Effects of Polarization-Mode Dispersion on Cross-Phase Modulation in Dispersion-Managed Wavelength-Division-Multiplexed Systems

Q. Lin and Govind P. Agrawal, *Fellow, IEEE, Fellow, OSA*

Invited Paper

***Abstract*—This paper develops a vector theory of cross-phase modulation (XPM) in optical fibers and use it to investigate the impact of polarization-mode dispersion (PMD) on the crosstalk induced by XPM in wavelength-division multiplexed lightwave systems. Under certain reasonable approximations, the theory permits us to obtain an analytic expression for the amplitude of probe fluctuations induced by a copropagating pump channel through XPM. We use this expression to calculate the average level of XPM-induced crosstalk together with its variance for several dispersion maps. We show that PMD not only reduces the crosstalk on average, but also impacts the efficiency of a commonly used polarization-interleaving technique.**

***Index Terms*—Cross-phase modulation (XPM), lightwave systems, optical communications, polarization-mode dispersion (PMD).**

I. Introduction

CROSS-PHASE modulation (XPM) is a nonlinear phenomenon occurring in optical fibers when two or more optical fields are transmitted through a fiber simultaneously [1]. It is known to impact the performance of modern wavelength-division-multiplexed (WDM) lightwave systems and has been studied extensively in this context [2]–[16]. The nonlinear phase modulation induced through XPM depends on the bit pattern of the inducing channel and is transferred as intensity fluctuations to neighboring channels through the group-velocity dispersion (GVD), resulting in interchannel crosstalk. The theory of XPM-induced crosstalk that was developed in previous work [2]–[15] is based on a scalar approach and ignores all polarization effects. However, residual birefringence fluctuations occurring in any optical fiber randomize the state of polarization (SOP) of all WDM channels and are thus likely to affect the phase-modulation efficiency of XPM. Indeed, several experiments have shown that polarization-mode dispersion (PMD) of a fiber plays an important role and affects the level of XPM-induced crosstalk [17], [18]. A scalar approach cannot explain such experimental results.

Manuscript received April 29, 2003; revised August 21, 2003. This work was supported by the U.S. National Science Foundation under Grants ECS-0320816 and ECS-0334982.

The authors are with the Institute of Optics, University of Rochester, Rochester 14627, NY 14627 USA (e-mail: gpa@optics.rochester.edu).

Digital Object Identifier 10.1109/JLT.2004.824858

In this paper, we develop a vector theory of XPM that is capable of including PMD-induced random changes in the SOP of various channels and use it to investigate the impact of PMD on the XPM crosstalk in WDM lightwave systems. In Section II, we drive the basic, coupled, vector nonlinear equations for a two-channel system using the Jones-matrix formalism and show that they can be simplified considerably in the pump–probe configuration when we make certain reasonable approximations. The simplified equations are solved in Section III to obtain an analytic expression for the modulation amplitude of the pump-induced interference in the probe channel. However, this modulation amplitude becomes a random quantity in the presence of PMD. We average over the PMD-induced fluctuations in Section IV to calculate the average level of XPM-induced crosstalk and introduce the modulation transfer function and calculate its spectrum for two specific dispersion maps. Section V focuses on the variance of the XPM-induced crosstalk. Here, we average on the random bit pattern of the pump channel as well and show how the variance depends on the PMD parameter and the channel spacing for the same two dispersion maps. The vector theory is extended to the case of multiple WDM channels in Section VI. The main results are summarized in the final concluding section.

II. Vector Theory of XPM

Although the following analysis can be generalized to the case of multiple channels, we first focus for simplicity on the XPM interaction between two channels. We use the general form of the third-order nonlinear polarization for silica glass for including the polarization effects [1], introduce the Jones vectors $|A_1\rangle$ and $|A_2\rangle$ associated with the two channels [19], and use the notation of [1]. As shown in Appendix A, one can then obtain the following two coupled vector equations governing the XPM process in optical fibers:

$$\begin{aligned}\frac{\partial|A_1\rangle}{\partial z}+\beta_{11}\frac{\partial|A_1\rangle}{\partial t}+\frac{i\beta_{21}}{2}\frac{\partial^2|A_1\rangle}{\partial t^2}+\frac{\alpha_1}{2}|A_1\rangle\\ +\frac{i}{2}\boldsymbol{B}_1\cdot\boldsymbol{\sigma}|A_1\rangle-\frac{1}{2}\boldsymbol{\delta}\cdot\boldsymbol{\sigma}\frac{\partial|A_1\rangle}{\partial t}\\ =\frac{i\gamma_1}{3}\left[2\langle A_1|A_1\rangle+|A_1^*\rangle\langle A_1^*|\right]|A_1\rangle\\ +\frac{2i\gamma_1}{3}\left[\langle A_2|A_2\rangle+|A_2\rangle\langle A_2|+|A_2^*\rangle\langle A_2^*|\right]|A_1\rangle \quad (1)\end{aligned}$$

$$\frac{\partial|A_2\rangle}{\partial z}+\beta_{12}\frac{\partial|A_2\rangle}{\partial t}+\frac{i\beta_{22}}{2}\frac{\partial^2|A_2\rangle}{\partial t^2}+\frac{\alpha_2}{2}|A_2\rangle +\frac{i}{2}\boldsymbol{B}_2\cdot\boldsymbol{\sigma}|A_2\rangle-\frac{1}{2}\boldsymbol{\delta}\cdot\boldsymbol{\sigma}\frac{\partial|A_2\rangle}{\partial t} =\frac{i\gamma_2}{3}\left[2\langle A_2|A_2\rangle+|A_2^*\rangle\langle A_2^*|\right]|A_2\rangle +\frac{2i\gamma_2}{3}\left[\langle A_1|A_1\rangle+|A_1\rangle\langle A_1|+|A_1^*\rangle\langle A_1^*|\right]|A_2\rangle \quad (2)$$

where ω_j, β_{1j}, β_{2j}, and $\alpha_j (j=1,2)$ are the optical carrier frequencies, inverse group velocities, GVD coefficients, and fiber losses for the two channels. These parameters are generally z dependent for dispersion-managed periodically amplified systems. Also, the parameter α_j includes both the gain and loss variations along the fiber.

In the Jones-matrix notation, $\langle A|$ is the Hermitian conjugate while $|A^*\rangle$ is the complex conjugate of $|A\rangle$. The random vector $\boldsymbol{B}_j \equiv \boldsymbol{B}(z,\omega_j)$ describes the residual fiber birefringence, while the vector $\boldsymbol{\sigma}$ has the Pauli matrices as its three elements. Because of a frequency difference between the two channels, the SOPs of the two channels evolve on the Poincaré sphere at different rates, as dictated by the magnitudes of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$. The vector $\boldsymbol{\delta}$ describes the *intrachannel* PMD effects resulting from random changes in the group velocities of the two polarization components of the same channel and producing pulse broadening [19], [20]. The effects of both the self-phase modulation (SPM) and XPM are included in (1) and (2) through the nonlinear parameter $\gamma_j \equiv \omega_j n_2/(c a_{\text{eff}})$, where n_2 is the material nonlinear parameter and a_{eff} is the effective core area of the fiber.

Equations (1) and (2) describe the XPM interaction between two channels in its most general form. They are simplified considerably when we consider the pump–probe configuration and assume that channel 2 is in the form of a low-power continuous-wave (CW) probe while channel 1 acts as a pump and imposes the XPM-induced phase shift on channel 2. Two approximations can then be made to simplify the following analysis. First, the probe is assumed to be weak enough that the XPM and SPM induced by it can be neglected. Second, we neglect the $\boldsymbol{\delta}$ terms responsible for *intrachannel* PMD. Although these terms broaden pulses in each channel, they barely affect the *interchannel* XPM interaction because channel spacing typically is much larger than channel bandwidth and the evolution of the SOP of two channels is mainly governed by $\boldsymbol{B}$. This approximation corresponds to the case when *interchannel* PMD dominates but the *intrachannel* PMD is negligible. With these simplifications, (1) and (2) reduce to

$$\frac{\partial|A_1\rangle}{\partial z}+\beta_{11}\frac{\partial|A_1\rangle}{\partial t}+\frac{i\beta_{21}}{2}\frac{\partial^2|A_1\rangle}{\partial t^2}+\frac{\alpha_1}{2}|A_1\rangle =-\frac{i}{2}\boldsymbol{B}_1\cdot\boldsymbol{\sigma}|A_1\rangle+\frac{i\gamma_1}{3}\left[2\langle A_1|A_1\rangle+|A_1^*\rangle\langle A_1^*|\right]|A_1\rangle \quad (3)$$

$$\frac{\partial|A_2\rangle}{\partial z}+\beta_{12}\frac{\partial|A_2\rangle}{\partial t}+\frac{i\beta_{22}}{2}\frac{\partial^2|A_2\rangle}{\partial t^2}+\frac{\alpha_2}{2}|A_2\rangle =-\frac{i}{2}\boldsymbol{B}_2\cdot\boldsymbol{\sigma}|A_2\rangle +\frac{2i\gamma_2}{3}\left[\langle A_1|A_1\rangle+|A_1\rangle\langle A_1|+|A_1^*\rangle\langle A_1^*|\right]|A_2\rangle. \quad (4)$$

XPM induces not only a time-dependent phase shift in the probe channel, but also a nonlinear polarization rotation of the probe channel. However, both the beat length ($\sim$10 m) and correlation length ($\sim$100 m) associated with residual birefringence are much shorter than the nonlinear length (>10 km depending on optical powers). As a result, the polarization rotation induced by residual birefringence is much faster than that induced by nonlinear polarization rotation. Rapid variations in the SOP's average over the SPM and XPM effects in (3) and (4) and eventually the nonlinear PMD becomes negligible [20], [21]. We can average over such rapid polarization variations to study the evolution of XPM on a length scale much longer than the correlation length [22] by adopting a rotating frame through a unitary transformation $|A_j\rangle = \overleftrightarrow{\mathbf{T}}|A_j'\rangle$, where the Jones matrix $\overleftrightarrow{\mathbf{T}}$ satisfies

$$\frac{d\overleftrightarrow{\mathbf{T}}}{dz}=-\frac{i}{2}\boldsymbol{B}_1\cdot\boldsymbol{\sigma}\overleftrightarrow{\mathbf{T}}. \quad (5)$$

After averaging over the fast variations induced by $\overleftrightarrow{\mathbf{T}}$, introducing the reduced time as $\tau = t-\beta_{12}z$ and dropping the primes for notational simplification, (3) and (4) reduce to

$$\frac{\partial|A_1\rangle}{\partial z}+\delta\beta_1\frac{\partial|A_1\rangle}{\partial\tau}+\frac{i\beta_{21}}{2}\frac{\partial^2|A_1\rangle}{\partial\tau^2}+\frac{\alpha_1}{2}|A_1\rangle =i\varepsilon_1 P_0|A_1\rangle \quad (6)$$

$$\frac{\partial|A_2\rangle}{\partial z}+\frac{i\beta_{22}}{2}\frac{\partial^2|A_2\rangle}{\partial\tau^2}+\frac{\alpha_2}{2}|A_2\rangle =-\frac{i}{2}\Omega\boldsymbol{b}\cdot\boldsymbol{\sigma}|A_2\rangle+\frac{i\varepsilon_2}{2}P_0(3+\widehat{P}\cdot\boldsymbol{\sigma})|A_2\rangle \quad (7)$$

(see Appendix B) where $P_0 = \langle A_1|A_1\rangle$ is the pump power, $\widehat{P}=\langle A_1|\boldsymbol{\sigma}|A_1\rangle/P_0$ is a unit vector representing the SOP of the pump on the Poincaré sphere, and $\Omega=\omega_2-\omega_1$ is the channel spacing. We have also introduced $\varepsilon_1 = 8\gamma_1/9$ and $\varepsilon_2 = 8\gamma_2/9$ as the effective nonlinear parameters for the two channels and $\delta\beta_1 = \beta_{11}-\beta_{12}$ as the group-velocity mismatch between the two channels. The birefringence vector $\boldsymbol{b}$ is related to $\boldsymbol{B}$ by a rotation. Since fiber length is typically much longer than the birefringence correlation length, we model $\boldsymbol{b}$ as a three-dimensional (3-D) stochastic process whose first- and second-order moments are given by

$$\langle\boldsymbol{b}(z)\rangle=0,\quad \langle\mathbf{b}(z_1)\mathbf{b}(z_2)\rangle=\frac{1}{3}D_p^2\overleftrightarrow{\mathbf{I}}\,\delta(z_2-z_1) \quad (8)$$

where $\overleftrightarrow{\mathbf{I}}$ is the second-order unit tensor and D_p is the PMD parameter of the fiber.

Equation (6) shows that pump polarization $\widehat{P}$ remains fixed in the rotating frame. However, PMD changes the relative orientation between the pump and probe Stokes vectors at a rate dictated by the magnitude of $\Omega\boldsymbol{b}$. The effectiveness of XPM depends not only on the group-velocity mismatch $\delta\beta_1$, but also on the relative orientation of the pump and probe SOPs. Notice also that fast polarization variations induced by PMD reduce the effective nonlinearity by a factor of 8/9 [23], [24].

III. XPM-Induced Crosstalk

In this section, we solve (6) and (7) approximately to study the temporal modulation of a CW probe induced by the combination of XPM and PMD. Since modulation amplitude is relatively small in general, we can linearize (7) by assuming that $|A_2\rangle = |A_{20}\rangle + |A_{21}\rangle$, where $|A_{20}\rangle$ and $|A_{21}\rangle$ are unperturbed and first-order perturbed probe fields, respectively. Substituting

this form into (7), we obtain two equations governing the evolution of the probe field as

$$\frac{\partial|A_{20}\rangle}{\partial z}+\frac{\alpha_2}{2}|A_{20}\rangle = -\frac{i}{2}\Omega\boldsymbol{b}\cdot\boldsymbol{\sigma}|A_{20}\rangle \tag{9}$$

$$\frac{\partial|A_{21}\rangle}{\partial z}+\frac{i\beta_{22}}{2}\frac{\partial^2|A_{21}\rangle}{\partial\tau^2}+\frac{\alpha_2}{2}|A_{21}\rangle = -\frac{i}{2}\Omega\boldsymbol{b}\cdot\boldsymbol{\sigma}|A_{21}\rangle+\frac{i\varepsilon_2}{2}P_0[3+\widehat{P}\cdot\boldsymbol{\sigma}]|A_{20}\rangle \tag{10}$$

where the dispersion term for $|A_{20}\rangle$ was set to zero for a CW input probe.

The total optical power of the probe field is found using

$$\langle A_2|A_2\rangle\approx\langle A_{20}|A_{20}\rangle+[\langle A_{20}|A_{21}\rangle+\text{c.c.}]\equiv S_0+\delta S_0 \tag{11}$$

where we neglected $\langle A_{21}|A_{21}\rangle$ because of its smallness and introduced the power and SOP of the unperturbed field as

$$S_0=\langle A_{20}|A_{20}\rangle \text{ and } \widehat{S}=\frac{\langle A_{20}|\boldsymbol{\sigma}|A_{20}\rangle}{S_0}. \tag{12}$$

Using (9), the power S_0 and the SOP $\widehat{S}$ are found to satisfy

$$\frac{dS_0}{dz}=-\alpha_2 S_0 \tag{13}$$

$$\frac{d\widehat{S}}{dz}=\Omega\boldsymbol{b}\times\widehat{S}. \tag{14}$$

These two equations can be solved analytically because of their linear nature.

The unperturbed probe power varies along the fiber as $S_0(z)=S_0(0)\exp(-\int_0^z\alpha_2(x)dx)$, but its SOP varies randomly. As is appropriate for problems with randomly varying birefringence, the stochastic differential equation (14) should be treated in the Stratonovich sense [25]. The mean value and correlation function of $\widehat{S}(z)$ can be obtained using the Ito calculus. Appendix C provides the mathematical details. The final result is found to be

$$\left\langle\widehat{S}(z)\right\rangle=\widehat{S}(0)\exp(-\eta z) \tag{15}$$

$$\left\langle\widehat{S}(z_1)\widehat{S}(z_2)\right\rangle=\left\{\frac{\overleftrightarrow{\mathbf{I}}}{3}+\left[\widehat{S}(0)\widehat{S}(0)-\frac{\overleftrightarrow{\mathbf{I}}}{3}\right]\exp(-3\eta z_1)\right\}\times\exp\left[-\eta(z_2-z_1)\right] \tag{16}$$

assuming $z_2>z_1$, where

$$\eta=\frac{1}{L_d}=\frac{D_p^2\Omega^2}{3} \tag{17}$$

and L_d is the PMD diffusion length.

To find the modulation amplitude δS_0, which is the measure of the XPM-induced crosstalk, we first find an equation for $\langle A_{20}|A_{21}\rangle$. Using (9) and (10), this quantity is found to satisfy

$$\frac{\partial\langle A_{20}|A_{21}\rangle}{\partial z}+\frac{i\beta_{22}}{2}\frac{\partial^2\langle A_{20}|A_{21}\rangle}{\partial\tau^2}=-\alpha_2\langle A_{20}|A_{21}\rangle+\frac{i\varepsilon_2}{2}P_0S_0\left[3+\cos\theta(z)\right] \tag{18}$$

where $\theta(z)=\widehat{P}\cdot\widehat{S}(z)$ is the random angle between the pump and probe SOPs. Equation (18) can be solved analytically in the frequency domain because of its linear nature. Using $\delta S_0\equiv\langle A_{20}|A_{21}\rangle+\langle A_{21}|A_{20}\rangle$ and introducing the normalized modulation amplitude as $\widetilde{\delta}_X(\omega)=\widetilde{\delta S}_0(L,\omega)/S_0(L)$, we obtain

$$\widetilde{\delta}_X(\omega)=-\int_0^L\varepsilon_2(z)\widetilde{P}_0(z,\omega)\left[3+\cos\theta(z)\right]\times\sin\left[\frac{1}{2}\int_z^L\omega^2\beta_{22}(z_1)dz_1\right]dz \tag{19}$$

where a tilde denotes the Fourier transform and $\widetilde{P}_0(z,\omega)$ is the Fourier spectrum of the pump power at a distance z inside the fiber. The effects of PMD enter in this equation through the angle $\theta(z)$. More precisely, PMD randomly changes the angle between $\widehat{P}$ and $\widehat{S}$ along the fiber and thus makes δ_X a random quantity.

It is known that the XPM-induced crosstalk depends on the pulse walkoff among channels [3]. However, in long-haul dispersion-managed WDM systems, the fiber link is composed of several periodic parts such that both the dispersion and losses are compensated after each part (or each map period) using optical amplifiers with dispersion-compensating modules. Although pulses in the two neighboring channels walk off from each other in each span, they walk back after the dispersion is compensated and coincide again at the beginning of the next span. As a result, the XPM crosstalk builds up from amplifier to amplifier and can become quite large for long fiber links. In the following, we consider one map period between two amplifiers and then add the contributions from all map periods to find the total crosstalk.

To proceed further, we need to find $\widetilde{P}_0(z,\omega)$ by solving (6) for the pump field. This equation cannot be solved in general because of the nonlinear term appearing on the right side. Strictly speaking, an analytical expression of $\widetilde{P}_0(z,\omega)$ is not known and one must follow a numerical approach. To make further analytical progress, we consider the worst-case situation from the XPM standpoint and assume that the effects of dispersion and nonlinearities do not significantly change the pulse shape of the pump channel along the fiber. From (6), we then obtain

$$\widetilde{P}_0(z,\omega)=\widetilde{P}_0(0,\omega)\exp\left\{\int_0^z[\mathrm{g}_1(z_1)-\alpha_1(z_1)+i\omega\delta\beta_1(z_1)]\,dz_1\right\} \tag{20}$$

where $\mathrm{g}_1(z)$ takes into account amplification of the pump channel explicitly. Substituting this expression into (19), we obtain the following analytic result for the XPM-induced crosstalk

$$\widetilde{\delta}_X(\omega)=\widetilde{P}(0,\omega)\int_0^L\left[3+\cos\theta(z)\right]F(z)dz \tag{21}$$

where the function $F(z)$ takes into account gain, loss, and dispersion variations along the link and is given by

$$F(z)=-\varepsilon_2(z)\exp\left\{\int_0^z[\mathrm{g}_1(z_1)-\alpha_1(z_1)+i\omega\delta\beta_1(z_1)]\,dz_1\right\}\times\sin\left[\frac{1}{2}\int_z^L\omega^2\beta_2(z_1)dz_1\right] \tag{22}$$

and β_{22} is denoted as β_2 to simplify the notation.

In most real systems, residual dispersion exists along the link for reducing the impact of SPM and XPM and it is post compensated at the receiver end [27]. For such a system, pulses in pump channel do not maintain their shape along the fiber link and the use of (20) is likely to overestimate the XPM-induced degradation. However, the theory presented here can be easily used to analyze these systems by using the numerical solutions of (6).

IV. Average Crosstalk Level

As we saw in Section III, PMD-induced fluctuations in the relative orientation between the pump and probe SOPs cause the XPM modulation amplitude to be random. Equation (21) shows that the total modulation amplitude is a sum of all the locally created differential modulations. According to the central limit theorem [25], $\widetilde{\delta}_X$ will follow a Gaussian distribution as long as the correlation between $\cos[\theta(z)]$ and $\cos[\theta(z')]$ goes to zero sufficiently rapidly as $|z-z'|$ increases, no matter what the statistics of $\cos[\theta(z)]$ is. Equation (16) shows that this correlation decays exponentially over a diffusion length $L_d = 1/\eta$. Thus, for any lightwave system of length $L \gg L_d$, $\widetilde{\delta}_X(\omega)$ in the frequency domain and $\delta_X(\tau)$ in the time domain follow a Gaussian distribution as long as the first-order perturbation theory remains valid. Since all statistical information about a Gaussian distribution is contained in the first two moments, the average and variance, we evaluate them in what follows.

The average value of the modulation amplitude is obtained by averaging over random birefringence fluctuations, responsible for PMD, along the fiber length. From (15), $\langle cos\theta(z)\rangle = cos\theta_0 e^{-\eta z}$ and the average value is found to be

$$\left\langle \widetilde{\delta}_X(\omega)\right\rangle_P = \widetilde{P}(0,\omega)\int_0^L [3+\cos\theta_0 e^{-\eta z}]F(z)dz \qquad (23)$$

where θ_0 is the value of θ at $z=0$ and the subscript P indicates that the PMD effects have been averaged out. Note that θ_0 is the relative angle between the pump and probe SOPs at the input end of fiber link. The repolarization effects induced by polarization-dependent gain and polarization-dependent loss are not included in our analysis [26]. The integral can be performed analytically for a two-section dispersion map with different fiber parameters for each section. Assuming that fiber losses are completely compensated at the end of each map period, as shown in Appendix D, the average modulation amplitude at the end of a fiber link of length L is given by

$$\left\langle \widetilde{\delta}_X(\omega)\right\rangle_P = \widetilde{P}(0,\omega)\{3[I(0,\omega)+I^*(0,-\omega)] + \cos\theta_0[I(\eta,\omega)+I^*(\eta,-\omega)]\} \qquad (24)$$

where $I(\eta,\omega)$ is given in Appendix D. In the absence of PMD and for copolarized pump and probe, $\eta = 0$, $\cos\theta_0 = 1$, and ε_2 are replaced by γ_2. Equation (24) then reduces to the result obtained in the scalar case [8], [9].

To characterize the XPM-induced crosstalk, it is common to introduce the modulation transfer function using the definition

$$H(\omega) = \frac{\left\langle \widetilde{\delta S}_0(L,\omega)\right\rangle_P}{\widetilde{P}(0,\omega)} = \frac{S_0(L)}{\widetilde{P}(0,\omega)}\left\langle \widetilde{\delta}_X(\omega)\right\rangle_P. \qquad (25)$$

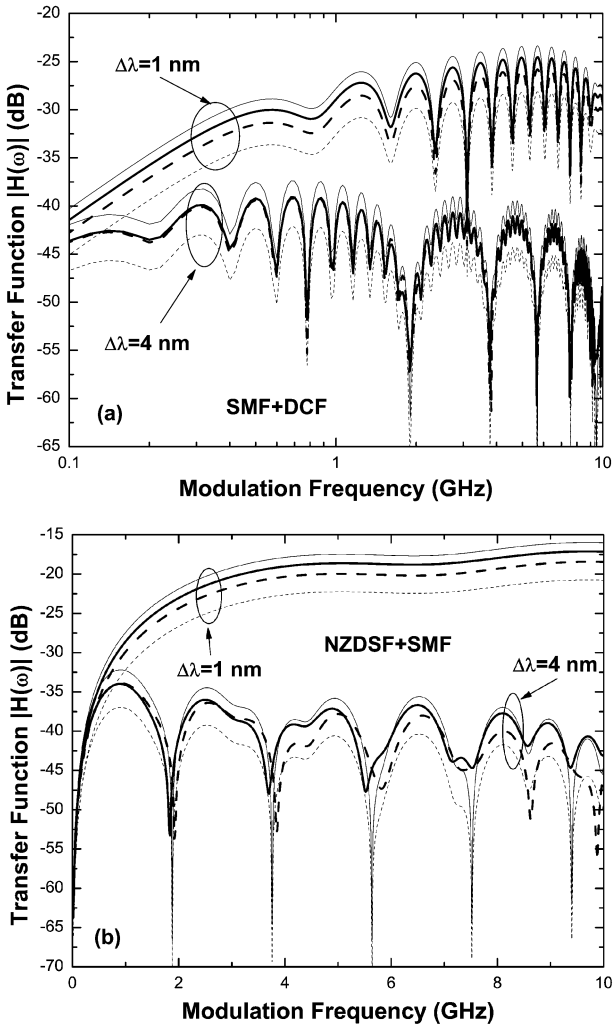

Fig. 1. Amplitude of the transfer function versus modulation frequency for channel spacings of 1 and 4 nm for two dispersion map denoted as (a) $\mathrm{SMF+DCF}$ and (b) $\mathrm{NZDSF+SMF}$. In each case, thin solid and dashed curves show the case without birefringence for copolarized and orthogonally polarized channels, respectively. Thick solid and dash curves correspond to $D_p = 0.1\ \mathrm{ps}/\sqrt{\mathrm{km}}$.

Fig. 1 shows the transfer function $H(\omega)$ for two multispan lightwave systems. The dispersion map in Fig. 1(a) consists of 80 km of standard single-mode fiber (SMF) ($D_a = 17$ ps/nm/km and $\gamma_a = 1.3\ \mathrm{W^{-1}/km}$), followed by 14.32 km of dispersion-compensating fiber (DCF) with $D_b = -95$ ps/nm/km and $\gamma_b = 6\ \mathrm{W^{-1}/km}$. The dispersion map in Fig. 1(b) consists of 85 km of nonzero-dispersion-shifted fiber (NZDSF) with $D_a = -2$ ps/nm/km and $\gamma_a = 2.35\ \mathrm{W^{-1}/km}$, followed by 10 km of SMF. Fiber losses and dispersion slope are taken to be the same and have values $\alpha = 0.21$ dB/km and $S = 0.07\ \mathrm{ps/(km\text{-}nm^2)}$. In both cases, the dispersion maps are chosen such that the average GVD is zero. The unperturbed probe power is 1 mW or 0 dBm.

For comparison, transfer function without residual birefringence is also shown in Fig. 1. In the absence of birefringence, there is 5-dB difference between the copolarized (thin solid line) and orthogonally polarized (thin dashed line) cases due to different XPM coupling efficiencies. However, PMD reduces this difference considerably (thick solid and dashed lines). For $D_p = 0.1\ \mathrm{ps}/\sqrt{\mathrm{km}}$, the difference is reduced to approximately 1.3 dB

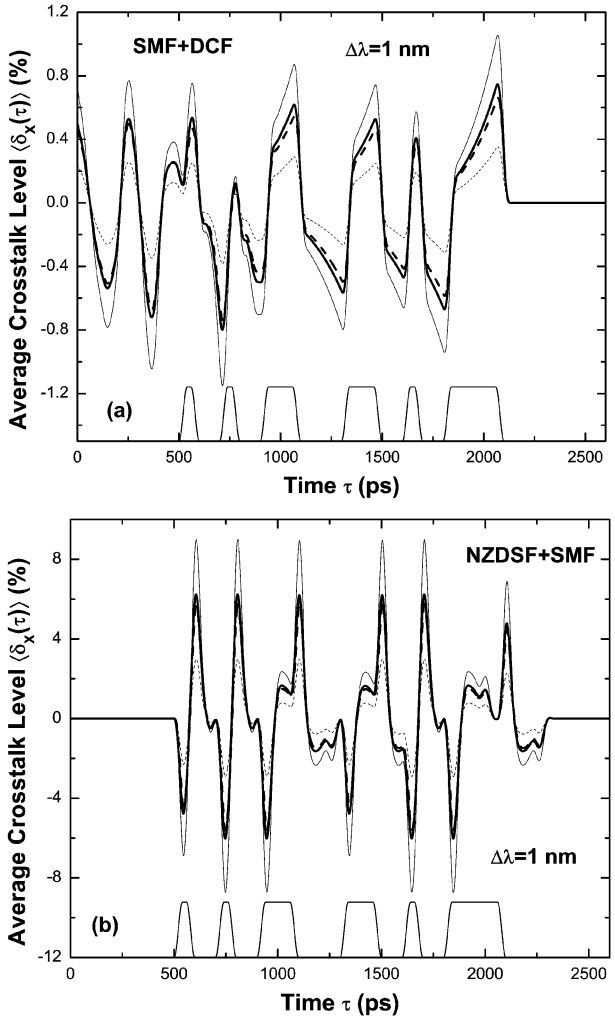

Fig. 2. Variations in average crosstalk level with time for 1-nm channel spacing for the same two dispersion maps. (a) SMF + DCF and (b) NZDSF + SMF. The bottom trace in each case shows the pump bit pattern. Thin solid and dashed curves show the case without birefringence for copolarized and orthogonally polarized channels, respectively. Thick solid and dashed curves correspond to $D_p = 0.2$ ps/$\sqrt{\text{km}}$.

when channel spacing is 1 nm and disappears almost completely when channel spacing increases to 4 nm in the SMF + DCF system due to increased PMD effects. Due to an increase in the walkoff length, the amplitude of transfer function increases significantly in the NZDSF + SMF system. The oscillatory structure in the transfer functions stems from an interference between the XPM-induced phase shifts in different sections of the fiber link and is determined by the walkoff length. The number of ripples is smaller in the NZDSF + SMF system compared with the SMF + DCF system due to increased walkoff length. However, since PMD changes the modulation efficiency of XPM, it also affects the interference condition. This can be seen clearly in Fig. 1(b) when channel spacing is 4 nm. The position of dips in the presence of PMD shifts as modulation frequency increases.

The average crosstalk level $\langle \delta_X(\tau) \rangle_P$ in the time domain is obtained by taking the inverse Fourier transform of (24). This last step is performed numerically for a given bit pattern in the pump channel. Fig. 2 shows the modulation amplitude in time domain for the two dispersion maps used in Fig. 1. The pump channel consists of a 10-Gb NRZ signal bit pattern (raised cosine pulses with a rise time equal to 25% of the bit slot), with a peak power of 8 dBm (corresponding to average power of 5 dBm). The PMD parameter is $D_p = 0.2\,\text{ps}/\sqrt{\text{km}}$ and all other parameters are identical to those used in Fig. 1. As expected, in the absence of birefringence, the modulation amplitude in the case of orthogonally polarized channels is one-third of that occurring for copolarized ones. However, the two curves approach each other in the presence of PMD and the difference becomes negligible. Physically speaking, the two channels cannot maintains their initial SOP (angle θ) after a few diffusion lengths and it becomes completely random. Although, in the first few map periods, the XPM effects are quite different for the copolarized and orthogonally polarized cases, the distinction between them disappears after a few diffusion lengths. Therefore, the difference between the copolarized and orthogonally polarized channels decreases in a long-haul fiber link as the link length increases. In the case of 1-nm channel spacing, the diffusion length is 122 km for $D_p = 0.2$ ps/$\sqrt{\text{km}}$. Thus, after the first two map periods, link length exceeds the diffusion length and the XPM effects become the same for the two cases.

V. Crosstalk Variance

As seen in Fig. 2, the average crosstalk level $\langle \delta_X(\tau) \rangle_P$ changes with time depending on the bit pattern in the pump channel. Thus, $\delta_X(\tau)$ fluctuates with time because of its doubly random nature. In the absence of residual birefringence, randomness comes only from the bit pattern in the pump channel. In this case, it is common to introduce the crosstalk variance, also called XPM-induced interference, using the definition

$$\begin{aligned}\sigma^2 &\equiv \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} [\delta_X(\tau)]^2\, d\tau \\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \widetilde{\delta}_X(\omega_1)\widetilde{\delta}_X^*(\omega_2) \\ &\quad \times \operatorname{sinc}\left[\frac{(\omega_1-\omega_2)T}{2}\right] \end{aligned} \tag{26}$$

where T is the time interval of measurement.

In our case, probe–power variations have two sources of randomness because both the bit pattern in the pump channel and the birefringence variations along the fiber are random. Moreover, PMD can vary with time on a time scale of milliseconds. However, usually the measurement time T is small compared with the fluctuation time of PMD and birefringence fluctuations remain frozen during measurement. In this case, it is appropriate to first average over θ in (21) and then average over a random bit pattern of the pump channel. The average value of the crosstalk variance thus given by

$$\begin{aligned}\sigma_m^2 = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 \left\langle \left\langle \widetilde{\delta}_X(\omega_1)\right\rangle_P \left\langle \widetilde{\delta}_X^*(\omega_2)\right\rangle_P \right\rangle_B \\ \times \operatorname{sinc}\left[\frac{(\omega_1-\omega_2)T}{2}\right] \end{aligned} \tag{27}$$

where the subscript B denotes an average over the pump bit pattern.

The input pump power of a random bit stream can be written as

$$P_0(0,\tau) = \sum_m C_m x\left(\tau - \frac{m}{B_0}\right) \tag{28}$$

where B_0 is the bit rate, $x(\tau)$ is the pulse shape, and the random variable $C_m = 0$ or 1 with equal probabilities, i.e., $p(C_m) = 1/2$. The correlation function of pump power in the frequency domain can be easily calculated and is found to be

$$\left\langle \widetilde{P}_0(0,\omega_1)\widetilde{P}_0^*(0,\omega_2)\right\rangle_B = \pi B_0 \widetilde{X}(\omega_1)\widetilde{X}^*(\omega_2) \times \sum_m \delta(\omega_2 - \omega_1 - 2\pi m B_0) \tag{29}$$

where $\widetilde{X}(\omega)$ is the Fourier transform of $x(\tau)$. Substituting (24) into (27) and using (29), we obtain σ_m^2. Assuming that the measurement time T is much longer than the time slot allocated to a single bit $\mathrm{sinc}(m\pi B_0 T) \ll 1$ for $m \neq 0$. For lightwave systems consisting of a two-section dispersion map, σ_m^2 is found to be given by

$$\sigma_m^2 = \frac{B_0}{4\pi}\int_{-\infty}^{\infty} d\omega \left|\widetilde{X}(\omega)\right|^2 \left|3\left[I(0,\omega) + I^*(0,-\omega)\right] + \cos\theta_0\left[I(\eta,\omega) + I^*(\eta,-\omega)\right]\right|^2 \tag{30}$$

(see Appendix D).

Fig. 3 shows the standard deviation of XPM crosstalk for two values of PMD parameters using the same two dispersion maps that were used earlier. The crosstalk is relatively small because of the complete dispersion compensation assumed here. Although there is a 5-dB difference between the copolarized and orthogonally polarized cases in the absence of birefringence, it decreases quickly with increased channel spacing in the presence of PMD. This significant dependence of XPM crosstalk on channel spacing comes from the fact that the PMD diffusion length L_d is inversely proportional to the square of channel spacing. When channel spacing is small (below 0.5 nm), the reduction in polarization sensitivity only comes from the lowering of the nonlinear parameter by a factor of $8/9$ because the diffusion length is relatively long; $L_d = 216$ km for $D_p = 0.1$ ps/$\sqrt{\text{km}}$ and $\Delta\lambda = 0.5$ nm. When channel spacing becomes larger than 1 nm, the two curves approach each other and eventually merge. If all channels are copolarized initially, PMD helps to reduce the XPM-induced crosstalk in WDM systems. However, this situation changes when polarization interleaving is used. PMD significantly reduces the efficiency of this technique because the neighboring channels do not remain orthogonally polarized. This degradation becomes more severe as the bit rate increases beyond 80 Gb due to increased channel spacing.

Fig. 4 shows the crosstalk variance as a function of PMD parameter for the same two dispersion maps using a channel spacing of 1 nm. When $D_p = 0$, the cases with and without residual birefringence are different by a factor of $1/9$ because

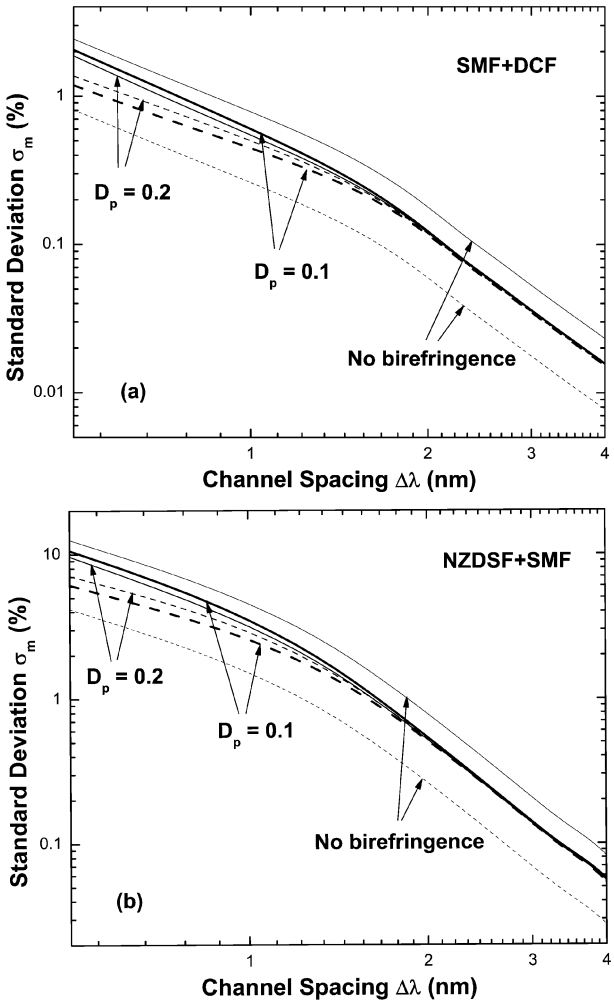

Fig. 3. Standard deviation of modulation interference as a function of channel spacing for two values of D_p for the same two dispersion maps: (a)SMF+DCF and (b) NZDSF + SMF. Solid and dashed curves correspond to copolarized and orthogonally polarized channels, respectively. Thin curves show the case without birefringence. The peak power in the pump channel is 8 dBm in all cases.

of the reduction in the effective nonlinear parameter. This difference decreases quickly with increasing D_p values and becomes negligible for $D_p > 0.3$ ps/$\sqrt{\text{km}}$. When the PMD parameter is not too large, say for $D_p = 0.1$ ps/$\sqrt{\text{km}}$, the difference between the orthogonally polarized channels and copolarized channels is not that large. In a fiber link of moderate PMD, the efficiency of the polarization-interleaving technique becomes questionable when bit rate becomes fairly high.

We briefly consider the impact of PMD fluctuations that typically occur on a time scale of milliseconds because of environmental changes. This case can be considered by averaging over both random processes simultaneously. The average value of crosstalk variance in this case is calculated using

$$\sigma_t^2 = \frac{1}{4\pi^2}\int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \left\langle\left\langle \widetilde{\delta}_X(\omega_1)\widetilde{\delta}_X^*(\omega_2)\right\rangle\right\rangle \times \mathrm{sinc}\left[\frac{(\omega_1 - \omega_2)T}{2}\right]. \tag{31}$$

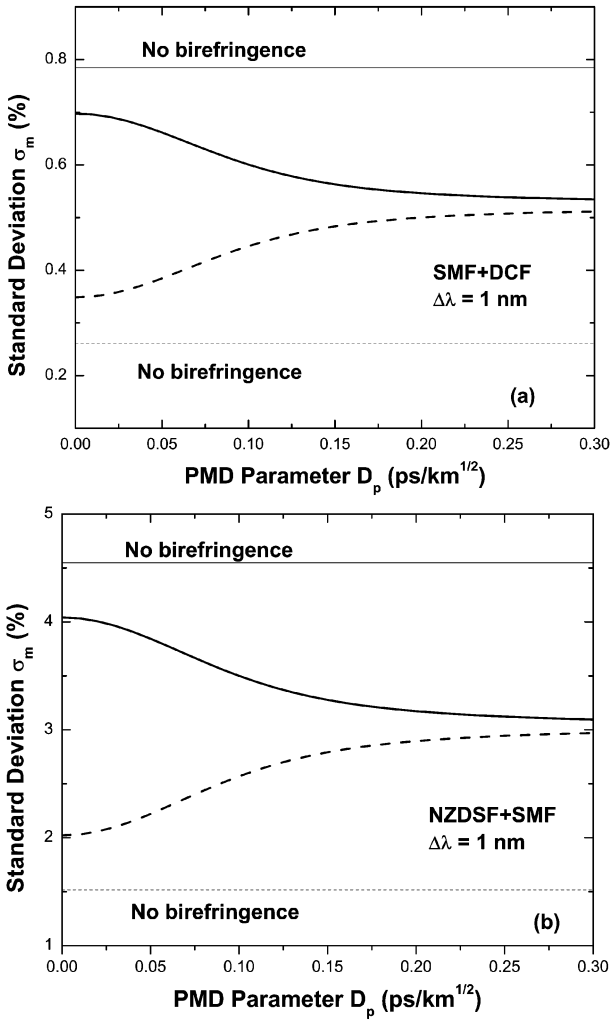

Fig. 4. Standard deviation of modulation interference as a function of PMD parameter for a channel spacing $\Delta\lambda = 1$ nm for the same two dispersion maps: (a) SMF + DCF and (b) NZDSF + SMF. Solid and dashed curves correspond to copolarized and orthogonally polarized channels, respectively. Thin horizontal lines show the case without birefringence.

The difference between σ_t^2 and σ_m^2 provides the crosstalk enhancement induced by long-term PMD fluctuations. Comparing (31) with (27), we find that

$$\sigma_t^2 - \sigma_m^2 = \frac{B_0}{4\pi}\int_{-\infty}^{\infty} d\omega \left|\widetilde{X}(\omega)\right|^2 \times \int_0^L\int_0^L dz_1 dz_2 F(z_1)F^*(z_2)\widehat{P} \cdot \left[\left\langle \widehat{S}(z_1)\widehat{S}(z_2)\right\rangle - \left\langle \widehat{S}(z_1)\right\rangle\left\langle \widehat{S}(z_2)\right\rangle\right]\cdot\widehat{P} \quad (32)$$

where $\langle\widehat{S}(z)\rangle$ and $\langle\widehat{S}(z_1)\widehat{S}(z_2)\rangle$ are given by (15) and (16), respectively. A detailed analysis shows that this difference is rather small for typical channel spacings (in the range $\Delta\lambda = 0.5 - 4$ nm) and for values of the PMD parameter such that $D_p \le 0.2$ ps/$\sqrt{\text{km}}$. The reason is that PMD-induced fluctuations only become important when diffusion length becomes comparable to the walkoff length. The walkoff length determined by the GVD, channel spacing $\Delta\lambda$, and rise time Δt of optical pulses [3] and is given by $L_w = \Delta t/(D\Delta\lambda)$. When $L_d \sim L_w$, $D_p \sim (\lambda^2/2\pi c)[3D/(\Delta t\Delta\lambda)]^{1/2}$. For typical system parameters, such values of D_p are beyond the normal value for existing fibers. For example, taking $\Delta t = 25$ ps for a 100-ps time slot and a channel spacing of 1 nm, this D_p should be approximately 1.8 and 0.62 ps/$\sqrt{\text{km}}$ for SMF and NZDSF, respectively, both of which are much larger than typical D_p values for such fibers. Only when the channel spacing becomes rather large does the long-term PMD-induced fluctuations become significant. However, the crosstalk itself then becomes negligible because of a small walkoff length. Therefore, we conclude that PMD reduces XPM-induced crosstalk in an average sense and induces no additional long-term fluctuations.

VI. Multichannel WDM Systems

In this section, we briefly discuss the more general case of multiple channels. If there are more than two channels, (3) and (4) can be easily extended as

$$\frac{\partial|A_2\rangle}{\partial z} + \beta_{12}\frac{\partial|A_2\rangle}{\partial t} + \frac{i\beta_{22}}{2}\frac{\partial^2|A_2\rangle}{\partial t^2} + \frac{\alpha_2}{2}|A_2\rangle = -\frac{i}{2}\boldsymbol{B}_2\cdot\boldsymbol{\sigma}|A_2\rangle + \frac{2i\gamma_2}{3} \times \sum_j \left[\langle A_j|A_j\rangle + |A_j\rangle\langle A_j| + \left|A_j^*\right\rangle\left\langle A_j^*\right|\right]|A_2\rangle \quad (33)$$

$$\frac{\partial|A_j\rangle}{\partial z} + \beta_{1j}\frac{\partial|A_j\rangle}{\partial t} + \frac{i\beta_{2j}}{2}\frac{\partial^2|A_j\rangle}{\partial t^2} + \frac{\alpha_j}{2}|A_j\rangle = -\frac{i}{2}\boldsymbol{B}_j\cdot\boldsymbol{\sigma}|A_j\rangle + \frac{i\gamma_j}{3}\left[2\langle A_j|A_j\rangle + \left|A_j^*\right\rangle\left\langle A_j^*\right|\right]|A_j\rangle \quad (34)$$

where we consider the XPM effects on channel 2 while all other channels act as the pump ($j \ne 2$). To average over the rapid polarization variations, it is convenient to choose a rotating frame in which the SOP of channel 2 remains frozen. The required unitary transformation corresponds to

$$\frac{d\overleftrightarrow{\mathbf{T}}}{dz} = -\frac{i}{2}\boldsymbol{B}_2\cdot\boldsymbol{\sigma}\overleftrightarrow{\mathbf{T}}. \quad (35)$$

After averaging over rapid variations induced by $\overleftrightarrow{\mathbf{T}}$ and using a reduced time variable as $\tau = t - \beta_{12}z$, we obtain

$$\frac{\partial|A_2\rangle}{\partial z} + \frac{i\beta_{22}}{2}\frac{\partial^2|A_2\rangle}{\partial\tau^2} + \frac{\alpha_2}{2}|A_2\rangle = \frac{i\varepsilon_2}{2}\sum_j P_{0j}[3 + \widehat{P}_j\cdot\boldsymbol{\sigma}]|A_2\rangle \quad (36)$$

$$\frac{\partial|A_j\rangle}{\partial z} + \delta\beta_{1j}\frac{\partial|A_j\rangle}{\partial\tau} + \frac{i\beta_{2j}}{2}\frac{\partial^2|A_j\rangle}{\partial\tau^2} + \frac{\alpha_j}{2}|A_j\rangle = -\frac{i}{2}\Omega_j\boldsymbol{b}\cdot\boldsymbol{\sigma}|A_j\rangle + i\varepsilon_j P_{0j}|A_j\rangle \quad (37)$$

where $P_{0j} = \langle A_j|A_j\rangle$ and $\widehat{P}_j = \langle A_j|\boldsymbol{\sigma}|A_j\rangle/P_{0j}$ represent the power and Stokes vector for the jth channel. Furthermore, $\Omega_j = \omega_j - \omega_2$, $\delta\beta_j = \beta_{1j} - \beta_{12}$, and $\varepsilon_j = 8\gamma_j/9$.

From (37), $\widehat{P}_j$ is found to evolve as

$$\frac{d\widehat{P}_j}{dz} = \Omega_j\boldsymbol{b}\times\widehat{P}_j. \quad (38)$$

This equation is the same as (14). The mean and correlation function of $\widehat{P}_j$ are thus given by (15) and (16), except that η is replaced by $\eta_j = 1/L_{dj} = D_p^2\Omega_j^2/3$.

Following the same procedure as discussed in the Section II, we can find the dynamic equation of the perturbation $\langle A_{20}|A_{21}\rangle$. Equation (18) is now replaced with

$$\frac{\partial\langle A_{20}|A_{21}\rangle}{\partial z} + \frac{i\beta_{22}}{2}\frac{\partial^2\langle A_{20}|A_{21}\rangle}{\partial\tau^2} = -\alpha_2\langle A_{20}|A_{21}\rangle + \frac{i\varepsilon_2}{2}S_0\sum_j P_{0j}[3+\widehat{P}_j\cdot\widehat{S}]. \quad (39)$$

This equation can again be solved in the Fourier domain and the total modulation amplitude for channel 2 is just a sum over all pump channels

$$\widetilde{\delta}_X(\omega) = \sum_{j\neq 2}\widetilde{\delta}_{Xj}(\omega) \quad (40)$$

where $\widetilde{\delta}_{Xj}$ has the form of (21). This equation can be used to calculate the average and variance of $\widetilde{\delta}_X(\omega)$. Clearly

$$\left\langle\widetilde{\delta}_X(\omega)\right\rangle_P = \sum_{j\neq 2}\left\langle\widetilde{\delta}_{Xj}(\omega)\right\rangle_P. \quad (41)$$

If we assume that bit patterns in different channels are independent of each other, it is easy to show that the variance is also a sum over individual channels such that $\sigma^2 = \sum_j \sigma_j^2$, where σ_j^2 is given by (30).

VII. Conclusion

In this paper, we have developed a vector theory of the XPM phenomenon that occurs inside optical fibers using the Jones-matrix formalism. We applied this theory to a pump-probe configuration in which a weak CW probe is perturbed by a pump channel carrying a random bit pattern. The birefringence fluctuations responsible for PMD change the relative orientation θ between the pump and probe Stokes vectors. Changes in θ affect the XPM interaction among various channels of a WDM system. We show that PMD changes the XPM efficiency and affects the interference condition among the XPM-induced nonlinear phase shifts in different sections of the fiber link. Even though the modulation amplitude of the XPM crosstalk is a random quantity, we show, by using the central limit theorem, that it follows a Gaussian distribution.

We show that one can average over birefringence fluctuations using a standard technique to obtain the average crosstalk level. We can also calculate variance of this crosstalk by performing a second average over the bit pattern of the pump channel. We illustrate our analysis for two types of dispersion maps commonly employed in practice. Our results show that PMD reduces the difference in the average crosstalk level between the cases of copolarized and orthogonally polarized channels. In fact, XPM crosstalk becomes polarization independent when channel spacing is large or when the fiber has a relatively large value of the PMD parameter. We thus conclude that when polarization interleaving is not used, PMD helps to reduce the XPM-induced crosstalk in WDM systems. The use of scalar theory in this case may lead to an overly pessimistic conclusion. We have verified that our analytical results agree with full numerical simulations based on the nonlinear Schrödinger equation.

In the case of polarization interleaving, neighboring channels are intentionally launched with orthogonal polarizations for reducing the impact of XPM crosstalk. In this case, PMD significantly reduces the advantage of this technique because neighboring channels do not remain orthogonally polarized along the fiber link. This reduction becomes more severe when channel spacing Ω becomes large because the diffusion length scales with Ω as $L_d = 3(D_p\Omega)^{-2}$ and becomes relatively small for a larger channel spacing for the same value of D_p. When polarization interleaving is used, channel spacing should be kept small enough to make diffusion length comparable to the total system length. In lightwave systems operating at bit rates larger than 40 Gb/s, the efficiency of polarization interleaving will become questionable because channel spacing is likely to exceed 100 GHz. The use of PMD compensation at the receiver will not solve this problem because of the distributed nature of the XPM-induced crosstalk. As XPM is the main source of crosstalk in WDM systems, optical fibers with ultra-low PMD may become essential for implementation of the polarization-interleaving technique.

Appendix A
Coupled Vector Equations

In this appendix, we provide the derivation of (1) and (2). Assuming that the instantaneous electronic response dominates for the XPM process, the third-order nonlinear polarization in a medium such as silica glass is given by [28]

$$\boldsymbol{P}^{(3)}(\boldsymbol{r},t) = \varepsilon_0\overleftrightarrow{\boldsymbol{\chi}}^{(3)}\vdots\boldsymbol{E}(\boldsymbol{r},t)\boldsymbol{E}(\boldsymbol{r},t)\boldsymbol{E}(\boldsymbol{r},t) \quad (42)$$

where $\overleftrightarrow{\boldsymbol{\chi}}^{(3)}$ is a measure of the instantaneous electronic response of the nonlinear medium.

In the case of two distinct optical fields propagating simultaneously inside an optical fiber, the total electric field can be written as

$$\boldsymbol{E} = \mathrm{Re}\,[\boldsymbol{E}_1\exp(-i\omega_1 t) + \boldsymbol{E}_2\exp(-i\omega_2 t)] \quad (43)$$

where Re stands for the real part and $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are the slowly varying (complex) amplitudes for the fields oscillating at frequencies ω_1 and ω_2, respectively. Writing $\boldsymbol{P}^{(3)}$ also in the same form as

$$\boldsymbol{P}^{(3)} = \mathrm{Re}\,[\boldsymbol{P}_1\exp(-i\omega_1 t) + \boldsymbol{P}_2\exp(-i\omega_2 t)] \quad (44)$$

the nonlinear polarization at the pump and signal frequencies is found to be

$$\boldsymbol{P}_j(\omega_j) = \frac{\varepsilon_0\chi^{(3)}_{1111}}{4}\left[(\boldsymbol{E}_j\cdot\boldsymbol{E}_j)\boldsymbol{E}_j^* + 2\left(\boldsymbol{E}_j^*\cdot\boldsymbol{E}_j\right)\boldsymbol{E}_j\right] + \frac{\varepsilon_0\chi^{(3)}_{1111}}{2} \times\left[(\boldsymbol{E}_m^*\cdot\boldsymbol{E}_m)\boldsymbol{E}_j + (\boldsymbol{E}_m\cdot\boldsymbol{E}_j)\boldsymbol{E}_m^* + (\boldsymbol{E}_m^*\cdot\boldsymbol{E}_j)\boldsymbol{E}_m\right] \quad (45)$$

where $j, m = 1$, or $2 (j \neq m)$ and the electronic response is assumed to be isotropic for silica such that $\chi^{(3)}_{1111} = 3\chi^{(3)}_{1122} = 3\chi^{(3)}_{1212} = 3\chi^{(3)}_{1221}$.

If we use (43) and (44) in the Maxwell equations, we find that $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ satisfy in the frequency domain a nonlinear Helmholtz equation of the form

$$\nabla^2\widetilde{\boldsymbol{E}}_j + \frac{\omega_j^2}{c^2}\overleftrightarrow{\boldsymbol{\varepsilon}}_j(\omega)\widetilde{\boldsymbol{E}}_j = -\frac{\omega_j^2}{\varepsilon_0 c^2}\widetilde{\boldsymbol{P}}_j \tag{46}$$

where the tilde denotes Fourier transform ε_0 is the vacuum permittivity and $\overleftrightarrow{\boldsymbol{\varepsilon}}_j(\omega)$ is the linear part of the dielectric constant resulting from the linear response of silica glass. Its tensorial nature is important to account for the PMD effects that have their origin in the birefringence of silica fibers, while its frequency dependence leads to chromatic dispersion.

Both $\widetilde{\boldsymbol{E}}_1$ and $\widetilde{\boldsymbol{E}}_2$ evolve along the fiber length. It is common to choose the z axis along the fiber axis and assume that $\widetilde{\boldsymbol{E}}_1$ and $\widetilde{\boldsymbol{E}}_2$ lie in the $x-y$ plane. This assumption amounts to neglecting the longitudinal component of the two vectors and is justified in practice as long as the spatial size of the fiber mode is larger than the optical wavelength. In the Jones-matrix notation [19], the two fields at any point $\boldsymbol{r}$ inside the fiber can be written as

$$\begin{aligned}\widetilde{\boldsymbol{E}}_1(\boldsymbol{r},\omega-\omega_1) &= F_1(x,y)\left|\widetilde{A}_1(z,\omega-\omega_1)\right\rangle \\ &\quad\times\exp\left[i\beta(\omega_1)z\right] \qquad (47)\\ \widetilde{\boldsymbol{E}}_2(\boldsymbol{r},\omega-\omega_2) &= F_2(x,y)\left|\widetilde{A}_2(z,\omega-\omega_2)\right\rangle \\ &\quad\times\exp\left[i\beta(\omega_2)z\right] \qquad (48)\end{aligned}$$

where $F_1(x,y)$ and $F_2(x,y)$ represent the fiber-mode profile, $\beta(\omega_1)$ and $\beta(\omega_2)$ are propagation constants at the two carrier frequencies, and the Jones vectors $|\widetilde{A}_1\rangle$ and $|\widetilde{A}_2\rangle$ are two-dimensional (2-D) column vectors representing the two components of the electric field in the $x-y$ plane. Since F_1 and F_2 do not change with z, we only need to consider the evolution of $|\widetilde{A}_1\rangle$ and $|\widetilde{A}_2\rangle$ along the fiber.

We substitute (47) and (48) back into (46), integrate over the transverse mode distribution in the $x-y$ plane, and assume $|\widetilde{A}_1\rangle$ and $|\widetilde{A}_2\rangle$ to be slowly varying functions of z so that we can neglect their second-order derivative with respect to z. The fiber-mode profiles can be taken to be nearly the same for typical channel spacings, i.e., $F_1(x,y)\approx F_2(x,y)\equiv F(x,y)$, which amounts to assuming the same effective core area a_{eff} for the two channels. With these simplifications, the equation governing the evolution of $|\widetilde{A}_1\rangle$ and $|\widetilde{A}_2\rangle$ takes the form

$$\frac{\partial|\widetilde{A}_j\rangle}{\partial z}+\left(\frac{\omega_j^2\overleftrightarrow{\boldsymbol{\varepsilon}}_j(\omega)}{2i\beta(\omega_j)c^2}+\frac{i}{2}\beta(\omega_j)\sigma_0\right)|\widetilde{A}_j\rangle = \frac{i\gamma_j}{3}\langle\widetilde{\boldsymbol{P}}_j\rangle_F \tag{49}$$

where the subscript F denotes the average over the mode profile, σ_0 is a unit matrix, and the nonlinear parameter at the carrier frequency ω_j is defined in the usual manner as [1]

$$\gamma_j \equiv \frac{n_2\omega_j}{(ca_{\text{eff}})} = \frac{3\omega_j^2\chi^{(3)}_{1111}}{[8c^2\beta(\omega_j)a_{\text{eff}}]}. \tag{50}$$

To proceed further, we write the dielectric constant tensor $\overleftrightarrow{\boldsymbol{\varepsilon}}$ in the basis of Pauli matrices as [19]

$$\left(\frac{\omega_j^2}{c^2}\right)\overleftrightarrow{\boldsymbol{\varepsilon}}_j(\omega) = \left[\beta_j(\omega)+\frac{i\alpha_j(\omega)}{2}\right]^2\sigma_0-\beta_j(\omega)\boldsymbol{B}(\omega)\cdot\boldsymbol{\sigma}. \tag{51}$$

The vector $\boldsymbol{\sigma}$ is formed as $\boldsymbol{\sigma}=\sigma_1\hat{\boldsymbol{e}}_1+\sigma_2\hat{\boldsymbol{e}}_2+\sigma_3\hat{\boldsymbol{e}}_3$, where $\hat{\boldsymbol{e}}_1$, $\hat{\boldsymbol{e}}_2$, and $\hat{\boldsymbol{e}}_3$ are the three unit vectors in the Stokes space and the Pauli matrices are given as

$$\sigma_1=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \sigma_2=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_3=\begin{pmatrix}0&-i\\i&0\end{pmatrix}. \tag{52}$$

The vector $\boldsymbol{B}(\omega)$ accounts for fiber birefringence. Its frequency dependence produces PMD.

If the channel bandwidth is not too large, we can assume $\alpha_j(\omega)\approx\alpha(\omega_j)$ and expand $\beta_j(\omega)$ and $\boldsymbol{B}(\omega)$ around ω_j in Taylor series as

$$\beta(\omega)\approx\beta(\omega_j)+\beta_1(\omega-\omega_j)+\beta_2\frac{(\omega-\omega_j)^2}{2}+,\ldots \tag{53}$$

$$\boldsymbol{B}(\omega)\approx\boldsymbol{B}(\omega_j)+(\omega-\omega_j)\boldsymbol{\delta}. \tag{54}$$

Using these expansions in (51) and substituting them into (49), we obtain the following vector equation in the frequency domain:

$$\begin{aligned}\frac{d|\widetilde{A}_j\rangle}{dz} &= \left[-\frac{\alpha_j}{2}+i\beta_1(\omega-\omega_j)+\frac{i\beta_2}{2}(\omega-\omega_j)^2\right]|\widetilde{A}_j\rangle\\ &\quad-\frac{i}{2}\left[\boldsymbol{B}(\omega_j)+(\omega-\omega_j)\boldsymbol{\delta}\right]\cdot\boldsymbol{\sigma}|\widetilde{A}_j\rangle+\frac{i\gamma_j}{3}\langle\widetilde{\boldsymbol{P}}_j\rangle_F. \qquad (55)\end{aligned}$$

As a final step, we write (55) in the time domain by using $(\omega-\omega_j)\rightarrow i(\partial/\partial t)$, use the form of nonlinear polarization in (45) and denote $\boldsymbol{B}(z,\omega_j)$ as simply $\boldsymbol{B}_j$ to obtain (1) and (2).

Appendix B
Pump–Probe Equations

In this section, we derive (6) and (7) from (3) and (4) using the transformation (5) and averaging over rapid PMD-induced changes in the SOP at the pump frequency. The unitary matrix $\overleftrightarrow{\mathbf{T}}$ in (5) corresponds to random rotations of the Stokes vector on the Poincaré sphere that do not change the vector length. In the Jones (SU2) space, an arbitrary unitary matrix can be written in the form

$$\overleftrightarrow{\mathbf{T}}=\begin{pmatrix}a_1&-a_2^*\\a_2&a_1^*\end{pmatrix} \tag{56}$$

where $|a_1|^2+|a_2|^2=1$. If we introduce a Jones vector $|a\rangle$ with its two elements as a_1 and a_2, this vector satisfies

$$\frac{d|a\rangle}{dz}=-\frac{i}{2}\boldsymbol{B}_1\cdot\boldsymbol{\sigma}|a\rangle. \tag{57}$$

Since random residual birefringence makes all SOPs equally likely, $|a\rangle$ can be expressed in its most general form as

$$|a\rangle = e^{-\frac{i\varphi_0}{2}}\begin{pmatrix}\cos\left(\frac{\theta}{2}\right)e^{-\frac{i\varphi}{2}}\\ \sin\left(\frac{\theta}{2}\right)e^{-\frac{i\varphi}{2}}\end{pmatrix} \tag{58}$$

where φ_0 and φ are uniformly distributed in the range $[0,2\pi]$ and $\cos\theta$ is uniformly distributed in the range $[-1,1]$. Thus, the most general form of the transformation matrix $\overleftrightarrow{\mathbf{T}}$ is given by

$$\overleftrightarrow{\mathbf{T}}=\begin{pmatrix}\cos\left(\frac{\theta}{2}\right)e^{-\frac{i(\varphi_0+\varphi)}{2}} & -\sin\left(\frac{\theta}{2}\right)e^{\frac{i(\varphi_0-\varphi)}{2}}\\ \sin\left(\frac{\theta}{2}\right)e^{-\frac{i(\varphi_0-\varphi)}{2}} & \cos\left(\frac{\theta}{2}\right)e^{\frac{i(\varphi_0+\varphi)}{2}}\end{pmatrix}. \tag{59}$$

Now, we make the transformation $|A_j\rangle = \overleftrightarrow{\mathbf{T}}|A_j'\rangle$ in (3) and (4) and make use of the relations

$$|A^*\rangle\langle A^*| = |A\rangle\langle A| - \langle A|\sigma_3|A\rangle\sigma_3 \tag{60}$$

$$|A\rangle\langle A| = \frac{[\langle A|A\rangle + \langle A|\boldsymbol{\sigma}|A\rangle\cdot\boldsymbol{\sigma}]}{2}. \tag{61}$$

In the rotating frame, (3) and (4) become

$$\frac{\partial|A_1'\rangle}{\partial z} + \beta_{11}\frac{\partial|A_1'\rangle}{\partial t} + \frac{i\beta_{21}}{2}\frac{\partial^2|A_1'\rangle}{\partial t^2} + \frac{\alpha_1}{2}|A_1'\rangle = \frac{i\gamma_1}{3}\left[3\langle A_1'|A_1'\rangle - \langle A_1'|\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}|A_1'\rangle\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}\right]|A_1'\rangle \tag{62}$$

$$\frac{\partial|A_2'\rangle}{\partial z} + \beta_{12}\frac{\partial|A_2'\rangle}{\partial t} + \frac{i\beta_{22}}{2}\frac{\partial^2|A_2'\rangle}{\partial t^2} + \frac{\alpha_2}{2}|A_2'\rangle = -\frac{i}{2}\Omega\boldsymbol{b}\cdot\boldsymbol{\sigma}|A_2'\rangle + \frac{2i\gamma_2}{3}[2\langle A_1'|A_1'\rangle + \langle A_1'|\boldsymbol{\sigma}|A_1'\rangle\cdot\boldsymbol{\sigma} - \langle A_1'|\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}|A_1'\rangle\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}]|A_2'\rangle \tag{63}$$

where $\boldsymbol{b}$ is related to $\boldsymbol{B}$ by a rotation as

$$\boldsymbol{b} = \overleftrightarrow{\boldsymbol{R}}^{-1}\frac{[\boldsymbol{B}(z,\omega_2) - \boldsymbol{B}(z,\omega_1)]}{\Omega} \equiv \overleftrightarrow{\boldsymbol{R}}^{-1}\frac{d\boldsymbol{B}}{d\omega} \tag{64}$$

where $\overleftrightarrow{\boldsymbol{R}}$ is the three-dimensional (3-D) rotation matrix in the Stokes space that is isomorphic to $\overleftrightarrow{\mathbf{T}}$ in the Jones space, i.e., $\overleftrightarrow{\boldsymbol{R}}\boldsymbol{\sigma} = \overleftrightarrow{\mathbf{T}}^{\dagger}\boldsymbol{\sigma}\overleftrightarrow{\mathbf{T}}$.

We now average (62) and (63) over θ, ϕ, and ϕ_0 to obtain the evolution behavior of the fields on a length scale much longer than the birefringence correlation length. It is easy to show that

$$\left\langle\langle A_1'|\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}|A_1'\rangle\overleftrightarrow{\mathbf{T}}^{\dagger}\sigma_3\overleftrightarrow{\mathbf{T}}\right\rangle = \frac{1}{3}\langle A_1'|\boldsymbol{\sigma}|A_1'\rangle\cdot\boldsymbol{\sigma}. \tag{65}$$

Substituting (65) into (62) and (63) and using the reduced time as $\tau = t - \beta_{12}z$ as the new temporal variable, we obtain (6) and (7).

Appendix C
First- and Second-Order Moments

In this section, we derive (15) and (16) by integrating and averaging (13) and (14). Following [25] with (14), we obtain the dynamic equations governing $\widehat{S}$ and $\widehat{S}\widehat{S}$ in the Ito sense as

$$d\widehat{S} = -\eta\widehat{S}dz + \Omega d\boldsymbol{W}\times\widehat{S} \tag{66}$$

$$d(\widehat{S}\widehat{S}) = -3\eta(\widehat{S}\widehat{S})dz + \eta\overleftrightarrow{\mathbf{I}}dz - \Omega(\widehat{S}\widehat{S})\times d\boldsymbol{W} + \Omega d\boldsymbol{W}\times(\widehat{S}\widehat{S}) \tag{67}$$

where $d\boldsymbol{W} = \boldsymbol{b}dz$ is a 3-D Wiener process. When we average (66) and (67) over $d\boldsymbol{W}$, the terms containing $d\boldsymbol{W}$ disappear. Thus, the evolution of the first- and second-order moments of $\widehat{S}$ is governed by

$$\frac{d\langle\widehat{S}\rangle}{dz} = -\eta\langle\widehat{S}\rangle \tag{68}$$

$$\frac{d\langle\widehat{S}\widehat{S}\rangle}{dz} = -3\eta\langle\widehat{S}\widehat{S}\rangle + \eta\overleftrightarrow{\mathbf{I}}. \tag{69}$$

Equation (68) can be solved easily to obtain (15) while (69) provides the second-order moment of $\widehat{S}$ at $z = z_1$ in the form

$$\left\langle\widehat{S}(z_1)\widehat{S}(z_1)\right\rangle = \frac{\overleftrightarrow{\mathbf{I}}}{3} + \left[\widehat{S}(0)\widehat{S}(0) - \frac{\overleftrightarrow{\mathbf{I}}}{3}\right]\exp(-3\eta z_1). \tag{70}$$

Noting that the correlation function is related to a conditional average as

$$\left\langle\widehat{S}(z_1)\widehat{S}(z_2)\right\rangle = \left\langle\widehat{S}(z_1)\left[\left\langle\widehat{S}(z_2)\right\rangle\right]_{\widehat{S}(z_1)}\right\rangle \tag{71}$$

where the subscript denotes average over $\widehat{S}(z_2)$ provided under the condition that $\widehat{S} = \widehat{S}(z_1)$ at $z = z_1(z_2 > z_1)$. From (68), the conditional average is easily found to be

$$\left\langle\widehat{S}(z_2)\right\rangle\Big|_{\widehat{S}(z_1)} = \widehat{S}(z_1)\exp\left[-\eta(z_2 - z_1)\right]. \tag{72}$$

Substituting (72) into (71) and using (70), we obtain the correlation function of $\widehat{S}$ given in (16).

Appendix D
Average Crosstalk Level

In this section, we calculate the average value of $\widetilde{\delta}_X(\omega)$ given in (24) when the whole link of length L is composed of N map periods of length $L_0 = L_a + L_b$, where L_a and L_b represent lengths of two fiber sections in each map period. The average values of the group-velocity mismatch and GVD for such a map are given by

$$\overline{\delta\beta}_1 = \frac{(\delta\beta_{1a}L_a + \delta\beta_{1b}L_b)}{L_0}, \quad \overline{\beta}_2 = \frac{(\beta_{2a}L_a + \beta_{2b}L_b)}{L_0}. \tag{73}$$

From (22), we notice that $F(z)$ can be separated by two conjugate parts as $F(z) = Q(z,\omega) + Q^*(z,-\omega)$ where

$$Q(z,\omega) = \frac{i\kappa\varepsilon_2(z)}{2}\exp\left[\int_0^z q(z_1)dz_1\right]. \tag{74}$$

$\kappa = \exp(i\omega^2\overline{\beta}_2L/2)$ is a constant and the function $q(z_1)$ is defined as

$$q(z_1) = g_1(z_1) - \alpha_1(z_1) + i\omega\delta\beta_1(z_1) - \frac{i\omega^2\beta_2(z_1)}{2}. \tag{75}$$

If we define an integral function as

$$I(\eta,\omega) = \int_0^L Q(z,\omega)\exp(-\eta z)dz \tag{76}$$

the average value of $\widetilde{\delta}_X(\omega)$ will then be given as (24).

We now evaluate $I(\eta,\omega)$ in a closed form. Since the fiber link is periodic, using $L = \sum_1^N mL_0$, the integral in (76) can be written as a sum over individual map periods as

$$I(\eta,\omega) = \sum_{m=1}^{N}\int_{(m-1)L_0}^{mL_0} Q(z,\omega)\exp(-\eta z)dz \tag{77}$$

where the integrand is of the form

$$Q(z,\omega)e^{-\eta z} = \begin{cases}\frac{i\kappa\varepsilon_{2a}\exp(q_az)h^m(\omega)}{2}, & mL_0 \le z < mL_0 + L_a\\ \frac{i\kappa\varepsilon_{2b}\exp(q_aL_a + q_bz)h^m(\omega)}{2}, & mL_0 + L_a \le z < (m+1)L_0\end{cases} \tag{78}$$

and we have defined

$$h(\omega) = \exp\left[L_0\left(-\eta + i\omega\overline{\delta\beta}_1 - \frac{i\omega^2\overline{\beta}_2}{2}\right)\right] \quad (79)$$

$$q_j = -\alpha_{1j} - \eta + i\omega\delta\beta_{1j} - \frac{i\omega^2\beta_{2j}}{2} \quad (j = a, b). \quad (80)$$

Performing the integrals, $I(\eta, \omega)$ can be written as

$$I(\eta, \omega) = U(\eta, \omega)W(\eta, \omega) \quad (81)$$

where

$$U(\eta, \omega) = \frac{i\varepsilon_{2a}}{2q_a}[e^{q_a L_a} - 1] + \frac{i\varepsilon_{2b}}{2q_b}e^{q_a L_a}[e^{q_b L_b} - 1] \quad (82)$$

$$W(\eta, \omega) = \frac{\exp\left(\frac{i\overline{\beta}_2\omega^2 L}{2}\right) - \exp\left[L(-\eta + i\overline{\delta\beta}_1\omega)\right]}{1 - \exp\left[L_0\left(-\eta + i\overline{\delta\beta}_1\omega - \frac{i\overline{\beta}_2\omega^2}{2}\right)\right]}. \quad (83)$$

REFERENCES

[1] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. San Diego, CA: Academic, 2001.

[2] T. K. Chiang, N. Kagi, M. E. Marhic, and L. G. Kazovsky, "Cross-phase modulation in fiber links with multiple optical amplifiers and dispersion compensation," *J. Lightwave Technol.*, vol. 14, pp. 249–260, 1996.

[3] M. Shtaif, "Analytical description of cross-phase modulation in dispersive optical fibers," *Opt. Lett.*, vol. 23, pp. 1191–1193, 1998.

[4] M. Shtaif and M. Eiselt, "Analysis of intensity interference caused by cross-phase modulation in dispersive optical fibers," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 979–981, July 1998.

[5] R. Hui, Y. Wang, K. Demarest, and C. Allen, "Frequency response of cross-phase modulation in multispan WDM optical fiber systems," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1271–1273, Sept. 1998.

[6] A. V. T. Cartaxo, "Impact of modulation frequency on cross-phase modulation effect in intensity modulation-direct detection WDM systems," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1268–1270, Sept. 1998.

[7] G. Bellotti, M. Varani, C. Francia, and A. Bononi, "Intensity distribution induced by cross-phase modulation and chromatic dispersion in optical-fiber transmissions with dispersion compensation," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1745–1747, Dec. 1998.

[8] A. V. T. Cartaxo, "Cross-phase modulation in intensity modulation-direct detection WDM systems with multiple optical amplifiers and dispersion compensators," *J. Lightwave Technol.*, vol. 17, pp. 178–190, 1999.

[9] R. Hui, K. R. Demarest, and C. T. Allen, "Cross-phase modulation in multispan WDM optical fiber systems," *J. Lightwave Technol.*, vol. 17, pp. 1018–1026, 1999.

[10] S. Bigo, G. Bellotti, and M. W. Chbat, "Investigation of cross-phase modulation limitation over various types of fiber infrastructures," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 605–607, May 1999.

[11] M. Shtaif, M. Eiselt, and L. D. Garrett, "Cross-phase modulation distortion measurement in multispan WDM systems," *IEEE Photon. Technol. Lett*, vol. 12, pp. 88–90, Jan. 2000.

[12] F. S. Yang, M. E. Marhic, and L. G. Kazovsky, "Nonlinear crosstalk and two countermeasures in SCM-WDM optical communication systems," *J. Lightwave Technol.*, vol. 18, pp. 512–520, 2000.

[13] R. I. Killey, H. J. Thiele, V. Mikhailov, and P. Bayvel, "Prediction of transmission penalties due to cross-phase modulation in WDM systems using a simplified technique," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 804–806, July 2000.

[14] S. Betti and M. Giaconi, "Effect of the cross-phase modulation on WDM optical systems: Analysis of fiber propagation," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 305–307, Apr. 2001.

[15] ——, "Analysis of the cross-phase modulation in dispersion compensated WDM optical fiber systems," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 1304–1306, Dec. 2001.

[16] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. New York: Wiley, 2002.

[17] L. Rapp, "Experimental investigation of signal distortions induced by cross-phase modulation combined with dispersion," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 1592–1594, Dec. 1997.

[18] H. J. Thiele, R. I. Killey, and P. Bayvel, "Investigation of cross-phase modulation-induced transmission penalties using the pump-probe technique," *Opt. Fiber Technol.*, vol. 8, pp. 71–81, 2002.

[19] J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," in *Proc. Natl. Acad. Sci. USA*, vol. 97, 2000, pp. 4541–4550.

[20] 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.

[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, pp. 1735–1746, 1997.

[22] C. R. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," *J. Eng. Math.*, vol. 36, pp. 113–136, 1999.

[23] P. K. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringence," *Opt. Lett.*, vol. 16, pp. 1231–1233, 1991.

[24] S. G. Evangelides Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightwave Technol.*, vol. 10, pp. 28–35, 1992.

[25] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. New York: Springer, 1985.

[26] 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, 1997.

[27] G. Bellotti, A. Bertaina, and S. Bigo, "Dependence of self-phase modulation impairments on residual dispersion in 10-Gb/s-based terrestrial transmissions using standard fiber," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 824–826, 1999.

[28] R. W. Boyd, *Nonlinear Optics*, 2nd ed. San Diego, CA: Academic, 2003.

Q. Lin received the B.S. degree in applied physics and the M.S. degree in optics from Tsinghua University, Beijing, China, in 1996 and 1999, respectively. He is currently working toward the Ph.D. degree in the Institute of Optics, University of Rochester, Rochester, NY.

His research interests include nonlinear fiber optics, polarization-mode dispersion, optical communication, and ultrafast optics.

Govind P. Agrawal (M'83–SM'96–F'96) received the B.S. degree from the University of Lucknow, Lucknow, India, in 1969 and the M.S. and Ph.D. degrees from the Indian Institute of Technology, New Delhi, in 1971 and 1974, respectively.

After holding positions at the Ecole Polytechnique, France, the City University of New York, New York, and AT&T Bell Laboratories, Murray Hill, NJ, he joined the faculty of the Institute of Optics at the University of Rochester, Rochester, NY, in 1989, where he is a Professor of Optics. His research interests focus on quantum electronics, nonlinear optics, and laser physics. In particular, he has contributed significantly to the fields of semiconductor lasers, nonlinear fiber optics, and laser physics. He is an author or coauthor of more than 300 research papers, several book chapters and review articles, and five books entitled *Semiconductor Lasers* (Norwell, MA: Kluwer, 2nd ed., 1993), *Fiber-Optic Communication Systems* (New York: Wiley, 3rd ed., 2002), *Nonlinear Fiber Optics* (New York: Academic, 3rd ed., 2001), *Applications of Nonlinear Fiber Optics* (New York: Academic, 2001), and *Optical Solitons: From Fibers to Photonic Crystals* (New York: Academic, 2003). He has also edited two books *Contemporary Nonlinear Optics* (New York: Academic, 1992) and *Semiconductor Lasers: Past, Present and Future* (New York: AIP, 1995).

Dr. Agrawal is a Fellow of the Optical Society of America (OSA). He has participated multiple times in organizing the technical conferences sponsored by the IEEE and OSA. He was the Program Co-Chair and the General Co-Chair of the Quantum Electronics and Laser Science Conference in 1999 and 2001, respectively. He was a Member of the program subcommittee in 2004 for the Conference on Lasers and Electro-Optics (CLEO).