Polarization Evolution Due to the Kerr Nonlinearity and Chromatic Dispersion

D. Wang and C. R. Menyuk

Abstract— **This paper numerically investigated the evolution of the degree of polarization of individual channels and their Stokes parameters in a wavelength division multiplexed (WDM) system in which Kerr nonlinearity and chromatic dispersion are taken into account but in which polarization mode dispersion as well as polarization-dependent loss and gain are neglected. We compared the results to a mean field model which assumes that the channels are strongly dispersion-managed so that each channel is only affected by the Stokes parameters of the others. This model predicts no change in the degree of polarization of each of the channels so that an initially polarization-scrambled channel does not repolarize. The full simulations showed that the repolarization of a polarization-scrambled signal is small for parameters corresponding to realistic communication systems, validating the use of the mean field model. However, we also found that the repolarization can become significant for low data rates and a small number of channels in a dispersion-managed system with a short length map, thus setting limits on the model's validity and indicating operating regimes that should be avoided in real communication systems.**

Index Terms—**Chromatic dispersion, evolution, modeling, non-linearity, optical fiber communication, polarization, repolarization, Stokes parameters, wavelength division multiplexing (WDM).**

I. INTRODUCTION

POLARIZATION effects in optical fiber transmission systems are due to birefringence in both the optical fiber itself and the components like the WDM couplers that are used in current amplifier systems. Birefringence in the optical fiber is due to the accidental loss of degeneracy of the two orthogonal polarization modes that exist in a single-mode fiber and is weak and randomly varying. By contrast the materials like $LiNbO_3$ that are used in optical components often have a strong polarization dependence in which case these components are strongly birefringent along a fixed axis. Polarization effects in both optical components and optical fiber have become important in recent years because the advent of the erbium-doped fiber amplifiers (EDFA's) implies that these effects accumulate over hundreds of kilometers in terrestrial systems and thousands of kilometers in undersea systems.

It is generally assumed that the principal polarization effects that lead to transmission impairments are polarization-dependent loss, polarization-dependent gain, and polarization mode dispersion (PMD) [1]. Polarization-dependent loss is caused by the strong polarization dependence in optical components such as WDM couplers, isolators, and optical switches. Polarization-dependent gain is caused by polarization hole-burning in the Er-doped fiber amplifiers. Finally, polarization mode dispersion is principally caused by the randomly varying birefringence in the optical fiber, although the differential group delay of the optical components can contribute under some circumstances. The combination of these effects contribute to channel fading [2] which may sometimes lead to the loss of an entire channel in a long-distance, high-data-rate wavelength division multiplexed (WDM) system. Polarization scrambling greatly ameliorates the effects of polarization impairments [3], but under some circumstances the signal can repolarize [4], [5], leading once more to polarization effects.

These effects couple in turn to the nonlinearity and dispersion in a complex way that is described by the Manakov-PMD equation [6], [7]. There is a general misconception that initially polarization-scrambled channels cannot repolarize unless some combination of polarization mode dispersion, polarization-dependent loss, and polarization-dependent gain are present. However, nothing could be farther from the truth! As we shall describe in detail in this article, nonlinearity and dispersion all by themselves can led to channel interactions that in turn cause significant repolarization when the channels are closely spaced. These interactions set limits on how closely channels can be spaced and other system parameters. These limits are of increasing importance in modern-day WDM systems.

It is now commonplace to study fiber system impairments using the coupled nonlinear Schrödinger equation that has been modified to include loss, amplification, spontaneous emission noise [8], and other effects. When polarization effects, including polarization mode dispersion, can be neglected and the signal is launched in a single state of polarization, then it is possible to study optical fiber impairments using the scalar nonlinear Schrödinger equation and its modifications [6], [7]. In fact, this scalar approach is far more commonly used than the complete vector equation [8].

There is a misconception that the nonlinear Schrödinger equation is more fundamental than the Manakov equation in which polarization mode dispersion is neglected [5]. In fact, just the opposite is true since polarization mode dispersion must be negligible and in addition the signal must be launched in a single polarization state for the nonlinear Schrödinger equation to hold. Polarization mode dispersion is a significant issue in terrestrial systems with optical fiber that is more

than a decade old, but it plays a nearly negligible role in undersea systems or terrestrial systems that use fiber that has been laid within the last five years. For example, if we assume a value for the polarization mode dispersion of 0.1 ps/(km)$^{1/2}$, then over 10 000 km, the differential group delay is only 10 ps. This value is not large enough to significantly distort individual channels, although it does lead to some relative drift of the polarization states of channels that are spaced far apart [6], [7]. We will show explicitly, later in this paper, that the effect of this additional polarization mode dispersion on signal repolarization is small. Modern-day undersea systems also use polarization scrambling or orthogonal polarization when launching their signals. In these systems, the Manakov equation is the correct equation to use to study of nonlinear and dispersive effects. Solving either the scalar nonlinear Schrödinger equation or vector equations such as the Manakov equation or the Manakov-PMD equation can be very computationally time-consuming—particularly when WDM systems that include more than ten channels are studied. A full simulation of a terrestrial system can take many central processor unit (CPU) hours for a single parameter set and a full simulation of an undersea system can actually take many CPU days [9]! To appropriately study polarization effects, one must additionally simulate many different realizations of the randomly varying polarization orientations in the optical fiber. Thus, it is not realistic to study the impairments introduced by polarization effects in modern-day WDM systems that may have many tens of channels using complete numerical solutions of the Manakov equation and its modifications. There is thus an urgent need for effective reduced models.

In previous work, we introduced a mean field approach in which one merely follows the Stokes parameters of the individual channels neglecting their detailed temporal evolution [5]. The potential computational saving in this approach is vast. We note that this approach presupposes that polarization effects like repolarization are dominated by the interplay of polarization mode dispersion, polarization-dependent loss, and polarization-dependent gain and that nonlinearity and dispersion play no significant role in these effects. This assumption is reasonable in modern-day dispersion-managed systems in which the local dispersion is large, although the average dispersion is kept small. So, the bits in two channels slide rapidly through one another, and each channel is only affected by the average properties of the others. However, this assumption will fail when channels are spaced too closely together. At this point nonlinearity and dispersion can lead to significant polarization impairment. Given the interest in spacing channels ever closer together in dense WDM system, it is of practical importance to determine this limit.

It is the purpose of this article to explore the conditions under which the mean field approach is valid for studying polarization effects. We will be focusing on the extent to which the combination of chromatic dispersion and the Kerr effect can induce changes in the polarization states of the individual channels beyond what the mean field model predicts. Our procedure will be to compare full split-step simulations with realistic temporal pulse profiles and with up to seven channels to the mean field model. Our purposes are threefold. First,

we will determine when the system parameter are such that nonlinearity and dispersion alone lead to significant polarization effects. These effects limit the channel spacing and other parameters in real systems. Second, we will set a baseline for further studies in which neglected effects like polarization mode dispersion, polarization-dependent loss, polarization-dependent gain and amplified spontaneous emission noise are included. Given the complex interplay between the different effects, it is important to study them in isolation before combining them. We will however determine the effect of adding moderate polarization in this paper. Third, we will determine when the mean field approach can be safely used to study polarization effects. This approach yields vast computational savings [5].

The remainder of this paper is organized as follows: Section II reviews the derivation of the mean field equations governing the evolution of the Stokes parameters. We add to previous work [5] by taking into account the effect of fiber attenuation and amplifier gain. In Section III, we compare the results of the mean field model to full split-step simulations for a typical set of parameters and up to seven channels. In Section IV, we examine the effects of varying the channel spacing, varying the dispersion, and adding filtering. Section V contains the conclusions.

II. DERIVATION OF THE MEAN FIELD EQUATION

Our starting point is the Manakov equation

$$i\frac{\partial \mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \gamma(\mathbf{U}^\dagger \mathbf{U})\mathbf{U} = 0 \tag{1}$$

where $\mathbf{U} = (u_x, u_y)^t$ represents the complex envelope of the two polarizations, where β'' is the dispersion coefficient, γ is the nonlinear coefficient, and z and t are distance along the fiber and retarded time. Earlier experimental [9] and theoretical [6], [7] work has shown that this equation accurately describes nonlinear and dispersive light propagation in standard communication fiber with rapidly and randomly varying birefringence when polarization mode dispersion can be neglected. There are several reasons for neglecting polarization mode dispersion in most of this study. First, with high-quality optical fibers and a limited number of channels, polarization mode dispersion is negligible. For example, assuming a value of 0.1 ps/(km)$^{1/2}$, polarization mode dispersion can be neglected over a bandwidth of about 4 nm over 500 km and about 1 nm over 10 000 km. These numbers are sufficient to include several channels with single channel rates of 2.5 Gb/s. Even at higher data rates and with channels spread over a larger bandwidth, the effect is small, as we will show explicitly in Section III, as long as the polarization mode dispersion is moderate. Second, it is of practical importance to determine when nonlinearity and chromatic dispersion lead to significant polarization effects on their own. Third, it is of critical importance in validating the mean field approach to determine its ability to accurately reproduce the impact of the Kerr effect and chromatic dispersion on the polarization states of the channels in a WDM systems in the absence of other

effects. Thus, this work will serve as a baseline against which the importance of polarization mode dispersion and other polarization effects in combination can be determined. We also note that (1) does not include the spatially varying gain and loss that is present in real systems; however, this effect is easily included and we will show how to do that at the end of this section. Finally, we note that $\gamma = (8/9)(k_0 n_2/A_{\text{eff}})$, where k_0 is the light's central wavenumber, n_2 is the Kerr coefficient, and A_{eff} is the optical fiber's effective area. The extra factor of 8/9 appears due to rapid mixing of the central frequency's polarization state on the Poincaré sphere [6], [7]. In principle, there can be an additional contribution to (1) due to incomplete mixing of the electric field on the Poincaré sphere that is referred to as nonlinear polarization mode dispersion [6], [7]. However, in standard communication fiber in which the scale length on which the field mixes on the Poincaré sphere is 100 m or less while the scale on which the Kerr nonlinearity and chromatic dispersion act is many hundreds of kilometers, this effect is completely negligible [7].

We next write $\mathbf{U}$ as a sum of contributions over n channels, obtaining

$$\mathbf{U} = \sum_{m=1}^{n} \mathbf{U}^{(m)} \exp\left[ik^{(m)}z - i\omega^{(m)}t\right] \qquad (2)$$

where $k^{(m)}$ and $\omega^{(m)}$ are the central wavenumber and frequency of the m-th channel with respect to the central wavenumber and frequency of $\mathbf{U}$, and $\mathbf{U}^{(m)}$ is the corresponding wave envelope. Substituting (2) into (1), we find

$$i\frac{\partial \mathbf{U}^{(m)}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}^{(m)}}{\partial t^2} + \gamma\left(\mathbf{U}^{\dagger(m)}\mathbf{U}^{(m)}\right)\mathbf{U}^{(m)}$$
$$+ \gamma \sum_{q=1,\neq m}^{n} \left(\mathbf{U}^{\dagger(q)}\mathbf{U}^{(q)}\right)\mathbf{U}^{(m)}$$
$$+ \gamma \sum_{q=1,\neq m}^{n} \left(\mathbf{U}^{\dagger(q)}\mathbf{U}^{(m)}\right)\mathbf{U}^{(q)} = 0 \qquad (3)$$

where, consistent with our assumption that the dispersion between channels is large, we neglect four wave mixing terms. We now define the Stokes parameters for each of the channels as

$$S_0^{(m)} = \frac{1}{T}\int_{t_1}^{t_2} \left[\left|u_x^{(m)}(t)\right|^2 + \left|u_y^{(m)}\right|^2\right] dt \qquad (4a)$$

$$S_1^{(m)} = \frac{1}{T}\int_{t_1}^{t_2} \left[\left|u_x^{(m)}(t)\right|^2 - \left|u_y^{(m)}\right|^2\right] dt \qquad (4b)$$

$$S_2^{(m)} = \frac{2}{T}\int_{t_1}^{t_2} \text{Re}\left[u_x^{(m)}(t)u_y^{(m)^*}(t)\right] dt \qquad (4c)$$

$$S_3^{(m)} = \frac{2}{T}\int_{t_1}^{t_2} \text{Im}\left[u_x^{(m)}(t)u_y^{(m)^*}(t)\right] dt \qquad (4d)$$

where we assume that $T = t_2 - t_1$ contains a very large number of individual bits. Using (3) to determine the evolution of the Stokes parameters, we find $dS_0^{(m)}/dz = 0$, and we find for $dS_1^{(m)}/dz$ that

$$\frac{dS_1^{(m)}}{dz} = i\frac{\gamma}{T}\int_{t_1}^{t_2} \left[\left(u_x^{(m)}u_y^{(m)^*} + u_x^{(m)^*}u_y^{(m)}\right)\right.$$
$$\times \sum_{q=1,\neq m}^{n} \left(u_x^{(q)}u_y^{(q)^*} - u_x^{(q)^*}u_y^{(q)}\right)$$
$$- \left(u_x^{(m)}u_y^{(m)^*} - u_x^{(m)^*}u_y^{(m)}\right)$$
$$\left.\times \sum_{q=1,\neq m}^{n} \left(u_x^{(q)}u_y^{(q)^*} + u_x^{(q)^*}u_y^{(q)}\right)\right] dt. \qquad (5)$$

In a highly dispersive system, the channels for which $q \neq m$ rapidly pass through the m-th channel in the time domain. Consequently, the evolution of the m-th channel is only affected by the averaged time variation of the $q \neq m$ channels so that we may effectively treat them as continuous waves. This assumption is the heart of the mean field approximation. We thus replace

$$u_x^{(q)}u_y^{(q)^*} - u_x^{(q)^*}u_y^{(q)} \rightarrow \frac{1}{T}\int_{t_1}^{t_2} \left(u_x^{(q)}u_y^{(q)^*} - u_x^{(q)^*}u_y^{(q)}\right) dt \quad (6)$$

from which we conclude

$$\frac{dS_1^{(m)}}{dz} = \gamma \sum_{q=1}^{n} \left(S_2^{(m)}S_3^{(q)} - S_3^{(m)}S_2^{(q)}\right). \qquad (7)$$

We can find similar expressions for $dS_2^{(m)}/dz$ and $dS_3^{(m)}/dz$, so that we finally obtain

$$\frac{d\mathbf{S}^{(m)}}{dz} = \gamma \mathbf{S}^{(m)} \times \sum_{q=1}^{n} \mathbf{S}^{(q)}. \qquad (8)$$

The effect of dispersion does not appear in (8); only the nonlinearity appears, and the equations are analogous to the equations that govern nonlinear polarization rotation of continuous wave beams [10]. However, the local dispersion plays a critical role since it must be large enough so that each channel only affects its neighbors through its mean field. We note as well that there is no change in $S_0^{(m)}$, i.e., $dS_0^{(m)}/dz = 0$. From (8), it follows that the only effect of combining the Kerr effect with large dispersion is to rotate the polarization states of the different WDM channels. In particular, there is no change in the degree of polarization! This result is significant because polarization scrambling which reduces the degree of polarization of a single channel to a small value is often used to combat fading. However, under some circumstances the signal can repolarize. We earlier showed that polarization-dependent loss can account for this repolarization [4], but it is important to verify that chromatic dispersion and the Kerr effect cannot also account for it. If the mean field approximation is valid, i.e., (6) is an accurate approximation, then these effects lead to no repolarization. Conversely, if chromatic dispersion and the Kerr effect do contribute significantly to repolarization, which can happen even in the absence of polarization mode dispersion and other polarization effects when the mean field approximation fails, then there is a large nonlinear interaction between the channels which is unattractive for communication

systems. Thus, examining the evolution of the degree of polarization in full simulations will allow us to determine the parameter regimes in which the mean field approximation is valid and in which chromatic dispersion and the Kerr effect do not contribute significantly to repolarization. We will do that in the following sections.

It is interesting to note that even though (8) is nonlinear, a complete analytical solution may be found

$$
\mathbf{S}^{(m)} = a_1 \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + a_2 \begin{pmatrix} r \\ s \\ t \end{pmatrix} e^{i\gamma cz} + a_3 \begin{pmatrix} r^* \\ s^* \\ t \end{pmatrix} e^{-i\gamma cz},
$$

$$
r = c_1 c_3 + i c c_2, \quad s = c_2 c_3 - i c c_1, \quad t = c_3^2 - c^2,
$$

$$
c_1 = \sum_{q=1}^{n} S_{10}^{(q)}, \quad c_2 = \sum_{q=1}^{n} S_{20}^{(q)}, \quad c_3 = \sum_{q=1}^{n} S_{30}^{(q)},
$$

$$
c^2 = c_1^2 + c_2^2 + c_3^2, \quad a_1 = D_1/\Delta, \quad a_2 = D_2/\Delta,
$$

$$
a_3 = D_3/\Delta, \tag{9}
$$

$$
\Delta = \begin{vmatrix} c_1 & r & r^* \\ c_2 & s & t \\ c_3 & t & t \end{vmatrix}, \quad D_1 = \begin{vmatrix} S_{10}^{(m)} & r & r^* \\ S_{20}^{(m)} & s & s^* \\ S_{30}^{(m)} & t & t \end{vmatrix},
$$

$$
D_2 = \begin{vmatrix} c_1 & S_{10}^{(m)} & r^* \\ c_2 & S_{20}^{(m)} & s^* \\ c_3 & S_{30}^{(m)} & t \end{vmatrix}, \quad D_3 = \begin{vmatrix} c_1 & r & S_{10}^{(m)} \\ c_2 & s & S_{20}^{(m)} \\ c_3 & t & S_{30}^{(m)} \end{vmatrix}
$$

where $S_{10}^{(q)}, S_{20}^{(q)}, S_{30}^{(q)}, S_{10}^{(m)}, S_{20}^{(m)}$, and $S_{30}^{(m)}$ are the initial values of the Stokes vector in the channels. This result is intrinsically significant because the number of large-dimensional nonlinear systems for which exact solutions can be found is small.

When fiber loss and lumped gain at the amplifier are introduced into the transmission line, (1) becomes

$$
i\frac{\partial \mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \gamma(\mathbf{U}^\dagger \mathbf{U})\mathbf{U}
$$

$$
= -i\alpha \mathbf{U} + ig \sum_{p=1}^{N_{\text{amp}}} \delta(z - pL_a)\mathbf{U} \tag{10}
$$

where α is the loss coefficient, g is the amplifier gain, L_a is the amplifier spacing, and N_{amp} is the total number of amplifiers. Since the loss and gain occur periodically in real systems on a length scale that is short compared to the scale on which the nonlinearity and average chromatic dispersion operate, they will induce periodic oscillations in the amplitude without changing the state of polarization. Since the Stokes vector components also periodically oscillate as a result, $\mathbf{S}^{(q)}$ in (8) will be smaller than its initial value right after an amplifier during most of the propagation, and $\mathbf{S}^{(m)}$ will evolve more slowly than if $\mathbf{S}^{(q)}$ always had this initial value. Defining $\mathbf{U}(z,t) = \bar{\mathbf{U}}(z,t)\exp[\alpha(z - pL_a)]$ in the interval $pL_a < z < (p+1)L_a$, so that we normalize the amplitude throughout the interval between amplifiers to its value at the beginning of the interval, and defining $\bar{\mathbf{S}}^{(m)}(z)$ from $\bar{\mathbf{U}}(z,t)$ analogous to (4), we find that

$$
\frac{d\bar{\mathbf{S}}^{(m)}}{dz} = \gamma \bar{\mathbf{S}}^{(m)} \times \sum_{q=1}^{n} \bar{\mathbf{S}}^{(q)} \exp[-2\alpha(z - pL_a)]. \tag{11}
$$

Defining now, $\bar{z} = \bar{z}_p + \int_{pL_a}^{z} \exp[-2\alpha(z' - pL_a)]\,dz' = \bar{z}_p + (1/2\alpha)[1 - \exp[-2\alpha(z - pL_a)]]$ where $\bar{z}_p$ is a constant, we conclude

$$
\frac{d\bar{\mathbf{S}}^{(m)}(\bar{z})}{d\bar{z}} = \gamma \bar{\mathbf{S}}^{(m)}(\bar{z}) \times \sum_{q=1}^{n} \mathbf{S}^{(q)}(\bar{z}) \tag{12}
$$

which has the same form as (8). If we let $\bar{z}_p = (p/2\alpha)[1 - \exp(-2\alpha L_a)] = pL_{\text{eff}}$ then $\bar{z}(z)$ varies continuously, and (12) can be interpreted as an evolution equation-dependent upon this new variable. The variation of $\bar{z}(z)$ is somewhat complicated by the exponential variation, but in most realistic settings, the nonlinear evolution is slow compared to the periodic evolution due to the fiber loss and lumped gain so that it is sufficient to approximate (12)

$$
\frac{d\bar{\mathbf{S}}^{(m)}(\bar{z})}{dz} = \frac{L_{\text{eff}}}{L_a}\gamma \bar{\mathbf{S}}^{(m)}(z) \times \sum_{q=1}^{n} \bar{\mathbf{S}}^{(q)}(z). \tag{13}
$$

We will validate this approximation in Section IV.

When we include the effect of polarization mode dispersion, (1) becomes

$$
i\frac{\partial \mathbf{U}}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \gamma(\mathbf{U}^\dagger \mathbf{U})\mathbf{U} = -i\Delta\beta'\bar{\sigma}\frac{\partial \mathbf{U}}{\partial t} \tag{14}
$$

where $\Delta\beta'$ indicate the inverse group velocity difference between the fast and slow axes while $\bar{\sigma}$ is a rapidly varying unitary matrix that takes into account the changes in orientation of the birefringent axes [8]. We solve this equation using the coarse step method [8].

III. BASIC NUMERICAL MODEL

In our simulations, we used standard split-step methods to solve (1). In each channel, we used synchronous phase modulation and in some cases synchronous amplitude modulation of nonreturn-to-zero (NRZ) signals, much as described by Bergano *et al.* [3], to both polarization scramble the signals and minimize pulse distortion. The functional form that we used for the initial field at the entry to the fiber is

$$
u_x^{(m)}(t) = A_x(t)\exp[i\phi_x(t)] \tag{15a}
$$

$$
u_y^{(m)}(t) = A_y(t)\exp[i\phi_y(t)] \tag{15b}
$$

when the signal has no amplitude modulation and

$$
u_x^{(m)}(t) = A_x(t)\exp[i\phi_x(t)]\cos(\Omega t/2 + \pi/2) \tag{16a}
$$

$$
u_y^{(m)}(t) = A_y(t)\exp[i\phi_y(t)]\cos(\Omega t/2 + \pi/2) \tag{16b}
$$

when the signal is amplitude-modulated, where $\phi_x(t) = \delta_x + a_x \cos(\Omega t + \psi_x + \pi/2)$ and $\phi_y(t) = \delta_y + a_y \cos(\Omega t + \psi_y + \pi/2)$. Here, we let $A_x(t) = c_x H(t)$ and $A_y(t) = c_y H(t)$, where c_x and c_y are constant coefficients, while $H(t) = 1$ in the time slots of the marks and $H(t) = 0$ in the time slots of the spaces except when making a transition from a space to a mark or from a mark to a space. When making a transition from a space to a mark, we set $H(t) = \{1 + \tanh[(t - t_0)/t_{\text{rise}}]\}/2$, where t_0 is the boundary between the space and the mark and $t_{\text{rise}} = 10$ ps for a 5 Gb/s signal and $t_{\text{rise}} = 5$ ps for a 10 Gb/s signal. When making a transition from a mark to a space, we

define $H(t) = \{1 + \tanh[-(t - t_0)/t_{\text{rise}}]\}/2$. The coefficients a_x and a_y give the strength of the phase modulation. For all the simulations in this paper, we set $a_x = 3.307$ and $a_y = 0.903$ so that the difference $a_x - a_y$ nearly equals to 2.405, the first zero of the zeroth Bessel function. With this choice, setting $c_x = c_y$, an ideal square pulse would be completely depolarized. In our simulations, since the pulses are never perfectly square, even without amplitude modulation, there is always a small residual degree of polarization like in the experiments of Bergano et al. The sum $a_x + a_y$ was chosen consistent with the experiments of Bergano et al. [3]. The phase modulation frequency Ω corresponds to the bit rate, and ψ_x and ψ_y describe the relative phase between the phase modulation and the data bits. By varying ψ_x, ψ_y, c_x, and c_y, we can adjust the degree of polarization to any desired value. We used a 64 bit pattern in each channel, chosen so that the number of marks and spaces is identical and so that strings of marks of varying sizes are in the pattern. We experimented with a number of different strings in several cases and verified that our results are insensitive to this choice. We used 8192 node points for each channel in all cases which experimentation showed was adequate.

We chose the system parameters as follows: the central wavelength $\lambda = 1.58$ μm, the Kerr coefficient $n_2 = 2.6 \times 10^{-20}$ m^2/W, and a total fiber length $L = 10\,000$ km. For our basic simulations, we used a dispersion map with a map length of 1000 km like that of Bergano et al. [3], but we also investigated the effect of shortening the map, and we will report these results in Section IV. In every map period, there is a span of length L_1 that consists of standard fiber with dispersion value $D_1(\lambda)$ and a span of length L_2 that consists of dispersion-shifted fiber with dispersion value $D_2(\lambda)$. We chose the lengths of these two portions so that $L_1 : L_2 = 1 : 8.5$. At the central frequency $\lambda_0 = 1.58$ μm, we choose the corresponding dispersion values so that they are in the ratio $8.5 : -1$, and the net cumulative dispersion is zero. Typical experimental systems that model undersea systems would have a value of $D_1 = -2$ ps/nm-km and $D_2 = 17$ ps/nm-km, but we often varied these values significantly in our simulations to compare our results to the mean field approach. We note that values in terrestrial systems are often substantially higher. When $\lambda \neq \lambda_0$, third-order dispersion implies that the net dispersion is nonzero. In our simulations, we used $dD/d\lambda = 0.07$ ps/nm^2-km. In our basic simulations, we did not use amplitude modulation, and the signaling rate is 5 Gb/s, corresponding to a bit period of 200 ps. The channel spacing is 0.5 nm.

In Fig. 1, we show a two-channel simulation. We note that because $a_x \neq a_y$, and ψ_x and ψ_y are different for the two channels, the two channels have slightly different values of S_2 and S_3 and hence slightly different degrees of polarization. The degree-of-polarization is nonzero in this case because $A_x \neq A_y$. Fig. 1(a) shows the theoretically predicted result from (9), and Fig. 1(b) shows the evolution of the Stokes parameters with standard values of the dispersion. There are significant quantitative discrepancies between the mean field theory and the simulation, and these discrepancies only completely disappear when the local dispersion becomes quite large as shown in Fig. 1(c). Nonetheless, the Stokes
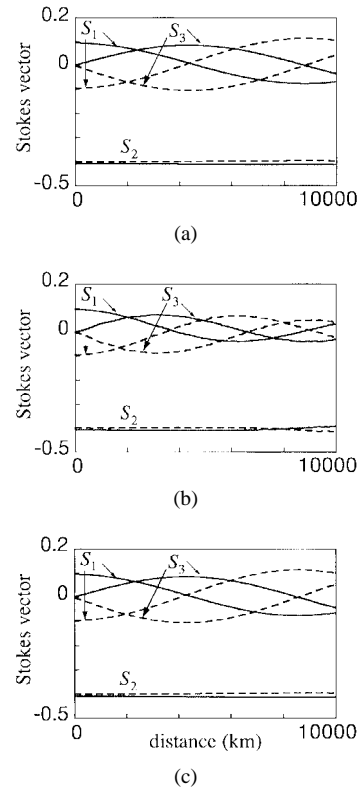

Fig. 1. Evolution of the Stokes vector components as a function of distance in a 5-Gb/s system. The dispersion map length is 1000 km, and the channel spacing is 0.5 nm. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. Other simulation parameters are $\lambda = 1550$ nm for channel one, $\lambda = 1550.5$ nm for channel two; $\psi_x = 0.7\pi$ and $\psi_y = 0$ for channel one, $\psi_x = 0$ and $\psi_y = 0.7\pi$ for channel two; the peak power in the x-polarization is 0.24 mW for channel one and is 0.2 mW for channel two; the peak power in the y-polarization is 0.2 mW for channel one and 0.24 mW for channel two.
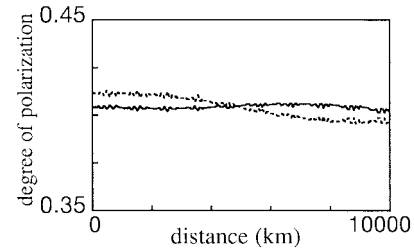

Fig. 2. Evolution of the degree of polarization as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. Other parameters are the same as in Fig. 1.

parameters still oscillate around their initial values in Fig. 1(b), although with somewhat different frequencies and amplitudes than in Fig. 1(a). There are no long-term drifts in the Stokes parameters from the theoretically-predicted values. Thus, we would anticipate that there is little change in the degree of polarization of the channels, and this expectation is borne out as shown in Fig. 2 where we show the degree of polarization, $[S_1^2 + S_2^2 + S_3^2]^{1/2}/S_0$, for each channel over 10 000 km. The change is only about 0.02. In particular, if the degree-of-polarization is initially near zero for both channels, which we obtain by setting $A_x = A_y$ in this case, then they undergo little repolarization as shown in Fig. 3.
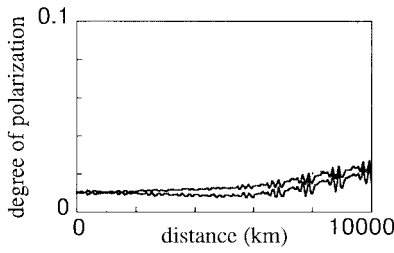
Fig. 3. Parameters are the same as in Fig. 1 except that the peak power in both the x- and y-polarization is 0.2 mW for both channels.
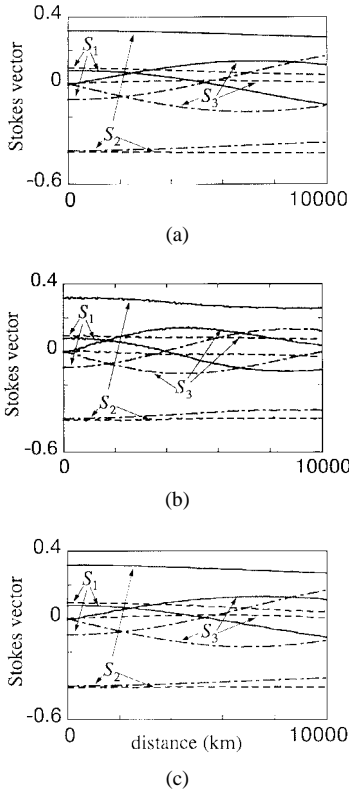

Fig. 4. Evolution of the Stokes vector components as a function of distance in a 5-Gb/s system. The dispersion map is 1000 km and the channel spacing is 0.5 nm. The solid lines are the Stokes components of channel one, the dashed lines are the Stokes components of channel two, and the dash-dotted lines are the Stokes parameters of channel one. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. Other simulation parameters are $\psi_x = 0.7\pi$, $\psi_y = 0$ for channel one; $\psi_x = 0$, $\psi_y = 0.7\pi$ for channel two; $\psi_x = 0.4\pi$, $\psi_y = 0$ in channel three; the peak power in the x-polarization is 0.24 mW for channel one, 0.2 mW for channel two, and 0.23 mW for channel three; the peak power in the y-polarization is 0.2 mW for channel one, 0.24 mW for channel two, and 0.23 mW for channel three. Other parameters are the same as in Fig. 1.

As the number of channel increases, the agreement between the mean field theory and simulation improves somewhat because the presence of multiple channels leads to better averaging over the different channels. In Fig. 4, we show the evolution with three channels. The oscillation periods of the Stokes parameters are reduced with respect to the case with two channels because we kept the power in each channel nearly the same as in the two-channel case. In Fig. 5, we show the evolution of the Stokes parameter S_1 with seven channels. We only show one of the Stokes parameters because the figure
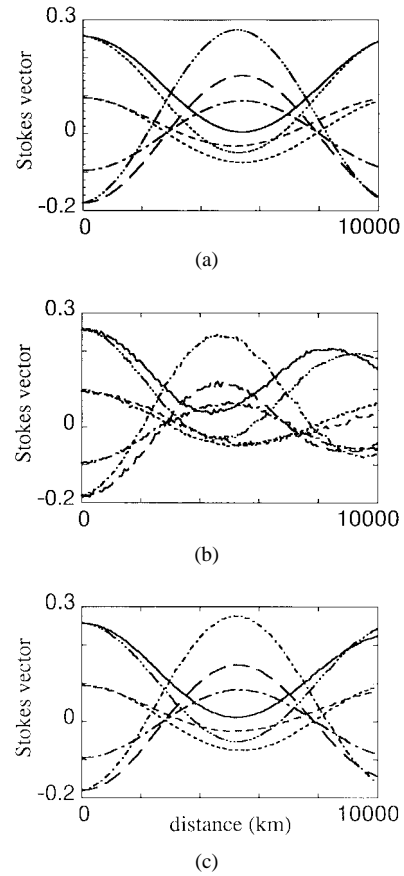

Fig. 5. Evolution of the Stokes vector component S_1 of seven channels as a function of distance in a 5-Gb/s system. The dispersion map length is 1000 km and the channel spacing is 0.5 nm. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km. The channels are centered around $\lambda = 1550\,\mu$m. Other parameters are: $\psi_x = 0$, $\psi_y = 0$ for all channels.

would be too busy if we included all three. Finally, in Fig. 6, we show the evolution of the degree of polarization of all seven channels. The worst channel changes by less than 0.03.

In Fig. 6, we also show the effect of adding a moderate polarization mode dispersion of 0.1 ps/km$^{1/2}$. We observe a slight increase in the repolarization in all of the channels. This increase occurs because even moderate polarization mode dispersion changes the polarization state slightly from what it would have been across the spectrum of the channel. As a consequence, the Stokes vectors are slightly shifted as a function of frequency from their values in the absence of polarization mode dispersion. We note that this increased repolarization is too small to affect our conclusions when repolarization due purely to nonlinearity and chromatic dispersion becomes too large to be tolerated, indicating an unacceptable parameter regime.

IV. EFFECTS OF PARAMETER VARIATION

A. Increasing the Data Rate

In Fig. 7, we show the evolution of a two-channel system when the signaling rate is increased to 10 Gb/s, corresponding to a bit period of 100 ps. In this case, the channel spacing is 1 nm. We see that the agreement between the mean field theory
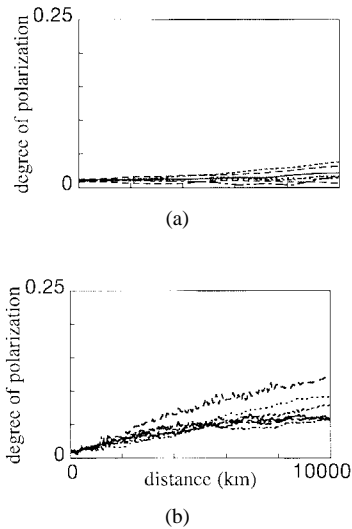
Fig. 6. Evolution of the degree of polarization as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The peak power in the x- and y-polarization is 0.2 mW for all channels. (a) PMD $= 0$ ps/km$^{1/2}$. (b) PMD $= 0.1$ ps/km$^{1/2}$.
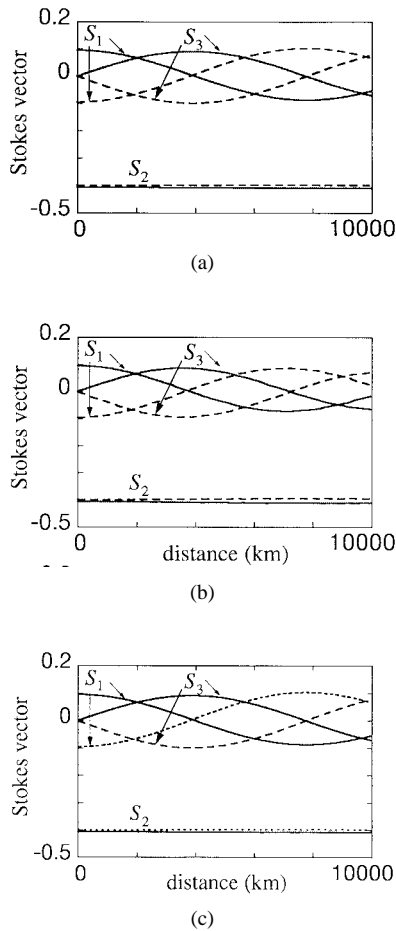

Fig. 7. Evolution of the Stokes vector components as a function of distance for a 10-Gb/s system. The dispersion map length is 1000 km, and the channel spacing is 1 nm. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km. Channel one is at $\lambda = 1550 \,\mu$m, and channel two is at $\lambda = 1551 \,\mu$m. Other parameters are the same as in Fig. 1.
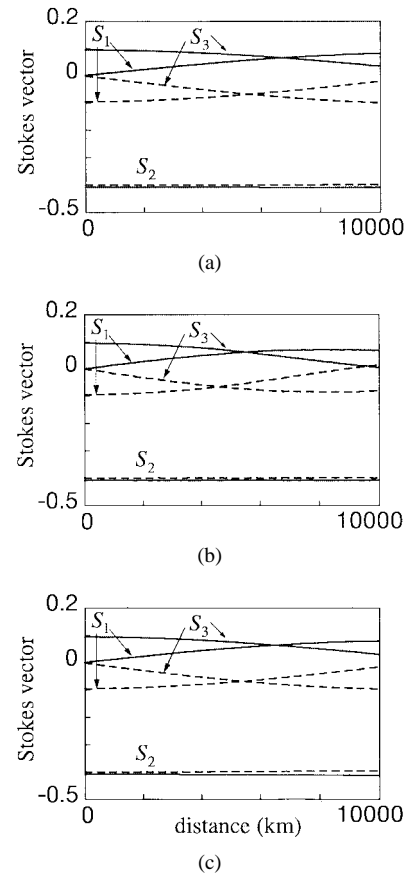

Fig. 8. Evolution of the Stokes vector components as a function of distance with gain and loss added. The system parameters are the same as in Fig. 1.

and simulations at a typical dispersion value is better at this data rate than it was at 5 Gbits/sec. Moreover, at dispersion values of $D_1 = -10$ ps/nm-km, $D_2 = 85$ ps/nm-km, the agreement is already complete. The wider channel spacing required to avoid intersymbol interference and the smaller bit period lead to better averaging. As before, when we increase the number of channels, we find that agreement improves.

B. Adding Gain and Loss

We add gain and loss to our basic model by assuming a loss coefficient $\alpha = 0.2$ db/km and an amplifier spacing of 50 km. As shown in Fig. 8, the results are essentially identical to Fig. 1, except that the oscillations are stretched out by a factor L_a/L_{eff} as predicted by (13). We have left out the bars above the Stokes vector in Fig. 8.

C. Adding Amplitude Modulation

We now add amplitude modulation to our basic model as shown in (15). The results are shown in Fig. 9. The behavior is similar to that shown in Fig. 1. At realistic values of the dispersion, shown in Fig. 9(b) the mean field theory differs quantitatively from the actual simulation, but shows no large drifts, and when the dispersion is 20 times larger, as shown in Fig. 9(c), the results become identical. We conclude that the detailed pulse shape is not important in the evolution of the polarization states.
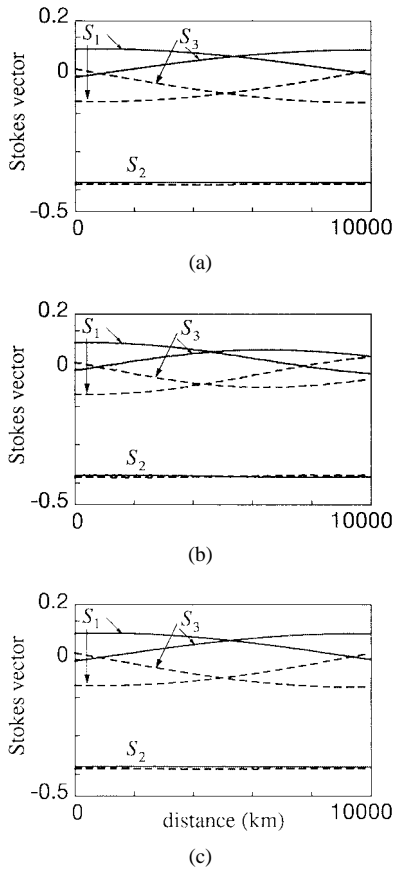
Fig. 9. Evolution of the Stokes vector components as a function of distance with amplitude modulation added. The system parameters are the same as in Fig. 1. The dispersion map is 1000 km, and the channel spacing is 0.5 nm. The amplitude modulation is added. The solid lines are the Stokes components of channel one; the dashed lines are the Stokes components of channel two. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. The other parameters are the same as Fig. 1.

D. Changing Channel Spacing

Increasing the channel spacing while keeping the dispersion map fixed increases the relative velocity of the channels and is nearly equivalent to increasing the magnitude of the local dispersion values by the same factor. In Fig. 10, we show the result of setting the channel spacing to 2.5 nm. Agreement between the theory and the simulation is almost complete.

However, when we reduce the channel spacing we find large changes in the polarization evolution due to the nonlinear interaction between channels. Phase modulation expands the spectral width of each channel, leading to channel crosstalk. Simulations indicate that the channel spacing is limited to 0.3 nm before these effects become intolerably large. From a purely theoretical standpoint, these results imply that the mean field approach can only be used for channel spacing of 0.3 nm or more.

E. Reducing the Map Length

In this case, the averaging becomes worse because the channels move back and forth through each other with a smaller period. Consequently, the agreement between the mean field theory and simulation deteriorates as shown in Fig. 11. While there is improvement as the number of channels
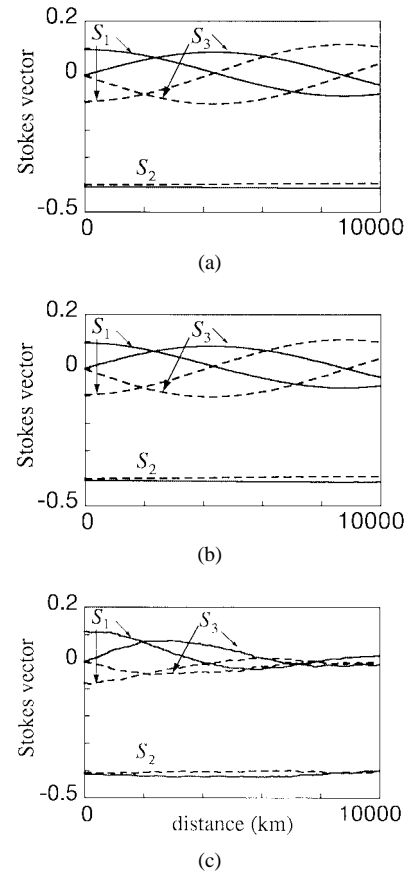

Fig. 10. Evolution of the Stokes vector components as a function of distance with variation of the channel spacing. (a) Analytical result. (b) Simulation result. Channel one is at $\lambda = 1550$ nm, and channel two is at $\lambda = 1552.5$ nm. (c) The channel spacing is 0.3 nm. Other parameters are the same as in Fig. 1.

increases, significant differences remain with up to seven channels. In Fig. 12, we show the degree of polarization for each channel in a seven channel system. We see that there is a nonnegligible repolarization in some cases, amounting to a change in the degree of polarization of nearly 0.1 in the worst case. Thus, the mean field approach fails in this case.

F. Adding Filtering

In soliton WDM systems, in-line Fabry–Perot filters are sometimes added to reduce the timing jitter due to the Gorden–Haus effect and soliton collisions [9]. Here, we examine its effect on the polarization state of a polarization-scrambled, NRZ WDM signal. When the filter acts on each channel, the effect is to disturb the phase relationship between the different frequency components which make up the channel's signal so that the signal will repolarize. The transmission function $\hat{f}(\omega)$ of the filter may be written

$$\hat{f}(\omega) = \frac{1 - R}{1 - R\exp[i(\omega - \omega_f)z_d/c\tau_0]} \tag{17}$$

where $R = 0.03$ is the reflectivity, ω_f is the central frequency of the filter which is centered on one of the channels, and the the free spectral range $z_d/c\tau_0$ is chosen equal to the channel spacing. As shown in Fig. 13, there is a large amount of repolarization, virtually ruling out the use of in-line filters
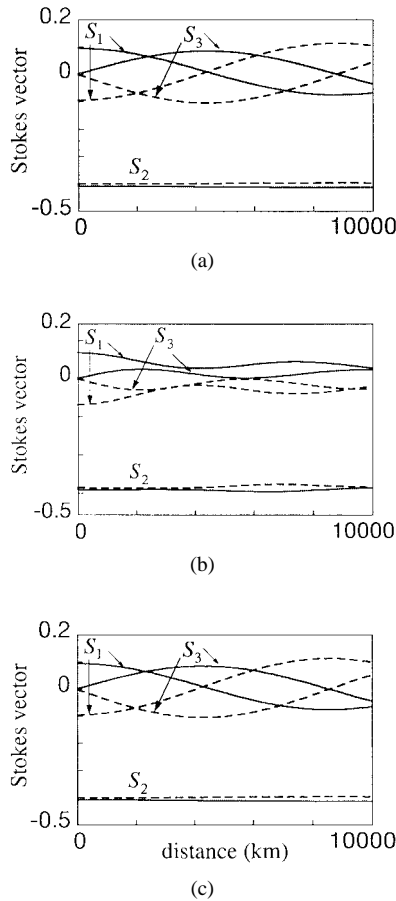
Fig. 11. Evolution of the Stokes vector components as a function of distance in a 5-Gb/s system. The dispersion map length is 200 km. (a) Analytical result. (b) Simulation result, $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. (c) Simulation result, $D_1 = -20$ ps/nm-km, $D_2 = 170$ ps/nm-km. The other parameters are the same as in Fig. 1.
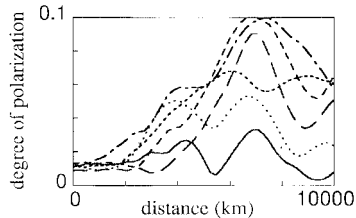

Fig. 12. Evolution of the degree of polarization of seven channels as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The dispersion map length is 200 km. Other system parameters are the same as in Fig. 6.
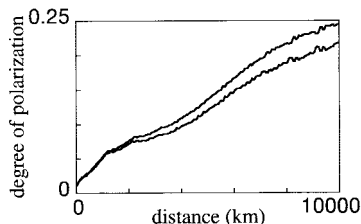

Fig. 13. Evolution of the degree of polarization of two channels as a function of distance with $D_1 = -2$ ps/nm-km, $D_2 = 17$ ps/nm-km. The Fabry–Perot filter is added. Other parameters are the same as in Fig. 1.

in real systems with an NRZ pulse format. From a purely theoretical standpoint, the mean field approach would have to be modified before it could be used with filters.

V. CONCLUSION

In this paper, we have investigated the range of validity of a mean field approach in which we follow only the Stokes parameters of WDM channels and ignore their detailed temporal behavior in determining the evolution of the polarization properties of each channel. The parameters that we used correspond closely to those used by Bergano, *et al.* [3] in system experiments. When the mean field approach fails, it indicates that there is a large nonlinear interaction among neighboring channels that would be bad for communication systems.

We began by deriving the basic equations governing the mean field approach, and we showed that this approach is expected to become exact in the limit of large local dispersion. We then compared the mean field approach to complete simulations. We found for typical parameters and only two channels that this approach does not yield exact quantitative agreement with the simulations. However, the qualitative agreement is good, and, in particular, there is little repolarization consistent with its predictions. We verified that as the local dispersion increases, its predictions ultimately become exact. We also showed that as the number of channels increases, the data rate increases, or the channel spacing increases, its predictions improve. As theoretically expected, we found that including realistic gain and loss only changes the rate of evolution of the Stokes parameters and has no effect on the mean field approach's validity. We also found that including amplitude modulation leads to changes in the approach's predictions, but it does not affect the intersymbol interference or the qualitative behavior. Finally, we found that significantly reducing the channel spacing, shortening the dispersion map, or adding filters leads to a deterioration in the approach's predictions. In both cases, the signals suffered a significant amount of repolarization, and the mean field approach fails. Our results indicate that communication systems should not operate in these regimes. We further showed that our conclusions are not affected by moderate polarization mode dispersion.

Our results serve three purposes. First, they demonstrate that the Kerr effect and chromatic dispersion induce little repolarization for typical system parameters. Thus, the repolarization that is sometimes observed experimentally must be due to something else. In earlier work, we showed that polarization-dependent loss can account for the observed repolarization [4]. Second, they determine when the nonlinear channel interactions become sufficiently strong to induce significant polarization effects. These regimes should be avoided in practice. Given the current trend toward denser channel spacing, this limit is an important one. Third, given the large number of channels that are used in current WDM systems, an unrealistically large amount of CPU time would be required to study these systems using standard simulation methods. There is a clear need for effective reduced models, and the mean field approach offers the prospect of vast computational savings. Our results indicate that the mean field approach captures the essential behavior of the system for

typical parameters, particularly as the number of channels increases, and is therefore a useful approximation. At the same time, it does not yield exact quantitative agreement with full simulations for typical parameters and can fail, for example, when the channel spacing becomes too small or the dispersion map period becomes too short. Thus, it should be used with caution and carefully validated.

In future work, we will extend our model to include the effects of polarization mode dispersion and amplifier spontaneous emission noise along with polarization-dependent loss and polarization-dependent gain. By comparing the results of the extended model to those presented here, we will ultimately be able to assess the real penalties that these different polarization effects introduce into real systems.

ACKNOWLEDGMENT

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, 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, 1991.

[3] N. S. Bergano, "Wavelength division multiplexing in long-haul transmission systems," *J. Lightwave Technol.*, vol. 14, pp. 1299–1308, 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.*, vol. 9, pp. 1247–1249, 1997.

[5] D. Wang and C. R. Menyuk, "A reduced model for the evolution of the polarization states," *Opt. Lett.*, vol. 23, pp. 1677–1679, 1998.

[6] G. P. Agrawal, *Nonlinear Fiber Optics*, San Diego, CA: Academic, 1995.

[7] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 14, pp. 148–157, 1996.

[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, 1997.

[9] L. F. Mollenauer, J. P. Gordon, and S. G. Evangelides, "The sliding-frequency guiding filter: An improved form of soliton jitter control," *Optics Lett.*, vol. 17, pp. 1575–1578, 1992.

[10] P. D. Maker and R. W. Terhune, "Study of optical effects due to an induced polarization third order in the electric field strength," *Phys. Rev.*, vol. 137, pp. A801–A818, 1965.

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

C. R. Menyuk, photograph and biography not available at the time of publication.