

Novel MSE Adaptive Control of Optical PMD Compensators

Enrico Forestieri, *Member, IEEE*, Giulio Colavolpe, *Associate Member, IEEE*, and Giancarlo Prati, *Member, IEEE*

Abstract—In this paper, we describe a new effective technique for the adaptive adjustment of the control parameters of any structure of polarization-mode dispersion compensator, such as, e.g., those based on a cascade of polarization controllers and polarization-maintaining fibers. This technique is based on the mean square error between the photodetected signal and the decided symbol and allows us to obtain fast convergence and a lower outage probability with a very limited complexity.

Index Terms—Adaptive equalizers, compensation, optical fiber communication, polarization-mode dispersion (PMD).

I. INTRODUCTION

IN OPTICAL transmission systems at the speed of 40 Gb/s and beyond, one of the most challenging impairments is represented by the signal distortions produced by polarization-mode dispersion (PMD). Due to the coherence time of the PMD phenomenon of the order of minutes or even longer, sequences of errors may be actually generated. These bursts of errors cannot be corrected by using forward-error correction (FEC) schemes. In fact, the relevant error distribution is not random and cannot be made as such, since a sufficient interleaving is totally impractical to realize. Therefore, specific PMD compensators must be adopted.

In the technical literature, several solutions, ranging from first- [1]–[3] to higher-order compensators exploiting planar lightwave circuits (PLCs) [4], [5] or cascaded polarization controllers (PCs) coupled with polarization-maintaining fibers (PMFs) or delay lines [6]–[9], have been proposed. The adaptive adjustment of the compensator parameters is often based on criteria aimed at the reconstruction of the undistorted transmitted waveform, by using either the degree of polarization, the orientation of the PMD vector, or the electrical spectrum of the signal at the photodetector output [1]–[3], [6]–[11]. The complexity of the resulting compensator driver can be high due to the needed error signals.

Since the goal of PMD compensation is lowering the bit-error rate (BER), the best control strategy would be based on its minimization. Due to the complexity of a BER measurement, this is not feasible, but it is intuitive to use a quantity closely related to the BER in order to perform this task. This quantity is the mean square error (MSE) between the quantized and unquantized signal sample and can be used to control the parameters

of a PMD compensator, irrespective of its physical structure. In this paper, we describe a simple and effective technique based on the minimization of the output MSE [12], where no optical or electrical filters are needed and the error signal is simply based on the photodetector output [13]. This technique allows us to obtain fast convergence, high stability, and a lower outage probability with a very limited complexity.

The paper is organized as follows. In Section II, we describe the structure of a possible PMD compensator and the parameters that must be controlled to perform the adjustment of the compensator. In Section III, the MSE criterion is described, along with two possible algorithms to implement it. Numerical results are presented in Section IV, and, finally, conclusions are drawn in Section V.

II. THE PMD COMPENSATOR

PMD compensation can be achieved both in optical or electronic domain, but optical compensation should be preferred, because, from a theoretical point of view, PMD causes a linear distortion of the optical signal, which can be more effective to compensate than the nonlinear distortion in which it manifests after photodetection. For illustration purposes, we consider the compensator shown in Fig. 1, but the extension to any other compensator structure is also possible. The compensator consists of a cascade of some optical devices. The first optical device is a PC which allows us to modify the polarization of the optical signal at its input. We then have in this example three PMFs separated by two other PCs or optical rotators. A PMF introduces a differential group delay (DGD) between the components of the optical signal on the two orthogonal states of polarization (SOP) corresponding to its slow and fast axes. The PMD compensator in Fig. 1 generates a polarization dispersion vector that combines with the one generated by the fiber in order to reduce the overall distortion. In order to be able to follow the fiber dispersion vector evolution, the DGD introduced by each PMF should be properly optimized, and the functional form of this delay was chosen to be τ_c , $\alpha\tau_c$ and $(1 - \alpha)\tau_c$, where $0 < \alpha < 1$ and τ_c are design parameters. As the length of the dispersion vector generated by the compensator depends on τ_c , seemingly this parameter should be chosen on the basis of the fiber mean DGD, but we will see that this may not be true, depending on the control strategy and compensator structure (i.e., when using PCs or optical rotators). α is a flexibility parameter introduced a priori in the structure to allow for one more degree of freedom in shaping the optical transfer function of the compensator. We will see that, in practice, $\alpha = 0.5$ turns out to be the best in the generality of cases.

We denote by $x_1(t)$ and $x_2(t)$ the components of any two orthogonal SOP of the signal at the compensator input, whereas, similarly, $y_1(t)$ and $y_2(t)$ are corresponding components of the

Manuscript received May 31, 2002; revised September 6, 2002. This work was carried out by CNIT, Conzorio Nazionale Interuniversitario per le Telecomunicazioni, under a grant by Marconi Communications SpA.

E. Forestieri and G. Prati are with Scuola Superiore Sant'Anna and CNIT, Photonic Networks National Lab, Pisa I-56124, Italy.

G. Colavolpe is with Dipartimento di Ingegneria dell'Informazione, University of Parma and CNIT Research Unit, Parma I-43100, Italy.

Digital Object Identifier 10.1109/JLT.2003.807963

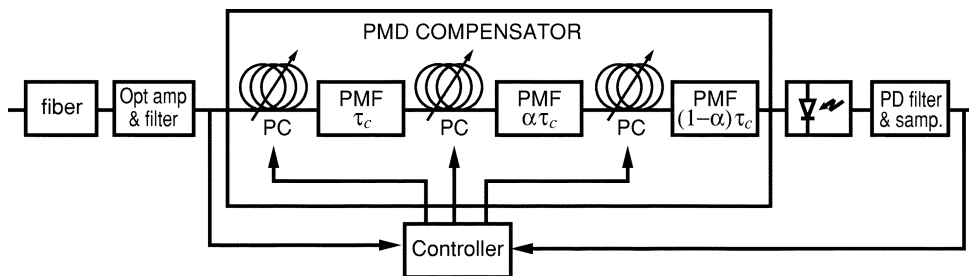


Fig. 1. Structure of the adaptive compensator.

optical signal at the compensator output. Signals $y_1(t)$ and $y_2(t)$ are sent to the input of the photodetector, which produces the signal $y(t)$, given by

$$y(t) = |y_1(t)|^2 + |y_2(t)|^2. \quad (1)$$

Notice that the output signal $y(t)$ remains unchanged irrespective of the reference SOP's choice. This signal may be filtered by means of a postdetection filter. Without loss of generality, we may assume that this filter is not present—the presence of this filter produces straightforward modifications in the adaptive adjustments of the compensator parameters.

We will describe the input/output behavior of each optical device through its Jones transfer matrix $\mathbf{H}(\omega)$ [14], which is a 2×2 matrix characterized by frequency-dependent components. Denoting by $W_1(\omega)$ and $W_2(\omega)$, the Fourier transforms of the components of the optical signal at the device input, the Fourier transforms $Z_1(\omega)$ and $Z_2(\omega)$ of the components of the optical signal at the device output are given by

$$\begin{pmatrix} Z_1(\omega) \\ Z_2(\omega) \end{pmatrix} = \mathbf{H}(\omega) \begin{pmatrix} W_1(\omega) \\ W_2(\omega) \end{pmatrix}. \quad (2)$$

The Jones transfer matrix of the i th PC (or optical rotator) is¹

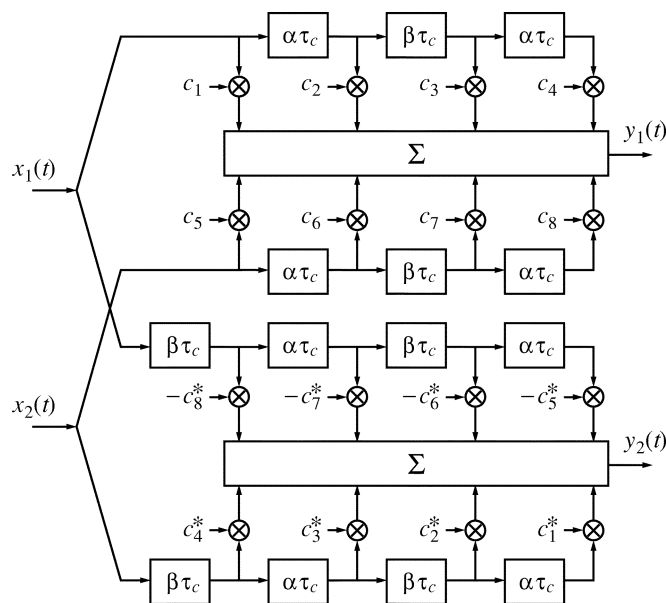
$$\begin{pmatrix} h_{i,1} & h_{i,2} \\ -h_{i,2}^* & h_{i,1}^* \end{pmatrix} \quad (3)$$

where $h_{i,1}$ and $h_{i,2}$ are frequency-independent and satisfy the condition $|h_{i,1}|^2 + |h_{i,2}|^2 = 1$. Denoting by ϕ_i the rotation angle about the S_3 axis in Stokes' space, for an optical rotator, we have that $h_{i,1} = \cos \phi_i$ and $h_{i,2} = \sin \phi_i$, i.e., $h_{i,1}$, $h_{i,2}$ are real quantities, whereas for a PC, they are complex quantities depending on at least two parameters. Thus, only one parameter for each optical rotator and two (or more) parameters for each PC must be controlled, depending on the implementation (e.g., the relative angles between waveplates, for waveplate-based PCs). Denoting by $\phi_{i,1}$ and $\phi_{i,2}$ these control parameters (in general some voltages) for the i th PC, $h_{i,1}$ and $h_{i,2}$ are a function of these parameters. The detailed analytical expression depends on the considered PC.

The Jones transfer matrix of a PMF with DGD τ_i may be expressed as \mathbf{RDR}^{-1} , where

$$\mathbf{D} \triangleq \begin{pmatrix} e^{j\omega\tau_i/2} & 0 \\ 0 & e^{-j\omega\tau_i/2} \end{pmatrix} \quad (4)$$

¹In the following, $(\cdot)^*$ denotes the complex conjugate, whereas $(\cdot)^T$ denotes "transpose."

Fig. 2. Equivalent model of the PMD compensator ($\beta = 1 - \alpha$).

and \mathbf{R} is a frequency-independent unitary rotation matrix accounting for the SOP's orientation. Without loss of generality, this matrix \mathbf{R} may be taken as the identity matrix \mathbf{I} .

As shown in Fig. 1, in order to control the PMD compensator, we need a controller supplied with the quantities necessary to update the compensator parameters. These quantities will be extracted from the signals at the compensator input and/or the samples at the photodetector output.

An important interpretation upon which we build up our control technique is that the compensator in Fig. 1 is equivalent to a two-dimensional transversal filter using some tapped delay lines (TDLs) combining together the signals on two orthogonal SOP [15]. This equivalent model is shown in Fig. 2, where

$$\begin{aligned} c_1 &\triangleq h_{1,1}h_{2,1}h_{3,1} \\ c_2 &\triangleq -h_{1,1}h_{2,2}^*h_{3,2} \\ c_3 &\triangleq -h_{1,2}^*h_{2,2}h_{3,1} \\ c_4 &\triangleq -h_{1,2}^*h_{2,1}^*h_{3,2} \\ c_5 &\triangleq h_{1,2}h_{2,1}h_{3,1} \\ c_6 &\triangleq -h_{1,2}h_{2,2}^*h_{3,2} \\ c_7 &\triangleq h_{1,1}^*h_{2,2}h_{3,1} \\ c_8 &\triangleq h_{1,1}^*h_{2,1}h_{3,2}. \end{aligned} \quad (5)$$

We denote by $\mathbf{c}(\boldsymbol{\phi})$ the vector whose components are the c_ℓ in (5). Note that the tap coefficients c_ℓ of the four TDLs are not independent from each other. Indeed, given four of them, the others are completely determined by (5).

The controller will operate by following the criterion described in Section III and will use one of the two algorithms described in Sections III-A or III-B. In order to illustrate the algorithms for the adaptive adjustment of the PMD compensator, we collect the control parameters of the PCs in a vector $\boldsymbol{\phi}$ defined as

$$\boldsymbol{\phi} \triangleq (\phi_{1,1}, \phi_{1,2}, \phi_{2,1}, \phi_{2,2}, \phi_{3,1}, \phi_{3,2})^T.$$

Being that the PMD is a slowly varying phenomenon, the adjustment of the compensator parameters will be performed at a rate lower than the transmit symbol rate $1/T$. We assume that this adjustment is performed at the discrete-time instants $t_{nL} = nLT$, where $L \geq 1$. We denote by

$$\boldsymbol{\phi}(nL) = (\phi_{1,1}(nL), \phi_{1,2}(nL), \phi_{2,1}(nL), \phi_{2,2}(nL), \phi_{3,1}(nL), \phi_{3,2}(nL))^T$$

the value of the compensator parameters after the n th update.

III. THE MEAN SQUARE ERROR CRITERION

In the MSE criterion, the compensator parameters $\boldsymbol{\phi}$ are adjusted to minimize the mean square value of the error [12], [16]

$$e(k) \triangleq y(t_k) - u(k) \quad (6)$$

where $u(k)$ is the transmitted information symbol in the k th symbol interval. This error is a function of $\boldsymbol{\phi}$ through $y(t_k)$. We explicit this dependence by defining $F(\boldsymbol{\phi}) \triangleq e(k)$. Therefore, the performance index to be minimized is the mean value of $F^2(P\boldsymbol{\phi})$.

The update rules that we use to control the parameters of the i th PC are the following [16]

$$\begin{aligned} \phi_{i,1} [(n+1)L] &= \phi_{i,1}(nL) - \gamma \left. \frac{\partial E \{F^2(\boldsymbol{\phi})\}}{\partial \phi_{i,1}} \right|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \\ &= \phi_{i,1}(nL) - 2\gamma F(\boldsymbol{\phi}) \left. \frac{\partial E \{F(\boldsymbol{\phi})\}}{\partial \phi_{i,1}} \right|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \\ &= \phi_{i,1}(nL) - 2\gamma [y(t_{nL}) - u(nL)] \\ &\quad \cdot \frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,1}} \\ \phi_{i,2} [(n+1)L] &= \phi_{i,2}(nL) - \gamma \left. \frac{\partial E \{F^2(\boldsymbol{\phi})\}}{\partial \phi_{i,1}} \right|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \\ &= \phi_{i,2}(nL) - 2\gamma F(\boldsymbol{\phi}) \left. \frac{\partial E \{F(\boldsymbol{\phi})\}}{\partial \phi_{i,1}} \right|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \\ &= \phi_{i,2}(nL) - 2\gamma [y(t_{nL}) - u(nL)] \\ &\quad \cdot \frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,2}} \end{aligned} \quad (7)$$

where $E\{\cdot\}$ denotes ‘‘expectation,’’ and $\gamma > 0$ is a scale factor that controls the amount of adjustment. In vector notation, this

means that the vector of the compensator parameters is updated by adding a new vector with its norm proportional to the norm of the gradient of $F^2(\boldsymbol{\phi})$ and with opposite direction, i.e., all of its components have the sign changed, as follows:

$$\begin{aligned} \boldsymbol{\phi} [(n+1)L] &= \boldsymbol{\phi}(nL) - \gamma \Delta E \{F^2(\boldsymbol{\phi})\} \Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \\ &= \boldsymbol{\phi}(nL) - 2\gamma F(\boldsymbol{\phi}) \nabla E \{F(\boldsymbol{\phi})\} \Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)}. \end{aligned} \quad (8)$$

In this way, we are sure to move toward a relative minimum of the functional $F^2(\boldsymbol{\phi})$.

Three variations of the basic updating algorithm (7) are obtained by using only sign information contained in the error $e(k)$ and/or in the partial derivative. Hence, the three possible variations are (considering, as an example, the updating rule related to $\phi_{i,1}$)

$$\begin{aligned} \phi_{i,1} [(n+1)L] &= \phi_{i,1}(nL) \\ &\quad - 2\gamma \text{sign} [y(t_{nL}) - u(nL)] \\ &\quad \times \frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,1}} \end{aligned} \quad (9)$$

$$\begin{aligned} \phi_{i,1} [(n+1)L] &= \phi_{i,1}(nL) \\ &\quad - 2\gamma [y(t_{nL}) - u(nL)] \\ &\quad \text{sign} \left[\frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,1}} \right] \end{aligned} \quad (10)$$

$$\begin{aligned} \phi_{i,1} [(n+1)L] &= \phi_{i,1}(nL) \\ &\quad - 2\gamma \text{sign} [y(t_{nL}) - u(nL)] \\ &\quad \text{sign} \left[\frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,1}} \right]. \end{aligned} \quad (11)$$

The following two algorithms illustrate how to compute the gradient of the functional of $F^2(\boldsymbol{\phi})$.

A. First Algorithm

Let us consider the updating rule in vector notation (8). In order to simplify this rule, in the error $F(\boldsymbol{\phi}) = e(k)$, we substitute the transmitted information symbol $u(k)$ with the corresponding decision $\hat{u}(k)$ (not necessarily correct), i.e., we substitute the error $e(k)$ with the estimated error $\hat{e}(k)$ defined as

$$\hat{e}(k) \triangleq y(t_k) - \hat{u}(k) \quad (12)$$

i.e., the difference between the unquantized and quantized sample at the photodetector output. When an accurate characterization of each PC is available, the updating rules may be expressed as a function of the estimated error and the signals on two orthogonal polarizations at the compensator input.

By using the *stochastic gradient* algorithm [16] and by substituting in (7) the error $e(nL)$ with the corresponding estimated error $\hat{e}(nL)$, the updating rule of the parameters of the i th PC becomes

$$\begin{aligned} \phi_{i,1} [(n+1)L] &= \phi_{i,1}(nL) - 2\gamma \hat{e}(nL) \frac{\partial y(t_{nL})}{\partial \phi_{i,1}} \\ \phi_{i,2} [(n+1)L] &= \phi_{i,2}(nL) - 2\gamma \hat{e}(nL) \frac{\partial E \{y(t_{nL})\}}{\partial \phi_{i,2}}. \end{aligned} \quad (13)$$

In vector notation, the algorithm (8) becomes

$$\boldsymbol{\phi}[(n+1)L] = \boldsymbol{\phi}(nL) - 2\gamma\hat{e}(nL)\nabla y(t_{nL}). \quad (14)$$

Recalling the equivalent model of the compensator, the partial derivatives of $y(t_{nL})$, which appear in (14), may be expressed as a function of the components on two orthogonal SOPs of the signal at the compensator input at some proper instants. As can be inferred from Fig. 2, the output sample $y(t_k)$ may be written as²

$$y(t_k) = \mathbf{c}^H \mathbf{A}(k) \mathbf{c} \quad (15)$$

where the Hermitian matrix $\mathbf{A}(k)$ is given by

$$\mathbf{A}(k) \triangleq \mathbf{a}^*(k)\mathbf{a}^T(k) + \mathbf{b}^*(k)\mathbf{b}^T(k) \quad (16)$$

with vectors $\mathbf{a}(k)$ and $\mathbf{b}(k)$ defined by

$$\mathbf{a}(k) \triangleq \begin{pmatrix} x_1(t_k) \\ x_1(t_k - \alpha\tau_c) \\ x_1(t_k - \tau_c) \\ x_1(t_k - \tau_c - \alpha\tau_c) \\ x_2(t_k) \\ x_2(t_k - \alpha\tau_c) \\ x_2(t_k - \tau_c) \\ x_2(t_k - \tau_c - \alpha\tau_c) \end{pmatrix}, \quad \mathbf{b}(k) \triangleq \begin{pmatrix} x_2^*(t_k - 2\tau_c) \\ x_2^*(t_k - \tau_c - \beta\tau_c) \\ x_2^*(t_k - \tau_c) \\ x_2^*(t_k - \beta\tau_c) \\ -x_1^*(t_k - 2\tau_c) \\ -x_1^*(t_k - \tau_c - \beta\tau_c) \\ -x_1^*(t_k - \tau_c) \\ -x_1^*(t_k - \beta\tau_c) \end{pmatrix}. \quad (17)$$

By computing the gradient of $y(t_{nL})$, the algorithm (8) becomes

$$\boldsymbol{\phi}[(n+1)L] = \boldsymbol{\phi}(nL) - 4\gamma\hat{e}(nL)\Re\{\mathbf{J}^H \mathbf{A}(nL)\mathbf{c}\} \quad (18)$$

where

$$\mathbf{J} \triangleq \begin{pmatrix} \frac{\partial c_1}{\partial \phi_{1,1}} & \frac{\partial c_1}{\partial \phi_{1,2}} & \frac{\partial c_1}{\partial \phi_{2,1}} & \frac{\partial c_1}{\partial \phi_{2,2}} & \frac{\partial c_1}{\partial \phi_{3,1}} & \frac{\partial c_1}{\partial \phi_{3,2}} \\ \frac{\partial c_2}{\partial \phi_{1,1}} & \frac{\partial c_2}{\partial \phi_{1,2}} & \frac{\partial c_2}{\partial \phi_{2,1}} & \frac{\partial c_2}{\partial \phi_{2,2}} & \frac{\partial c_2}{\partial \phi_{3,1}} & \frac{\partial c_2}{\partial \phi_{3,2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial c_8}{\partial \phi_{1,1}} & \frac{\partial c_8}{\partial \phi_{1,2}} & \frac{\partial c_8}{\partial \phi_{2,1}} & \frac{\partial c_8}{\partial \phi_{2,2}} & \frac{\partial c_8}{\partial \phi_{3,1}} & \frac{\partial c_8}{\partial \phi_{3,2}} \end{pmatrix} \quad (19)$$

is the Jacobian matrix of the transformation $\mathbf{c} = \mathbf{c}(\boldsymbol{\phi})$

When the control parameters are different from those assumed in this paper, we will have different relationships between these control parameters and coefficients c_ℓ . For instance, if the PC is controlled by means of some other voltages, given the relationship between these voltages and the coefficients $h_{i,1}$ and $h_{i,2}$ which appear in (3), we will be able to express, by using equations (5), the coefficients c_ℓ as a function

²(\mathbf{B})^H denotes the transpose conjugate of the matrix \mathbf{B} .

of these new control parameters. As a consequence, in the computation of the gradient of $y(t_{nL})$, the only modification we must take into account will be the expression of the Jacobian matrix \mathbf{J} , which has to be modified accordingly.

Finally, note that when this algorithm is used, the controller must receive the optical signals at the compensator input and the estimated error $\hat{e}(nL)$.

B. Second Algorithm

In this case, the expected value of the gradient of $F^2(\boldsymbol{\phi})$ is estimated by a trial-and-error procedure. In fact, the controller tentatively updates the compensator parameters, one by one by a fixed step size, and computes the corresponding gradient components by averaging a given number of values of the estimated error. It will be shown that, in this case, the relationship between the control parameters of each PC and the corresponding Jones matrix is not needed to compute the gradient components.

Defining $G[\boldsymbol{\phi}(nL)] \triangleq E\{\hat{e}^2(nL)\}$, the updating rule (8) becomes

$$\boldsymbol{\phi}[(n+1)L] = \boldsymbol{\phi}(nL) - \gamma\nabla G(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)}. \quad (20)$$

The partial derivatives of $G(\boldsymbol{\phi})$ for $\boldsymbol{\phi} = \boldsymbol{\phi}(t_n)$ can be computed by using the following seven-step procedure:

- **Step 1: Find the value $G[\boldsymbol{\phi}(nL)]$ at iteration n .**

In the time interval $(nLT, nLT + (LT/7))$ an estimate of $G_1 \triangleq G[\boldsymbol{\phi}(nL)]$ is computed through time averaging, i.e., by averaging $L/7$ values of the square estimated error

$$G_1 = G[(\boldsymbol{\phi}(nL))] = \frac{\sum_{i=0}^{\frac{L}{7}-1} \hat{e}(nL+i)}{\frac{L}{7}}. \quad (21)$$

- **Step 2 ℓ ($\ell = 2, \dots, 7$): Find the partial derivative $(\partial G(\boldsymbol{\phi})/\partial \phi_{i,j})|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)}$ at iteration n .**

In order to do this, parameter $\phi_{i,j}$ ($i = 1, 2, 3$, and $j = 1, 2$) is temporarily set to the value $\phi_{i,j}(nL) + \Delta$, whereas the other parameters are left unchanged. The corresponding value of $G(\boldsymbol{\phi})$, denoted by G_ℓ , is computed as in Step 1 in the time interval $(nLT + (\ell - 1)LT/7, nLT + \ell LT/7)$.

The estimate of the partial derivative of $G(\boldsymbol{\phi})$ with respect to $\phi_{i,j}$ is computed as

$$\left. \frac{\partial G(\boldsymbol{\phi})}{\partial \phi_{i,j}} \right|_{\boldsymbol{\phi}=\boldsymbol{\phi}(nL)} \simeq \frac{G_\ell - G_1}{\Delta}. \quad (22)$$

After the gradient estimation has been completed, the definitive parameter update is performed.

Note that, in this case, it is not necessary to know the relationship between the control parameters of each PC and the corresponding Jones matrices. In fact, the partial derivatives of the functional with respect to the compensator control parameters are computed without knowledge of this relationship.

Finally, note that when this algorithm is used, the controller must receive the estimated error only.

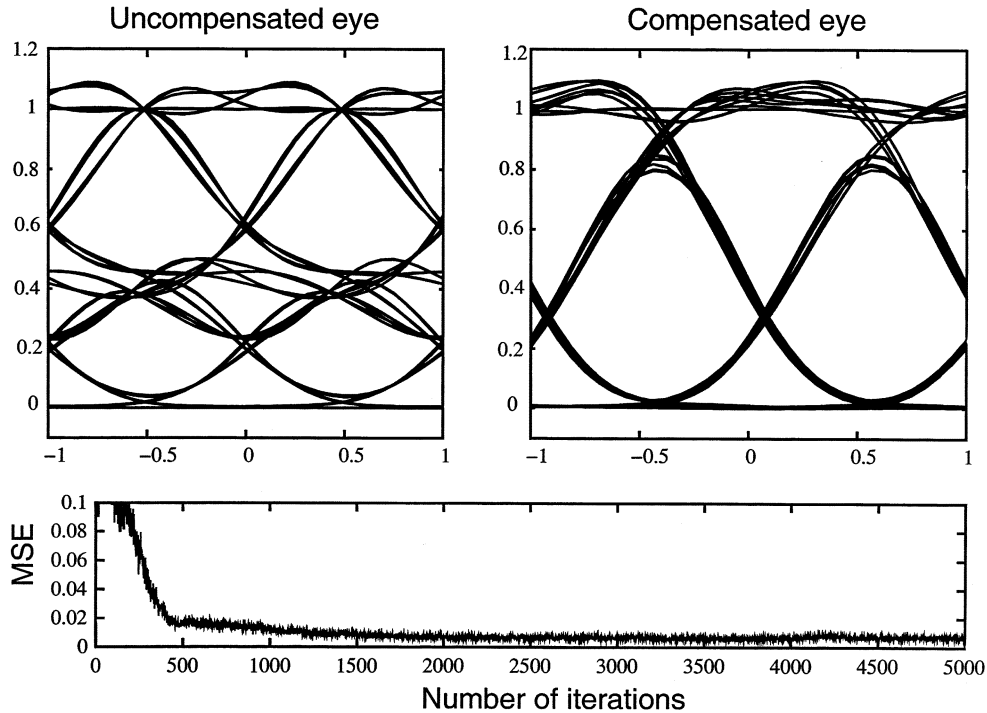


Fig. 3. Eye diagrams and MSE evolution.

IV. NUMERICAL RESULTS

We simulated the system in Fig. 1 in order to test its dynamical behavior. A Gaussian-shaped optical filter with 3-dB bandwidth equal to 1.875 times the bit rate and a fifth-order Bessel postdetection filter with 3-dB bandwidth equal to 0.75 times the bit rate were used. The DGD of the three PMFs in the compensator was taken equal to τ_c , $\tau_c/2$ and $\tau_c/2$, respectively, i.e., parameter α was chosen equal to 0.5, as we have verified that the MSE performance is best for α in the range of 0.4 and 0.6 and not sensitive to the actual value. Choosing $\alpha = 0.5$ generates equal delays in the equivalent model of Fig. 2, making it easier to deal with the equivalent fractionally spaced adaptive filter, with no loss of generality. In Fig. 3, the uncompensated and compensated eye diagrams are shown for $\tau_c = T$, T being the bit time, fixed PMD such that the second-order parameters [fiber DGD $\Delta\tau$, DGD derivative $\Delta\tau_\omega$, and principal states of polarization (PSP) rotation rate q_ω] are $\Delta\tau = T$, $\Delta\tau_\omega = 0.1T^2$, $q_\omega = 0.4T$, respectively, and there is equal signal power splitting among the PSP. The compensator adjustment is performed by using the second algorithm described in Section III-B. Also reported in Fig. 3 is the MSE evolution as a function of the number of PMD compensator parameters iterations, showing the convergence from the uncompensated to the compensated eye diagram.

The compensator parameters are updated after a number of signal samples sufficiently high to obtain a good estimate for the MSE. We found that $100 \cdot N$ samples are more than enough, if N is the number of compensator parameters. Since we can take into account, as an example, only one of every four signal samples, at 40 Gb/s, the compensator could react in $500 \cdot 100 \cdot N \cdot 4 \cdot 25 \text{ ps} = 5 \cdot N \text{ } \mu\text{s}$ to any PMD change and thus is only limited by the PC speed.

The results shown in Fig. 3 were obtained by using a nearly truly random bit sequence (pseudorandom sequence with a period of $2^{144} - 1$) and show the compensator effectiveness when using our control algorithm. However, the ultimate figure of merit is the outage probability. The outage probability is defined here as the probability that the BER exceeds 10^{-12} , given 3-dB sensitivity penalty with respect to the case of PMD absence. A 2^5 -bit de Bruijn sequence [17] was chosen in order to account for the intersymbol interference (ISI) due to four adjacent bits for the BER evaluation, performed by taking into account the exact postdetection noise statistics [18]. The outage probability was evaluated through a Monte Carlo approach by using the random waveplate model for the fiber. An extensive optimization was carried out in the case of one PC and two rotators, and it turned out that the optimum value for parameter τ_c is about $0.4T$ for both MSE and Stokes' parameters criterion. As an example, we show in Fig. 4, the outage probability as a function of τ_c for a fixed mean DGD.

We report in Figs. 5 and 6 the results of two different configurations for the compensator, corresponding to $\tau_c = T$ and $\tau_c = 0.4T$. The case $\tau_c = T$ was chosen because it is a special case, as explained subsequently. Fig. 5 refers to the case of one PC and two optical rotators, while Fig. 6 is relative to the case of three PCs. For comparison, in Figs. 5 and 6, the outage probabilities for the uncompensated (dash-dotted) and compensated (dashed line) system are shown. Also for comparison, another controlling strategy (solid line), based on the constancy with the frequency of the signal Stokes' parameters, is considered. This strategy corresponds to the strategy of aligning the overall PMD vector $\vec{\Omega}$ with the signal SOP \vec{s} , according to

$$\frac{d\vec{s}}{d\omega} = \vec{\Omega} \times \vec{s} = 0. \quad (23)$$

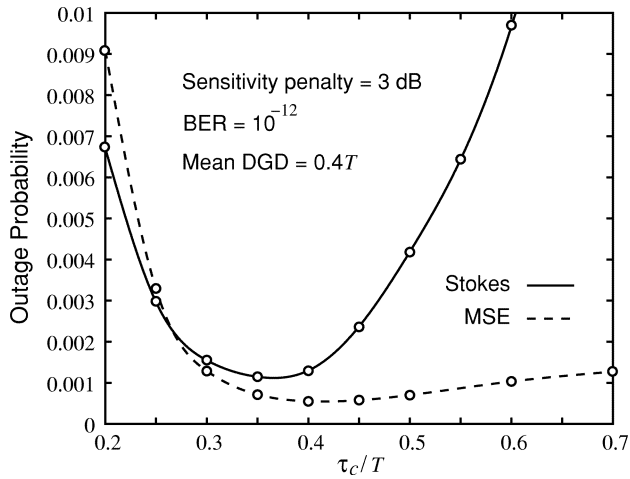


Fig. 4. Outage probability as a function of τ_c for $\alpha = 0.5$ and compensator with one PC and two optical rotators.

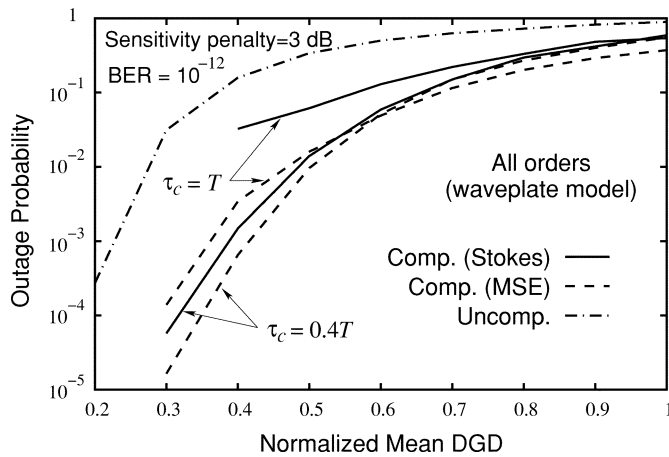


Fig. 5. Outage probability for compensator with one PC and two optical rotators.

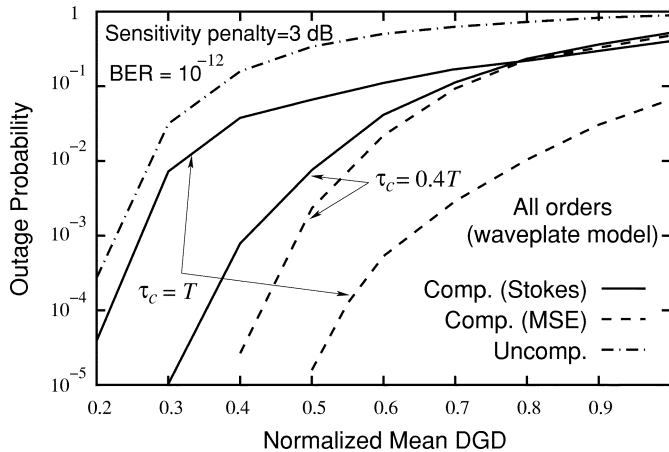


Fig. 6. Outage probability for compensator with three PCs.

Following this criterion, we apply the stochastic gradient algorithm to find the minimum of a suitable functional [19]. The functional we use is based on the fact that N vectors are parallel when the sum of the squared modulus of the difference of all possible couple of vectors is minimum, as follows:

$$\sum_{i=2}^N \sum_{j=1}^{i-1} |\vec{s}_i - \vec{s}_j|^2 \quad (24)$$

where \vec{s}_i and \vec{s}_j are the signal SOPs at frequencies ω_i and ω_j , $i, j = 1, 2, \dots, N$, respectively. These frequencies are chosen to be equally spaced by $\Delta\omega$, and it turns out that there is a minimum value of $\Delta\omega$ and a maximum value of N , giving the asymptotic behavior for $\Delta\omega \rightarrow 0$ and $N \rightarrow \infty$. In practice, it suffices to consider a small number of equispaced frequencies (5–7) in a narrow bandwidth (about half the signal bandwidth) around the optical carrier.

It can be seen that the MSE criterion gives a lower outage probability in both cases $\tau_c = T$ and $\tau_c = 0.4T$ and that it is much less sensitive to the τ_c value than the other criterion in the case of one PC and two rotators. In this case, the sensitivity of the Stokes' parameters criterion to the τ_c value is explained by the fact that, given the particular compensator structure, the compensator PMD vector magnitude is not independent from its orientation, and thus directions exist for which the compensator DGD cannot be less than τ_c , for example. This means that for some overall PMD vector components due to the fiber $\vec{\Omega}_f$ and signal SOP \vec{s} , the compensator may not be able to generate the needed dispersion vector $\vec{\Omega}_c$ such that $\vec{\Omega} = \vec{\Omega}_f + \vec{\Omega}_c$ is aligned with \vec{s} , and this happens with higher probability when τ_c is significantly greater than the fiber mean DGD. The MSE criterion is less sensitive to the τ_c value because it is a criterion aiming at the maximization of the eye opening, regardless of the signal waveform, which is allowed to be different from the undistorted transmitted one. In the case of three PCs, the criterion outperforms the Stokes' criterion, as expected, for $\tau_c = T$, because the increased number of parameters also increases the degree of freedom for the tap coefficients c_ℓ , and the structure in Fig. 2 is equivalent to a fractionally spaced equalizer with spacing $T/2$, a structure which is known to perform very well [12] as it is able to combine the operations of matched filtering and equalization of ISI into a single filter.

By comparing Fig. 5 and Fig. 6, we can also see that when using a criterion whose goal is the inversion of the channel Jones matrix, such as the Stokes' parameters criterion, the improvement in the outage probability obtainable by using three PCs instead of one PC and two optical rotators, is very limited. On the contrary, the MSE criterion is able to take advantage from the increased number of degrees of freedom.

As a final note, notice that the stochastic gradient algorithm can only guarantee the reach of a local minimum and not a global one. Nevertheless, whichever PC structure we tried, all local minima turned out to be almost equivalent, and the use of algorithms that are able to reach a global minimum (such as the simulated annealing algorithm) produce nearly equal results in terms of outage probability.

V. CONCLUSION

In this paper, a new simple and effective technique for the adaptive adjustment of the parameters of a PMD compensator has been presented. This technique is based on the minimization of the MSE between the photodetector output and the corresponding symbol decision. Two algorithms based on this criterion have been proposed. The first one needs an accurate characterization of the optical compensator but is faster and more accurate. On the contrary, the second algorithm does not require

the knowledge of the relationship between the control parameters and the Jones matrices of the compensator components. In any case, the resulting compensator is able to react to any PMD change and is only limited by the speed of the compensator's physical components.

With respect to other algorithms for the adaptive control of PMD compensators, with the proposed algorithm, the resulting compensator is characterized by a better convergence speed and steady-state behavior with a lower complexity as only the squared difference between the unquantized and quantized signal sample is needed.

REFERENCES

- [1] F. Heismann, D. A. Fishmann, and D. L. Wilson, "Automatic compensation of first-order polarization mode dispersion in a 10 Gbit/s transmission system," in *Proc. ECOC'98*, Madrid, Spain, 1998, pp. 529–530.
- [2] C. Francia, F. Bruyère, J.-P. Thierry, and D. Penninckx, "Simple dynamic polarization mode dispersion compensator," *Electron. Lett.*, vol. 35, no. 5, pp. 414–416, 1999.
- [3] H. Ooi, Y. Akiyama, and G. Ishikawa, "Automatic polarization-mode dispersion compensation in 40 Gbit/s transmission," in *Proc. OFC'99*, San Diego, CA, 1999, pp. 86–87.
- [4] M. Bohn, G. Mohs, C. Scheerer, C. Glingener, C. Wree, and W. Rosenkranz, "An adaptive optical equalizer concept for single channel distortion compensation," in *Proc. ECOC'01*, Amsterdam, The Netherlands, 2001, pp. 6–7.
- [5] T. Saïda, K. Takiguchi, S. Kuwahara, Y. Kisaka, Y. Miyamoto, Y. Hashizume, T. Shibata, and K. Okamoto, "Planar lightwave circuit polarization mode dispersion compensator," in *Proc. ECOC'01*, Amsterdam, The Netherlands, 2001, pp. 10–11.
- [6] J. Patscher and R. Eckhardt, "Component for second-order compensation of polarization-mode dispersion," *Electron. Lett.*, vol. 33, no. 13, pp. 1157–1159, 1997.
- [7] R. Noé, D. Sandel, M. Y.-Dierolf, S. Hinz, V. Mirvoda, A. Schöpflin, C. Glingener, E. Gottwald, C. Scheerer, G. Fischer, T. Weyrauch, and W. Haase, "Polarization mode dispersion compensation at 10, 20, and 40 Gbit/s with various optical equalizers," *J. Lightwave Technol.*, vol. 17, pp. 1602–1616, Sept. 1999.
- [8] Q. Yu, L. Yan, S. Lee, Y. Xie, M. Hauer, Z. Pan, and E. Willner, "Enhanced higher order PMD compensation using a variable time delay between polarizations," in *Proc. ECOC*, vol. 2, Munich, Germany, 2000, pp. 47–48.
- [9] M. Karlsson, C. Xie, H. Sunnerud, and A. Andrekson, "Higher order polarization mode dispersion compensator with three degrees of freedom," in *Proc. OFC'01*, Anaheim, CA, 2001.
- [10] T. Ono, S. Yamakazi, H. Shimizu, and K. Emura, "Polarization control method for suppressing polarization mode dispersion influence in optical transmissions systems," *J. Lightwave Technol.*, vol. 12, pp. 891–898, May 1994.
- [11] N. Kikuchi and S. Sasaki, "Polarization Mode Dispersion (PMD) detection sensitivity of degree of polarization method for PMD compensation," in *Proc. ECOC'99*, vol. 2, Nice, France, 1999, pp. 8–9.
- [12] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. London: Prentice-Hall, 1987.
- [13] G. Prati, E. Forestieri, and G. Colavolpe, "Method Based on the Mean Square Error for the Adaptive Adjustment of Optical PMD Compensators," Italian Patent no. MI2001A002 631, Dec. 2001.
- [14] R. C. Jones, "A new calculus for the treatment of optical systems (I)," *J. Opt. Soc. Amer. A. Opt. Image Sci.*, vol. 31, pp. 488–493, 1941.
- [15] E. Forestieri, "Impact of Polarization Mode Dispersion on Direct Detection Lightwave Systems Performance," University of Parma, Tech. Rep., 2001.

- [16] G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.
- [17] W. Golomb, *Shift Register Sequences*. San Francisco: Holden-Day, 1967.
- [18] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and post-detection filtering," *J. Lightwave Technol.*, vol. 18, pp. 1493–1503, Nov. 2000.
- [19] G. Colavolpe and E. Forestieri, "Method Based on the Stokes Parameters for the Adaptive Adjustment of Optical PMD Compensators," Italian Patent no. MI2001A002 632, Dec. 2001.

Enrico Forestieri (S'91–M'91) 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 has been a postdoctoral scholar at the University of Parma, Parma, Italy, working in optical communication systems. From 1991 to 2000, he was a Research Scientist and Faculty Member of the University of Parma. He is now an Associate Professor of Telecommunications at the Scuola Superiore Sant'Anna in Pisa, Italy. His research interests are in the area of digital communication theory and optical communication systems.

Giulio Colavolpe (S'96–A'00) was born in Cosenza, Italy, in 1969. He received the Dr. Ing. degree in telecommunications engineering (cum laude) from the University of Pisa, Pisa, Italy, in 1994 and the Ph.D. degree in information technology from the University of Parma, Parma, Italy, in 1998.

Since 1997, he has been at the University of Parma, where he is now an Associate Professor of Telecommunications. In 2000, he was a Visiting Scientist at the Insitut Eurécom, Valbonne, France. His main research interests include digital transmission theory, channel coding, and signal processing.

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

From 1975 to 1978, he was Associate Professor of Electrical Engineering at the University of Pisa. From 1978 to 1979, he was on a NATO-supported Fellowship Leave in the Department of Electrical Engineering, University of Southern California, Los Angeles, working in optical communications. From 1976 to 1986, he was a Research Scientist of the Italian National Research Council (CNR) at the Centro di Studio per Metodi e Dispositivi di Radiotrasmissione, Pisa. In 1982, he was Visiting/Associate Professor in the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. From 1986 to 1988, he was Professor of Electrical Engineering at the University of Genoa, Italy. From 1988 to 2000, he was Professor of Telecommunications Engineering at the University of Parma, Italy, where he served as Dean of the Engineering Faculty from 1992 to 1998. Since 1995, he has been Director of CNIT, Italian Interuniversity Consortium for Telecommunications, incorporating 28 universities. He is now Professor of Telecommunications at the Scuola Superiore Sant'Anna in Pisa, Italy. His professional and academic interests are in telecommunication systems and digital signal processing in communications. The activity has 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, and fiber communications.

Dr. Prati has been a Member of the Scientific Committee of the Italian Space Agency (ASI) and of the Technical Program Committee of the European Conference on Optical Communications (ECOC) since 1997.