CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Archivio istituzionale della ricerca - Università di Brescia

Mitigation of Nonlinear and PMD Impairments by Bit-Synchronous Polarization Scrambling

S. Wabnitz, *Member, IEEE*, and K. S. Turitsyn

***Abstract*—Our statistical study by the importance sampling method shows that a significant performance improvement may be achieved by bit-synchronous polarization scrambling in the presence of polarization mode dispersion and fiber nonlinearity in dispersion managed, single-channel, or wavelength-division-multiplexed NRZ 10 Gbit/s fiber optic transmission systems.**

***Index Terms*—Optical fiber communication, optical fiber polarization, optical modulation, nonlinear optics, wavelength division multiplexing.**

I. Introduction

THE fiber Kerr nonlinearity, polarization mode dispersion (PMD), and their interaction represent major sources of impairment in wavelength-division-multiplexed (WDM) optical transmissions [1], [2]. Even in the single-channel case the interplay of PMD [or differential group delay (DGD)] and self-phase modulation (SPM) may lead to large performance degradations whenever a PMD compensator (PMDC) is used at the receiver [3]. In WDM systems, the nonlinear cross-polarization modulation (XPolM) among wavelength channels introduces a signal depolarization that varies within the bit period time scale [2], [4], [5]. As a result, both intra- and interchannel nonlinearities combined with PMD may lead to significant system penalties when a polarization sensitive receiver is used [6], [7]. On the other hand, it was also pointed out that nonlinear effects may even be beneficial in mitigating PMD impairments in dense WDM systems [8].

In this paper, we present a numerical statistical study of the use of different fast (i.e., bit-synchronous) polarization scrambling schemes to mitigate nonlinear impairments in the presence of PMD. As a matter of fact, the polarization scrambling technique is extensively used in lightwave systems [9]. Transmitter polarization scrambling to depolarize the signal within a forward error correction (FEC) frame was proposed to average out PMD-induced error bursts [10]–[12]. Polarization scrambling also reduces noise accumulation from polarization hole-burning (or polarization-dependent gain) in erbium-doped amplifiers (EDFAs) [13]. Indeed, Bergano *et al.* [14], [15] have successfully exploited a scrambling frequency equal to the transmitter bit rate. Such bit-synchronous polarization scrambling (BSPS) also permits to introduce a common phase modulation (PM) to the input data, which was shown to be important in order to reduce SPM impairments [14]–[16]. A different type of fast polarization scrambling (operating at half the data rate) involves the periodic alternation of the state of polarization of subsequent bits, or alternate-polarization ON–OFF keying (OOK) format. Such input polarization rotation reduces the nonlinear interaction among subsequent marks whenever a return-to-zero (RZ) modulation format is used [17], [18].

Whenever nonlinear effects and PMD are simultaneously present, experiments have revealed that BSPS could suffer from significant penalty (with respect to slow, e.g., at 10 kHz, polarization scrambling) owing to the PMD-induced conversion of polarization modulations into intensity modulations [9]. Moreover, it was shown that BSPS could also lead to significant PMD-induced signal repolarization [9].

In order to clarify the impact of PMD on the performance of the BSPS technique, in this study we numerically evaluate by the importance sampling (IS) method [19] the statistical distribution of the Q factor in the presence of PMD for a typical 10 Gbit/s single-channel long-distance dispersion managed transmission link using the nonreturn-to-zero (NRZ) OOK format. Our study permits us to estimate the relative improvement of the outage probability with different polarization scrambling schemes. We will also briefly discuss the role of interchannel impairments in WDM transmissions using BSPS.

II. Model

Let us consider the polarization interaction of N WDM channels. Our simulation model takes into account all orders of PMD, four-wave mixing, SPM, cross-phase modulation, first-order group-velocity dispersion (GVD), second-order GVD (or third-order dispersion), and fiber attenuation. In the coarse-step method, fibers with random birefringence are divided into a series of scattering sections of length, say, $\Delta z_{\mathrm{scatt}}$ with uniform birefringence; propagation within each section is described by the two coupled nonlinear Schrödinger equations [20], [21]

$$\left[\frac{\partial}{\partial z}+\frac{\alpha_{x,y}}{2}+\beta_{1x,y}\frac{\partial}{\partial t}-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}-\frac{\beta_3}{6}\frac{\partial^3}{\partial t^3}\right]E_{x,y}(t,z)$$
$$=-j\gamma\frac{8}{9}[|E_{x,y}(z,t)|^2+|E_{y,x}(z,t)|^2]E_{x,y}(t,z) \quad (1)$$

Manuscript received December 12, 2011; revised May 17, 2012; accepted May 18, 2012. Date of publication June 04, 2012; date of current version July 02, 2012. This work was supported in part by the Italian Ministry of Education, University and Research (MIUR) under Grant 2008MPSSNX and in part by the UniBS-MIT-MechE faculty exchange program cosponsored by the CARIPLO Foundation, Italy, under Grant 2008-2290.

S. Wabnitz is with the Department of Information Engineering, University of Brescia, Brescia 25123, Italy (e-mail: stefano.wabnitz@ing.unibs.it).

K. S. Turitsyn is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: turitsyn@gmail.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.2012.2201209

0733-8724/$31.00 © 2012 IEEE

where $\alpha_{x,y}$ are linear attenuations along the two fiber axes, $\beta_{x,y}$ are the respective group velocities, $\beta_{2,3}$ are the second- and third-order group velocity dispersions, and $\gamma = 2\pi n_2/(\lambda A_{\rm eff})$ is the fiber nonlinear coefficient, where n_2 is the nonlinear refractive index, λ is the wavelength, and $A_{\rm eff}$ is the fiber effective area. The lengths of the scattering sections $\Delta z_{\rm scatt}$ are chosen from a random Gaussian distribution with a given mean value $\mu = 1$ km and standard deviation $\sigma = 100$ m. The group velocities of the two polarization modes are related to the fiber PMD coefficient $D_{\rm PMD}$ as [21]

$$\beta_{1x} = -\beta_{1y} = \frac{D_{\rm PMD}}{2\sqrt{\Delta z_{\rm scatt}}}$$

and $L_{\rm corr} = \Delta z_{\rm scatt}/2$ is the fiber correlation length. Linear propagation in each scattering section under the action of fiber birefringence is defined in terms of a frequency-dependent Jones matrix, whose eigenvectors define the rotation axis in Stokes space. The angle θ between the rotation axis (or direction of the PMD vector) of a subsequent section with respect to the direction of PMD vector at the output of the previous section is chosen as $\cos\theta = 2x^{1/\alpha} - 1$ where x is a uniform random variable in [0,1]. Note that if $\alpha = 1$, then $\cos\theta$ has uniform distribution between $[-1, 1]$ (unbiased case). Whenever $\alpha > 1$, one obtains an angle distribution which is biased toward large DGD values [19].

As we shall demonstrate in the next section, the performance degradations resulting from the interplay of nonlinearity and PMD can be effectively mitigated by means of input BSPS. This involves high-frequency polarization modulation of input signals according to the expression

$$E_{x,y}(t, z = 0) = A(t)e^{-i(\omega t + \phi_{x,y}(t))}. \tag{2}$$

We will compare the performance of different types of BSPS [22]. Consider first the "pure" BSPS corresponding to a sawtooth phase variation in time

$$\phi_{x,y}(t) = \varphi_{x,y} + a_{x,y}\left(\frac{t}{nT} - \left\lfloor\frac{t}{nT}\right\rfloor\right). \tag{3}$$

By setting, for example, $\varphi_{x,y} = 0$, $a_x = 0$, $a_y = 2\pi n$, the input Stokes vector performs a complete rotation about the S_1 axis during each bit time T (see Fig. 1), namely

$$\begin{aligned} S_1(t, z = 0) &= 0 \\ S_2(t, z = 0) &= \cos(2\pi t/T) \\ S_3(t, z = 0) &= \sin(2\pi t/T). \end{aligned} \tag{4}$$

Note that for high channel rates (e.g., a bit rate $B = 1/T = 10$ Gbit/s), one should use in practice a relatively long phase modulation period (e.g., $n = 8$) in order to synthesize a sufficient number of harmonics to represent the sawtooth modulation, where B is the highest harmonic frequency.

Next we consider the sinusoidal polarization modulation scheme that was introduced in [14], [15]

$$\phi_{x,y}(t) = \varphi_{x,y} + a_{x,y}\cos(2\pi t/T + \Psi). \tag{5}$$

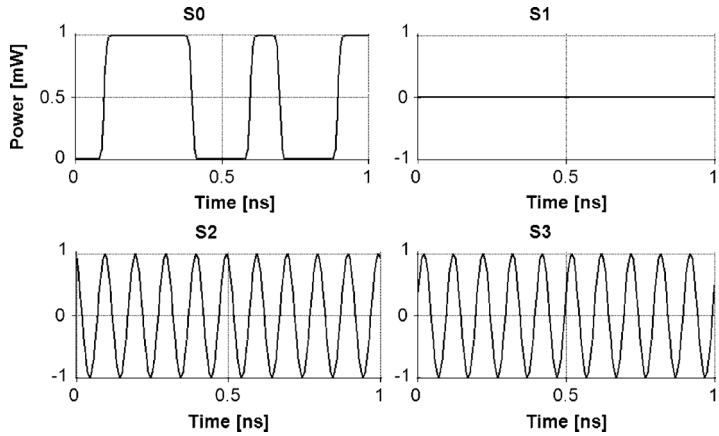

Fig. 1. Input temporal profile of Stokes parameters with sawtooth polarization modulation.

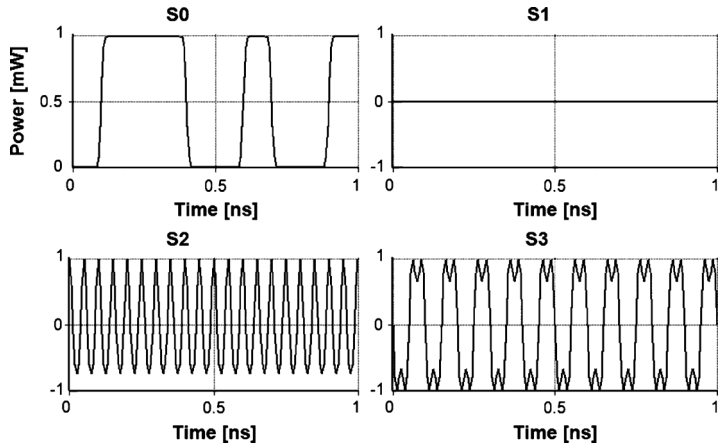

Fig. 2. Input temporal profile of Stokes parameters with sinusoidal polarization modulation.

Fig. 2 shows the corresponding temporal variation of the input Stokes parameters (e.g., as in [14], [15], and [20], we set here $\varphi_x = \varphi_y = 0$, $a_x = 3.307$, $a_y = 0.903$, and $\Psi = 0$): the input Stokes vector does not fully rotate across the entire Poincaré sphere, even though the input time averaged degree of polarization virtually vanishes. Note that the input polarization rotation due to the differential phase modulation $a_x - a_y = 0.765\pi$ is accompanied by a common PM $(a_x + a_y)/2 = 1.34\pi$ that leads to useful nonlinear pulse reshaping and enhanced eye opening [14], [15].

III. Simulations

For assessing the impact of BSPS on both intra- and interchannel nonlinear effects in the presence of PMD, we numerically simulated the transmission in a typical 10 Gbit/s NRZ-OOK WDM 1600 km long dispersion managed link involving 20 spans of standard single-mode fiber (SSMF) with dispersion compensating fibers (DCF). The SSMF span length was 80 km and its 0.2 dB/km loss was compensated by an EDFA; the dispersion (and dispersion slope) compensating fiber length was 16 km: its 0.5 dB/km loss was compensated by a second EDFA. EDFAs had a noise figure of 4 dB; the

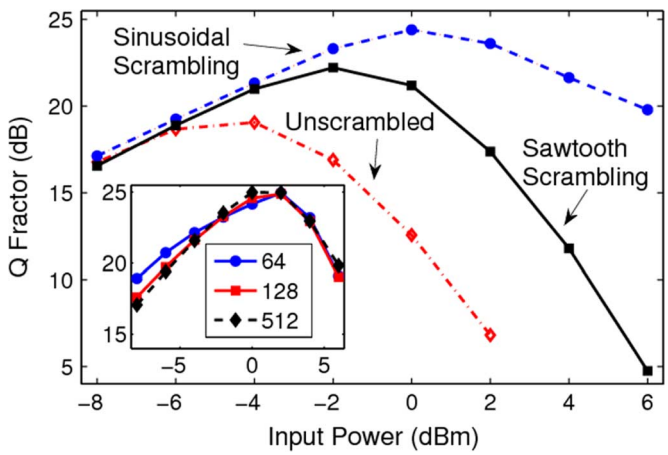

Fig. 3. Output Q factor versus average channel power without polarization scrambling (dash–dotted red curve), with sawtooth scrambling (solid black curve), or with sinusoidal scrambling (dashed blue curve) in the absence of PMD; inset: input power dependence of the output Q factor (with sinusoidal BSPS and $\Psi = 0$) when the number of signal bits was increased from 64 up to 128 or 512.

SSMF and DCF nonlinear index was $n_2 = 2.6 \times 10^{-20}$ m^2/W, and their core area was 80 and 50 μm^2, respectively. In our simulations, we used 64 signal bits per channel, and the Q factor was estimated by assuming Gaussian statistics for the optical beat noise.

Let us consider first the role of BSPS on the mitigation of single-channel nonlinear effects, or SPM, in the absence of PMD. To this end, we set the PMD of both SSMFs and DCFs to zero and compare the propagation of 10 Gbit/s NRZ signals with or without BSPS. Fig. 3 shows the output Q factor as a function of the input channel average power: the optimal operating power depends on whether polarization scrambling is applied or not, as well as on the scrambling method and additional PM (here as in [20] sinusoidal scrambling is applied with $\Psi = \pi/2$).

Indeed, Fig. 3 shows that BSPS leads to a large performance improvement in the high power regime, in particular when sinusoidal scrambling (5) is employed. In the inset of Fig. 3 we also show an example of the variation of the power dependence of the output Q factor as a function of the bit sequence length. As can be seen, with only 64 bits one already obtains Q factor fluctuations (about the values obtained for long sequences) of less than half dB for powers between -4 dBm and $+4$ dBm, i.e., around the optimal input power of 0 dBm. This justifies our choice of using 64 bits per channel only in our subsequent simulations, an approximation that permits us to significantly reduce the computational burden in our study of the statistical properties of the Q factor distribution in the presence of PMD.

In order to understand the physical reason for the significant performance improvement brought about by BSPS, it proves convenient to compare (within the last of the 20 SSMF/DCF spans) the temporal profiles of the optical field after the EDFA and before the last DCF, and after the second EDFA and the last DCF, without or with sinusoidal BSPS, respectively. We set the input average power at 0 dBm, in order to have a situation that leads to a large performance improvement by BSPS as

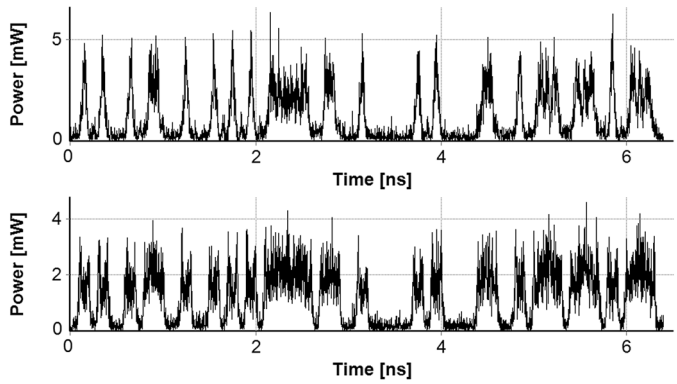

Fig. 4. Power profile of NRZ signal at the 20th span before (top) or after (bottom) the last DCF without polarization scrambling.

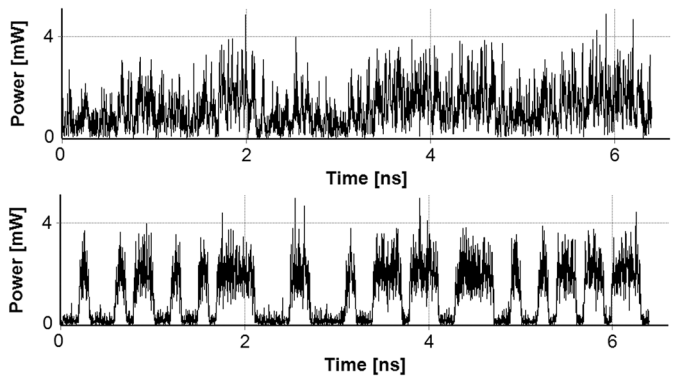

Fig. 5. Power profile of NRZ signal at the 20th span before (top) or after (bottom) the last DCF with sinusoidal polarization scrambling.

shown in Fig. 3. In Fig. 4, which refers to the case of no polarization scrambling, the top plot shows the pulse train before the DCF. As can be seen, the NRZ pulses are compressed in time by the combined effect of nonlinearity and anomalous GVD in the SMF: the pulse peak power is doubled from the input value up to 4 mW. Whereas the bottom plot in Fig. 4 shows the pulse train after the DCF and the second EDFA: here the marks suffer from significant amplitude fluctuations, which lead to the low Q factor values in Fig. 3. On the other hand, the temporal power profiles in Fig. 5 refer to the case of sinusoidal BSPS: the top plot shows that before the last DCF instead of compressing the pulses appear to substantially broaden in time. The associated spectra also exhibit a substantial spectral broadening due to the initial periodic PM associated with sinusoidal BSPS [see (2) and (5)]. Indeed, the field has a seemingly chaotic behavior, with a peak power that is rather evenly distributed around the original 2 mW value. In spite of the strong signal distortion before the DCF, the bottom plot of Fig. 5 shows that the pulses largely recover their shape after the last DCF. The significant amplitude fluctuations on the marks remain, but they are sufficiently separated from the zero levels so that the resulting error probability remains error-free. Of course, the amplitude jitter in Figs. 4 and 5 is mostly of high frequency, so it is largely removed by the electrical filter at the detector.

The comparison between the eye diagrams (not shown here) obtained after detection and electrical filtering (with a 7 GHz

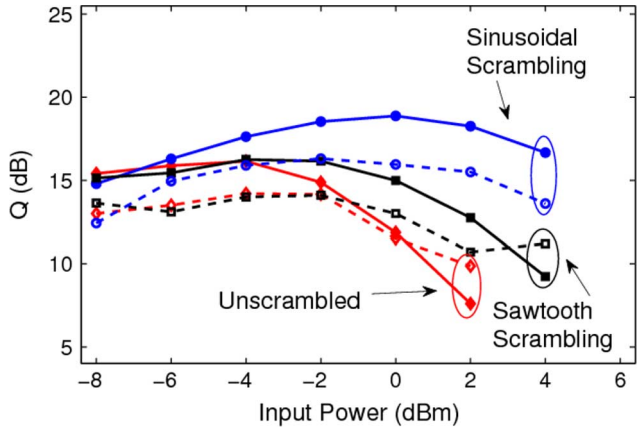

Fig. 6. Average (solid curves) or worst case after PMDC (dashed curves) Q factor versus input average channel power without polarization scrambling (red curve and diamonds), with sawtooth scrambling (black curve and squares), or with sinusoidal scrambling (blue curve and dots) in the presence of PMD.

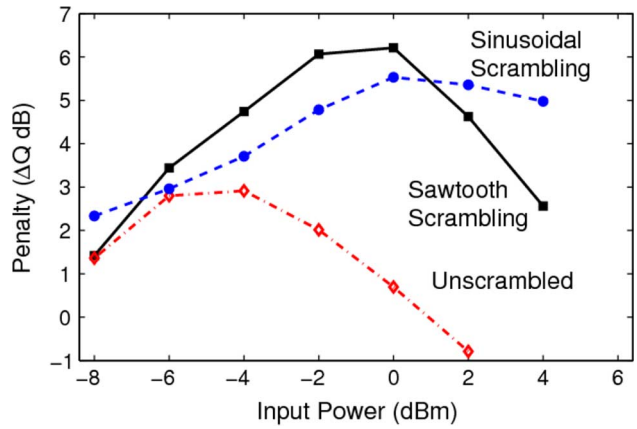

Fig. 7. Average Q factor PMD penalty versus input average channel power without polarization scrambling (dash–dotted red curve), with sawtooth scrambling (solid black curve), or with sinusoidal scrambling (dashed blue curve).

low-pass filter) without or with sinusoidal polarization scrambling reveals that in the first case there are significant low-frequency pulse amplitude fluctuations which cannot be filtered away from the signal; hence, they provide the major source of performance degradation. The origin for these low-frequency distortions is the larger nonlinearity that is seen at the entrance of the DCFs without polarization scrambling, because of the almost twice larger power of the compressed NRZ pulses in the unscrambled case when compared with the scrambled case.

The previous simulation results have shown that in the nonlinear propagation regime a large difference exists in the transmission performance between the polarization unscrambled and scrambled cases. These simulations were carried out in the ideal case of vanishing fiber PMD. Therefore, it is important to study next what is the additional degradation that is observed in the averaged (over a set of random realizations of the randomly birefringent fibers) Q factor when the PMD coefficient is raised from zero to a realistic finite value. To this end, let us consider the case when both the SSMFs and the DCFs have a common relatively large PMD coefficient, namely $D_{\rm PMD} = 1$ ps/km$^{1/2}$. As outlined in the previous section, our statistical modeling of the birefringent fibers used the coarse-step method including all orders of PMD, as well biasing through IS [19] the PMD statistics in order to explore rare events associated with relatively high DGD realizations.

At first, we computed the average Q factor after performing 30 runs with different realization of the randomly birefringence fibers (no PMD compensator was inserted at the end of the link). The resulting input power dependence of the average Q factor is illustrated by the solid curves in Fig. 6. In Fig. 6, we also show by dashed curves the Q factor with or without BSPS whenever a three-degree-of-freedom PMDC with optimum orientation is inserted before the receiver. For each type of input modulation, we selected among the 30 runs the particular PMD realization that led to a worst transmission performance (i.e., lowest Q factor) after the PMDC. Note that such realizations do not coincide, in general, with those giving worst performance before the PMDC. Interestingly, Fig. 6 shows that the worst-case Q factor remains generally lower than its average value, in spite of the inclusion of an optimized PMDC.

As can be seen by comparing Fig. 6 with 3, in all cases PMD leads to an average performance degradation. This degradation is relatively larger when polarization scrambling is employed, due to the PMD-induced conversion of input polarization fluctuations into output intensity fluctuations as outlined in [9]. It is interesting to connect our numerically observed degradation with previously known results. In the present case, we have a total (for 20 spans of 80 km SSMF and 16 km DCF) average DGD of about 44 ps, or 44% of the bit time. In the case of the RZ modulation format with a 50% duty cycle, such high value of DGD is known to lead in the linear propagation regime to about 1 dBQ of PMD-induced performance degradation [23]. For the NRZ format that we have studied so far, we may expect to obtain a slightly larger penalty level.

Indeed, by subtracting the average Q values in Fig. 6 from the Q values obtained with zero PMD (as in Fig. 3), one obtains the plot of the power-dependent average PMD-induced Q degradation or penalty which is shown in Fig. 7. Here, we show that at relatively low powers (i.e., -8 dBm) or in the linear regime the PMD-induced degradation is as expected between 1 and 2 dBs. However, because of SPM the larger performance degradations occur near the input power values that give optimal performance; next, the Q factor rapidly drops in the highly nonlinear regime where the performance is severely degraded. Figs. 6 and 7 show that the peak PMD-induced degradation is relatively larger when polarization scrambling is employed. Note that whenever a PMDC is employed, although linear PMD degradations can be in principle eliminated, nonlinear degradations (which already occur in the absence of PMD) as well as their interactions with PMD cannot be compensated for.

In order to further assess the transmission performance and its improvement by BSPS in the presence of both SPM and PMD, it proves convenient to consider a situation where fibers with a relatively low PMD coefficient $D_{\rm PMD}$ are used. In this way, we may ensure that the PMD-induced degradations remain relatively small (say, below 0.5 dBQ in the linear regime when assuming a safety margin as dictated by the Maxwellian distribution for the linear DGD of the fiber as discussed in [23]). Next we will compare the degradations with and without polarization scrambling in this case. In fact, Chowdhury *et al.* [23] have

shown that in order to transmit at with the NRZ-OOK format over 1600 km (in a configuration involving 20 spans of 80 km of SMF) with an outage probability of less than 10^{-7} in the presence of PMD, it is necessary that the PMD coefficient of the SSMFs is $D_{\rm PMD} < 0.2$ ps\km$^{1/2}$ (for a given and typical DCF PMD coefficient $D_{\rm PMD} = 0.35$ ps\km$^{1/2}$).

The results of Fig. 3 have shown that in the absence of PMD, the NRZ format with sinusoidal polarization scrambling (accompanied by PM) has superior performance with respect to both cases of either no polarization scrambling or pure (sawtooth) polarization scrambling. Therefore, we would like to study the effects of PMD in a situation where no PMD compensation is required at the system output. In other words, we shall investigate what is the maximum level of PMD that still leads to an acceptable level of outage probability without the need for PMD compensation at the receiver. In order to be able to estimate with good accuracy the outage probability, in particular in the presence of nonlinear effects, it is necessary to postulate or to calculate by numerical (Monte-Carlo) computation the probability density function (PDF) of the output Q factor in the presence of PMD. In practice, it is not possible with simple Monte-Carlo simulations to numerically evaluate the PDF, since in order to evaluate the tails of the PDF with reasonable accuracy at probability levels equal to or less than 10^{-10} one needs a prohibitively large number of simulations. On the other hand, it is well known that the tails of the PDF at levels as low as 10^{-20} can be computed with just a few thousand of simulations whenever the IS technique is employed [19].

We thus applied the IS technique to numerically compute the PDF of the output Q factor for the NRZ format with sinusoidal polarization scrambling and the input average power of either -4 dBm or $+4$ dBm. As can be seen in Fig. 3, without PMD in these two situations the output Q factor has nearly the same value of $Q = 21$ dB. Whereas in the first case the performance is ASE noise limited, and in the second case the performance is nonlinearity limited. Therefore, we may directly compare the effect of PMD on the statistics of the Q factor in either a noise-limited or a nonlinearity-limited transmission system. For the SSMFs, we have set the PMD coefficient $D_{\rm PMD} = 0.4$ ps\km$^{1/2}$ (which is twice the limiting value for the unscrambled NRZ format case), and kept fixed the DCFs the PMD coefficient at $D_{\rm PMD} = 0.35$ ps\km$^{1/2}$. The likelihood ratio L of each individual simulation was computed as the product of the likelihood ratios of each fiber in the 20-span link, i.e.,

$$L = \prod_{i=1}^{20} L_i^{\rm SSMF} L_i^{\rm DCF}$$

where for each fiber $L = (2/(\cos\theta + 1))^{1-\alpha}/\alpha$ [19].

The top plot in Fig. 8 illustrates the results of the IS simulations in the ASE-noise limited case (i.e., $P_{\rm in} = -4$ dBm): the PDF of the output Q factor was obtained by combining (using a simplified balance heuristic method [19]) the results of four different PDFs obtained with either no bias ($\alpha = 1$) or with three different bias levels, respectively (i.e., $\alpha = 1/(1 - {\rm C})$ with $C = 0.05, 0.1$, and 0.2). For each set of simulations, we have collected 5000 simulations, for a total of 20 000 runs. As can be

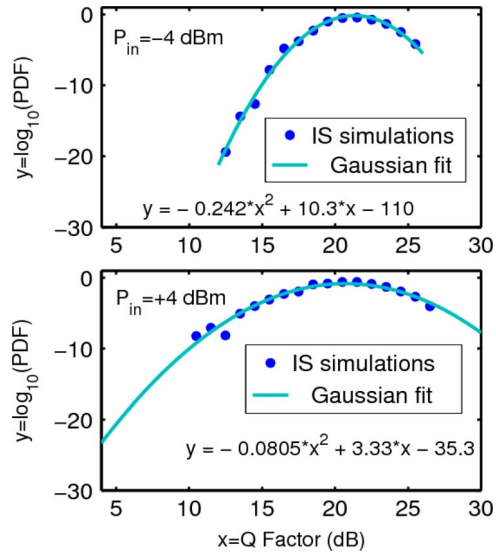

Fig. 8. Numerically computed PDF via IS (blue dots) and its Gaussian approximation (solid light blue curves) with sinusoidal polarization scrambling and input average power of -4 dBm (top) or $+4$ dBm (bottom).

seen, PDF levels as low as 10^{-20} can be easily reached by the IS technique for the PDF of the output Q factor. Moreover, it can be noticed that a good fit of the Q factor PDF is provided by a simple Gaussian distribution (continuous line). By using the analytical fit for the PDF and its associated cumulative distribution, we may estimate an outage probability (OP) (that is, the probability that the BER $> 10^{-9}$, or $Q > 15.6$ dB) as low as ${\rm OP} = 2.5 \times 10^{-9}$.

The bottom panel in Fig. 8 shows the result of the IS computations in the nonlinearity-limited case (i.e., $P_{\rm in} = +4$ dBm). Here, the PDF was numerically obtained by the IS technique and using three sets of simulations (one unbiased and two biased), with 5000 runs each for a total of 15000 simulations. As can be seen from the bottom panel in Fig. 8, the PDF is again well fitted by a Gaussian function. However in this case the average Q factor is degraded by about 1 dB with respect to the no-PMD case. More importantly, the standard deviation of the Gaussian distribution in the nonlinearity-limited case of Fig. 9, bottom, is about 1.73 times larger than that in the ASE-limited case of Fig. 8, top. Both facts contribute to a significant increase of the outage probability in the nonlinearity-limited case, namely we calculate that the ${\rm OP} = 5.6 \times 10^{-4}$.

After evaluating the PDF for the output Q factor for the two cases of either a noise or a nonlinearity-limited transmission (i.e., with the input power of -4 or $+4$ dBm) with sinusoidal polarization scrambling, let us consider now the PDF for the optimal case of an input power of 0 dBm, where in the absence of PMD the Q factor is the highest, thus indicating that the noise and nonlinearity-induced transmission impairments equally balance each other. The corresponding PDF as was obtained by collecting three sets of either unbiased or IS-biased simulations with 5000 runs each is shown in the top panel of Fig. 9. As can be seen, in this case the standard deviation is increased by SPM of 1.3 times with respect to the case with $P_{\rm in} = -4$ dBm. On the other hand, the outage probability is as low as

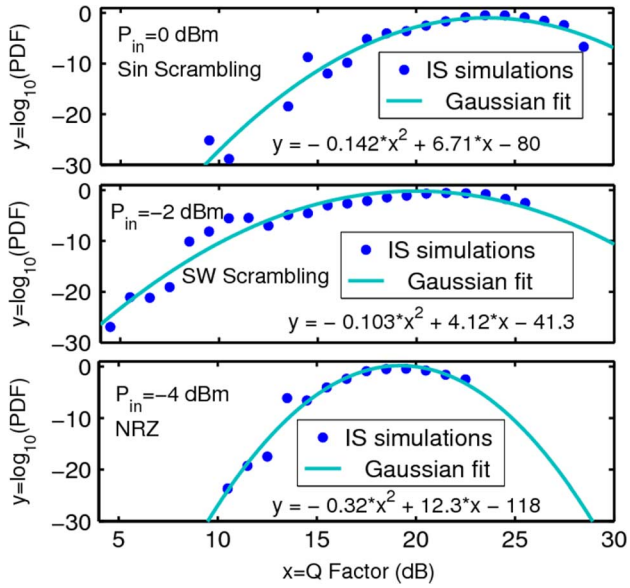

Fig. 9. Numerically computed PDF via IS (blue dots) and its Gaussian approximation (solid light blue curves) with: (top) sinusoidal polarization scrambling and input average power of 0 dBm; (middle) sawtooth polarization scrambling and input average power of -2 dBm; (bottom) no polarization scrambling and input average power of -4 dBm.

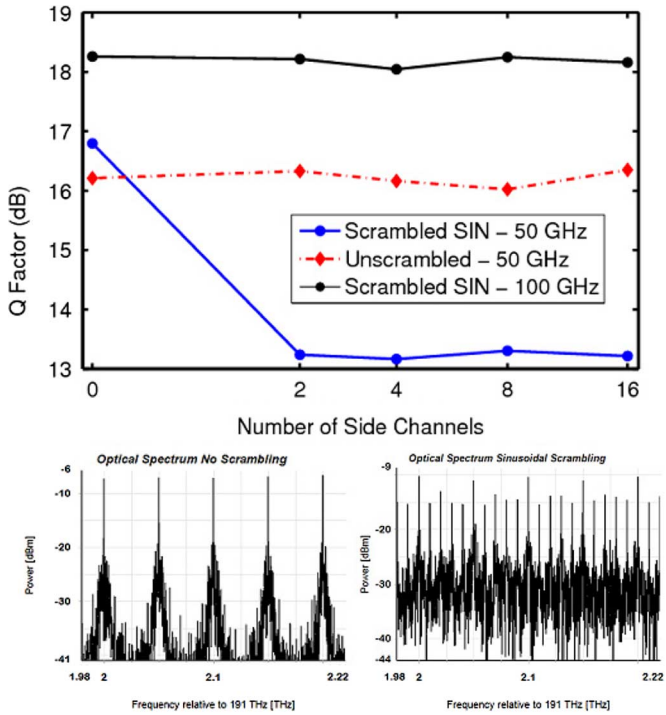

Fig. 10. Top: Q factor of central test channel after 1600 km transmission versus the number of side channels with or without input BSPS, for different channel spacing. Bottom: output WDM spectrum without or with BSPS.

OP $= 5.4 \times 10^{-11}$, owing to the overall improvement associated with the higher average Q factor.

Let us consider next the numerical estimation of the PDF in the case of a bit-synchronous scrambling of the input polarization state with a sawtooth temporal profile as in (3). If we take the input power $P_{\rm in} = -2$ dBm which is the optimal value in the absence of PMD, and by computing three different numerical frequency distributions with either an unbiased or a biased selection (including 5000 samples each) of the random birefringence with two different bias values, we obtain that the PDF may again be fitted by a Gaussian curve (see the middle panel in Fig. 9). The standard deviation of this Gaussian PDF with sawtooth polarization scrambling is 1.17 times that of sinusoidal scrambling. Moreover, the reduced peak (or average) value of the distribution, which is of only about $Q = 20$ dB with sawtooth scrambling, leads to a relatively large value of the outage probability, namely OP $= 3.5 \times 10^{-3}$.

Let us complete our comparison of the different PDF distributions in the presence of SPM and PMD by considering the case of a NRZ input modulation format without any input polarization scrambling. In this case, Fig. 3 shows that the optimal input power without PMD is equal to -4 dBm: the corresponding numerical PDF (again obtained from three distributions, one unbiased and two biased with 5000 samples each) illustrated in the bottom panel of Fig. 9 shows a relatively good fit with a Gaussian function. The standard deviation of the Q distribution is 0.67 times that of the optimal sinusoidal scrambling, and 0.57 that of the optimal sawtooth scrambling. Because of the reduced width of the distribution, even if the peak or average Q value is limited to about 19.5 dB, the associated outage probability is estimated at OP $= 1.8 \times 10^{-5}$, which is reduced with respect to the sawtooth polarization scrambling case (OP $= 3.5 \times 10^{-3}$), although it remains well above the outage probability with sinusoidal scrambling (OP $= 5.4 \times 10^{-11}$).

These results clearly show that when comparing the transmission performance of different modulation formats in the presence of fiber nonlinearity and PMD, it is essential to consider not only the average values of the Q factor but also its overall statistical distribution, which is ultimately necessary for the estimation of the outage probability. In particular, our analysis has shown that although pure (or sawtooth polarization scrambling) leads to higher average Q values with respect to the unscrambled case, the associated increase of the width of the PDF completely overturns the situation whenever the outage probability is computed.

So far, we have limited our study to the possibility of mitigating (by BSPS) SPM impairments in single-channel transmissions in the presence of PMD. Let us now briefly consider the impact of the additional nonlinear impairment brought about by XPolM in the case of WDM transmissions. To this end, we computed the average Q factor of the central channel in a WDM comb at the output of the 20 SSMF/DCF spans for 30 different realizations of the random PMD fiber link. For the estimation of the XPolM-induced performance degradation, we considered both single-channel propagation (zero side channels) and WDM transmissions with either 2, 4, 8, and 16 side channels, symmetrically placed around the central test channel. For consistency with the single-channel case that was shown in Figs. 6 and 7, we have set the PMD coefficient of both the SSMFs and the DCFs to the relatively large value of $D_{\rm PMD} = 1$ ps/km$^{1/2}$.

Fig. 10 illustrates the effect of BSPS on the dependence of the Q factor versus the number of side channels, averaged over 30 runs. The filled black and blue round points (connected by a solid curve as a guide to the eye) in Fig. 10 were obtained with sinusoidal scrambling, with an input average power per

channel of -2 dBm, and a channel spacing of 100 or 50 GHz, respectively. Moreover, the red-filled diamonds were obtained without any polarization scrambling, a channel power of -4 dBm, and 50 GHz channel spacing. The output test channel was selected by placing an optical filter (of width equal to the channel spacing) before the receiver. Fig. 10 shows that the presence of the side channels does not introduce additional penalties in the unscrambled case, whereas a large penalty is obtained with scrambling whenever the channel spacing is reduced from 100 to 50 GHz. The origin of such penalty is XPolM which is particularly severe because of the enhanced bandwidth of the individual channels due to the fast temporal polarization scrambling, as shown by the plots in the bottom part of Fig. 10. This result shows that although BSPS may improve the transmission performance in the single-channel case, in WDM transmissions BSPS is not compatible with dense channel spacing due to its higher sensitivity to interchannel impairments.

IV. Conclusion

We proposed and verified by statistical simulations based on the importance sampling method that polarization scrambling involving bit-synchronous sinusoidal polarization modulation and phase modulation may strongly mitigate nonlinear SPM impairments in 10 Gbit/s OOK-NRZ transmission systems. Indeed, we estimated that the PMD-induced outage probability may be reduced from $\mathrm{OP} = 1.8 \times 10^{-5}$ down to $\mathrm{OP} = 5.4 \times 10^{-11}$ by simultaneous polarization and phase modulation in a 1600 km long dispersion managed link involving 80 km SSMF spans. On the other hand, due to the associated spectral broadening, BSPM leads to performance degradation in WDM systems with 50 GHz channel spacing.

Indeed, the performance of a dispersion-managed transmission system using BSPS and PM is determined as a tradeoff among two competing effects. First, scrambling enhances the temporal breathing of pulses within the dispersion map, which spreads pulses more evenly over the time domain. This reduces the variability of the nonlinear phase shift on pulses, thus effectively reducing the nonlinear interaction of a single pulse with itself. Therefore, there is an overall improvement for single-channel transmissions. On the other hand, scrambling increases the spectral width of pulses, which increases their interaction with neighboring channels. The value of channel spacing in a WDM system will determine if nonlinear interactions between channels are more important than the interaction of a single pulse with itself. Thus, BSPS and PM are expected to be beneficial in situations where the intrachannel interactions represent a dominating source of distortion with respect to interchannel effects.

Our statistical approach based on the IS method is also relevant to coherent communication systems, where it has been shown that nonlinear interchannel penalties may be mitigated by increasing the fiber PMD [24]. The results of our analysis of the combined effect of PMD and nonlinearity in polarization-multiplexed, quadrature phase shift keying systems operating at the channel rate of 112 Gbit/s will be reported elsewhere.

Acknowledgment

The author would like to thank G. Biondini for helpful discussions on the importance sampling method. They are grateful to V. V. Kozlov for many helpful discussions, and to the anonymous reviewers for their constructive comments. Numerical simulations used the VPITransmissionMaker 8.6–8.7 software.

References

[1] C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *J. Lightwave Technol.*, vol. 14, no. 7, pp. 2806–2826, Jul. 2006.

[2] M. Karlsson and H. Sunnerud, "Effects of nonlinearities on PMD-induced system impairments," *J. Lightwave Technol.*, vol. 24, no. 11, pp. 4127–4137, Nov. 2006.

[3] M. Karlsson and H. Sunnerud, "PMD impact on optical systems: Single and multichannel effects," *J. Opt. Fiber Commun. Rep.*, vol. 1, no. 2, pp. 123–140, 2004.

[4] L. F. Mollenauer, J. P. Gordon, and F. Heismann, "Polarization scattering by soliton-soliton collisions," *Opt. Lett.*, vol. 20, no. 20, pp. 2060–2062, Oct. , 1995.

[5] A. Bononi, A. Vannucci, A. Orlandini, E. Corbel, S. Lanne, and S. Bigo, "Degree of polarization degradation due to cross-phase modulation and its impact on polarization-mode dispersion compensators," *J. Lightwave Technol.*, vol. 21, no. 9, pp. 1903–1913, Sep. 2003.

[6] C. Xie, L. Möller, D. Kilper, and L. Mollenauer, "Impact of cross-phase modulation induced polarization scattering on optical PMD compensation in WDM systems," *Opt. Lett.*, vol. 28, pp. 2303–2305, 2003.

[7] J. Lee, K. Park, C. Kim, and Y. Chung, "Effects of nonlinear crosstalk in optical PMD compensation," *IEEE Photon. Technol. Lett.*, vol. 14, no. 8, pp. 1082–1084, Aug. 2002.

[8] G. Zhang, J. T. Stango, X. Zhang, and C. Xie, "Impact of fiber nonlinearity on PMD penalty in DWDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 17, pp. 501–503, Feb. 2005.

[9] P. K. J. Park, J. H. Lee, and Y. C. Chung, "Optimization of polarization scrambling frequency in lightwave systems," *Opt. Eng.*, vol. 47, p. 035005, Mar. 2004.

[10] B. Wedding and C. N. Haslach, "Enhanced PMD mitigation by polarization scrambling and forward error correction," in *Proc. Optical Fiber Communication Exhibit*, 2001, Paper WAA1.

[11] C. R. Davidson, H. Zhang, Y. Cai, L. Liu, J.-X. Cai, A. Pilipetskii, M. Nissov, and N. S. Bergano, "Direct measure of system margin enhancement by polarization scrambling," in *Proc. Optical Fiber Commun. Conf.*, Los Angeles, CA, Feb. 2004, paper WE1.

[12] Z. Li, J. Mo, Y. Wang, and C. Lu, "Experimental evaluation of the effect of polarization scrambling speed on the performance of PMD mitigation using FEC," in *Proc. Optical Fiber Commun. Conf.*, Los Angeles, CA, Feb. 2004, paper MF69.

[13] N. S. Bergano, V. J. Mazurczyk, and C. R. Davidson, "Polarization scrambling improves SNR performance in a chain of EDFAs," in *Proc. Opt. Fiber Commun. Conf., ThR4*, San José, CA, 1994.

[14] N. S. Bergano, C. R. Davidson, and F. Heismann, "Bit-synchronous polarization and phase modulation scheme for improving the transmission performance of optical amplifier transmission systems," *Electron. Lett.*, vol. 32, no. 1, pp. 52–54, Jan. 4, 1996.

[15] N. S. Bergano and C. R. Davidson, "Wavelength division multiplexing in long-haul transmission systems," *J. Lightwave Technol.*, vol. 14, no. 6, pp. 1299–1308, Jun. 1996.

[16] Y. K. Park, T. V. Nguyen, O. Mizuhara, C. D. Chen, L. D. Tzeng, P. D. Yeates, F. Heismann, Y. C. Chen, D. G. Ehrenberg, and J. C. Feggeler, "Field demonstration of 10 Gb/s line-rate transmission on an installed transoceanic submarine lightwave cable," *IEEE Photon. Technol. Lett.*, vol. 8, no. 3, pp. 425–427, Mar. 1996.

[17] T. Ito, Y. Inada, K. Fukuchi, and T. Suzaki, "Suppression of nonlinear waveform distortion using alternate polarization modulation for long distance 10 Gb/s based WDM transmission systems," in *Proc. Opt. Fiber Commun. Conf., ThQ7*, San Diego, CA, 1999.

[18] C. Xie, I. Kang, A. H. Gnauck, L. Möller, L. F. Mollenauer, and A. R. Grant, "Suppression of intrachannel nonlinear effects with alternate-polarization formats," *J. Lightwave Technol.*, vol. 22, no. 3, pp. 806–812, Mar. 2004.

[19] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion: Techniques and applications," *J. Lightwave Technol.*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.

[20] D. Wang and C. R. Menyuk, "Polarization evolution due to the Kerr nonlinearity and chromatic dispersion," *J. Lightwave Technol.*, vol. 17, no. 12, pp. 2520–2529, Dec. 1999.

[21] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 15, no. 9, pp. 1735–1745, Sep. 1997.

[22] F. Heismann, D. A. Gray, B. H. Lee, and R. W. Smith, "Electrooptic polarization scramblers for optically amplified long-haul transmission systems," *IEEE Photon. Technol. Lett.*, vol. 6, no. 9, pp. 1156–1158, Sep. 1994.

[23] D. Chowdhury, M. Mlejnek, and Y. Mauro, "Numerical modeling of PMD," *J. Opt. Fiber Commun. Rep.*, vol. 1, pp. 141–149, 2004.

[24] P. Serena, N. Rossi, O. Bertran-Pardo, J. Renaudier, A. Vannucci, and A. Bononi, "Intra- versus inter-channel PMD in linearly compensated coherent PDM-PSK nonlinear transmissions," *J. Lightwave Technol.*, vol. 29, no. 11, pp. 1691–1700, 2011.

Author biographies not included at author request due to space constraints.