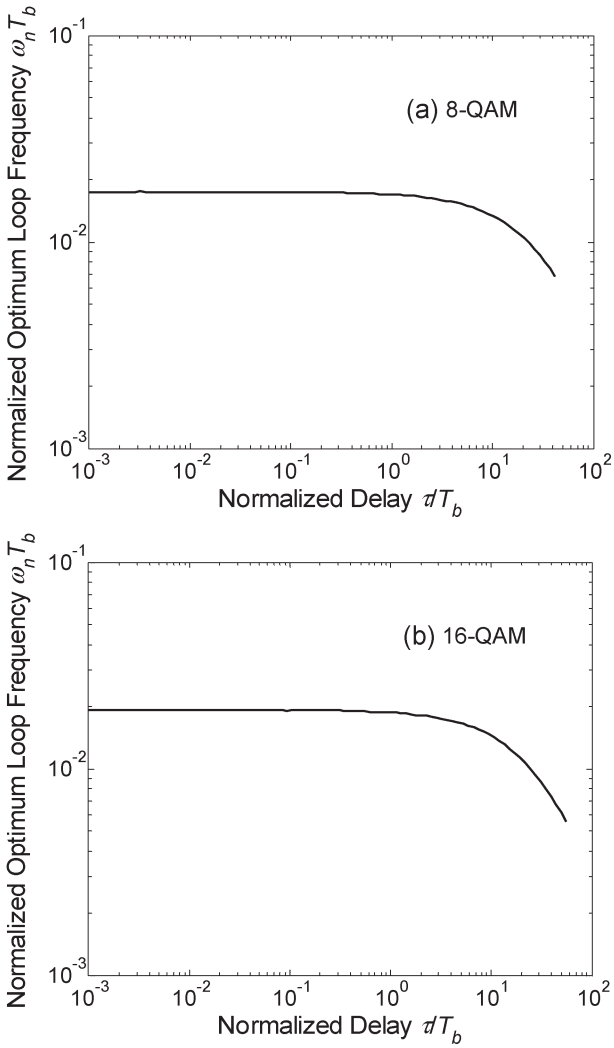

Fig. 19. Optimal PLL natural frequency versus feedback delay for (a) 8-QAM and (b) 16-QAM.

Finally, in Fig. 20, we plot the maximum allowable laser linewidth as a function of feedback delay assuming the PLL always uses an optimized value for its loop natural frequency. As τ increases, the system is less able to track the instantaneous phase noise, so the tolerable linewidth is decreased. The rate of this decrease is also well modeled by the magnitude response of first-order LPF [7]. When feedback delay is 0, we observe that for 8- and 16-QAM, the beat linewidths may not exceed 180 and 13.7 kHz, respectively.

V. Conclusion

We presented a detailed analysis of the performance of 8- and 16-quadrature amplitude modulation (QAM) in the presence of phase noise and additive white Gaussian noise (AWGN). We numerically evaluated bit-error-rate (BER) performance for each transmission scheme assuming the receiver employs maximum-likelihood (ML) detection. These results may be used to determine the maximum phase-error standard deviations that 8- and 16-QAM can tolerate while satisfying power-penalty constraints at a given BER. We analyzed two PLLs that are suitable for phase synchronization of 8- and 16-QAM: the discrete-time decision-directed loop and the analog decision-directed loop. We found that the contribution of AWGN to the phase-error variance is proportional to $\eta T_{\rm b}/\overline{n_{\rm b}}$, where $\overline{n_{\rm b}}/T_{\rm b}$ is the photon arrival rate at the receiver. We computed η for LO-spontaneous beat-noise-limited systems and found it was proportional to the constellation penalty, which is a function of the arrangement of signal points in the constellation.

Assuming the use of a proportional-plus-integral loop filter, we computed the maximum tolerable laser linewidth and optimal PLL natural frequency as functions of delay in the PLL feedback path. We found that at a bit rate of 10 Gb/s, lasers having a beat linewidth $\Delta\nu = 30$ kHz can be used for 8-QAM but not for 16-QAM, even with zero delay.

Appendix

In Section IV-A, we found that the PLL's estimate of the phase error is corrupted by an AWGN term $w(t)$. It was claimed

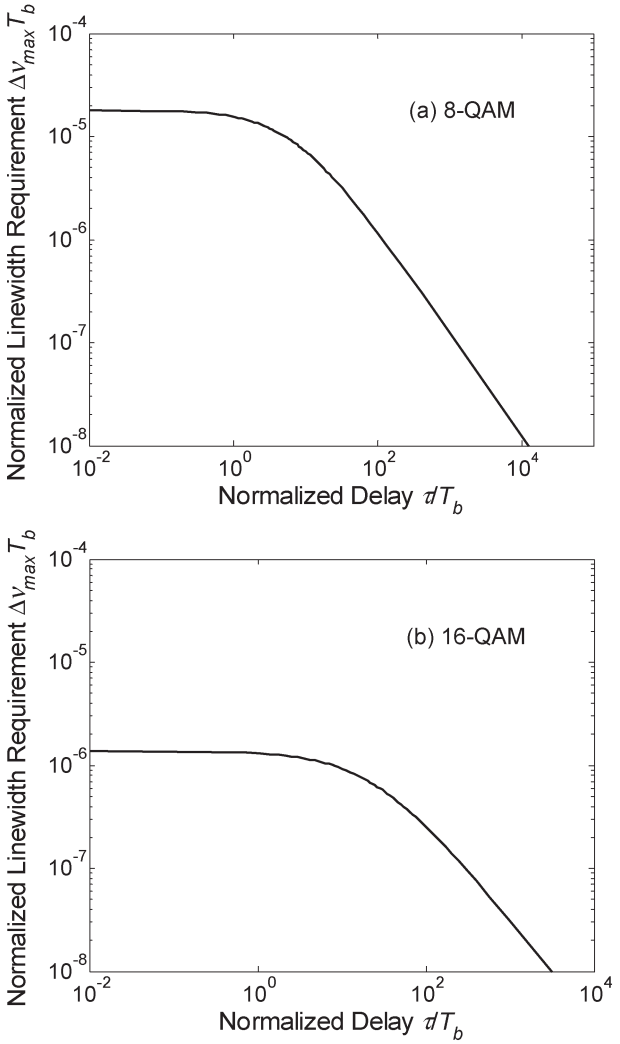

Fig. 20. Maximum tolerable beat linewidth versus feedback delay for 0.5 dB phase-error penalty: (a) 8-QAM and (b) 16-QAM.

that $w(t)$ is stationary, whose PSD is well approximated by a white noise spectrum. In this Appendix, we justify these assumptions and derive the form of the scaling factor η for both shot-noise-limited and LO-spontaneous beat-noise-limited systems.

We first consider a signal $w'(t)$ that is related to $w(t)$ by dropping the $\mathrm{Im}\{\cdot\}$ in (43). Substituting the definition of $\overline{n_k}$ provided in (37), we have

$$w'(t) = \frac{\mathrm{e}^{-j\theta_k}}{RT\sqrt{P_{\mathrm{LO}}P_k}} \int_{kT}^{(k+1)T} n(t')dt' \quad \text{for } kT < t \le (k+1)T. \tag{53}$$

We assume the transmitted symbols is a stationary process that is independent of white noise $\overline{n}(t)$. This is a valid assumption provided the LO power at the receiver is much higher than the signal power, as it results in $\overline{n}(t)$ being a zero-mean Gaussian random process whose variance depends only on the LO power, as per (29) or (30). $w'(t)$ has an autocorrelation function

$$\begin{aligned} &E\left[w'(t)w'^{*}(t-\tau)\right] \\ &\quad = \frac{1}{R^2T^2P_{\mathrm{LO}}} \\ &\qquad \cdot E\left[\left(\int_{kT}^{(k+1)T}\int_{jT}^{(j+1)T} n(t')n(u')dt'du'\right)\frac{\mathrm{e}^{-j(\theta_k-\theta_j)}}{\sqrt{P_kP_j}}\right] \\ &\quad = \frac{1}{R^2T^2P_{\mathrm{LO}}}\left(\int_{kT}^{(k+1)T}\int_{jT}^{(j+1)T} E\left[n(t')n(u')\right]dt'du'\right) \\ &\qquad \cdot E\left[\frac{\mathrm{e}^{-j(\theta_k-\theta_j)}}{\sqrt{P_kP_j}}\right] \end{aligned} \tag{54}$$

where $jT < t-\tau \le (j+1)T$ and $kT < t \le (k+1)T$.

Since $\overline{n}(t)$ has a white spectrum, its autocorrelation function is

$$E\left[\overline{n}(t')\overline{n}(u')\right] = \frac{S_{\overline{n}\,\overline{n}}}{2}\delta(t'-u') \tag{55}$$

where $S_{\overline{n}\,\overline{n}}$ is the one-sided PSD of $\overline{n}(t)$.

The double integral in (54) is 0 if t and $t-\tau$ are not in the same symbol period ($j \neq k$); and is $S_{\overline{n}\,\overline{n}}/2$ otherwise ($j = k$). The probability that two time samples drawn randomly separated by τ are both in the same symbol period is $1-|\tau|/T$. Hence

$$\begin{aligned} E\left[w'(t)w'^{*}(t-\tau)\right] &= \frac{S_{\overline{n}\,\overline{n}}}{2R^2T^2P_{\mathrm{LO}}}\Pr\{j=k\}\cdot E\left[\frac{1}{P_k}\right] \\ &= \frac{S_{\overline{n}\,\overline{n}}}{2R^2T^2P_{\mathrm{LO}}}\left(1-\frac{|\tau|}{T}\right)\cdot E\left[\frac{1}{P_k}\right] \end{aligned} \tag{56}$$

depends only on the time difference τ and not on t. Similarly, one can show that the process $w(t)$ has zero mean and is therefore independent of t. $w(t)$ is thus at least wide-sense stationary. However, $w(t)$ is Gaussian, so it is in fact a strict-sense-stationary process. Taking the Fourier transform of (56) yields the PSD of $w'(t)$

$$S_{W'W'}(\omega) = \frac{S_{\overline{n}\,\overline{n}}}{2R^2T^2P_{\mathrm{LO}}}E\left[\frac{1}{P_k}\right]\cdot\frac{\sin^2\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)^2}. \tag{57}$$

In practice, the bandwidth of the optimal loop $B_{\mathrm{L,opt}}$ is much smaller than the bandwidth of the signal: $B_{\mathrm{L,opt}}T \ll 1$ [7]. Consequently, $S_{W'W'}(\omega)$ is well modeled by a white Gaussian spectrum whose PSD is given by (57), evaluated at $\omega = 0$. Since the sinc function has value 1 at $\omega = 0$, and $w(t)$ is the imaginary part of $w'(t)$ and so has half the power, we have

$$S_{WW'}(\omega) \approx \frac{1}{2}S_{W'W'}(0) = \frac{S_{\overline{n}\,\overline{n}}}{4R^2P_{\mathrm{LO}}}\cdot E\left[\frac{1}{P_k}\right]. \tag{58}$$

For a shot-noise-limited system, the one-sided PSD $S_{\overline{n}\,\overline{n}}$ is given by (30)

$$S_{WW}^{\mathrm{SN}}(\omega) = \frac{q}{2RP_{\mathrm{r}}} \cdot P_{\mathrm{r}} E\left[\frac{1}{P_k}\right]. \tag{59}$$

We note that the received power P_{r} is equal to the average power of the detected symbols $E[P_k]$. We can further define $m = RP_{\mathrm{r}}T_{\mathrm{b}}/q$ as the mean number of detected photoelectrons per bit period. This leads to

$$S_{WW}^{\mathrm{SN}}(\omega) = \eta_{\mathrm{SN}} \frac{T_{\mathrm{b}}}{m} \tag{60}$$

where η_{SN} is

$$\eta_{\mathrm{SN}} = \frac{1}{2} E[P_k] E\left[\frac{1}{P_k}\right] = \frac{1}{2}\eta_{\mathrm{c}}. \tag{61}$$

For constant-envelop modulation formats such as QPSK, $\eta_{\mathrm{c}} = 1$, so (60) reduces to the result derived in [7].

For an LO-spontaneous beat-noise-limited system, $S_{\overline{n}\,\overline{n}}$ is given by (29). We have

$$\begin{aligned} S_{WW}^{\mathrm{LO-spont}}(\omega) &= \frac{N_A n_{\mathrm{sp}}(G-1)h\nu}{2} \cdot E\left[\frac{1}{P_k}\right] \\ &= \frac{N_A n_{\mathrm{sp}} h\nu}{2P_{\mathrm{r}}} \left(\frac{G-1}{G}\right) \cdot GP_{\mathrm{r}} E\left[\frac{1}{P_k}\right]. \end{aligned} \tag{62}$$

Owing to preamplification at the receiver, the average power of the detected symbols is $E[P_k] = GP_{\mathrm{r}}$. Noting that $\overline{n_{\mathrm{b}}} = P_{\mathrm{r}}T_{\mathrm{b}}/h\nu$ is the average number of received photons per bit at the input of the heterodyne receiver, we have

$$S_{WW}^{\mathrm{LO-spont}}(\omega) = \eta_{\mathrm{LO-spont}} \frac{T_{\mathrm{b}}}{\overline{n_{\mathrm{b}}}} \tag{63}$$

where the premultiplier in this case is

$$\begin{aligned} \eta_{\mathrm{LO-spont}} &= \frac{N_A n_{\mathrm{sp}}}{2} \left(\frac{G-1}{G}\right) \cdot E[P_k] E\left[\frac{1}{P_k}\right] \\ &= \frac{\eta_{\mathrm{c}} N_A n_{\mathrm{sp}}}{2} \left(\frac{G-1}{G}\right). \end{aligned} \tag{64}$$

We have thus proved (45).

ACKNOWLEDGMENT

The authors would like to acknowledge A. Kumar for his contribution to the early stages of this work.

REFERENCES

[1] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 259–272, Mar./Apr. 2004.

[2] R. A. Griffin and A. C. Carter, "Optical differential quadrature phase-shift keying (oDQPSK) for high-capacity optical transmission," in *Proc. Optical Fiber Communication Conf.*, Anaheim, CA, 2002, pp. 367–368.

[3] M. Ohm, "Optical 8-DPSK and receiver with direct detection and multilevel electrical signals," in *Proc. IEEE/LEOS Workshop Advanced Modulation Formats*, San Francisco, CA, Jul. 2004, pp. 45–46.

[4] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, no. 6841, pp. 1027–1030, Jun. 2001.

[5] 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, Dec. 1990.

[6] S. Norimatsu and K. Iwashita, "Linewidth requirements for optical synchronous detection systems with nonnegligible loop delay time," *J. Lightw. Technol.*, vol. 10, no. 3, pp. 341–349, Mar. 1992.

[7] J. R. Barry and J. M. Kahn, "Carrier synchronization for homodyne and heterodyne detection of optical quadriphase-shift keying," *J. Lightw. Technol.*, vol. 10, no. 12, pp. 1939–1951, Dec. 1992.

[8] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "On the selection of a two-dimensional signal constellation in the presence of phase jitter and Gaussian noise," *Bell Syst. Tech. J.*, vol. 52, pp. 927–965, Jul./Aug. 1973.

[9] K. P. Ho, "Generation of arbitrary quadrature-amplitude modulation signals using a single dual-drive modulator," in *Proc. IEEE/LEOS Workshop Advanced Modulation Formats*, San Francisco, CA, Jul. 2004, pp. 11–12.

[10] V. K. Prabhu, "PSK performance with imperfect carrier phase recovery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-12, no. 2, pp. 275–285, Mar. 1976.

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

[12] 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. Lightw. Technol.*, vol. 10, no. 5, pp. 660–671, May 1992.

[13] D.-S. Ly-Gagnon, K. Katoh, and K. Kikuchi, "Unrepeatered optical transmission of 20 Gb/s quadrature phase-shift keying signals over 210 km using homodyne phase-diversity receiver and digital signal processing," *Electron. Lett.*, vol. 41, no. 4, pp. 206–207, Feb. 2005.

[14] ——, "Coherent demodulation of differential 8-phase-shift keying with optical phase diversity and digital signal processing," in *Proc. 17th Annu. Meeting IEEE Lasers and Electro-Optics Society (LEOS)*, Rio Grande, PR, Nov. 2004, vol. 2, pp. 607–608.

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

[16] M. A. Grant, W. C. Michie, and M. J. Fletcher, "The performance of optical phase-locked loops in the presence of nonnegligible loop propagation delay," *J. Lightw. Technol.*, vol. LT-5, no. 4, pp. 592–597, Apr. 1987.

[17] E. Ip, J. M. Kahn, D. Anthon, and J. Hutchins, "Linewidth measurements of MEMS-based tunable lasers for phase-locking applications," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2029–2031, Oct. 2004.

Ezra Ip received the B.E. (Hons) degree in electrical and electronics engineering from the University of Canterbury, Christchurch, New Zealand, in 2002, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2004. He is currently working towards the Ph.D. degree in electrical engineering at Stanford University.

In 2002, he was a Research Engineer at Industrial Research Ltd., New Zealand. His research interests include single-mode optical fiber communications.

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 (U.C.), Berkeley, in 1981, 1983, and 1986, respectively.

From 1987 to 1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, in Holmdel, NJ. He demonstrated multigigabit-per-second 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 U.C. Berkeley, performing research on optical and wireless communications. In 2000, he helped found StrataLight Communications, where he served as Chief Scientist from 2000 to 2003. Since 2003, he has been a Professor of Electrical Engineering at Stanford University, Stanford, CA. His current research interests include single- and multimode optical fiber communications, free-space optical communications, and microelectromechanical systems (MEMS) for optical communications.

Prof. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993–2000, he served as a Technical Editor of *IEEE Personal Communications Magazine.*