Nonlinear Propagation in Multimode and Multicore Fibers: Generalization of the Manakov Equations

Sami Mumtaz, René-Jean Essiambre, *Fellow, IEEE*, and Govind P. Agrawal, *Fellow, IEEE*

***Abstract*—We investigate theoretically nonlinear transmission in space-division multiplexed (SDM) systems using multimode fibers exhibiting rapidly varying birefringence. A primary objective is to generalize the Manakov equations, well known in the case of single-mode fibers. We first investigate the case where linear coupling among spatial modes of the fiber is weak and derive new Manakov equations after averaging over random birefringence fluctuations. Such an averaging reduces the number of intermodal nonlinear terms drastically since all four-wave-mixing terms vanish. Cross-phase modulation terms still affect multimode transmission but their effectiveness is reduced. We verify the accuracy of new Manakov equations by simulating the transmission of multiple 114-Gb/s bit streams in the PDM-QPSK format over different modes of a multimode fiber and comparing the numerical results with those obtained by solving the full stochastic equations. The agreement is excellent in all cases studied. A major benefit of the new Manakov equations is that they typically reduce the computation time by more than a factor of 10. Our results show that birefringence fluctuations improve system performance by reducing the impact of fiber nonlinearities. The extent of improvement depends on the fiber design and how many spatial modes are used for SDM transmission. We also consider the case where all spatial modes experience strong random linear coupling modeled using a random matrix. We derive new Manakov equations in this regime and show that the impact of some nonlinear effects can be reduced relatively to single-modes fibers. Finally, we extend our analysis to multicore fibers and show that the Manakov equations obtained in the strong- and weak-coupling regimes can still be used depending on the extent of coupling among fiber cores.**

***Index Terms*—Birefringence, fiber nonlinearity, Manakov equation, multicore fiber, multimode fiber, space-division multiplexing.**

I. Introduction

RECENT work on the fiber-capacity limit [1]–[3] has indicated that the continued use of single-mode fibers for modern telecommunication systems may not be able to support the growing data-traffic demand in the near future [4], [5]. Although more single-mode fibers can be deployed as a short-term solution, it is important to look for alternative solutions for the next generation of optical transmission systems. The impressive bit-rate increase of the last two decades was made possible by exploiting diverse properties of electromagnetic fields through wavelength-division multiplexing (WDM) in combination with phase modulation formats and polarization-division multiplexing (PDM). Recently, multimode fibers have become the focus of attention [6], [7]; they permit multiplexing over different spatial modes of the same fiber through space-division multiplexing (SDM). Recent work has focused on demonstrating the feasibility of the SDM technique and how digital signal processing (DSP) can help to compensate linear distortions of the received signal [8], [9]. However, nonlinear penalties are often the limiting factor for modern telecommunication systems and their understanding in the case of multimode fibers is still very limited compared to the case of single-mode fibers.

In this paper we follow an approach similar to that in [10] and derive in Section II a new set of multimode nonlinear equations by representing the two polarization components of each spatial mode through a Jones vector. We use them to study the effects of fiber's random birefringence between the two polarization states of the same mode using random Jones matrices in two important cases of practical interest that we refer to as the weak- and strong-coupling regimes. In the weak-coupling regime, linear coupling among distinct spatial modes is small compared to the birefringence-induced coupling between the two polarization components of the same spatial mode. In contrast, two types of random coupling are comparable in the strong-coupling regime. Even though no mode coupling occurs in ideal multimode fibers, real fibers always exhibit some coupling between any pair of spatial modes because of external factors such as cabling, bending, twisting, and core-size variations occurring during fiber fabrication. In practice, some modes of multimode fibers may be weakly coupled while others are strongly coupled. For instance, it was observed in [8] that the LP11a spatial mode was strongly coupled to the LP11b mode but only weakly coupled to the LP01 mode. The strong- and weak-coupling regimes studied in this work represent two extreme cases for practical systems.

It is well known in the case of single-mode fibers that rapidly varying birefringence results in an averaging effect that reduces the strength of nonlinearities: mathematically, this averaging is described by the Manakov equations [11]–[13] whose use also reduces computation time during numerical simulations. The single-mode Manakov equations cannot be used for multimode fibers as they do not include the intermodal nonlinear effects. In Section III, we derive a new set of Manakov equations for multimode fibers in the weak-coupling regime by averaging over all possible polarization states [11]. In Section IV, the impact of rapidly varying birefringence on a specific SDM transmission system is studied through numerical simulations. In Section V, we consider the strong-coupling regime that has

Manuscript received July 17, 2012; revised October 24, 2012 and November 27, 2012; accepted November 27, 2012. Date of publication December 03, 2012; date of current version December 28, 2012. This work was supported by Alcatel-Lucent in the framework of GreenTouch.

S. Mumtaz and G. P. Agrawal are with the Institute of Optics, University of Rochester, Rochester, NY 14627 USA.

R.-J. Essiambre is with Bell Laboratories, Alcatel-Lucent, Holmdel, NJ 07733 USA.

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

Digital Object Identifier 10.1109/JLT.2012.2231401

attracted theoretical attention to obtain an understanding of stochastic effects such as random differential group delays [14] or random mode-dependent losses [15]. Mecozzi *et al.* recently obtained Manakov equations for multimode fibers in the strong coupling regime [16]. Following a different approach based on random matrix theory, we obtain similar results and compare the system performance in the weak- and strong-coupling regimes. In Section VI, we consider the case of a multicore fiber where linear coupling depends on the core spacing. We show that Manakov equations obtained here in the weak-coupling regime are valid as long as the length scale of the birefringence fluctuations is 100 times shorter than that of linear coupling. We also show that Manakov equations obtained in the strong-coupling regime can be used when the coupling length is comparable to the length scale of birefringence fluctuations.

II. Nonlinear Propagation in Multimode Fibers

We simplify the analysis in [10] by neglecting longitudinal components of modal fields and assume that two linearly polarized components of each spatial mode have the same spatial distribution (see [8]). We write the total electric field in the spectral domain as a sum over M distinct spatial modes of the fiber:

$$\tilde{\mathbf{E}}(x,y,z,\omega) = \sum_{m=1}^{M} e^{\imath \boldsymbol{\beta}_m(\omega) z} \tilde{\mathbf{A}}_m(z,\omega) \mathrm{F}_m(x,y)/\sqrt{\mathrm{N}_m}, \quad (1)$$

where $\tilde{\mathbf{A}}_m(z,\omega) = [\tilde{\mathrm{A}}_{m\mathrm{x}}(z,\omega)\tilde{\mathrm{A}}_{m\mathrm{y}}(z,\omega)]^{\mathrm{T}}$ is the Fourier transform of the slowly varying field envelope of the mth mode in the Jones-vector notation. The superscript $^{\mathrm{T}}$ denotes the transpose operation and a tilde denotes a frequency-domain variable. The mth spatial mode has the spatial distribution $\mathrm{F}_m(x,y)$ and the propagation constant $\boldsymbol{\beta}_m(\omega)$, expressed in the form of a 2×2 diagonal matrix to account for fiber birefringence, i.e., $\boldsymbol{\beta}_m = \mathrm{diag}[\beta_{m\mathrm{x}}\ \beta_{m\mathrm{y}}]$. The modal distributions in (1) satisfy the orthogonality condition and are normalized such that

$$\iint \mathrm{F}_p^*(x,y)\mathrm{F}_m(x,y)dx\,dy = \frac{\bar{n}_{\mathrm{eff}}}{\bar{n}_m}\mathrm{I}_m\delta_{mp}. \quad (2)$$

$\mathrm{N}_m = (1/2)\epsilon_0\bar{n}_{\mathrm{eff}}\,c\,\mathrm{I}_m$ in (1) represents the power carried by the mth mode, $\mathrm{I}_m = (\bar{n}_m/\bar{n}_{\mathrm{eff}})\iint |\mathrm{F}_m|^2(x,y)dx\,dy$, ϵ_0 is the vacuum permittivity, $\bar{n}_m$ the effective index of the mode m, and $\bar{n}_{\mathrm{eff}}$ is the effective index of the fundamental mode. In this paper, we employ the terminology of LP modes obtained in the weakly-guiding approximation, but allow for the possibility of different propagation constants for the modes that are degenerate in that approximation (such as LP11a and LP11b modes).

Each frequency component of the optical field satisfies the following equation:

$$\nabla^2\tilde{\mathbf{E}}(\omega) + n_0^2(x,y)\frac{\omega^2}{c^2}\tilde{\mathbf{E}}(\omega) = -\omega^2\mu_0\tilde{\mathbf{P}}^{\mathrm{NL}}(\omega), \quad (3)$$

where $n_0(x,y)$ is the refractive index profile of the multimode fiber and $\tilde{\mathbf{P}}^{\mathrm{NL}}(\omega)$ is the Fourier transform of the third-order nonlinear response [18]:

$$\mathbf{P}^{\mathrm{NL}}(t) = \frac{\epsilon_0}{4}\chi^{(3)}[(\mathbf{E}^{\mathrm{T}}\mathbf{E})\mathbf{E}^* + 2(\mathbf{E}^{\mathrm{H}}\mathbf{E})\mathbf{E}]. \quad (4)$$

Here, $\chi^{(3)}$ is the third-order nonlinear susceptibility of silica and the superscript $^{\mathrm{H}}$ denotes the Hermitian conjugate.

We assume that the modal spatial distribution is not perturbed by $\mathbf{P}^{\mathrm{NL}}$ or by small variations of the refractive index along z. In that case, $\mathrm{F}_m(x,y)$ satisfies the eigenvalue equation

$$\nabla^2\mathrm{F}_m(x,y) + \left(\frac{\omega^2}{c^2}n_0^2(x,y) - \beta_{mj}^2\right)\mathrm{F}_m(x,y) = 0, \quad (5)$$

where $j = \mathrm{x},\mathrm{y}$ refers to the two orthogonal polarizations of the mth spatial mode. In the case of an ideal fiber with perfect cylindrical symmetry, $\beta_{m\mathrm{x}} = \beta_{m\mathrm{y}}$. In practice, all fibers exhibit some birefringence that fluctuates along the fiber length on a length scale $\sim$100 m. In the remainder of this section, we consider a sufficiently short fiber section and assume birefringence remains constant along its length. Multiple randomly oriented birefringent fibers segments are treated in subsequent sections.

To isolate the evolution of a specific spatial mode, say the pth mode, we substitute (1) in (3), multiply both sides with F_p^*, and integrate over the transverse x-y plane. The resulting equation is then converted to the time domain by following a standard procedure and expanding $\boldsymbol{\beta}_p(\omega)$ in a Taylor series around the carrier frequency ω_0 [19]. The final result can be written in the form

$$\frac{\partial\mathbf{A}_p}{\partial z} = \imath(\boldsymbol{\beta}_{0p} - \beta_r)\mathbf{A}_p - \left(\boldsymbol{\beta}_{1p} - \frac{1}{v_{g_r}}\right)\frac{\partial\mathbf{A}_p}{\partial t} - \imath\frac{\beta_{2p}}{2}\frac{\partial^2\mathbf{A}_p}{\partial t^2} + \imath\sum_{lmn} f_{lmnp}\frac{\gamma}{3}\left[\left(\mathbf{A}_n^{\mathrm{T}}\mathbf{A}_m\right)\mathbf{A}_l^* + 2\left(\mathbf{A}_l^{\mathrm{H}}\mathbf{A}_m\right)\mathbf{A}_n\right], \quad (6)$$

where $\mathbf{A}_p(z,t)$ is the slowly varying envelope of the pth mode expressed in a reference moving frame at a group velocity v_{g_r} and β_r is a reference propagation constant. Further, $\boldsymbol{\beta}_{0p} = \boldsymbol{\beta}_p(\omega_0)$, $\boldsymbol{\beta}_{1p} = \partial\boldsymbol{\beta}_p/\partial\omega|_{\omega_0}$, and $\beta_{2p} = \partial^2\boldsymbol{\beta}_p/\partial\omega^2|_{\omega_0}$ are respectively the propagation constant, inverse group velocity and group-velocity dispersion (GVD) of the pth spatial mode. We assume here that the two polarization components of a spatial mode may have different group velocities but have the same GVD.

The nonlinear parameter in (6) is defined in a fashion similar to single-mode fibers, $\gamma = 3k_0\chi^{(3)}/(4\epsilon_0 c\bar{n}_{\mathrm{eff}}^2 A_1^{\mathrm{eff}})$, where $k_0 = \omega_0/c$ and A_1^{eff} is the effective area of the fundamental mode. The nonlinear coupling among spatial modes is governed by

$$f_{lmnp} = \frac{A_1^{\mathrm{eff}}}{(\mathrm{I}_l\mathrm{I}_m\mathrm{I}_n\mathrm{I}_p)^{1/2}}\iint \mathrm{F}_l^*\mathrm{F}_m\mathrm{F}_n\mathrm{F}_p^*\,dx\,dy. \quad (7)$$

Equation (6) governs propagation of arbitrarily polarized light in the spatial mode p within the Jones-matrix formalism. It includes all third-order nonlinear effects (both intramodal and intermodal types) as well as dispersive and birefringence effects. Fiber losses can also be included by adding the term $-(\alpha_p/2)A_p$ on its right side, where the loss parameter α_p can be different for different modes.

III. Manakov Equations in Weak-Coupling Regime

Equation (6) does not prescribe how constant birefringence evolve along the fiber length. In practice, owing to fiber imperfections such as a nonuniform core whose shape varies along the fiber, birefringence varies rapidly and seemingly randomly on a length scale that is expected to be short compared to the effective lengths associated with the GVD and various nonlinearities. This feature can be implemented numerically by rotating the principal axes of the fiber periodically after a distance shorter than the length scale of birefringence fluctuations. Such an approach is needed when polarization-mode dispersion (PMD) is of concern. However, a numerical solution of (6) requires a relatively small step size and is quite time-consuming.

In the case of single-mode fibers, this problem has been solved by adopting the well-known Manakov-PMD equations [11], [12]. The idea is that, as a rapidly varying birefringence changes continually the state of polarization (SOP) of propagating light in a random fashion, one can average the propagation equation itself over all polarization states. We can follow the same procedure for multimode fibers by assuming that the SOP of each spatial mode evolves randomly and *independently* of other modes. This approximation can be justified by noting that, as spatial distributions of various fiber modes are different, the influence of local stress and fiber imperfections may also be different from one spatial mode to another. Even if some correlation exists in the SOP evolution of different spatial modes over short distances, it is unlikely that it will persist after a sufficiently long propagation distance. Fiber imperfections also cause some linear coupling among spatial modes. In this section we assume this coupling to be so weak that it can be neglected. We have verified numerically that our results remain unaffected when small random mode couplings are included.

A rapidly varying fiber birefringence can be tracked using the transformation

$$\mathbf{A}_p(z) = \mathbf{R}_p(z)\bar{\mathbf{A}}_p(z) \tag{8}$$

where $\mathbf{R}_p(z)$, a unitary matrix belonging to the group U(2), is a Jones matrix of the form

$$\mathbf{R}_p(z) = \begin{bmatrix} r_{11p} & r_{12p} \\ r_{21p} & r_{22p} \end{bmatrix} \tag{9}$$

where r_{ijp} are random variables. An important issue is how to construct $\mathbf{R}_p$ to ensure that the SOP probability for the pth mode is uniformly distributed over the entire Poincaré sphere. We use the following procedure applicable to the general case of $\mathrm{n}\times\mathrm{n}$ matrices [27]: pick a matrix $\mathbf{M}$ whose all elements are normally distributed (with zero mean) and apply a QR decomposition, $\mathbf{M} = \mathbf{RT}$, such that $\mathbf{T}$ is an upper triangular matrix with positive diagonal entries. Then, $\mathbf{R}$ is a Haar matrix whose elements are uniformly distributed over U(n).

The use of (8) in (6) leads to the following equation for $\bar{\mathbf{A}}_p$:

$$\frac{\partial \bar{\mathbf{A}}_p}{\partial z} = \imath\,\boldsymbol{\delta\beta}_{0p}\,\bar{\mathbf{A}}_p - \boldsymbol{\delta\beta}_{1p}\,\frac{\partial \bar{\mathbf{A}}_p}{\partial t} - \imath\frac{\boldsymbol{\beta}_{2p}}{2}\frac{\partial^2 \bar{\mathbf{A}}_p}{\partial t^2} + \mathcal{N}_p, \tag{10}$$

where

$$\boldsymbol{\delta\beta}_{0p} = \mathbf{R}_p^{\mathrm{H}}\left(\boldsymbol{\beta}_{0p} - \beta_r\right)\mathbf{R}_p - \imath\mathbf{R}_p^{\mathrm{H}}\frac{\partial \mathbf{R}_p}{\partial z}, \tag{11}$$

$$\boldsymbol{\delta\beta}_{1p} = \mathbf{R}_p^{\mathrm{H}}\left(\boldsymbol{\beta}_{1p} - 1/v_{g_r}\right)\mathbf{R}_p, \tag{12}$$

and the nonlinear term is given by

$$\mathcal{N}_p = \imath\sum_{lmn} f_{plmn}\frac{\gamma}{3}\left(\left[\bar{\mathbf{A}}_l^{\mathrm{T}}\mathbf{R}_l^{\mathrm{T}}\mathbf{R}_m\bar{\mathbf{A}}_m\right]\mathbf{R}_p^{\mathrm{H}}\mathbf{R}_n^*\bar{\mathbf{A}}_n^* + 2\left[\bar{\mathbf{A}}_l^{\mathrm{H}}\mathbf{R}_l^{\mathrm{H}}\mathbf{R}_m\bar{\mathbf{A}}_m\right]\mathbf{R}_p^{\mathrm{H}}\mathbf{R}_n\bar{\mathbf{A}}_n\right). \tag{13}$$

Equation (10) is stochastic because $\mathbf{R}_m(z)$ is a random matrix that changes along the fiber length on a length scale associated with birefringence fluctuations. As a result, the birefringence parameters appearing there vary randomly. In addition, the intramodal and intermodal nonlinear couplings also become random. To obtain the Manakov equations, we average (10) over all possible realizations of the matrix $\mathbf{R}_m$. Averaging over birefringence amounts to assuming that other phenomena that produce z-dependent variations occur over a much longer length scale in comparison to the length scale of polarization fluctuations within each spatial mode.

Let us focus first on the nonlinear terms in (13). The matrix multiplications appearing there can be expressed as

$$\mathbf{R}_l^{\mathrm{T}}\mathbf{R}_m = \begin{bmatrix} \hat{a}_{lm} & \hat{b}_{lm} \\ \hat{b}_{ml} & \hat{c}_{lm} \end{bmatrix} \quad \mathbf{R}_l^{\mathrm{H}}\mathbf{R}_m = \begin{bmatrix} a_{lm} & b_{lm} \\ b_{ml}^* & c_{lm}, \end{bmatrix} \tag{14}$$

where we have defined the following quantities:

$$a_{lm} = r_{11l}^*\,r_{11m} + r_{21l}^*\,r_{21m} \tag{15}$$
$$b_{lm} = r_{11l}^*\,r_{12m} + r_{21l}^*\,r_{22m} \tag{16}$$
$$c_{lm} = r_{21l}^*\,r_{21m} + r_{22l}^*\,r_{22m} \tag{17}$$
$$\hat{a}_{lm} = r_{11l}\,r_{11m} + r_{21l}\,r_{21m} \tag{18}$$
$$\hat{b}_{lm} = r_{11l}\,r_{12m} + r_{21l}\,r_{22m} \tag{19}$$
$$\hat{c}_{lm} = r_{21l}\,r_{21m} + r_{22l}\,r_{22m}. \tag{20}$$

Using (14) in (13), the nonlinear terms for the x-polarized pth mode become:

$$\begin{aligned}\mathcal{N}_{p\mathrm{x}} = \imath\sum_{lmn} f_{plmn}\frac{\gamma}{3} \\ &\times\Big(2\big[a_{lm}a_{pn}\bar{\mathrm{A}}_{l\mathrm{x}}^*\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{x}} + b_{ml}^*a_{pn}\bar{\mathrm{A}}_{l\mathrm{y}}^*\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{x}} \\ &+ b_{lm}a_{pn}\bar{\mathrm{A}}_{l\mathrm{x}}^*\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{x}} + c_{lm}a_{pn}\bar{\mathrm{A}}_{l\mathrm{y}}^*\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{x}} \\ &+ a_{lm}b_{pn}^*\bar{\mathrm{A}}_{l\mathrm{x}}^*\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{y}} + b_{ml}^*b_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}^*\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{y}} \\ &+ b_{lm}b_{pn}\bar{\mathrm{A}}_{l\mathrm{x}}^*\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{y}} + c_{lm}a_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}^*\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{y}}\big] \\ &+ \Big[\hat{a}_{lm}\hat{a}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{x}}\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{x}}^* + \hat{b}_{ml}\hat{a}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{x}}^* \\ &+ \hat{b}_{lm}\hat{a}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{x}}\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{x}}^* + \hat{c}_{lm}\hat{a}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{x}}^* \\ &+ \hat{a}_{lm}\hat{b}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{x}}\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{y}}^* + \hat{b}_{ml}^*\hat{b}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}\bar{\mathrm{A}}_{m\mathrm{x}}\bar{\mathrm{A}}_{n\mathrm{y}}^* \\ &+ \hat{b}_{lm}\hat{b}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{x}}\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{y}}^* + \hat{c}_{lm}\hat{b}_{pn}^*\bar{\mathrm{A}}_{l\mathrm{y}}\bar{\mathrm{A}}_{m\mathrm{y}}\bar{\mathrm{A}}_{n\mathrm{y}}^*\Big]\Big)\end{aligned} \tag{21}$$

We are interested in finding the ensemble average, or the expectation value denoted by $\langle\,\rangle$, of various coefficients appearing in (21). When the subscripts l, m, p, n correspond to different spatial modes ($l \neq m \neq p \neq n$), the coefficient $a_{lm}a_{pn}$ is a sum of products of functions $r_{11l}, r_{11m}, r_{11n}, r_{11p}$ and $r_{21l}, r_{21m}, r_{21n}, r_{21p}$, all of which are independent random variables with zero mean. As a result, $\langle a_{lm}a_{pn}\rangle = 0$. With the same kind of reasoning, we find that all terms in (21) average to zero, except those containing the combinations

$$\langle\hat{b}_{pp}\hat{b}_{pp}^*\rangle = 1 - \langle\hat{a}_{pp}\hat{a}_{pp}^*\rangle = \frac{1}{3} \tag{22}$$

$$c_{mm}a_{pp} = c_{mm}a_{pp} = a_{pp}a_{pp} = c_{pp}a_{pp} = 1 \tag{23}$$

$$\langle\hat{a}_{pm}\hat{a}_{pm}^*\rangle = \langle\hat{b}_{pm}\hat{b}_{pm}^*\rangle = \frac{1}{2} \tag{24}$$

$$\langle a_{mp}a_{pm}\rangle = \langle b_{pm}b_{pm}^*\rangle = \frac{1}{2}. \tag{25}$$

These equations are obtained using the properties of Haar matrices, which are given by (35) to (37) in the case of $2M \times 2M$ matrices and assuming that $\mathbf{R}_i$ are independent random unitary matrices. Finally, by gathering the nonzero terms both in $\mathcal{N}_{p\mathrm{x}}$ and $\mathcal{N}_{p\mathrm{y}}$, we obtain the following Manakov equation for the pth mode:

$$\frac{\partial \bar{\mathbf{A}}_p}{\partial z} + \langle \boldsymbol{\delta\beta}_{0p} \rangle \bar{\mathbf{A}}_p + \langle \boldsymbol{\delta\beta}_{1p} \rangle \frac{\partial \bar{\mathbf{A}}_p}{\partial t} + \imath \frac{\beta_{2p}}{2} \frac{\partial^2 \bar{\mathbf{A}}_p}{\partial t^2} = \imath\gamma \left(f_{pppp} \frac{8}{9} |\bar{\mathbf{A}}_p|^2 + \sum_{m \neq p} f_{mmpp} \frac{4}{3} |\bar{\mathbf{A}}_m|^2 \right) \bar{\mathbf{A}}_p, \quad (26)$$

with

$$\langle \boldsymbol{\delta\beta}_{0p} \rangle = \frac{1}{2}(\beta_{p\mathrm{x}} + \beta_{p\mathrm{y}}) - \beta_r, \quad (27)$$

$$\langle \boldsymbol{\delta\beta}_{1p} \rangle = \frac{1}{2} \left(\left.\frac{\partial \beta_{p\mathrm{x}}}{\partial \omega}\right|_{\omega_0} + \left.\frac{\partial \beta_{p\mathrm{y}}}{\partial \omega}\right|_{\omega_0} \right) - \frac{1}{v_{g_r}}. \quad (28)$$

These generalized Manakov equations are deterministic, as they do not contain any rapidly fluctuating terms. Averaging over birefringence fluctuations amounts to evaluating the overall effect of random birefringence after a sufficiently long propagation distance. For instance, the average group velocity of the two polarization components is given by (28). In practice, they propagate with group velocities varying around this average value, resulting in the PMD effects. The fluctuations around the average are given by $\boldsymbol{\delta\beta}_{1p} - \langle \boldsymbol{\delta\beta}_{1p} \rangle$ and can be included if PMD effects are of interest. In the same way, nonlinear PMD [12] can be studied by adding the term $\mathcal{N}_p - \langle \mathcal{N}_p \rangle$. By including these rapidly varying terms to (26), one can obtain the generalized Manakov-PMD equation.

Equations (26) represents an extension of the standard Manakov equations for multimode fibers that takes into account random polarization birefringence within each spatial mode. The first nonlinear term represents the *intramodal* nonlinear effects, it occurs for single-mode fibers as well with the same coefficient 8/9. The second nonlinear term is new. Its origin lies in the *intermodal* nonlinearities among various fiber spatial modes. The averaging over random birefringence fluctuations reduces the factor of 2 that is associated with XPM to a lower value of 4/3.

The new Manakov equations (26) obtained for multimode fibers can be solved numerically much faster than (10) and they should help considerably in understanding nonlinear propagation of SDM systems. However, before it can be used with confidence, we need to ensure that its predictions agree with those of (10). In the next section we verify this by considering a specific SDM system.

IV. Numerical Simulations

In this section we compare the performance of a specific SDM system under three different conditions. In the case of no birefringence, we solve the deterministic equation (6) numerically with the split-step Fourier method. Birefringence is included either by solving explicitly the stochastic equation (10) using new random matrices at each step [20] or by solving the new set of generalized Manakov equations (26). We use a step size of 100 m in all cases although larger step sizes could have been used in the Manakov case. This step size corresponds to an average nonlinear phase shift of less than 0.01 rad per step and it was verified that smaller step sizes lead to the same results.

In the following simulations, we assume that each spatial mode carries two single-wavelength channels through PDM (no WDM) at a bit rate of 100 Gb/s with 14% FEC overhead, resulting in a symbol rate of $R_s = 28.5$ Gbaud for the QPSK format. Each PDM-QPSK signal is constructed from an independent random bit stream that is Gray-mapped onto the QPSK symbols. The pulse shape corresponds to a raised-cosine spectrum with a roll-off factor of 0.2. Each PDM-QPSK signal is made of 2^{15} modulated symbols with 2^{17} samples per polarization. The injected power per spatial mode is $P_{\mathrm{in}} = 7$ dBm. This value is deliberately high in order to observe large nonlinear impairments so as to represent a very stringent test of the new Manakov equations.

The 1000-km transmission line consists of 10 sections of 100-km multimode fiber, each section being followed by an ideal erbium-doped fiber amplifier (EDFA) that can compensate for the span losses of 0.2 dB/km (assumed to be the same for each spatial mode). To save computation time, we do not add noise during amplification but add the total noise at the end of the fiber link such that it corresponds to the amplified spontaneous emission (ASE) added by all EDFAs. To justify this procedure, we compared the system performance with full simulations that added noise after each amplifier and found that the location of the noise did not make a noticeable difference.

We focus on a step-index multimode fiber with a core diameter of 12 μm and a numerical aperture of 0.2 ($\Delta = 0.01$). Such a fiber has the V parameter of 5 at 1550 nm, and it supports LP01, LP02, LP11, LP21 modes, resulting in a total of 6 spatial modes when we take into account two-fold degeneracy of the LP11 and LP21 modes [17]. Table I presents the spatial distributions $\mathrm{F}_m(x, y)$ of different spatial modes supported by the fiber. Each spatial mode has its own propagation constant, modal group velocity and chromatic dispersion D. We choose the reference values β_r and v_{gr} to correspond to the fundamental mode of the fiber. With this identification, $\langle \delta\boldsymbol{\beta}_{1p} \rangle$ is the differential modal group delay (DMGD) of the pth mode. Table II lists the values of DMGD together with D and the effective mode area A_m^{eff} for all spatial modes. The value of nonlinear coefficient is $\gamma = 1.4\ \mathrm{W}^{-1}\mathrm{km}^{-1}$ in all our simulations.

At the receiving end, the symbol stream is coherently detected by polarization-diversity 90-degree hybrid, followed by analog to digital converters. We assume an ideal DSP scheme that compensates perfectly all linear impairments (such as chromatic dispersion, group-velocity mismatch, mode mixing, and polarization mixing). This is justified in practice since modern equalizers can remove most, if not all, linear impairments [21]. It also lets us present our results in a way that is independent of the type of equalizer employed.

Fig. 1 shows the bit-error rate (BER), averaged over all propagating spatial modes when SDM is employed, as a function of the optical signal-to-noise ratio (OSNR) [1] after 1000 km when 114-Gb/s bit streams are transmitted on different combinations of spatial modes. Solid and dashed curves compare predictions of our new Manakov equations (26) with of the full stochastic

TABLE I
SPATIAL DISTRIBUTIONS OF THE MODES SUPPORTED BY THE STEP-INDEX AND GRADED-INDEX FIBERS USED IN NUMERICAL SIMULATIONS

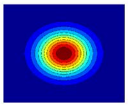
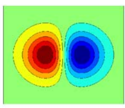
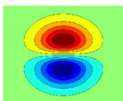
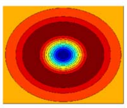
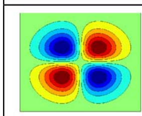
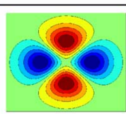
Spatial distribution shapes	Step-index fiber	Graded-index fiber
	LP01	HG00
 	LP11a, LP11b	HG01, HG10
	LP02	HG02+HG20
 	LP21a, LP21b	HG11a, HG11b

TABLE II
DMGD, DISPERSION D, AND EFFECTIVE MODE AREA A_m^{eff} OF VARIOUS SPATIAL MODES SUPPORTED BY THE STEP-INDEX MULTIMODE FIBER

	DMGD [ns/km]	D [ps/(km-nm)]	A_m^{eff} [μm^2]
LP01	0	25	80
LP11	6.5	27.3	76
LP02	9.9	-2.3	83
LP21	12	20.8	86

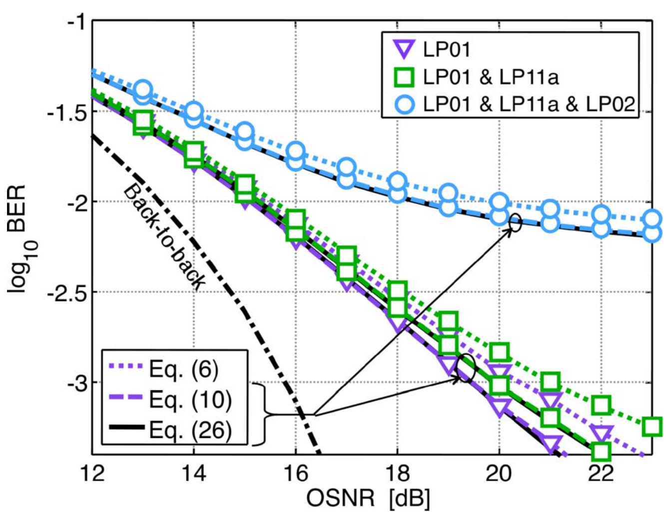


Fig. 1. Calculated BER versus OSNR when a step-index few-mode fiber is used to transmit 114-Gb/s PDM-QPSK signals over 1000 km using one (triangles), two (squares), or three (circles) spatial modes (BER is averaged in the case of multiple spatial modes). Solid curves obtained using Manakov equations (26) overlap completely with dashed curves obtained using the stochastic equation (10); the dotted curves show the zero-birefringence case using (6). Back-to-back case (no fiber) is also shown for reference.

equations (10); their complete overlap indicates that our new Manakov equations should prove useful for SDM systems as its use reduces considerably the computation time.

To study the impact of random birefringence on system performance, we also show in Fig. 1 with dotted curves the case of an ideal multimode fiber (no birefringence) by solving (6) numerically. It is evident that birefringence actually improves system performance in all cases by lowering the BER slightly. This improvement in BER is easily understood if we recall that the averaging over rapidly varying birefringence leads to (26) in which the nonlinear coefficient associated with the SPM intramodal nonlinearities is reduced from 1 to 8/9. Note also that, for the chosen configurations of transmitted spatial modes, intermodal XPM effects are averaged out because of a large differential group delay among spatial modes [22], [23]. The reason behind the large BER degradation observed in Fig. 1 in the case of 3 co-propagating spatial modes $(\mathrm{LP01}+\mathrm{LP11}+\mathrm{LP02})$ is related to a relatively small value of the dispersion D for the LP02 mode that degrades the performance of that mode (see Table II).

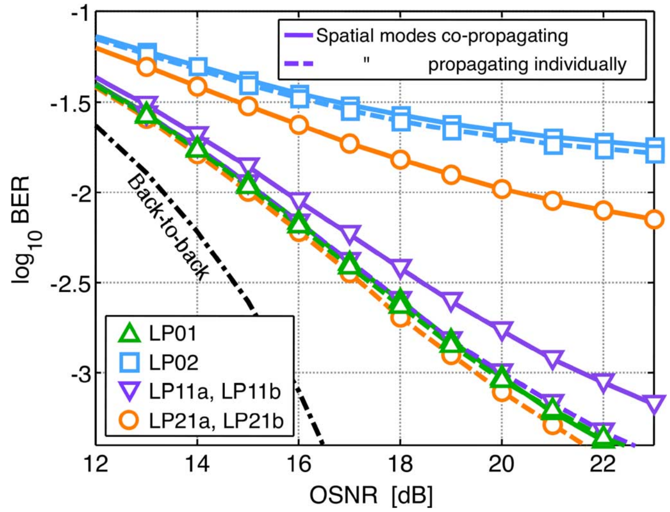


Fig. 2. Calculated BER versus OSNR for individual spatial modes after 1000 km transmission of 114-Gb/s PDM-QPSK signals. Solid curves show the case when all 6 spatial modes are used through SDM, while dashed curves are obtained when signal is launched into one spatial mode of a multimode fiber (no SDM).

Fig. 2 shows the calculated BER for each spatial mode in two cases. Solid curves show the case when all 6 spatial modes are used through SDM, while dashed curves are obtained when signal is launched into one spatial mode of a multimode fiber (no SDM). The BER degradation between these two cases is due to the inter-modal XPM. We observe that LP01 and LP02 spatial modes suffer little from inter-modal XPM because of their nondegenerate nature. In contrast, both LP11 and LP21 modes suffer considerably from inter-modal XPM because of their degenerate nature that allows mode pairs such as LP11a and LP11b to propagate at the same group velocity. The nonlinear inter-modal XPM coefficient is found to be 0.35 for the LP11 pair and 0.92 for the LP21 pair. This relatively large value for the LP21 pair is responsible for the large BER degradation seen in Fig. 2. We should stress that these results neglect any linear coupling among degenerate modes. In practice, degenerate modes can couple strongly, resulting in an averaging effect that will reduce such large inter-modal XPM penalties.

In a step-index multimode fiber, various spatial modes propagate with different group velocities. This DMGD introduces linear degradations (similar to PMD) that must be compensated at the receiver. In practice, the equalization complexity grows in proportion to the magnitude of DMGDs. One way to reduce the DMGD problem is to employ a graded-index fiber for which

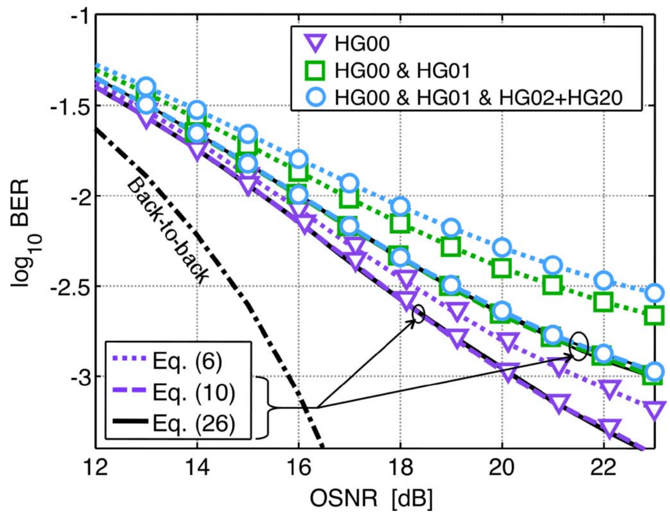


Fig. 3. Same as Fig. 1, but using a graded-index multimode fiber.

DMGD nearly vanishes for all modes supported by the fiber. It is thus important to study how the system performance in Fig. 1 will change if the step-index multimode fiber used there were replaced with a graded-index multimode fiber. For this purpose, we focus on a graded-index fiber whose refractive index varies quadratically inside the core (with a parabolic shape). In such idea fibers, all spatial modes propagate with the same group velocity [17].

For numerical simulations, we consider a graded-index fiber with a core diameter of 17.4 μm. Such a fiber supports several Hermite-Gauss modes listed in Table I with $D = 21.5$ ps/(km-nm) for all modes and $\gamma = 1.4$ W^{-1}km^{-1}. As before, we assume no linear coupling occurs among the fiber modes. For a fair comparison with the LP02 mode of a step-index fiber, we consider a spatial distribution corresponding to the sum of the HG02 and HG20 spatial modes. Fig. 3 shows the BER as a function of OSNR under the same conditions used in Fig. 1, except that the graded-index fiber replaces the step-index fiber. A comparison of two figures shows several interesting features. First, a rapidly varying birefringence results in a larger improvement in the case of graded-index fibers. Since all spatial modes have the same group velocity, the inter-modal XPM effects are enhanced in a graded-index fiber. The effect of rapidly varying birefringence results in the reduction of the XPM coefficient from 2 to 4/3, which improves system performance. Second, nonlinear penalties for graded-index fibers are larger in the two-mode case but become considerably smaller in the three-mode case. This can be explained by noting that all spatial modes have here the same GVD of 21.5 ps/(km-nm). In the case of our step-index fiber, the LP02 had a very small dispersion parameter, resulting in poor performance for this spatial mode.

Finally, we come to possibly the main advantage of using the new Manakov equations (26). We have seen that its numerical predictions give virtually identical results compared to those obtained by solving (10). From a practical standpoint, the use of Manakov equations is preferable because it requires much less computational time. One reason is that the number of nonlinear terms is drastically reduced. For instance in the case of 3 propagating modes, there are 81 nonlinear terms in (10) but only 3 nonlinear terms are needed in (26). In general, there are $M^3 - 1$ intermodal nonlinear terms for a M-mode fiber while there are $M - 1$ intermodal nonlinear terms for each generalized Manakov equation. A second reason is that a solution of (10) requires computation of an exponential matrix of size $2M \times 2M$, M being the number of spatial modes, which is a time-consuming numerical operation. As a result, the reduction in computing time depends on the value of M. This is evident in Table III, which shows the computation time obtained on a desktop computer using the MATLAB software (version 2007b) for 1000-km propagation of 114-Gb/s channels with a step size of 100 m and a temporal grid size of 2^{17}. Computation time is reduced by nearly a factor of 10 even for $M = 2$, with larger reductions occurring for larger values of M. For the sake of comparison, identical step sizes were used in both cases. However, a much larger step size can be used in the Manakov case because step size is set by the nonlinear effects which usually occur on a length scale much longer than that of birefringence fluctuations. Taking this into account, the use of the new Manakov equations should reduce the computing time by a factor of more than 100 in most cases of practical interest.

TABLE III
COMPUTING TIME IN SECONDS FOR 1000-KM TRANSMISSION OF 114-GB/S CHANNELS OVER MULTIPLE MODES WITH A 100-M STEP SIZE

Number of modes, M	1	2	3	4
Manakov Eq. (26)	270	800	1600	2800
unaveraged Eq. (10)	720	7600	16000	31000

V. MANAKOV EQUATIONS IN STRONG-COUPLING REGIME

The Manakov equations (26) are obtained in the limiting case of no linear coupling among spatial modes. As mentioned before, in practice, some linear coupling occurs invariably in any multimode fiber. In general, the strength of coupling may vary for all mode pairs and is likely to be strongest for nearly degenerate modes. Because it is difficult to address the general problem of mixed coupling, it is assumed often that all modes are strongly coupled [3], [14], [15]. We also adopt the strong-coupling regime and derive new Manakov equations in this situation. More specifically, we employ $2M \times 2M$ random unitary matrices to account for random couplings among M spatial modes, with two polarization states per mode.

Before applying such matrices to the nonlinear propagation equations (6), we need to rewrite it in a fully vectorial form. By introducing $\mathcal{A} = [\mathbf{A}_1^{\mathrm{T}} \ldots \mathbf{A}_M^{\mathrm{T}}]^{\mathrm{T}}$ as the column vector containing $2M$ field envelopes. We obtain

$$\frac{\partial \mathcal{A}}{\partial z} = \imath(\mathcal{B}_0 - \beta_r)\mathcal{A} - \left(\mathcal{B}_1 - \frac{1}{v_{g_r}}\right)\frac{\partial \mathcal{A}}{\partial t} - \imath\frac{\mathcal{B}_2}{2}\frac{\partial^2 \mathcal{A}}{\partial t^2} + \iint \frac{\imath\gamma}{3}\Big[\left(\mathcal{A}^{\mathrm{T}}\mathcal{F}^{(1)}\mathcal{A}\right)\mathcal{F}^{(1)*}\mathcal{A}^* + 2\left(\mathcal{A}^{\mathrm{H}}\mathcal{F}^{(2)}\mathcal{A}\right)\mathcal{F}^{(2)}\mathcal{A}\Big]\,dx\,dy \tag{29}$$

where $\mathcal{F}^{(k)}$ is a $2M \times 2M$ matrix such that

$$\mathcal{F}^{(k)} = \begin{bmatrix} \mathbf{F}_{11}^{(k)} & \cdots & \mathbf{F}_{1M}^{(k)} \\ \vdots & \ddots & \\ \mathbf{F}_{M1}^{(k)} & & \mathbf{F}_{MM}^{(k)} \end{bmatrix} \tag{30}$$

whose M^2 elements are 2×2 block matrices such that

$$\mathbf{F}_{ij}^{(1)} = \frac{1}{(\mathrm{I}_i \mathrm{I}_j)^{1/2}} \mathrm{F}_i \, \mathrm{F}_j \, \mathbf{I}_2, \quad \mathbf{F}_{ij}^{(2)} = \frac{1}{(\mathrm{I}_i \mathrm{I}_j)^{1/2}} \mathrm{F}_i^* \, \mathrm{F}_j \, \mathbf{I}_2, \tag{31}$$

where $i, j = 1 \ldots M$ and $\mathbf{I}_2$ is the 2×2 identity matrix.

The matrices $\mathcal{B}_0$, $\mathcal{B}_1$ and $\mathcal{B}_2$ are $2M \times 2M$ diagonal matrices containing respectively the propagation constants, inverse group velocities, and dispersion parameters of each mode. $\beta_r = \mathrm{trace}(\mathcal{B}_0)/2M$ and $1/v_{g_r} = \mathrm{trace}(\mathcal{B}_1)/2M$ represent respectively the average values of the propagation constant and inverse group velocity.

Following the procedure used in Section III, we apply the substitution $\mathcal{A} = \mathcal{R}\bar{\mathcal{A}}$ in (29) and obtain

$$\begin{aligned} \frac{\partial \bar{\mathcal{A}}}{\partial z} &= \imath \delta \mathcal{B}_0 \bar{\mathcal{A}} - \delta \mathcal{B}_1 \frac{\partial \bar{\mathcal{A}}}{\partial t} - \imath \frac{\mathcal{B}_2}{2} \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} \\ &+ \iint \frac{\imath \gamma}{3} \Big[\left(\bar{\mathcal{A}}^{\mathrm{T}} \mathcal{R}^{\mathrm{T}} \mathcal{F}^{(1)} \mathcal{R} \bar{\mathcal{A}} \right) \mathcal{F}^{(1)*} \mathcal{R}^{\mathrm{H}} \mathcal{R}^* \bar{\mathcal{A}}^* \\ &+ 2 (\bar{\mathcal{A}}^{\mathrm{H}} \mathcal{R}^{\mathrm{H}} \mathcal{F}^{(2)} \mathcal{R} \bar{\mathcal{A}}) \mathcal{F}^{(2)} \bar{\mathcal{A}} \Big] \, dx \, dy \end{aligned} \tag{32}$$

where

$$\delta \mathcal{B}_0 = \mathcal{R}^{\mathrm{H}} (\mathcal{B}_0 - \beta_r) \mathcal{R} - \imath \mathcal{R}^{\mathrm{H}} \frac{\partial \mathcal{R}}{\partial z}, \tag{33}$$

$$\delta \mathcal{B}_1 = \mathcal{R}^{\mathrm{H}} (\mathcal{B}_1 - \frac{1}{v_{g_r}}) \mathcal{R}. \tag{34}$$

Here, $\mathcal{R}$ is a random $2M \times 2M$ rotation matrix whose elements r_{ij} satisfy the following relations [28]:

$$\langle |r_{ik}|^2 |r_{ik'}|^2 \rangle = \frac{1 + \delta_{kk'}}{2M(2M+1)}, \tag{35}$$

$$\langle |r_{km}|^2 |r_{k'm'}|^2 \rangle = \frac{1}{(2M+1)(2M-1)}, \quad \begin{matrix} k \neq k' \\ m \neq m' \end{matrix} \tag{36}$$

$$\langle r_{ki}^* r_{kj} r_{k'i} r_{k'j}^* \rangle = \frac{-1}{(2M+1)2M(2M-1)} \quad \begin{matrix} k \neq k' \\ i \neq j \end{matrix} \tag{37}$$

After averaging (32) we obtain the following Manakov equation in the strong-coupling regime:

$$\frac{\partial \bar{\mathcal{A}}}{\partial z} + \imath \frac{\bar{\beta}_2}{2} \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} = \imath \gamma \kappa |\bar{\mathcal{A}}|^2 \bar{\mathcal{A}} \tag{38}$$

where $\bar{\beta}_2 = \mathrm{trace}(\mathcal{B}_2)/2M$ is the average GVD and

$$\kappa = \sum_{k \leq l}^{M} \frac{32}{2^{\delta_{kl}}} \frac{f_{kkll}}{6M(2M+1)}. \tag{39}$$

Equations (38) and (39) are in agreement with the results in [16] obtained using a different procedure in a more general case.

These Manakov equations result in having all spatial modes propagate, on average, with the same group velocity v_{g_r}. However, in practice, fluctuations of the mode velocities around v_{g_r} could be significant, due to incomplete averaging of group velocities. One can account for them by adding rapidly varying terms that correspond to the DMGD fluctuations given by $\mathcal{R}^{\mathrm{H}} \mathcal{B}_1 \mathcal{R} - 1/v_{g_r}$ as in Eq. (34) (similar to the PMD case discussed in Section III).

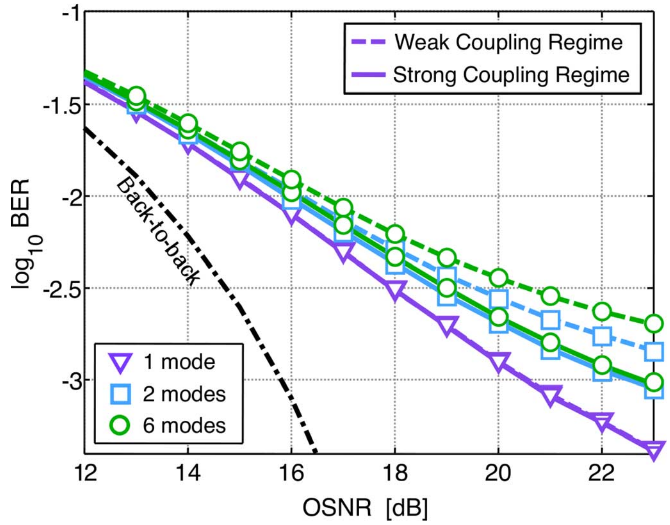


Fig. 4. Averaged BER versus OSNR after 1000 km of propagation over a graded-index fiber when 114-Gb/s PDM-QPSK signals are transmitted using one, three, and six spatial modes. Square and diamond markers represent the weak- and strong-coupling regimes, respectively. The dotted curve shows the back-to-back case (no fiber).

A comparison of the magnitudes of various nonlinear coefficients appearing in (26) and (38) reveals that system performance improves in the high-coupling regime as the number of spatial modes increases. This feature suggests that multimode fibers supporting $2M$ modes exhibiting strong-coupling regime may perform better than M single-mode fibers supporting 2 polarization modes under certain conditions.

Using the same parameter values used in Section IV, Fig. 4 presents numerical results for 1000-km transmission over a graded-index fibers supporting one, three, and seven Hermite-Gaussian HG_{mn} spatial modes. In the case of a three-mode fiber, data are transmitted on the HG00, HG01 and HG10 spatial modes. In the case of a seven mode-fiber, data are transmitted on the six spatial modes presented in Table I. Comparison of the weak- and strong-coupling regimes shows that, as expected, the BER curves in the two cases coincide when only one spatial mode of the fiber is used. However, the high-coupling regime results in better BER performance when data is transmitted using multiple spatial modes. Another difference is that the performance is degraded in the weak-coupling regime as the number of spatial modes increases whereas, somewhat surprisingly, performance is nearly the same for three or six spatial modes in the high-coupling regime.

VI. Multicore Fibers

Future SDM systems may also make use of multicore fibers in which each core supports a single spatial mode, but modes in different cores can couple strongly if these cores are relatively close to each other. In this case, the coupling strength depends on the physical distance among its cores and its magnitude varies exponentially with this distance.

For numerical simulations, we consider an ideal 3-core fiber having identical, and equidistant single-mode cores in an equilateral triangular configuration. There are three spatial modes F_m, each associated with the field propagating in one core. The

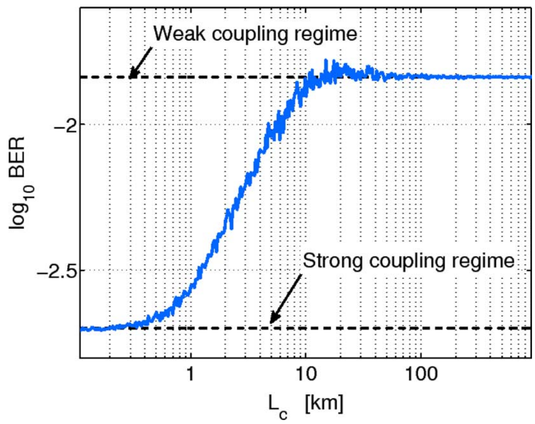


Fig. 5. BER (averaged over the cores) as a function of coupling length L_c when 114-Gb/s PDM-QPSK bit streams are transmitted over 1000 km (solid curve). Dashed lines show the limiting BER in the weak- and strong-coupling regimes.

nonlinear propagation equation is given by (6), provided we add the linear-coupling term $\imath \sum_m q_{mp}\mathbf{A}_m$ on the right side of this equation. The coupling coefficients are defined as

$$q_{mp} = \frac{k_0^2}{2\beta_0(\mathrm{I}_m\mathrm{I}_p)^{1/2}} \iint \left(n^2 - n_m^2\right) \mathrm{F}_m\mathrm{F}_p^* \, dx\, dy, \quad (40)$$

where $n_m(x, y)$ is the spatial distribution of the refractive index of the isolated cores in which the spatial mode m propagates and β_0 is the propagation constant.

Equation (40) is obtained using a coupled-core approach, which assumes that spatial distribution of the electric field propagating in one core, is not perturbed by the presence of other cores [24]. Because of the symmetry of the three-core fiber considered here, all coupling coefficients are identical and we can replace q_{mp} with a single parameter q. Each core of the fiber has a core diameter of 8.2 μm, with $\Delta = 0.003$ and $\gamma = 1.4$ W^{-1}km^{-1}. With these design parameters, intermodal nonlinear coefficients are negligible [25]. We modify the strength of linear coupling by adjusting the distance between fiber cores. It is useful to present the results in terms of the coupling length $L_c = \pi/(3q)$, which represents the distance at which maximum linear power transfer occurs among the cores. A rapidly varying birefringence in each core is included by applying (8) and changing the rotation matrix randomly after each 100 m step. These intra-core perturbations affect the linear coupling such that the exchange of power between the cores becomes stochastic.

Fig. 5 shows the BER as a function of coupling length (solid blue curve) when 114-Gb/s independent random bit streams are simultaneously transmitted through all three cores. The limiting BERs in the weak- and strong-coupling regimes are shown for comparison by dashed black lines. The injected power per core is 8 dBm, and the OSNR at the receiver is set to 20 dB in all cases.

When coupling length is comparable to the length scale L_f of birefringence fluctuations (here 0.1 km), there is a large performance improvement compared to the single-core case. This improvement can be understood by noting that a rapid power transfer among various modes induced by a strong linear coupling and birefringence-induced polarization mixing helps to mitigate nonlinear effects. It can be observed that performance is identical to that of the strong-coupling regime. When the coupling length becomes much larger than L_f, the effects of linear coupling are averaged out by rapidly varying birefringence, and the nonlinear propagation in each core of the 3-core fiber behaves similarly to that in single-core fiber [25]. This occurs in Fig. 5 for $L_c > 10$ km and the BER approaches the limit set in the weak-coupling regime.

The main point of Fig. 5 is that the new Manakov equations derived in this paper can be used with success for multicore fibers in certain cases. When cores are far enough apart that $L_c > 100L_f$, Manakov equations in the weak-coupling regime are valid. When $L_c \lesssim L_f$, Manakov equations in the strong-coupling regime can be used. These results also suggest that in the case of multimode fibers, Manakov equations (26) are also valid when the coupling between spatial modes, resulting from fiber imperfections, occurs on a length scale 100 times larger than that of birefringence fluctuations and that (38) is valid when they occur on the same or shorter length scale.

VII. Conclusion

In this paper we have focused on SDM systems using multimode or multicore fibers and derived a set of nonlinear propagation equations satisfied by the bit stream transmitted through each optical mode in the presence of fiber dispersion, random birefringence, and nonlinearity. We first expressed these equations in Jones vector form in the slowly varying envelope approximation so that both polarization components of each mode can be treated simultaneously and we then used them to investigate the performance of SDM systems designed with multimode fibers exhibiting modal birefringence that varies randomly on a length scale smaller than the length over which nonlinear effects become important. The effects of fluctuating birefringence were included through random Jones matrices. As expected, intermodal nonlinearities are present and can degrade system performance for multimode fibers. There are as many as $M^3 - 1$ intermodal nonlinear terms for an M-mode fiber.

Since numerical simulations based on such stochastic nonlinear propagation equations are generally computationally intensive, we generalized the Manakov equations, well known in the case of single-mode fibers, to the case of multimode fibers. We considered the weak linear coupling case first and averaged over birefringence fluctuations following a procedure similar to that used in the original derivation of the Manakov equations for single-mode fibers. We found that averaging over birefringence fluctuations reduces the $M^3 - 1$ intermodal nonlinear terms to only $M - 1$ terms because all FWM-type terms disappear. The XPM-type terms remain but their effectiveness is reduced because the well-known factor of 2 is reduced to 4/3.

We verified the validity of our generalized Manakov equations by simulating the transmission of multiple 114-Gb/s bit streams in the PDM-QPSK format over different modes of a multimode fiber and comparing the numerical results with those obtained by solving the full set of stochastic equations explicitly. The agreement between the two methods was excellent in all cases studied. We found that the use of the generalized Manakov equations can reduce computation time by at least a factor of 10 and by a factor in excess of 100 as the number of modes grows. Our numerical results show that birefringence fluctuations improve system performance by reducing the impact of fiber nonlinearities. The extent of improvement depends on the fiber design and how many modes are used for SDM transmission. We investigated both the step-index and graded-index multimode fibers and considered up to three spatial modes in each case.

We extended our theory to the case of strong random coupling among various spatial modes of a multimode fiber and derive generalized Manakov equations for this regime. We showed though numerical simulations that the strong-coupling regime leads to better system performance than the weak-coupling regime when the number of spatial modes is at least two. The theory also predicts that performance can improve for a large number of co-propagating spatial modes.

Finally we studied the case of multicore fibers and discussed the validity of the generalized Manakov equations by studying the impact of linear coupling among various cores. We found that the Manakov equations in the weak-coupling regime are valid when the length scale of the random birefringence of each core is about 100 times smaller than the linear coupling length among the cores of a multicore fiber. We also found that the Manakov equations in the strong-coupling regime are valid when the length scale of the random birefringence is nearly the same as the linear coupling length.

REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[2] R.-J. Essiambre, "Impact of fiber parameters on nonlinear fiber capacity," in *Proc. Opt. Fiber Commun. Conf. Expo./Nat. Fiber Opt. Eng. Conf.*, Mar. 2011, OTuJ1.

[3] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudo-linear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2010–2024, 2012.

[4] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, 2010.

[5] R.-J. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks," *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, May 2012.

[6] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle, "Space-division multiplexing over 10 km of three-mode fiber using coherent 6×6 MIMO processing," in *Proc. Opt. Fiber Commun. Conf. Expo./Nat. Fiber Opt. Eng. Conf.*, Mar. 2011, pp. 1–3.

[7] E. Ip, N. Bai, Y.-K. Huang, E. Mateo, F. Yaman, M.-J. Li, S. Bickham, S. Ten, J. Linares, C. Montero, V. Moreno, X. Prieto, V. Tse, K. Chung, A. Lau, H.-Y. Tam, C. Lu, Y. Luo, G.-D. Peng, and G. Li, "$88 \times 3 \times 112$-Gb/s WDM transmission over 50 km of three-mode fiber with inline few-mode fiber amplifier," in *Proc. 37th Eur. Conf. Opt. Commun.*, Sep. 2011, pp. 1–3.

[8] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012.

[9] C. Koebele, M. Salsi, L. Milord, R. Ryf, C. Bolle, P. Sillard, S. Bigo, and G. Charlet, "40 km transmission of five mode division multiplexed data streams at 100 Gb/s with low MIMO-DSP complexity," in *Proc. 37th Eur. Conf. Opt. Commun.*, Sep. 2011, pp. 1–3.

[10] F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," *J. Opt. Soc. Am. B*, vol. 25, pp. 1645–1654, Sep. 2008.

[11] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, 1997.

[12] D. Marcuse, C. R. Manyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, 1997.

[13] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Sov. Phys. JETP*, vol. 38, pp. 248–253, 1974.

[14] K. Ho and J. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *J. Lightw. Technol.*, vol. 29, no. 21, pp. 3119–3128, Nov. 2011.

[15] K. Ho and J. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Express*, vol. 19, pp. 16612–16635, 2011.

[16] A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multi-mode fiber in the strong coupling regime," *Opt. Express*, vol. 20, pp. 11673–11678, 2012.

[17] A. Ghatak and K. Thyagarajan, *Introduction to Fiber Optics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[18] R. W. Boyd, *Nonlinear Optics*, 2nd ed. New York: Academic Press, 2003.

[19] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. San Diego, CA: Academic, 2007.

[20] S. G. Evangelides, L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightw. Technol.*, vol. 10, no. 1, pp. 28–35, 1992.

[21] R. Ryf, R.-J. Essiambre, S. Randel, A. H. Gnauck, P. J. Winzer, T. Hayashi, T. Taru, and T. Sasaki, "MIMO-based crosstalk suppression in spatially multiplexed 3 56-Gb/s PDM-QPSK signals for strongly coupled three-core fiber," *IEEE Photon. Technol. Lett.*, vol. 23, no. 20, pp. 1469–1471, Oct. 2011.

[22] C. Koebele, M. Salsi, G. Charlet, and S. Bigo, "Nonlinear effects in mode-division-multiplexed transmission over few-mode optical fiber," *IEEE Photon. Technol. Lett.*, vol. 23, no. 18, pp. 1316–1318, Sep. 2011.

[23] R.-J. Essiambre, R. Ryf, M. A. Mestre, A. H. Gnauck, R. W. Tkach, A. R. Chraplyvy, S. Randel, Y. Sun, X. Jiang, and R. Lingle, Jr, "Intermodal nonlinear interactions between well separated channels in spatially-multiplexed fiber transmission," in *Proc. 38th Eur. Conf. Opt. Commun.*, Sep. 2012, paper Tu.1.C.4.

[24] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd ed. New York: Academic, 1991.

[25] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Reduction of nonlinear penalties due to linear coupling in multicore optical fibers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 18, pp. 1574–1576, 2012.

[26] M. B. Shemirani, W. Mao, R. A. Panicker, and J. M. Kahn, "Principal modes in graded-index multimode fiber in presence of spatial- and polarization-mode coupling," *J. Lightw. Technol.*, vol. 27, no. 10, pp. 1248–1261, May 2009.

[27] F. Mezzadri, "How to generate random matrices from the classical compact groups," *Notices of AMS*, vol. 54, no. 5, pp. 592–604, May 2007.

[28] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

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