

PMD Effects in Fiber Optic Transmission Systems

Julien Poirier, Michel Gadonna, Laurent Dupont

▶ **To cite this version:**

Julien Poirier, Michel Gadonna, Laurent Dupont. PMD Effects in Fiber Optic Transmission Systems. Fiber and Integrated Optics, 2008, 27 (6), pp.559-578. 10.1080/01468030802272476 . hal-00365826

HAL Id: hal-00365826
https://hal.science/hal-00365826

Submitted on 15 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

PMD Effects in Fiber Optic Transmission Systems

J. POIRRIER,[1] M. GADONNA,[2] and L. DUPONT[3]

[1]France Telecom R&D, Lannion, France
[2]Institut Telecom, Telecom-Bretagne, Centre Commun Lannionnais d'Optique (CCLO), UMR CNRS FOTON6082, Lannion, France
[3]Institut Telecom, Telecom-Bretagne, Optics Department, UMR CNRS FOTON 6082, Brest, France

Abstract *Polarization mode dispersion (PMD) remains the most important limiting impairment for high bit rate transmission over optical fibers. In this article, after a brief description of PMD model and the associated statistic, we study the pulse distortion due to PMD using a powerful time-frequency diagram. This approach allows a very simple illustration of the second and high order PMD effects through effective pulse broadening. Finally, the PMD emulation issue and both electronic and optical compensation techniques are presented, that is a current field of cooperation between ENST-Bretagne and France Telecom R&D.*

Keywords optical transmission impairments, PMD compensation, PMD emulation, polarization mode dispersion, time-frequency analysis

1. Introduction

Polarization mode dispersion (PMD) has become the most limiting issue for high data rate optical transmission due to the statistics of its temporal variations that require dynamical compensation system. Many challenges remain to tackle for a good PMD management; among them are a simple PMD model well suited to approximate real PMD fiber, understanding of PMD impact on optical data transmission, and finally, design of the most efficient compensation system. These different aspects have been jointly studied in both France Telecom R&D and ENST Bretagne in the frame of the European e-photon one+ network of excellence.

2. PMD Principles

PMD can be understood as a dispersion, function of polarization. To the first order, it can simply be understood as a two mode transmission. Accordingly, the equivalent channel can be understood as generating a single echo, as is described next.

2.1. First Order PMD

Basically, the PMD results from the dependence of the vectorial optical propagation operator on the frequency and it induces pulse distortion [1]. The optical transmission

Address correspondence to M. Gadonna, GET/ENST Bretagne/CCLO UMR CNRS, Foton 6082, France. E-mail: gadonna@enssat.fr

properties of an optical fiber on the electrical field are well described by the Jones matrix that is frequency dependent (each dimension represents a polarization):

$$e^{i\phi(\omega)}T(\omega), \tag{1}$$

where ω represents the angular frequency deviation from the central frequency and $\phi(\omega)$ represents the isotropic phase accumulated by the mode along the propagation. The output state of polarization (SOP) dispersion is expressed as:

$$\widehat{s}(\omega) = e^{i\phi}T(\omega)T^{-1}(0)\widehat{s}(0). \tag{2}$$

Using the Taylor expansion of the correlation matrix, $T(\omega)T^{-1}(0)$, and considering only the first order, it is interesting to extract the polarization states $\widehat{p}$ that remain invariant with frequency (principal state of polarization (PSP)). Straightforward calculations show that these steady polarizations are the eigenvectors of the matrix,[1] $T_\omega(0)T^{-1}(0)$, which are orthogonal in the absence of polarization dependent losses. The associated eigenvalues are $\pm\tau/2$ and they are identified to the variation to the average group delay for each PSP. This eigenvector analysis, first proposed by Poole and Wagner [2], is a powerful representation of the first order PMD in time domain that gives a very simple interpretation of pulse impairment by the introduction of two different travelling modes in optical fiber (the "echo channel" mentioned earlier). The repartition of energy between the two modes or replica of the input pulse is determined by the projection of the input polarization onto both PSP. The difference τ between the arrival times of both replicas is called the differential group delay (DGD), while their respective amplitudes are directly related to the power sharing between the PSP. We notice that the operator, $T_\omega(0)T^{-1}(0)$, can be expanded with the Pauli spin matrices [3]:

$$T_\omega T^{-1} = -\frac{i}{2}\vec{\tau}.\vec{\sigma}, \tag{3}$$

where $\vec{\tau}$ is the PMD vector defined as $\vec{\tau} = \tau\vec{p}$, with $\vec{p}$ the eigenvector defined in Stokes space and $\vec{\sigma}$ is the spin vector. Using this formalism, the dispersion of polarization with frequency, at first order, is assimilated to a rotation where the rotation axis is given by the PMD vector, $\vec{\tau}$, and the angular velocity of polarization versus the frequency is given by the modulus, $|\vec{\tau}|$. This evolution of polarization versus frequency can be summarized by the simple vectorial equation:

$$\frac{d\vec{s}}{d\omega} = \vec{\tau} \times \vec{s}(\omega). \tag{4}$$

2.2. Higher Order PMD

In Stokes space, the high order description is the consequence of the limitation to the (first order) rotation description of output polarization with frequency. We observe that the rotation axis and the angular velocity do not remain constant over the full spectrum. We introduce then small variations of PMD vector versus frequency through the second order limited Taylor expansion of this vector:

$$\vec{\tau}(\omega) = \vec{\tau}(0) + \vec{\tau}_\omega(0)\omega + \frac{1}{2}\vec{\tau}_{\omega\omega}(0)\omega^2 + (0)\omega^2 + \dots . \tag{5}$$

[1]The subscript ω stands for the derivation with respect to the variable ω.

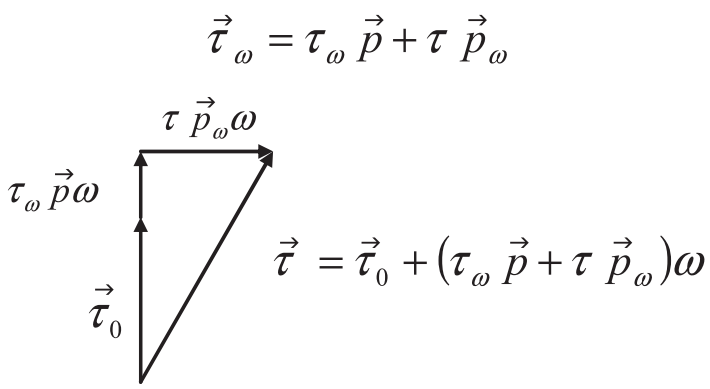


Figure 1. Vectorial second order PMD representation.

This description gives us an interesting geometrical approach of PMD dispersion (Figure 1). For example, the second order, taking into account the PMD vector definition, is expressed by the vector derivative versus ω:

$$\vec{\tau}_\omega = \tau_\omega \vec{p} + \tau \vec{p}_\omega. \tag{6}$$

The first term describes the polarization dependant chromatic dispersion (PCD). This term is parallel to the first order PMD vector and geometrically, it describes a change of SOP rotation velocity versus the frequency. The second is orthogonal to the first order PMD vector and describes the depolarization of PSP with the frequency. This classical representation of PMD should be carefully handled because the cut-off of the Taylor expansion of PMD vector inevitably induces divergence of PMD magnitude with frequency that is in contradiction with what is observed experimentally in optical fiber with bounded magnitude PMD parameters. These aspects have been extensively studied by F. Heismann [3].

3. PMD Statistics

We have seen that PMD limited to the first order is fully described by a vector: the PMD vector. Its length is the differential group delay and its direction fixed along the slow PSP. The fast PSP is associated to the antipodal vector. We also mentioned that PMD is a random process. Let us now recall the statistics of this process to derive the statistics of its system impact.

3.1. Maxwellian Statistics

The PMD vector of a fiber can be viewed as the concatenation of elementary PMD vectors each representing an elementary section of birefringent fiber. The coupling angle between two elementary sections is random; the global PMD problem is therefore a three-dimensional random walk problem. As a consequence, the total length of the vector follows a Maxwellian distribution, while it can take any direction equiprobably [1]. The average DGD is referred as to PMD: it must be noted that this acronym represents the physical phenomenon as well as the average DGD.

On the other hand, the impact of PMD on the signal is related to both the value of the DGD and the factor γ characterizing the repartition of energy between the modes. The repartition of energy γ is a random variable bounded between 0 and 1. γ is not a uniform but is more likely to take a value close to 0.5.

The impact can actually be expressed as:

$$g(\theta) = aDGD^2.\gamma(\theta)(1-\gamma(\theta)) = aDGD^2.\sin^2(\theta), \tag{7}$$

where θ follows a uniform distribution over the interval $[0\ 2\pi]$ and a is a constant.

As the DGD is theoretically not bounded, so is the penalty. To reach a reasonable system design, it is typically specified that the system can not work correctly for a limited and specified amount of its lifetime. This is referred as to the system unavailability or outage. It is usually specified that PMD should not cause unavailability for more than 5 min/year (probability of 10^{-5}). Accordingly the probability that the PMD induced penalty exceeds a given penalty is bounded to this outage time.

3.2. From Statistics to Penalty

Equation (7) expresses the penalty as a function of DGD and γ. How can we translate this relationship to a penalty given an outage and a (*DGD*)? Estimating the outage implies integrating the joint probability of the bi-dimensional random variable (*DGD*, γ) over the region where (*DGD*, γ) induces a penalty greater than the margin. Fortunately, *DGD* and γ being two independent random variables, the joint probability is simply the product of the individual densities.

A crude approximation consists in neglecting the statistics of γ. It is then only necessary to integrate over the probability density function (pdf) of *DGD*. We prefer a method that implicitly takes into account the γ dependency [4]. Let us call τ the DGD and $\tau_\perp$ the projection of the PMD vector along the axis perpendicular to the launch SOP; we have the expressions:

The penalty formula is:

$$P_{\tau_\perp,dB} = \frac{A_{PMD}}{4T_{bit}^2}(\tau_\perp)^2 \tag{8}$$

The random variable $\tau\perp$ follows a Rayleigh distribution:

$$p_{\tau_\perp}(x) = \frac{x}{\sigma^2}\cdot\exp\left(-\frac{x^2}{2\sigma^2}\right) \qquad \text{if } x>0, \qquad 0 \quad \text{elsewhere}$$
$$\sigma = \frac{1}{2}\cdot\sqrt{\frac{\pi}{2}}\langle\tau\rangle \tag{9}$$

translates in a penalty (ε) pdf

$$p_\varepsilon(\varepsilon) = p_{\Delta\tau}(\Delta\tau(\varepsilon))\cdot\frac{d\Delta\tau}{d\varepsilon} = \frac{1}{\langle\varepsilon\rangle}\cdot\exp\left(-\frac{\varepsilon}{\langle\varepsilon\rangle}\right)$$
$$\langle\varepsilon\rangle = \frac{\pi A_{PMD}}{16T_{bit}^2}\langle\tau\rangle^2 \tag{10}$$

The probability for a penalty to exceed N dB:

$$P_{out} = P_{\varepsilon\geq N} = \int_N^\infty p_\varepsilon(\varepsilon)d\varepsilon = \exp\left(-\frac{N}{\langle\varepsilon\rangle}\right) \tag{11}$$

Equivalently the maximum acceptable PMD given a penalty of N dB is:

$$\frac{\langle\tau\rangle}{T_{bit}} = 4 \cdot \sqrt{\frac{N}{\pi A_{PMD} \ln(1/P_{out})}} \tag{12}$$

For transponder testing it is easier to measure in terms of instantaneous DGD and then to relate the DGD limit to the average DGD:

$$\frac{\langle\tau\rangle}{\tau_{\lim it}} = 2 \cdot \sqrt{\frac{1}{\pi \ln(1/P_{out})}} \cong 3 \tag{13}$$

3.3. Enhanced Statistics: Hinge Model

The underlying assumption of a Maxwellian DGD distribution has been questioned [5]. Tending towards Maxwellian statistics requires that all the elementary sections composing a fiber are evolving with time. It has, however, been observed that most of them are stable. They belong to buried fibers and therefore do not sense ground temperature modifications. In the absence of external time varying stresses, these sections appear as "dead." Only fiber sections exposed to temperature fluctuations such as when crossing a bridge or amplifier site induce time-varying fluctuations. The total PMD vector appears then as the concatenation of large vectors (representing the segments embedding all the sections between two sensing points) that are connected by hinges around which they rotate. These hinges correspond to the sensing points.

When the number of hinges is limited, the central theorem cannot be applied and the DGD distribution diverges from a Maxwellian distribution. First, it implies that the distributions (one per wavelength) are truncated with respect to a Maxwellian of same first moment. Second, different wavelengths exhibit different distributions, because the DGD in each dead section is wavelength dependent. Consequently, different DGD distributions are observed at different wavelengths, as presented on Figure 2. There the wavelength dependent DGD cumulative probability for 80 wavelengths and for a 12 spans, 11 hinges

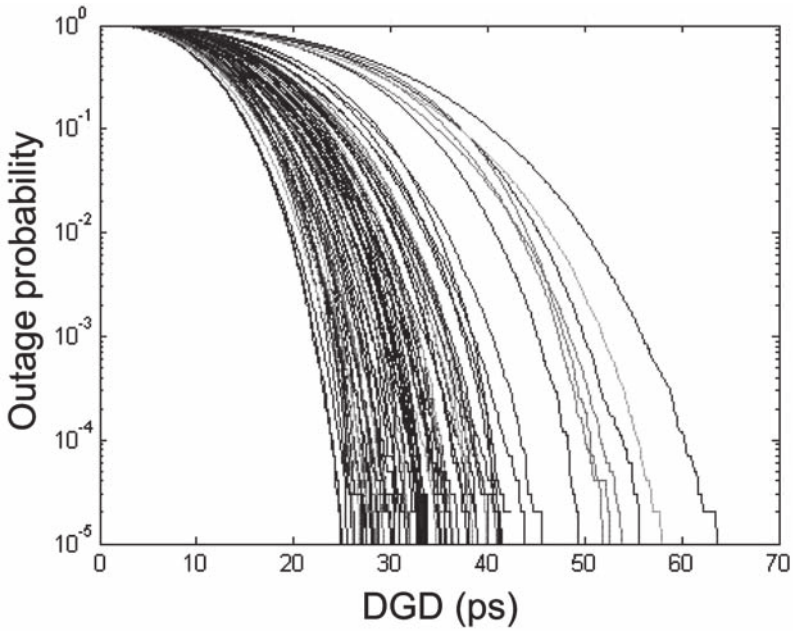


Figure 2. Outage probability as a function of DGD for 80 wavelengths, 10^5 independent hinges draws.

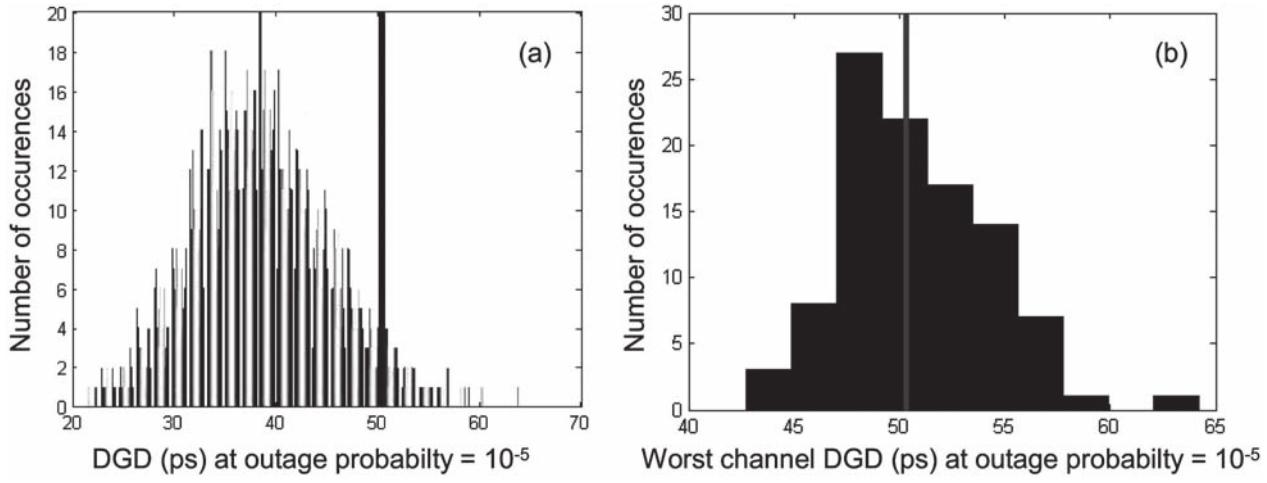


Figure 3. Distribution of the DGD at an outage probability of 10^{-5} for 100 independent segments DGD draws. (a) Outage DGD for all wavelengths (a color represents a wavelength) and (b) distribution of the worst channel DGD out of the 80 wavelengths. Histogram for the 100 draws. The bold line represents the outage probability expected under the Maxwellian assumption.

link are presented; 10^5 different hinges draws were realized. The outage is understood as the probability of observing a DGD greater than the value found on the x-axis. The PMD values per segment are [5.7, 8.3, 2.8, 0.6, 5.4, 0.2, 8.4, 3.9, 5, 0.5, 0.3, 0.5]. The maximum DGD observed for a 10^{-5} outage varies from 28 to over 60 ps. The previous simulation (based on hinge draws) was repeated for 100 draws of DGD per segment and per wavelength. The observed DGDs at a cumulative probability of 10^{-5}, referred as to "outage DGD" for the 80 wavelengths and the 100 independent segment DGD draws is presented on Figure 3a, where grayscale represents wavelengths. For the present example, it appears that the actual outage DGD is often smaller than the outage DGD predicted using the Maxwellian distribution (thick vertical line). This is a rather good piece of news. Unfortunately, the counterpart of this observation is that in some cases and for some wavelengths, the Maxwellian outage DGD is exceeded by the actual outage DGD. On a system point of view, this means that for the considered wavelength its DGD will exceed its "nominal outage DGD" during a longer time than expected/accounted for. At this wavelength the system does not respect its availability specification.

Although only few channels per draw are not compliant with their specifications, there is typically half of the draws for which at least a channel is not compliant Figure 3b. Actually, the section DGD draw presented on Figure 3b exhibits 6 noncompliant channels, and the 52nd out of the 100 section DGD draws had 13 noncompliant channels out of 80.

4. System Impact

4.1. Classical

The prediction of PMD impact is a great challenge for optical transmission systems. Usually the impact of PMD is quantified with system penalties and the main difficulty is to find the relevant and simple parameter that can be used to express the law for the system penalties. This parameter should include all PMD orders to simplify the mathematical expression of the law and to simplify experimental measurement. Finally, another remark should be pointed out: the truncation of PMD expansion (till second order for example) induces always an overestimated penalty due to the unbounded magnitude of the second order PMD in frequency. This subtle problem of penalty estimation is obviously linked

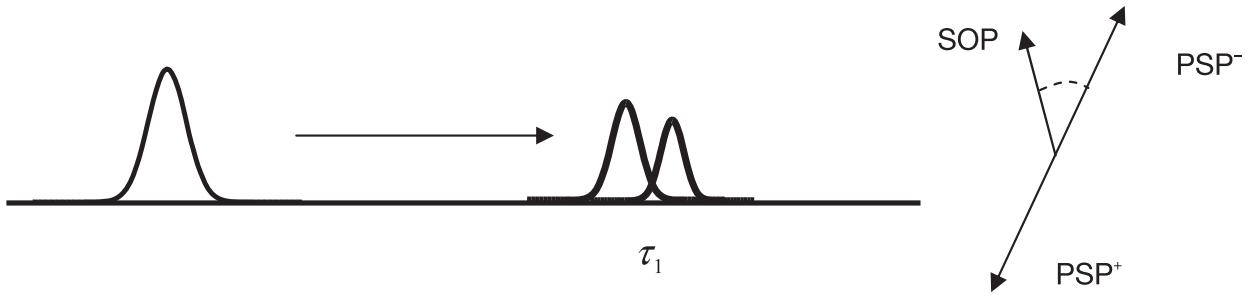


Figure 4. Projection in Stokes space of the input polarization on the two PSPs.

to that of PMD representation (or PMD emulation). Two parameters give interesting results and are usually taken to measure the impact of PMD: the rms pulse spreading and the string length (polarization state trajectory on the Poincaré sphere over the signal bandwidth). The degree of polarization (DOP) has also been extensively studied but the correlation between DOP and the bit error rate (BER) become evanescent in presence of high order PMD. This parameter is no longer used for the penalties estimation.

4.2. PMD Pulse Impairment

The impact of PMD on pulse distortion remains a very informative way to understand the physics of PMD. Basically, the propagation of a pulse in a first order section induces a pulse splitting in two replicas (see Figure 4).

The first and high order PMD can be emulated by a two sections emulator; Figure 5 shows the four replicas observed at the output with this configuration (for concatenation

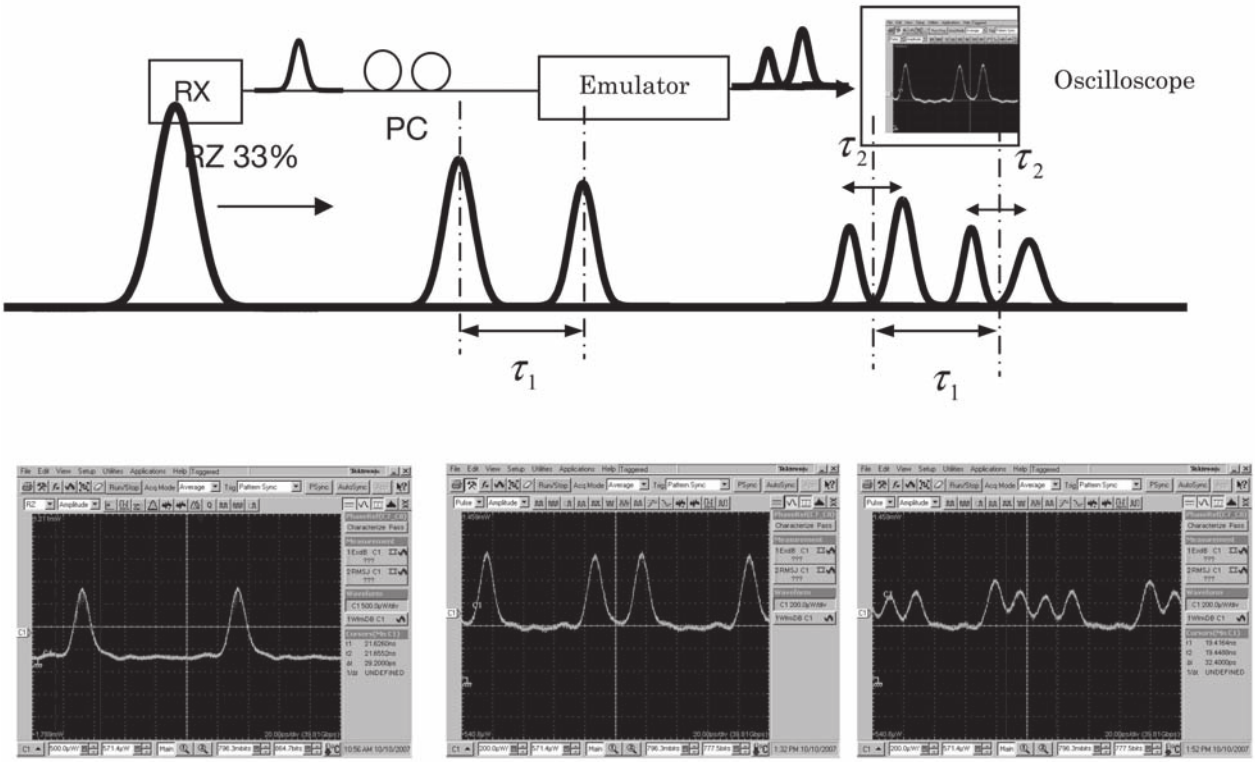


Figure 5. (a) Experimental set-up. (b) Principle of successive splitting of pulse with a two-section PMD emulator. (c) Oscillogram with 33% NRZ pulse. The pulse distortion is studied using a bit pattern: [1,0,0] at 40 Gb/s. The pulse duration is 8 ps. Right same pulse splitting with first order PMD (DGD = 37 ps).

rule see emulator section). The two replicas generated by the first section are equally split into two replicas in the second section.

4.3. Time Frequency Representation of Pulse Distortion Due to the PMD

The aim of this section is to illustrate the first and second order PMD effects on pulse propagation. The study of optical pulse propagation in optical fiber with PMD is interesting from both fundamental and applied views. For these studies the signals can be represented in either the time or frequency domains. In this work, pulse shape analysis is based on representation in a joint time-frequency (TF) plane [6]. TF representations have been used for the analysis of short light pulses and a well know application in optical domain is the spectrogram based on short time Fourier transform [7] and the Wigner distribution [8–9]. Here we will use the spectrogram which is a very simple tool from an implementation point of view and provides a reliable and easily interpretable representation of the TF energy distribution of an optical pulse. The TF representation does not provide more information than the time or the frequency representation but give this information in a more illustrative form.

Given a pulse envelope $A(z,t)$, this signal is represented in the frequency domain by:

$$\tilde{A}(z,\omega) = \mathrm{TF}[A(z,t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} A(z,t) \exp(-j\omega t) dt. \tag{14}$$

The spectrogram provides a short time Fourier analysis. A small portion of the signal centered at time t' is selected and Fourier transformed, giving then a "time dependent spectrum," the modulus of which is called the spectrogram. In order to select a portion of the signal, we use a gate or window function $G(t)$. The spectrogram $SP(\omega,t')$ and the distribution $W(\omega,t')$ are:

$$SP(\omega,t') = |W(\omega,t')| = \left| \int A(z,t) G(t-t') \exp(-j\omega t) dt \right|. \tag{15}$$

The input optical pulse is assumed to be Gaussian and a Gaussian function is selected as a mobile window. This choice for the window is known to optimize the resolution between time and frequency:

$$\begin{aligned} A(0,t) &= \exp\left(-\frac{1}{2}\left(\frac{t}{\delta t_s}\right)^2\right); \\ G(t) &= \exp\left(-\frac{1}{2}\left(\frac{t}{\delta t_g}\right)^2\right). \end{aligned} \tag{16}$$

The practical implementation will use discrete Fourier transform. The resolution in time and frequency depends on the sampling frequency used for the discrete Fourier transform and the window width. With this basic analysis without any specific corrections, no re-assignments of time and frequency are used; we obtain only an estimated value for time and frequency, but the general pulse shape is correct.

In this study, the PMD effects analysis is based on the Bruyère's model [10–12]. For simplicity we consider that PSPs at the central frequency are aligned with S_1 in Stokes

space. Following Bruyère's model, the PMD operator is described by the matrix $M(\omega)$:

$$M(\omega) = R(\omega)^{-1} D(\omega) R(\omega)$$

$$M(\omega) = \begin{bmatrix} \cos(k\omega) & -\sin(k\omega) \\ \sin(k\omega) & \cos(k\omega) \end{bmatrix} \begin{bmatrix} \exp(-j(\tau_0\omega + \tau_\omega\omega^2)/2) & 0 \\ 0 & \exp(j(\tau_0\omega + \tau_\omega\omega^2)/2) \end{bmatrix} \cdot \begin{bmatrix} \cos(k\omega) & \sin(k\omega) \\ -\sin(k\omega) & \cos(k\omega) \end{bmatrix} \quad (17)$$

where τ_0 is the DGD, τ_ω is the PCD, and k is linked to the depolarization $|\vec{p}_\omega| = 4k$ at central frequency.

In our work, the input pulse polarization is linear and its orientation makes an angle θ with the PSPs at the central frequency.

First we calculate the input pulse spectrogram, depicted in Figure 6a. The first order PMD is only considered in a first step, therefore the matrix $R(\omega)$ reduces to the identity matrix. The operator $D(\omega)$ represents a delay of $\pm\tau/2$ for the projections onto the PSPs, giving at the output two replicas of the input pulse (Figure 6b). In a second step, the PCD is set to zero while $4k = 30$ ps. We can see on the spectrogram of Figure 7a that in this case the effect of matrix $R(\omega)$ and $R(\omega)^{-1}$ appears as a periodic amplitude variation with frequency. Figure 7b presents the same pulse after Fourier Transform inversion; therefore the frequency domain is transformed back in time domain. The six replicas as determined by the Bruyère's model are clearly visible on Figure 7b.

PCD and chromatic dispersion correspond to a frequency dependent time delay. Now we fix also a value for the PCD and the output pulse spectrogram is shown in Figure 8a and after Fourier transform inversion in Figure 8b. Figures 9a and 9b show the same configuration with a non zero chromatic dispersion. It can be observed that combination of PCD and chromatic dispersion changes the pulse distortion and therefore the PMD impact.

Figure 10 presents the pulse shape when only depolarization is present and the input pulse injection is aligned with a PSP at the central frequency. The output pulse is then split up into two replicas at $\pm\tau_0/2$, which indicates that energy is present on the two PSPs (otherwise only one pulse would be observed).

This time frequency analysis allows illustrating the impacts of the different PMD parameters on the pulse shape. We have to carry on more work in order to optimize the

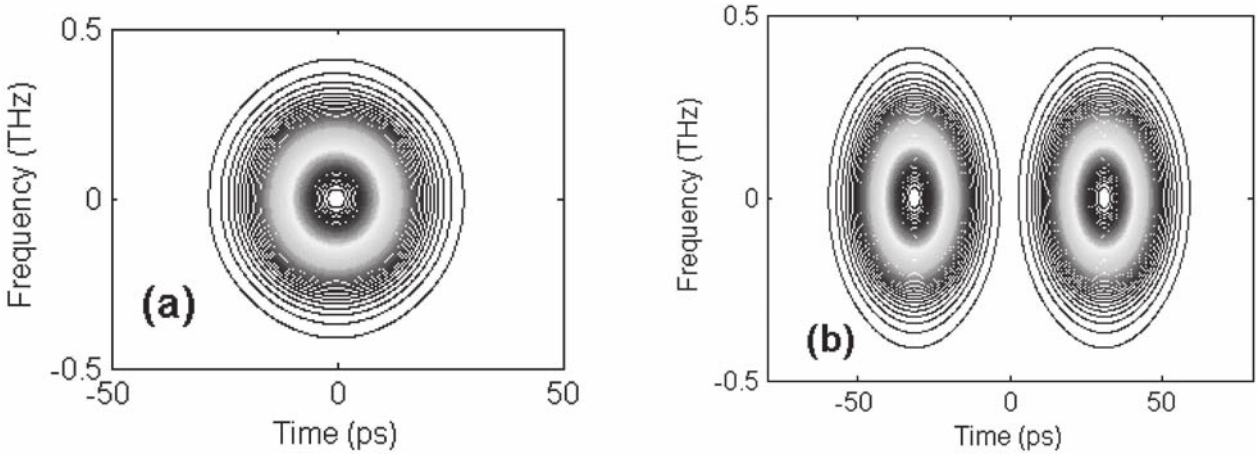


Figure 6. (a) Input pulse spectrogram and (b) output pulse spectrogram with DGD = 60 ps, PCD = 0, $p_\omega = 0$.

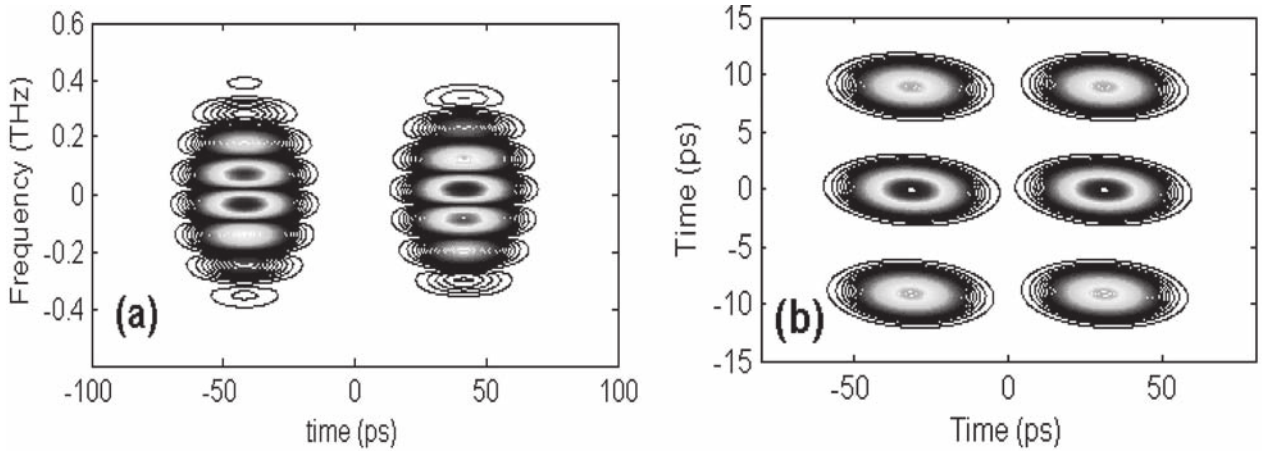


Figure 7. (a) Output pulse spectrogram and (b) Inverse transform of spectrogram with DGD = 60 ps, PCD = 0, $p_\omega = 20$ ps.

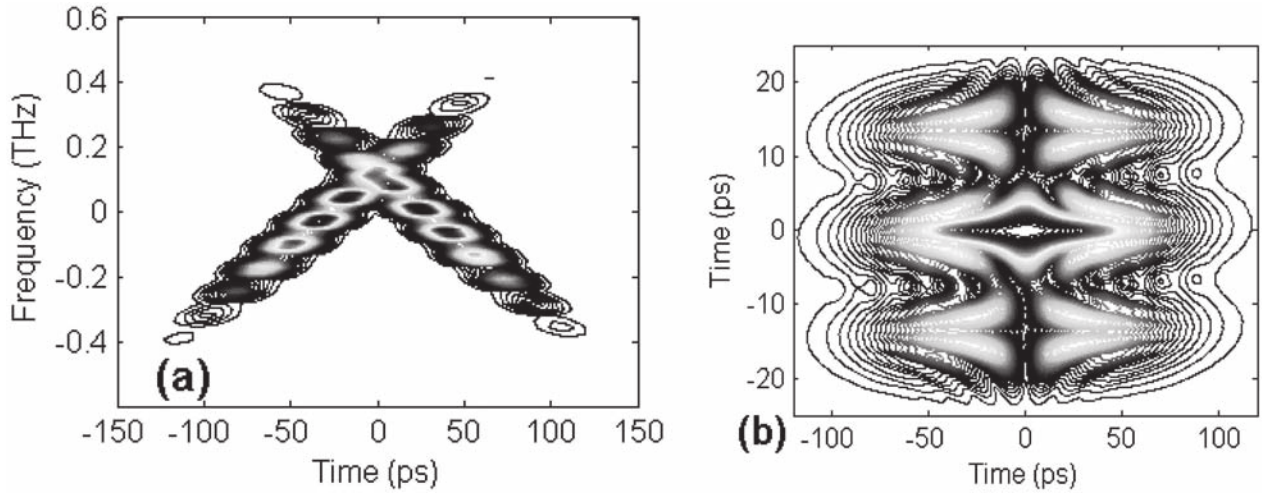


Figure 8. (a) Output pulse spectrogram and (b) inverse transform of spectrogram with DGD = 60 ps, PCD = 5 10^3 ps^2, $p_\omega = 20$ ps.

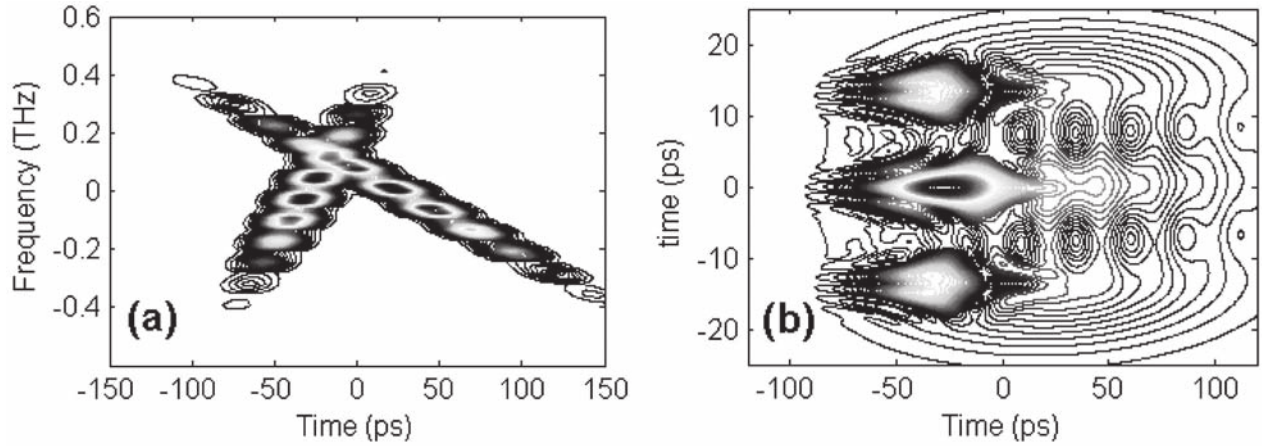


Figure 9. (a) Output pulse spectrogram and (b) inverse transform of spectrogram with DGD = 60 ps, PCD = 5 10^3 ps^2, $p_\omega = 20$ ps, chromatic dispersion = 400 ps/nm.

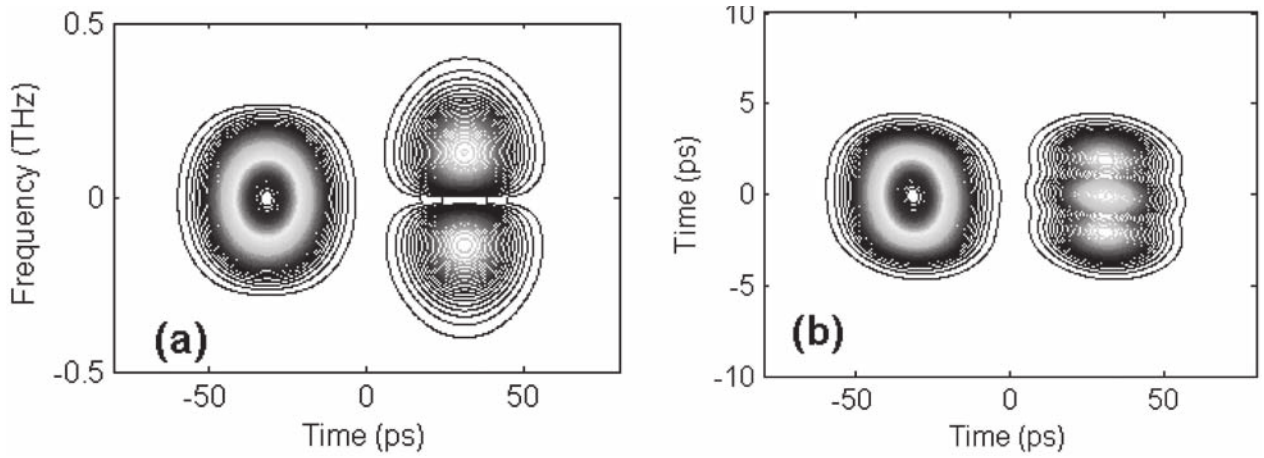


Figure 10. (a) Output pulse spectrogram and (b) inverse transform of spectrogram with DGD = 60 ps, PCD = 0, $p_\omega = 5$ ps, input polarization aligned with a PSP.

resolution and precision on time and frequency and to study in more details the relation between the output and input distribution in different cases.

5. Emulation

The basis of PMD emulation is to create a realistic, reproducible and steady PMD source with a simple set of parameters to test optical transmission system sensitivity to PMD impairments.

5.1. Principle of PMD Emulation

The issue of PMD emulation is how to generate a realistic frequency dependent polarimetric operator which simulates the polarimetric effects of optical fiber [12]. To do that, the first step is to define the relevant parameters of PMD (first and second order for example) and then using a mathematical inversion process to obtain the emulation operator. This operator could be defined for example with his Jones matrix, $T(\omega)$. Extracting the emulator Jones matrix with only first order PMD and PCD terms is formally very simple by using the equation:

$$T_\omega T^{-1} = -\frac{i}{2}\vec{\tau}.\vec{\sigma}. \tag{18}$$

Finally, we find the Jones matrix of the emulator:

$$T(\omega) \equiv e^{-\frac{i\omega}{2}\vec{\tau}(\omega).\vec{\sigma}}. \tag{19}$$

In the case of high order PMD with frequency-dependent cross coupling between PSP vectors (e.g., depolarization terms), the calculations are significantly more complex. A method has been proposed [13] to solve rigorously this equation, including second order PMD. However, this method cannot be used to carry out practicable PMD emulator architectures.

PMD emulators are basically constituted by the assembling of group delay blocks and birefringent waveplates and all architectures use the concatenation rules to generate predictable, stable, and repeatable PMD parameters (first and second order PMD). Emulators can be divided in two groups: those which are compatible with a WDM use and gives a realistic polarimetric behavior over a large spectral range, and those which

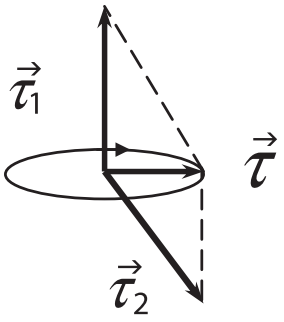


Figure 11. Geometrical concatenation of PMD vectors for the planar sweep model.

are devoted to a single channel use and should emulate PMD behavior over the signal spectral bandwidth.

5.2. Architectures of Single Channel of PMD Sources

Focusing on single channel PMD sources, two kinds of architecture that can emulate higher orders PMD are identified [14–15]. The first architecture represented on Figure 11, called "planar sweep model," derives directly from the application of concatenation rule of two birefringent sections: $\vec{\tau}_{out} = \vec{\tau}_2 + R(\omega\tau_2)\vec{\tau}_1$.

The resulting PMD vector rotates around the PMD vector of the second birefringent section. The first and second orders are simply expressed as

$$\begin{aligned} \vec{\tau} &= \vec{\tau}_1 + \vec{\tau}_2 \\ \vec{\tau}_\omega &= \vec{\tau}_1 \times \vec{\tau}_2 \end{aligned} \tag{20}$$

This model is able to emulate first order and depolarization term. The PCD term is absent due to the fact that the module of the resulting PMD vector does not depend on frequency.

The second architecture derives from the Bruyère's model [15]. This emulator is made up of the concatenation of three elements corresponding to the three terms of the product in Eq. (17). The circular birefringent waveplates introduce a rotation versus frequency of the linear birefringence. Straightforward calculations give a spectral dependence of the PMD vector as depicted is the Figure 12. The eigenvectors of the Jones matrix T and the PMD vector do not coincide over all spectral range but only at the central frequency, and then the trajectory of PMD vector no longer stays in the equatorial plan as the birefringent vector.

This emulator is able to simulate all PMD orders (PCD and depolarization + higher orders). A particular design of PMD compensator has been proposed by Shtaif et al. [10]. This compensator can be obviously used as a PMD emulator with a decreasing degree of freedom numbers. His architecture is based on the Bruyère's model and it offers the possibility to induce simultaneously first and higher order PMD parameters that emulate realistic PMD fiber.

6. OPMDC

Compensation of channel with distorting effects is a natural step in transmission engineering. In the case of PMD this compensation can be done by two ways, either in optical layer or by electronic means.

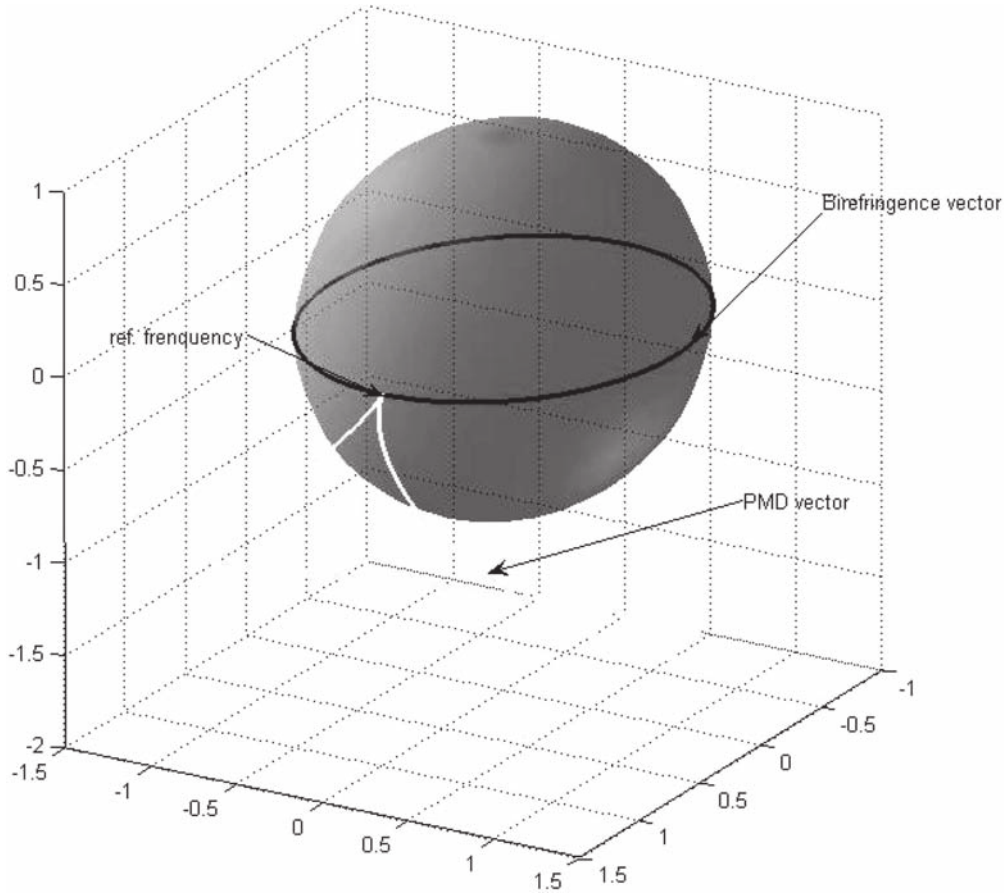


Figure 12. Trajectory of PMD vector in the Bruyère based model emulator. All sections have the same DGD.

6.1. Principle of PMD Optical Compensation

The concept of optical PMD compensation (OPMDC) is simple: it consists to add a "birefringent" element that can be, more or less, the inverse operator to the one of fiber. This function is carried out by the concatenation of a polarization controller (PC) and first order PMD sections. However a deeper insight lets numerous difficulties to appear carrying out this highly complex functionality.

The first difficulty is linked to the architecture of the compensator and this problem appears very similar to that of the emulator: how to create the compensation operator as the fiber with a reduced number of relevant parameters (degrees of freedom). The second difficulty is linked to the constraint on the number of degrees of freedom (DOF) to maintain a sufficient efficiency to the compensator. Too many parameters induce a great complexity of the tracking algorithms. Moreover, the response time of the compensator obviously grows when increasing the DOF number. Another difficulty is linked to the nature of the monitoring signal: DOP, RF (Radio-frequency, i.e., tones in the electrical spectrum of the received spectrum), BER, Eyes opening, etc. More precisely, the great challenge to carry out an efficient OPMDC is to extract a relevant monitoring signal strongly correlated to the BER, which determines the quality of the received digital data stream, seen from the operator. The third difficulty is related to the complexity of efficient and endless tracking algorithms with several sub-optimums (operating points) and related possible false locks.

The last difficulty is to obtain the full endless working for the optical compensator. This ability results from both the polarization controller (PC) behavior and from the algorithm efficiency. The PC should be able to continuously transform any input polarization to each other. This property is either intrinsic to the PC (endless rotatable waveplates) or

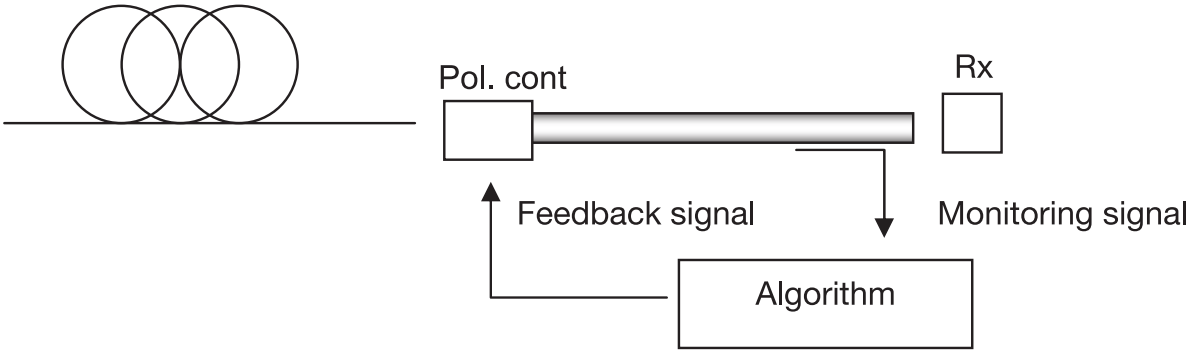


Figure 13. Architecture of first order OPMDC.

is obtained by specific driving algorithms that include some reset operations compatible with the endless control.

6.2. OPMDC Architectures

The simple architecture for the PMD compensator is the first order OPMDC made up of a PC and a fixed first order PMD section as on Figure 13.

The PC orients the PMD vector of OPMDC. The resulting PMD vector is then the concatenation of two PMD vectors and the objective is to align the resulting PMD vector onto the input SOP. With this PMD compensation scheme, the monitoring signal classically used is the DOP.

First and high order OPMDC has been also proposed based on two or three PMD sections. These devices are designed to compensate a larger range of PMD effects, but the driving of such a system becomes significantly more complex due to the large number of degrees of freedom.

7. Electronic PMD Compensation (EPMDC)

Digital signal processing (DSP) takes place either at the emitter between the Forward Error Correction (FEC) coder and the modulator, or at the receiver between the detection stage and the FEC decoder. For PMD compensation, it is implemented at the receiver, because it is required to adaptively track the channel changes. Doing so at the emitter would require a feedback control signal from the receiver side to the emitter which is quite complex. Information about the current system performance is either internally determined by the DSP or received (typically from the FEC).

The goal of EPMDC is to make the full digital channel from coder to decoder more robust to PMD (and any distortion) induced by the fiber optical channel. Different strategies can be embraced to reach this goal, the theoretical background of which is not specific to the optical channels, because these equalizers are widely used in a lot of modern communication systems. First DSP techniques, feed forward equalizer (FFE) and decision feedback equalizer (DFE) aim at removing the inter symbol interference (ISI) resulting from pulse overlapping found at the decision time. It is equivalent to re-opening the eye at the decision instant. Other techniques aim at making the decision process more robust towards ISI, through a more optimal processing of the received signal. These techniques are nonlinear cancellation (NLC) and maximum likelihood sequence estimation (MLSE).

7.1. *Linear Transversal Filter: Feed Forward Equalizer and Decision Feedback Equalizer*

A linear transversal filter realizes the addition of weighted and delayed copies of the received signal (Figure 14, left).

The goal of the linear addition is to remove as much ISI as beneficial at the sampling time. It does not mean that the equalizer tries to recover exactly the sent eye but rather to open it at the sampling time.

The FFE treats similarly ISI induced by previous or future symbols. Nevertheless, previous symbols could be advantageous processed because they are associated to already decided bits (assuming errorless decision). Decided signal samples are not corrupted by noise. Then the ISI produced by these symbols can be suppressed without any noise enhancement. A DFE implements the observed advantages by addressing the ISI caused by previously decided symbols. The linear addition of weighted copies of (noiseless) signal samples after decision replaces the linear addition of signal copies. Usually FFE and DFE are associated for optimal performance.

A drawback of the DFE is its catastrophic behavior at high BER. An erroneous decision weakens the ISI cancellation by removing/adding a wrong amount of energy, causing then the so-called error propagation phenomenon. Such behavior appears however at BER, which transmission systems in normal conditions do not operate at, around 10^{-2}.

Furthermore, high speed implementation of such a device is quite challenging as it requires analogue feeding back and addition signal within one bit period without distorting too much of the signal. This becomes quite tedious already for 10 Gbit/s applications. That is why a high speed implementation was proposed that allows further nonlinear cancellation of ISI.

7.2. *Nonlinear Canceller*

The nonlinear canceller (NLC) [16] is an adaptation of a DFE. It constitutes an elegant way of overcoming the analogue feedback issue required in the typical DFE implementation. It offers further nonlinear equalization as the previous bits dependent subtraction/addition is any function of these previous bits (not just linear).

The NLC is built as 2^N independent decision flip-flops, having independent thresholds (Figure 15), where N represents the number of previously decided bits that are taken into account. The 2^N decision signals feed a switch that selects the current correct decision device according to the previously decided bits. The previous decided sequence can be understood as an address. There is no analogue feedback but a digital one used to address the switch. 2^N decisions take place at each clock cycle but only one is delivered after the switch. Adaptation of the equalization reduces to adaptation of the

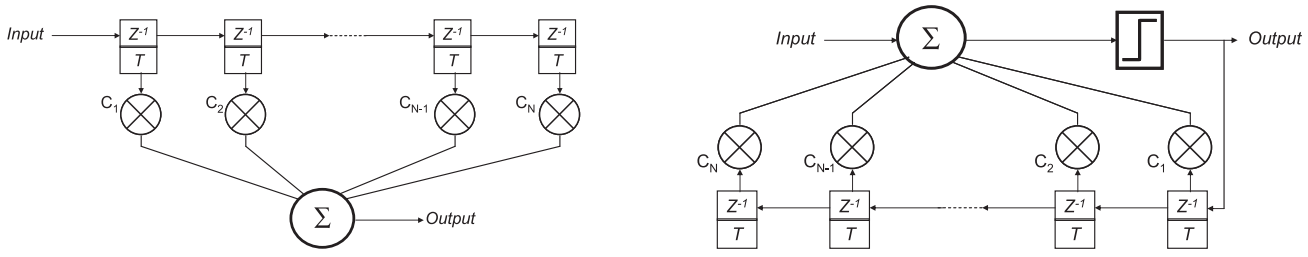


Figure 14. Schematic of a feed forward equalizer (left) and a decision feedback equalizer (right), T represents the bit period.

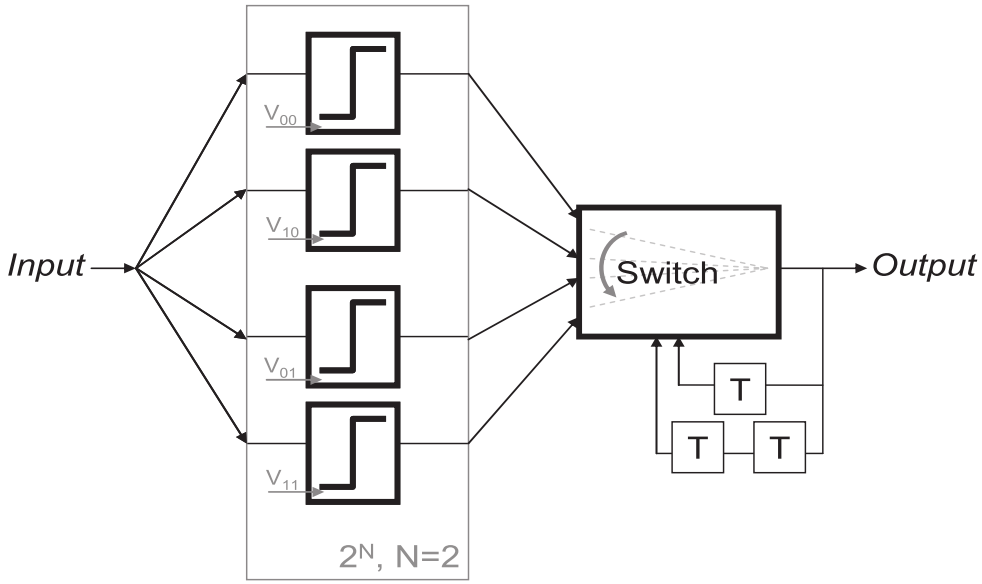


Figure 15. Schematic of a 2 bit nonlinear canceller.

2^N thresholds. These can be arbitrarily chosen offering nonlinear cancellation. The 2^N independent thresholds replace the weighting coefficients as providers of the mitigation dynamicity. The NLC catastrophic behavior at high BER is the same as for DFE.

The main drawback of a NLC resides in its limitation to the cancellation of ISI induced by previously received symbols. In a sense, it can only cope with half of the cases. On the other hand, the NLC provides nonlinear ISI cancellation and does not cause noise enhancement. It is therefore attractive to implement a pseudo nonlinear canceller addressing the ISI due to symbols to be decided. This requires estimating future symbols. This can be done for example as a tentative decision that runs in parallel of the main decision scheme. Figure 16 represents a possible implementation of such a pseudo NLC that we refer as to multi threshold receiver (MTR). For the sake of clarity, a 1-1 bit MTR

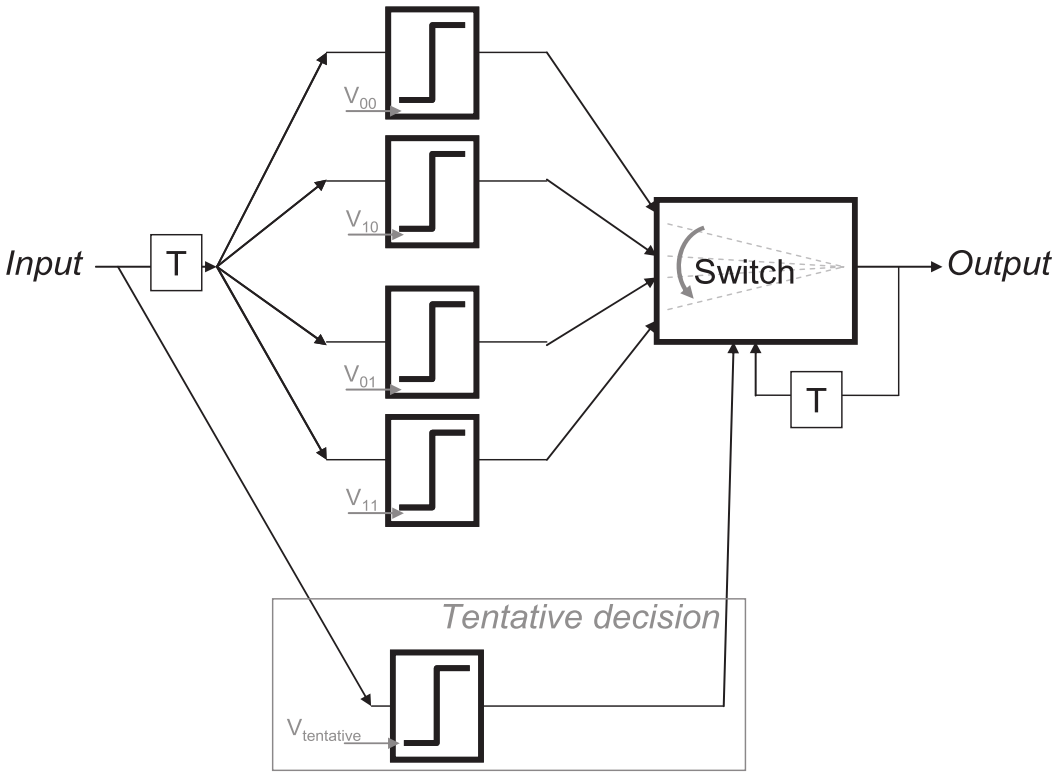


Figure 16. Schematic of a MTR.

is presented. 1-1 means that the MTR takes decisions conditioned to one previously decided bit and to one estimated future bit.

Guessing the next bit contributes also to the catastrophic behavior and in a stronger manner. Well-optimized thresholds ensure that the performance is not any worse than the performance of a NLC. The threshold of the tentative decision circuit can typically be chosen as the average of V_{01} and V_{10}.

7.3. Maximum Likelihood Sequence Estimation

In a sense, linear filtering prepares signals for the decision. On the other hand, NLC principle resides in making the decision "context aware" by adapting the decision threshold according to the surrounding bits and to the ISI level. A MLSE receiver goes one step further in taking into account the context. As the transmission channel spreads the impulse response over several timing intervals, the decision process can take advantage of the signal in the surrounding timing intervals, which contains information useful for the decision. The MLSE receiver does so by determining the most likely (or most probable) data, conditionally to the sequence at the receiver input.

MLSE aims at finding the most probable/likely sent sequence given the received signal. It moves away from a bitwise decision to a sequence decision. Thinking in terms of "most likely" induces the quantitative evaluation of the likelihood of each sequence. In the ordinary binary case (absence of ISI), no calculation is needed, although the receiver maximizes also the likelihood (maximization reduces to the comparison to a threshold). The MLSE is reduced to the classical bit by bit decision receiver in this trivial case.

The goal of MLSE is to associate to an N bits long received sequence the most likely binary sequence given the received signal of duration NT. It is equivalent to comparing the received sequence to the 2^N possible reference sequences and to compute for each of them the "distance" between received and reference sequences (measuring the sequence likelihood conditionally to the received signal). The closest or most likely is chosen and the associated binary sequence is the decision result.

Let us take an example with $N = 3$. This specializes in finding the most likely sequence out of 8 [000, 001, . . . , 111]. If *xyz* represents any of the triplets, each sequence is the corner of a unitary cube in a 3D space. A received $3T$ long sequence appears after sampling as a triplet of real for example [0.1 0.9 0.2]. Under additive white Gaussian noise (AWGN) assumption the most likely sequence is just the closest in terms of Cartesian distance. Since direct detection the fiber optic channel is not AWGN, Cartesian distance is suboptimal, but it will be used first to ease understanding. Figure 17 presents a 2^7 PRBS noisy sampled sequence (crosses) in this 3D representation.

If ISI is present, the reference triplets (black circles) move and are no more on the cube corners (Figure 18). Decision is equivalent to determining for each cross the closest cube corner rather than considering just the y direction position.

7.4. Comparative Performance

An experimental test bed was built to determine the relative performance of the two most promising DSP techniques that are NLC and MLSE. The test bed includes noise, CD, and PMD (up to second order) emulation. Polarization control ensures further the correct estimation of the launch SOP impact. The joint statistics of first and second order PMD were used to generalize the measured performance to any PMD. Figure 19 presents the

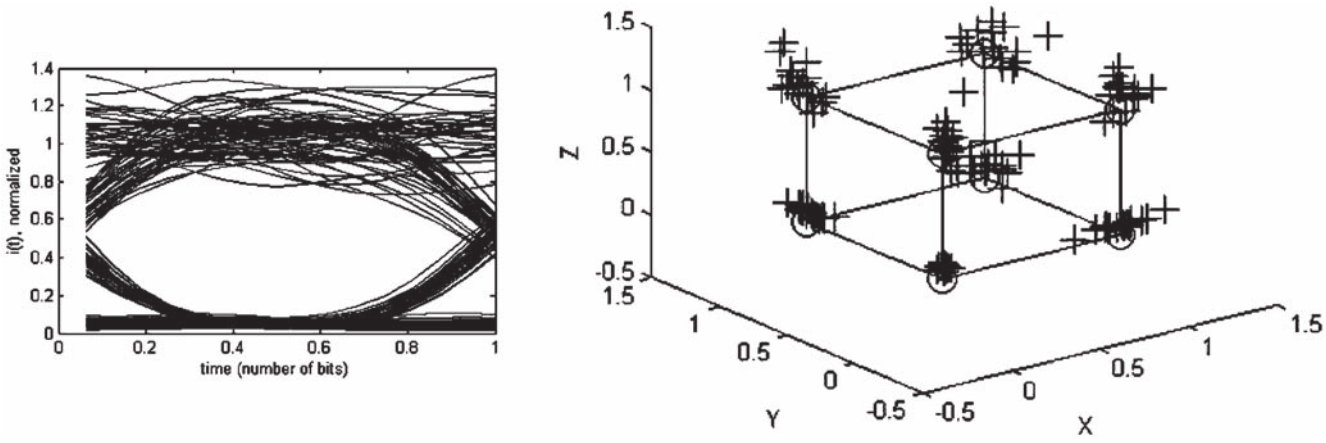


Figure 17. Eye diagram for a noisy 2^7 PRBS, NRZ. Left hand side corresponding triplet samples representation in Cartesian coordinates. Circles are the reference triplet, the red crosses are the samples.

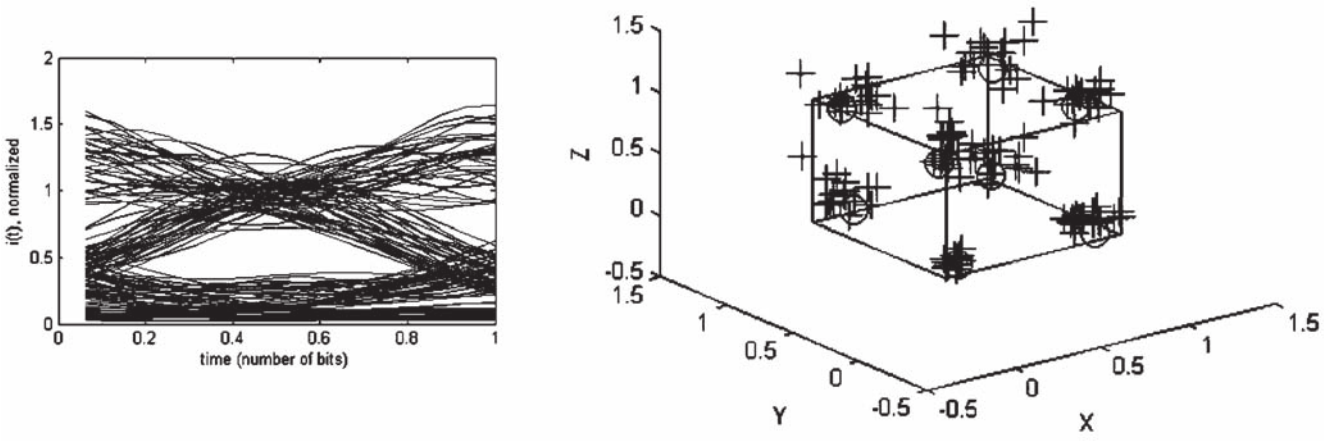


Figure 18. Similar to Figure 17 for non-negligible ISI.

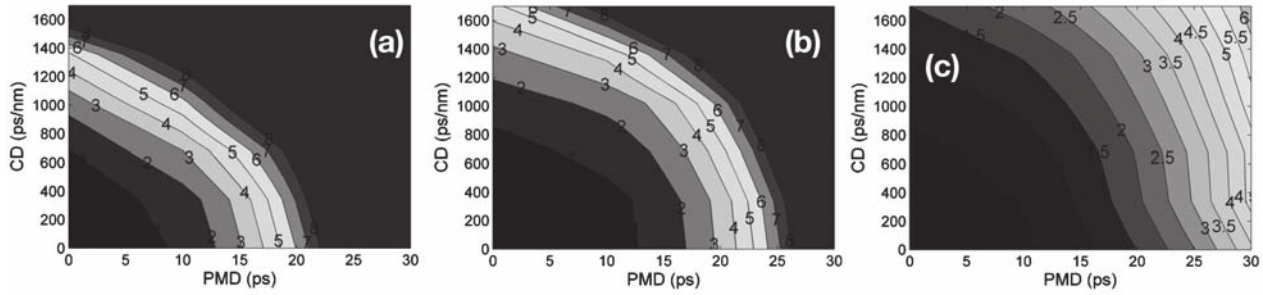


Figure 19. OSNR penalty (dB) for 10^{-5} outage probability as a function of CD and PMD; for three different transponders: Standard, eNLC, and MLSE. Other conditions: Back to back, 10.7 Gbit/s NRZ, SOP penalty dependency taken into account.

induced CD and PMD penalty for three receivers: the standard (optimized threshold) receiver (used as reference), the MTR, and the MLSE.

8. Conclusions

In this article we have introduced some fundamental aspects of PMD: its system impact, statistics, and emulation. The core of this article resides in the proposition of a new representation of its impact by time-frequency analysis. This powerful tool proved to deepen our understanding of PMD impact, especially higher order. Finally, solutions

to overcome PMD, either optical or electrical, were described. For the most promising solutions (assuming 10 Gbit/s systems), their robustness was assessed. The maximum likelihood sequence estimator appeared as the most powerful electronic dispersion compensator.

Next, studies should assess the performance of electronic mitigation when associated to optical coherent detection. In this context, the optical channel is linear again and equalization should provide much better performance.

Finally, the time-frequency representation of PMD impact has only been touched on and further work is necessary to extract all the benefits of this new tool.

References

1. Kaminov, I., and Li, T. 2002. *Optical Fiber Telecommunications IVB, Systems and Impairments*, chapter 15. Academic Press.
2. Brodsky, M., Frigo, N. J., Boroditsky, M., and Tur, M. 2006. Polarization mode dispersion of installed fibers. *Journal of Lightwave Technology* 24(11)4584–4599.
3. Gordon, J. P., and Kogelnik, H. 2000. PMD fundamentals: Polarisation mode dispersion in optical fibers. *Proceedings of the National Academy of Sciences* 97(9):4541–4550.
4. Poole, C. D., and Wagner, R. E. 1986. Phenomenological approach to polarisation dispersion in long single-mode fibres. *Electronics Letters* 22(19):1029–1030.
5. Francia, C., Bruyère, F., Penninckx, D., and Chbat, M. 1998. PMD second order effects on pulse propagation in single mode optical fibers. *Photonics Technology Letters, IEEE* 10(12):1739–1741.
6. Kogelnik, H., Nelson, L. E., and Gordon, J. P. 2003. Emulation and inversion of polarisation mode dispersion. *Journal of Lightwave Technology* 21(2):482–495.
7. Kogelnik, H., Nelson, L. E., Gordon, J. P., and Jopson, R. M. 2000. Jones matrix for second order polarisation mode dispersion. *Optics Letters* 25(1):19–21.
8. Forestieri, E., and Vincetti, L. 2001. Exact evaluation of the Jones matrix of a fiber in the presence of polarisation mode dispersion of any order. *Journal of Lightwave Technology* 19(12):1898–1909.
9. Orlandini, A., and Vincetti, L. 2001. A simple and useful model for Jones matrix to evaluate higher order polarisation-mode dispersion effects. *Photonics Technology Letters, IEEE* 13(11):1176–1178.
10. Shtaif, M., Mecozzi, A., Turr, M., and Nagel, J. A. 2000. A compensator for the effects of high-order polarisation mode dispersion in optical fibers. *IEEE PTL* 12(4):434–436.
11. Kim, N.-Y., and Park, N. 2005. Independently tunable first- and second-order polarisation-mode dispersion emulator. *Photonics Technology Letters, IEEE* 17(3):576–578.
12. Orlandini, A., and Vincetti, L. 2003. Comparison of the Jones matrix analytical models applied to optical system affected by high-order PMD. *Journal of Lightwave Technology* 21(6):1456–1464.
13. Kasturia, S., and Winters, J. H. 1991. Techniques for high-speed implementation of nonlinear cancellation. *IEEE Journal on Selected Areas in Communications* 9(5):711–717.
14. Cohen, L. 1989. Time-frequency distributions—A review. *Proceedings of the IEE* 88(7):941–981.
15. Kane, D. J., and Trebino, R. 1993. Single-shot measurement of the intensity and phase of an arbitrary ultrashort pulse by using frequency-resolved optical gating. *Optics Letters* 18(10):823–825.
16. Azana, J. 2005. Time-frequency (Wigner) analysis of linear and nonlinear pulse propagation in optical fibers. *Eurasip Journal on Applied Signal Processing* 2005(10):1554–1565.
17. Helczinski, L., Anderson, D., Fedele, R., Hall, B., and Lisak, M. 2002. Propagation of partially incoherent light in nonlinear media via the Wigner transform distribution. *IEEE Journal of Selected Topics in Quantum Electronics* 8(3):408–412.