Calculation of Penalties Due to Polarization Effects in a Long-Haul WDM System Using a Stokes Parameter Model

D. Wang and C. R. Menyuk, *Fellow, IEEE*

***Abstract*—We derive a Stokes parameter model to calculate the penalties due to the combination of polarization mode dispersion (PMD), polarization dependent loss (PDL), and polarization dependent gain (PDG) in long-haul, dense wavelength divison multiplexed (WDM) systems. In this model, we follow the Stokes parameters for the signal and the noise in each channel instead of following the full time domain behavior of each channel. This approach allows us to determine the statistical distribution of penalties with up to 10^5 fiber realizations and 40 channels. We validate this model to the extent possible by comparison to full numerical simulations. Using this model, we find that the interaction of PMD and PDL is the major source of penalties and that the effect of PDG is negligible in WDM systems with more than ten channels.**

***Index Terms*—Optical fiber transmission, outage probabilities, polarization effects, reduced models, Stokes parameters, wavelength division multiplexing (WDM).**

I. Introduction

DUE TO THE rapid increase in the demand for bandwidth, wavelength divison multiplexed (WDM) systems have been widely deployed in trans-oceanic links as well as continental and metropoliton networks. There are three polarization effects that lead to impairments in the long-haul optical fiber transmission systems: polarization mode dispersion (PMD), polarization dependent loss (PDL), and polarization dependent gain (PDG) [1]–[3]. Reference [4] shows how these three effects can combine to produce signal impairments in single-channel systems. In particular, PDG can lead to excess noise in the polarization orthogonal to the signal, and it therefore plays an important role in determining the degradation and variance of the Q factor. By contrast, dense WDM systems whose channels are spread over a large bandwidth rapidly change their relative polarization states due to PMD so that the overall degree of polarization of the system is nearly zero, and PDG is ineffective. At the same time, different channels experience different amounts of PDL, and, since the amplifiers maintain the total signal power nearly constant, individual channels undergo a kind of random walk so that it is possible for some channels to fade.

Manuscript received March 21, 2000. This work was supported by the Department of Energy, the National Science Foundation, and AT&T Bell Laboratories.

D. Wang is with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA.

C. R. Menyuk is with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA, and also with the Laboratory for Telecommunications Sciences, c/o USARL, Adelphi, MD 20783-1197 USA (e-mail menyuk@umbc.edu).

Publisher Item Identifier S 0733-8724(01)00740-X.

Calculating the impairments due to the combination of PMD, PDL, and PDG in WDM systems is a formidable theoretical challenge. Running a complete split-step simulation with just one fiber realization can cost many minutes or even hours of CPU time [5], and it is necessary to know the penalties for many thousands of realizations in order to accurately determine the outage probabilities [4].

In this paper, we propose a reduced model that is based on following the four Stokes parameters for the signal and the four Stokes parameters for the noise for every channel in the WDM system. Using this model, we calculate the penalties due to polarization effects for up to 10^5 fiber realizations and 40 channels. We then determine the outage probabilities for margins of 2.5 and 3.0 dB. We validate this reduced model to the extent possible by comparison to full, split-step simulations. While we cannot run enough realizations of the full model to completely validate the reduced model (the very difficulty that prompted us to develop the reduced model in the first place), we do find that they are consistent. We also have validated this model by comparison to recirculating loop experiments, and the results will be reported elsewhere [6].

From an intuitive standpoint, it is reasonable to expect this reduced model to work well. PMD, PDL, and PDG are slow time effects that will affect a whole channel in the same way, as long as the PMD is not so large that it distorts a single channel. By contrast, nonlinearity and chromatic dispersion are fast effects that affect each bit separately. Since the effects of PMD, PDL, and PDG on the one hand and effects of nonlinearity and dispersion on the other hand exist on different time scales, it is reasonable to anticipate that the penalties due to these two kinds of effects will be separable. We have partially verified this assumption in earlier work [7], [12], and we do so again here.

The remainder of this paper is organized as follows: In Section II, we derive the Stokes model and describe the numerical procedure that we use to solve it. In Section III, we validate this model to the extent possible by comparison to full, split-step simulations. In Section IV, we apply this model to calculate the outage probability for parameters that correspond to a realistic trans-oceanic system with margins of 2.5 and 3.0 dB for polarization effects. We show that the effect of PDG is negligible in WDM systems with more than ten channels.

II. The Stokes Model

Since we are interested in the evolution of the polarization of an entire communication channel, we will be focusing on the

evolution of the averaged Stokes parameters for each separate channel m in a WDM system. We first define $\mathbf{U}(z, t)$ as the Jones vector in the time domain. We next write $\mathbf{U}$ as a sum of contributions over n channels, obtaining

$$\mathbf{U} = \sum_{m=1}^{n} \mathbf{U}^{(m)} \exp\left[ik^{(m)}z - i\omega^{(m)}t\right] \tag{1}$$

where $k^{(m)}$ and $\omega^{(m)}$ are the central wavenumber and frequency of the mth channel measured with respect to the central wavenumber and frequency of $\mathbf{U}$. The variable $\mathbf{U}^{(m)}$ is the corresponding wave envelope. The definition of the Stokes parameters for each channel is

$$\begin{aligned}
S_0^{(m)} &= \frac{1}{T}\int_{t_1}^{t_2}\left[\left|u_x^{(m)}(t)\right|^2 + \left|u_y^{(m)}\right|^2\right]\,dt \\
S_1^{(m)} &= \frac{1}{T}\int_{t_1}^{t_2}\left[\left|u_x^{(m)}(t)\right|^2 - \left|u_y^{(m)}\right|^2\right]\,dt \\
S_2^{(m)} &= \frac{2}{T}\int_{t_1}^{t_2}\mathrm{Re}\left[u_x^{(m)}(t)u_y^{(m)^*}(t)\right]\,dt \\
S_3^{(m)} &= \frac{2}{T}\int_{t_1}^{t_2}\mathrm{Im}\left[u_x^{(m)}(t)u_y^{(m)^*}(t)\right]\,dt
\end{aligned} \tag{2}$$

where $T = t_2 - t_1$, while $u_x^{(m)}(t)$ and $u_y^{(m)}(t)$ are the wave envelopes in two orthogonal polarizations. We are assuming that T is very large compared to a single bit period and that the channel becomes statistically stationary when T is large so that this definition is meaningful. We will treat the last three Stokes parameters $S_1^{(m)}$, $S_2^{(m)}$, and $S_3^{(m)}$ just like the three components of a Stokes vector at a single frequency $\omega = \omega^{(m)}$, the central frequency of channel m.

Since we will only be following one set of Stokes parameters for the signal and another set for the noise in each channel in the WDM system, we must convert from the Jones representation to the Mueller representation so that we can deal with partially polarized channels.

A. Stokes Model for PMD

PMD in the fiber will cause the polarization states of different channels to evolve differently. However, when the PMD is too small to cause distortion in a single channel, then the polarization states in a single channel all evolve uniformly. In our model, we simply calculate the evolution due to PMD of the last three Stokes parameters, $S_1^{(m)}$, $S_2^{(m)}$, and $S_3^{(m)}$, in each channel as if they correspond to a single Stokes vector at the channel's central frequency $\omega^{(m)}$. This approximation is reasonable as long as the accumulated differential group delay in each channel is not large compared to the pulse duration. Since there is no PDL in the optical fiber, in contrast to the amplifiers where we will take into account the PDL separately, and since the polarization-independent loss must be exactly compensated by the gain over the length of the transmission line, so that we may ignore the spatially varying gain and loss, we find that $S_0^{(m)}$ is unaffected by PMD.

We calculate the evolution of the Stokes parameters using a variant of the coarse step method [8]. Referring to the step size along the fiber transmission path between amplifiers as ζ, which we typically took to be 1 km, we find $S_0^{(m)}(z+\zeta) = S_0^{(m)}(z)$. In this expression, $z = z_0 + j\zeta$, where z_0 is the location of one amplifier and $j = 1, \cdots, N$, where N is the total number of steps to the next amplifier. Also, we find that the Stokes vector portion of the Stokes parameters $\mathbf{S}^{(m)} = (S_1^{(m)}, S_2^{(m)}, S_3^{(m)})^t$ transforms on the jth step according to the relationship

$$\mathbf{S}^{(m)}(z+\zeta) = M_R^{(m)} M_j(z)\mathbf{S}^{(m)}(z) \tag{3}$$

where $M_R^{(m)}$ and M_j in (4a) and (4b) shown at the bottom of the page. The quantity $\Delta\beta'$ is proportional to the average inverse group velocity difference along the two polarization axes due to the birefringence. It should be chosen so that if $\langle\tau\rangle$ is the expected value of the differential group delay due to PMD over a length $Z = N\zeta$, then $\Delta\beta' = (3\pi/32\zeta Z)^{1/2}\langle\tau\rangle$ [8], [9]. The ϕ_j and ψ_j are random variables, chosen independently at each j, from uniform distributions in the range $[0, 2\pi]$, while θ_j is a random variable chosen independently at each j such that $\cos(\theta_j)$ is uniformly distributed in the range $[-1, 1]$. This choice corresponds to a random rotation with a uniform probability distribution on the Poincaré sphere. We stress that θ_j, ϕ_j, and ψ_j are the same for all channels. A detailed demonstration that this model accurately reproduces the linear PMD may be found in [8] and [9].

B. Stokes Model for PDL

PDL is due to the polarization dependence of the transmission in some devices, notably the WDM couplers in the amplifiers. Typical values are less than 0.1 dB in a single amplifier, and it is important to keep this value low [4]. The effect of a polarization dependent loss element is to cause excess loss in one of two orthogonal polarizations. Using the Jones vector notation defined earlier, where we take the second component to be in the direction of maximum loss, we may write

$$\begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\text{after}} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\text{before}} \tag{5}$$

$$M_R^{(m)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\Delta\beta'\omega^{(m)}\zeta) & -\sin(\Delta\beta'\omega^{(m)}\zeta) \\ 0 & \sin(\Delta\beta'\omega^{(m)}\zeta) & \cos(\Delta\beta'\omega^{(m)}\zeta) \end{pmatrix} \tag{4a}$$

$$M_j = \begin{pmatrix} \cos\theta_j & \sin\theta_j\cos\psi_j & -\sin\theta_j\sin\psi_j \\ -\sin\theta_j\cos\phi_j & \cos\theta_j\cos\phi_j\cos\psi_j - \sin\phi_j\sin\psi_j & -\cos\theta_j\cos\phi_j\sin\psi_j - \sin\phi_j\cos\psi_j \\ -\sin\theta_j\sin\phi_j & \cos\theta_j\sin\phi_j\cos\psi_j + \cos\phi_j\sin\psi_j & -\cos\theta_j\sin\phi_j\sin\psi_j + \cos\phi_j\cos\psi_j \end{pmatrix}. \tag{4b}$$

where "before" and "after" refer to the values before and after the PDL element, and α is related to $x_{\rm PDL}$, the PDL measured in dB, through the relationship, $x_{\rm PDL}$ (in dB) $= -20\log_{10}\alpha$. From (2) and (5) we find

$$\begin{aligned} S^{(m)}_{0,\,\rm after} &= \frac{1+\alpha^2}{2}\,S^{(m)}_{0,\,\rm before} + \frac{1-\alpha^2}{2}\,S^{(m)}_{1,\,\rm before} \\ S^{(m)}_{1,\,\rm after} &= \frac{1-\alpha^2}{2}\,S^{(m)}_{0,\,\rm before} + \frac{1+\alpha^2}{2}\,S^{(m)}_{1,\,\rm before} \\ (S_2+iS_3)^{(m)}_{\rm after} &= \alpha(S_2+iS_3)^{(m)}_{\rm before} \end{aligned} \tag{6}$$

where we recall that the Stokes parameters are averaged over time.

C. Stokes Model for PDG

PDG is due to polarization hole burning induced by the incoming signal to an erbium-doped fiber amplifier (EDFA). The gain in the polarization orthogonal to the incoming signal is larger than the gain in the polarization of the incoming signal. The amount of PDG in a single amplifier is only about 0.07 dB for an EDFA with 3 dB of gain compression and becomes larger as the amplifier goes deeper into gain compression. The magnitude of the polarization hole burning is proportional to the degree of polarization, $d_{\rm pol}$, of the incoming signal.

We will model the PDG much like the PDL, except that the direction of maximum gain must be chosen self-consistently with the existing signal in a given system. Thus, if we ignore the noise contribution for the moment, returning to it later, and we write

$$S_0^{\rm total} = \sum_{m=1}^{n} S_0^{(m)} \quad \mathbf{S}^{(\rm total)} = \sum_{m=1}^{n} \mathbf{S}^{(m)} \tag{7}$$

we find the total degree of polarization $d_{\rm pol} = |\mathbf{S}^{(\rm total)}|/S_0^{(\rm total)}$ and the total state of polarization, $\mathbf{s} = \mathbf{S}^{(\rm total)}/|\mathbf{S}^{(\rm total)}|$. We now write

$$\begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\rm after} = R \begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix} R^{-1} \begin{pmatrix} u_x(t)^{(m)} \\ u_y(t)^{(m)} \end{pmatrix}_{\rm before} \tag{8}$$

where g is the polarization dependent gain, normalized to the gain in the polarization state of the input signal. The value of g is related to $x_{\rm PDG}$, the PDG measured in dB, through the relationship [1]

$$x_{\rm PDG}\, d_{\rm pol} = 20\log_{10} g. \tag{9}$$

The rotation matrix R is determined by the overall polarization state of the signal and noise since it is this polarization state that determines the orientation of the PDG, while R^{-1} is the inverse of R. It is useful to define angles φ, ψ, and ϕ that are related to R *via* the relationship as shown in (10) at the bottom of the page. We find that the elements of R are related to $\mathbf{s}$ through the relationships

$$\begin{aligned} s_1 &= |r_{11}|^2 - |r_{12}|^2 = \cos\varphi \\ s_2 + is_3 &= 2r_{11}r_{12}^* = \sin\varphi\exp(-i\phi). \end{aligned} \tag{11}$$

We note that it is not possible to determine ψ from $\mathbf{s}$. However, the angle ψ does not affect the evolution of the Stokes parameters of the individual channels in any way and so can be safely ignored.

We once again transform from the Jones representation to the Stokes representation using (2) to do the appropriate time average for each channel. We obtain

$$\begin{aligned} \mathbf{S}^{(m)}_{\rm after} &= -\frac{g^2-1}{2}\,S^{(m)}_{0,\,\rm before}\cdot\mathbf{s} + \frac{g^2+1}{2}\,\mathbf{S}^{(m)}_{\rm before} \\ &\quad + \frac{(g-1)^2}{2}\,\mathbf{s}\times\left(\mathbf{s}\times\mathbf{S}^{(m)}_{\rm before}\right) \\ S^{(m)}_{0,\,\rm after} &= \frac{g^2+1}{2}\,S^{(m)}_{0,\,\rm before} - \frac{g^2-1}{2}\,\mathbf{s}\cdot\mathbf{S}^{(m)}_{\rm before}. \end{aligned} \tag{12}$$

D. Combining PDG with PMD, PDL, and ASE Noise

We account for the ASE noise by following four noise Stokes parameters ($S^{(m)}_{0,\,\rm noise}$, $\mathbf{S}^{(m)}_{\rm noise}$) at each m. We must track these Stokes parameters separately from the signal Stokes parameters because they are random variables while the signal Stokes parameters are deterministic. Since the ASE noise is unpolarized, each amplifier will cause the following change in the Stokes parameters:

$$\begin{aligned} S^{(m)}_{0,\,\rm noise,\,after} &= S^{(m)}_{0,\,\rm noise,\,before} + 2n_{\rm sp}(G-1)B^{(m)}h\nu \\ \mathbf{S}^{(m)}_{\rm noise,\,after} &= \mathbf{S}^{(m)}_{\rm noise,\,before} \end{aligned} \tag{13}$$

where

- $n_{\rm sp}$ spontaneous emission factor;
- G amplifier gain;
- $h\nu$ energy of a single photon;
- $B^{(m)}$ optical bandwidth of the mth channel.

These Stokes parameters are affected by the PMD, PDL, and PDG in exactly the same way as the signal Stokes parameters and participate in determining the degree of polarization and total Stokes parameters. Additionally, if there is any part of the gain bandwidth of the EDFA that is not included in one of the optical channels, then this noise energy will participate in the total energy balance. We may write for this additional portion

$$S^{(\rm add)}_{0,\,\rm noise,\,after} = S^{(\rm add)}_{0,\,\rm noise,\,before} + 2n_{\rm sp}(G-1)B^{(\rm add)}h\nu \tag{14}$$

$$R = \begin{pmatrix} \cos(\varphi/2)\exp[-i(\psi+\phi)/2] & -\sin(\varphi/2)\exp[i(\psi-\phi)/2] \\ \sin(\varphi/2)\exp[-i(\psi-\phi)/2] & \cos(\varphi/2)\exp[i(\psi+\phi)/2] \end{pmatrix}. \tag{10}$$

and we will assume that this contribution is unpolarized. We now write

$$S_0^{(\text{total})} = \sum_{m=1}^{n} S_0^{(m)} + \sum_{m=1}^{n} S_{0,\,\text{noise}}^{(m)} + S_{0,\,\text{noise}}^{(\text{add})}$$
$$\mathbf{S}^{(\text{total})} = \sum_{m=1}^{n} \mathbf{S}^{(m)} + \sum_{m=1}^{n} \mathbf{S}_{\text{noise}}^{(m)}. \tag{15}$$

The degree of polarization may now be written $d_{\text{pol}} = |\mathbf{S}^{(\text{total})}|/S_0^{(\text{total})}$. The final step in the procedure is to take into account the effect of gain saturation by assuming that the total power at the output of the amplifier is fixed at a value $\mathcal{S}$. We then renormalize $S_0^{(m)}$, $\mathbf{S}^{(m)}$, $S_{0,\,\text{noise}}^{(m)}$, $\mathbf{S}_{\text{noise}}^{(m)}$, $\mathbf{S}_0^{(\text{add})}$ by the factor $\mathcal{S}/S_0^{(\text{total})}$ which takes into account the renormalization of the total power that occurs in real systems due to gain saturation in the amplifiers.

From the calculated signal and noise Stokes parameters, it is possible to determine $Q^{(m)}$—the so-called Q factor—for each channel m and from that to infer the penalty due to PDL and PDG. To calculate this penalty, we first note that a $Q^{(m)}$ that we calculate from this model is not meaningful by itself because this reduced model does not take into account the degradation due to nonlinearity and dispersion. What is meaningful is the *difference* between the $Q^{(m)}$ values that we calculate when PDL and PDG are present and when they are absent for a fixed value of PMD.

To calculate $Q^{(m)}$ for a particular choice of PMD, PDL, and PDG, we first obtain the signal-to-noise (SNR) ratio of channel m which is equal to

$$\text{SNR}^{(m)} = \frac{P_{\text{peak}}}{P_{\text{ave}}} \frac{S_{0,\,\text{signal}}^{(m)}}{S_{0,\,\text{noise}}^{(m)}} \tag{16}$$

where the ratio $P_{\text{peak}}/P_{\text{ave}}$ is the ratio of the peak power in a mark to the average power in the signal channel. For the standard nonreturn-to-zero (NRZ) format, this ratio is two; for the standard return-to-zero (RZ) format, this ratio is four; and, for the chirped returned-to-zero (CRZ) format of Bergano, *et al.* [10], this ratio is approximately 5.3. In principle, these ratios are lowered somewhat by the electrical filtering in the receiver; however, we have found that our results are insensitive to these corrections. We may then use a formula relating the Q-factor to the signal-to-noise ratio (SNR) assuming that the noise is Gaussian distributed [11], [13]

$$Q^{(m)} = \frac{\text{SNR}^{(m)}}{\sqrt{2\text{SNR}^{(m)}+1}+1} \sqrt{\frac{2B_{\text{opt}}}{B_{\text{elec}}}} \tag{17}$$

where B_{opt} is the optical bandwidth and B_{elec} is the electrical bandwidth. We note that this expression includes contributions from both the spontaneous-spontaneous beat noise and the signal-spontaneous beat noise. Physically, the electrical detector at the end of transmission line receives $2B_{\text{opt}}/B_{\text{elec}}$ noise modes. Therefore, the signal-spontaneous beat noise $S_{0,\,\text{sig}-\text{spon}}$ is given by $(P_{\text{peak}}/P_{\text{ave}})^{1/2}(S_0 S_{0,\,\text{noise}})^{1/2}(B_{\text{elec}}/2B_{\text{opt}})^{1/2}$, while the spontaneous-spontaneous beat noise $S_{0,\,\text{spon}-\text{spon}}$ just equals $S_{0,\,\text{noise}}$. The noise power in the marks is given by $S_{0,\,\text{sig}-\text{spon}} + S_{0,\,\text{spon}-\text{spon}}$ while the noise power in the spaces is just given by $S_{0,\,\text{spon}-\text{spon}}$.

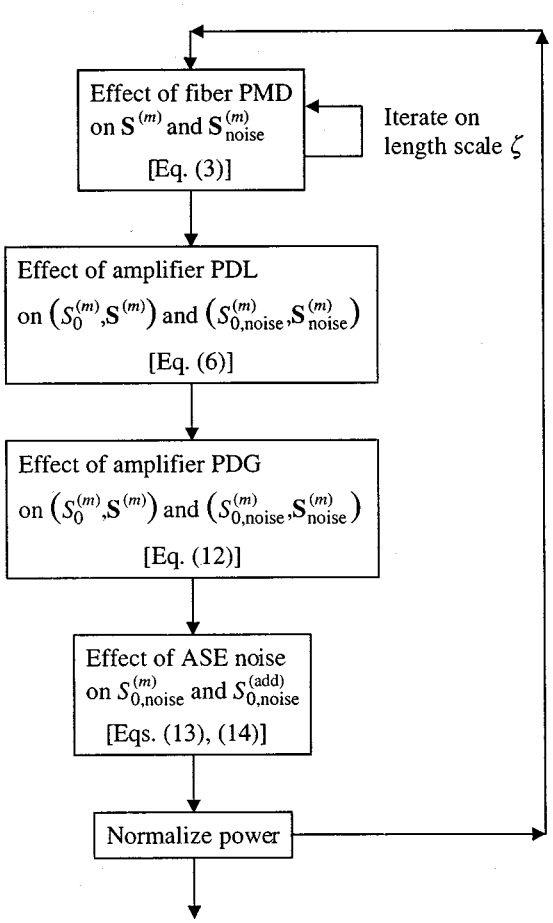


Fig. 1. Schematic illustration of the Stokes model procedure. One iteration between amplifiers is shown.

We summarize the complete procedure schematically in Fig. 1. This procedure is repeated iteratively from amplifier to amplifier.

III. Validation

In order to validate the Stokes model of polarization effects, we built a full model using the Manakov-PMD equation [8]. We then compared the Stokes model to the full model for both single-channel and eight-channel systems at a data rate of 10 Gbits/sec per channel. In the WDM studies, we used a 1 nm channel spacing. We have also carried out comparisons at 5 Gbits/s that we will not present here, and the results were similar. In the full model, we studied NRZ, RZ, and chirped returned to zero CRZ data formats, although we will only present the results for the RZ systems in this paper. The results for NRZ and CRZ are similar and can be found in [9].

We used a periodic dispersion map that consisted of one section of a single mode fiber whose dispersion D_1 at $\lambda_0 = 1.55$ μm is 16 ps/nm-km and whose length is 264 km, and another section of dispersion-shifted fiber whose dispersion D_2 at $\lambda_0 = 1.55$ μm is -2 ps/nm-km and whose length is 33 km. In both sections, we used a dispersion slope of 0.07 ps/nm^2-km. We

used pre and postdispersion compensation, split equally, to compensate for the excess dispersion in channels for which $\lambda \neq \lambda_0$. At the end of the transmission line, we optically filtered the separate channels with a 60 GHz, tenth-order Bessel filter in the WDM simulations, and we then electrically filtered each channel using a 10 GHz tenth-order Bessel filter. We used clock recovery to determine the boundaries of the time slots in each channel. PMD is included using the coarse step method [8], PDL and PDG are included using (5) and (8) respectively, and ASE noise is included using standard Monte Carlo methods [9]. For each set of parameters we ran 20 cases, each of which corresponds to a different realization of the random variations of the birefringence axes of the fiber and a different realization of the ASE noise. However, we chose the bit string to be the same in each channel in all 20 cases in order to avoid Q-variations from case to case due to pattern dependences in the limited strings of 64 bits per channel that we could keep in the simulations.

For each set of parameters, we set the decision level for the marks and spaces empirically to obtain the best average SNR. We determined the SNR for each of our 20 realizations and then calculated the Q factor using (17). Using the Q factor when the PDL and PDG are zero, we then determined $\Delta Q^{(m)}$ for all 20 realizations and from that we found the mean $\langle \Delta Q^{(m)} \rangle$ and the standard deviation $\sigma_Q^{(m)}$ for comparison to the reduced model.

Given the large random variation of the signal-spontaneous beat noise from bit to bit which leads to significant variations from realization to realization, 20 realizations is not really sufficient to accurately determine $\langle \Delta Q^{(m)} \rangle$ and $\sigma_Q^{(m)}$. We used only 20 realizations because for each set of parameters in which we ran an eight-channel simulation, the simulation required between 26 and 27 CPU hours on an SGI Onyx. Additionally, there can be a large amount of variation due to pattern dependences since we only keep 64 bits in each channel. Thus, a comparison of the reduced model to the full model should be viewed as a demonstration of consistency rather than a complete check of the reduced model.

We note that in each of our comparisons with the reduced model, we used 2000 realizations of the reduced model. Additionally, we note that we have also validated this model by comparison to recirculating loop experiments, and these results will be presented separately [6]. This comparison showed excellent agreement between the Stokes model and the experiments.

A. Single Channel Comparisons

We first compare the full model to the Stokes model in the simple case when the pulse modulation format is RZ and there is only a single channel. We show the results as a function of the PDL in Figs. 2 and 3, setting the PMD $= 0.1$ ps/km$^{1/2}$ and the PDG $= 0.0$ dB and 0.06 dB respectively. The agreement between the two models is quite good. The PDL values that we compared are 0.1, 0.2, $\cdots$, 0.6 dB. We note that when $\sigma_Q = 1$, the expected deviation of the full model from its mean is approximately $1/\sqrt{19} = 0.23$ since we only have 20 realizations at each value of the PDL. Thus, we find that the deviation between the full model and the Stokes model lies within the expected statistical error of the full model. Indeed, given the small number of realizations for the full model and the small number of bits in each realization, we consider the Stokes model to be at least as reliable as the full model. We note that the difference between the two models appears to be systematic rather than purely random since the full model consistently yielded either higher or lower values than the Stokes model for both $\langle \Delta Q \rangle$ and σ_Q in every plot as we varied the PDL. This systematic deviation is not surprising because the realizations with different

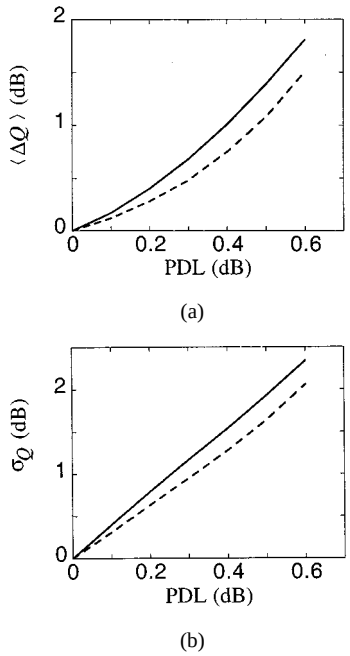


Fig. 2. Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD $= 0.1$ ps/km$^{1/2}$, PDG $= 0.0$ dB. (a) $\langle \Delta Q \rangle$. (b) σ_Q. Solid lines indicate the Stokes model, and dashed lines indicate the average of the full model.

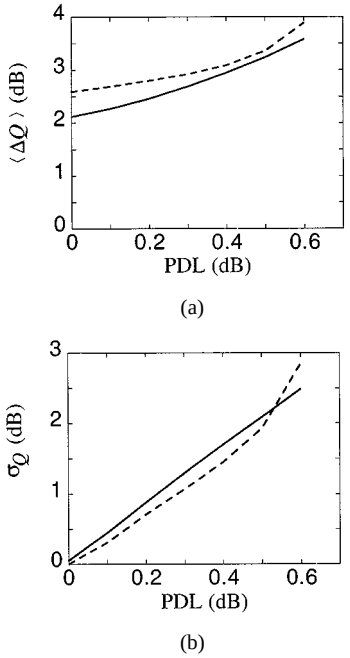


Fig. 3. Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD $= 0.1$ ps/km$^{1/2}$, PDG $= 0.06$ dB. (a) $\langle \Delta Q \rangle$. (b) σ_Q. Solid lines indicate the Stokes model, and dashed lines indicate the average of the full model.

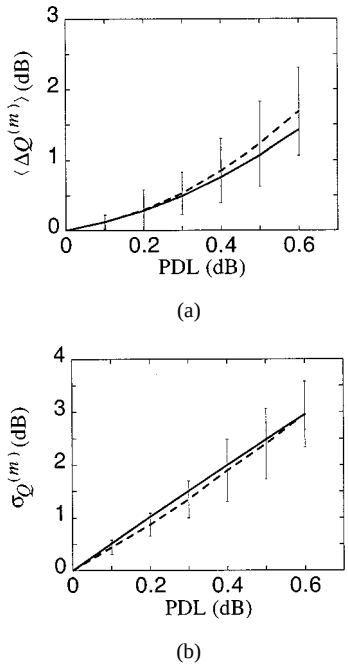


Fig. 4. Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD $= 0.1$ ps/km$^{1/2}$, PDG $= 0.0$ dB. (a) $\langle \Delta Q^{(m)} \rangle$. (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of the values for all eight channels.

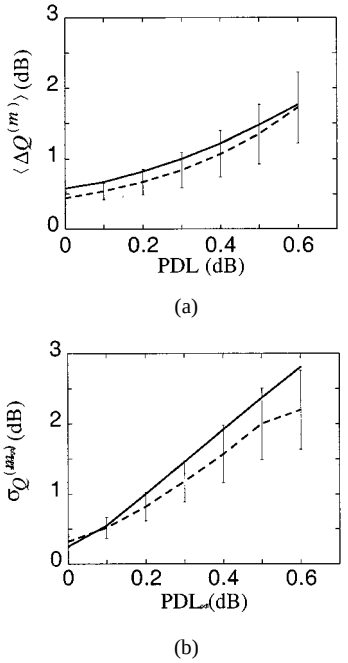


Fig. 5. Comparison of the signal degradation as a function of PDL in the Stokes model and in the full model, PMD $= 0.1$ ps/km$^{1/2}$, PDG $= 0.06$ dB. (a) $\langle \Delta Q^{(m)} \rangle$. (b) $\sigma_Q^{(m)}$. Solid lines indicate the Stokes model, dashed lines indicate the average of the full model, and the error bars indicate the standard deviation of the values for all eight channels.

values of PDL are not truly independent. First, we used the same bit pattern for all realizations at all values of PDL in order to avoid artificially enlarging σ_Q because of pattern dependences. Second, we reinitialized the random number generator for each new value of the PDL, so that we used the same 20 fiber realizations for each value. If we use different fiber realizations, then we find that the sign of the deviation between the full model and the Stokes model changes. For the RZ simulations that we present here, we found that the choice of fiber realizations was more significant than the pattern dependences in creating a systematic deviation between the full model and the Stokes model. We found the same result for NRZ simulations without polarization scrambling; however, we found that when we add polarization scrambling then the effect of pattern dependences becomes very significant.

Comparing Figs. 2 and 3, it is apparent that PDG adds a substantial penalty almost independent of the PDL. When the PDL is 0.6 dB but the PDG is 0.0 dB, $\langle \Delta Q \rangle$ is under 2 dB. By contrast, when the PDG is 0.06 dB, $\langle \Delta Q \rangle$ is consistently above 2 dB regardless of the PDL and almost reaches 4 dB when the PDL is 0.6 dB. However, σ_Q only increases slightly with nonzero PDG. We found similar results with the NRZ format without polarization scrambling; however, polarization scrambling substantially reduces the effect of the PDG, as expected [9].

B. Eight Channel Comparisons

We now discuss an eight channel system that uses the RZ format. We show the comparison between the full model and the reduced model in Fig. 4 when the PDG is 0 dB. The error bars on the dashed line show the standard deviation of the 8 channels. Comparison to Fig. 2 shows that the degradation $\langle \Delta Q^{(m)} \rangle$ is almost the same as with a single channel, but the $\sigma_Q^{(m)}$ are larger. With PDG included, we show the comparison in Fig. 5. In contrast to the single channel system considered in Section III-A, the effect of PDG is negligible. Again, we found similar results for NRZ systems [9].

IV. Applications

Having validated the Stokes model, we will now use it to calculate the outage probability in trans-oceanic systems assuming a system margin for polarization effects of either 2.5 dB or 3.0 dB. We used 10^5 fiber realizations for each choice of parameters, and, when necessary to compute the outage probability, we fitted a Gaussian distribution to the tail of the numerically-determined probability distribution function. To calculate the probability of failure per unit time, it is necessary to know the rate at which the polarization states of the transmission line change and become uncorrelated. This number is not well-known, but it has been estimated that an undersea system will pass through on the order of $10^5 - 10^6$ independent states in a 20-year lifetime [4]. Since the outage probability is a rapidly decreasing function of $\Delta Q^{(m)}$, there is little ambiguity in simply demanding that the outage probability be less than 10^{-6}.

The number of WDM channels in trans-oceanic systems has grown rapidly in recent years. While the effect of PMD on a single channel is typically small, the PMD does change the polarization states of the different channels with respect to one another. In other words, the PMD changes the angular separation of the channels on the Poincaré sphere. As a consequence of the interaction of the PMD and the PDL, different channels will undergo different amounts of loss when they pass through a device with PDL. Since the gain saturation in the amplifiers is

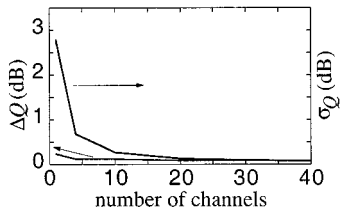


Fig. 6. The degradation and variance of the Q factor as a function of the number of channels.

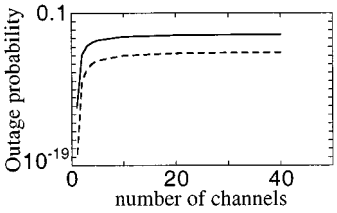


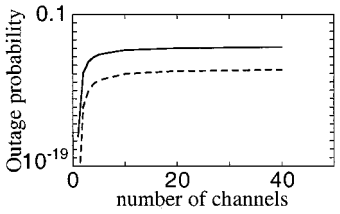
Fig. 7. Outage probability as a function of number of channels. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

tuned to effectively restore the total signal power in all the channels, some channels gain power at the expense of others. This effect leads to a random walk in the power of each channel and can cause one or more channels to fade. We will show that this mechanism is the primary cause of fading in systems with more than approximately ten channels, in contrast to single-channel systems in which PDG is the primary cause of fading.

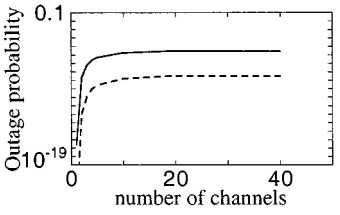
To investigate this issue, we considered a system in which the channel spacing and the optical filter bandwidth equaled 0.6 nm. We set the other system parameters as follows: PMD $= 0.1$ ps/km$^{1/2}$, PDL $= 0.0$ dB, and PDG $= 0.06$ dB. Fig. 6 shows that as the number of channels increases, the importance of PDG decreases as expected from the argument in the preceding paragraph.

Next, we set the PDG equal to zero, leaving only the effects of PMD and PDL in the model. In this case, we set the channel spacing to 1.0 nm and the optical filter bandwidth to 0.5 nm. We set PMD $= 0.1$ ps/km$^{1/2}$ and PDL $= 0.1$ dB. Increasing the number of channels, the result is shown in Fig. 7. We find that if $\Delta Q_{\text{allowed}}$, the allowed degradation level for any single channel, is set equal to 2.5 dB, then the outage probability dramatically increases from 6.5×10^{-13} in the case of a single channel to 3.0×10^{-4} when there are many channels. With only three channels, the outage probability already exceeds 10^{-5}. If we raise $\Delta Q_{\text{allowed}}$ to 3.0 dB, then the maximum outage probability falls to 2.3×10^{-6}, a decrease of more than two orders of magnitude.

When we increase the amplifier spacing from 33 to 45 km, and then 50 km, we find that the average value of Q decreases due to the additional ASE noise that is added to the total signal. However, the outage probability decreases because the number of PDL elements along the transmission line is reduced, as shown in Fig. 8. When the number of channels is 40, the outage probability drops from 3.0×10^{-4} to 1.3×10^{-5} and 2.8×10^{-6}, respectively. So, when one designs a WDM system and chooses the amplifier spacing, one has to take into



(a)



(b)

Fig. 8. Outage probability as a function of the number of channels. Amplifier spacing equals (a) 45 km, (b) 50 km. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

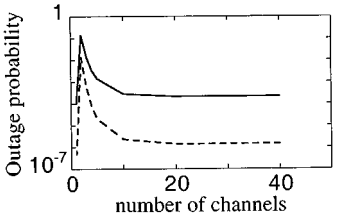


Fig. 9. Outage probability as a function of the number of channels, PDG $=$ 0.07 dB. Solid line is for the 2.5 dB decision level; dashed line is for the 3 dB decision level.

account both noise-induced and polarization-induced penalties. If the PDL is the same in each amplifier, then a short amplifier spacing will introduce less noise but will increase the outage probability. By contrast, a long amplifier spacing will introduce more noise but will decrease the outage probability.

We showed in Fig. 6 that the effect of PDG becomes insignificant when there are more than approximately ten channels in a WDM system. To further investigate this issue, we added a PDG of 0.07 dB to the case we showed in Fig. 7. We show these new results in Fig. 9. Instead of a small outage probability when the number of channels is small, we find that the outage probability becomes large for a small number of channels and then decreases to its final value. The dramatic increase in the outage probability due to PDG when the number of channels is small is due to the faster growth of ASE noise that is induced. The outage probability then decreases as the number of channels becomes larger because the PMD between the channels leads to an averaging of the polarization states so that the DOP for the *total* signal is nearly zero, and the PDG leads to almost no excess ASE growth. When the number of channels equals 40, the outage probability is 2.2×10^{-4} which is actually smaller than the corresponding value of 3.0×10^{-4} when there is no PDG. The reason for this paradoxical decrease in the outage probability is that the PDG tends to compensate for the effects of PDL on channels that experience excess loss.

V. Conclusion

We have proposed a Stokes parameter model to study the combined effects of PMD, PDL, and PDG in long-haul, dense WDM systems. We then presented validations of this reduced model by comparison to full simulations. More extensive numerical validations and an experimental validation are presented elsewhere [6], [9]. We next used this model to calculate the outage probabilities in trans-oceanic systems. We found that PDG plays a key role in single-channel systems, but its importance decreases as the number of channels increases and becomes negligible beyond ten channels. When PDG is not present, the outage probability increases rapidly with the number of channels and saturates beyond ten channels. By contrast, when PDG is present, the outage probability spikes at about three channels and then falls to its final asymptotic value. When the spacing between amplifiers increases, the penalty due to polarization effects decreases in WDM systems with more than about ten channels because the number of amplifiers in the transmission line is smaller, decreasing the effect of PDL.

Acknowledgment

The authors are grateful to F. Kerfoot and P. Runge for encouraging us to study this problem and arranging for its partial support at an early stage.

References

[1] E. Lichtman, "Limitations imposed by polarization-dependent gain and loss on all-optical ultralong communication systems," *J. Lightwave Technol.*, vol. 13, pp. 906–913, May 1995.

[2] C. D. Poole, R. W. Tkach, A. R. Chraplyvy, and D. A. Fishman, "Fading in lightwave systems due to polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 68–70, Jan. 1991.

[3] N. S. Bergano, "Wavelength division multiplexing in long-haul transmission systems," *J. Lightwave Technol.*, vol. 14, pp. 1299–1308, June 1996.

[4] C. R. Menyuk, D. Wang, and A. N. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 1247–1249, Sept. 1997.

[5] C. R. Menyuk, "Impairments due to nonlinearity and birefringence in optical fiber transmission systems," in *System Technologies*, A. E. Wilner and C. R. Menyuk, Eds. Washington, DC: Optical Society of America, 1997, vol. 12, Trends in Optics and Photonics, pp. 196–203.

[6] Y. Sun, D. Wang, P. Sinha, G. Carter, and C. R. Menyuk, "Polarization evolution of both signal and noise in a 100 km dispersion managed recirculating loop," IEEE Photon. Technol. Lett., submitted for publication.

[7] D. Wang and C. R. Menyuk, "Reduced model of the evolution of the polarization states in wavelength-division-multiplexed channels," *Opt. Lett.*, vol. 23, pp. 1677–1679, 1998.

[8] 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, pp. 1735–1746, Sept. 1997.

[9] D. Wang, "Polarization effects in dense WDM systems," Ph.D. dissertation, Univ. of Maryland Baltimore County, Mar. 2000.

[10] N. S. Bergano, C. R. Davidson, M. Ma, A. Pilipetskii, S. G. Evangelides, H. D. Kidorf, J. M. Darcie, E. Golovchenko, K. Rottwitt, P. C. Corbett, R. Menges, M. A. Mills, B. Pedersen, D. Peckham, A. A. Abramov, and A. M. Vengsarkar, "320 Gb/s WDM transmission (64×5 Gb/s) over 7,200 km using large mode fiber spans and chirped return-to-zero signals," . San Jose, CA, Feb. 1998, paper PD12-1.

[11] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 8, pp. 1816–1823, Dec. 1990.

[12] D. Wang and C. R. Menyuk, "Polarization evolution due to the Kerr nonlinearity and chromatic dispersion," *J. Lightwave Technol.*, vol. 17, pp. 2520–2529, Dec. 1999.

[13] G. P. Agrawal, *Fiber-Optic Communication Systems*. New York: Wiley, 1997.

D. Wang, photograph and biography not available at time of publication.

C. R. Menyuk (SM'88–F'98), photograph and biography not available at time of publication.