Polarization Modulated Direct Detection Optical Transmission Systems

S. Betti, G. De Marchis, and E. Iannone

Abstract—**In this paper a novel polarization modulated direct detection (PM-DD) system is presented suitable both for binary and multilevel transmission. At the transmitter the optical field is polarization modulated by means of a standard polarization modulator. The receiver is based on the estimation of the Stokes parameters of the received optical field by means of a direct detection optical front end and a baseband electrical processing. The Poincaré sphere rotation induced by the fiber is compensated by means of a pure electronic algorithm and the decision is performed in the Stokes space.**

The system performance is evaluated by means of an analytical model both if the only relevant noise source is the receiver thermal noise and in case that erbium doped optical amplifiers are present so to introduce amplified spontaneous emission (ASE) noise.

The complete compatibility of the proposed system with a direct detection based optical network and the possibility to implement efficient multilevel modulation formats make it promising for applications in the future telecommunication environment. Moreover it can be useful in the design of future parallel supercomputers.

I. Introduction

In the last years the introduction of high efficient, low noise, reliable optical amplifiers based on erbium-doped optical fibers has led to a great improvement in the performance of direct detection optical transmission systems. Due to this new potentiality direct detection systems seem to maintain a central role in the future communication environment.

As a matter of fact the use of optical amplifiers allows long haul trunks to be carried out without signal regenerators and maintaining a simple system structure [1]–[3]. Moreover the large fluorescence bandwidth of erbium doped fibers [4], [5] permits the amplification of WDM signals [6], [7] so that very long distance, large capacity networks can be designed.

However, it is to be noted that with IM-DD systems efficient multilevel modulation formats cannot be implemented. As a matter of fact the information is coded exploiting the optical field intensity so that only multilevel intensity modulation can be adopted. This modulation format is known to be one of the less efficient for digital communication [8], [9] so to be of poor interest for optical applications. On the other hand efficient multilevel modulation offers some important potentialities even for optical telecommunication networks.

Manuscript received December 2, 1991; revised February 24, 1992. Work carried out in the framework of the agreement between Fondazione Ugo Bordoni and the Italian P.T. Administration and under the partial financial support of the National Research Council (CNR) in the frame of the Telecommunications Project.

The authors are with Fondazione Ugo Bordoni, Viale Europa 190–00144, Rome, Italy.

IEEE Log Number 9200836.

First of all the huge optical bandwidth available on conventional single-mode fiber can be exploited by means of frequency division multiplexing (FDM) techniques. On the other hand, the single channel bandwidth is limited by detectors and electronic components and determines the maximum information rate that can be delivered by the network to a single user. A possible solution to improve the capacity of a single optical channel at the expense of receiver sensitivity is multilevel transmission.

Secondly, more sophisticated multilevel codes can be realized that are suitable for power limited systems such as optical transmission systems. These allow an efficient modulation, both in terms of bandwidth and sensitivity, to be achieved at the expense of a greater system complexity. Generally this kind of modulation is based on continuous phase modulation (CPM) [8], [9] but recently a novel technique based on polarization modulation has been proposed [10].

Finally multilevel optical signaling can be of interest for high-speed parallel data transmission between processing units of the same supercomputer. As a matter of fact, at the present in parallel computers the calculation speed bottleneck is often represented by data transmission between processing units [11]. As a matter of fact, if parallel processing is adopted within a single processing unit, a high-speed parallel to serial conversion is required to transmit data using a binary system. This operation is no more required if a multilevel transmission system is used. It is to be noted that multilevel optical system tailored for high-speed data transmission must fulfill quite different requirements with respect to a system embedded in a standard communication environment. First of all, since the transmission distance is quite short (a maximum of a few meters), there are not strong sensitivity constraints. On the other hand systems with a great number of levels (e.g. 1024 or 2048 levels) transmitting at high symbol rate (of the order of 500 M symbols/s or more) must be implemented.

In this paper a novel polarization modulated direct detection (PM-DD) system is presented suitable both for binary and multilevel transmission. At the transmitter the optical field is polarization modulated by means of a standard polarization modulator [12]. The receiver is based on the estimation of the Stokes parameters of the received optical field by means of a direct detection optical front end and a baseband electrical processing. The Poincaré sphere rotation induced by the fiber is compensated by means of a pure electronic algorithm and the decision is performed in the Stokes space.

The system performance is evaluated by means of an analytical model both if the only relevant noise source is the re-

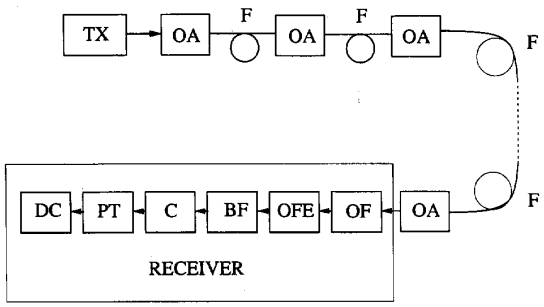

Fig. 1. General block scheme of the PM-DD system with power booster at the transmitter and optical preamplifier at the receiver. These last components can in case be absent (TX = Transmitter, OA = Optical Amplifier, F = Fiber Cable, OF = Optical Filter, OFE = Optical Front End, BF = Baseband Filter, C = Multiplication by the Matrix C, PT = Polarization Tracking, DC = Decision Circuit).

ceiver thermal noise and in the case that erbium doped optical amplifiers are present so to introduce amplified spontaneous emission (ASE) noise.

The complete compatibility of the proposed system with a direct detection based optical network and the possibility to implement efficient multilevel modulation formats make it promising for applications in the future telecommunication environment. Moreover it can be useful in the design of future parallel supercomputers.

In Section II the system structure is presented and its basic operation principle is discussed.

In Section III the system performance is analyzed if the only relevant noise term is the receiver Gaussian thermal noise. The system parameters are optimized so to achieve the lowest error probability and the obtained results are shown and commented on.

In Section IV the performance analysis is carried out in the case that both thermal and ASE noise are present. The system parameters are firstly optimized if the ASE noise is dominant then when the two noise sources are comparable.

Finally the main results are summarized in Section V.

II. System Structure

A. General Structure

The general block diagram of the system is shown in Fig. 1.

The polarization modulated field is generated by a polarization modulator starting from a linearly polarized field emitted by a semiconductor laser.

The modulated field propagates along a fiber link in which a number of erbium doped fiber amplifiers is allocated. Moreover a fiber amplifier can be put either immediately after the transmitter to work as power booster, or before the receiver so to achieve ASE limited receiver working.

An optical filter is placed between the fiber output and the receiver input so to limit the ASE power that affects the receiver performance.

The first stage of the receiver is an optical front end that performs the functions of a direct detection polarimeter. The average values of the four electrical currents at the optical front end output are proportional to four independent linear combinations of the input field Stokes parameters. Indicating with s_k ($k = 0, 1, 2, 3$) the received field Stokes parameters, which are assumed to be normalized in such a way that s_0 is the received optical power and $(s_1)^2 + (s_2)^2 + (s_3)^2 = 1$, and the currents at the optical front end output with j_k ($k = 0, 1, 2, 3$) the following equation holds:

$$\langle j_h \rangle = R a_h^k s_k \tag{1}$$

where R is the responsivity of the photodiodes that perform the optical-to-electrical conversion, $\langle \cdot \rangle$ indicates the ensemble average and the non-singular polarimeter matrix $\boldsymbol{A} = \{a_h^k\}$ characterizes the front end structure [13], [14]. Einstein's convention is used in (1) and it will be adopted in the whole paper. A detailed explanation of this notation and of some related topics is reported in Appendix A.

The vector of the currents generated by the front end is processed by an instantaneous device that multiplies it by the inverse of the matrix $\boldsymbol{A}$, denominated $\boldsymbol{C}$, so to obtain four baseband currents z_k ($k = 0, 1, 2, 3$) whose average values are proportional to the received field Stokes parameters. The current z_0 provides an estimate of the received power s_0 and can therefore be used by the automatic gain control (AGC) but not for the information retrieval. The currents z_1, z_2, and z_3 represent instead the estimates of the components of the vector in the Stokes space associated with the received optical field.

To take into account the random polarization fluctuations induced by the fiber birefringence and coupling, the current vector $Z = \{z_k\}$ is processed by an electronic polarization tracking stage so to obtain an estimate of the Stokes vector of the transmitted field.

After polarization fluctuations compensation, the decision process allows the transmitted symbol to be estimated.

B. Polarization Modulation and Fiber Propagation

Several different implementations have been proposed for polarization modulation systems both in case of binary [15], [16] and multilevel [12] modulation.

In the case of binary modulation, starting from the linearly polarized laser field, the modulator generates two orthogonal polarization states, which correspond to "1" and "0" bits. These states can be represented in the Stokes space as the two extremes of a diameter of the transmitter Poincaré sphere so that this modulation format can be envisaged as "antipodal" in the Stokes space [16].

In the case of multilevel modulation format each transmitted symbol is associated to a polarization state of the transmitted optical field, i.e., to a point in the Stokes space. This point constellation completely defines the modulation format.

In general, the expression of the slowly varying part of the transmitted electrical field is given by

$$\vec{E}_t = A_t \left[\cos \alpha_t \vec{x} + \sin \alpha_t e^{i\delta_t} \vec{y}\right] \tag{2}$$

where the information is coded by means of the parameters A_t, α_t, and δ_t.

During fiber propagation the field state of polarization experiences a slow random evolution due to the fiber birefringence

and coupling so that the output polarization will be different from the transmitted one. Neglecting both fiber dichroism and nonlinear propagation effects, the polarization evolution can be simply represented by means of the time dependent Jones matrix [17], [18], a unitary operator relating the input and the output optical fields.

Moreover a further noise source due to the phenomenon of ASE in optical amplifiers must be considered. Since in all the realistic communication systems the modulation bandwidth is by far narrower than the fluorescence linewidth of the amplifier, ASE noise can be supposed a zero mean, Gaussian distributed white process with equal average power on both polarization components [19], [20].

The optical filter before the receiver input is supposed to be an ideal bandpass filter of bandwidth W_o. The signal distortion induced by this filter is considered negligible so that its only rule is to limit the amount of ASE noise affecting the receiver performance.

With the above assumptions the slowly varying part of the field $\vec{E}$ at the receiver input can be written, as a function of the transmitted field $\vec{E}_t = E_{tx}\vec{x} + E_{ty}\vec{y}$, as follows:

$$\begin{aligned}\vec{E} = e^{-\gamma}\boldsymbol{J}\vec{E}_t + \vec{\eta} &= e^{-\gamma}[\vec{x}\,\vec{y}]\begin{bmatrix} u_1 & u_2 \\ -u_2^* & u_1^* \end{bmatrix}\begin{bmatrix} E_{tx} \\ E_{ty} \end{bmatrix} \\ &\quad + [\vec{x}\,\vec{y}]\begin{bmatrix} \eta_{xI} + i\eta_{xQ} \\ \eta_{yI} + i\eta_{yQ} \end{bmatrix} \\ &= \big[A\cos\alpha\cos\gamma_i + \eta_{xI} \\ &\quad + iA\cos\alpha\sin\gamma_i + i\eta_{xQ}\big]\vec{x} \\ &\quad + \big[A\sin\alpha\cos(\gamma_i+\delta) + \eta_{yI} \\ &\quad + iA\sin\alpha\sin(\gamma_i+\delta) + i\eta_{yQ}\big]\vec{y} \quad (3)\end{aligned}$$

where the complex constant $\gamma = \gamma_r + i\gamma_i$, with γ_r and γ_i real, takes into account all the polarization independent propagation effects and the amplifiers gain, $\boldsymbol{J}$ is the unitary Jones matrix whose complex components u_1 and u_2 are related by $|u_1|^2 + |u_2|^2 = 1$. The vector $\vec{\eta}$ represents the ASE noise whose quadratures η_{xI}, η_{xQ}, η_{yI}, and η_{yQ} are Gaussian, independent, bandlimited white processes of unilateral bandwidth W_o and unilateral power spectral density N_A. The received average optical power is given by $A^2 = e^{-2\gamma_r}A_t^2$ and the parameters α and δ can be easily obtained as functions of α_t and δ_t substituting (2) into (3).

In the above hypothesis the signal to noise ratio Θ_A at the receiver input is given by

$$\Theta_A = \frac{A^2}{N_A W_o}. \quad (4)$$

The described propagation model can be viewed even in the Stokes space since the Stokes vector at the receiver end can be derived from the transmitted one multiplying it by a real unitary operator, the Müller matrix [21], whose elements are functions of u_1 and u_2, and adding to it a noise vector composed by non-Gaussian processes that are nonlinear functions of the ASE noise terms. Therefore the vector in the Stokes space that represents the fiber output field $\{s_k(\alpha)\}$ can be expressed, as a function of that at the fiber input $\{s_k(\tau)\}$, as follows:

$$s_L(\alpha) = m_L^H s_H(\tau) + v_L \quad (5)$$

where $\boldsymbol{M} = \{m_L^H\}$ is the Müller matrix and v_L the noise terms. The uppercase indexes range from 1 to 3 to distinguish them from the lowercase indices, which range from 0 to 3, and the explicit expressions of the Müller matrix elements can be found in [21].

C. Optical Front End

Conceptually there are several ways to carry out a direct detection polarimeter that can be used as an optical front end in a PM-DD system; for example by realizing a suitable interferometer before optical-to-electrical conversion, by analyzing the incoming field by means of quarter and half wave plates and polarization beam splitters or exploiting the property that, if a light beam incides on a partially reflecting surface, the normal and parallel polarization components are differently reflected and transmitted.

A block scheme of an interferometric polarimeter is shown in Fig. 2. The input electrical field is divided by means of a polarization beam splitter into its linear polarization components and then these are combined by means of a couple of standard beam splitters of splitting ratio α_s, a couple of standard beam splitters of splitting ratio β_s and a couple of polarization beam splitters as shown in the figure. If the optical path is accurately controlled, the two polarization components are combined in phase before detection of the second photodiode and in quadrature on the third one.

The matrix of the polarimeter, as defined by (1), is given by (6) (see bottom of page).

A direct detection polarimeter based on the analysis of the incoming field by means of half and quarter wave plates is reported and analyzed in detail in [32]. This structure may result more stable and feasible than that based on an interferometer so to be promising for practical system implementation. The matrix $\boldsymbol{A}$ relative to this implementation of the optical

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{\alpha_s^2} & 0 & 0 & \frac{1}{\alpha_s^2} \\ \frac{1}{\alpha_s^2} & 0 & 0 & -\frac{1}{\alpha_s^2} \\ -\frac{\beta_s}{\alpha_s^2\sqrt{1-\beta_s^2}} & \frac{1}{(1-\alpha_s^2)\beta_s\sqrt{1-\beta_s^2}} & 0 & -\frac{\sqrt{1-\beta_s^2}}{\alpha_s^2\beta_s} \\ -\frac{\sqrt{1-\beta_s^2}}{\alpha_s^2\beta_s} & 0 & \frac{1}{(1-\alpha_s^2)\beta_s\sqrt{1-\beta_s^2}} & \frac{\beta_s}{\alpha_s^2\sqrt{1-\beta_s^2}} \end{bmatrix}. \quad (6)$$

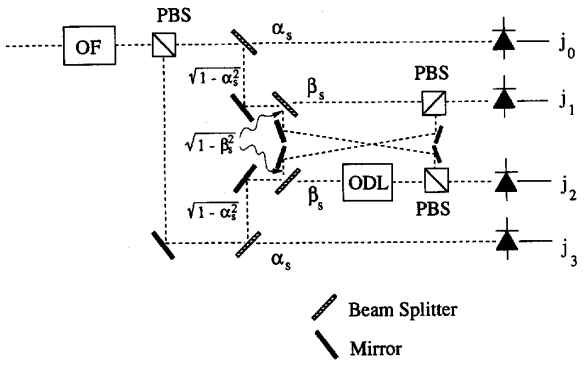

Fig. 2. Interferometric optical front end block scheme (OF = Optical Filter, PBS = Polarization Beam Splitter, ODL = Optical Delay Line).

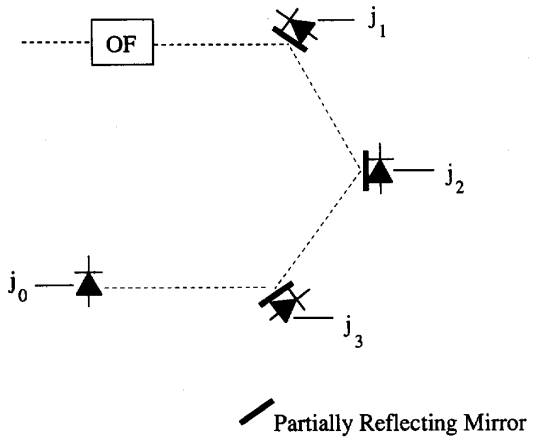

Fig. 3. Reflectometric optical front end block scheme (OF = Optical Filter).

front end depends on the power splitting ratio of the two beam splitters. It is similar to that reported in (6) in that $a_0^1 = a_0^2 = a_1^1 = a_1^2 = a_2^2 = a_3^1 = 0$, while the dependence of the nonzero elements on the two parameters is different.

The structure of the polarimeter based on reflection properties has been widely analyzed in literature [13], [14] since it is used for optical laboratory measures and optical sensors. The main advantage of such a scheme is represented by minor constraints due to the optical path control which is instead a severe requirement of the interferometric polarimeter. Its schematic block diagram is shown in Fig. 3. Three photodetectors PD_0, PD_1, and PD_2 are coated with a reflection coating characterized by reflection parameters r_k, Ψ_k, and Δ_k $(k = 0, 1, 2,)$ while the fourth photodetector, PD_3, is antireflection coated so that all the optical power incident on it is detected. In particular, r_k indicates the reflectance of the surface for circular or unpolarized light, $\tan \Psi_k \exp(i\Delta_k)$ is the ratio between the complex reflection coefficient for linearly polarized light along parallel and perpendicular directions to the local incidence plane. It can be demonstrated that (1) holds even for this polarimeter [14]: in this case the matrix $\boldsymbol{A}$ can be expressed as function of the above parameters and the angles between the light incidence planes and the photodetector surfaces.

D. Electrical Processing

The electrical processing needed to perform the decision correctly can be divided into three steps: a baseband filtering of the four currents at the output of the optical front end, the inversion of the matrix $\boldsymbol{A}$ and the polarization tracking.

The four baseband filters are assumed to be integrators with integration time T_s equal to the inverse of the symbol rate R_s. Moreover the signal distortion induced by the filters is neglected.

Since the matrix of the polarimeter is function only of the system structure, the electrical block that multiplies the vector $J = \{j_k\}$ by the matrix $\boldsymbol{C} = \boldsymbol{A}^{-1}$ can be easily realized. At the output of this circuit a vector $Z = \{z_k\}$ is obtained whose components are the estimates of the Stokes parameters of the received field.

In order to track the fiber induced polarization fluctuations, which can be represented as a slow random rotation of the Poincaré sphere in the Stokes space [15], two different classes of algorithms can be used:

- algorithms that use the received signal to estimate the channel transfer matrix, in this case the Müller matrix that characterizes the polarization evolution in the Stokes space;
- algorithms that track the position of the points representing the transmitted symbols (reference points) in the Stokes space so to follow the Poincaré sphere rotation.

Both classes of algorithms have been studied both for binary [22], [23] and multilevel systems [12], [24] in the case of coherent communication systems and they can be applied to this case without any modifications. In the following it will be assumed that an algorithm of the second class be used so that the estimate of the Stokes vector, to be compared with the reference points, is proportional to Z. Anyway it is demonstrated that in the case of coherent systems the choice of the algorithm for the polarization tracking does not affect the system performance if the tracking is carried out without errors. Even if this assumption cannot be directly verified in the case of PM-DD, it seems that the dependence of the performance on the tracking algorithm, if any, should not be prominent.

E. Decision Process

Starting from the estimate of the Stokes parameters of the transmitted field Ξ and from the knowledge of the position of the reference points in the Stokes space, the decision circuit has to provide an estimate of the transmitted symbol. Assuming that a modulation without memory is applied [12], the only kind of modulation considered in this paper, a distance based decision criterion can be used. This means that the nearest reference point to Ξ in the Stokes space is supposed to represent the transmitted symbol.

In general, assuming that a multilevel transmission with N symbols is performed, this is equivalent to calculate the N decision variables $Q(\beta)$'s given by

$$\begin{aligned} Q(\beta) &= \left[\xi^R - s^R(\beta)\right]\left[\xi_R - s_R(\beta)\right] \\ &= \xi^R\xi_R + s^R(\beta)s_R(\beta) - 2s^R(\beta)\xi_R \end{aligned} \tag{7}$$

and to decide that the α-th symbol has been transmitted if $Q(\alpha) < Q(\beta) \quad \forall \beta \neq \alpha$.

III. Performance Evaluation in the Absence of Optical Amplifiers

A. Expression of the Bit Error Probability

Dealing with a multilevel modulation without memory, the first step in calculating the system bit error probability P_e is to relate it to the symbol error probability P_s. The exact expression of P_e as a function of P_s is quite involved and depends on a large number of parameters, however, assuming for sake of simplicity $N = 2^M$, a simple and accurate approximation can be provided by means of the expression $P_e = (\bar{e}Ps)/M$, where $\bar{e}$ is the average number of wrong bits in a wrong symbol. The approximation leads to an upperbound for the bit error probability if the symbols are codified so that binary words that differ only for few bits correspond to symbols neighbors in the Stokes space.

The average number of wrongly estimated bits in a wrong symbol is given by

$$\bar{e} = \frac{1}{2^M - 1} \sum_{k=1}^{M} \binom{M}{k} k = \frac{M 2^{M-1}}{2^M - 1} = \frac{MN}{2(N-1)} \tag{8}$$

so that

$$P_e = \frac{N}{2(N-1)} P_s. \tag{9}$$

In order to calculate the symbol error probability the following general expression can be used

$$P_s = \sum_{\alpha=1}^{N} \pi(\alpha) P_s(\alpha) \tag{10}$$

where $\pi(\alpha)$ is the probability of transmitting the α-th symbol and $P_s(\alpha)$ the symbol error probability conditioned to the transmission of the α-th symbol. The evaluation of the terms of (10) requires the computation of the probability density function of the decision variables conditioned to the transmitted symbol, the identification in the Stokes space of the reference points constellation and of the decision zone around each point and then the integration of the probability density function in each decision zone [25]. This procedure is quite involved even in the case of N-SPSK coherent receiver [26] for the complex expression of the noise terms in the Stokes space and in all those cases it requires a particular reference points constellation to be chosen. The above arguments lead to apply the union bound approximation [25], [17] to evaluate $P_s(\alpha)$. It leads to an expression of the symbol error probability that holds for every reference points constellation and, at the low error rates typical of optical systems, it is quite tie [27].

In the considered case the union bound approximation is expressed by the following formula:

$$\begin{aligned} P_s(\alpha) &= \sum_{\beta=1,\beta\neq\alpha}^{N} Pr\{Q(\beta) < Q(\alpha)/\alpha\} \\ &= \sum_{\beta=1,\beta\neq\alpha}^{N} Pr\{q(\alpha,\beta) < 0\} \end{aligned} \tag{11}$$

where $q(\alpha,\beta) = Q(\beta) - Q(\alpha)$ if the α-th symbol is transmitted.

In the following paragraph the bit error probability will be evaluated starting from (9)–(11).

B. Error Probability in the Union Bound Approximation

In the absence of optical amplifiers the input field is deterministic and the most relevant noise source is the thermal noise at the receiver. In this case the four currents after the optical front end and the baseband filters have the expression

$$j_h = RA^2 a_h^k s_k(\alpha) * h(t) + n_h' * h(t) \approx RA^2 a_h^k s_k(\alpha) + n_h \tag{12}$$

where $s_h(\alpha)$ $(h = 0,1,2,3)$ are the Stokes parameters of the received field that coincide, under the hypothesis of perfect polarization tracking, with the coordinates of the point in the Stokes space that represents the α-th symbol. The points representing the transmitted symbols in the Stokes space at the receiver will be indicated as reference points. The pulse response of the baseband filters, which are assumed all perfectly equal, is indicated with $h(t)$. It is supposed that the baseband filters are ideal with a bandwidth large enough to neglect signal distortion. This assumption justifies the approximation performed in (12). Moreover the noise terms n_h, obtained by filtering the thermal noise contributions n_h', result to be zero mean, bandlimited, Gaussian processes whose power is given by

$$\sigma_n^2 = \frac{4KT_k}{R_c} B_n F_a \tag{13}$$

where K is the Boltzman constant, T_k the absolute temperature, R_c the load resistance of the electrical front end following the photodiodes, B_n the noise bandwidth and F_a the noise figure of the amplifier of the electrical front end. The definition of the noise figure adopted in this paper is implicit in (13): it is the ratio between the noise power spectral density at the amplifier output and the thermal power spectral density that can be calculated starting from the room temperature and the amplifier input resistance.

The components of the vector Z are given by

$$z_L = c_L^h j_h = RA^2 s_L(\alpha) + c_L^h n_h \tag{14}$$

where the uppercase index ranges from 1 to 3 while the lowercase one from 0 to 3.

Starting from its definition the expression of $q(\alpha,\beta)$ can be easily calculated, obtaining

$$\begin{aligned} q(\alpha,\beta) &= \left[s^L(\beta) - \frac{1}{R} z^L\right]\left[s_L(\beta) - \frac{1}{R} z_L\right] \\ &\quad - \left[s^L(\alpha) - \frac{1}{R} z^L\right]\left[s_L(\alpha) - \frac{1}{R} z_L\right] \\ &= \left[s_0(\beta)\right]^2 - \left[s_0(\alpha)\right]^2 - \frac{2}{R} z^L \left[s_L(\beta) - s_L(\alpha)\right] \end{aligned} \tag{15}$$

where $s_L(\alpha)$ and $s_L(\beta)$ are the Stokes parameters of the α-th and the β-th reference point respectively.

Limiting the analysis to the "pure polarization modulation" in which the power of the transmitted field remains constant during transmission, from (15) it can be obtained

$$\Pr\{q(\alpha,\beta)<0\}=\Pr\left\{\frac{2}{R}z^L\left[s_L(\beta)-s_L(\alpha)\right]>0\right\}. \quad (16)$$

Since in the above approximations Z is a Gaussian vector the probability in (16) can be easily calculated. In particular the signal average S is given by

$$S=\left\langle\frac{2}{R}z^L\left[s_L(\beta)-s_L(\alpha)\right]\right\rangle=A^2d^2(\alpha,\beta) \quad (17)$$

where A^2 is the received optical power and $d(\alpha,\beta)$ is the distance in the Stokes space between the α-th and the β-th reference points.

The signal variance σ_s^2 can be similarly computed obtaining

$$\sigma_s^2=\frac{\sigma_n^2}{R^2}\left[s^L(\beta)-s^L(\alpha)\right]\left[s_H(\beta)-s_H(\alpha)\right]c_L^jc_j^H. \quad (18)$$

From (18) it can be noted that the additive noise variance is related to dependence on time of the polarization fluctuations. As a matter of fact the reference points in the Poincaré sphere are updated by the polarization tracking algorithm in order to compensate the polarization fluctuations at the fiber output. It means that, indicating the coordinates of the reference points at a given starting instant with $\hat{s}_k(\gamma)$ $(\gamma=1,\cdots,N)$ and the Müller matrix as $\boldsymbol{M}=\{m_H^L\}$, (18) can be rewritten as

$$\sigma_s^2=\frac{\sigma_n^2}{R^2}m_F^H\left[\hat{s}^F(\beta)-\hat{s}^F(\alpha)\right]c_H^jm_L^I\left[\hat{s}_I(\beta)-\hat{s}_I(\alpha)\right]c_j^L. \quad (19)$$

Finally the single term of the union bound, expressed by (11), can be calculated, starting from (17) and (18), by means of the following expression:

$$\Pr\{q(\alpha,\beta)<0\}=\operatorname{erfc}\left(\frac{S}{\sqrt{2}\sigma_s}\right). \quad (20)$$

C. Comments of the Results

In this paragraph the performance of a PM-DD system based on an interferometric optical front end is analyzed. Only this kind of front end will be considered in this paper since it can be optimized only acting on two parameters, α_s and β_s, that is on the power splitting ratio of the beam splitters of the front end. Anyway the performance of PM-DD systems using the optical front end based on quarter and half wave plate is equivalent to that of systems using the interferometric front end. The only difference is that the optimum values of the front end parameters, that is the splitting ratio of the beam splitters, are different. About the front end configuration based on the reflection properties of partially reflecting surfaces, it is simpler to be implemented but the system optimization is by far more complex, due to the great number of front end parameters and, from the first results, it seems to offer worse performance.

Even if the union bound gives a general expression of the error probability, evaluating the bit error probability requires a reference points constellation to be fixed. The reference points constellation could be optimized to obtain the lowest error probability for a given number of points. However this is a quite difficult goal to be achieved since the system performance depends not only on the points constellation but also on a lot of other parameters. A simpler choice is to consider the constellation characterized by the largest distance between first neighbors. For example, for a binary system this is equivalent to choose antipodal points in the Stokes space, which even represents the optimum choice for this case. For $N=4$, $N=8$, and $N=16$ the spherical coordinates of the points and the distance matrixes for the three constellations are reported in Tables I to VI [12].

The values of the other system parameters assumed in the following are reported in Table VII.

As shown by (19) the system performance depends on the elements of the Müller matrix. Therefore two main parameters can be introduced to characterize the front end behavior against the fiber birefringence and coupling fluctuations: the worst case bit error probability P_{ew}, for a given input optical power, and the sensitivity fluctuation ΔS for a reference error probability of 10^{-9}. The optimum front end parameters with respect to ΔS and P_{ew} are reported in Tables VIII and IX, respectively for $N=2$, 4, 8, and 16. Moreover for low error probabilities the optimum value of α_s and β_s with respect to P_{ew} can be shown not to depend on the input power.

From the above tables it can be noted that, if the front end structure is carefully chosen, the sensitivity changes due to fiber birefringences and coupling fluctuations can be maintained within 1 dB. Moreover the optimum values with respect to ΔS and P_{ew} are not the same so that the front end can be optimized only with respect to one parameter. In the following the optimization with respect to P_{ew} is chosen.

The worst case error probability is shown against the received optical power in Fig. 4 for a binary PM-DD system and, for sake of comparison, for a conventional IM-DD system under the same hypotheses and reporting on the abscissa axis the received peak power. The sensitivity of the PM-DD system for an error probability of 10^{-9} is about 2.8 dB better than that of the IM-DD system, since a binary polarization modulated signal can be conceptually assumed as equivalent to a couple of in-quadrature intensity modulated signals carried by orthogonal channels. This observation would lead to foresee a sensitivity gain of 3 dB for a PM-DD system, however the actual gain of 2.8 dB when considering the worst case sensitivity can be caused by the dependence of the PM-DD receiver sensitivity on the received field state of polarization.

The worst case error probability is shown against the received optical power for multilevel systems with $N=4$, 8, and 16 in Fig. 5. It is to be pointed out that the worst case sensitivity is –30.8 dBm for $N=2$ and –31.4 dBm for $N=4$ so that the performance of 4-level system result better than that of the binary system both in terms of optical bandwidth and power sensitivity. This can be explained by the following observations:

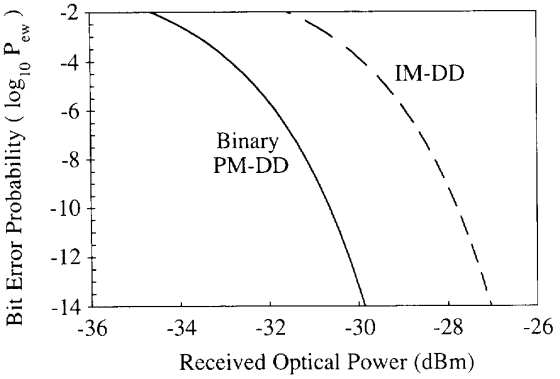

Fig. 4. Worst-case bit error probability (P_{ew}) for a binary system in the case of thermal noise limited operation versus the input optical power. The optical front end is interferometric and its parameters are chosen so to minimize the worst case error probability. The values of the other system parameters are reported in Table VII. The error probability for a conventional IM-DD system is also shown for comparison under the same conditions. In this case the received peak power is shown on the abscissa axis.

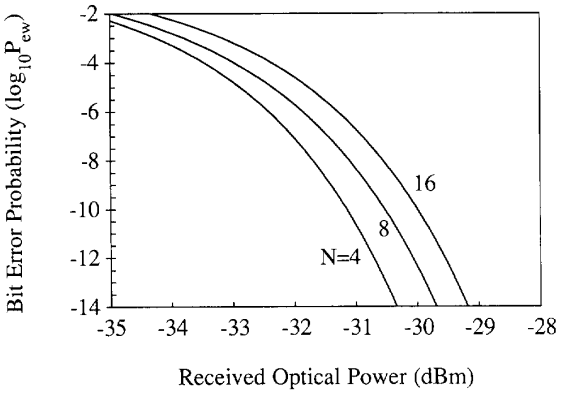

Fig. 5. Worst-case bit error probability (P_{ew}) for systems with $N = 4$, 8, and 16 in the case of thermal noise limited operation versus the input optical power. The optical front end is interferometric and its parameters are chosen so to minimize the worst case error probability. The values of the other system parameters are reported in Table VII.

- the normalized distance between first neighbors decreases from 2 to 1.63 when N ranges from 2 to 4, as shown in Table II; this induces a penalty of 1.9 dB for the 4-level system as shown from (17) and (20);
- the electrical noise bandwidth ranges from R_b to $R_b/2$ so to induce a 3-dB gain in the case of 4-level system;
- in the case of binary constellation each point has only one first neighbors while in the case of 4-level system they are three so to induce a sensitivity penalty for the 4-level system of about 0.18 dB for $P_{ew} = 10^{-9}$;
- in the expression of the noise variance a constellation form factor is present, given by $m_F^L[\hat{s}^F(\beta) - \hat{s}^F(\alpha)]\, c_L^j m_H^I[\hat{s}_I(\beta) - \hat{s}_I(\alpha)]c_j^H$, which is 1.08 times greater in the case of 4-level system than in the binary case, inducing a penalty of 0.32 dB.

From the above considerations the sensitivity gain of 4-level system with respect to binary system is easily stated.

IV. Performance in the Presence of Optical Amplifiers

A. Error Probability in the Union Bound Approximation

In the presence of optical amplifiers the expression of the error probability, obtained by using the union bound approximation, as a function of the probabilities $\Pr\{q(\alpha,\beta) < 0\}$, maintains unchanged. However, with reference to the previous analysis, another noise source is introduced: the ASE from optical amplifiers.

In this case the optical field at the receiver input is the sum of the signal field and the ASE noise so that it can be represented by the random process described in (3). Indicating with $n_k(t)$ the thermal noise contributions before the baseband filters, indicated with n'_h in Section III, the four electrical currents after baseband filtering can be written as

$$j_h = RA^2 a_h^k \xi_k * h(t) + n_h * h(t) \tag{21}$$

where the components of the vector $\Xi = \{\xi_k\}$ are the transmitted Stokes parameters, modified by the fiber induced polarization fluctuations and corrupted by ASE due to optical amplifiers, $h(t)$ is the pulse response of the baseband filter.

The vector Z is given by

$$z_H = c_H^h j_h = \left[RA^2 \xi_H + c_H^h n_h\right] * h(t) \tag{22}$$

so that the argument of the probability at the last member of (11) becomes, in the case of "pure polarization modulation,"

$$q(\alpha,\beta) = \left\{ A^2 \xi^H \left[s_H(\beta) - s_H(\alpha)\right] + \frac{1}{R}\left[s^H(\beta) - s^H(\alpha)\right] c_H^k n_k \right\} * h(t). \tag{23}$$

Since the noise terms n_k are Gaussian random processes, the second addendum of $q(\alpha,\beta)$ is a Gaussian variable while the first one is not Gaussian since it contains, for example, the squares of the ASE quadratures. In this situation the simple theory used in the previous section, where the decision variable distribution is supposed Gaussian, cannot be any more applied. To obtain an approximate expression of the error probability it is useful to express the components of the vector Ξ as functions of the four quadratures $x_k (k = 0, 1, 2, 3)$, which completely characterize the received field [24].

Starting from (3) the vector $X = \{x_k\}$ is given by

$$X = \begin{pmatrix} A\cos\alpha\cos\gamma_i + \eta_{xI} \\ A\cos\alpha\sin\gamma_i + \eta_{xQ} \\ A\sin\alpha\sin(\gamma_i + \delta) + \eta_{yI} \\ A\sin\alpha\sin(\gamma_i + \delta) + \eta_{yQ} \end{pmatrix} \tag{24}$$

and the vector Ξ can be expressed, as a function of the components of X, as

$$\Xi = \begin{pmatrix} (x_0)^2 + (x_1)^2 + (x_2)^2 + (x_3)^2 \\ (x_0)^2 + (x_1)^2 - (x_2)^2 - (x_3)^2 \\ 2(x_0x_2 + x_1x_3) \\ 2(x_0x_3 - x_1x_2) \end{pmatrix}. \tag{25}$$

Equation (25) can be more usefully rewritten by introducing the two covariant, one contravariant tensor $\boldsymbol{B} = \{b^r_{hk}\}$, whose nonzero components are given by

$$\{b^r_{hk}\} = \left\{ \begin{array}{l} b^0_{00} = b^1_{10} = b^2_{20} = b^3_{30} = 1 \\ b^0_{01} = b^1_{11} = -b^2_{21} = -b^3_{31} = 1 \\ b^2_{02} = b^0_{22} = b^3_{12} = b^1_{32} = 1 \\ b^3_{03} = b^0_{33} = -b^2_{13} = -b^1_{23} = 1 \end{array} \right\}. \tag{26}$$

By the above definitions the following equation holds:

$$\xi_h = x^k b^r_{kh} x_r. \tag{27}$$

Substituting (27) into (23) and assuming

$$f^r_k(\alpha,\beta) = A^2\Big[s^H(\beta) - s^H(\alpha)\Big]b^r_{kH}$$
$$\mu = \frac{1}{R}\Big[s^H(\beta) - s^H(\alpha)\Big]c^k_H n_k \tag{28}$$

(23) can be rewritten as

$$q(\alpha,\beta) = \big[x^k f^r_k(\alpha,\beta)x_r + \mu\big] * h(t) \tag{29}$$

that gives an expression of $q(\alpha,\beta)$ as function of only Gaussian variables, i.e., the input field quadratures and the additive noise.

In order to carry out an estimate of the system error probability starting from (29), the convolution with the baseband filter pulse response must be calculated. To obtain a simpler analytical model the baseband integrator is supposed to be a digital filter with integration time T_s equal to the inverse of the symbol rate and with sampling time $T_o = 1/2W_o$ related to the system optical bandwidth. Under such an assumption equation (29) can be rewritten as

$$q(\alpha,\beta) = \sum_{\zeta=1}^{n_s} v(t_\zeta) + v_n$$
$$= \sum_{\zeta=1}^{n_s} x^k(t_\zeta) f^r_k(\alpha,\beta) x_r(t_\zeta) + \sum_{\zeta=1}^{n_s} \mu(t_\zeta) \tag{30}$$

where n_s is the number of input signal samples during the integration time, i.e., the integer part of $W_o T_s$, the instants t_ζ are the sampling instants during the integration time, v_n is the noise term after integration and $v(t_\zeta)$ are quadratic forms of the input field quadratures.

The characteristic matrix $f^h_k(\alpha,\beta)$ of the quadratic forms $v(t_\zeta)$ depends only on the received power and on the α-th and β-th reference points coordinates in the Stokes space. It is simple to demonstrate that it is a Hermitian matrix. The characteristic function of Hermitian quadratic forms of real Gaussian random variables can be calculated in closed form [28], [29] obtaining

$$G_v(\lambda) = \frac{\exp\left\{-\frac{A^2 d^2(\alpha,\beta)}{4N_A W_0}\lambda \frac{1-4\lambda^2}{1-\lambda^2 d^2(\alpha,\beta)}\right\}}{\big[1-\lambda^2 d^2(\alpha,\beta)\big]^2} \tag{31}$$

where $\lambda = \lambda_r + i\lambda_i$ is the Laplace variable associated with $v(t_\zeta)$, normalized multiplying it by $2N_A W_o$.

Since the addenda of the first summation into (30) are obtained at a sampling rate $T_o = 1/2W_o$, they can be assumed to be independent so that the characteristic function of the summation is the product of the characteristic functions of the addenda. Moreover v_n is a Gaussian random variable with zero mean and variance given by (18) so that its characteristic function, taking into account the normalization adopted in defining the variable λ, is given by

$$G_{v_n}(\lambda) = \exp\left\{\frac{\sigma_s^2}{8N_A^2 W_o^2}\lambda^2\right\}. \tag{32}$$

Combining (31), (32) and the observation that the samples $v(t_\zeta)$ are statistically independent, the characteristic function of $q(\alpha,\beta)$ can be evaluated obtaining

$$G_q(\lambda) = \frac{\exp\left\{-\frac{A^2 d^2(\alpha,\beta) n_s \lambda}{4N_A W_0}\frac{1-4\lambda^2}{1-\lambda^2 d^2(\alpha,\beta)} + \frac{\sigma_s^2\lambda^2}{8W_o^2 N_A^2}\right\}}{\big[1-\lambda^2 d^2(\alpha,\beta)\big]2_{n_s}}. \tag{33}$$

It is to be noted that the characteristic function given by (33) depends only on four parameters:

- $d(\alpha,\beta)$ — distance between the α-th reference point that corresponds to the transmitted symbol and the β-th reference point;
- $n_s = W_o T_s$ — the ratio between the optical and the electrical bandwidth;
- Θ_A — signal-to-noise ratio at the input of the optical front end;
- $\Theta_N = \sigma_s^2/N_A^2 W_o^2$ — additive noise power to ASE noise equivalent current power ratio.

The first parameter characterizes the reference points constellation that has been chosen, the second one is a system design parameter, the third one characterizes the noise introduced by the optical amplifiers and the fourth one is a comparison measure of the influence of the additive noise at the receiver and the ASE noise on the system performance. In particular, if Θ_N is by far greater than one, the ASE noise is negligible with respect to the additive noise while the opposite situation occurs when $\Theta_N << 1$.

Expressing the characteristic function directly as function of the above four parameters the following equation is obtained

$$G_q(\lambda) = \frac{\exp\left\{-\frac{\Theta_A d^2(\alpha,\beta) n_s \lambda}{4}\frac{1-4\lambda^2}{1-\lambda^2 d^2(\alpha,\beta)} + \frac{\Theta_N}{8}\lambda^2\right\}}{\big[1-\lambda^2 d^2(\alpha,\beta)\big]2_{n_s}}. \tag{34}$$

To compute the conditional symbol error probability $P_s(\alpha)$, the addenda into equation (11) can be calculated starting from the characteristic function of $q(\alpha,\beta)$ by means of the Cauchy formula [29]

$$\Pr\{q(\alpha,\beta) < 0\} = -\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{G_q(\lambda)}{\lambda - i\varepsilon}\, d\lambda \tag{35}$$

where the integration path is the imaginary axis and the ε at the denominator means that the pole in the origin must be kept on the right of the integration path.

TABLE I
SPHERICAL COORDINATES IN THE STOKES SPACE OF THE POINTS OF THE 4-LEVEL CONSTELLATIONS. USING NORMALIZED STOKES PARAMETERS SUCH THAT $(s_1)^2+(s_2)^2+(s_3)^2=1$ THE ANGLES Φ AND Ψ ARE DEFINED AS FOLLOWS:

$$\Phi = \arccos(s_1) = \arcsin(s_2) \qquad (-\pi < \Phi \leq \pi)$$

$$\Psi = \arccos(s_3) \qquad (0 < \Psi \leq \pi)$$

N	Φ	Ψ
1	0	0
2	250.56	0
3	125.29	−54.73
4	125.21	54.72

The integral into (35) cannot be solved analytically so that the system error probability must be calculated either using numerical methods or by introducing some approximations.

In this paper the second way is adopted. The low value of the standard error probability $\left(P_e = 10^{-9}\right)$, used to define the system sensitivity S, makes possible to evaluate S with a good accuracy using methods that are asymptotically exact in the limit of high signal-to-noise ratio. Moreover the use of approximate methods has the advantage to provide an analytical expression for the approximate error probability that is easier to be used in practice than a numerical solution of (35).

The adopted approximation is based on the saddle point approximation (SAP) [29]–[31] a well known method to obtain a tight upper bound for an error probability expressed as the integral of the tail of the probability density function of the decision variable.

An estimation of the accuracy of the error probability approximation, provided by the SAP in the present case, can be obtained computing the bit error probability by means of this approximation when only the thermal noise is relevant, and then comparing the results with those provided in the previous section. Such condition is obtained passing to the limit for $\Theta_N \to \infty$ while Θ_A^2/Θ_N, i.e., the signal to thermal noise ratio, is maintained constant. The sensitivity error ΔS induced by the SAP is within 0.2 dB for $N = 2$, 4, 8, and 16, so that the adopted approximation results quite good.

B. Comments on the Results

A first evaluation of the PM-DD system performance in the presence of optical amplification can be drawn considering a receiver based on an optical fiber preamplifier, placed in front of the optical front end. In this case the optical power incident on the photodiodes is quite high so that thermal noise effect can be neglected. This is equivalent to assume $\Theta_N = 0$ into (34) and in the following derivation of the bit error probability.

Under such an assumption the system performance is independent both of the fiber birefringence and coupling fluctuations and the optical front end parameters. As a consequence both the interferometric and the reflectometric front ends have the same performance and, once a front end structure has been chosen, its parameters can be set only on the basis of technological requirements, without affecting the system performance.

TABLE II
DISTANCE MATRIX FOR THE 4-LEVEL CONSTELLATION

	1	2	3	4
1	0.000	1.633	1.633	1.633
2	1.633	0.000	1.633	1.633
3	1.633	1.633	0.000	1.633
4	1.633	1.633	1.633	0.000

TABLE III
SPHERICAL COORDINATES IN THE STOKES SPACE OF THE POINTS OF THE 8-LEVEL CONSTELLATION. THE DEFINITIONS OF Φ AND Ψ ARE REPORTED IN THE CAPTION OF TABLE I

N	Φ	Ψ
1	0	0
2	218.15	0
3	288.44	−37.74
4	−70.31	37.67
5	59.889	−59.91
6	73.92	14.67
7	144.22	−14.67
8	158.28	59.97

TABLE IV
DISTANCE MATRIX FOR THE 8-LEVEL CONSTELLATION

	1	2	3	4	5	6	7	8
1	0.00	1.89	1.21	1.21	1.21	1.21	1.89	1.71
2	1.89	0.00	1.21	1.21	1.71	1.89	1.21	1.21
3	1.21	1.21	0.00	1.21	1.21	1.89	1.71	1.89
4	1.21	1.21	1.21	0.00	1.89	1.71	1.89	1.21
5	1.21	1.21	1.21	1.89	0.00	1.21	1.21	1.89
6	1.21	1.89	1.89	1.71	1.21	0.00	1.21	1.21
7	1.89	1.21	1.71	1.89	1.21	1.21	0.00	1.21
8	1.71	1.21	1.89	1.21	1.89	1.21	1.21	0.00

TABLE V
SPHERICAL COORDINATES IN THE STOKES SPACE FOR THE 16-LEVEL CONSTELLATION. THE DEFINITIONS OF THE ANGLES Φ AND Ψ ARE GIVEN IN THE CAPTION OF TABLE I

N	Φ	Ψ
1	0.	0
2	180.4	0
3	218.2	−37.3
4	−49.2	17.6
5	−38.9	−36.1
6	50.2	11.2
7	35.2	−40.0
8	177.1	51.9
9	269.4	−13.5
10	91.6	−20.0
11	283.4	−86.7
12	82.0	56.1
13	147.0	−41.1
14	298.4	68.6
15	129.3	15.9
16	230.5	20.8

Under the above condition the system performance has been calculated assuming the points constellations in the Stokes space detailed in Tables I–VI, a bit rate $R_b = 1$ Gb/s and an electrical baseband filter bandwidth equal to the symbol

TABLE VI
DISTANCE MATRIX FOR THE 16-LEVEL CONSTELLATION

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.00	2.00	1.80	0.86	0.86	0.86	0.86	1.80	1.42	1.42	1.42	1.35	1.80	1.28	1.80	1.80
2	2.00	0.00	0.86	1.80	1.80	1.80	1.80	0.86	1.42	1.42	1.42	1.46	0.86	1.53	0.90	0.90
3	1.80	0.86	0.00	1.56	1.24	1.94	1.56	1.48	0.86	1.56	0.86	1.90	0.90	1.74	1.50	0.90
4	0.86	1.80	1.56	0.00	0.90	1.46	1.50	1.52	0.86	1.90	1.583	1.48	1.94	0.86	1.90	1.21
5	0.86	1.80	1.24	0.90	0.00	1.46	0.95	1.93	0.86	1.60	0.86	1.85	1.56	1.59	1.96	1.56
6	0.86	1.80	1.94	1.46	1.46	0.00	0.90	1.56	1.90	0.86	1.56	0.86	1.55	1.37	1.24	1.92
7	0.86	1.80	1.56	1.50	0.95	0.90	0.00	1.93	1.60	0.86	0.86	1.56	1.24	1.80	1.56	1.95
8	1.80	0.86	1.50	1.52	1.93	1.56	1.93	0.00	1.56	1.56	1.90	0.86	1.50	0.86	0.6	0.86
9	1.42	1.42	0.86	0.86	0.86	1.90	1.60	1.56	0.00	1.90	1.21	1.86	1.56	1.34	1.90	0.86
10	1.42	1.42	1.56	1.90	1.60	0.86	0.86	1.56	1.90	0.00	1.21	1.249	0.864	1.802	0.870	1.889
11	1.42	1.42	0.86	1.60	0.86	1.56	0.86	1.90	1.21	1.21	0.00	1.92	0.86	1.95	1.62	1.62
12	1.28	1.46	1.90	1.46	1.90	0.86	1.56	0.86	1.86	1.24	1.92	0.00	1.62	0.86	0.90	1.50
13	1.80	0.86	0.86	1.94	1.56	1.56	1.24	1.50	1.56	0.86	0.86	1.62	0.00	1.92	1.00	1.50
14	1.28	1.53	1.74	0.86	1.60	1.42	1.80	0.86	1.34	1.80	1.90	0.86	1.90	0.00	1.46	1.00
15	1.80	0.86	1.50	1.91	1.96	1.24	1.56	0.86	1.90	0.86	1.62	0.90	1.00	1.46	0.00	1.46
16	1.80	0.86	1.00	1.21	1.56	1.90	1.90	0.86	0.86	1.90	1.62	1.50	1.50	1.00	1.46	0.00

rate R_s. In Fig. 6 the bit error probability P_e is shown for a binary PM-DD system against the input signal-to-noise ratio Θ_A assuming an optical filter bandwidth $W_o = 10$ GHz. For sake of comparison also the performance of a conventional IM-DD system in the presence of optical amplification is reported in the hypotheses of ASE limited receiver operation. Such results have been evaluated by means of the analytical model proposed by Marcuse [19]. Moreover, since the optical field at the input of an IM-DD receiver has a given polarization state while the ASE noise is completely unpolarized, the performance of the IM-DD system can be improved by setting a polarization filter in front of the receiver. Such operation, requiring an active polarization control, allows only the signal field polarization to be detected by the receiver so to reduce to one half the ASE power affecting the receiver performance. The error probability of this kind of IM-DD receiver is also reported on Fig. 6. It is to be observed that the performance of a PM-DD system and of an IM-DD system using a polarization filter are quite near (within 0.5 dB of sensitivity difference). As a matter of fact, for a given signal peak power, in a PM-DD system 3 dB are gained by constant power signalling while 3 dB are lost since the overall ASE power affects detection. The performance of these two systems are not exactly the same because of the difference in the decision variable probability distribution.

On the other hand PM-DD system shows a sensitivity gain of about 3.3 dB with respect to a conventional IM-DD system.

The performance of multilevel PM-DD systems assuming an optical filter bandwidth $W_o = 10$ GHz are shown in Fig. 7 for $N = 4$, 8, and 16. It is to be observed that even in this case, for a bit error probability of 10^{-9}, the performance of a 4-level system is better than that of a binary system. However the difference is less evident (≈ 0.3 dB).

The performance of a 4-level system is shown in Fig. 8 against the input signal to noise ratio Θ_A for different values of the ratio between the optical and the electrical bandwidth n_s. From the figure it would seem that, if the optical bandwidth is increased, the system performance gets better but this is not a correct evaluation. As a matter of fact systems with different

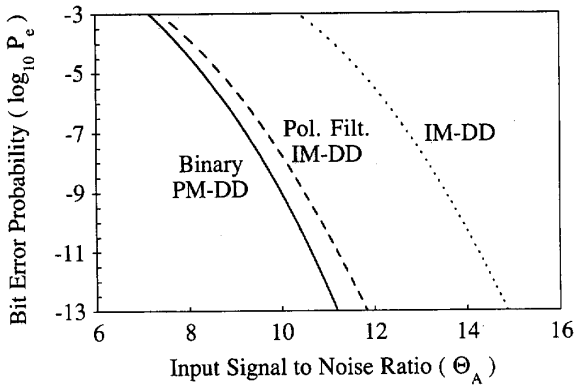

Fig. 6. Bit error probability (P_e) for a binary system in the case of ASE noise limited operation versus the optical signal to noise ratio at the receiver input Θ_A. The ratio n_s between the optical and the electrical bandwidth is assumed equal to 10. For sake of comparison the performance of a conventional IM-DD system and an IM-DD system using a polarization filter in front of the receiver (indicated with Pol. Filt. IM-DD) are also reported under the same assumptions.

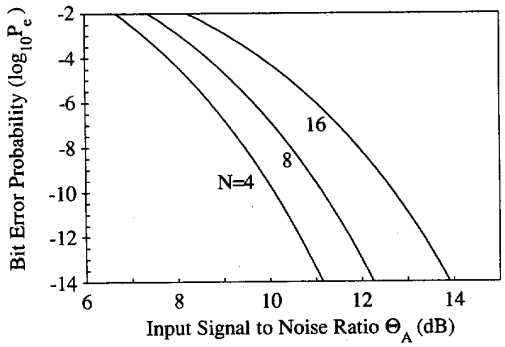

Fig. 7. Bit error probability (P_e) for systems with $N = 4$, 8, and 16 in the case of ASE noise limited operation versus the optical signal-to-noise ratio at the receiver input Θ_A. The ratio n_s between the optical and the electrical bandwidth is assumed equal to 10.

values of n_s, the same symbol rate, the same received optical power and the same value of $\Theta_A = (A^2)/(R_s n_s N_A)$ are

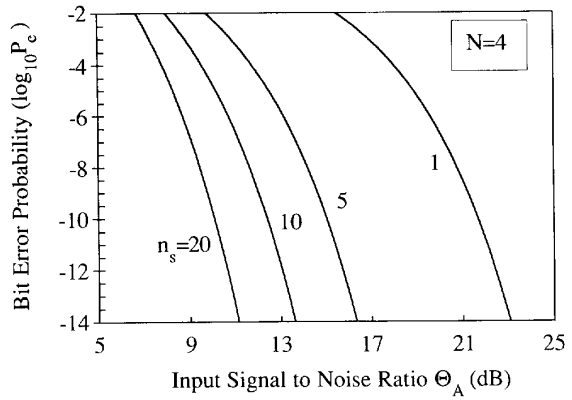

Fig. 8. Bit error probability (P_e) for a 4-level system in the case of ASE noise limited operation versus the optical signal-to-noise ratio at the receiver input Θ_A. Different values of the ratio n_s between the optical and the electrical bandwidth are considered.

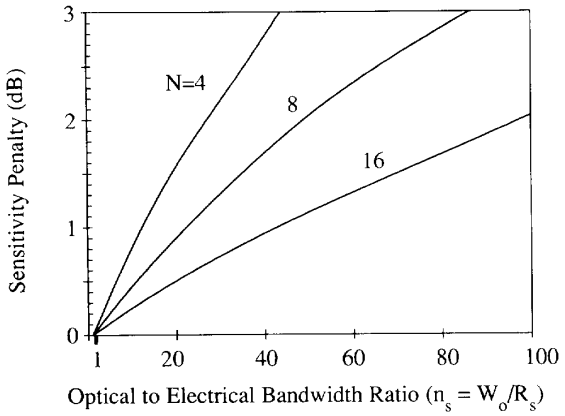

Fig. 10 Sensitivity penalty due to the optical bandwidth enlargement versus the ratio n_s between the optical and the electrical bandwidth for systems with $N = 4$, 8, and 16. ASE noise limited operation is assumed.

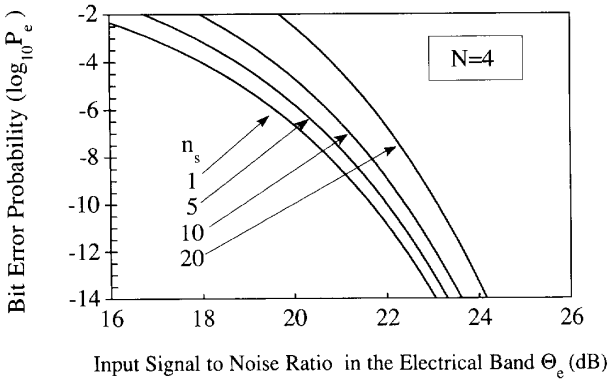

Fig. 9. Bit error probability (P_e) for a 4-level system in the case of ASE noise limited operation versus the optical signal-to-noise ratio reported to the electrical bandwidth Θ_e. Different values of the ratio n_s between the optical and the electrical bandwidth are considered.

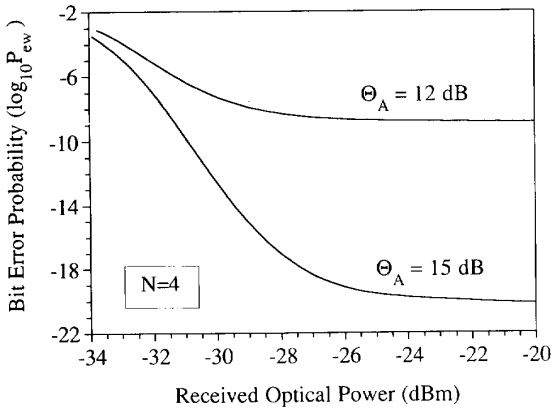

Fig. 11. Worst-case bit error probability (P_{ew}) for a 4-level system versus the optical power at the receiver input. Both ASE and thermal noise are considered, the ratio n_s between the optical and the electrical bandwidth is assumed equal to 10, two values of the optical signal-to-noise ratio at the receiver input Θ_A are considered and the values of the other system parameters are reported in Table VII.

characterized by a different value of the ASE noise spectral density N_A so that such a comparison is not practically significant.

A more significant comparison between such systems can be provided by showing the bit error probability dependence on the input signal to noise ratio referred to the electrical bandwidth, i.e., $\Theta_e = (A^2)/(R_s N_A)$, which does not depend on n_s. Such a comparison is shown in Fig. 9 for $N = 4$. From this figure it can be drawn that, if the optical bandwidth is increased over the value $W_o = R_s$, i.e., $n_s = 1$, the system performance gets worse. Anyway, passing from $n_s = 1$ to $n_s = 20$, the induced sensitivity penalty, for $P_e = 10^{-9}$, is only of 1.5 dB showing that in practical cases, in which the optical bandwidth is greater than the electrical one, the induced penalty is sufficiently low.

The penalty due to an enlargement of the optical filter bandwidth above $W_o = R_s$ is shown in Fig. 10 for $N = 4$, 8, and 16. In practice the curve for $N = 2$ coincides with that for $N = 4$ so that it has not been reported in the figure. It can be observed that increasing the number of levels the optical bandwidth is less critical.

In conclusion an evaluation of the system performance, both when the thermal noise or the ASE noise are dominant and when they are comparable, is shown in Figs. 11 and 12. In these figures the worst case error probability is reported versus the optical power at the receiver input for $N = 4$ and 8, respectively. The system parameters have the values reported in Table VII, n_s is assumed equal to 10 and different values of Θ_A are considered.

V. Conclusions

A novel direct detection optical system is presented based on polarization modulation (PM-DD). At the transmitter the optical field is polarization modulated by means of a standard polarization modulator. At the receiver the Stokes parameters of the input field are estimated by means of a direct detection optical front end and a baseband electrical processing. The Poincaré sphere rotation induced by the fiber is compensated by means of a pure electronic algorithm and the decision is performed in the Stokes space. In particular, among the points

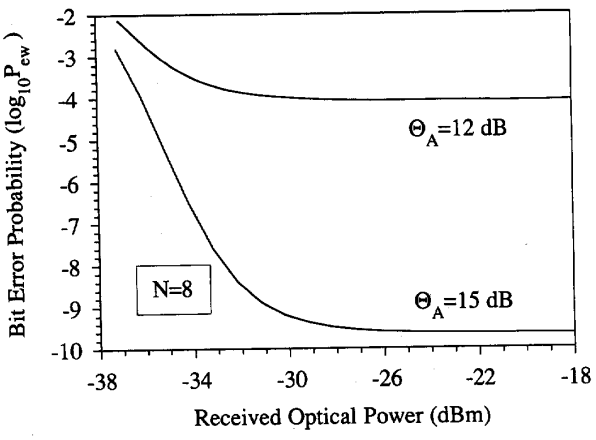

Fig. 12. Worst-case bit error probability (P_{ew}) for an 8-level system versus the optical power at the receiver input. Both ASE and thermal noise are considered, the ratio n_s between the optical and the electrical bandwidth is assumed equal to 10, two values of the optical signal-to-noise ratio at the receiver input Θ_A are considered and the values of the other system parameters are reported in Table VII.

associated to the possible symbols, that nearest to the point representing the received field is assumed to represent the transmitted symbol.

Both binary and multilevel modulations are considered and the points constellations in the Stokes space are chosen in such a way to minimize the distance between adjacent symbols.

If the dominant noise source is the receiver thermal noise the system performance depends on the instantaneous value of the fiber birefringence, anyway such dependence can be maintained quite limited if the optical front end parameters are carefully chosen.

In the particular case of the interferometric optical front end, the front end parameters that allow the minimum worst case bit error probability to be achieved are computed and it is shown that they also permit to limit the sensitivity fluctuations due to fiber birefringence within 1 dB.

Using the above parameters the system performance is calculated for the different multilevel systems obtaining that the 4-level system achieves the best performance. This result is carefully justified and commented on since, for conventional modulation formats, binary systems achieve the best sensitivity.

When considering doped fiber optical amplifiers, the system performance is firstly estimated assuming that ASE noise be dominant. In this case the performance is independent of both the fiber birefringence and the optical front end structure whose parameters depend only on the optical signal-to-noise ratio and the optical bandwidth.

In conclusion the system performance is evaluated considering both receiver thermal noise and amplifiers induced ASE noise to obtain a complete description of the system behavior.

For its good sensitivity and high bandwidth efficiency the proposed system seems promising for applications in high capacity WDM networks and in high speed parallel data transmission, even between different processing units of the same supercomputer.

TABLE VII
ASSUMED VALUES OF THE SYSTEM PARAMETERS (R_s INDICATES THE SYMBOL RATE)

R_b	Bit Rate	1	Gb/s
—	Optical wavelength	1.55	μm
R	Photodiode Responsivity	1	A/W
F_a	Electrical front end noise figure	1.5	—
R_c	Electrical front end load resistance	50	Ω
T_k	Temperature	300	°K
B_n	Thermal noise bandwidth	R_s	Hz

TABLE VIII
VALUES OF THE INTERFEROMETRIC OPTICAL FRONT END PARAMETERS α_s AND β_s THAT MINIMIZE THE POWER SENSITIVITY FLUCTUATION ΔS DUE TO THE FIBER MÜLLER MATRIX EVOLUTION. MODULATION FORMATS WITH DIFFERENT NUMBER OF LEVELS N ARE CONSIDERED

N	α_s	β_s	ΔS (dB)
2	0.2	0.72	0.004
4	0.63	0.81	0.8
8	0.46	0.81	0.66
16	0.72	0.36	0.7

TABLE IX
VALUES OF THE INTERFEROMETRIC OPTICAL FRONT END PARAMETERS α_s AND β_s THAT MINIMIZE THE WORST-CASE BIT ERROR PROBABILITY P_{ew}. MODULATION FORMATS WITH DIFFERENT NUMBER OF LEVELS N ARE CONSIDERED, ΔS REPRESENTS THE POWER SENSITIVITY FLUCTUATION DUE TO THE FIBER MÜLLER MATRIX EVOLUTION AND P_{opt} THE RECEIVED OPTICAL POWER

N	α_s	β_s	P_{opt} (dBm)	$\log_{10} P_{ew}$	ΔS (dB)
2	0.54	0.72	−31	−8.57	0.13
4	0.72	0.81	−32	−7.06	0.8
8	0.72	0.28	−30	−12.3	1.0
16	0.72	0.9	−31	−6.70	0.8

APPENDIX A

When operating with vectors and linear operators in Euclidean vector spaces Einstein's convention results to be very useful to simplify complex formulas and to allow the most important elements in an equation to be pointed out.

In this notation a vector is referred by means of its coordinates in a given base; for example the vector $v = \{v_k\}$ is simply referred through its generic coordinate v_k. The subscript index, denominated covariant index, indicates that v is a vector. On the other hand a linear form (i.e., a linear operator mapping the elements of a vector space on the related scalar field) is indicated with $w = \{w^k\}$ or simply with w^k, where k at superscript is denominated contravariant index. It is to be noted that in some cases subscripts are used not as index but simply to distinguish different variables indicated with the same character, for example P_e and P_s. Anyway the difference can be drawn from the context.

The difference between v and w can be easily understood considering v as a "column" of coordinates and w as a "row". Applying w to the vector v is equivalent to multiply w and v as matrices, operation that in this case produces a scalar.

The above operation can be indicated as follows:

$$c = w^k v_k = \sum_{k=1}^{MS} w^k v_k \tag{A1}$$

where MS is the vector space dimension, c is a scalar and whenever the index k is repeated, both as covariant and contravariant index, indicates a summation. In general whenever the same index is present in an expression both as covariant and contravariant index, a summation on the whole space dimension MS is implicit.

A form ν can be associated to each vector v so that the scalar $c = \nu v$ is the norm of the vector v. In the case of Euclidean spaces the coordinates of v and ν are the same, so that the norm of a vector v is given by the expression $|v|^2 = v^k v_k$ and, in general, the scalar product c between two vectors v and u is provided by $c = u^k v_k = v^k u_k$.

The above formalism for vectors and linear forms can be extended to any kind of linear and multilinear operators. A linear operator mapping a Euclidean vector space on itself can be represented by the set of elements of the representing matrix A in a given base, i.e., as $\boldsymbol{A} = \{a_h^k\}$. With this formalism the covariant index corresponds to the rows and the contravariant one to the columns. Using this formalism the product $\boldsymbol{C}$ between two matrixes $\boldsymbol{A}$ and $\boldsymbol{B}$ can be indicated as $\boldsymbol{C} = \{c_h^k\}$ with $c_h^k = a_h^j b_j^k$.

Finally the formalism can be applied to tensors, viewed as multilinear operators. For example, given two tensors $\boldsymbol{A} = \{a_{hj}^k\}$ and $\boldsymbol{B} = \{b_h^{kj}\}$, the inner product between $\boldsymbol{A}$ and $\boldsymbol{B}$, which provides the matrix $\boldsymbol{C}\{c_h^k\}$, can be written as $c_h^k = a_{rj}^k b_h^{rj}$ while the outer product, which gives the tensor $\boldsymbol{D} = \{c_{jnm}^{khr}\}$ as $c_{jnm}^{khr} = a_{jn}^k b_m^{hr}$. It is to be noted that in this last expression no summation is implicit since any index is replayed neither as covariant nor as contravariant.

REFERENCES

[1] E. Edagawa *et al.*, "904 km, 1.2 Gbit/s non-regenerative optical fiber transmission using 12 Ed-doped fibre amplifiers," *Electron. Lett.*, vol. 26, no. 1, pp. 66–67, 1990.

[2] A. Righetti *et al.*, "11 Gbit/s, 260 km transmission experiment using a directly modulated 1536 nm DFB laser with two Er-doped fibre amplifiers and clock recovery," *Electron. Lett.*, vol. 26, no. 5, pp. 330–331, 1990.

[3] L.T. Blair and H. Nakano, "High sensitivity 10 Gbit/s optical receiver using two cascaded EDFA preamplifiers," *Electron. Lett.*, vol. 27, no. 10, pp. 835–836, 1990.

[4] E. Desurvire, "Study of complex atomic susceptibily of erbium-doped fiber amplifiers," *J. Lightwave Technol.*, vol. 8, no. 10, pp. 1517–1527, 1990.

[5] C. Lin and W. Way, "Optical amplifiers for multiwavelength broadband distribution networks," in *Proc. 17-th European Conf. Opt. Commun., ECOC'91* (Paris, France), Sept. 9–12, 1991, We.C9.1, vol. 2, pp. 125–133.

[6] H. Taga, Y. Yoshida, E. Edagawa, S. Yamamoto, and H. Wakabayashi, "459 km, 2.4 Gbit/s four wavelength multiplexing optical fiber transmission experiment using six Er-doped fiber amplifiers," *Electron. Lett.*, vol. 26, no. 8, pp. 500–501, 1990.

[7] J.P. Blondel *et al.*, "Erbium doped fibre amplifier behavior at various signal wavelength in transoceanic links," in *Proc. 17-th European Conf. Opt. Commun. ECOC'91* (Paris, France), Sept. 9–12, 1991, We.A6.3, vol. 1 part 2, pp. 389–392.

[8] J. G. Proakis, *Digital Communications.* New York: McGraw Hill, 1989, ch. 3.

[9] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory.* Englewood Cliffs, NJ: Prentice Hall, 1987, ch. 5.

[10] S. Benedetto and P. Poggiolini, "Highly bandwidth efficient transmission through continuous polarization modulation," *Electron. Lett.*, vol. 26, no. 17, pp. 1392–1394, 1990.

[11] Optical Computing, *IEEE Spectrum*, vol. 23, no. 8, 1986.

[12] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Multilevel coherent optical system based on Stokes parameters modulation," *J. Lightwave Technol.*, vol. 8, no. 7, pp. 1127–1136, 1990.

[13] R.M.A. Azzam, "Arrangement of four photodetectors for measuring the state of polarization of light," *Opt. Lett.*, vol. 10, no. 4, pp. 309–311, 1985.

[14] R.M.A. Azzam, I.M. Elminyawi, and A.M. El-Sabe, "General analysis and optimization of the four-detector polarimeter," *J. Opt. Soc. Amer. A*, vol. 5, no. 5, pp. 681–689, 1988.

[15] E. Dietrich, B. Enning, R. Gross, and H. Knupke, "Heterodyne transmission of a 560 Mbit/s optical signal by means of polarization shift keying," *Electron. Lett.*, vol. 23, no. 8, pp. 421–422, 1987.

[16] S. Betti, F. Curti, B. Daino, G. De Marchis, and E. Iannone, "State of polarization and phase noise independent coherent optical transmission system based on Stokes parameters detection," *Electron. Lett.*, vol. 24, no. 23, pp. 1460–1461, 1988.

[17] A. Simon and R. Ulrich, "Evolution of polarization along a single mode fibre," *Appl. Phys. Lett.*, vol. 31, no. 8, pp. 1460–1461, 1977.

[18] S.C. Rashleigh, "Origins and control of the polarization effects in single-mode fibers," *J. Lightwave Technol.*, vol. 1, no. 2, pp. 312–321, 1983.

[19] D. Marcuse, "Calculation of bit-error probability for a lightwave system with optical amplifiers and post detection gaussian noise," *J. Lightwave Technol.*, vol. 9, no. 4, pp. 505–513, 1991.

[20] O.K. Tonguz and L.G. Kazowsky, "Theory of direct-detection lightwave receivers using optical amplifiers," *J. Lightwave Technol.*, vol. 9, no. 2, pp. 174–180, 1991.

[21] T. Eftimov and T. Kortensky, "State of polarization in open- and closed-loop elices of single-mode fiber," *J. Modern Opt.*, vol. 36, no. 3, pp. 287–304, 1989.

[22] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Phase-noise and polarization state insensitive optical coherent systems," *J. Lightwave Technol.*, vol. 8, no. 5, pp. 756–767, 1990.

[23] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Double FSK coherent optical system exploiting the orthogonal polarization modes of a single mode optical fiber," *Microwave and Optical Technol. Lett.*, vol. 2, no. 9, pp. 325–3276, 1989.

[24] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "A novel multilevel coherent optical system: 4-quadrature signalling," *J. Lightwave Technol.*, vol. 9, no. 4, pp. 514–523, 1991.

[25] J. G. Proakis, *Digital Communications.* New York: McGraw Hill, 1989, ch. 4.

[26] S. Benedetto and P. Poggiolini, "Performance evaluation of multilevel polarization shift keying modulation schemes," *Electron. Lett.*, vol. 26, no. 4, pp. 244–246, 1990.

[27] S. Betti, G. De Marchis, E. Iannone, M. Pietroselli, and M. Todaro, "Optical coherent transmission systems: Towards a general model for performance evaluation," submitted for publication in *IEEE Trans. Inform. Theory.*

[28] M. Schwartz, W.R. Bennett, and S. Stein, *Communication Systems and Techniques.* New York: McGraw-Hill, 1966, Appendix B.

[29] J.E. Mazo and J. Salz, "Probability of error for quadratic detectors," *Bell Syst. Tech. J.*, vol. 44, no. 11, pp. 2165–2186, 1965.

[30] K. Shumacher and J.J. O'Reilly, "Relationship between the saddle-point approximation and the modified Chernov bound," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 270–272, 1990.

[31] C.W. Helstrom, "Performance analysis of optical receivers by the saddlepoint approximation," *IEEE Trans. Commun.*, vol. Com-27, no. 1, pp. 186–191, 1979.

[32] T. Pikaar, A.C. van Bochove, M.O. van Deventer, H.J. Fraukene, and F.H. Groen, "Fast complete polarimeter for optical fibers," in *Proc. EFOC/LAN'89* (Amsterdam, The Netherlands), June 12–16, 1989, pp. 206–209.

S. Betti, photograph and biography not available at the time of publication.

G. De Marchis, photograph and biography not available at the time of publication.

E. Iannone, photograph and biography not available at the time of publication.