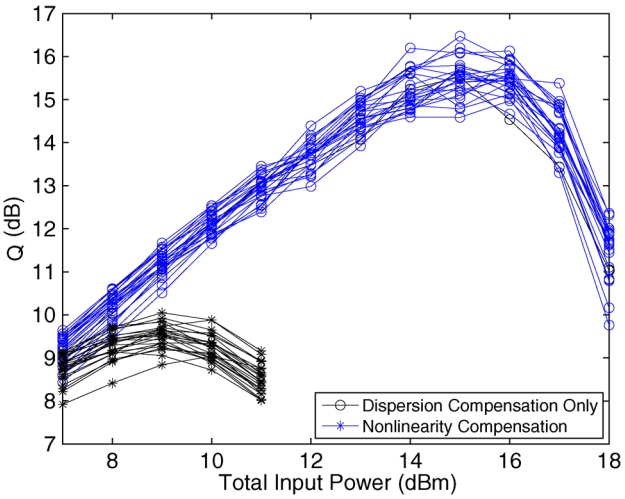
IEEE Photonics Journal
An IEEE Photonics Society Publication

Invited Paper

Nonlinear Impairment Compensation for Polarization-Division Multiplexed WDM Transmission Using Digital Backward Propagation

Volume 2, Number 5, October 2010

Fatih Yaman
Guifang Li, Senior Member, IEEE



DOI: 10.1109/JPHOT.2010.2071860
1943-0655/$26.00 ©2010 IEEE

Nonlinear Impairment Compensation for Polarization-Division Multiplexed WDM Transmission Using Digital Backward Propagation

Fatih Yaman[1,2] **and Guifang Li,**[2] ***Senior Member, IEEE***

(Invited Paper)

[1]NEC Laboratories America, Inc., Princeton, NJ 08540 USA
[2]College of Optics & Photonics/CREOL&FPCE, University of Central Florida, Orlando, FL 32816-2700 USA

DOI: 10.1109/JPHOT.2010.2071860
1943-0655/$26.00 ©2010 IEEE

Manuscript received August 13, 2010; accepted August 24, 2010. Date of publication August 30, 2010; date of current version September 17, 2010. Corresponding author: F. Yaman (e-mail: fyaman@nec-labs.com).

Abstract: A comprehensive treatment of digital backward propagation (DBP) accounting for the vectorial nature of optical transmission is presented. Experimental results show that self-phase and cross-phase modulation are the major sources of nonlinear impairments, even for small channel spacings and for transmission in low dispersion fibers. It is verified that compensating only the incoherent nonlinear impairments not only has the advantage of requiring lower computational load but also removes the necessity of using phase-locked carriers for the signal or phase-locked local oscillators. Simulation results show that polarization-mode dispersion has to be taken into account for practical wavelength division multiplexing systems for DBP to work properly. It is found that to compensate interchannel nonlinear impairments, the changes in the polarization states of channels have to be followed at every span.

Index Terms: Nonlinear impairments, digital signal processing.

1. Introduction

Optical signals suffer linear and nonlinear impairments during transmission in optical fiber, which limits the transmission capacity [1]–[6]. With the introduction of wavelength division multiplexing (WDM) to increase fiber capacity, it became clear that not only dispersion, but fiber nonlinearity as well, could significantly degrade signal quality [7]. Therefore, mitigating or compensating these impairments is a significant research component for optical communication. As the methods for compensating linear impairments including dispersion and polarization-mode dispersion (PMD) matured [8]–[10], nonlinear impairments became the limitation on the fiber capacity. Increasing the number of WDM channels or decreasing the channel spacing results in higher nonlinear penalties. Increasing the spectral efficiency by using higher order modulation formats and increasing the transmission distance require increasing the optical signal-to-noise ratio (OSNR). Increasing OSNR for a given noise level requires increased signal power, which in turn produces more severe nonlinear impairments [4]. Therefore, reducing or compensating nonlinear impairments can significantly increase the fiber capacity.

Several methods have been used to reduce nonlinear impairments, such as optimizing fiber dispersion and dispersion maps [2], [11]–[13], using new modulation formats [14]–[18], polarization interleaving [19]–[21], advanced amplification methods [15], optical phase conjugation [22], [23], and nonlinear post processing [24], [25]. With the exception of optical phase conjugation and

nonlinear post processing, all of these methods have been aimed at designing a more nonlinearity tolerant system rather than removing the nonlinear impairments.

In addition to the above methods, which are all optical in nature, digital compensation methods have been proposed recently. Initially, lumped nonlinear post compensation was proposed based on the assumption that dispersion was low [26]–[29]. There are many reasons for this shift as digital compensation methods are much more flexible and less costly in comparison. Thanks to the recent developments in coherent optical communication [30]–[32], the complete electric field of the signal became accessible in the digital domain. With access to the complete electric field it became possible to compensate nonlinear impairments even when fiber has large dispersion so that the changes in the signal waveform along the fiber cannot be neglected [33]–[47]. Moreover, it became possible to compensate interchannel nonlinear impairments, such as cross-phase modulation (XPM) and four-wave mixing (FWM), as long as the neighboring channels are also received. Among these techniques, the digital backward propagation (DBP) method, which is based on solving the nonlinear propagation equation in the backward direction starting with the received signal as the input and producing the signal at the transmitter at its output, proved to be quite promising.

Compensation of nonlinear impairments in WDM systems using DBP is technically challenging. In WDM systems, where the system as a whole experiences considerable amount of nonlinearity and dispersion, it is impossible to compensate the nonlinear impairments in a single step. Rather, the electric field of all WDM channels as a whole must be propagated backward along the fiber. This requires significantly more computation compared with compensating for only linear impairments [36], [38], [39]. Therefore, for digital techniques to be practical in compensating nonlinear impairments, efficient techniques have to be developed that can still compensate nonlinear impairments but, at the same time, require an affordable computation load. As a way of reducing the number of computations required for DBP, it has been suggested that only the impact of incoherent nonlinear processes including self-phase modulation (SPM) and XPM be compensated [38], [41].

In addition to requiring large amounts of computational load, the DBP technique also has hardware requirements at the receiver. The nonlinear processes such as FWM can be extremely phase sensitive, meaning that when WDM channels are received, their phases in relation to other channels must be preserved. This is extremely difficult to achieve because in practice different WDM channels have different local oscillators. To retain the phase relation between different channels, these local oscillators must be locked in phase.

As the DBP technique relies on reproducing the transmitted signal along the fiber faithfully until reaching the transmitter, it is crucial to know how this technique performs in the presence of random perturbations along the fiber, such as those caused by PMD [48]. As a result of PMD, polarizations of different WDM channels rotate differently [49]. Since nonlinear processes depend on the polarizations of the individual channels, these rotations are expected to affect the results obtained by DBP. The effect of PMD on DBP has yet to be studied.

In this paper, we provide a comprehensive treatment of DBP for polarization-division multiplexed (PDM) WDM systems and in the presence of PMD. In Section 2, the governing equation for vectorial DBP, namely, the computationally efficient Manakov equation, is presented, together with its coupled-mode formalism, which allows selective compensation of individual nonlinear effects. In Section 3, the coupled Manakov equations are employed to perform DBP for a PDM WDM transmission system. It is shown experimentally that indeed, compensating for only SPM and XPM can increase the signal quality significantly, even for small WDM channel spacing. It is also verified by the experimental results that in this case, there is no need to phase lock the WDM carriers or the local oscillators. In Section 4, the impact of PMD on DBP is studied by simulating a PDM WDM system. It is found that even low levels of PMD significantly impair the DBP technique. However, by monitoring PMD on a per-span basis and incorporating PMD into DBP, DBP can still remain effective in the presence of PMD.

2. Manakov Equation

In a coherent optical communication system, the complex field of each WDM channel including amplitude, phase, and state of polarization (SOP) can be fully recovered at the receiver. The complete

knowledge of the received signal makes it possible to calculate the signal at the transmitter; distortions caused by the fiber can be removed from the received signal. To recover the signal at the transmitter, the evolution of the signal through the fiber should be calculated. The evolution of the electric field in the fiber can be described by the vectorial nonlinear Schrodinger equation [50], [51]. Because of the random residual birefringence in the fibers, a simpler equation called the Manakov equation can be used instead of the vectorial nonlinear Schrodinger equation. The Manakov equation can be written as follows [52], [53]:

$$\frac{\partial A_x}{\partial z} = -\frac{\alpha}{2}A_x + i\sum_{n>1}\frac{i^n\beta_n}{n!}\frac{\partial^n}{\partial t^n}A_x + i\gamma_m\left(|A_x|^2 + |A_y|^2\right)A_x$$
$$\frac{\partial A_y}{\partial z} = -\frac{\alpha}{2}A_y + i\sum_{n>1}\frac{i^n\beta_n}{n!}\frac{\partial^n}{\partial t^n}A_y + i\gamma_m\left(|A_y|^2 + |A_x|^2\right)A_y \quad (1)$$

where A_x and A_y are the two polarization components of the total electric field, α is the loss coefficient, β_n is the nth-order dispersion parameter, $\gamma_m = 8\gamma/9$, and γ is the nonlinearity parameter. The Manakov equation is invariant under polarization rotations, and it is written in the x- and y-polarization basis only for convenience.

The Manakov equation describes the evolution of the total electric field according to fiber dispersion and nonlinearity. In the absence of nonlinearity, dispersion modifies the optical waveform in time domain according to the spectral content of the signal. In the absence of dispersion, nonlinearity induces a time dependent chirp, which modifies the spectrum of the total field. The time-dependent chirp is proportional to the sum of the optical power in both polarizations, as given by the last term in (1). When neither dispersion nor nonlinearity can be neglected, the electric field changes along the fiber as a result of the interplay between the dispersive and nonlinear effects, and these two effects cannot be separated anymore.

2.1. Coupled Manakov Equations

The Manakov equation and the description of the nonlinear process above include all the intrachannel as well as interchannel nonlinear interactions as the equation pertains to the whole field, including all the WDM and polarization channels. An alternative to (1) is to write the evolution of each individual WDM channels coupled through the interchannel nonlinear interactions, such as XPM and FWM, as follows:

$$\begin{aligned}\frac{\partial A_{i(p)}}{\partial z} = &-\frac{\alpha}{2}A_{i(p)} + i\sum_{n>0}\frac{i^n\beta_{in}}{n!}\frac{\partial^n}{\partial t^n}A_{i(p)} + i\gamma_m\left(\left|A_{i(p)}\right|^2 + \left|A_{i(|p-1|)}\right|^2\right)A_{i(p)}\\
&+ i\gamma_m\left(\sum_{j\neq i}2\left|A_{j(p)}\right|^2 + \left|A_{j(|p-1|)}\right|^2\right)A_{i(p)} + i\gamma_m\left(\sum_{j\neq i}A^*_{j(|p-1|)}A_{j(p)}\right)A_{i(|p-1|)}\\
&+ i\gamma_m\sum_{q,s,t}\sum_{k\neq l,m}\sum_{l\neq i}\sum_{m\neq i}A^*_{k(q)}A_{l(s)}A_{m(t)}\delta(p+q-s-t)\\
&\times\delta(\omega_i+\omega_k-\omega_l-\omega_m)\exp(i\Delta k_{i,k,l,m}z) \quad (2)\end{aligned}$$

where $A_{i(p)}$ is the electric field amplitude for the ith WDM channel associated with the carrier frequency ω_i, and the subscript in the parenthesis takes on values of 0 or 1 to denote the x or y polarization component, respectively. β_{in} is the nth dispersion coefficient experienced by the ith channel. In other words, it is the nth coefficient of the Taylor expansion of the propagation constant $\beta(\omega)$ expanded around the carrier frequency of the channel ω_i. Here, in deriving (2), it is assumed that the total electric field is composed of the WDM channels as follows:

$$A_x = \sum_i A_{i(0)}\exp[-i\beta(\omega_i)z - i\omega_i t]$$
$$A_y = \sum_i A_{i(1)}\exp[-i\beta(\omega_i)z - i\omega_i t]. \quad (3)$$

In writing the coupled equations (CEs) in (2), it is possible to identify the individual terms responsible for different nonlinear processes. The third term on the right-hand side of (2) is the SPM term. Up to the third term, this equation is exactly the same as (1), which is simply the Manakov equation written for a single channel. The rest of the nonlinear terms are responsible for the interchannel nonlinear effects. The fourth term is the XPM term, which is similar in form to the XPM term that occurs in the scalar case where all the channels have the same polarization. This term induces nonlinear chirp proportional to the power of the neighboring channels, which does not depend on their phases. Because of the absence of the explicit dependence on the phases of the neighboring channels, we will call this term the incoherent XPM (IncXPM). The IncXPM is twice as strong for the channels with the same polarization compared with the orthogonal polarization. The fifth term is also a XPM term in the sense that it involves only two WDM channels. Unlike the IncXPM term, these terms are not completely independent of the phases of the interacting fields. Phase changes originating from processes such as PMD where the phase depends on both the polarization and frequency affects these terms directly. On the other hand, the phase that is common to both polarizations such as group-velocity dispersion or that is common to both WDM channels such as birefringence cancels out. This term can also be viewed in terms of the FWM process between two WDM channels. As a result of this process, one photon from *x* polarization of the first channel and one photon from the y polarization of the second channel are transferred to the *y* polarization component of the first channel and the *x* polarization of the second channel or *vice versa*. This term is also partially responsible for the polarization dependence of XPM.

The final term is the FWM term, which can cause transfer of power between different channels. In the notation of (2), the channels A_i, A_k, A_l, and A_m are coupled through FWM process, as long as their frequencies satisfy the condition $\omega_i + \omega_k - \omega_l - \omega_m = 0$, and their polarization states are such that if A_i and A_k are parallel, A_l, A_m are also parallel to them, and if A_i and A_k are orthogonal to one another, A_l, A_m are also orthogonal to one another. As a result of FWM an energy transfer occurs between the A_i, A_k pair and A_l, A_m pair. The magnitude and the direction of the energy transfer are determined by the phase relation between the four interacting channels. If the system is not working in the extremely high nonlinear regime, the FWM requires a certain length to accumulate enough energy transfer between the channels to cause significant amount of distortion. If the phase relation between the channels is not maintained for long distances in the fiber, distortions caused by FWM do not grow, and it can be neglected. The condition for maintaining the phase relation between the channels is called the phase matching condition, and it is given by $\Delta k_{i,k,l,m} + \theta_{NL} = 0$. $\Delta k_{i,k,l,m}$ and θ_{NL} represent the change in the phase relation between the channels per unit length as a result of fiber dispersion and accumulated nonlinear phase, respectively.

For a given WDM system with certain dispersion parameters, channel spacing, nonlinearity coefficient, and power per channel, it is possible to estimate whether FWM is negligible or not. In WDM systems with equal channel spacing, FWM is expected to be strongest between neighboring channels. In other words, if dispersion is strong enough to suppress FWM between neighboring channels, it is strong enough to suppress FWM between channels that are farther apart. For FWM effects to grow considerably, the phase relation between the interacting channels must be maintained for a length scale comparable with the nonlinear length given by $L_{NL} = 1/(\gamma_e P_{ch})$, where $\gamma_e = \gamma_m[1 - \exp(-\alpha L_{sp})]/\alpha$ is the effective nonlinear parameter, and L_{sp} is the span length. Here, nonlinear length is defined for power per channel to estimate FWM interactions between only the neighboring channels. The change in the phase relation between the interacting channels in a single nonlinear length should be small so that $\Delta k L_{NL} \ll \pi$. For a given dispersion and channel spacing, this condition becomes $\beta_2 \Delta\omega^2/(\gamma_e P_{ch}) \ll \pi$. This condition can be used as a rule of thumb for the simplest cases as several assumptions have been made, for instance, the effect of polarization is ignored, as well as the possibility of efficient FWM coupling between farther channels through modulational instability.

As given in (2), the CEs leave out some of the second-order processes captured by the full Manakov equation (1). For instance, (2) cannot account for energy transfer from WDM channels outside of the WDM bandwidth. To capture those effects, extra channels beyond the WDM band should be added to

the CEs. CEs also neglect the energy transfer through FWM to the parts of the spectrum that falls between the WDM channels.

XPM terms do not depend on the phases of the interacting channels. However, they are still sensitive to the SOPs of the channels. As a major source of nonlinear impairment, it is crucial to understand the nature of this polarization dependence. XPM can be divided into two parts: one part that is independent of the SOPs and one part that is completely polarization dependent. To see this distinction explicitly the CEs are presented below in the vector notation, and FWM terms are omitted:

$$\frac{\partial|A_i\rangle}{\partial z} = \mathcal{L}(\omega_i)|A_i\rangle + i\gamma_m\langle A_i|A_i\rangle|A_i\rangle + i\gamma_m\sum_{j\neq i}\left(\langle A_j|A_j\rangle + |A_j\rangle\langle A_j|\right)|A_i\rangle \tag{4}$$

where $|A_i\rangle = [A_{i(x)}\ A_{i(y)}]^T$ is the amplitude of the ith WDM channel in the Jones vector notation, and $\mathcal{L}(\omega_i)$ represents the linear terms in (2), including the loss and dispersion terms. It can be seen by looking at (4) that the SPM term and the first term of the XPM term have no polarization dependence. The second term in the XPM term is a matrix with nonzero trace, meaning that it has a component which has polarization dependence. This term can be separated into the polarization dependent and independent parts by using the Pauli matrices as follows: $|A_j\rangle\langle A_j| = [P_j + \vec{\boldsymbol{P}}_j\cdot\vec{\sigma}]/2$, $P_j = \langle A_j|A_j\rangle$, and $\vec{\boldsymbol{P}}_j$ are the optical power and the Stokes vector of the jth channel, $\vec{\sigma} = [\sigma_1, \sigma_2, \sigma_3]$ is the Pauli spin vector, and σ_1, σ_2, and σ_3 are the Pauli spin matrices [49], [50]. Equation (4) can now be separated into polarization independent and dependent parts as

$$\frac{\partial|A_i\rangle}{\partial z} = \mathcal{L}(\omega_i)|A_i\rangle + i\gamma_m P_i|A_i\rangle + \frac{3i\gamma_m}{2}\sum_{j\neq i}P_j|A_i\rangle + \frac{i\gamma_m}{2}\sum_{j\neq i}\vec{\boldsymbol{P}}_j\cdot\vec{\sigma}|A_i\rangle. \tag{5}$$

In (5), the first term in the second line is the part of XPM that do not depend on the SOPs of the interacting channels. This term induces phase shift that is proportional to the total powers of WDM channels in both polarizations. The last term, however, is completely dependent on the SOPs of the interacting channels. Similar to the polarization independent part, all channels also contribute to the polarization dependent part. The polarization dependent part has two distinct contributions. First, just as the polarization independent part of XPM, it contributes to the nonlinear phase shift. Second, it causes nonlinear rotation of the channel. To see these contributions more clearly, (5) can be put into the following form:

$$\frac{\partial|A_i\rangle}{\partial z} = \mathcal{L}(\omega_i)|A_i\rangle - i\gamma_m P_i|A_i\rangle + \frac{3i\gamma_m}{2}P_T|A_i\rangle + \frac{i\gamma_m}{2}\hat{\boldsymbol{s}}_i\cdot\vec{\boldsymbol{P}}_T|A_i\rangle + \frac{\gamma_m}{2}(\hat{\boldsymbol{s}}_i\times\vec{\boldsymbol{P}}_T)\cdot\vec{\sigma}|A_i\rangle \tag{6}$$

where $P_T = \sum_i P_i$ is the total power in all the WDM channels, $\vec{\boldsymbol{P}}_T = \sum_i \vec{\boldsymbol{P}}_i$ is the sum of the Stokes vectors of all the channels, e.g., the collective vector, and $\hat{\boldsymbol{s}}_i$ is the unit Stokes vector of the ith channel. As it can be seen in (6), the polarization dependent term is split into two parts. The first part contributes only to the nonlinear phase shift. This term changes between the extremes $\pm\gamma_m P_T/2$, depending on the relative orientation of the channel with respect to the collective vector at a given point in time. The second term, on the other hand, causes nonlinear rotation of the ith channel around the collective vector at a rate proportional to the part of the collective vector that is orthogonal to the ith channel. Note that the cross product is maximum when the Stokes vectors are orthogonal, which occurs when the ith channel and the collective vector are at 45° in Jones space. Note also that the polarization independent part of XPM is proportional to the sum of the powers of all the channels. However, the polarization dependent part is proportional to the sum of the Stokes vectors of all the channels. In the extreme case when all the channels have the same SOP, the two parts of XPM are similar in magnitude. Nevertheless, in a PDM WDM system with a large number of channels, this term is expected to be smaller compared with the polarization independent portion because of the vectorial sum.

As long as the signal at the receiver is obtained with adequate accuracy, the signal at the transmitter can be obtained by numerically solving either the full Manakov equation or the CEs. These equations can be solved in the backward direction, and the received signal is used as the initial condition. The most common numerical method for solving the nonlinear propagation equations is the split-step

Fourier method. In the split step method the fiber is divided into steps. As an approximation, only the nonlinear part or the dispersive part of the equations is solved in each step. The step size should be small enough to follow changes in the propagating field that are caused by both dispersion and nonlinearity. The step size should be small enough so that the changes in the spectrum caused by the fiber nonlinearity could be followed accurately, and therefore, in the next step, the dispersive propagation can be calculated accurately. The step size required by the fiber nonlinearity should be much smaller than the nonlinear length that is defined as $L_{NL} = (\gamma_e P)$, where P is the total power. Step size should also be small enough to follow how the optical fields evolve along the fiber as a result of dispersion so that in the following step, the impact of nonlinear interactions can be followed accurately.

Calculating the dispersive and nonlinear evolution over and over in small steps along the fiber increases the computational load. To avoid unnecessarily increasing the computational load, only changes in the optical field that are relevant to accurate calculation of nonlinear processes should be followed in small steps. For instance, in a transmission system where FWM is strong, step size should be small enough to follow the phases of the interacting fields relative to one another because FWM is sensitive to such changes. However, if FWM is negligible, and XPM is the dominant source of nonlinear distortions, there is no need to follow the changes in the phases of the channels in small steps since XPM is insensitive to relative phases of the channels. In this case, the step size is still limited because XPM is sensitive to the walk off between channels as a result of the group delay. However, it is much larger compared with what is required by the FWM process. It is useful to define the FWM length $L_{\mathrm{FWM}} = 4(\beta_2 N^2 \Delta\omega^2)^{-1}$ and the XPM length $L_{\mathrm{XPM}} = (\beta_2 N \Delta\omega B)^{-1}$, where N is the total number of WDM channels, and B is the baud rate. L_{FWM} is the length scale at which the relative phases of the farthest channels interacting through FWM changes by 1 rad because of dispersion, and L_{XPM} is simply the length at which the farthest WDM channels walk off by one another by a bit period. These length scales should not be understood as the length scales at which FWM or XPM becomes important; rather, they should be understood as the length scales that limit the maximum step size that can be used in the split-step method, which is imposed by the fiber dispersion to calculate FWM or XPM accurately [41].

Comparing the step size requirement imposed by FWM and XPM shows that for a transmission system where the impact of FWM is small, computational load can be decreased by a factor roughly given by the ratio of the FWM and XPM lengths: $L_{\mathrm{FWM}}/L_{\mathrm{XPM}} = N\Delta\omega/B$. It is important to note that, if the FWM is weak and, therefore, the step size is chosen to accommodate the XPM length, the FWM terms should be removed from the calculation. Otherwise, the impact of FWM terms will not be calculated correctly. During the backward propagation, the contribution of FWM can be miscalculated, and it may grow strong, even though FWM is weak in the forward propagation. One of the advantages of using the CEs instead of the full Manakov equation is that in the CEs, FWM terms appear explicitly, and they can be removed from the equation. This is not possible when full Manakov equation is used. Therefore, Manakov equation requires the step-size to be small enough to accurately calculate FWM, even though FWM is very weak and negligible [41].

Another advantage of using the CEs in terms of computational load is that it requires a lower sampling rate. Solving the Manakov equation accurately requires the sampling rate to cover twice the total bandwidth of the WDM system, even though there is no significant energy present outside of the WDM bandwidth. Therefore, the required number of samples per symbol becomes $S_M = 2N\Delta\omega/(2\pi B)$. This is because FWM can couple WDM channels to parts of the spectrum outside of the WDM bandwidth. If the sampling rate is not large enough, FWM causes aliasing leading to errors. CEs, on the other hand, require correct sampling of individual channels, which is just two samples per channel per symbol. Therefore, the number of samples per symbol becomes $S_{\mathrm{CE}} = 2N$. As a result, CEs always require lower number of samples as long as channel spacing is larger than the symbol rate [41].

When the effect of FWM is small and it is negligible, CEs have another significant advantage from a practical point of view. Because FWM is sensitive to the relative phases of the WDM channels, the relative phases of the WDM channels must be accurately measured at the receiver. Otherwise, during the back propagation, the contribution of FWM cannot be calculated accurately. To make sure that the relative phases of the WDM channels are measured correctly, the local oscillators used at the coherent

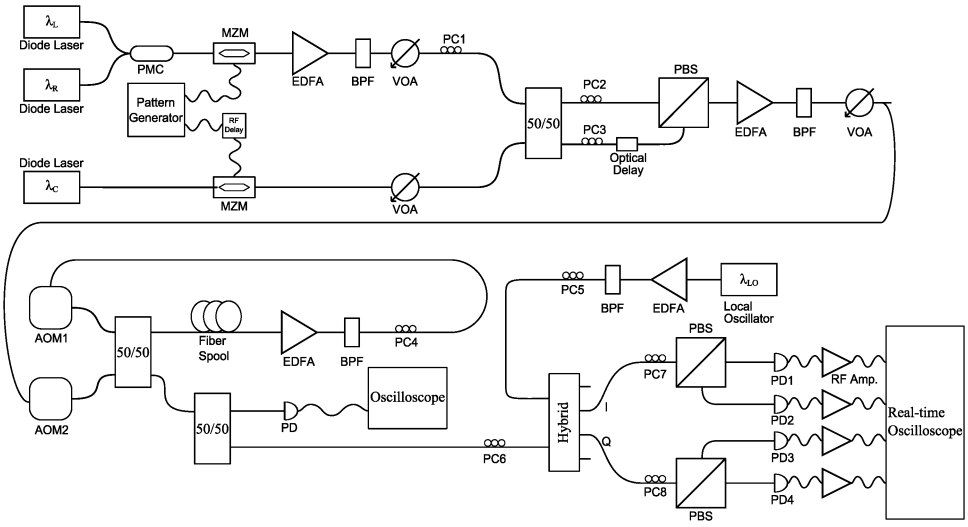


Fig. 1. Experimental setup for a polarization-division multiplexed, three-channel OWDM system consisting of BPSK transmitters, a re-circulating loop, and coherent polarization diversity receivers.

receivers for each WDM channel must be locked in phase. Alternatively, if the nonlinear compensation is implemented at the transmitter using DBP, it is necessary that the carriers of different WDM channels be locked in phase. Satisfying these conditions becomes extremely difficult as the number of the channels is increased. If the impact of FWM is negligible, however, neither the local oscillators nor the carriers need to be locked in phase since XPM does not depend on the phases of the channels [41].

3. Digital Backward Propagation

This section presents experimental results of DBP using the Manakov equations. For completeness and for comparison purposes, the results of DBP using the full Manakov equation will be summarized first [40]. Then, the results of DBP using the coupled Manakov equations will be presented.

3.1. Experimental Setup

The discussion in Section 2 shows that nonlinear impairments are deterministic and that they depend on the signal and fiber parameters. Even with this deterministic nature of the impairments, it is difficult to determine how effective DBP would be in a practical system in the presence of ASE noise, nonzero carrier and local oscillator linewidth, fiber dispersion that is typically not constant along the fiber, receiver imperfections such as detector and analog to digital converter (DAC) bandwidth, imbalance and delay in the I and Q arms, as well as polarization arms and DAC quantization. In a recent experiment, the possibility of compensating the nonlinear impairments experienced in a PDM WDM system was demonstrated by using DBP based on solving the full Manakov equation. In this section, this experiment and its main results will be summarized first, and then, it will be compared with the proposed XPM only compensation.

The experimental set up is shown in Fig. 1. At the transmitter side, three distributed feedback lasers are used as the carriers. The same pseudo random BPSK pattern with a length of $2^{23} - 1$ is used to modulate the side channels at 6 GBd/s, whereas the center channel is modulated with the delayed version of the same pattern. Polarization multiplexing is achieved by combining the delayed copies of the combined channels in the orthogonal polarization state by using a polarizing beam splitter.

The loop consists of a single span of nonzero dispersion shifted fiber, a single EDFA to compensate for the total loss of the loop, a band pass filter, and a polarization controller. Two acoustooptic modulators are used as optical switches. The fiber is 80 km long, with second-order dispersion parameter of $\beta_2 = -4.84$ ps^2/km and a nonlinearity parameter of $\gamma = 1.5$ W^{-1}km^{-1}. The band pass

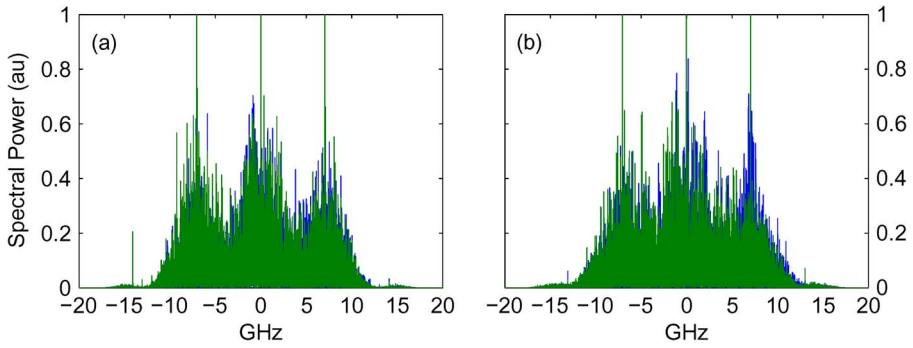


Fig. 2. Spectra of the measured WDM channels (a) right after the transmitter and (b) after 18 loops. The total input power is 6 dBm, and the green and blue curves show the two orthogonal polarization components.

filter has a 1-nm bandwidth. After the loop, a portion of the signal is used to monitor the loss/gain balance in the loop, and the rest is sent to the coherent receiver.

At the receiver, a single local oscillator is tuned to the carrier frequency of the center channel. The received signal and the local oscillator are combined at the 90° hybrid. The two quadratures leaving the hybrid are divided into the two polarization components and detected by four photodiodes, each measuring one polarization component of one quadrature. The electrical signal is amplified by linear radio frequency amplifiers and received by the real-time oscilloscope (RTO). The RTO has an analog bandwidth of 12 GHz, and it samples all four ports at 50 Gsa/s. Since balanced detection is not used, the DC component of the electrical signal is blocked before entering the RTO. For the same reason, the signal power entering the hybrid is attenuated so that the power level of the signal is 23 dB below the local oscillator power. To make sure that the photodetectors operate in the linear regime, the power of the local oscillator is adjusted so that total optical power entering the photodetectors is below 5 dBm.

Orthogonal WDM (OWDM) is used to pack three WDM channels into 25-GHz bandwidth limited by the analog bandwidth of the RTO. Using OWDM, it was possible to fit three channels at 6 Gsym/s Bd rate separated by 7 GHz into the RTO bandwidth. Note that three is the minimum number of WDM channels necessary to see the full effect of both XPM and FWM. An RF delay at the transmitter is used to make sure that the bit slots of all channels are aligned in time, which is a requirement of OWDM. Having only four DACs (four ports of the RTO) also made it necessary to use the same local oscillator for all the channels.

Fig. 2 shows the spectrum of the measured field before and after 1400 km of fiber transmission. Comparison of the spectrum after transmission with the back-to-back case shows how severely spectrum is modified through nonlinear propagation. The major change in the spectrum is due to spectral broadening.

3.2. DBP Using the Full Manakov Equation

When the full Manakov equation [see (1)] is solved in the DBP, the whole electric field is used as the starting point, and the evolution of the electric field is backtracked step by step using the split step method until the transmitter. After the backward propagation, the channel to be recovered is filtered using an appropriate OWDM filter. OWDM filtering is followed by carrier phase estimation, polarization demultiplexing, and, finally, Q value estimation.

After DBP is performed, the Q value is calculated and plotted as a function of transmission distance for both X and Y polarizations shown by circles and stars, respectively, in Fig. 3. The black curves show the Q value obtained when the full Manakov equation is solved in DBP. The red curves in contrast show when only dispersion compensation is used. Clearly, at the 6 dBm of total power and distances beyond 1000 km, dispersion compensation is not enough as the signal accumulates a significant amount of nonlinear impairments. With all the experimental imperfections and uncertainties, DBP can

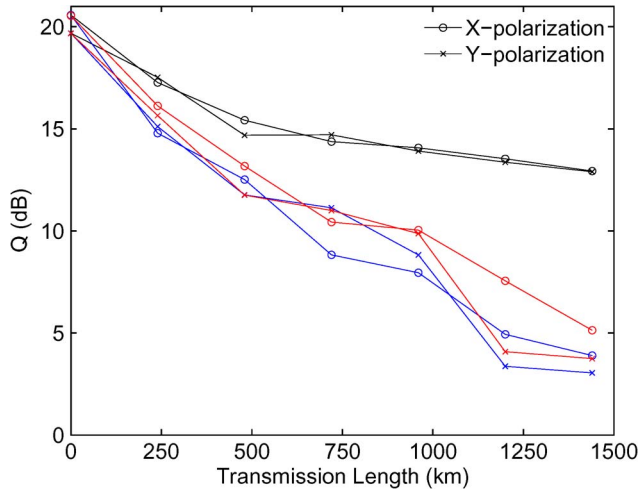


Fig. 3. Q value of the center channel after DBP as a function of transmission length. The black curves show after DBP using full Manakov equation, the blue curves show when the impairment coming from the orthogonal polarization is ignored, and the red curves show when only dispersion compensation is used. The circles and the stars mark the X and Y polarization.

still remove most of the nonlinear impairments successfully. The blue curves are obtained when the impairments resulting from the orthogonal polarization are ignored. For instance, for the *X* polarization, the power in the *Y* polarization is set to zero. In this case, the Q value is even worse than dispersion compensation case, showing that the impairments resulting from the orthogonal polarization is significant.

3.3. DBP Using the Coupled Manakov Equations

To demonstrate the feasibility and the advantage of using CEs and compensating only for XPM and SPM effects, the CEs [see (2)] instead of full Manakov equations are solved in DBP. Note that the CEs are derived based on the assumption that the total field can be decomposed into WDM channels, which can be separated from one another based on their spectrum. When OWDM is used, the channel spacing is so small that individual channels cannot be separated based on their spectrum. Indeed, to access the data carried by each OWDM channel orthogonality of the channels are used and not their separability by spectrum. Nevertheless, the outline of the three channels is still visible, as can be seen in Fig. 2. To use the CEs the spectrum obtained at the receiver is split into three channels before DBP. The digital filters are centered around the carriers of the OWDM channels, and the filter bandwidth is 6 GHz. The spectra of the filtered channels along with the corresponding filters are shown in Fig. 4. After filtering, the channels are brought to the baseband to be used as the input for the CEs, which takes the following form in the case of three WDM channels:

$$\begin{aligned}
\frac{\partial A_{1x}}{\partial z} &= -\frac{\alpha}{2}A_{1x} + i\sum_{n>0}\frac{i^n\beta_n(\omega_1)}{n!}\frac{\partial^n}{\partial t^n}A_{1x} + i\gamma_m\left(|A_{1x}|^2 + |A_{1y}|^2\right)A_{1x} \\
&\quad + i\gamma_m\left(2|A_{2x}|^2 + 2|A_{3x}|^2 + |A_{2y}|^2 + |A_{3y}|^2\right)A_{1x} + i\gamma_m\left(A_{2y}^*A_{2x} + A_{3y}^*A_{3x}\right)A_{1y} \\
&\quad + i\gamma_m\left(A_{2x}^2A_{3x}^*\exp(i\Delta k_{1,3,2,2}z) + A_{2x}A_{2y}A_{3y}^*\right)\exp(iz\Delta k_{1,3,2,2}) \\
\frac{\partial A_{2x}}{\partial z} &= -\frac{\alpha}{2}A_{2x} + i\sum_{n>0}\frac{i^n\beta_n(\omega_2)}{n!}\frac{\partial^n}{\partial t^n}A_{2x} + i\gamma_m\left(|A_{2x}|^2 + |A_{2y}|^2\right)A_{2x} \\
&\quad + i\gamma_m\left(2|A_{1x}|^2 + 2|A_{3x}|^2 + |A_{1y}|^2 + |A_{3y}|^2\right)A_{2x} + i\gamma_m\left(A_{1y}^*A_{2x} + A_{3y}^*A_{3x}\right)A_{2y} \\
&\quad + i\gamma_m\left(2A_{1x}A_{3x}A_{2x}^*\exp(i\Delta k_{1,3,2,2}z) + A_{1x}A_{3y}A_{2y}^* + A_{1y}A_{3x}A_{2y}^*\right)\exp(iz\Delta k_{2,2,1,3})
\end{aligned} \tag{7}$$

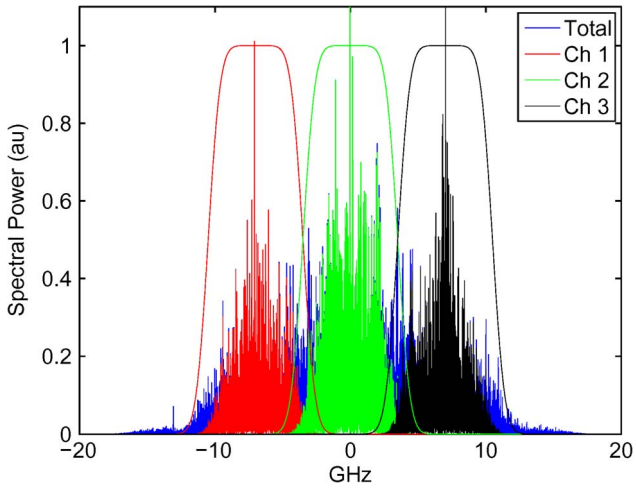


Fig. 4. Spectra of the three channels (red, green, and black) obtained by filtering the spectrum of the optical field measured after 1400 km propagation (blue). The corresponding filters are overlaid on the spectra. Only the x-polarization components are shown for clarity.

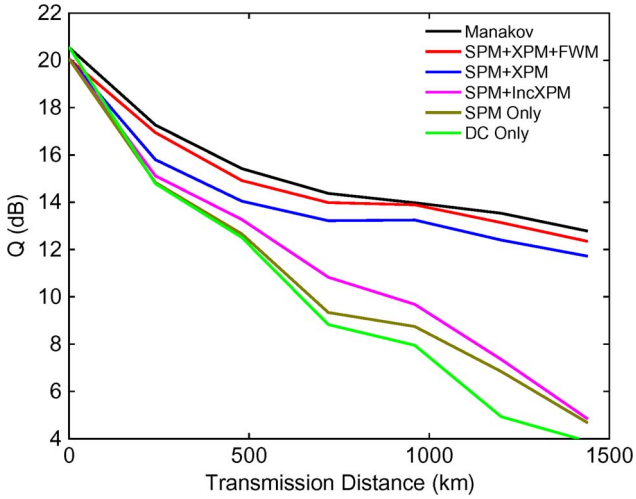


Fig. 5. Q value after DBP as a function of transmission distance. Black: Full Manakov equation. Red: CE with all terms retained. Blue: FWM is neglected. Pink: FWM and CohXPM neglected. Brown: Only the SPM term is retained. Green: Dispersion compensation only.

where A_2 is the electric field of the center channel and A_1 and A_3 are the side channels. The equation for A_3 can be obtained by interchanging subscripts 1 and 3 in the first equation of (7). The equations for the y-polarization component can be found by interchanging the subscripts x and y.

After receiving the electrical field, the CEs are solved in the backward direction instead of the full Manakov equation to remove the nonlinear impairments. After the backward propagation, the channel to be recovered is filtered using an appropriate OWDM filter. In the case of CEs, in order to retain the orthogonality of the channels, the side channels are multiplied by their carriers and added to the center channel before the OWDM filter is applied. OWDM filtering is followed by carrier phase estimation, polarization demultiplexing, and, finally, Q value estimation, as before.

After DBP is performed, the Q value is calculated and plotted as a function of transmission distance in Fig. 5. To compare the efficiency of using the full Manakov equation and to see the effect of omitting different nonlinear processes, six different cases are considered. The black curve shows the Q values

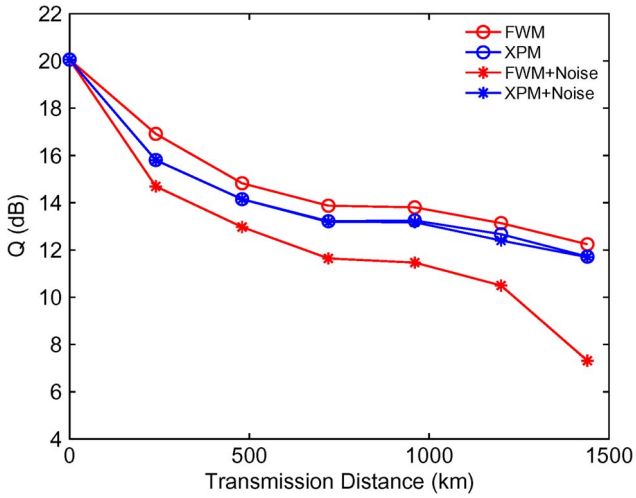


Fig. 6. Q value after DBP as a function of transmission distance. Red and blue lines show when the CEs are solved with and without the FWM terms. Stars show when the phase noise is added to the channels to mimic the use of independent local oscillators, and circles show when the local oscillators are phase locked.

obtained when the full Manakov equation is solved, and the green curve shows when only the dispersion compensation is used. The red curve shows the Q values obtained when the CE of (7) is solved while keeping all the terms. The slight difference between solving the CEs and the Manakov equation can be attributed to the filtering since the same difference in Q value is observed also back to back. The rest of the curves are also obtained by solving the CEs, but in each case, different nonlinear processes are neglected, i.e., removed from (7). The blue curve shows the Q values obtained when the FWM terms are neglected. Comparing it with the red curve shows that the penalty from ignoring the FWM process is less than 1 dB. In this particular setup, this penalty is not significant enough to justify including FWM process in to the calculations. Next, to test how big the contribution of the CohXPM is, this term is omitted from (7), while the IncXPM terms are retained, and the results are shown with the pink curve. Clearly, CohXPM terms produce significant nonlinear impairment, and they should be included in the compensation. The brown curve shows the case when only SPM terms are retained among the nonlinear terms. At high power levels, compensating for only SPM terms produces no advantage over compensating dispersion only.

Fig. 5 shows that it is possible to use CEs instead of Manakov equations and that compensating for only XPM and SPM is enough even when the channel spacing is very low. In WDM systems with larger channel spacings, the difference between the two cases is expected to be even less. However, because the channel spacing is of the same order as the channel bandwidth in the experiment, we were not able to verify the advantage of ignoring the FWM in terms of being able to use larger step size or using lower sampling rate.

The advantage of compensating only for XPM in terms of not needing phase-locked oscillators is not clear from Fig. 5 since only one local oscillator is used for all channels. The most significant effect of using independent local oscillators is that each channel will be received with a phase that is random relative to the others. This can be simulated easily by multiplying each channel by a random phase after the channels are filtered and before they are put into back propagation. Fig. 6 compares the Q values obtained when the CEs are solved in DBP with (red) and without (green) the FWM terms as well as with (circles) and without (stars) the added random phase noise. The Q values obtained with the random phase noise are the worst Q values obtained among 50 different runs. Fig. 6 shows that there is no penalty arising from using independent local oscillators when FWM is dropped from the CEs. However, when FWM is included in the equations, the random phases added by the local oscillators induce a significant Q penalty because in this case, the FWM terms

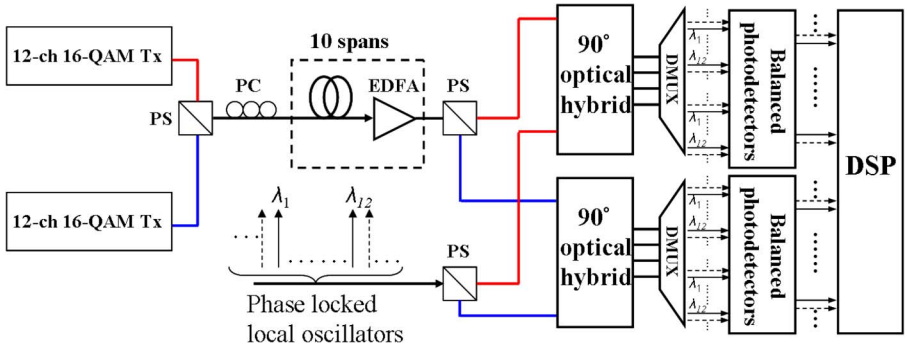


Fig. 7. Schematic of the simulation setup used for forward propagation of a 12-channel 16-QAM PDM WDM system.

are not computed accurately. If independent local oscillators are used, including FWM terms results in worse Q values.

4. Polarization-Mode Dispersion

For DBP to remove the nonlinear impairments effectively, it is necessary that the evolution of the signal from the transmitter to the receiver must be backtracked faithfully. For instance, the evolution of the signal as a result of fiber dispersion must be tracked carefully so that the nonlinear processes can be calculated accurately. Another linear process that occurs in fibers and affects the nonlinear interactions is PMD. PMD is a consequence of residual birefringence present in fibers that varies in magnitude and direction randomly along the fiber. PMD causes two polarization components of the signal to travel with slightly different group velocities [48], [49]. If the delay becomes comparable with the bit period, PMD causes outages. For individual WDM channels, the impact of PMD can be avoided by keeping the baud rate low, or it can be removed using several optical or digital signal processing techniques. However, when nonlinear interactions among the channels are strong, the impact of PMD over several channels has to be considered. From another perspective, the effect of PMD can be described as frequency dependent birefringence. Polarizations of different channels rotate at different rates or direction because of PMD. In a WDM system with many channels, even though the baud rate of individual channels may be small, PMD can still induce random rotations of distant WDM channels with respect to one another and affect the nonlinear interactions among them.

Because the nonlinear processes depend on the relative orientations of the interacting channels, it is necessary to follow these random rotations caused by PMD along the fiber to compensate nonlinear impairments effectively. Just as in the case of dispersive effects, the random rotations caused by PMD should be followed in the backward direction frequently enough that at each step, the amount of relative rotations should be small. Similar to the dispersion length or nonlinear length where these effects are expected to be significantly large, a PMD diffusion length L_{PMD} can be defined that corresponds to the length scale at which the random rotations caused by PMD are large. Because both the magnitude and the direction of residual birefringence fluctuate along the fiber as well as in time, the magnitude and direction of PMD also vary randomly in time. Therefore, the effects of PMD can only be described statistically. $L_{\mathrm{PMD}} = 3/(D_p \Delta\omega)^2$ is the length scale at which two WDM channels separated in frequency by $\Delta\omega$ cannot be expected in an average sense to retain their relative orientations [54], [55]. D_p is the PMD parameter which depends on the properties of a given fiber. Modern fibers typically have PMD parameters lower than 0.5 ps/$\sqrt{\mathrm{km}}$ and can be as low as 0.01 ps/$\sqrt{\mathrm{km}}$ [56].

The impact of PMD on DBP depends on both the fiber PMD parameter and the bandwidth of the WDM system. To see the impact of PMD it is necessary to evaluate a WDM system with much larger bandwidth than the experimental setup. For this purpose, a 12-channel PDM WDM system is simulated using the commercial software VPI. Fig. 7 shows the schematic of the simulation setup. A total of 12 channels per polarization are modulated at 25 GBd, and channel spacing is 50 GHz.

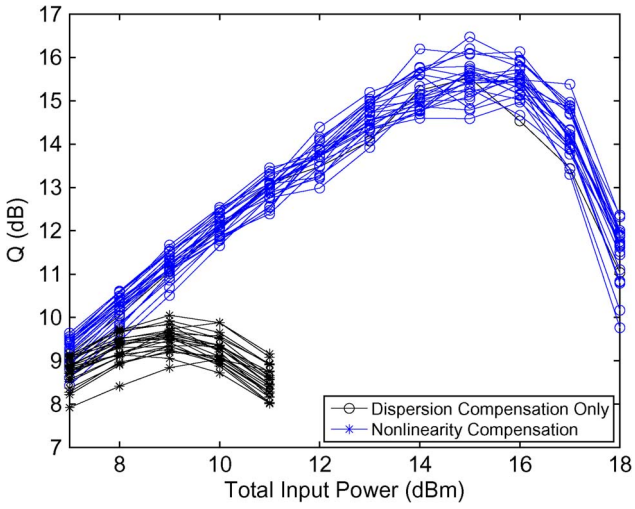


Fig. 8. Q value calculated as a function of total average input power. Blue shows the case when DBP based on Manakov equation is used to compensate the nonlinear impairments. Black curves show the Q values when only dispersion is compensated. Each case consists of 24 curves corresponding to 12 WDM channels in two PDM channels.

16-QAM is chosen as the modulation format. The two polarization channels are combined, and then, their SOPs are adjusted with a polarization controller to be at 45° with the principal state of polarizations (PSPs) of the transmission fiber at the central wavelength. This is the case of the worst alignment which guarantees sampling the worst-case results. The transmission system consists of 10 spans of 100-km-long NZDSF fibers separated by EDFAs with a 5-dB noise figure. The dispersion and dispersion slope of the fibers are 4.4 ps/km/nm and 0.045 ps/km/nm^2. The nonlinear parameter is 1.3 W^{-1}km^{-1}, and loss coefficient is 0.2 dB/km. EDFAs are running in the power mode, and they fully compensate the fiber loss. After forward propagation, the transmitted signals are separated into two polarizations, and each polarization component is mixed with 12 phase-locked local oscillators of the same polarizations in 90° hybrids. After coherent detection, back propagation and the following DSP is performed with Matlab.

After polarization diversity coherent detection, electric fields of all the channels in both polarizations are available. These fields are used as the initial condition to solve either the CEs of (2), or they can be combined to obtain the whole field to be used as the initial condition for the Manakov equation (1). VPI is not used in the backward propagation, and these equations are solved using the split-step Fourier transform method in Matlab.

First, average optical power levels where the transmission is significantly nonlinear are determined by changing the input power in the forward propagation. For each power, the field is back propagated, and the Q value is calculated as a figure of merit. PMD is artificially turned off so that the impact of PMD does not interfere with the determination of the nonlinear regime. In this case, all the channels experience the same fast rotation due to fiber birefringence so that Manakov equations are valid. Fig. 8 shows the Q value calculated for each channel and polarization as a function of total launched power. For the particular system that is simulated, the optimum total power is found to be 15 dBm when DBP is used for impairment compensation. For comparison, the Q values obtained when only dispersion compensation is applied are also shown in Fig. 8. In this case, the optimum power is only 9 dBm, and more importantly, the maximum achievable Q value is 6 dB lower compared with what can be achieved by backward propagation. Impairment compensation by DBP is justified beyond 9 dBm. If the impact of PMD can be taken into account perfectly, the Q value should approach the blue lines and approximately 16 dB at 15 dBm input power.

To see how much PMD can render DBP ineffective, the simulations are repeated by adding PMD in the forward propagation. To obtain statistically meaningful results, simulations are repeated for

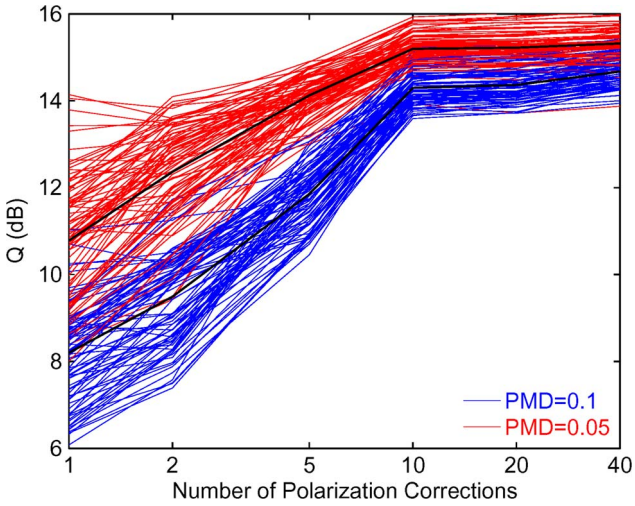


Fig. 9. Q value of the center channel as a function of the number of corrections made to the SOPs of the channels during back propagation. Red and blue curves are for $D_p = 0.05$ and 0.1 ps/$\sqrt{\text{km}}$. Individual curves correspond to different realizations of PMD. The black lines are the average of the 40 random realizations.

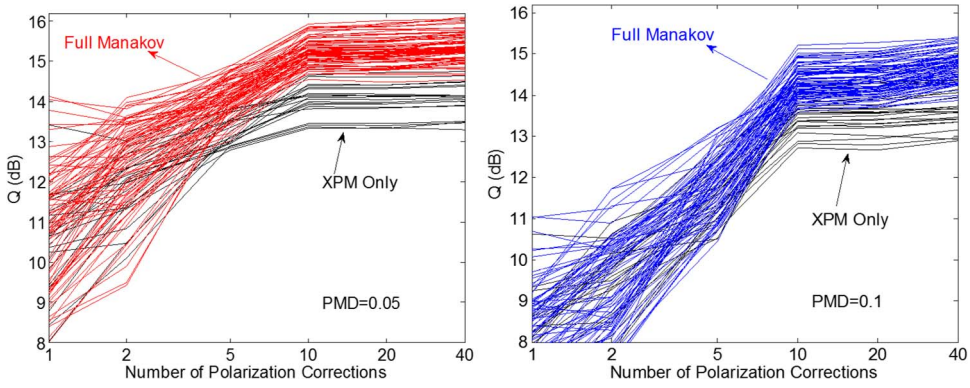


Fig. 10. Q value of the center channel as a function of number of corrections made to the SOPs of the channels during the back propagation for (a) $D_p = 0.05$ ps/$\sqrt{\text{km}}$ and (b) $D_p = 0.1$ ps/$\sqrt{\text{km}}$. The red and blue curves are when Manakov equation is solved, and black curves show the results when CEs are solved and the FWM terms are ignored.

40 random realizations of PMD, while in each case, the input SOP of the center channel is aligned at a 45° angle with the PSPs. To determine how significant the impact of PMD is, as well as to make sure that the deviation from the ideal compensation is due to PMD, the SOPs of individual channels are corrected periodically during the back propagation calculation. The results are shown in Fig. 9. The case when the number of polarization corrections equals 1 means the received field is back propagated as it is until the transmitter and the SOPs of the channels are corrected to separate the data in two polarizations. When number of polarization corrections is 10, which is the number of spans, SOPs of each channel are corrected at every span to match the SOP of that channel at that location in forward propagation. This correction is possible in a simulation as the transmission matrix as a function of wavelength is recorded along the fiber. In a real transmission system, the polarization rotation matrix per span has to be monitored for each WDM channel. It should be noted that monitoring the polarization rotation matrix is expected to be more difficult than current efforts in monitoring only the PSPs and differential group delay (DGD).

Fig. 9 shows that for this highly nonlinear system PMD significantly deteriorates the capability of the DBP technique to compensate nonlinear impairments if the relative rotations of the SOPs of WDM channels are not taken into account. Even for a highly nonlinear system, correction of channel polarization at every span can be enough to achieve the same level of nonlinear compensation that can be achieved when there is no PMD. PMD is expected to be a less of a problem for transmission systems with lower nonlinearity as well as when fibers with lower PMD parameters significantly lower than 0.05 $\text{ps}/\sqrt{\text{km}}$ are used.

The simulation results are also used to see the penalty arising from neglecting the FWM terms. Fig. 9 is reproduced by solving the CEs in (2) with FWM terms dropped. The results are compared with the case when the full Manakov equation is solved in DBP, and they are shown in Fig. 10(a) for $D_p = 0.05\ \text{ps}/\sqrt{\text{km}}$ and in Fig. 10(b) for $D_p = 0.1\ \text{ps}/\sqrt{\text{km}}$. CEs are capable of compensating for nonlinearity, even though they perform slightly worse when FWM is not included in the equations as expected.

5. Conclusion

Nonlinear impairments are currently limiting fiber-optic transmission capacity. DBP has emerged as a potential method of removing nonlinear impairments and thereby increasing the capacity of fiber-optic transmission. The effectiveness of DBP using the full Manakov equation has been shown not only by simulations but also by experiments. This paper further demonstrates that DBP can be applied using the CEs and compensates only for XPM and SPM, which is possible when FWM is small. The improvement obtained by compensating all the nonlinear impairments versus compensating only the incoherent processes is presented. The comparison showed that impact of FWM is small even for channel spacing as low as 7 GHz. The additional advantage of compensating only for incoherent processes is that it does not require the use of phase-locked local oscillators. This advantage is also demonstrated based on the experimental results.

A major concern for the applicability of DBP is the presence of PMD. Since nonlinear processes couple WDM channels far from one another spectrally, PMD plays an important role even if individual channel baud rates are low enough to avoid impairment from PMD. The impact of PMD on DBP is estimated by extensive simulations. It is shown that PMD significantly hinders the ability to remove nonlinear impairments by DBP because of the random nature of the PMD. Nevertheless, effective compensation of nonlinear impairments using DBP is still possible in the presence of PMD, provided that the polarization rotation matrix of the transmission fiber is monitored on a per-span basis.

References

[1] A. R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *J. Lightw. Technol.*, vol. 8, no. 10, pp. 1548–1557, Oct. 1990.
[2] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, and R. M. Derosier, "8 × 10 Gb/s transmission through 280 km of dispersion-managed fiber," *IEEE Photon. Technol. Lett.*, vol. 5, no. 10, pp. 1233–1235, Oct. 1993.
[3] R. W. Tkach, A. R. Chraplyvy, F. Forghieri, A. H. Gnauck, and R. M. Derosier, "Four-photon mixing and high-speed WDM systems," *J. Lightw. Technol.*, vol. 13, no. 5, pp. 841–849, May 1995.
[4] N. S. Bergano, "Wavelength division multiplexing in long-haul transoceanic transmission systems," *J. Lightw. Technol.*, vol. 23, no. 12, pp. 4125–4139, Dec. 2005.
[5] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, "Information capacity of optical fiber channels with zero average dispersion," *Phys. Rev. Lett.*, vol. 91, no. 20, pp. 1–4, Nov. 2003.
[6] R. J. Essiambre, G. J. Foschini, P. J. Winzer1, G. Kramer, and E. C. Burrows, "The capacity of fiber-optic communication systems," in *Proc. OFC*, 2008, pp. 1–3.
[7] N. S. Bergano and J. C. Feggeler, "TuB4 four-wave mixing in long lengths of dispersion-shifted fiber using a circulating loop," in *Proc. OFC*, 1992, p. TuB4.
[8] N. Takachio and K. Iwashita, "Compensation of fiber dispersion in optical heterodyne detection," *Electron. Lett.*, vol. 24, no. 2, pp. 759–760, Jan. 1988.
[9] S. Boehm, K. Schumacher, D. Goelz, and P. Meissner, "PMD compensation with coherent reception and digital signal processing," in *Proc. Lasers Electro-Opt. CLEO*, 2007, pp. 1–2, JTuA132.
[10] A. T. Erdogan, A. Demir, and T. M. Oktem, "Automatic PMD compensation by unsupervised polarization diversity combining coherent receivers," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1823–1834, Jul. 2008.

[11] C. Kurtzke, "Suppression of fiber nonlinearities by appropriate dispersion management," *IEEE Photon. Technol. Lett.*, vol. 5, no. 10, pp. 1250–1252, Oct. 1993.
[12] N. S. Bergano, C. R. Davidson, B. M. Nyman, S. G. Evangelides, J. M. Darcie, J. D. Evankow, P. C. Corbett, M. A. Mills, G. A. Ferguson, J. A. Nagel, J. L. Zyskind, J. W. Sulhoff, A. J. Lucero, and A. A. Klein, "40 Gb/s WDM transmission of eight 5 Gb/s data channels over transoceanic distances using the conventional NRZ modulation format," in *Proc. OFC*, 1995, p. PD19.
[13] T. Naito, T. Terahara, T. Chikama, and M. Suyama, "Four 5-Gb/s WDM transmission over 4760-km straight-line using pre- and post-dispersion compensation and FWM cross talk reduction," in *Proc. OFC*, 1996, p. WM3.
[14] A. H. Gnauck, G. Raybon, S. Chandrasekhar, J. Leuthold, C. Doerr, L. Stulz, A. Agarwal, S. Banerjee, D. Grosz, S. Hunsche, A. Kung, A. Marhelyuk, D. Maywar, M. Movassaghi, X. Liu, C. Xu, X. Wei, and D. M. Gill, "2.5 tb/s (64×42.7 Gb/s) transmission over 40×100 km NZDSF using RZ-DPSK format and all-Raman-amplified spans," in *Proc. OFC*, 2002, pp. FC2-1–FC2-3.
[15] T. Mizuochi, K. Ishida, T. Kobayashi, J. Abe, K. Kinjo, K. Motoshima, and K. Kasahara, "A comparative study of DPSK and OOK WDM transmission over transoceanic distances and their performance degradations due to nonlinear phase noise," *J. Lightw. Technol.*, vol. 21, no. 9, pp. 1933–1943, Sep. 2003.
[16] I. Neokosmidis, T. Kamalakis, A. Chipouras, and T. Sphicopoulos, "New techniques for the suppression of the four-wave mixing-induced distortion in nonzero dispersion fiber WDM systems," *J. Lightw. Technol.*, vol. 23, no. 3, pp. 1137–1144, Mar. 2005.
[17] C. Behrens, R. I. Killey, S. J. Savory, and P. Bayvel, "Reducing the impact of intrachannel nonlinearities by pulse-width optimisation in multi-level phase-shift-keyed transmission," in *Proc. ECOC*, 2009, pp. 1–2.
[18] B. Châtelain, Y. Jiang, K. Roberts, X. Xu, J. Cartledge, and D. V. Plant, "Impact of pulse shaping on the SPM tolerance of electronically pre-compensated 10.7 Gb/s DPSK systems," in *Proc. OFC*, 2010, pp. 1–3, OTuE6.
[19] K. Inoue, "Fiber four-wave mixing suppression using two incoherent polarized lights," *J. Lightw. Technol.*, vol. 11, no. 12, pp. 2116–2122, Dec. 1993.
[20] D. van den Borne, S. L. Jansen, S. Calabro, N. E. Hecker-Denschlag, G. D. Khoe, and H. de Waardt, "Reduction of nonlinear penalties through polarization interleaving in 2×10 GB/s polarization-multiplexed transmission," *IEEE Photon. Technol. Lett.*, vol. 17, no. 6, pp. 1337–1339, Jun. 2005.
[21] C. Xie, "WDM coherent PDM-QPSK systems with and without inline optical dispersion compensation," *Opt. Express*, vol. 17, no. 6, pp. 4815–4823, Mar. 2009.
[22] S. Watanabe, T. Chikama, G. Ishikawa, T. Terahara, and H. Kuwahara, "Compensation of pulse shape distortion due to chromatic dispersion and Kerr effect by optical phase conjugation," *IEEE Photon. Technol. Lett.*, vol. 5, no. 10, pp. 1241–1243, Oct. 1993.
[23] O. Kuzucu, Y. Okawachi, R. Salem, M. A. Foster, A. C. Turner-Foster, M. Lipson, and A. L. Gaeta, "Spectral phase conjugation via temporal imaging," *Opt. Express*, vol. 17, no. 22, pp. 20 605–20 614, Oct. 2009.
[24] X. Liu, X. Wei, R. E. Slusher, and C. J. McKinstrie, "Improving transmission performance in differential phase-shift-keyed systems by use of lumped nonlinear phase-shift compensation," *Opt. Lett.*, vol. 27, no. 18, pp. 1616–1618, Sep. 2002.
[25] C. Xu and X. Liu, "Postnonlinearity compensation with data-driven phase modulators in phase-shift keying transmission," *Opt. Lett.*, vol. 27, no. 18, pp. 1619–1621, Sep. 2002.
[26] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.
[27] G. Charlet, N. Maaref, J. Renaudier, H. Mardoyan, P. Tran, and S. Bigo, "Transmission of 40 Gb/s QPSK with coherent detection over ultra-long distance improved by nonlinearity mitigation," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, 2006.
[28] A. J. Lowery, "Fiber nonlinearity mitigation in optical links that use OFDM for dispersion compensation," *IEEE Photon. Technol. Lett.*, vol. 19, no. 19, pp. 1556–1558, Oct. 2007.
[29] K. Kikuchi, "Electronic post-compensation for nonlinear phase fluctuations in a 1000-km 20-Gb/s optical quadrature phase-shift keying transmission system using the digital coherent receiver," *Opt. Express*, vol. 16, no. 2, pp. 889–896, Jan. 2008.
[30] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.
[31] Y. Han and G. Li, "Coherent optical communication using polarization multiple-input-multiple-output," *Opt. Express*, vol. 13, no. 19, pp. 7527–7534, 2005.
[32] T. Pfau, S. Hoffmann, O. Adamczyk, R. Peveling, V. Herath, M. Porrmann, and R. Noé, "Coherent optical communication: Towards realtime systems at 40 Gb/s and beyond," *Opt. Express*, vol. 16, no. 2, pp. 866–872, Jan. 2008.
[33] K. Roberts, L. Chuandong, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 18, no. 2, pp. 403–405, Jan. 2006.
[34] R.-J. Essiambre, P. J. Winzer, X. Q. Wang, W. Lee, C. A. White, and E. C. Burrows, "Electronic predistortion and fiber nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 18, no. 17, pp. 1804–1806, Sep. 2006.
[35] E. Yamazaki, F. Inuzuka, K. Yonenaga, A. Takada, and M. Koga, "Compensation of interchannel crosstalk induced by optical fiber nonlinearity in carrier phase-locked WDM system," *IEEE Photon. Technol. Lett.*, vol. 19, no. 1, pp. 9–11, Jan. 2007.
[36] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Opt. Express*, vol. 16, no. 2, pp. 880–888, Jan. 2008.
[37] G. Goldfarb, M. G. Taylor, and G. Li, "Experimental demonstration of fiber impairment compensation using the split-step finite-impulse-response filtering method," *IEEE Photon. Technol. Lett.*, vol. 20, no. 22, pp. 1887–1889, Nov. 2008.
[38] E. Mateo, L. Zhu, and G. Li, "Impact of XPM and FWM on the digital implementation of impairment compensation for WDM transmission using backward propagation," *Opt. Express*, vol. 16, no. 20, pp. 16 124–16 137, 2008.

[39] E. Ip and J. Kahn, "Compensation of dispersion and nonlinear effects using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 2008.
[40] F. Yaman and G. Li, "Nonlinear impairment compensation for polarization-division multiplexed WDM transmission using digital backward propagation," *IEEE Photon. J.*, vol. 1, no. 2, pp. 144–152, Aug. 2009.
[41] E. F. Mateo and G. Li, "Compensation of interchannel nonlinearities using enhanced CEs for digital backward propagation," *Appl. Opt.*, vol. 48, no. 25, pp. F6–F10, 2009.
[42] E. Yamazaki, H. Masuda, A. Sano, T. Yoshimatsu, T. Kobayashi, E. Yoshida, Y. Miyamoto, R. Kudo, K. Ishihara, M. Matsui, and Y. Takatori, "Multi-staged nonlinear compensation in coherent receiver for 12 015 km WDM transmission of 10-ch $\times$ 111 Gb/s no-guard-interval co-OFDM," *Electron. Lett.*, vol. 45, no. 13, pp. 695–697, Jun. 2009.
[43] D. S. Millar, S. Makovejs, V. Mikhailov, R. I. Killey, P. Bayvel, and S. J. Savory, "Experimental comparison of nonlinear compensation in long-haul PDM-QPSK transmission at 42.7 and 85.4 Gb/s," in *Proc. Eur. Conf. Opt. Commun.*, 2009, pp. 1–2, Paper 9.4.4.
[44] T. Tanimura, S. Oda, T. Tanaka, and C. Ohsima, "Systematic analysis on multi-segment dual-polarisation nonlinear compensation in 112 Gb/s DP-QPSK coherent receiver," in *Proc. ECOC*, Vienna, Austria, 2009, pp. 1–2, Paper Th9.4.5.
[45] S. Oda, T. Tanimura, T. Hoshida, C. Ohshima, H. Nakashima, Z. Tao, and J. C. Rasmussen, "112 Gbps DP-QPSK transmission using a novel nonlinear compensator in digital coherent receiver," in *Proc. OFC/NFOEC*, San Diego, CA, Mar. 22–26, 2009, Paper OThR6.
[46] S. J. Savory, G. Gavioli, E. Torrengo, and P. Poggiolini, "Impact of interchannel nonlinearities on a split-step intrachannel nonlinear equalizer," *IEEE Photon. Technol. Lett.*, vol. 22, no. 10, pp. 673–675, May 2010.
[47] L. Du, B. Schmidt, and A. Lowery, "Efficient digital backpropagation for PDM-CO-OFDM optical transmission systems," in *Proc. OFC*, 2010, pp. 1–3.
[48] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, no. 19, pp. 1029–1030, Sep. 1986.
[49] J. P. Gordon and H. Kogelnik, "PMD fundamentals," *Proc. Nat. Acad. Sci.*, vol. 97, pp. 4541–4550, 2000.
[50] G. P. Agrawal, *Nonlinear Fiber Optics,* 3rd ed. San Diego, CA: Academic, 2001.
[51] B. Crosignani and P. Di Porto, "Intensity induced rotation of the polarization ellipse in low-birefringence, single mode optical fibres," *Opt. Acta*, vol. 32, no. 9/10, pp. 1251–1258, 1985.
[52] 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. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.
[53] P. K. A. Wai, W. L. Kath, C. R. Menyuk, and J. W. Zhang, "Nonlinear polarization-mode dispersion in optical fibers with randomly varying birefringence," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 14, no. 11, pp. 2967–2979, 1997.
[54] M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," *Opt. Lett.*, vol. 24, no. 14, pp. 939–941, 1999.
[55] F. Yaman, Q. Lin, and G. P. Agrawal, "Effects of polarization-mode dispersion in dual-pump fiber-optic parametric amplifiers," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 431–433, Feb. 2004.
[56] X. Zhou, J. Yu, M. Huang, Y. Shao, T. Wang, P. Magill, M. Cvijetic, L. Nelson, M. Birk, G. Zhang, S. Y. Ten, H. B. Matthew, and S. K. Mishra, "32 Tb/s (320 $\times$ 114 Gb/s) PDM-RZ-8QAM transmission over 580 km of SMF-28 ultra-low-loss fiber," presented at the OFC, San Diego, CA, 2009, Paper PDPB4.