Robust Multilevel Coherent Optical Systems With Linear Processing at the Receiver

Giulio Colavolpe, *Member, IEEE*, Tommaso Foggi, Enrico Forestieri, *Member, IEEE*, and Giancarlo Prati, *Fellow, IEEE*

Abstract—This paper investigates optical coherent systems based on polarization multiplexing and high-order modulations such as phase-shift keying (PSK) signals and quadrature amplitude modulations (QAM). It is shown that a simple linear receiver processing is sufficient to perfectly demultiplex the two transmitted streams and to perfectly compensate for group velocity dispersion (GVD) and polarization mode dispersion (PMD). In addition, in the presence of a strong phase noise of the lasers at the transmitter and receiver, a symbol-by-symbol detector with decision feedback is able to considerably improve the receiver robustness with a limited complexity increase. We will also discuss the channel estimation and the receiver adaptivity to time-varying channel conditions as well as the problem of the frequency acquisition and tracking. Finally, a new two-dimensional (polarization/time) differential encoding rule is proposed to overcome a polarization-ambiguity problem. In the numerical results, the receiver performance will be assessed versus the receiver complexity.

Index Terms—Electrical equalization, group velocity dispersion (GVD), intersymbol interference (ISI), optical coherent transmission systems, phase-shift keying (PSK), polarization mode dispersion (PMD), quadrature amplitude modulation (QAM).

I. INTRODUCTION

COHERENT optical systems, already investigated in early nineties, were lately abandoned when the advent of optical amplifiers favored the development and spreading of intensity-modulation/direct-detection systems [1]. Nowadays, however, the ever increasing data rates have stimulated a renewed interest toward coherent systems since they can allow to (more easily) use high-order modulation formats, thus reducing the symbol rate for a given bit rate and hence the speed of the required electronic processing and the impact of group velocity dispersion (GVD) and polarization mode dispersion (PMD) [2], [3]. Another advantage of coherent systems is represented by the absence of nonlinear transformations at the receiver that degrade the information content of the received signal. Hence, all postdetection processing techniques result more effective.

Theoretical and experimental works have been recently carried out to demonstrate the feasibility of coherent receivers for next generation optical communication systems. A first class of these works mainly focuses on the problem of carrier synchronization. In fact, although homodyne detection allows for a simple and effective postdetection processing [4], the demand for a stable and accurate local oscillator locked to the optical carrier still entails problems. On the other hand, if heterodyne schemes do not need an optical phase-locked loop (PLL), the necessary two frequency conversion steps result more expensive. For these reasons, the so called intradyne schemes may represent the solution [5]. For them, feedforward carrier recovery [6]–[12], digital PLLs [13], or automatic frequency control (AFC) loops [14] have been proposed or experimentally demonstrated. With the exception of [11], [12], [14], a relatively low frequency mismatch between the received signal and the local laser is assumed. In addition, except for the algorithms described in [10] and [14], a very low tolerance toward phase noise is achieved.

In [6] and [7], electronic polarization control for intradyne receivers is also discussed. A more sophisticated processing is required to compensate for GVD and PMD. The effectiveness of postdetection processing in coherent optical systems was known since these receivers were firstly devised and derived from communication-theoretical arguments. It was shown in [15] and experimentally demonstrated in [13], how simple fractionally spaced equalizers, in both homodyne and heterodyne systems, can completely compensate for the effect of GVD with any modulation formats. This is due to the fact that GVD is a "phase distortion" only and hence a filter, matched to the received pulse and adaptively adjusted, simply compensates for it. All these considerations can be extended to PMD and a polarization multiplexed transmission. In fact, in this case, a single-mode fiber (SMF) can be seen as a two-input/two-output channel whose frequency response, the so-called Jones transfer matrix, is unitary irrespective of the amount of GVD and the amount or model (1st, 2nd, or higher) of PMD. Hence, in this case a two-dimensional (2-D) matched filter can achieve perfect compensation.[1] This kind of solution was recently proposed or experimentally demonstrated by many authors (see e.g., [17]–[21] and references therein).

[1]A more involved processing is required to achieve the same goal when, instead of a coherent receiver, a different front end based on an interferometer is employed [16].

In this paper, we discuss and analyze in depth all theoretical and practical aspects related to the implementation of intradyne receivers. In particular, since the frequency offset between the optical carrier and the free-running local oscillator can be a notable fraction of the symbol rate, thus preventing the receiver from properly working, after the optoelectronic conversion an AFC is to be included in the processing unit to keep such offset within an acceptable range. This task can be performed by resorting to a couple of classical synchronization techniques borrowed from wireless communications and adapted to this new scenario. The core of the electronic receiver processing is then constituted by the above-mentioned 2-D matched filter, which is able to perfectly compensate for GVD and PMD. It will be implemented in the form of an adaptive 2-D fractionally spaced feedforward equalizer (FFE). In fact, we will demonstrate that by adopting the minimum mean square error (MMSE) criterion for the adaptation of its coefficients, the FFE converges to the 2-D matched filter. We will also provide a method for computing the FFE coefficients and the equivalent channel impulse response after the FFE, thus allowing to analytically compute the receiver performance, at least in the absence of phase noise and assuming that the uncompensated frequency offset after the AFC is small enough that the performance is not degraded. By using this tool, we are able to investigate in depth the receiver performance for a fixed FFE complexity in terms of outage probability curves. The ambiguities occurring when the adjustment of the FFE coefficients is performed in a decision-directed fashion will be then discussed and a countermeasure to them is also indicated in the form of a 2-D differential encoding rule. We will also face the major problem of coherent optical systems, namely, the receiver sensitivity to the phase noise introduced by the transmit and receive lasers. We will show that the robustness to this phenomenon can be considerably increased by using an asynchronous detection strategy and an asynchronous filter adjustment algorithm for the FFE taps, still borrowed from wireless communications. In particular, the detection strategy exploits decision feedback and does not increase significantly the receiver complexity.

The paper is organized as follows. In Section II, we describe the system model and the receiver structure. The major component of the receiver, the adaptive 2-D fractionally spaced FFE is analyzed in Section III, which focuses on its structure, adaptivity, analytical performance evaluation, and on an asynchronous detection strategy and an asynchronous filter adjustment algorithm to increase the robustness toward phase noise and frequency offset. Despite the adoption of this asynchronous strategies, the FFE cannot work properly without a preliminary frequency estimation and compensation, which must be performed in nondata-aided and nonclock-aided fashion. Hence, an AFC strategy is described in Section IV. To overcome the problem of a possible exchange of the two data streams transmitted over the two orthogonal polarizations, a 2-D differential encoding strategy and the corresponding decoding is described in Section V. Numerical results are then shown in Section VI and finally, Section VII provides some concluding remarks.

II. SYSTEM MODEL AND RECEIVER STRUCTURE

The system model (excluding the receiver part) and its low-pass equivalent are shown in Fig. 1. In the considered system, polarization multiplexing is employed. To this end, two independent sequences $\{a_{i,k}, k = 1, 2, \ldots, K - 1\}$, $i = 1, 2$, of $K - 1$ complex symbols belonging to an M-ary complex alphabet undergo separate differential encoding [22], thus generating two sequences $\{d_{i,k}, k = 0, 1, \ldots, K - 1\}$, $i = 1, 2$, of K complex symbols each, belonging to the same alphabet. For conciseness, we will use the notation[2]

$$\begin{aligned} \mathbf{a}_k &= [a_{1,k}, a_{2,k}]^T \\ \mathbf{d}_k &= [d_{1,k}, d_{2,k}]^T. \end{aligned} \tag{1}$$

Without loss of generality, in the numerical results we will consider classical phase-shift keying (PSK) signals, for which the standard differential encoding rule is employed, and square quadrature amplitude modulations (QAMs) for which quadrant differential encoding [22] (see also [23, Section V-A] for a concise description) is adopted. However, our derivations can also be applied to other alphabets, e.g., amplitude- and phase-shift keying (APSK) modulations, whose signal constellations are composed of more concentric rings of PSK points.

These two symbol streams are launched, after linear modulation, on two orthogonal states of polarization (SOPs) of an SMF. We can express the low-pass equivalent of the transmitted signal components as

$$[s_1(t), s_2(t)]^T = \sum_\ell \mathbf{d}_\ell p(t - \ell T) = \sum_\ell \mathbf{P}(t - \ell T)\mathbf{d}_\ell \tag{2}$$

where T is the symbol interval, $p(t)$ is the transmitted pulse, and $\mathbf{P}(t) = p(t)\mathbf{I}$, having denoted by $\mathbf{I}$ the 2×2 identity matrix. Without loss of generality, we suppose that $p(t) \otimes p^*(-t)|_{t=kT} = \delta_k$, where δ_k is the Kronecker delta and $\otimes$ denotes "convolution," that is we suppose that the transmitted pulse, after its corresponding matched filter, satisfies the condition for the absence of intersymbol interference (ISI). This ensures that in the back-to-back (b2b) case the optimal detector is the symbol-by-symbol one. Hence,[3]

$$\mathbf{P}^H(-t) \otimes \mathbf{P}(t)|_{t=kT} = \delta_k \mathbf{I}. \tag{3}$$

The SMF introduces PMD and GVD. We will quantify the amount of GVD by using the dimensionless chromatic dispersion index γ [24], defined as

$$\gamma = 2\lambda_0 R_b^2 D_r / \omega_0 \tag{4}$$

where ω_0 and λ_0 are the optical carrier frequency and wavelength, respectively, R_b is the bit rate, and D_r is the residual fiber dispersion (usually expressed in ps/nm), accounting for

---

[2]In the following, $(\cdot)^T$ denotes transpose, $(\cdot)^H$ transpose conjugate, and $(\cdot)^*$ complex conjugate.

[3]In (3), the integral of a matrix $\mathbf{A}(t)$ is defined as a new matrix whose entries are the integrals of the entries of the original matrix $\mathbf{A}(t)$.
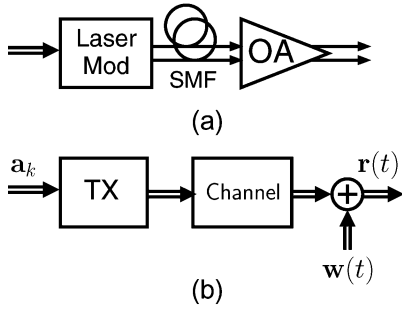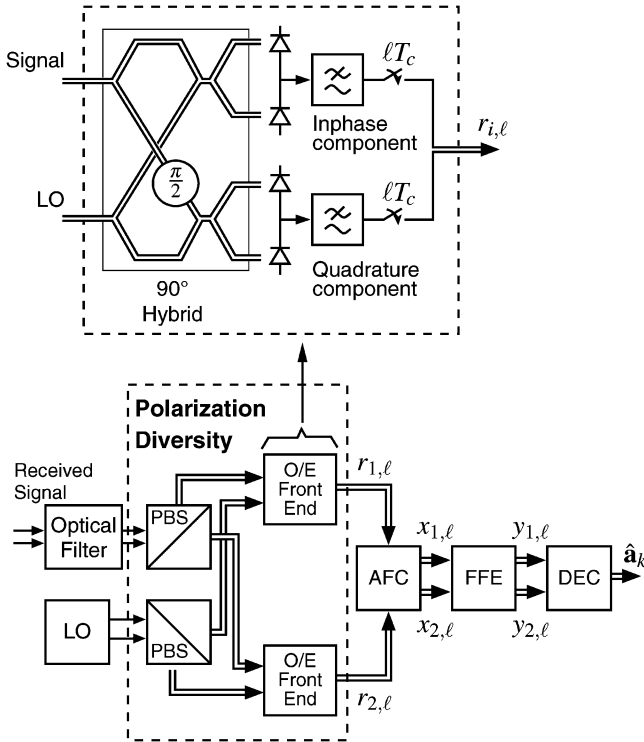
Fig. 1. (a) System model. (b) Low-pass equivalent.

Fig. 2. Receiver structure.

different pieces of fiber whose chromatic dispersion D has alternating sign, such as in a chromatic dispersion compensated link. In the case of a second-order model, PMD can be characterized by the values of the signal power splitting ρ among the principal states of polarization (PSPs), the differential group delay (DGD) τ, the DGD derivative $\Delta\tau_\omega$, and the PSPs rotation rate q_ω [25]. We denote by $\mathbf{H}(t)$ a 2×2 Jones matrix representing the 2-D impulse response of the SMF, accounting for both GVD and PMD and a possible constant unknown phase shift due to the phase uncertainty of the transmit and receive lasers.[4] Its entrywise Fourier transform is a unitary matrix irrespective of the amount of GVD and the amount or model (1st, 2nd, or higher) of PMD [25]. Hence

$$\mathbf{H}(t) \otimes \mathbf{H}^H(-t) = \mathbf{H}^H(-t) \otimes \mathbf{H}(t) = \delta(t)\mathbf{I} \qquad (5)$$

having denoted by $\delta(t)$ the Dirac delta function.

[4]The phase noise due to the transmit and receive lasers will be taken into account later.

The transmitted signal also experiences optical amplification before the receiver end. The low-pass equivalent of the amplified spontaneous emission (ASE) noise can be modeled as a couple of independent complex noise components, each with two-sided power spectral density (PSD) equal to N_0, taking into account the noise components on two orthogonal SOPs. We consider the receiver as composed of an analog part, the optoelectronic (O/E) front end, devoted to signal demodulation and conversion from the optical to the electrical domain, and a digital part devoted to electronic processing (see Fig. 2). After a preliminary optical filtering, two orthogonal SOPs are split through a polarization beam slitter (PBS). They are then separately combined with the optical field of a local oscillator laser (LO) in a 2×4 $90°$ hybrid [2] and detected with two balanced photodetectors. In this way, the two received signals, one for each SOP, are converted in the electrical domain, in practice performing a frequency conversion. We suppose, and this is certainly true since otherwise wavelength division multiplexing would not be possible, that the frequency offset F between the incoming signal and the LO laser is at most equal to the symbol rate.[5] This allows to use, at the receiver, a free-running LO laser without resorting to a complex optical PLL, thus delegating to the electronic processing part the task of a fine frequency recovery. In other words, an intradyne scheme is implemented [26].[6] Hence, the received signal can be expressed as

$$\mathbf{r}(t) = [r_1(t), r_2(t)]^T = \sum_\ell \mathbf{Q}(t - \ell T)\mathbf{d}_\ell e^{j2\pi Ft} + \mathbf{w}(t) \quad (6)$$

where $F \leq 1/T$ is the above mentioned frequency offset between the incoming signal and the local oscillator, $\mathbf{Q}(t)$ is a 2×2 matrix given by $\mathbf{Q}(t) = \mathbf{H}(t) \otimes \mathbf{P}(t)$, and $\mathbf{w}(t) = [w_1(t), w_2(t)]^T$ collects the noise signal components on the above-mentioned orthogonal SOPs. We assume that the ASE noise is dominant over thermal and shot noise.[7]

Without loss of generality, we assume that the further processing is fully digital, although an equivalent analog processing can be devised [15]. To this purpose, a possible way of extracting sufficient statistics from the received signal $\mathbf{r}(t)$ is by means of sampling at the Nyquist rate [28]. In the following, we will assume that η samples per symbol interval are extracted from the signal, that is the sampling interval is $T_c = T/\eta$. This number of samples depends on the bandwidth of the received useful signal and the value of F. We assume that the optical filter has no effect on the useful signal and that the electrical filters in Fig. 2 have squared amplitude response with vestigial symmetry around $\eta/2T$ [28]. This latter condition ensures that the noise samples are independent and identically distributed

[5]A fraction of $10 \div 20\%$ of the symbol rate is a more likely value of the frequency offset.

[6]An equivalent heterodyne scheme may also be conceived. In this case, the LO laser will perform the conversion from the optical to the electrical domain at an IF. A further LO in the electrical domain is then necessary. The frequency discriminator described later can still be adopted, provided that the frequency offset between the incoming signal converted to IF and the electrical LO is at most equal to the symbol rate.

[7]Although in any case the statistics of the overall noise are still Gaussian (shot noise has a Poisson distribution but can be well approximated as Gaussian [2], [3]), we consider the case of dominant ASE noise since in this case homodyne, heterodyne, and intradyne are equivalent in terms of performance [27].

complex Gaussian random variables with mean zero and variance $\sigma^2 = N_0\eta/T$ [28]. The samples of $\mathbf{r}(t)$ at discrete-time instants $\ell T_c = kT + nT_c = (k\eta + n)T_c$, $k = 0, 1, \ldots, K-1$, $n = 0, 1, \ldots, \eta - 1$, will be denoted as $\mathbf{r}_\ell = \mathbf{r}(\ell T_c) = [r_{1,\ell}, r_{2,\ell}]^T$.

The fine frequency recovery is then performed by means of an electrical AFC loop, which performs closed-loop frequency estimation and compensation assuming that neither data nor clock information is available. The samples at its output will be denoted as $\mathbf{x}_\ell = [x_{1,\ell}, x_{2,\ell}]^T$. As will be demonstrated in the next section, an adaptive 2-D fractionally spaced FFE of sufficient length is then able to perfectly compensate for GVD and PMD, thus allowing a simple classical symbol-by-symbol detection. In the presence of phase noise, a more robust symbol-by-symbol detection strategy with decision feedback will also be described.

III. ADAPTIVE 2-D FRACTIONALLY SPACED FFE

A. Two-Dimensional Fractionally-Spaced FFE

The adaptive 2-D fractionally spaced FFE will now be described, deferring the discussion on the frequency estimation and compensation to Section IV. In the hypothesis that the AFC block perfectly compensates for the frequency offset F, the samples $\mathbf{x}_{k\eta+n}$ can be expressed as

$$\mathbf{x}_{k\eta+n} = \sum_\ell \mathbf{Q}_{\eta(k-\ell)+n}\mathbf{d}_\ell + \mathbf{w}_{k\eta+n}$$

$$= \sum_{\ell=0}^{L_s-1} \mathbf{Q}_{\ell\eta+n}\mathbf{d}_{k-\ell} + \mathbf{w}_{k\eta+n} \qquad (7)$$

having assumed that $\mathbf{Q}_\ell = \mathbf{Q}(\ell T_c)$ has a length of L samples and $L_s = \lceil L/\eta \rceil$. Samples $\mathbf{w}_{k\eta+n}$ after the frequency compensation are statistically equivalent to the samples of the noise after the electrical filters. Hence, the two components $w_{1,k\eta+n}$ and $w_{2,k\eta+n}$ of $\mathbf{w}_{k\eta+n}$ are independent and identically distributed complex Gaussian random variables each with mean zero and variance $\sigma^2 = N_0\eta/T$, i.e., $E\{|w_{1,k\eta+n}|^2\} = E\{|w_{2,k\eta+n}|^2\} = \sigma^2 = N_0\eta/T$. Let us now suppose to filter the discrete 2-D signal $\mathbf{x}_{k\eta+n}$ with a 2-D filter with impulse response $\mathbf{C}_\ell$, $\ell = -L_c+1, \ldots, -1, 0$.[8] The output of this fractionally spaced FFE is

$$\mathbf{y}_{k\eta+n} = \sum_{\ell=-(L_c-1)}^{0} \mathbf{C}_\ell\mathbf{x}_{k\eta+n-\ell}. \qquad (8)$$

Assuming now that $\mathbf{Q}_\ell$ is known, since from (3) and (5)

$$\mathbf{Q}_{-\ell}^H \otimes \mathbf{Q}_\ell|_{\ell=k\eta} = \delta_k \mathbf{I} \qquad (9)$$

it is sufficient to choose $L_c = L$ and $\mathbf{C}_\ell = \mathbf{Q}_{-\ell}^H$, such that the FFE output at symbol time will be

$$\mathbf{y}_{k\eta} = \mathbf{d}_k + \mathbf{w}_{k\eta}', \qquad (10)$$

where $\mathbf{w}_{k\eta}'$ *is statistically equivalent* to $\mathbf{w}_{k\eta}$ since the FFE channel impulse response satisfy condition (9) and hence it does not color the noise. This is obvious, since PMD and GVD are

[8]Without loss of generality, this finite impulse response (FIR) filter is supposed to be anticausal. In this way, we can avoid to consider the delay of $L-1$ samples at its output, thus simplifying the notation.
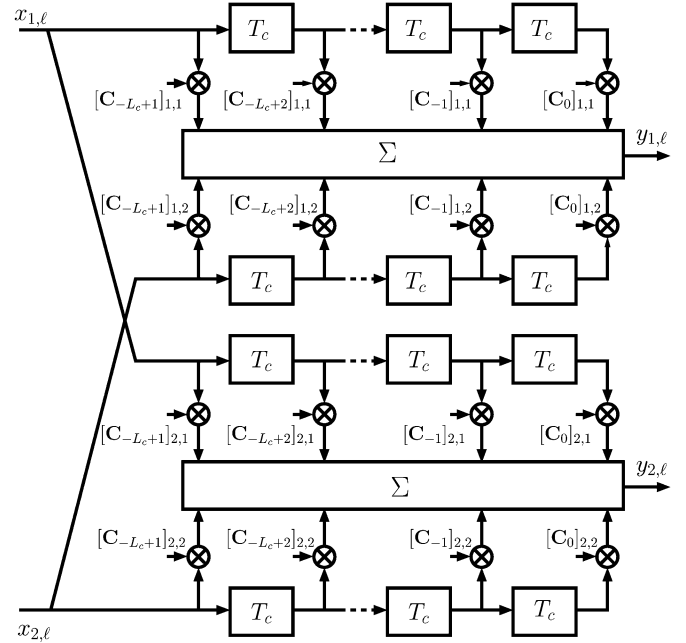

Fig. 3. FFE scheme.

phase distortions only (see also [16]). Hence, on each symbol stream, a symbol-by-symbol detector based on the strategy

$$\hat{d}_{i,k} = \arg\min_{d_{i,k}} \left[|y_{i,k\eta} - d_{i,k}|^2\right] \qquad (11)$$

and followed by differential decoding is sufficient to obtain a decision on the transmitted symbols. This is not surprising, since we are implementing a 2-D matched filter and the discrete-time channel impulse response $\mathbf{Q}_\ell$ satisfy the condition for the absence of ISI [see (9)]. We can also state that by filtering each of the two components of $\mathbf{x}_{k\eta+n}$ with a filter matched to $p(t)$, in the b2b case we would obtain an output statistically equivalent to $\mathbf{y}_{k\eta}$. Hence, the described receiver structure is able to attain a *perfect compensation for GVD and PMD*, provided that the AFC is able to compensate for the frequency offset and a perfect knowledge of $\mathbf{Q}_\ell$ is available. Notice that this FFE also performs the compensation for a constant phase shift, possibly introduced by transmit and receive lasers, which is implicit in the knowledge of the channel impulse response $\mathbf{Q}_\ell$. Hence, an explicit phase estimation, such as that used in [17], is not necessary, at least when phase noise is absent (i.e., assuming that the transmit and receive lasers only introduce a constant phase shift). The FFE structure is shown in Fig. 3 where $[\mathbf{C}_\ell]_{k,n}$ denotes the (k, n) entry of $\mathbf{C}_\ell$.

We would like to emphasize here that the capability of this scheme to perfectly compensate for PMD and GVD is not due to the fact that these phenomena are *linear* distortions. In fact, a linear distortion inducing ISI after the matched filter always degrade the performance despite the optimal processing based on the Viterbi algorithm is adopted at the receiver [29]. In this case, however, we have *phase* distortions, mathematically described by (5). Hence, the 2-D matched filter, implemented by the FFE, will perfectly compensate for them.

In a practical implementation, the channel must be estimated, or in other words, an adaptive FFE is necessary. This can be

done in a decision-directed fashion with the algorithm described in Section III-B. However, it must be noticed that when a decision-directed algorithm is employed, the channel will be estimated except for a couple of ambiguities. The first one affects each signal component and is related to the rotation symmetry angle φ of the employed alphabet. As an example, $\varphi = \pi/2$ for quaternary PSK (QPSK) and QAM alphabets. The second kind of ambiguity is related to a possible exchange of the transmitted streams. In other words, in the case of a decision-directed channel estimation the FFE output at symbol time will be

$$\mathbf{y}_{k\eta} = \mathbf{R}\mathbf{d}_k + \mathbf{w}'_{k\eta} \tag{12}$$

where

$$\mathbf{R} = \begin{bmatrix} e^{j\varphi_1} & 0 \\ 0 & e^{j\varphi_2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^i \tag{13}$$

φ_1 and φ_2 are multiple of φ, and $i = 0, 1$—here, $i = 1$ means that the two polarizations have been exchanged. The first kind of ambiguity does not represent a problem since, as mentioned in Section II, the two streams are differentially encoded. Regarding the second kind of ambiguity, this problem can be faced by inserting some known (pilot) symbols in the transmitted sequence. However, since a training sequence is usually inserted to ease the convergence of the FFE, these known symbols will also allow to correctly align the two streams at the transmission beginning.[9] One may wonder whether this is sufficient, i.e., once the two streams are correctly aligned, if it is possible to avoid the further insertion of known symbols without a subsequent exchange of the two polarizations due to the noise or to channel variations. In order to investigate this aspect, we run several computer simulations in which, after the initial training period that allows the FFE to reach convergence, simultaneous abrupt variations of the azimuth and ellipticity angles of the PSPs up to 10° have been injected every one hundred symbol intervals and never observed an exchange of the two polarizations. Hence, one may conjecture that unless a very strong channel variation occurs—and in this case a new training period is necessary for the FFE also—there is no risk to exchange the two polarizations even in the absence of other known symbols. Notice that this is not true for the first kind of ambiguity in the sense that, although after the training sequence the FFE is able to converge ensuring $\varphi_1 = \varphi_2 = 0$ in (13), due to the ASE noise and the lasers' phase noise, phase slips are highly likely [31]. Hence, the differential encoding described in Section II cannot be avoided. In any case, a 2-D (polarization/time) differential encoding rule that solves both the ambiguity problems, thus avoiding the insertion of known symbols, will also be described in Section V.

B. Filter Adjustment

The filter adjustment can be performed by resorting to the zero-forcing (ZF) criterion. However, we must force to zero the

[9]The transmission of the training sequence, although not necessary for small constellations such as QPSK since convergence can still be achieved, could be avoided by resorting to some of the available algorithms for blind equalization (see [30] and references therein). An example of such a technique is the constant modulus algorithm (CMA), tailored in [21] for these applications. However, after the acquisition phase, it is more convenient to adopt the algorithm described in Section III-B, since it exhibits a better performance, converging to the matched filter, and has a lower complexity, as the reader can easily notice.

2-D channel impulse response after the FFE, namely $\mathbf{C}_\ell \otimes \mathbf{Q}_\ell$, only at the time instants of the form $k\eta$, with $k \neq 0$, by neglecting its values at time instants $k\eta + n$, with $n \neq 0$. Hence, the MMSE criterion must be preferred since 1) it converges to the 2-D matched filter when the FFE has a length $L_c \geq L$, as shown in the Appendix, and 2) it guarantees a better performance with respect to the ZF criterion when the FFE has a length lower than $L_c < L$ [29]. In the Appendix, a method to compute the expression of the optimal MMSE coefficients $\mathbf{C}_\ell$ is provided for each value of L_c.

The MMSE criterion is based on the minimization of

$$J = E\left\{\|\mathbf{y}_{k\eta} - A\mathbf{d}_k\|^2\right\} = E\left\{\|\mathbf{e}_k\|^2\right\} \tag{14}$$

having defined $\mathbf{e}_k = \mathbf{y}_{k\eta} - A\mathbf{d}_k$, where $A = 1 + (\sigma^2/\sigma_d^2)$ and $\sigma_d^2 = E\{|d_{i,k}|\}$. For the orthogonality principle [32], this is equivalent to the following conditions

$$E\left\{\mathbf{e}_k\mathbf{x}_{k\eta-\ell}^H\right\} = 0, \quad \ell = -(L-1), \ldots, -1, 0.$$

Therefore, the filter coefficients $\mathbf{C}_\ell$ are iteratively adjusted according to the following updating rule (decision-directed least-mean-square (LMS) algorithm with a symbol-time adjustment) [29]

$$\hat{\mathbf{C}}_\ell^{(k+1)} = \hat{\mathbf{C}}_\ell^{(k)} - \alpha\hat{\mathbf{e}}_k\mathbf{x}_{k\eta-\ell}^H \tag{15}$$

having denoted by $\hat{\mathbf{C}}_\ell^{(k)}$ the estimate of the filter coefficients at the kth symbol interval. Because of the adaptive algorithm, the receiver is able to track the phase variations due to the lasers' phase noise, provided they are sufficiently slow, as will be shown in the numerical results. However, in order to avoid the use of expensive transmit and receive lasers, characterized by a very small linewidth, it is more convenient to increase the receiver robustness by using the asynchronous algorithm described in Section III-D.

C. FFE Performance for $L_c < L$

When $L_c < L$, the FFE will not be able to perfectly remove ISI and to perfectly demultiplex the transmitted streams. In this case, a method to compute the FFE coefficients is described in the Appendix. Hence, the equivalent channel impulse response $\mathbf{F}_\ell = \mathbf{C}_\ell \otimes \mathbf{Q}_\ell$ after the FFE can be derived, and the samples $\mathbf{y}_{k\eta}$ can be expressed as

$$\mathbf{y}_{k\eta} = \sum_{\ell=0}^{L_F} \mathbf{F}_{\ell\eta}\mathbf{d}_{k-\ell} + \mathbf{w}'_{k\eta} \tag{16}$$

where L_F is the length (in symbols) of $\mathbf{F}_\ell$. In this case, the FFE changes the noise statistics. In fact, the noise is no more white and, in addition, the noise components on the two streams are correlated. With straightforward manipulations, we obtain

$$E\left\{\mathbf{w}'_{k\eta}\mathbf{w}'^H_{(k+m)\eta}\right\} = \sum_\ell \mathbf{C}_\ell\mathbf{C}^H_{m\eta+\ell}.$$

The receiver performance could be improved by taking into account the correlation among consecutive noise samples (thus resorting to a Viterbi algorithm) or simply by considering a symbol-by-symbol detector, which jointly detect the symbols

on the two streams at the same instant taking into consideration their correlation. However, this approach is not pursued here since we are interested in still considering, after the FFE, a receiver similar to that adopted in the case of perfect compensation (i.e., $L_c = L$).

When computing the receiver performance, we have to take into account, on each stream, the intersymbol arising from other symbols of the same stream and the intersymbol arising from the other stream. The performance in terms of bit error rate (BER) can be easily computed by using the exhaustive method (with exponential complexity in the number of interferers) or, when the number of interfering terms is very large, the method based on the computation of the ISI moments [33]. These analytical methods will be used in the results reported in Section VI.

D. Asynchronous Detection Strategy and Asynchronous Filter Adjustment

As mentioned, the FFE is able to compensate for an unknown phase shift and, because of its adaptive adjustment, is also able to cope with a slowly varying phase noise also. However, in the presence of a strong phase noise, in order to make the FFE able to follow the channel phase variations, its step-size α in (15) should be increased, thus degrading its steady-state performance. Hence, in this case, it is more convenient to delegate to a couple of digital PLLs, placed after the FFE, the task of channel phase estimation and compensation. In this way, the equivalent bandwidths of these PLLs can be optimized according to the phase noise at hand, whereas the FFE step-size can be properly reduced, since the PMD is slowly varying, to improve the overall receiver performance.

The receiver robustness against phase noise, which is one of the main issues in coherent optical systems, can be further increased by resorting to a symbol-by-symbol asynchronous detection strategy. In particular, one of the asynchronous strategies in [23] and [34] can be employed which entail no performance degradation w.r.t. a synchronous detection strategy when the phase noise is absent.[10] According to them, symbols $d_{i,k}$, $i = 1, 2$, are detected as

$$\hat{d}_{i,k} = \arg \max_{d_{i,k}} \left[\left| y_{i,k\eta} d_{i,k}^* + g_{i,k} \right| - \frac{|d_{i,k}|^2}{2} \right] \quad (17)$$

where

$$g_{i,k} = \sum_{n=1}^{N} y_{i,(k-n)\eta} \hat{d}_{i,k-n}^* \quad (18)$$

[10]There is some confusion between the terminology used in the optical community and that used in the context of wireless communication systems. In wireless systems, the term "coherent" refers to the knowledge of the phase of the received signal, that is a coherent receiver is designed assuming that the phase of the received signal is known or separately estimated. On the contrary, a noncoherent receiver is designed assuming that the channel phase is unknown and modeled as a random variable or a stochastic process. This terminology is employed in [23], [34], and [35]. In optical systems, the term "coherent" refers to the coherence of the optical carrier whereas the terms "synchronous" and "asynchronous" refer to a processing, which assumes or not the knowledge of the channel phase.

in the case of the strategy in [23], whereas $g_{i,k}$ is recursively computed as

$$g_{i,k} = \beta g_{i,k-1} + y_{i,(k-1)\eta} \hat{d}_{i,k-1}^* \quad (19)$$

in the case of the strategy in [34]. The integer parameter N in (18) and the real parameter $0 \leq \beta \leq 1$ in (19) are design parameters that can be optimized for the phase noise at hand. The receiver robustness can be further increased by resorting to the approach, based on linear prediction and still working in a symbol-by-symbol fashion, described in [36]. The prediction coefficients are, in this case, *a priori* computed based on the knowledge of the phase noise at hand. Alternatively, the approach described in [10] could be adopted but, despite the increased complexity, since a preliminary channel estimate is then refined by filtering past and future estimates, there are no performance advantages with respect to the solutions in [23], [34], and [36].

When an asynchronous detection strategy is adopted, it is more convenient to substitute the described filter adjustment rule (20) with the following one

$$\hat{\mathbf{C}}_\ell^{(k+1)} = \hat{\mathbf{C}}_\ell^{(k)} - \alpha(\mathbf{y}_{k\eta} - A\mathbf{g}_k \odot \hat{\mathbf{d}}_k)\mathbf{x}_{k\eta-\ell}^H \quad (20)$$

where $\mathbf{g}_k = (g_{1,k}, g_{2,k})^T$ and $\odot$ denotes the Hadamard product (entrywise product or Schur product). This update rule is the extension to the case of two polarizations of that described in [37] and converges to the same minimum of the rule (15), although it is more robust in the presence of phase noise. The adoption of the asynchronous detection strategy and the asynchronous filter adjustment has also a convenient side effect. In fact, in this case, the FFE must track the variations due to channel only without taking into account those due to phase noise. This allows for a more relaxed adjustment of the FFE taps.

The described symbol-by-symbol asynchronous detection strategies (17), (18), and (17)–(19) and the updating rule (20) can be equivalently expressed as a function of the information symbols $\{a_{i,k}\}$. This is certainly possible, since the sequence $\{d_{i,k}\}$ is, after all, a function of the sequence $\{a_{i,k}\}$. As an example, in the case of a PSK alphabet, the symbol-by-symbol asynchronous strategy (17), (18) can be equivalently expressed as (taking also into account that $|a_{i,k}| = |d_{i,k}| = 1$ and that $d_{i,k} = d_{i,k-1}a_{i,k}$ according to the differential encoding rule) [23]

$$\hat{a}_{i,k} = a_{i,k} \left| y_{i,k\eta} a_{i,k}^* + g_{i,k}' \right| \quad (21)$$

where

$$g_{i,k}' = y_{i,(k-1)\eta} + \sum_{n=2}^{N} y_{i,(k-n)\eta} \prod_{m=1}^{n-1} \hat{a}_{i,k-m}. \quad (22)$$

In this way, an explicit differential decoding is not necessary.

IV. FREQUENCY ESTIMATION AND COMPENSATION

Frequency estimation and compensation may be performed by using a closed-loop nondata-aided and nonclock-aided algorithm. The AFC scheme is shown in Fig. 4. We denote by $\hat{F}_\ell$ the frequency estimate at time ℓT_c. Using this estimate, the AFC
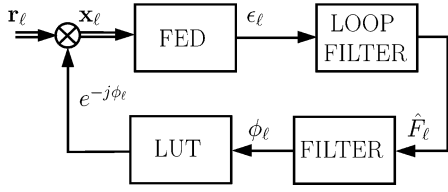
Fig. 4. AFC scheme.

input is derotated by a phase ϕ_ℓ related to $\hat{F}_\ell$ by the recursive equation

$$\phi_{\ell+1} = \phi_\ell + 2\pi \hat{F}_\ell T_c$$

obtaining the signal $\mathbf{x}_\ell = \mathbf{r}_\ell e^{-j\phi_\ell}$ sent at the FFE input. The phasor $e^{-j\phi_\ell}$ is obtained through the lookup table (LUT) shown in the figure. The signal $\mathbf{x}_\ell$ is also sent to the input of the frequency error detector (FED), which computes the error signal ϵ_ℓ used to update the frequency estimate according to the recursion

$$\hat{F}_\ell = \hat{F}_{\ell-1} + \zeta \epsilon_\ell$$

where ζ is a proper step-size.

Several classical FEDs can be chosen according to the desired performance/complexity tradeoff. In our simulations, we considered the delay&multiply (DM) FED and a FED derived from the maximum likelihood (ML) criterion [31]. The DM FED is also proposed and experimentally demonstrated in [14]. However, no polarization diversity is considered there nor the effects of both GVD and PMD on the FED performance are addressed. Both DM and ML FEDs suffice to our purposes although, as known, the DM FED has a lower complexity but a worse performance. In this case, the error signal ϵ_ℓ is obtained as [31]

$$\epsilon_\ell = \Im \left\{ x_{1,\ell} x_{1,\ell-1}^* + x_{2,\ell} x_{2,\ell-1}^* \right\}$$

whereas in the case of the ML-based FED, by following the same manipulations in [31] and using also (9), we obtain

$$\epsilon_\ell = \Im \left\{ x_{1,\ell}' x_{1,\ell}''^* + x_{2,\ell}' x_{2,\ell}''^* \right\}$$

where $x_{i,\ell}'$ and $x_{i,\ell}''$ are obtained by filtering $x_{i,\ell}$ with two finite impulse response (FIR) filters with impulse responses $p_{-\ell} = p(-\ell T_c)$ and $\ell p_{-\ell}$, respectively.[11]

V. POLARIZATION-TIME (2-D) DIFFERENTIAL ENCODING

To solve the ambiguity problem related to an exchange of the two data streams, in this section, we propose a simple encoding rule to be used before the two streams are separately encoded with the classical differential encoding rule. The resulting overall 2-D encoder is shown in Fig. 5. The joint encoder can be used regardless the modulation format and the adopted detection strategy (synchronous or asynchronous).

With reference to the figure, let us denote by $b_{1,k}$ and $b_{2,k}$ the couple of symbols at the output of the joint encoder. Let us also establish an arbitrary one-to-one correspondence between

the alphabet of symbols and the integer set $\{0, 1, \ldots, M-1\}$ and denote by $\overline{a}_{i,k}$ and $\overline{b}_{i,k}$ the integer representations of $a_{i,k}$ and $b_{i,k}$, respectively. The joint encoder is a finite-state machine with state $\sigma_k \in \{\pm 1\}$. The "output" and "next state" functions are, respectively,

$$(\overline{b}_{1,k}, \overline{b}_{2,k}) = \begin{cases} (\overline{a}_{1,k}, \overline{a}_{2,k}) & \text{if } \sigma_k = -1 \\ (\overline{a}_{2,k}, \overline{a}_{1,k}) & \text{if } \sigma_k = 1 \end{cases} \quad (23)$$

$$\sigma_{k+1} = \begin{cases} -1 & \text{if } \overline{b}_{1,k} < \overline{b}_{2,k} \\ \sigma_k & \text{if } \overline{b}_{1,k} = \overline{b}_{2,k} \\ 1 & \text{if } \overline{b}_{1,k} > \overline{b}_{2,k}. \end{cases} \quad (24)$$

Denoting by $\hat{\overline{b}}_{i,k}$ the integer representations of symbols of the stream i after the classical differential decoding at the receiver, the integer representation of the original information symbols will be

$$(\hat{\overline{a}}_{1,k}, \hat{\overline{a}}_{2,k}) = \begin{cases} (\hat{\overline{b}}_{1,k}, \hat{\overline{b}}_{2,k}) & \text{if } \hat{\sigma}_k = -1 \\ (\hat{\overline{b}}_{2,k}, \hat{\overline{b}}_{1,k}) & \text{if } \hat{\sigma}_k = 1 \end{cases} \quad (25)$$

where $\hat{\sigma}_k$ is the estimate of the state σ_k, recursively updated as

$$\hat{\sigma}_{k+1} = \begin{cases} -1 & \text{if } \hat{\overline{b}}_{1,k} < \hat{\overline{b}}_{2,k} \\ \hat{\sigma}_k & \text{if } \hat{\overline{b}}_{1,k} = \hat{\overline{b}}_{2,k} \\ 1 & \text{if } \hat{\overline{b}}_{1,k} > \hat{\overline{b}}_{2,k}. \end{cases} \quad (26)$$

VI. NUMERICAL RESULTS

In order to assess the performance of the proposed receivers, we use standard Monte Carlo simulations and the analytical method described in Section III-C. The considered modulation formats are QPSK with the standard differential encoding rule and 16-QAM with quadrant differential encoding. Gray mapping is employed to map bits onto M-ary symbols. A nonreturn-to-zero (NRZ) pulse filtered through an electrical baseband Gaussian filter with -3 dB bandwidth equal to $1/T$ is adopted at the transmitter. At the receiver, we use an optical fourth-order Gaussian filter whose low-pass equivalent has a -3 dB bandwidth equal to $1/T$. The receiver works using two samples per symbol interval (i.e., $\eta = 2$) and unless phase noise is present or otherwise stated, after the FFE, the standard symbol-by-symbol detection rule (11) is adopted.

A. Performance in the Absence of Phase Noise

In Figs. 6 and 7, the BER curves for polarization-multiplexed QPSK and 16-QAM, respectively, versus E_b/N_0, E_b being the received signal energy per information bit, are shown.[12] In all these figures, the presence of GVD with $\gamma = 16$, and second-order PMD with $\rho = 0.5$, $\Delta\tau = 6T_b$, $\Delta\tau_\omega = 1.2T_b^2$, and $q_\omega = 0.8T_b$, is considered. Different values of L_c are considered. In the figures, we used marks for Monte Carlo simulations and lines for the analytical results. The excellent agreement deserves no further comment. It can be observed that, provided the value of L_c is sufficiently large, the performance of the b2b

---

[11]For the commonly used transmitted waveforms, these FIR filters have a very small number of taps.

[12]As shown in [38], the ratio E_b/N_0 represents the number of detected photons per bit at the input of the optical amplifier and is related to the optical signal-to-noise ratio (OSNR) by $E_b/N_0 = 2 \times \text{OSNR} \times WT$, with W being the reference measurement bandwidth usually taken equal to 12.5 GHz (0.1 nm).
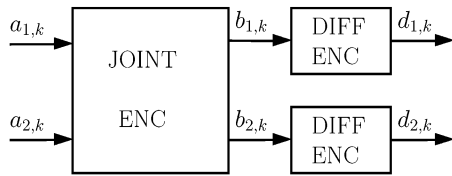
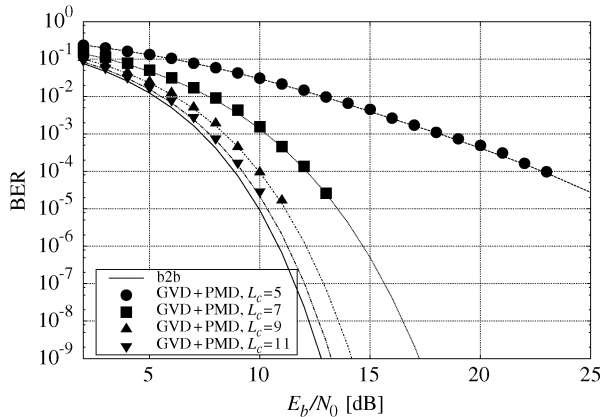Fig. 5. Overall polarization/time differential encoder.
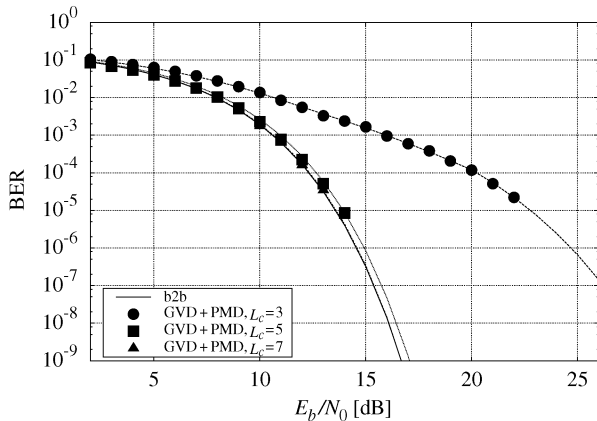

Fig. 6. Performance for a QPSK transmission.

Fig. 7. Performance for a 16-QAM transmission.

Fig. 8. Values of E_b/N_0 necessary to obtain a BER of 10^{-5} versus the value of γ, when only GVD is present.

Fig. 9. Outage probability due to first- and second-order PMD. The considered modulation format is a QPSK.

case, also shown for comparison, can be attained. This minimal value of L_c is 11 for the QPSK case and 5 for the 16-QAM. This is not surprising since we have shown the performance for the same values of GVD and PMD parameters. Since these values are normalized to the bit interval, for a given value of T_b, the symbol interval is larger for 16-QAM and the amount of dispersion is lower. To obtain the simulation results, the FFE was adaptively adjusted and the error count started after the acquisition phase. During the acquisition phase, a training sequence was employed. The number of necessary training symbols obviously depends on the value of the adopted step-size and on the possible use of a time-varying step-size—to reduce the acquisition phase the FFE step-size can be progressively reduced. For the analytical results, the steady-state value of the FFE coefficients was computed with the method described in the Appendix and this value then used, as explained in Section III-C.
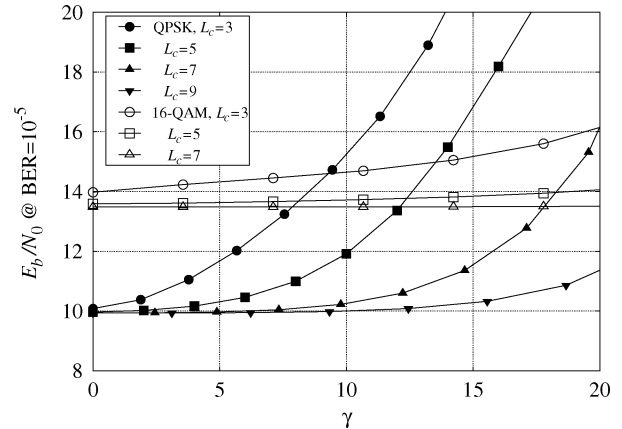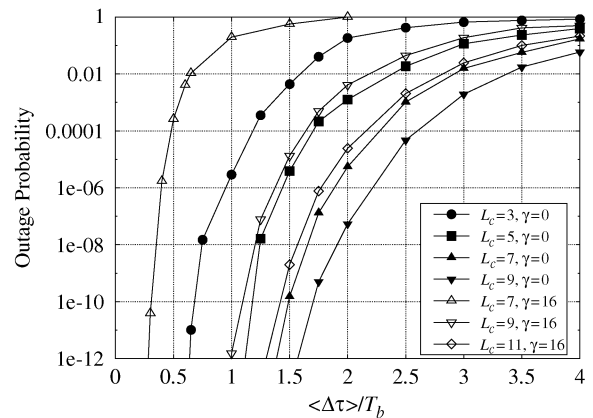
To assess the robustness to GVD of the proposed schemes, in Fig. 8 we show, for both QPSK and 16-QAM, the values of E_b/N_0 necessary to obtain a BER of 10^{-5} for a different amount of dispersion. We can see that, provided the value of L_c is sufficiently large, an arbitrarily amount of GVD can be compensated. Finally, for the QPSK modulation, in Fig. 9, we show (both in the absence and presence of GVD with $\gamma = 16$) the outage probability due to second-order PMD, defined here as the probability that the penalty exceeds the margin of 3 dB when the BER is fixed at 10^{-12}, versus the normalized mean value of the DGD. There is no closed-form approximation for the outage probability due to second-order PMD, and we analytically evaluated it as described in [39], using the model in [40] for the fiber Jones matrix. In our experience, the analytical second-order models for the fiber PMD are always more pessimistic than the random waveplate model, which accounts for all PMD orders, thus we are confident that our results also remain valid when taking into account higher PMD orders.

Finally, in Fig. 10, the robustness of the proposed receivers to GVD and first-order PMD is reported by showing the contour curves corresponding to an E_b/N_0 penalty of 3 dB with respect to the b2b case and for a BER of 10^{-12}. As can be observed,
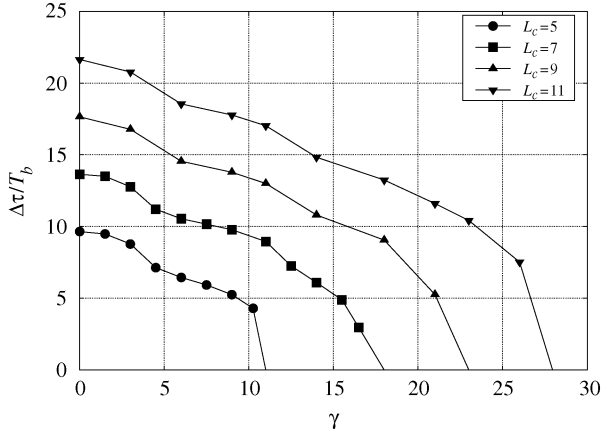
Fig. 10. Contour curves corresponding to an E_b/N_0 penalty of -3 dB for a BER $= 10^{-12}$, versus $\Delta\tau/T_b$ and γ. QPSK modulation is considered.
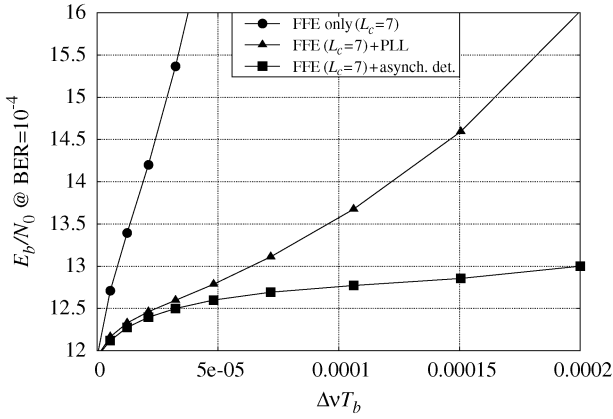

Fig. 12. Values of E_b/N_0 necessary to obtain a BER of 10^{-4} versus the value of the laser linewidth (FWHM, normalized to the bit rate). The considered modulation format is a 16-QAM.

Fig. 11. Values of E_b/N_0 necessary to obtain a BER of 10^{-4} versus the value of the laser linewidth (FWHM, normalized to the bit rate). The considered modulation format is a QPSK.

Fig. 13. Robustness to an uncompensated frequency offset.

the receiver robustness is impressive, provided the value of L_c is sufficiently large.

B. Robustness to Phase Noise

In order to assess the robustness to phase noise, Figs. 11 and 12 show, for the same amount of PMD and GVD considered in Figs. 6 and 7, the values of E_b/N_0 necessary to obtain a BER of 10^{-4} for different values of the sum of the (full-width half-maximum) linewidth $\Delta\nu$ of the transmit and receive lasers normalized to the bit rate. In all cases, the FFE has length $L_c = 5$. Fig. 11 refers to the QPSK, whereas Fig. 12 refers to the 16-QAM. In both figures, we considered the case of a FFE alone, and in this case its coefficients are adjusted every symbol time to follow the phase variations, and the cases when a PLL or the asynchronous strategy (17), (18) are employed. The PLL equivalent bandwidth and the value of N for the asynchronous strategy (17), (18), have been optimized for every value of the lasers' linewidth. As previously mentioned, the asynchronous strategy exhibits an impressive robustness and is immune from the occurrence of losses of lock which on the contrary affect the PLL.
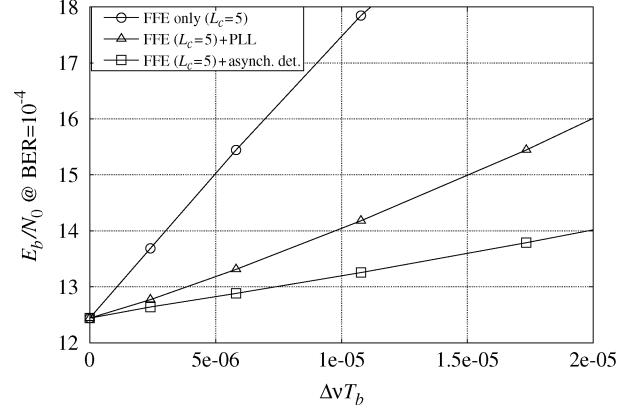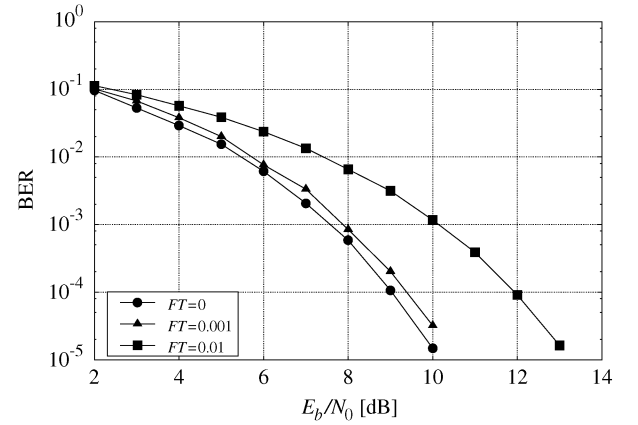
C. Frequency Estimation and Compensation

We now address the problem of the frequency estimation and compensation. First of all, we consider the asynchronous detection strategy (17), (18) and assess its robustness to an uncompensated frequency offset. To this purpose, we consider the QPSK modulation, although similar considerations also hold for the 16-QAM. For the same amount of PMD and GVD used in Fig. 6, and assuming $L_c = 11$, in Fig. 13 we show the performance for different values of the normalized uncompensated (i.e., after the AFC) frequency offset FT. As can be observed, the AFC must be able to reduce the normalized frequency offset below 10^{-3} in order to assure a limited performance loss.

The behavior of the considered FEDs is illustrated by the so-called S-curve, defined as

$$S(F) = E\{\epsilon_\ell | F\}$$

and computed when the loop is open. This function, that in general looks like an "S" (rotated by $90°$), allows us to understand if the estimator is biased and if it can track some possible wrong values (false locks). Moreover, it shows the acquisition range of the closed loop—the loop will eventually lock on the incoming carrier frequency provided that F is within this range. In Fig. 14 the S-curves for both FEDs are reported. These curves were
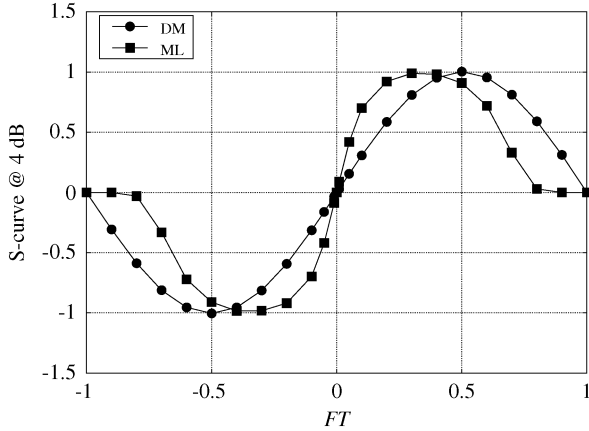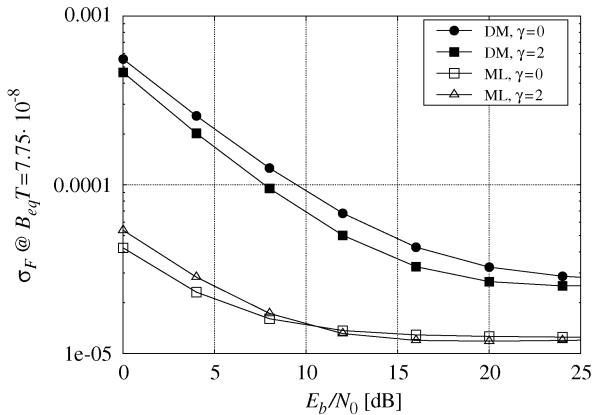
Fig. 14. S-curves for the DM and ML FEDs.

Fig. 15. Standard deviation of the frequency error for the DM-and ML-based AFCs.

Fig. 16. Performance of the proposed 2-D differential encoding rule.

E. Fixed-Point Receiver Implementation

We now report some investigations about the fixed-point implementation of the fractionally spaced 2-D FFE. In particular, the aim of this study is the analysis of the performance degradation, with respect to the unquantized floating-point version, due to the quantization of the received samples and of the FFE coefficients. We found that, although a limited number of bits is necessary to represent the received samples and the FFE coefficients when involved in the filtering operation (8), a larger number of bits is necessary to represent the FFE coefficients in the updating rules (15) or (20). However, the channel coefficients need to be updated very slowly and hence a larger number of bits to represent the channel coefficients when updated does not represent a significant increase in complexity. Once updated, a reduced precision representation can then be derived to be used in the filtering operation. In Fig. 17, we consider the same modulation formats and the same channel conditions of Fig. 6 and the receiver for $L_c = 9$. Along with the performance of the unquantized floating-point implementation of the FFE, we show the performance for different fixed-point implementations. Each curve is labeled with a couple of integers. The first one is the number of bits necessary to represent each component (real or imaginary) of received samples and FFE coefficients when involved in the filtering operation. The second one is, instead, the number of bits necessary to represent the component (real or imaginary) of the FFE coefficients in the updating rule. It can be observed that a limited performance degradation is obtained with at least 7 bits for the filtering operation and 16 bits for the updating rule.

VII. CONCLUSION

The design of intradyne coherent optical communication systems has been addressed, facing all practical and theoretical aspects. The main receiver component is a 2-D matched filter, implemented in the form of an adaptive 2-D fractionally spaced feedforward equalizer, which is able to perfectly compensate for GVD and PMD and to demultiplex two data streams transmitted over two orthogonal polarizations. The receiver robustness to phase noise is increased by using an asynchronous symbol-by-symbol detection strategy and asynchronous filter adjustment, and a novel 2-D differential encoding rule has been proposed

computed for a value of $E_b/N_0 = 4$ dB, although they are insensitive to the signal-to-noise ratio. From this figure, it is clear that the acquisition range is $-(1/T) \leq F \leq (1/T)$ for the DM and $-(0.8/T) \leq F \leq (0.8/T)$ for the ML FED. The AFC loop is unbiased and we verified that this is also true in the presence of ISI induced by the GVD and/or PMD.

The performance of the AFC loop in terms of standard deviation of the frequency error (which coincides with the square root of the mean square error since the loop is unbiased) is shown in Fig. 15 versus E_b/N_0 for a value of the normalized loop equivalent bandwidth $B_{eq}T = 7.75 \times 10^{-8}$ and for two values of the chromatic dispersion index γ. This figure confirms that, for both FEDs, provided the step-size is properly chosen, the frequency offset after the AFC can be reduced as low as required.

D. Two-Dimensional Differential Encoding

The performance, in terms of BER, of the proposed 2-D differential encoding rule is illustrated in Fig. 16 and compared with that of the classical differential encoding. We considered the b2b case since the FFE performance is irrelevant for this purpose. As can be noted, the immunity to an exchange of the two data streams is obtained at the expense of a very limited performance loss (less than 0.1 dB).
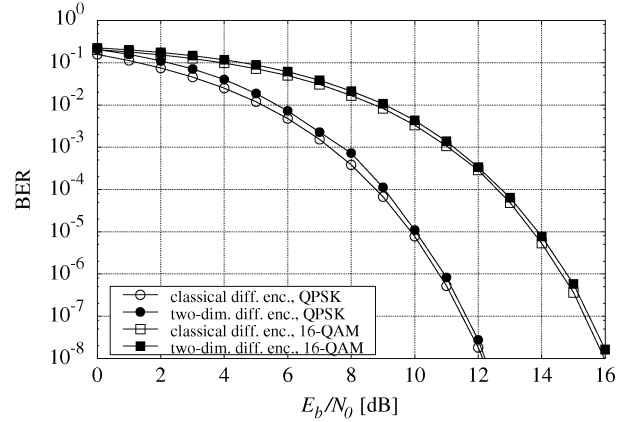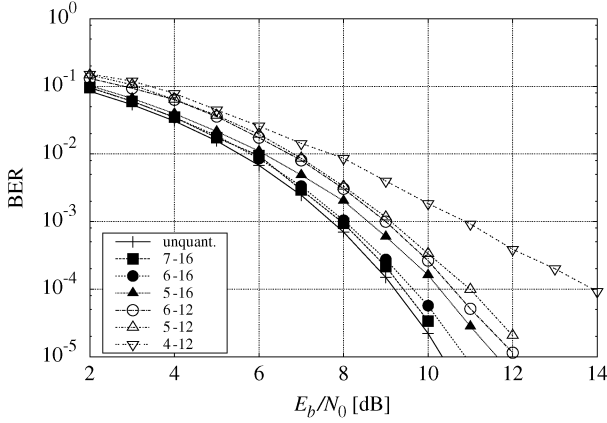
Fig. 17. Performance for different fixed-point receiver implementations.

to avoid an exchange of the data streams transmitted over two orthogonal polarizations. The details related to the frequency acquisition and compensation have also been discussed.

APPENDIX

In this appendix, we show how to compute the FFE coefficients following the MMSE criterion. We also show that when $L_c = L$ the FFE implements the matched filter.

To simplify the notation, we consider the output of the FFE at discrete-time 0 whose expression is

$$\mathbf{y}_0 = \sum_{\ell=-(L_c-1)}^{0} \mathbf{C}_\ell \mathbf{x}_{-\ell} = \sum_{i=0}^{L_c-1} \mathbf{C}_{-i} \mathbf{x}_i$$

where from (7)

$$\mathbf{x}_i = \sum_m \mathbf{Q}_{i-m\eta} \mathbf{d}_m + \mathbf{w}_i.$$

The mean square error (MSE) to be minimized is

$$J = E\left\{\|\mathbf{y}_0 - A\mathbf{d}_0\|^2\right\}. \tag{27}$$

For the orthogonality principle [32], coefficients $\mathbf{C}_\ell$ must satisfy the condition

$$E\left\{(\mathbf{y}_0 - A\mathbf{d}_0)\mathbf{x}_\ell^H\right\} = \mathbf{0}, \quad \ell = 0, 1, \ldots, L_c - 1$$

or, equivalently

$$E\left\{\mathbf{y}_0 \mathbf{x}_\ell^H\right\} = A E\left\{\mathbf{d}_0 \mathbf{x}_\ell^H\right\}$$

that is

$$\sum_{i=0}^{L_c-1} \mathbf{C}_{-i} E\left\{\mathbf{x}_i \mathbf{x}_\ell^H\right\} = A E\left\{\mathbf{d}_0 \mathbf{x}_\ell^H\right\}. \tag{28}$$

Let us compute the term

$$E\left\{\mathbf{x}_i \mathbf{x}_\ell^H\right\} = \sum_m \sum_n \mathbf{Q}_{i-m\eta} E\left\{\mathbf{d}_m \mathbf{d}_n^H\right\} \mathbf{Q}_{\ell-n\eta}^H$$
$$+ E\left\{\mathbf{w}_i \mathbf{w}_\ell^H\right\}$$
$$= \sigma_d^2 \sum_m \mathbf{Q}_{i-m\eta} \mathbf{Q}_{\ell-m\eta}^H + \sigma^2 \delta_{i-\ell} \mathbf{I}$$

having exploited the equalities

$$E\left\{\mathbf{d}_m \mathbf{d}_n^H\right\} = \sigma_d^2 \delta_{m-n} \mathbf{I}$$
$$E\left\{\mathbf{w}_i \mathbf{w}_\ell^H\right\} = \sigma^2 \delta_{i-\ell} \mathbf{I}.$$

Similarly

$$E\left\{\mathbf{d}_0 \mathbf{x}_\ell^H\right\} = \sigma_d^2 \mathbf{Q}_\ell^H.$$

As a consequence, by substituting in (28) and remembering the definition of A, we have

$$\sigma_d^2 \sum_{i=0}^{L_c-1} \mathbf{C}_{-i} \sum_m \mathbf{Q}_{i-m\eta} \mathbf{Q}_{\ell-m\eta}^H + \sigma^2 \mathbf{C}_{-\ell} = \left(1 + \frac{\sigma^2}{\sigma_d^2}\right) \sigma_d^2 \mathbf{Q}_\ell^H$$

i.e,

$$\sigma_d^2 \sum_{i=0}^{L_c-1} \mathbf{C}_{-i} \sum_m \mathbf{Q}_{i-m\eta} \mathbf{Q}_{\ell-m\eta}^H + \sigma^2 \mathbf{C}_{-\ell} = \sigma_d^2 \mathbf{Q}_\ell^H + \sigma^2 \mathbf{Q}_\ell^H. \tag{29}$$

Hence, a linear system with $4L_c$ equations and $4L_c$ complex unknown results, whose solution provides the desired coefficients $\mathbf{C}_\ell$. The MMSE can then be computed in closed form by substituting these coefficients in (27).

When $L_c = L$, it results $\mathbf{C}_\ell = \mathbf{Q}_{-\ell}^H$. In order to prove this statement, it is sufficient to verify that when $\mathbf{C}_\ell = \mathbf{Q}_{-\ell}^H$, (29) is satisfied. In this case, in fact

$$\sigma_d^2 \sum_{i=0}^{L-1} \mathbf{Q}_i^H \sum_m \mathbf{Q}_{i-m\eta} \mathbf{Q}_{\ell-m\eta}^H = \sigma_d^2 \mathbf{Q}_\ell^H$$

or equivalently

$$\sum_m \left[\sum_{i=0}^{L-1} \mathbf{Q}_i^H \mathbf{Q}_{i-m\eta}\right] \mathbf{Q}_{\ell-m\eta}^H = \mathbf{Q}_\ell^H.$$

Since $\sum_{i=0}^{L-1} \mathbf{Q}_i^H \mathbf{Q}_{i-m\eta} = \delta_m \mathbf{I}$, the equality is satisfied.

REFERENCES

[1] P. S. Henry, "Error-rate performance of optical amplifiers," presented at the presented at the Optical Fiber Communication Conf., Houston, TX, Feb. 1989, paper THK3, unpublished.
[2] L. G. Kazowsky, S. Benedetto, and A. Willner, *Optical Fiber Communication Systems*. Norwood, MA: Artec House, 1996.
[3] G. P. Agrawal, *Fiber-Optic Communications Systems*, 3rd ed. New York: Wiley, 2002.
[4] L. G. Kazovsky, G. Kalogerakis, and W. T. Shaw, "Homodyne phase-shift-keying systems: Past challenges and future opportunities," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4876–4884, Dec. 2006.
[5] L. G. Kazovsky, "Phase- and polarization-diversity coherent optical techniques," *J. Lightw. Technol.*, vol. 7, no. 2, pp. 279–292, Feb. 1989.
[6] R. Noè, "Phase noise-tolerant synchronous QPSK/BPSK baseband-type intradyne receiver concept with feedforward carrier recovery," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 802–808, Feb. 2005.

[7] R. Noè, "PLL-free synchronous QPSK polarization multiplex/diversity receiver concept with digital I&Q baseband processing," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 887–889, Apr. 2005.

[8] S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent demodulation of optical multilevel phase-shift-keying signals using homodyne detection and digital signal processing," *IEEE Photon. Technol. Lett.*, vol. 18, no. 5, pp. 1131–1133, May 2006.

[9] D.-S. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature-phase-shift-keying signal with carrier phase estimation," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 12–21, Jan. 2006.

[10] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, Sep. 2007.

[11] A. Leven, N. Kaneda, U.-V. Koc, and C. Y.-K. , "Frequency estimation in intradyne reception," *IEEE Photon. Technol. Lett.*, vol. 19, no. 6, pp. 366–368, Mar. 2007.

[12] L. Li, Z. Tao, S. Oda, T. Hoshida, and J. C. Rasmussen, "Wide-range, accurate and simple digital frequency offset compensator for optical coherent receivers," presented at the presented at the Optical Fiber Communication Conf., San Diego, CA, Mar. 2008, paper OWT4, unpublished.

[13] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.

[14] Z. Tao, H. Zhang, A. Isomura, L. Li, T. Hoshida, and J. C. Rasmussen, "Simple, robust, and wide-range frequency offset monitor for automatic frequency control in digital coherent receivers," presented at the presented at the Eur. Conf. on Optical Communication, Berlin, Germany, Sep. 2007, paper 3.5, unpublished.

[15] J. H. Winters, "Equalization in coherent lightwave systems using a fractionally spaced equalizer," *J. Lightw. Technol.*, vol. 8, no. 10, pp. 1487–1491, Oct. 1990.

[16] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "Multilevel optical systems with MLSD receivers insensitive to GVD and PMD," *J. Lightw. Technol.*, vol. 26, no. 10, pp. 1263–1273, May 2008.

[17] D. Crivelli, H. Carter, and M. Hueda, "Adaptive digital equalization in the presence of chromatic dispersion, PMD, and phase noise in coherent fiber optic systems," in *Proc. IEEE Global Telecommun. Conf.*, Nov./Dec. 2004, vol. 4, pp. 2545–2551.

[18] S. J. Savory, G. Gavioli, K. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Opt. Exp.*, vol. 15, no. 5, pp. 2120–2126, Mar. 2007.

[19] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.

[20] G. Charlet, J. Renaudier, M. Salsi, H. Mardoyan, P. Tran, and S. Bigo, "Efficient mitigation of fiber impairments in an ultra-long haul trasmission of 40 Gb/s polarization-multiplexed data, by digital signal processing in a coherent receiver," presented at the presented at the Optical Fiber Communication Conf., Anaheim, CA, Mar. 2007, paper PDP17, unpublished.

[21] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[22] W. J. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.*, vol. 26, no. 3, pp. 385–391, Mar. 1978.

[23] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1376–1385, Sep. 1999.

[24] A. F. Elrefaie, R. E. Wagner, D. A. Atlas, and D. G. Daut, "Chromatic dispersion limitations in coherent lightwave transmission systems," *J. Lightw. Technol.*, vol. 6, no. 5, pp. 704–709, May 1988.

[25] E. Forestieri and L. Vincetti, "Exact evaluation of the Jones matrix of a fiber in the presence of polarization mode dispersion of any order," *J. Lightw. Technol.*, vol. 17, no. 12, pp. 1898–1909, Dec. 2001.

[26] F. Derr, "Coherent optical QPSK intradyne system: Concept and digital receiver realization," *J. Lightw. Technol.*, vol. 10, no. 9, pp. 1290–1296, Sep. 1992.

[27] 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.

[28] H. Meyr, M. Oerder, and A. Polydoros, "On sampling rate, analog prefiltering, and sufficient statistics for digital receivers," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3208–3214, Dec. 1994.

[29] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[30] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a 'stop-and-go' decision directed algorithm," *IEEE Trans. Commun.*, vol. 35, no. 9, pp. 877–887, Sep. 1987.

[31] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers (Applications of Communications Theory)*. New York: Plenum, 1997.

[32] A. Papoulis, *Probability, Random Variables and Sthocastic Processes*. New York: McGraw-Hill, 1991.

[33] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[34] R. Schober and W. H. Gerstacker, "Metric for noncoherent sequence estimation," *IEE Electronics Letters*, vol. 35, no. 25, pp. 2178–2179, Nov. 1999.

[35] R. Schober and W. H. Gerstacker, "Noncoherent adaptive channel identification algorithms for noncoherent sequence estimation," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 229–234, Feb. 2001.

[36] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection Algorithms for Wireless Communications*. New York: Wiley, 2004.

[37] R. Schober and W. H. Gerstacker, "Adaptive noncoherent DFE for MDPSK signals transmitted over ISI channels," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1228–1139, Jul. 2000.

[38] T. Foggi, E. Forestieri, G. Colavolpe, and G. Prati, "Maximum likelihood sequence detection with closed-form metrics in OOK optical systems impaired by GVD e PMD," *J. Lightw. Technol.*, vol. 24, no. 8, pp. 3073–3087, Aug. 2006.

[39] E. Forestieri and G. Prati, "Exact analytical evaluation of second-order PMD impact on the outage probability for a compensated system," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 988–996, Apr. 2004.

[40] H. Kogelnik, L. E. Nelson, J. P. Gordon, and R. M. Jopson, "Jones matrix for second-order polarization mode dispersion," *Opt. Lett.*, vol. 25, no. 1, pp. 19–21, 2000.

Giulio Colavolpe (M'99) was born in Cosenza, Italy, in 1969. He received the Dr. Ing. degree (*cum laude*) in telecommunications engineering from the University of Pisa, Pisa, Italy, in 1994, and the Ph.D. degree in information technologies from the University of Parma, Parma, Italy, in 1998.

Since 1997, he has been with the University of Parma, where he is currently an Associate Professor of telecommunications. He is also with the Italian Interuniversity Consortium for Telecommunications (CNIT) Research Unit, Parma, Italy. In 2000, he was a Visiting Scientist at the Institut Eurécom, Valbonne, France. He is also the Principal Investigator of several research projects funded by the European Space Agency-European Space Research and Technology Centre (ESA-ESTEC) and important telecommunications companies. He is the author or coauthor of several scientific publications in leading international journals and conference proceedings and a few industrial patents. He is also a coauthor of the book *Detection Algorithms for Wireless Communications, With Applications to Wired and Storage Systems* (Wiley, New York, 2004). His current research interests include digital transmission theory, adaptive signal processing, channel coding, and information theory.

Dr. Colavolpe received the best paper award at the 13th International Conference on Software, Telecommunications and Computer Networks (SoftCOM'05) in 2005, and at the IEEE International Conference on Communications (ICC 2008). He is currently serving as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS

Tommaso Foggi was born in Parma, Italy, on May 7, 1978. He received the Dr. Ing. degree in telecommunications engineering in 2003 from the University of Parma, Parma, where he is working toward the Ph.D. degree at the Dipartimento di Ingegneria dell'Informazione (DII).

From July 2003 to July 2004, he was granted a National Interuniversity Consortium for Telecommunications (CNIT) scholarship at the Photonic Networks National Laboratory, Pisa, Italy, and then another CNIT scholarship from July 2005 to July 2006 at the Dipartimento di Ingegneria dell'Informazione (DII), Università Degli Studi di Parma, Parma. He is also with the Italian Interuniversity Consortium for Telecommunications (CNIT) Research Unit, Parma, Italy. His current research interests include electrical compensation of impaired optical transmission systems and satellite navigation and positioning.

Dr. Foggi received the best paper award at the the IEEE International Conference on Communications (ICC 2008).

Enrico Forestieri (S'91–M'92) was born in Milazzo, Italy, in 1960. He received the Dr. Ing. degree in electronics engineering from the University of Pisa, Pisa, Italy, in 1988.

From 1989 to 1991, he was a Postdoctoral Scholar at the University of Parma, Parma, Italy, where he was a Research Scientist and a faculty member from 1991 to 2000, and he was engaged in research on optical communication systems. Since 2001, he has been with the Scuola Superiore Sant'Anna di Studi Universitari e di Perfezionamento, Pisa, Italy, where he is currently a Professor of telecommunications. He is the author or coauthor of numerous scientific publications in leading international journals and conference proceedings, and he also holds a few patents. He is also the Leader of the "Optical Transmission Theory and Techniques" area at the Integrated Research Center for Photonic Networks and Technologies (IRCPhoNeT), Pisa. His current research interests include the general area of digital communication theory and optical communication systems, with special attention to adaptive optical and electronic equalization, channel coding, and advanced modulation formats for optical systems.

Prof. Forestieri was the General Chairman of the Tyrrhenian International Workshop on Digital Communications in 2004.

Giancarlo Prati (M'81–F'03) was born in Rome, Italy, on November 13, 1946. He received the Dr. Ing. degree (*cum laude*) in electronics engineering from the University of Pisa, Pisa, Italy, in 1972.

From 1975 to 1978, he was an Associate Professor of electrical engineering with the University of Pisa. From 1978 to 1979, he was on a North Atlantic Treaty Organization (NATO)-supported Fellowship Leave in the Department of Electrical Engineering, University of Southern California, Los Angeles, where he was engaged in research on optical communications. In 1982, he was a Visiting Associate Professor in the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. From 1976 to 1986, he was a Research Scientist with the Italian National Research Council (CNR) at the Centro di Studio per Metodi e Dispositivi di Radiotrasmissione, Pisa, . From 1986 to 1988, he was a Professor of electrical engineering with the University of Genoa, Genoa, Italy. From 1988 to 2000, he was a Professor of telecommunications engineering at the University of Parma, Parma, Italy, where he was the Dean of the Engineering Faculty from 1992 to 1998. He is currently a Professor of telecommunications with the Scuola Superiore S.Anna, Pisa. His current research interests include telecommunication systems and digital signal processing in communications. He has also focused on optical communications and radiofrequency communications, with application to satellite communications, high-capacity terrestrial digital radio links, mobile radio, modems for switched telephone lines, fiber communications.

He is also the President of Italian Interuniversity Consortium for Telecommunications (CNIT), Pisa, incorporating 35 universities. From 1997 to 2006, he was a member of the Technical Program Committee of the European Conference on Optical Communications (ECOC).