New Branch Metrics for MLSE Receivers Based on Polarization Diversity for PMD Mitigation

Gabriella Bosco, *Member, IEEE*, Pierluigi Poggiolini, *Member, IEEE*, Monica Visintin, *Member, IEEE*, Li Liangchuan, and Chen Ming

***Abstract*—In this paper, we analyze the performance of polarization-diversity receivers employing maximum-likelihood sequence estimation (MLSE) to combat the effects of polarization mode dispersion (PMD) for both the intensity-modulation direct detection (IMDD) and duobinary modulation formats at 43 Gb/s. We propose and study several parametric branch metrics, showing that the use of MLSE receivers based on polarization diversity with appropriate branch metrics allows to cancel out the high penalty incurred by standard MLSE receivers in the presence of PMD.**

***Index Terms*—Maximum-likelihood sequence estimation (MLSE), optical communications, polarization mode dispersion (PMD).**

I. Introduction

MAXIMUM-LIKELIHOOD SEQUENCE-ESTIMATION (MLSE) receivers (Rx) have been shown to be effective for the mitigation of chromatic dispersion and single-channel nonlinear effects [1]–[8]. On the contrary, they are not equally effective in mitigating polarization mode dispersion (PMD) [9]–[11], which has emerged as one of the main limiting factors in optical fiber transmission at data rates of 40 Gb/s and beyond. In particular, a very large worst case penalty of about 5.5–6 dB is incurred by MLSE Rxs when the differential group delay (DGD) approaches one bit time, irrespective of the number of MLSE states. In this paper, we show that the use of a suitable polarization diversity Rx [12], [13], also called "Stokes receiver," together with an MLSE processor based on appropriate parametric branch metrics, can overcome this problem.

In Section II, we analyze the performance of a standard direct-detection MLSE receiver, establishing a reference benchmark for a comparison with the polarization diversity solutions. We consider two different modulation formats: intensity-modulation direct detection (IMDD) and duobinary, in its phase-shaped binary transmission (PSBT) configuration [14], both running at 43 Gb/s.

In Section III, we describe the polarization diversity receivers, which are able to extract the information about the state of polarization (SOP) of the signal at their input and to use this information to improve the system performance.

In Section IV, we propose several parametric branch metrics to be used with the receivers described in Section III. The channel estimation procedure is described in Section V.

Section VI is devoted to the analysis of the performance of the polarization diversity Rx. Some results are shown on the dependence of the system performance on the SOP of the signal at the input of the receiver, which may be present for some kind of branch metrics, and on the direction of the PMD axis in the Stokes space. After identifying the worst case scenario for each kind of analyzed MLSE processor structure, we finally show the system performance in terms of OSNR needed to achieve bit error rate (BER) $= 10^{-3}$ for both IMDD and PSBT modulation formats.

The choice of the number of states N_s used by the MLSE processors is a compromise between performance and complexity. The higher the number of states, the greater is the amount of DGD that can be tolerated by the system. The following relationship approximately holds between N_s and the maximum DGD $\tau_{\max}$

$$\tau_{\max} \simeq \log_2(N_s) \cdot T \tag{1}$$

where T is a bit time. In our analysis, we chose to use eight-state MLSE processors, which yield a theoretical insensitivity to DGD up to $3T$.

II. Standard Single-Input MLSE Receiver

As a first step, to establish a precise benchmark, we look at the case of a standard Rx using an MLSE processor. We do so by carrying out detailed simulations. The system schematics is shown in Fig. 1. The transmitter (Tx) consists of a 43-Gb/s $(2^{16}-1)$ pseudorandom binary sequence (PRBS) generator, followed by a five-pole Bessel filter with bandwidth equal to 0.28 R_b for PSBT and 0.75 R_b for IMDD, where $R_b = 43$ Gb/s is the bit rate. The filter is followed by an ideal Mach–Zehnder modulator with infinite extinction ratio, driven in the range $[-V_\pi, V_\pi]$ for PSBT and $[0, V_\pi]$ for IMDD.

The SOP of the optical field can be represented by a point on the Poincaré sphere (see Fig. 2), characterized by an azimuth angle α and a zenith angle θ. The PMD channel introduces a

Manuscript received January 15, 2009; revised June 12, 2009. First published July 07, 2009; current version published September 10, 2009. This work was supported by the European Union within the BONE-project ("Building the Future Optical Network in Europe"), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

G. Bosco, P. Poggiolini, and M. Visintin are with OptCom, Dipartimento di Elettronica, Politecnico di Torino, 10129 Torino, Italy (e-mail: gabriella.bosco@polito.it; pierluigi.poggiolini@polito.it; monica.visintin@polito.it).

L. Liangchuan is with the Optical Network Research Department, Huawei Technologies Company, Ltd., Shenzhen 518129, China (e-mail: leelc@huawei.com).

C. Ming is with European Research Center, Huawei Technologies Deutschland GmbH, Bonn 53129, Germany (e-mail: minc@huawei.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2009.2026720

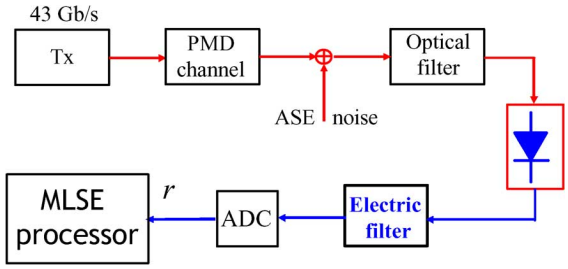

Fig. 1. Standard single-input MLSE Rx.

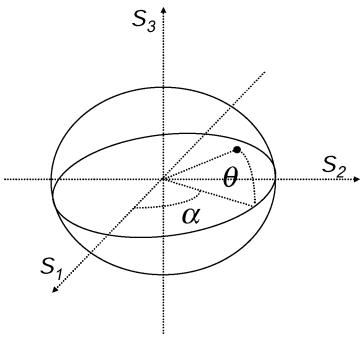

Fig. 2. SOP representation on the Poincaré sphere. $\alpha \in \{-180°, 180°\}$ is the azimuth angle and $\theta \in \{-90°, 90°\}$ is the zenith angle.

DGD equal to τ between the two principal states of polarization (PSPs) [16], emulating first-order effects only.

Amplified spontaneous emission (ASE) noise loading is performed at the Rx to obtain the desired optical signal-to-noise ratio (OSNR), which is computed over a 0.1-nm bandwidth, taking into account noise on both polarizations. A 50-GHz second-order super-Gaussian optical filter is used as Rx optical filter. The Rx photodetector and electrical circuitry are assumed ideal and noiseless. The postdetection (PD) filters are five-pole Bessel filters with bandwidth equal to R_b. The output of the PD filter is sampled and quantized and then sent to an eight-state MLSE processor. The total number of simulated bits is $(1.25 \cdot 2^{18})$ and performance is evaluated through direct error counting. The first 2^{16} bits are used for channel estimation, while the last 2^{18} have been considered for performance evaluation.

The MLSE processor uses the following Euclidean branch metric [20]–[22]:

$$m_i = \sum_{k=1}^{K} [(r_k)^{\alpha} - \mu_{i,k}]^2 \tag{2}$$

where α is a nonlinear distortion exponent (to be optimized), and $\mu_{i,k}$ are the average values of the signal r_k corresponding to branch i and the kth signal sample within the bit.

Most of the results presented in this paper have been obtained using 2 samples/bit $(K = 2)$ in order to avoid the problem of sample clock phase optimization, which arises when 1 sample/bit is used. At 43 Gb/s, the use of 2 samples/bit in an actual Rx would be unrealistic, and we did so only as a means to speed up the simulations. In Section VI, we present a subset of simulations run at one sample/bit, with clock optimization, showing that the results with 1 or 2 samples/bit are very similar and lead to identical conclusions.

Fig. 3 shows the worst case performance of the standard MLSE receiver as a function of the DGD, normalized to the bit time, in terms of the OSNR needed to achieve the target BER of 10^{-3} for both IMDD and PSBT modulation formats. The value of the nonlinear exponent α is equal to 0.35.

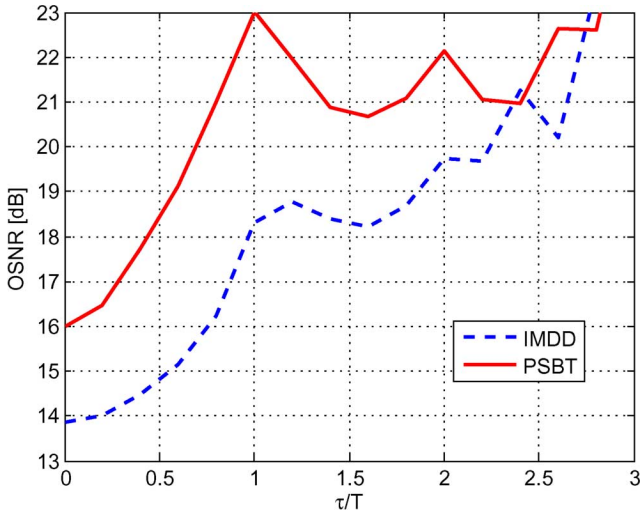

Fig. 3. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus DGD normalized to the bit time using the standard single-input MLSE Rx at 43 Gb/s.

The performance is independent of the SOP of the signal at the input of the Rx, and the worst case corresponds to the axis of DGD orthogonal to the input SOP, i.e., to a power splitting ratio between the two PSPs equal to 0.5. Our results, which indicate a large OSNR penalty (5–7 dB) for DGD values exceeding a bit time, are in accordance with those presented in [9]–[11]. The further increase in the penalty for values of DGD larger than two bit times is due to the fact that the memory of the channel starts exceeding the memory of the trellis.

The results reported in this section demonstrate that the use of a standard MLSE receiver is not effective in mitigating the effects of PMD, since the OSNR penalty with respect to back-to-back may be very high (5–7 dB).

III. MLSE Polarization Diversity Receivers

In order to improve the performance, we tried and used a receiver which is able to extract the information about the SOP of the signal at its input. Specifically, we chose a Stokes receiver [13].

A Stokes Rx can be implemented in two slightly different ways, shown in Figs. 4 and 5. The first configuration (Rx A) is composed of three polarizers corresponding to any three orthogonal axes on the Poincaré sphere, such as $\hat{s}_1$, $\hat{s}_2$, and $\hat{s}_3$. Each polarizer is followed by a PIN diode and a PD electrical filter. A fourth branch is used to detect the instantaneous optical power. This receiver requires the use of four analog-to-digital converters (ADCs). The samples of the electrical PD signals are then fed to a four-input MLSE processor. The components on the Poincaré sphere of the Stokes vector associated to the received signal samples are given by

$$\begin{aligned} r_1 &= 2\widetilde{r}_1 - \widetilde{r}_0 \\ r_2 &= 2\widetilde{r}_2 - \widetilde{r}_0 \\ r_3 &= 2\widetilde{r}_3 - \widetilde{r}_0. \end{aligned} \tag{3}$$

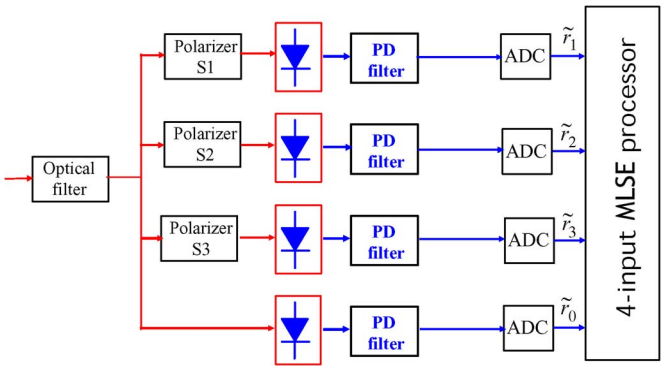

Fig. 4. Full Stokes polarization diversity receiver—type A.

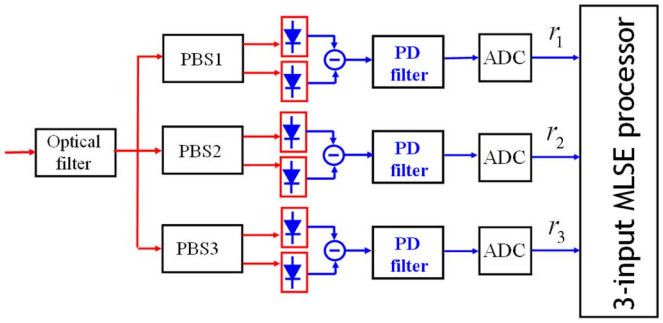

Fig. 5. Full Stokes polarization diversity receiver—type B.

Note that the same quantities r_1, r_2, and r_3 can be directly obtained by using the second receiver configuration (Rx B), shown in Fig. 5. It is composed of three mutually orthogonal (in Stokes space) polarizing beam splitters (PBSs), each followed by a pair of balanced photodetectors and a PD filter. The same postprocessing can thus be used for the two Rx configurations, including the MLSE branch metrics.

In order to reduce the complexity, we also analyzed the possibility of using a simpler configuration of the receiver (shown in Fig. 6), where only three branches and ADCs are used. This kind of Rx is not equivalent to the full Stokes receivers previously described, since it can be shown that it is not possible to derive the components (r_1, r_2, r_3) of the Stokes vector of the received signal starting from the knowledge of the three signals $\widetilde{r}_1, \widetilde{r}_2$, and $\widetilde{r}_3$ at the output of the simplified Stokes Rx. We show in the Appendix that there is a fundamental ambiguity that can be solved only by adding the fourth branch, present in the full Rx of Fig. 4, which provides the information about the total intensity of the received field.

IV. Branch Metrics

In this section, we propose several branch metrics that can be used by the Stokes receiver described in Section III. All metrics, except for the last one, are designed to work with the *full* Rx configuration of Figs. 4 or 5, while the last metric is designed for the simplified Rx of Fig. 6.

A. Exact Branch Metric

Under the assumptions that the optical signal is affected by ASE noise, modeled as white Gaussian noise with power spectral density $N_0/2$ on each polarization, and that the PD lowpass

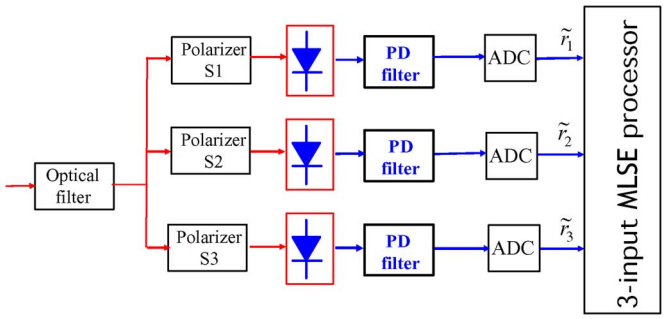

Fig. 6. Simplified Stokes polarization diversity receiver.

filters are wide enough not to alter the photodetected signal, the statistical properties of the Rx signal in Stokes space can be evaluated analytically [17]. We define: $\vec{R} = (r_1, r_2, r_3)$, the noisy received Stokes vector; $\vec{S} = (s_1, s_2, s_3)$, the Stokes vector that would be received in the absence of noise; $R = |\vec{R}|$, $S = |\vec{S}|$; θ the angle between $\vec{R}$ and $\vec{S}$

$$\theta = \arccos\left(\frac{\vec{S}\cdot\vec{R}}{|\vec{S}||\vec{R}|}\right). \tag{4}$$

Then, the exact joint probability density function (pdf) of the received Stokes vector $\vec{R} = (r_1, r_2, r_3)$, expressed in the polar coordinates R and θ, is [17]

$$f_{\vec{R}}(R,\theta) = \frac{1}{16\pi\sigma^4 R} e^{-\frac{S+R}{2\sigma^2}} I_0\left(\frac{\sqrt{RS}}{\sigma^2}\cos\frac{\theta}{2}\right) \tag{5}$$

where σ^2 is the variance of noise after the Rx optical filter ($\sigma^2 = N_0 B_o$, with B_o the equivalent noise bandwidth of the optical filter) and I_0 is the modified Bessel function of the first kind and order zero.

The optimum branch metrics for the Viterbi algorithm are given by [19]

$$m_i = -\sum_{k=1}^{K} \log\left[f_{\vec{R}}(R_k, \theta_{i,k})\right] \tag{6}$$

where K is the number of samples per bit used by the MLSE processor, and $(R_k, \theta_{i,k})$ are the values of R and θ for the kth signal sample of the received bit and the ith trellis branch. Substituting (5) into (6), and deleting the term $1/16\pi\sigma^4 R$ which is irrelevant because it is common to all branches, we get

$$m_i = \sum_{k=1}^{K} \log\left[e^{-\frac{S_{i,k}+R_k}{2\sigma^2}} I_0\left(\frac{\sqrt{R_k S_{i,k}}}{\sigma^2}\cos\frac{\theta_{i,k}}{2}\right)\right] \tag{7}$$

where $S_{i,k}$ is the modulus of the Stokes vector of the kth sample of the noiseless received signal corresponding to the ith trellis branch.

Since

$$\begin{aligned} R_k S_{i,k}\cos^2\frac{\theta_{i,k}}{2} &= \frac{1}{2} R_k S_{i,k}(1+\cos\theta_{i,k}) \\ &= \frac{1}{2}(R_k S_{i,k} + \vec{S}_{i,k}\cdot\vec{R}_k) \end{aligned} \tag{8}$$

it can be shown that metric (7) can be equivalently written as

$$m_i = \sum_{k=1}^{K} \left\{ \log \left[I_0 \left(\frac{1}{\sqrt{2}\sigma^2} \sqrt{R_k S_{i,k} + \vec{S}_{i,k} \cdot \vec{R}_k} \right) \right] - \frac{S_{i,k}}{2\sigma^2} \right\}. \tag{9}$$

Note that the use of the polarization diversity Rx with the exact metric (9) would require the knowledge of the received samples in the absence of noise, i.e., the vector $\vec{S}_{i,k}$. This makes the channel estimation procedure more complex and the dynamic update difficult. For this reason, in the following, we propose a set of simplified branch metrics, whose goal is to minimize the number of parameters to be estimated and updated during the channel estimation procedure. The performance of the polarization diversity Rx with the exact metric will be used in Section VI as a benchmark for the other receivers.

B. Approximated Metric

Using the following approximation of the modified Bessel functions [18]:

$$I_m(x) \simeq \frac{1}{\sqrt{2\pi x}} \mathrm{e}^x \quad x \gg m \tag{10}$$

and neglecting the logarithmic terms, metric (9) can be rewritten as

$$m_i \simeq \sum_{k=1}^{K} \left\{ \frac{1}{\sqrt{2}\sigma^2} \sqrt{R_k S_{i,k} + \vec{S}_{i,k} \cdot \vec{R}_k} - \frac{S_{i,k}}{2\sigma^2} \right\}. \tag{11}$$

Deleting the factors common to all branches

$$m_i \simeq -\sum_{k=1}^{K} \{ S_{i,k} - \sqrt{2} \cdot \sqrt{R_k S_{i,k} + \vec{S}_{i,k} \cdot \vec{R}_k} \}. \tag{12}$$

Finally, replacing the noiseless samples $S_{i,k}$, which are difficult to estimate, with the mean values of the noisy signal samples $\vec{\rho}_{i,k} = E\{\vec{R}_{i,k}\} = (\mu_{i,k}^{S1}, \mu_{i,k}^{S2}, \mu_{i,k}^{S3})$, $\rho_{i,k} = |\rho_{i,k}|$, we obtain

$$m_i \simeq -\sum_{k=1}^{K} \{ \rho_{i,k} - \sqrt{2} \cdot \sqrt{R_k \rho_{i,k} + \vec{\rho}_{i,k} \cdot \vec{R}_k} \}. \tag{13}$$

We will show in Section VI that the performance of an MLSE Rx using metric (13) is only slightly worse than the one using the exact metric (7). The advantage is that the channel estimation procedure is now much simpler, since only three parameters need to be estimated and updated, i.e., the average values of the noisy signal samples $\mu_{i,k}^{S1}$, $\mu_{i,k}^{S2}$, and $\mu_{i,k}^{S3}$.

C. Gaussian Metric

Another approximated branch metric can be obtained by assuming that the received electrical signal samples have a Gaussian statistical distribution. Assuming that the signal samples $r_{1,k}$, $r_{2,k}$, and $r_{3,k}$ are independent, the metrics can be evaluated as [25]

$$m_i = \sum_{k=1}^{K} \frac{\left(r_{1,k} - \mu_{i,k}^{S1} \right)^2}{\left(\sigma_{i,k}^{S1} \right)^2} + \log \left\{ \left(\sigma_{i,k}^{S1} \right)^2 \right\} + \frac{\left(r_{2,k} - \mu_{i,k}^{S2} \right)^2}{\left(\sigma_{i,k}^{S2} \right)^2} + \log \left\{ \left(\sigma_{i,k}^{S2} \right)^2 \right\} + \frac{\left(r_{3,k} - \mu_{i,k}^{S3} \right)^2}{\left(\sigma_{i,k}^{S3} \right)^2} + \log \left\{ \left(\sigma_{i,k}^{S3} \right)^2 \right\} \tag{14}$$

where $r_{1,k}$, $r_{2,k}$, and $r_{3,k}$ are obtained from $\widetilde{r}_{1,k}$, $\widetilde{r}_{2,k}$, and $\widetilde{r}_{3,k}$ using (3); $\mu_{i,k}^{S1}$, $\mu_{i,k}^{S2}$, and $\mu_{i,k}^{S3}$ are the average values of the signals $r_{1,k}$, $r_{2,k}$, and $r_{3,k}$, corresponding to branch i; and the kth signal sample within the bit and $(\sigma_{i,k}^{S1})^2$, $(\sigma_{i,k}^{S2})^2$, and $(\sigma_{i,k}^{S3})^2$ are the corresponding variances.

D. Variance-Stationary Gaussian Metric

The number of channel parameters to be estimated using (14) is six: three mean values and three variances. This number can be reduced to three by assuming that the variance is the same for all samples in all branches [variance-stationarity (VS) hypothesis]. In this case, metric (14) simplifies to

$$m_i = \sum_{k=1}^{K} \left[r_{1,k} - \mu_{i,k}^{S1} \right]^2 + \left[r_{2,k} - \mu_{i,k}^{S2} \right]^2 + \left[r_{3,k} - \mu_{i,k}^{S3} \right]^2. \tag{15}$$

E. Nonlinear Exponent Metric With Full Stokes Rx

In general, the variance of the samples has a nonnegligible dependence on the considered trellis branch. This means that the VS hypothesis is not very accurate and can lead to a substantial performance degradation. In [22], it was shown that the application of a nonlinear transformation, such as the square root (SQRT) operator, to the received samples can increase the stationarity of the variance. The SQRT operator, as well as other similar operators using an exponent α different from 0.5, can be directly applied to positive signals only, thus we chose to apply it to the signals $\widetilde{r}_{0,k}$, $\widetilde{r}_{1,k}$, $\widetilde{r}_{2,k}$, and $\widetilde{r}_{3,k}$, whose noise statistical distribution is similar to that of a conventional IMDD Rx. The branch metrics can be then evaluated as

$$m_i = \sum_{k=1}^{K} \left[(\widetilde{r}_{0,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S0} \right]^2 + \left[(\widetilde{r}_{1,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S1} \right]^2 + \left[(\widetilde{r}_{2,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S2} \right]^2 + \left[(\widetilde{r}_{3,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S3} \right]^2 \tag{16}$$

where α is a nonlinear distortion exponent (to be optimized), and $\widetilde{\mu}_{i,k}^{S0}$, $\widetilde{\mu}_{i,k}^{S1}$, $\widetilde{\mu}_{i,k}^{S2}$, and $\widetilde{\mu}_{i,k}^{S3}$ are the average values of the signals $\widetilde{r}_{0,k}$, $\widetilde{r}_{1,k}$, $\widetilde{r}_{2,k}$, and $\widetilde{r}_{3,k}$ corresponding to branch i and the kth signal sample within the bit. Note that each term in (16) is similar to the IMDD Rx metrics proposed in [20] and [21].

F. Nonlinear Exponent Metric With Simplified Stokes Rx

In the simplified Stokes Rx shown in Fig. 6, the noise statistical distribution on each of the three signals $\widetilde{r}_{1,k}$, $\widetilde{r}_{2,k}$, and $\widetilde{r}_{3,k}$ is similar to that of a conventional IMDD Rx. Therefore, a metric similar to (16) can be used. Assuming that the signal samples on the three branches are independent, the metrics can be evaluated as

$$m_i = \sum_{k=1}^{K} \left[(\widetilde{r}_{1,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S1} \right]^2 + \left[(\widetilde{r}_{2,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S2} \right]^2 + \left[(\widetilde{r}_{3,k})^{\alpha} - \widetilde{\mu}_{i,k}^{S3} \right]^2 \tag{17}$$

where α is the nonlinear distortion exponent (to be optimized), and $\widetilde{\mu}_{i,k}^{S1}$, $\widetilde{\mu}_{i,k}^{S2}$, and $\widetilde{\mu}_{i,k}^{S3}$ are the average values of the signal samples $\widetilde{r}_{1,k}$, $\widetilde{r}_{2,k}$, and $\widetilde{r}_{3,k}$ corresponding to branch i and the kth signal sample within the bit.

V. Channel Estimation and Metric Update

The parameters to be estimated when using the simplified parametric algorithms described in the previous sections are as follows.

1) Approximated metric (13). Parameters: $\mu_{i,k}^{S1}$, $\mu_{i,k}^{S2}$, and $\mu_{i,k}^{S3}$.
2) Gaussian metric (14). Parameters: $\mu_{i,k}^{S1}$, $\mu_{i,k}^{S2}$, $\mu_{i,k}^{S3}$, $(\sigma_{i,k}^{S1})^2$, $(\sigma_{i,k}^{S2})^2$, and $(\sigma_{i,k}^{S3})^2$.
3) VS Gaussian metric (15). Parameters: $\mu_{i,k}^{S1}$, $\mu_{i,k}^{S2}$, and $\mu_{i,k}^{S3}$.
4) Nonlinear exponent (NLE) metric (16). Parameters: $\widetilde{\mu}_{i,k}^{S0}$, $\widetilde{\mu}_{i,k}^{S1}$, $\widetilde{\mu}_{i,k}^{S2}$, and $\widetilde{\mu}_{i,k}^{S3}$.
5) Simplified NLE metric (17). Parameters: $\widetilde{\mu}_{i,k}^{S1}$, $\widetilde{\mu}_{i,k}^{S2}$, and $\widetilde{\mu}_{i,k}^{S3}$.

In all cases, the parameters to be estimated are the mean values $\mu_{n,k}$ and/or variances $(\sigma_{n,k})^2$ of the received signal samples

$$\begin{aligned}\mu_{n,k} &= E\{r_{i,k}|i=n\}\\ \sigma_{n,k}^2 &= E\left\{r_{i,k}^2|i=n\right\} - \mu_{n,k}^2\end{aligned} \tag{18}$$

where E is the "mean" operator and $r_{i,k}$ is the kth signal sample inside the current bit slot corresponding to the transmitted bit sequence $[b_{c-\delta}\ldots b_{c-1}, b_c]$, where b_c is the current bit and $\delta = 2^{N_s}$ is the trellis memory, with N_s being the number of trellis states. i is the number in decimal format corresponding the binary $(\delta+1)$-bit sequence $[b_{c-\delta}\ldots b_{c-1}, b_c]$, i.e., $i = b_c + b_{c-1}\cdot 2^1 + \ldots + b_{c-\delta}\cdot 2^\delta$. Note that in (18) the variable $r_{i,k}$ should be replaced by $\widetilde{r}$ for NLE metrics (16) and (17). However, for convenience, from now on we use the variable r with the meaning of a generic signal sample.

A. Channel Parameters Update

At the end of the training procedure, the channel parameters are continuously updated using a *decision-directed* algorithm, based on the following updating laws:

$$\begin{cases}\mu_{n,k}[j+1] = \mu_{n,k}[j] + \beta\cdot(r_{i,k} - \mu_{n,k}[j])\\ \mu_{n,k}^2[j+1] = \mu_{n,k}^2[j] + \beta\cdot\left(r_{i,k}^2 - \mu_{n,k}^2[j]\right), \quad \text{if } n=i.\\ \sigma_{n,k}^2[j+1] = \mu_{n,k}^2[j+1] - (\mu_{n,k}[j+1])^2\end{cases} \tag{19}$$

At each step, only the metric corresponding to branch i is updated, using the information carried by the signal samples $r_{i,k}$ inside the current bit slot. The value of i is deduced using the decoded bit sequence $[\hat{b}_{c-W-\delta}\ldots\hat{b}_{c-W-1}, \hat{b}_{c-W}]$, where b_c is the current bit and W is the size of the decision window of the trellis. For instance, at step j, a decision is taken on the bit received at step $(j-W)$. $\beta > 0$ is a convergence factor controlling the amount of adjustment.

VI. Performance Analysis

In a back-to-back configuration, using a standard direct-detection MLSE Rx, the OSNR values needed to achieve a BER value equal to 10^{-3} are 16 dB for PSBT and 13.9 dB for IMDD, regardless of the SOP of the signal at the input of the receiver.

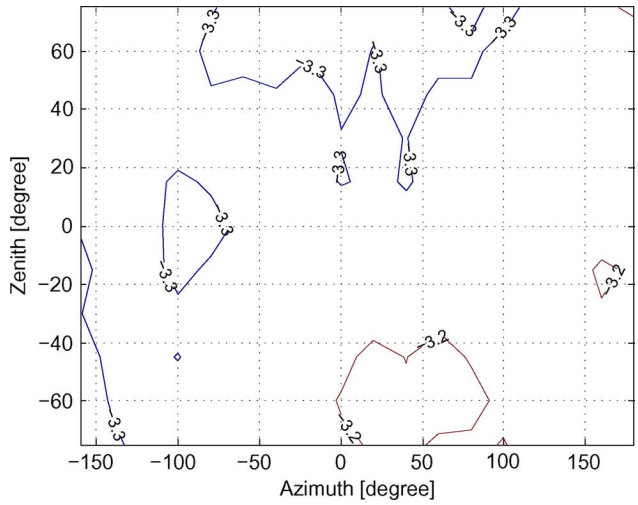

Fig. 7. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the received SOP over the Poincaré sphere. PSBT of 43 Gb/s with full Stokes receiver with Gaussian metric in back-to-back (OSNR = 17 dB over 0.1 nm).

While the performance of a standard direct-detection receiver is independent of the input SOP, the performance of polarization diversity receivers may depend on it. The goal of Section VI-A is to analyze this dependence for the MLSE receivers described in Sections III and IV.

Section VI-B is then devoted to the analysis of the dependence of the performance on the direction of the DGD axis. The goal of this analysis is to find the worst case scenarios for all analyzed MLSE receivers to be used in subsequent Sections VI-C and VI-D for the overall comparison of system performance.

A. Performance Dependence on the Signal SOP in Back-to-Back

Using the same PSBT setup described in Section II and eight-state MLSE processors with 2 samples/bit, we performed several simulation runs of a back-to-back optical system, varying the azimuth and zenith angles of the SOP of the transmitted signal. We evaluated the performance in terms of BER after MLSE processing using metrics (7) and (13)–(17). For all cases, the OSNR was fixed at 17 dB.

The performance of metrics exact (7), approximated (13), Gaussian (14), and VS Gaussian (15) turned out to be insensitive to the input SOP (see Fig. 7 as an example). The very small BER changes are due to Monte Carlo error counting. On the contrary, the performance of the NLE metric (16) of its simplified version (17) shows a dependence on it. The results are displayed in Figs. 8 and 9. Assuming that the three polarizers are directed along the axes $\hat{s}_1$, $\hat{s}_2$, and $\hat{s}_3$, the worst case SOP corresponds to an azimuth $\alpha = -135°$ and a zenith $\theta = -32.264°$, i.e., to coordinates in the Stokes space equal to $S_{wc} = (-(1/\sqrt{3}), -(1/\sqrt{3}), -(1/\sqrt{3}))$.

Similar results are obtained using the IMDD modulation format, i.e., the worst case SOP when using NLE metrics (16) and (17) is the same as for PSBT, while the performance is independent of the SOP when using the other metrics.

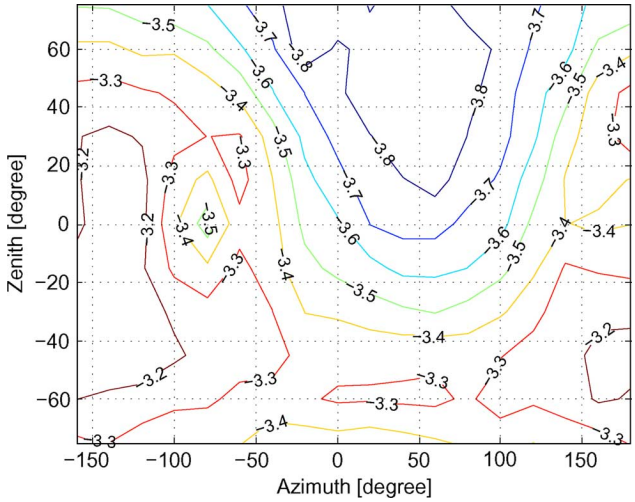

Fig. 8. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the received SOP over the Poincaré sphere. PSBT of 43 Gb/s with full Stokes receiver with NLE metric in back-to-back (OSNR = 17 dB over 0.1 nm).

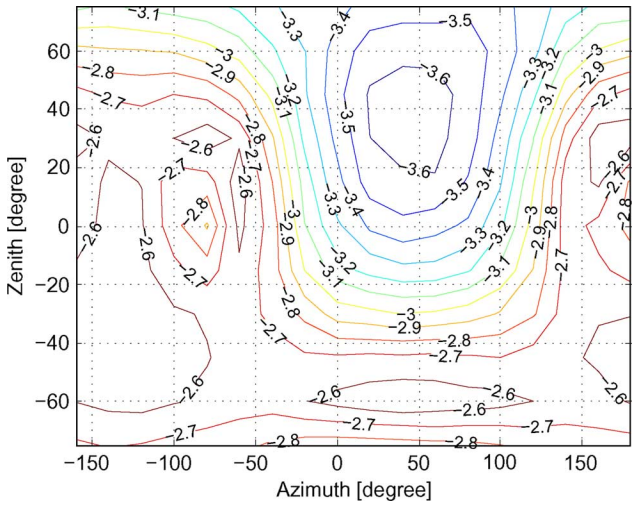

Fig. 9. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the received SOP over the Poincaré sphere. PSBT of 43 Gb/s with simplified Stokes receiver with NLE metric in back-to-back (OSNR = 17 dB over 0.1 nm).

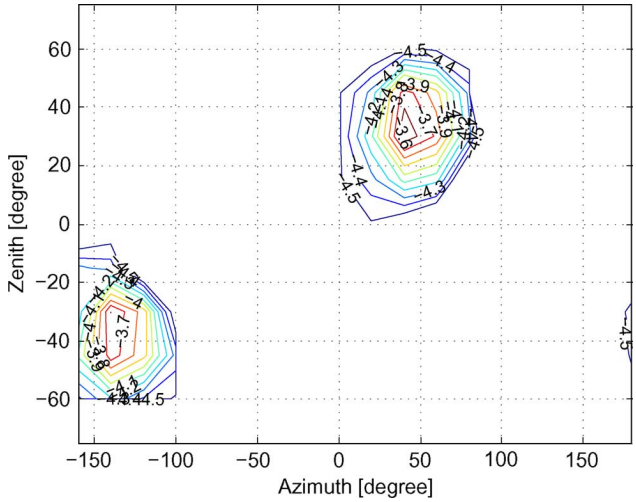

Fig. 10. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the direction of the DGD axis over the Poincaré sphere. PSBT transmission of 43 Gb/s with full Stokes receiver and approximated metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

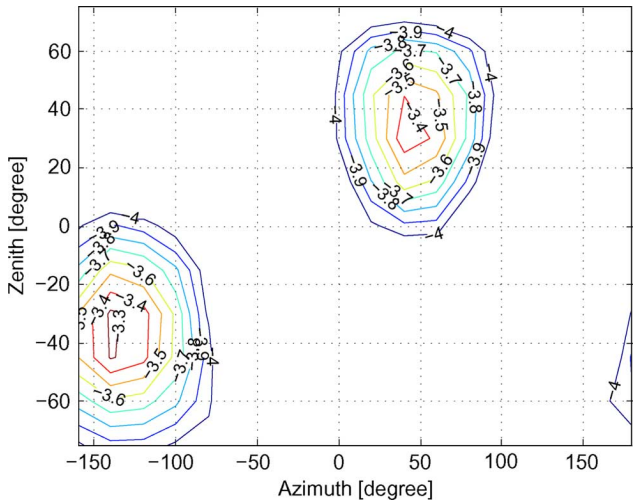

Fig. 11. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the direction of the DGD axis over the Poincaré sphere. PSBT transmission of 43 Gb/s with full Stokes receiver and Gaussian metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

B. Performance Dependence on the DGD Axis

After verifying that the performance of the MLSE processor with NLE metrics is dependent on the signal SOP, we analyzed its dependence on the direction of the DGD axis when first-order PMD is taken into account. We chose as SOP of the transmitted signal the one which gave the worst performance without PMD, i.e., $S_{wc} = (-(1/\sqrt{3}), -(1/\sqrt{3}), -(1/\sqrt{3}))$, after verifying by simulation that this configuration is the one giving the worst performance in the presence of PMD as well. We evaluated the performance in terms of BER after MLSE processing using metrics (13)–(17), varying the azimuth and zenith angles of the axis of DGD.

1) PSBT With Approximated and Gaussian Metrics: The results obtained using the approximated metric (13) are shown in Fig. 10, for a DGD amount equal to one bit time ($\tau = T$). The two points in the contour plot corresponding to the worst performance, i.e., the highest BER values, have coordinates $(\alpha, \theta) =$ 45° and 32.264° and −135° and −32.264°, respectively. In this cases, the DGD axis is parallel to the signal SOP, i.e., the SOP is coincident to one of the two DGD PSPs and the power splitting ratio between the two PSPs is equal to either 0 or 1.

The results obtained using the Gaussian and VS Gaussian metrics are shown in Figs. 11 and 12, respectively. Again, the DGD amount has been fixed to one bit time. The contour plots present a similar behavior to the one of the approximated metric (Fig. 10), i.e., the worst performance is achieved when the DGD axis is aligned with the signal SOP. In this case, the DGD has no effect on the propagating pulses and the performance is the same as in back-to-back.

Increasing the amount of DGD, the worst performance does not change until the memory of the channel exceeds the memory of the trellis. When this happens, the MLSE processor fails in compensating for the intersymbol interference generated by the effects of DGD and the worst case performance corresponds to

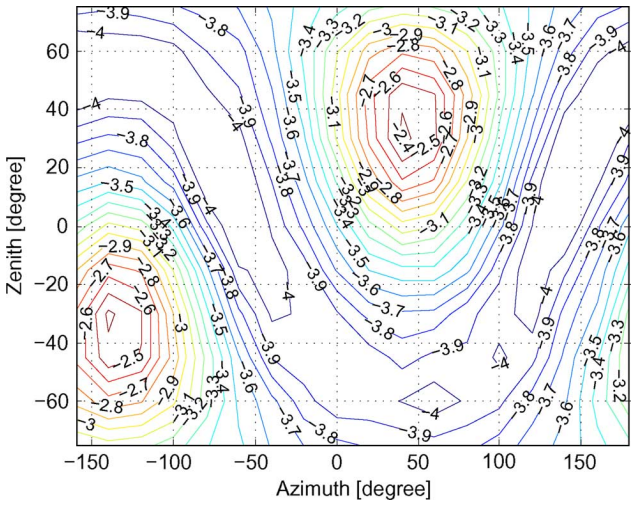

Fig. 12. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the direction of the DGD axis over the Poincaré sphere. PSBT transmission of 43 Gb/s with full Stokes receiver and VS Gaussian metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

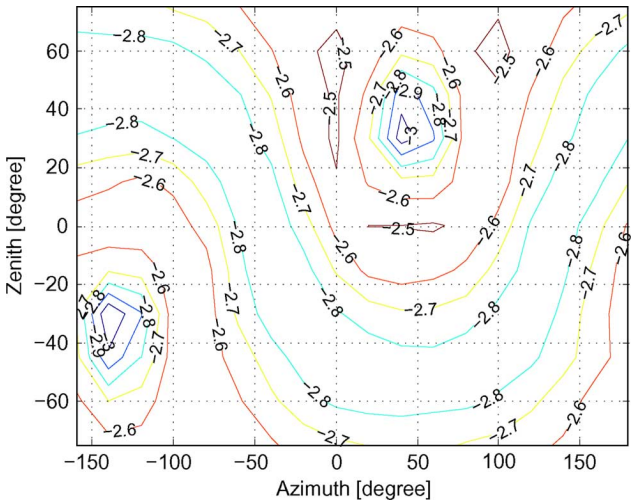

Fig. 13. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the direction of the DGD axis over the Poincaré sphere. PSBT transmission of 43 Gb/s with full Stokes receiver and NLE metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

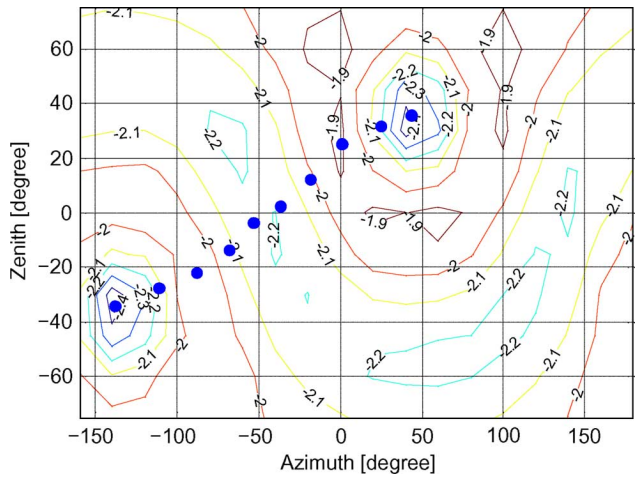

Fig. 14. Contour plot of $\log_{10}(\mathrm{BER})$ versus azimuth and zenith angles identifying the direction of the DGD axis over the Poincaré sphere. PSBT transmission of 43 Gb/s with simplified Stokes receiver with NLE metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

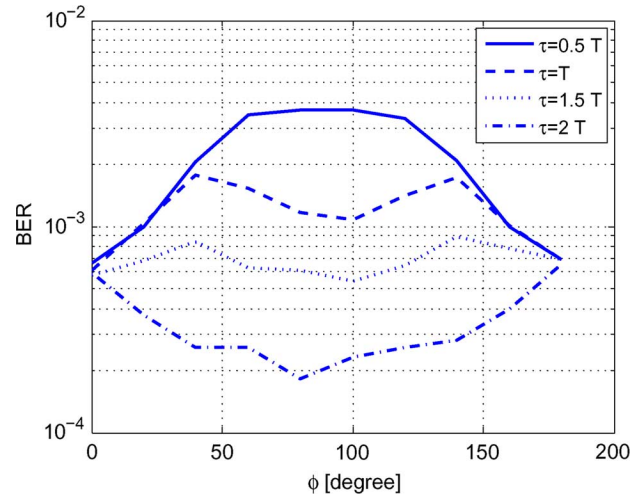

Fig. 15. BER versus angle ϕ between the DGD axis and the SOP for four different values of DGD. PSBT transmission of 43 Gb/s with full Stokes receiver and NLE metric (OSNR = 17 dB over 0.1 nm, DGD equal to one bit time).

a DGD axis orthogonal to the SOP, i.e., to a power splitting ratio between the two PSPs equal to 0.5. The contour plots obtained for $\tau > T$ are not shown here due to lack of space.

2) PSBT With NLE Metrics: The results obtained using the NLE metrics are shown in Figs. 13 and 14 for the full and simplified receivers, respectively. The DGD value is equal to one bit time ($\tau = T$). Differently from the other metrics, the worst performance is achieved in the region of points where the angle between the DGD axis and the SOP is $\phi = 45°$.

We have run more simulations increasing the value of DGD, obtaining different values of ϕ for the worst performance scenario. For instance, for DGD equal to $2T$ and $3T$, the worst case performance is achieved with $\phi = 0°$ and $\phi = 90°$, respectively.

This means that, for each value of DGD, a search for the worst case performance angle ϕ should be carried out. Since generating a complete contour map requires several hours of simulation, we reduced the number of simulation runs by choosing a limited number of DGD axis to be explored, corresponding to the points inside the contour plot of Fig. 14. Each point corresponds to a different value of ϕ, ranging from 0° to 180°. Examples of the performance curves as a function of ϕ obtained for four values of DGD are shown in Fig. 15.

3) IMDD: Differently from PSBT, when using the IMDD modulation format the worst case is always coincident with a DGD axis orthogonal to the input SOP (corresponding to a power splitting ratio between the two PSPs equal to $\gamma = 0.5$).

C. Worst Case Performance

Using the worst case scenarios obtained in Section VI-B, the performance curves shown in Fig. 16 have been obtained for PSBT, using the same Tx/Rx setup described in Section II and an eight-state MLSE processor with 2 samples/bit. This means that each point in the plots of Fig. 15 is a worst case results over all possible SOPs and DGD axis orientations. Similar results

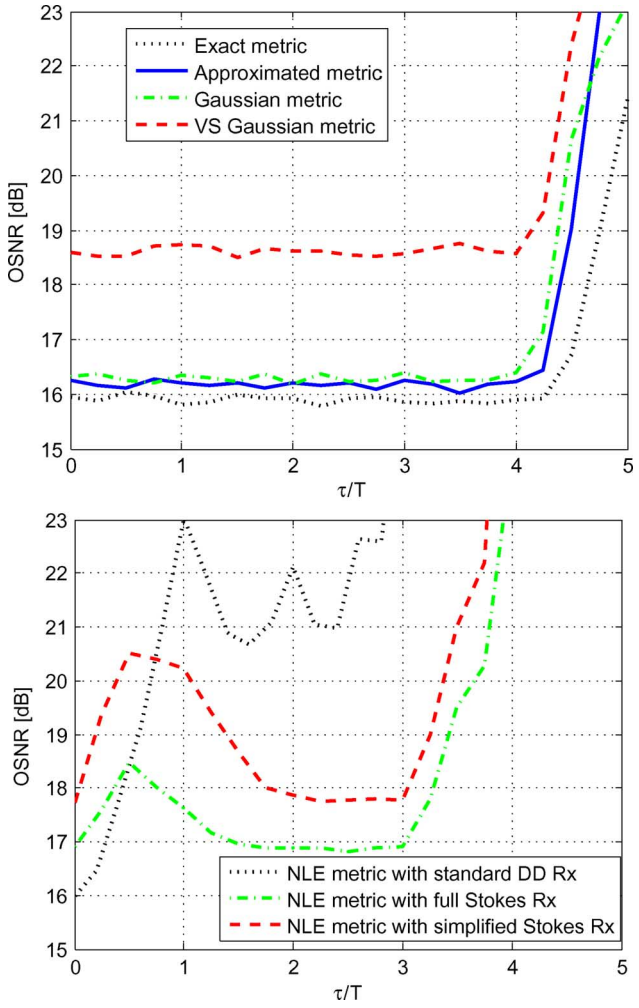

Fig. 16. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus DGD normalized to the bit time for PSBT. The seven curves are shown in two separate figures to increase the readability.

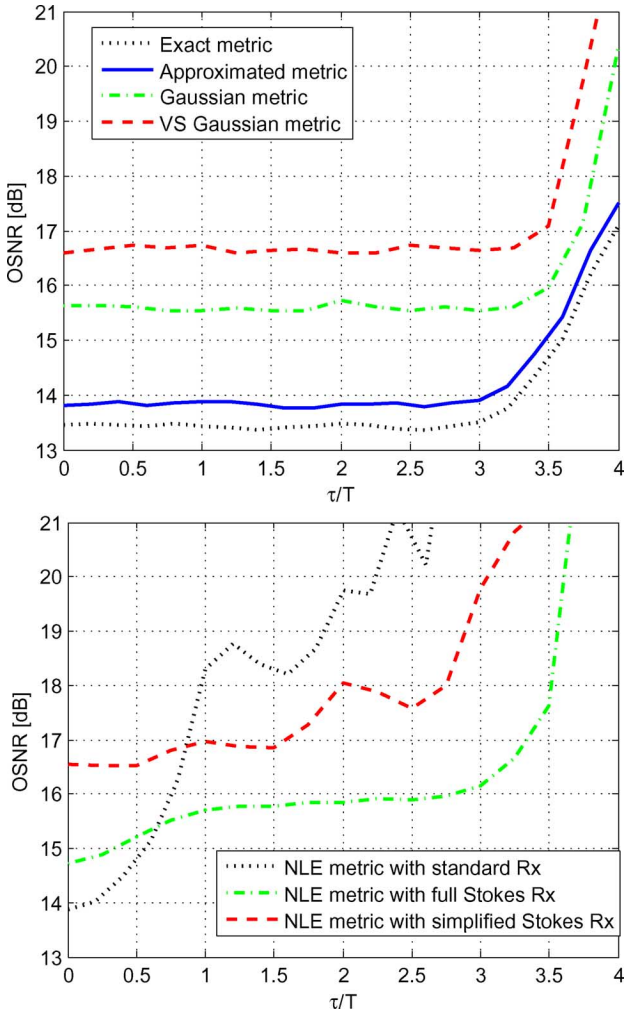

Fig. 17. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus DGD normalized to the bit time for IMDD. The seven curves are shown in two separate figures to increase the readability.

obtained using the IMDD modulation formats are reported in Fig. 17 for comparison.

The performance is shown in terms of OSNR needed to achieve a target BER of 10^{-3} as a function of the DGD value τ normalized to the bit time T. The OSNR is computed over a 0.1-nm bandwidth, taking into account noise on both polarizations. The performance of the standard single-input direct detection MLSE receiver is also shown for comparison.

When using the polarization diversity Rx with the exact metric, the system incurs no penalty at all up to a value of DGD equal to $3T$ for IMDD and $4T$ for PSBT. After that value, the penalty starts to increase due to the fact that the channel memory becomes larger than the Viterbi processor memory. A similar behavior is obtained when using the receiver with Gaussian or VS Gaussian metrics. For PSBT, the penalty in using a Gaussian metric with respect to the exact one is less than 0.5 dB, while it increases to 2.8 dB when using the VS Gaussian metric. For IMDD, the penalty is around 2 and 3 dB for the Gaussian and VS Gaussian metrics, respectively.

The performance of the same polarization diversity Rx (type A) with the NLE metric is worse than the one achieved using the approximated metric for both IMDD and PSBT, with a maximum penalty around 2 dB. The use of the three-input Rx would further decrease the complexity (only three parameters to be estimated and only three ADCs needed), but its performance is even poorer: its penalty with respect to the Gaussian metric may be as high as 4 dB for both PSBT and IMDD. Note that the nonlinear distortion exponent in NLE metrics (16) and (17) was optimized through simulation, finding optimum values around $\alpha = 0.3$–0.5 (not critical), depending on the amount of DGD.

D. Results With 1 Sample/Bit

All previous results have been obtained assuming that the MLSE processor uses 2 samples/bit, to avoid sampling clock phase optimization. On the other hand, in 43-Gb/s systems, the maximum sampling rates of the state-of-the-art ADCs provide 1 sample/bit only. Therefore, we also analyzed the performance of the MLSE receivers using 1 sample/bit, verifying that the penalty with respect to the 2 samples/bit case is limited (of the order of fractions of dB).

As an example, we show in Fig. 18 the results for the IMDD modulation formats when the full Stokes Rx with approximated metric is used. The penalty due to the use of 1 sample/bit only is around 0.2 dB. Similar results are obtained for PSBT. As a consequence, we can conclude that the general performance of

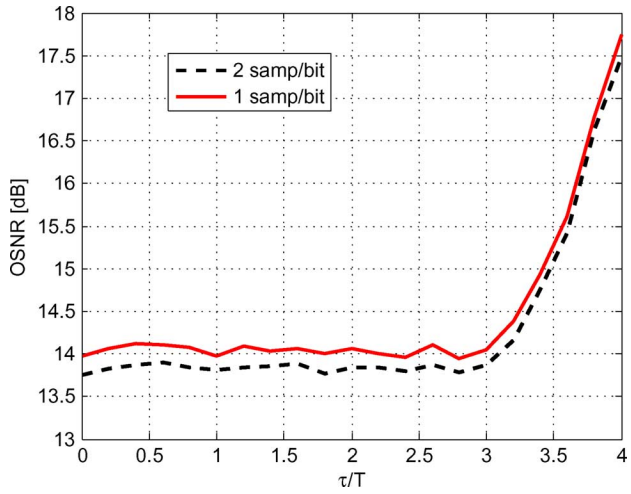

Fig. 18. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus DGD normalized to the bit-time for IMDD using the full Stokes Rx with approximated metric. Dashed line: 2 samples/bit; solid line: 1 sample/bit.

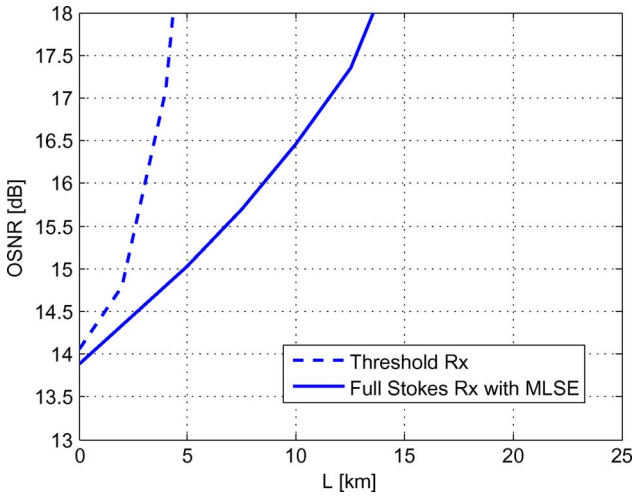

Fig. 19. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus fiber length L for IMDD. Solid line: full Stokes Rx with approximated metric (eight states). Dashed line: standard threshold Rx.

the Stokes Rx and all the metrics change very little (a small fraction of a decibel) if assessed using 1 or 2 samples/bit.

E. Tolerance to Chromatic Dispersion

Similarly to standard direct-detection receivers, the presence of the MLSE processor gives the further advantage of an increased tolerance to chromatic dispersion (CD). We show as an example the performance of the full Stokes Rx followed by an eight-state MLSE processor using the approximated branch metric (13), in terms of the OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus the fiber length L. The dispersion parameter is $D = 16.7$ ps/mk/km. The results are shown in Figs. 19 and 20 for IMDD and PSBT, respectively. The performance of a standard threshold Rx is also shown for comparison.

F. Higher Order PMD

As shown in the previous sections, the analysis of the various Rx configurations with multiple metrics was extremely complex. Compounding this first-order PMD study with a thorough higher order study would not have been possible.

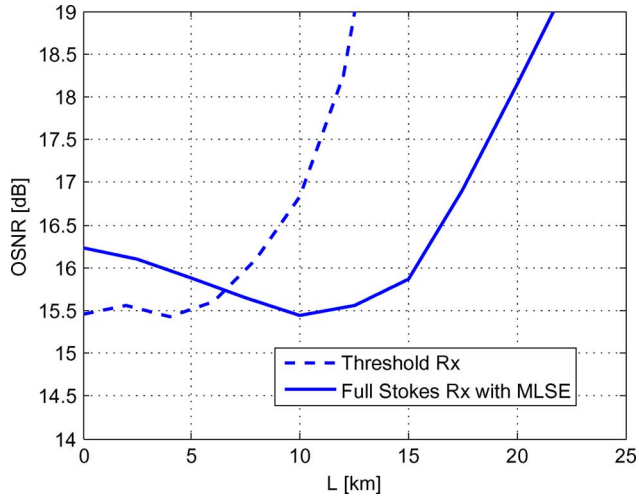

Fig. 20. OSNR (over 0.1 nm, in decibels) needed to achieve BER $= 10^{-3}$ versus fiber length L for PSBT. Solid line: full Stokes Rx with approximated metric (eight states). Dashed line: standard threshold Rx.

On the other hand, we were willing to probe the Rx's performance with higher order PMD to make sure there were no intrinsic weaknesses. As a result, we did a series of tests using an all-order wave-plate model [26], tuned to yield an average DGD of 8 ps. We ran 600 simulations, each of which generating a different realization of the wave-plate model. The first-order DGDs created by the all-order model actually ranged between 1 and 20 ps.

We considered the type A/B Rxs. The result was a maximum degradation of the needed OSNR for BER $= 10^{-3}$ equal to 0.2 dB. Even though the test is clearly not exhaustive, it provides an indication that higher order PMD should be tolerated without substantial excess penalty.

VII. Conclusion

We analyzed the performance of polarization-diversity receivers using MLSE to combat the effects of PMD in 43-Gb/s optical systems employing IMDD and duobinary modulation formats. We proposed and studied several parametric branch metrics, and compared their performance to the one achieved using the optimum metric based on the exact statistics of the Stokes vector components of the received signal samples.

Out of all the metrics analyzed in this paper, the best compromise between complexity and performance is the approximated metric (13), whose complexity is much lower than the exact metric (7) and whose OSNR penalty with respect to (7) is limited (lower than 0.5 dB for both IMDD and PSBT). As pointed out throughout this paper, the approximated metric is also characterized by a simple channel identification procedure, requiring the estimation of only three parameters to fully characterize each trellis branch sample. In addition, such parameters are very easily estimated as they consist of the average of the observed Stokes vector components samples.

As for the other analyzed metrics, the VS Gaussian, NLE, and simplified NLE have the advantage of a simple channel estimation and metric update procedure, as well. On the other hand,

their performance is worse with respect to the exact and approximated metrics, especially in conjunction with the simplified Stokes Receiver.

Finally, even though the Gaussian metric has a very good performance, at least in the PSBT system, it is characterized by a more complex channel estimation procedure compared to the approximated metric, since the number of parameters to be estimated is six, three of them being the variances of the observed Stokes vector components samples.

In conclusion, we have shown that the use of a full Stokes Rx (either type A or B), followed by a multiple-input MLSE processor using an appropriate branch metric, can virtually eliminate any PMD penalty for IMDD and PSBT. This is in contrast to standard direct-detection Rx's with single-input MLSE processors whose penalty, despite MLSE, peaks at 5–6 dB above back-to-back every time the DGD nears a multiple of the bit time. The complexity of the multiple-input MLSE processor for the Stokes Rx is reasonable, since the required number of states can be low (4–8 for a max DGD equivalent to 2–3 T). Also, thanks to proper branch metric approximations, the effort needed for branch metric computation and channel estimation is reasonable as well, whereas the penalty incurred because of metric suboptimality is almost negligible. As a positive side effect, the presence of MLSE also improves the chromatic dispersion tolerance of the system, quite substantially.

From the viewpoint of actual exploitation, this solution could be of interest for 43-Gb/s IMDD/PSBT systems. It might compete with DQPSK on one hand, and with coherent detection on the other. Among its main advantages, there is the fact that the Tx remains simple and does not need any change with respect to conventional IMDD/PSBT. In fact, the proposed MLSE Stokes Rx is fully backward compatible with any existing conventional IMDD/PSBT Tx. Among its drawbacks is the need for four ADCs, sampling at least at the bit rate.

In contrast, DQPSK needs a more complex Tx, and the Rx complexity is of the order of the Stokes Rx. DQPSK has a better "native" PMD insensitivity than standard IMDD/PSBT due to the twice longer symbol time, but to achieve very high PMD insensitivity, it may need multiple-input MLSE postprocessing as well, including some form of polarization diversity which would further increase the Rx complexity [24].

As for coherent detection, systems based on it are quite more complex and they can probably be justified only in the context of 100-Gb/s transmission, where there is no alternative to using complex constellations to substantially reduce the symbol rate, or to exploit their capability to electronically compensate for arbitrarily large amounts of chromatic dispersion.

APPENDIX

Using (3) and considering that the intensity parameter r_0 of the Stokes components of a fully polarized optical field is related to the three components (r_1, r_2, r_3) of the Stokes vector by the relationship: $r_0 = \sqrt{r_1^2 + r_2^2 + r_3^2}$, we obtain

$$\begin{cases} r_1 = 2\widetilde{r}_1 - \widetilde{r}_0 \\ r_2 = 2\widetilde{r}_2 - \widetilde{r}_0 \\ r_3 = 2\widetilde{r}_3 - \widetilde{r}_0 \\ r_0^2 = r_1^2 + r_2^2 + r_3^2 \end{cases} \tag{20}$$

where r_0 is the intensity of the overall optical field and is coincident to the signal $\widetilde{r}_0$ at the output of the lowest branch in the full Stokes Rx of Fig. 4.

Substituting the first three equations in the last one, we obtain

$$\widetilde{r}_0^2 - 2(\widetilde{r}_1 + \widetilde{r}_2 + \widetilde{r}_3)\widetilde{r}_0 + 2\left(\widetilde{r}_1^2 + \widetilde{r}_2^2 + \widetilde{r}_3^2\right) = 0. \tag{21}$$

The solution of (21) is

$$\widetilde{r}_0 = (\widetilde{r}_1 + \widetilde{r}_2 + \widetilde{r}_3) \pm \Delta \tag{22}$$

with $\Delta = \sqrt{(\widetilde{r}_1 + \widetilde{r}_2 + \widetilde{r}_3)^2 - 2((\widetilde{r}_1^2 + \widetilde{r}_2^2 + \widetilde{r}_3^2))} \geq 0$. The Stokes parameters r_1, r_2, and r_3 can thus be derived from $\widetilde{r}_1$, $\widetilde{r}_2$, and $\widetilde{r}_3$ as

$$\begin{cases} r_1 = \widetilde{r}_1 - \widetilde{r}_2 - \widetilde{r}_3 \pm \Delta \\ r_2 = \widetilde{r}_2 - \widetilde{r}_1 - \widetilde{r}_2 \pm \Delta \\ r_3 = \widetilde{r}_3 - \widetilde{r}_1 - \widetilde{r}_3 \pm \Delta \end{cases}. \tag{23}$$

The ambiguity due to the $\pm$ is fundamental and cannot be resolved without adding more information (for a formal proof, see [23]). A way of obtaining the Stokes parameters is to add a fourth branch in the Rx, which provides the information about the total intensity of the received field.

REFERENCES

[1] H. Bulow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 158–167, Jan. 2008.

[2] O. E. Agazzi, M. R. Hueda, H. S. Carrer, and D. E. Crivelli, "Maximum-likelihood sequence estimation in dispersive optical channels," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 749–763, Feb. 2005.

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

[4] N. Alic, G. C. Papen, R. E. Saperstein, R. Jiang, C. Marki, Y. Fainman, S. Radic, and P. A. Andrekson, "Experimental demonstration of 10 Gb/s NRZ extended dispersion-limited reach over 600 km-SMF link without optical dispersion compensation," presented at the Opt. Fiber Commun., Anaheim, CA, Mar. 5–10, 2006, paper OWB7.

[5] P. Poggiolini, G. Bosco, J. Prat, R. Killey, and S. Savory, "Branch metrics for effective long-haul MLSE," presented at the Eur. Conf. Opt. Commun., Cannes, France, Sep. 24–29, 2006, paper We2.5.

[6] S. Chandrasekhar, "Performance of MLSE receiver in a dispersion-managed experiment at 10.7 Gb/s under non-linear transmission," *IEEE Photon. Technol. Lett.*, vol. 18, no. 23, pp. 2448–2450, Dec. 1, 2006.

[7] M. R. Hueda, D. E. Crivelli, and H. S. Carrer, "Performance of MLSE-based receivers in lightwave systems with nonlinear dispersion and amplified spontaneous emission noise," in *Proc. IEEE GLOBECOM*, Nov. 29–Dec. 3 2004, vol. 1, pp. 299–303.

[8] P. Poggiolini, G. Bosco, M. Visintin, S. J. Savory, Y. Benlachtar, P. Bayvel, and R. I. Killey, "MLSE-EDC versus optical dispersion compensation in a single-channel SPM-limited 800 km link at 10 Gb/s," presented at the Eur. Conf. Opt. Commun., Berlin, Germany, Sep. 2007.

[9] F. Buchali, G. Thielecke, and H. Bulow, "Viterbi equalizer for mitigation of distortions from chromatic dispersion and PMD at 10 Gb/s," presented at the Opt. Fiber Commun., Los Angeles, CA, Feb. 23–27, 2004, paper MF85.

[10] J. M. Gené, P. J. Winzer, S. Chandrasekhar, and H. Kogelnik, "Simultaneous compensation of polarization mode dispersion and chromatic dispersion using electronic signal processing," *J. Lightw. Technol.*, vol. 25, no. 7, pp. 1735–1741, Jul. 2007.

[11] T. Kupfer, J. Whiteaway, and S. Langenbach, "PMD compensation using electronic equalization particular maximum likelihood sequence estimation," presented at the Opt. Fiber Commun., Anaheim, CA, Mar. 2007, paper OMH1.

[12] A. O. Lima, I. T. Lima, T. Adali, Jr., and C. R. Menyuk, "A novel polarization diversity receiver for PMD mitigation," *IEEE Photon. Technol. Lett.*, vol. 14, no. 4, pp. 465–467, Apr. 2002.
[13] S. Benedetto, R. Gaudino, and P. Poggiolini, "Direct detection of optical digital transmission based on polarization shift keying modulation," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 3, pp. 531–542, Apr. 1995.
[14] D. Penninckx, M. Chbat, L. Pierre, and J.-P. Thiery, "The Phase-shaped Binary Transmission (PSBT): A new technique to transmit far beyond the chromatic dispersion limit," in *Proc. Eur. Conf. Opt. Commun.*, Oslo, 1996, vol. 2, pp. 173–176.
[15] T. Kupfer, C. Dorschky, M. Ene, and S. Langenbach, "Measurement of the performance of 16-states MLSE digital equalizer with different optical modulation formats," presented at the Opt. Fiber Commun., San Diego, CA, Feb. 2008, paper PDP13.
[16] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode .bers," *J. Lightw. Technol.*, vol. 9, no. 11, pp. 1439–1456, Nov. 1991.
[17] S. Benedetto and P. Poggiolini, "Theory of polarization shift keying modulation," *IEEE Trans. Commun.*, vol. 40, no. 4, pp. 708–721, Apr. 1992.
[18] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge: Cambridge Univ. Press, 1992.
[19] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
[20] M. R. Hueda, D. E. Crivelli, H. S. Carrer, and O. E. Agazzi, "Parametric estimation of IM/DD optical channels using new closed-form approximations of the signal PDF," *J. Lightw. Technol.*, vol. 25, no. 3, pp. 957–975, Mar. 2007.
[21] M. Franceschini, G. Ferrari, R. Raheli, F. Meli, and A. Castoldi, "Post-detection nonlinear distortion for efficient MLSD in optical links," *Opt. Express*, vol. 15, no. 18, pp. 11750–11755, Sep. 3, 2007.
[22] G. Bosco, P. Poggiolini, and M. Visintin, "Performance analysis of MLSE receivers based on the square-root metric," *J. Lightw. Technol.*, vol. 26, no. 14, pp. 2098–2109, Jul. 2008.
[23] J. N. Damask, *Polarization Optics in Telecommunications*. New York: Springer-Verlag, 2004.
[24] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "Multilevel optical modulations with closed-form optimal metrics for MLSE receiver insensitive to GVD and PMD," presented at the Opt. Fiber Commun., San Diego, CA, Feb. 2008, paper JWA57.
[25] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991, p. 329.
[26] C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurement based on transmission spectra through a polarizer," *J. Lightw. Technol.*, vol. 12, no. 6, pp. 917–929, June 1994.

Gabriella Bosco (S'00–M'02) was born in Ivrea, Italy, in 1973. She received the Degree in telecommunication engineering (thesis on the nonlinear effect of the propagation in WDM optical systems) and the Ph.D. degree in electronic and communication engineering (thesis on the performance analysis of optical communication systems) from Politecnico di Torino, Torino, Italy, in 1998 and 2002, respectively.

In 2000, she was a Visiting Researcher at the Optical Communication and Photonic Network (OCPN) group, University of California at Santa Barbara, directed by Prof. Blumenthal, working on polarization-mode dispersion monitoring techniques. She currently holds a postdoctoral position in the Optical Communication Group at the Department of Electronics, Politecnico di Torino. Her current research interests are focused on advanced coherent modulation formats and electronic mitigation techniques in optical links.

Pierluigi Poggiolini (S'90–M'93) received the M.S. and Ph.D. degrees from Politecnico di Torino, Torino, Italy, in 1988 and 1993, respectively.

From 1990 to 1995, he was first a Visiting Scholar and then a Postdoctoral Fellow at the Optical Communications Research Laboratory, Stanford University, Stanford, CA, where he worked on the STARNET and CORD all-optical packet network projects. Since 1998 he has been an Associate Professor at Politecnico di Torino, where he is the Coordinator of the Optical Communications Group. He is involved in several EU-funded projects, including e-Photon/ONe, Nobel II, BONE, and Euro-FOS. He has published over 150 papers in international journals and conference proceedings. His current research interests include coherent modulation formats and electronic mitigation techniques for optical broadband transmission.

Dr. Poggiolini is an elected member of the Academic Senate of Politecnico di Torino and of the National University Council, an official body of the Italian Ministry of University and Research.

Monica Visintin (M'01) received the Laurea in ingegneria elettronica (*summa cum laude*) from Politecnico di Torino, Torino, Italy, in 1986, the M.S. degree in electrical engineering from the University of California at Los Angeles (UCLA), in 1988, and the Ph.D. degree in electrical engineering from Politecnico di Torino, in 1990.

Since 1990 she has been with the Dipartimento di Elettronica, Politecnico di Torino, where she is currently an Associate Professor. Her current interests are in the design and simulation of digital transmitters and receivers for space applications.

Li Liangchuan received the Ph.D. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2007.

He is currently a Researcher with the Department of Optical Networks Research, Huawei Advanced Technology, Shenzhen, China. His research works include 100-Gb/s optical transmission and electronic signal processing at high-speed optical communications systems.

Chen Ming, photograph and biography not available at the time of publication.