

Theory of polarization attraction in parametric amplifiers based on telecommunication fibers

Massimiliano Guasoni,^{1,*} Victor V. Kozlov,^{1,2} and Stefan Wabnitz¹

¹*Department of Information Engineering, Università di Brescia, Via Branze 38, 25123 Brescia, Italy*

²*Department of Physics, St.-Petersburg State University, Petrodvoretz, St.-Petersburg 198504, Russia*

*Corresponding author: massimiliano.guasoni@ing.unibs.it

Received May 25, 2012; revised July 30, 2012; accepted August 6, 2012;
posted August 6, 2012 (Doc. ID 169394); published September 11, 2012

We develop from first principles the coupled wave equations that describe polarization-sensitive parametric amplification based on four-wave mixing (FWM) in standard (randomly birefringent) optical fibers. We show that in the small-signal case these equations can be solved analytically, and permit us to predict the gain experienced by the signal beam as well as its state of polarization (SOP) at the fiber output. We find that, independently of its initial value, the output SOP of a signal within the parametric gain bandwidth is solely determined by the pump SOP. We call this effect of pulling the polarization of the signal towards a reference SOP the polarization attraction, and we call the parametric amplifier the FWM polarizer (which can equivalently be called the fiber-optic parametric amplifier polarizer). Our theory is valid beyond the zero polarization mode dispersion (PMD) limit, and it takes into account moderate deviations of the PMD from zero. In particular, our theory is capable of analytically predicting the rate of degradation of the efficiency of the parametric amplifier, which is caused by the detrimental PMD effect. © 2012 Optical Society of America

OCIS codes: 230.5440, 060.4370, 230.1150, 230.4320.

1. INTRODUCTION

Recent years have witnessed a substantial growth of interest in developing nonlinear-optical techniques for the control of the state of polarization (SOP) of light beams. The motivation behind such research activities is twofold. First of all, nonlinear-optical techniques may permit replacing the inefficient and lossy method of polarizing a light beam with conventional passive linear polarizers with the lossless polarization attraction of an arbitrary initial SOP towards the desired SOP at the output of a nonlinear medium. A key advantage of using lossless polarization attraction is that, in contrast with passive linear polarizers, input signal SOP changes do not lead to output signal intensity fluctuations or relative intensity noise (RIN). The second goal is to find efficient ways to exercise all-optical control over the SOP of a signal beam by exploiting its nonlinear interaction with a pump beam with a well-determined SOP. Here we analyze a novel method for achieving the all-optical control of the SOP of a signal beam, namely exploiting the four-wave-mixing (FWM)-mediated process of parametric amplification in a standard telecom optical fiber.

In short, nonlinear-optical methods allow for designing novel types of polarizers with much greater functionality than conventional passive linear polarizers. So far, two distinctly different types of nonlinear-optical polarizers were proposed. The first class comprises the so-called nonlinear lossless polarizers (NLPs), which are based on the cross-polarization modulation (XPoIM) of two intense beams in a Kerr medium. To the second class belong the so-called Raman polarizers, which are based on the polarization-sensitive Raman amplification of a signal beam in a Raman-active medium. These two types of polarizers exploit the two complementary manifestations of the cubic nonlinearity of fibers—conservative for inducing XPoIM effect, and dissipative, which is responsible for

the Raman effect. Here, we exploit the same cubic nonlinearity, more precisely its conservative part, for initiating the process of polarization-sensitive FWM between three beams.

The first NLP was proposed and experimentally demonstrated by Heebner *et al.* in [1]. It was based not on the Kerr nonlinearity, but on a photorefractive effect. This polarizer was capable of transforming, in a lossless manner, a light beam with an arbitrary initial SOP into a beam with one and the same SOP towards its output. The principle of operation of this device was the conversion of energy from one polarization component of the beam into its orthogonal polarization component. Photorefractive materials are characterized by a nonlinear response that is far too slow to be useful in contemporary ultrafast optics. In contrast, the Kerr nonlinearity of silica is virtually instantaneous, which makes optical fibers a promising medium for implementing lossless polarizers within high-bit-rate telecom networks. The progress in developing fiber-based NLPs started from impractical isotropic fibers [2–4] and evolved towards cheap and reliable telecom fibers [5–8] or specialty fibers such as highly birefringent and spun fibers [9]. The mathematical aspects of the problem were studied in [10–14], and allowed us to get further insight into the physics of fiber-based NLPs, whose principle of operation is different from that of photorefractive lossless polarizers. Instead of the self-interaction of a single beam in a photorefractive material, a two-beam cross-interaction (namely, XPoIM) is used in the Kerr medium. Namely, an auxiliary pump beam with a well-defined SOP is employed, serving as a polarization reference for the signal beam with arbitrary initial SOP. As previously outlined, when using lossless polarizers, input signal SOP fluctuations do not lead to output RIN [5].

Another type of nonlinear-optical polarizer is the Raman polarizer. It is different from conventional Raman amplifiers by its sensitivity to the SOP of the pump beam. The signal that experiences Raman amplification acquires an SOP that is dictated by the SOP of the pump. In this way we may exercise an all-optical control over the polarization of the signal beam. Note that conventional fiber-optic Raman amplifiers operate in the regime where the output SOP of the signal is independent on the pump SOP. Moreover, most conventional fiber-optic Raman amplifiers are driven by unpolarized pumps. The first Raman polarizer was demonstrated by Martinelli *et al.* in [15], followed by a number of theoretical papers [16–23]. Similar polarization-sensitive amplification was predicted theoretically and confirmed experimentally in [24] for the Brillouin amplification of a signal beam in standard optical fibers. These devices can be similarly called Brillouin polarizers. Since they are based on a gain mechanism, which is maximum whenever the signal and pump SOPs are aligned and zero when they are orthogonal, in general both Raman and Brillouin-based polarizers suffer from severe output RIN in the presence of input signal SOP fluctuations.

A common feature uniting all of these nonlinear fiber-optic polarizers is that they can operate efficiently only in the limit of vanishing polarization mode dispersion (PMD). PMD is the effect that is caused by random variations of the magnitude and/or orientation of the birefringence along the fiber length, and it is acquired as a result of inevitable technical imperfections in the process of drawing a fiber from a preform. Recent progress in fiber manufacturing brought to the market fibers with much lower values of PMD than was previously available. It is this technological breakthrough that made possible the observation of the previously discussed polarization-sensitive effects in optical fibers. Theoretical estimates show that the smaller the PMD coefficient, the shorter the total fiber length, and the smaller the frequency separation of the signal and the pump beams, the better the performance of all of the above described polarizers. It is one of the main goals of a theory to be able to predict the degradation rate of useful polarization attraction effects, which is caused by PMD. Such degradation rates for NLPs and Raman polarizers have been calculated analytically in [22,25].

It is important to note that the concept of all of these smart polarizers is not limited to fiber-optics applications only. Indeed, nonlinear polarizers can be implemented with any optical waveguide exhibiting Kerr and/or Raman nonlinearity. Using integrated optics waveguides may lift the problems that are associated with fiber PMD, and even make nonlinear polarizers less bulky and more compact, providing that the waveguide material exhibits nonlinear coefficients that are much larger than silica. For example, the silicon-based Raman polarizer proposed in [26] is free of the PMD-induced degradation and has a centimeter-long size as compared to the kilometer-long fibers, thanks to 3–4 orders of magnitude Raman gain enhancement in silicon with respect to silica.

The present theoretical study extends the concept of nonlinear polarizers to the FWM process in telecom fibers. The goal here is to find the conditions upon which the process of parametric amplification is sensitive to the SOP of the pump beam. In this way we arrive to the notion of an FWM polarizer, meaning that the SOP of the amplified signal beam is determined by the SOP of the fully polarized pump beam. We derive

here the coupled wave equations for the pump, idler, and signal beams. In the limit of zero PMD, these equations reduce to the equations that were previously derived by McKinstrie *et al.* in [27] for describing degenerate FWM in standard fibers. The major advantage of our theory is its applicability (slightly) beyond the zero-PMD limit, in the sense that it is capable of predicting the degradation rate of the efficiency of the FWM polarizer for low-PMD fibers as well. Knowing this degradation rate allows one to properly design practical fiber-based nonlinear polarizers. The present work substantially extends to the case of random birefringence telecom fibers a previous study of polarization attraction in deterministic, high-birefringence optical fibers [28]. Note that the polarization-sensitive parametric amplification in optical fibers was studied theoretically by Lin and Agrawal [29,30], and also theoretically and experimentally by Freitas *et al.* in [31], and resulted in a proposal of a fiber-based polarization switch. As discussed in [28], FWM-based polarizers are based on the polarization sensitivity of parametric gain. Such gain is maximum for a signal SOP that is aligned with that of the pump, and zero for a signal SOP orthogonal to the pump. Thus FWM polarizers are not immune from output RIN resulting from input signal polarization fluctuations. Nevertheless, since parametric gain is generally larger than Raman gain in silica fibers, FWM polarizers may employ shorter fibers or lower pump powers than Raman polarizers. In addition, since the repolarization capability of FWM polarizers is based on parametric gain, these devices provide a more flexible control over the gain and repolarization bandwidth. Indeed, such bandwidth may be extended up to 70 nm and even include the normal dispersion regime by properly engineering the wavelength dependence of the fiber dispersion and by adjusting the pump power [32].

2. EQUATIONS OF THE MODEL

We shall consider the process of degenerate FWM. This process involves three continuous waves with frequencies that satisfy the matching condition $2\omega_p = \omega_s + \omega_i$. Pump, signal, and idler waves are labeled correspondingly as p , s , i . All three waves are copropagating along the z direction in a telecommunication (i.e., randomly birefringent) fiber. The vectorial theory of parametric amplification in fibers was developed in [27,29,30], basing on the tensorial properties of silica in the telecom band. The starting equation is derived under standard for nonlinear optics approximations, from Maxwell's equation with a polarization that takes into account the nonlinear cubic response of silica and the birefringence of the fiber. Utilizing the Jones representation, the equations for the Jones vectors of the pump and the signal read as

$$\begin{aligned}
 i \frac{\partial U_p}{\partial z} + \Delta B(\omega_p, z) U_p + \frac{2}{3} \gamma \left[(U_p \cdot U_p^*) U_p + \frac{1}{2} (U_p \cdot U_p) U_p^* \right] \\
 + \frac{2}{3} \gamma [(U_s \cdot U_s^*) U_p + (U_s \cdot U_p) U_s^* + (U_p \cdot U_s^*) U_s \\
 + (U_i \cdot U_i^*) U_p + (U_i \cdot U_p) U_i^* + (U_p \cdot U_i^*) U_i] \\
 + \frac{2}{3} \gamma \exp(i \Delta k z) [(U_i \cdot U_s) U_p^* + (U_i \cdot U_p^*) U_s \\
 + (U_s \cdot U_p^*) U_i] = 0,
 \end{aligned} \tag{1}$$

$$\begin{aligned}
i \frac{\partial U_s}{\partial z} + \Delta B(\omega_s, z) U_s + \frac{2}{3} \gamma \left[(U_s \cdot U_s^*) U_s + \frac{1}{2} (U_s \cdot U_s) U_s^* \right] \\
+ \frac{2}{3} \gamma \left[(U_p \cdot U_p^*) U_s + (U_p \cdot U_s) U_p^* + (U_s \cdot U_p^*) U_p \right. \\
+ (U_i \cdot U_i^*) U_s + (U_i \cdot U_s) U_i^* + (U_s \cdot U_i^*) U_i \left. \right] \\
+ \frac{2}{3} \gamma \exp(-i \Delta k z) \left[\frac{1}{2} (U_p \cdot U_p) U_i^* + (U_p \cdot U_i^*) U_p \right] = 0, \quad (2)
\end{aligned}$$

while the idler equation is obtained from Eq. (2) by exchanging labels s and i .

The Jones vectors $U_f = (u_{xf}, u_{yf})^T$ (with $f = \{p, s, i\}$) are two-component vectors with $u_{xf}(z)$ and $u_{yf}(z)$ being the amplitudes of the polarization components in a fixed laboratory reference frame (x, y) . Note that the last terms in the left-hand sides of Eqs. (1) and (2) represent the so-called energy-exchange terms. They are responsible for the transfer of energy between different waves, and as such are most important for our analysis of parametric amplification. The wave vector mismatch $\Delta k = \beta_{ei} + \beta_{es} - 2\beta_{ep} = \beta_{oi} + \beta_{os} - 2\beta_{op}$, where β_{ef} and β_{of} are the propagation constants of the modes aligned with extraordinary (e), or slow, and ordinary (o), or fast, axes at frequency ω_f , respectively. The 2×2 matrix $\Delta B(\omega_f) =$

reference frame, which is defined by the polarization modes \mathbf{e}_x and \mathbf{e}_y . $\sigma_3 = \text{diag}(1, -1)$ and $\sigma_1 = \text{adiag}(1, 1)$ are known as Pauli matrices, where *diag* and *adiag* stand for diagonal and antidiagonal matrices, respectively.

The orientation angle θ is randomly varying in fibers used for telecommunication applications, which explains the term randomly birefringent fibers that is applied to them. In principle, the magnitude of the birefringence $b(\omega_f)$ also varies stochastically along z . However, as noticed in [33], the two approaches, one in which θ is the only stochastic variable, and the second, where both θ and b are stochastic variables, produce nearly identical results. Thus, here we shall develop our theory by assuming the single stochastic variable θ . The angle θ is driven by a white noise process $\partial_z \theta = g_\theta(z)$, where $\langle g_\theta(z) \rangle = 0$ and $\langle g_\theta(z) g_\theta(z') \rangle = 2L_c^{-1} \delta(z - z')$. Here L_c is the correlation length that characterizes the typical distance over which θ changes randomly. The theory developed below is the natural extension of the one-beam theory of Wai and Menyuk in [33] and the two-beam theory of Kozlov *et al.* [8] to the case of three interacting beams. All details of the derivations of the final equations of motion with deterministic coefficients starting from Eqs. (1) and (2) with stochastic coefficients, as well as the approximations that appeared on the way, can be found in Appendices A and B. Here we write down the final result:

$$\begin{aligned}
i \frac{\partial \phi_{1p}}{\partial z} + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{1i} \phi_{1p} \phi_{1i}^* + \frac{8}{9} \phi_{1p}^2 \phi_{1p}^* + \frac{2}{3} \left(2 C_a(z) + \frac{2}{3} C_b(z) \right) \phi_{1i} \phi_{1s} \phi_{1p}^* e^{i \Delta k z} + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{1p} \phi_{1s} \phi_{1s}^* \\
+ \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{1p} \phi_{2i} \phi_{2i}^* + \frac{8}{9} C_a(z) \phi_{1i} \phi_{2p} \phi_{2i}^* + \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{1s} \phi_{2i} \phi_{2p}^* e^{i \Delta k z} + \frac{8}{9} \phi_{1p} \phi_{2p} \phi_{2p}^* \\
+ \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{1i} \phi_{2s} \phi_{2p}^* e^{i \Delta k z} + \frac{8}{9} C_a(z) \phi_{1s} \phi_{2p} \phi_{2s}^* + \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{1p} \phi_{2s} \phi_{2s}^* = 0, \quad (3)
\end{aligned}$$

$$\begin{aligned}
i \frac{\partial \phi_{1s}}{\partial z} + \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{1p}^2 \phi_{1i}^* e^{-i \Delta k z} + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{1i} \phi_{1s} \phi_{1i}^* + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{1p} \phi_{1s} \phi_{1p}^* + \frac{8}{9} \phi_{1s}^2 \phi_{1s}^* \\
+ \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{1s} \phi_{2i} \phi_{2i}^* + \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{1p} \phi_{2p} \phi_{2i}^* e^{-i \Delta k z} + \frac{8}{9} C_a(z) \phi_{1i} \phi_{2s} \phi_{2i}^* + \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{1s} \phi_{2p} \phi_{2p}^* \\
+ \frac{8}{9} C_a(z) \phi_{1p} \phi_{2s} \phi_{2p}^* + \frac{8}{9} \phi_{1s} \phi_{2s} \phi_{2s}^* = 0, \quad (4)
\end{aligned}$$

$$\begin{aligned}
i \frac{\partial \phi_{2s}}{\partial z} + \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{2p}^2 \phi_{2i}^* e^{-i \Delta k z} + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{2i} \phi_{2s} \phi_{2i}^* + \frac{2}{3} \left(2 + \frac{2}{3} C_a(z) \right) \phi_{2p} \phi_{2s} \phi_{2p}^* + \frac{8}{9} \phi_{2s}^2 \phi_{2s}^* \\
+ \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{2s} \phi_{1i} \phi_{1i}^* + \frac{2}{3} \left(C_a(z) + \frac{1}{3} C_b(z) \right) \phi_{2p} \phi_{1p} \phi_{1i}^* e^{-i \Delta k z} + \frac{8}{9} C_a(z) \phi_{2i} \phi_{1s} \phi_{1i}^* \\
+ \frac{2}{3} \left(2 - \frac{2}{3} C_a(z) \right) \phi_{2s} \phi_{1p} \phi_{1p}^* + \frac{8}{9} C_a(z) \phi_{2p} \phi_{1s} \phi_{1p}^* + \frac{8}{9} \phi_{2s} \phi_{1s} \phi_{1s}^* = 0. \quad (5)
\end{aligned}$$

$b(\omega_f)(\cos \theta \sigma_3 + \sin \theta \sigma_1)$ represents the birefringence tensor, where $b(\omega_f) = \frac{1}{2}(\beta_{ef} - \beta_{of})$ is half of the value of the fiber birefringence at frequency ω_f . Moreover, θ is the angle of orientation of the axis of the birefringence with respect to the fixed

Here $C_a(z) = \exp[-(8/3)(\Delta^{(-)})^2 L_c z]$ and $C_b(z) = \exp[-(1/3)(\Delta^{(-)})^2 L_c z]$, where $\Delta^{(-)} = b(\omega_p) - b(\omega_i)$. The equation for ϕ_{1i} is obtained from Eq. (4) by exchanging the labels i and s ; the equations for ϕ_{2p} and ϕ_{2i} are obtained from the equations for

ϕ_{1p} and ϕ_{1i} , respectively, by exchanging the labels 1 and 2. The polarization components ϕ_{1f} and ϕ_{2f} are obtained from the original components u_{xf} and u_{yf} by means of a unitary transformation of the reference frame; see Appendices A and B for details.

When the value of $\Delta^{(-)}$ is zero, the z -dependent coefficients $C_a(z)$ and $C_b(z)$ are both equal to unity, and we restore the model equations derived in [27] starting from the Manakov equation. This limit corresponds to vanishing PMD, and it is quite natural to call it the Manakov limit. In this limit the conversion of the pump energy into the signal is maximally efficient, and therefore parametric amplifiers should be designed in such a way that the PMD diffusion length $L_d \equiv [(8/3)(\Delta^{(-)})^2 L_c]^{-1}$ is much longer than the fiber length L . To the best of our knowledge, our theory for the first time analytically predicts the length scale of degradation of the process of parametric amplification in telecom fibers. Strictly speaking, our theory is valid in two limits: $L \ll L_d$ and $L \gg L_d$. In the opposite limit (the limit of large PMD, which we call the diffusion limit) where $L \gg L_d$, the FWM process is totally suppressed. Therefore this regime is not interesting from the viewpoint of frequency conversion. Most likely, the intermediate case of $L \sim L_d$ can be adequately treated only numerically; however, we believe that the exponential decay of the nonlinear coefficients provides a qualitatively correct description of the rate of degradation.

Note that the PMD diffusion length L_d also enters the theory of two-beam nonlinear interactions: it was introduced by Lin and Agrawal in [25] in the context of fiber-optic Raman amplifiers, and it was identified as the typical length at which the mutual orientation of the states of polarization of the pump and signal beams is scrambled as a result of PMD. It is quite remarkable that the same length scale not only characterizes the ‘‘polarization memory’’ of Raman interactions, but also the degradation of XPolM mediated Kerr interactions of the two beams, as shown in [22]. In order to overcome the PMD-induced degradation of FWM efficiency, an experimentalist needs to select a low-PMD highly nonlinear fiber. As the degradation rate depends quadratically on the frequency difference between signal and pump, the effect of PMD can be also viewed as setting an upper limit to the bandwidth of the simultaneous parametric amplification and repolarization process.

In order to bring the definition of the PMD diffusion length to a more standard form, we introduce the PMD coefficient as proposed by Wai and Menyuk in [33]:

$$D_p = \frac{2\sqrt{2}\pi}{\omega_p} \frac{\sqrt{L_c}}{L_B}. \quad (6)$$

Here $L_B(\omega_p) = 2\pi / (\beta_{ep} - \beta_{op})$ is the beat length at the pump frequency. Then, we can write $L_d^{-1} = \frac{1}{3}(D_p \Delta\omega)^2$, where $\Delta\omega = \omega_p - \omega_i$.

Note that Eqs. (3)–(5) are applicable in the undepleted as well as pump-depleted regimes. In this paper we are aiming at the proof-of-principle demonstration of the FWM polarizer, and therefore limit ourselves to the study of the undepleted-pump regime only.

3. POLARIZATION ATTRACTION: ANALYTICAL RESULTS

In this section, we shall apply Eqs. (3)–(5) to describe the effect of polarization attraction of a signal or idler wave towards the SOP of a copropagating pump beam by means of FWM in a randomly birefringent telecom fiber. We will limit our analysis here to the small-signal case; i.e., we make the undepleted-pump approximation. As we shall see, this approximation permits us to obtain relatively simple analytical results for the effective bandwidth and gain of the polarization attraction process. From Eq. (3) one obtains the two polarization components of the pump amplitude as

$$\begin{aligned} \phi_{1p} &= \sqrt{P} \exp(i\theta_{1p0} + i\gamma(8/9)P_{\text{tot}}z), \\ \phi_{2p} &= \sqrt{Q} \exp(i\theta_{2p0} + i\gamma(8/9)P_{\text{tot}}z), \end{aligned} \quad (7)$$

where $P = |\phi_{1p}(0)|^2$, $Q = |\phi_{2p}(0)|^2$, $\sqrt{P} \exp(i\theta_{1p0})$, and $\sqrt{Q} \exp(i\theta_{2p0})$ are the input pump amplitudes in the fiber and $P_{\text{tot}} = P + Q$ is the conserved total pump power. For vanishing PMD (Manakov limit) $\Delta^{(-)} \cong 0$, $C_a(z) \cong C_b(z) \cong 1$; peak sideband gain is obtained at a frequency detuning $\Delta\omega_p = \Delta\omega_{pm}$ such that the dispersive mismatch is compensated by the pump-induced nonlinear phase shift, i.e., $\beta_2 \Delta\omega_{pm}^2 + \gamma(16/9)P_{\text{tot}} = 0$. In the absence of higher-order dispersion, this condition can only be reached in the anomalous dispersion regime (i.e., with $\beta_2 \leq 0$). In the opposite case of large PMD (diffusion limit), one has $C_a(z) \cong C_b(z) \cong 0$, so that the FWM terms are effectively suppressed. Yet, approaching the diffusion limit is equivalent to reducing the effective pump power to zero, which correspondingly leads to peak gain for sideband detunings $\Delta\omega_p \cong 0$. Let us consider now the intermediate case of $L \sim L_d$, where peak gain is observed for sideband detunings $\Delta\omega_p$ such that $0 \leq \Delta\omega_p \leq \Delta\omega_{pm}$. In order to quantify the sideband gain and evaluate their SOP relative to the pump, we need to solve Eqs. (3)–(5). Let us apply the change of variables $\phi_{1,2(i,s)} = \tilde{\phi}_{1,2(i,s)} \exp(-ivz/2 + i\theta_{1p0,2p0})$, where $v = \beta_2 \Delta\omega^2 - (16/9)\gamma P_{\text{tot}}$. By linearizing Eqs. (4) and (5) for the sidebands, one obtains

$$\frac{\partial \tilde{\phi}}{\partial z} = i \frac{8}{9} M(z) \tilde{\phi}, \quad (8)$$

where $\tilde{\phi} = [\tilde{\phi}_{1i}, \tilde{\phi}_{1s}^*, \tilde{\phi}_{2i}, \tilde{\phi}_{2s}^*]^T$, and

$$M(z) = \begin{bmatrix} F_A(z)P + F_B(z)Q + v/2 & F_C(z)P & F_D(z)\sqrt{PQ} & F_C(z)\sqrt{PQ} \\ -F_B(z)P & -F_A(z)P - F_B(z)Q - v/2 & -F_C(z)\sqrt{PQ} & -F_D(z)\sqrt{PQ} \\ F_D(z)\sqrt{PQ} & F_C(z)\sqrt{PQ} & F_B(z)P + F_A(z)Q + v/2 & F_C(z)Q \\ -F_C(z)\sqrt{PQ} & -F_D(z)\sqrt{PQ} & -F_C(z)Q & -F_B(z)P - F_A(z)Q - v/2 \end{bmatrix}$$

with $F_A(z) = 4\gamma/3(1 + C_a(z)/3)$, $F_B(z) = 4\gamma/3(1 - C_a(z)/3)$, $F_C(z) = 2\gamma/3(C_a(z) + C_b(z)/3)$; $F_D(z) = 8\gamma C_a(z)/9$. The solution of Eqs. (8) may be written as

$$\vec{\phi}(z=L) = \exp(\Omega(L))\vec{\phi}(z=0),$$

where $\Omega(z)$ is constructed from $M(z)$ as a Magnus series expansion [34]. Whenever the z -dependent coefficients $F_{(A,B,C,D)}(z)$ are slowly varying over L (i.e., $L_d \geq L$), we may truncate the expansion after the first term

$$\Omega(L) \cong \Omega_1(L) = \int_{z=0}^L M(z) dz \equiv \bar{M} \quad (9)$$

so that we simply replace $F_{(A,B,C,D)}(z)$ with their average values

$$\bar{F}_{(A,B,C,D)} = \frac{1}{L} \int_{z=0}^L F_{(A,B,C,D)}(z) dz, \quad (10)$$

which can be analytically calculated since $\bar{C}_a = k_a^{-1}(1 - \exp(-k_a L))$ and $\bar{C}_b = k_b^{-1}(1 - \exp(-k_b L))$, where $k_a = (8/3)(\Delta^{(-)})^2 L_c$ and $k_b = (1/3)(\Delta^{(-)})^2 L_c$. In the anomalous dispersion regime and for sideband frequency detunings $\Delta\omega$ below a certain cutoff value $\Delta\omega_c$, \bar{M} has an eigenvalue with positive imaginary part, leading to the effective (or average) sideband gain coefficient g_e

$$g_e^2 = \frac{4}{90} \gamma^2 P_{\text{tot}}^2 (-4 + 5\bar{C}_a^2 + \bar{C}_b^2 - 8\bar{C}_a + 6\bar{C}_a \bar{C}_b) - \frac{1}{4} \beta_2^2 \Delta\omega^4 - \frac{4}{9} \beta_2 (1 + \bar{C}_a) \gamma P_{\text{tot}} \Delta\omega^2. \quad (11)$$

From Eq. (11), we obtain the cutoff frequency $\Delta\omega_c$ of the gain band, the peak frequency detuning $\Delta\omega_p$, and the effective gain $g_{e,\text{peak}}$ as

$$\begin{aligned} \Delta\omega_c^2 &= 4c \left(\frac{6\gamma L_c^{-1} P_{\text{tot}}}{27|\beta_2|^2 c^2 L_c^{-1} + 41D_p^2 \gamma P_{\text{tot}} c^2 L_c^{-1} L} \right), \\ \Delta\omega_p &= \Delta\omega_c / \sqrt{2}, \\ g_{e,\text{peak}}^2 &= \frac{8}{3} \left(\frac{\gamma^2 P_{\text{tot}}^2 (3|\beta_2|^2 c^2 L_c^{-1} - D_p^2 \gamma P_{\text{tot}} L c^2 L_c^{-1})}{27\beta_2^2 c^2 L_c^{-1} + 41D_p^2 \gamma P_{\text{tot}} c^2 L_c^{-1} L} \right), \end{aligned} \quad (12)$$

where c is the speed of light in vacuum. Let us briefly discuss the role of the different physical parameters in determining the sideband gain and its bandwidth. First of all, increasing the PMD coefficient D_p or the fiber length L reduces the peak gain coefficient as well as the optimal sideband detuning. In order to study the polarization properties of the sidebands, we need to consider the eigenvectors of \bar{M} . For any frequency detuning $\Delta\omega$ within the gain band, let us denote by \vec{p} the eigenvector of \bar{M} , which grows as $\exp(g_e z)$. After a relatively short distance into the fiber, we may well approximate the sideband fields as $\vec{\phi} \approx C p \exp(g_e z)$, where C is the projection or scalar product (which we suppose nonzero for simplicity) of the input sidebands polarization vector $\vec{\phi}(z=0)$ on \vec{p} . The components of \vec{p} are such that $\vec{p}[1]/\vec{p}[3] = \vec{p}[2]/\vec{p}[4] = \sqrt{P/Q}$. Idler amplitudes $\vec{\phi}_{1i}$ and $\vec{\phi}_{2i}$ correspond to the first and third components of $\vec{\phi}$, respectively. Thus their ratio can be expressed as $\phi_{1i}/\phi_{2i} = \vec{p}[1]/\vec{p}[3] = \sqrt{P/Q}$. Since $\phi_{1i}/\phi_{2i} = (\tilde{\phi}_{1i}/\tilde{\phi}_{2i})e^{i\theta_{1i}-i\theta_{2i}}$, we obtain that

$$\frac{\phi_{1i}}{\phi_{2i}} = \frac{\sqrt{P}}{\sqrt{Q}} e^{i\theta_{1i}-i\theta_{2i}} = \frac{\phi_{1p}(z=0)}{\phi_{2p}(z=0)}. \quad (13)$$

A similar treatment can be developed for the signal amplitudes too, which proves the polarization attraction of both the signal and the idler to the input polarization of the pump.

In practice, since for a given sideband frequency detuning the effective gain coefficient g_e decreases as the fiber length L grows larger, the corresponding strength of polarization attraction will be reduced whenever the fiber length approaches L_d . As a matter of fact, in the diffusion limit $g_e = 0$ and FWM-induced polarization attraction is no longer observed. In the next section, we will provide a quantitative description of the fiber length dependence of the polarization attraction efficiency.

4. POLARIZATION ATTRACTION: EXAMPLES

Let us study the efficiency of polarization attraction as a function of fiber length L , hence of PMD. Consider a fiber with the nonlinear coefficient $\gamma = 11.9 \text{ W}^{-1} \text{ km}^{-1}$ and dispersion $\beta_2 = -0.5 \text{ ps}^2 \text{ km}^{-1}$ at the wavelength $\lambda = 1550 \text{ nm}$; the PMD correlation length is set to $L_c = 10 \text{ m}$. The chosen parameters are typical for highly nonlinear-optical fibers. As is well known, the SOP of each interacting wave may be represented by means of its corresponding unitary dimensionless Stokes vector as $\vec{S}_j = [S_{1j} = S_{0j}^{-1}(\phi_{1j}^* \phi_{2j} + \phi_{1j} \phi_{2j}^*), S_{2j} = S_{0j}^{-1}(i\phi_{1j}^* \phi_{2j} - i\phi_{1j} \phi_{2j}^*), S_{3j} = S_{0j}^{-1}(|\phi_{1j}|^2 - |\phi_{2j}|^2)]$, ($j = i, p, s$), where $S_{0j} = (|\phi_{1j}|^2 + |\phi_{2j}|^2)^{1/2}$.

In this notation the principal SOPs $[1, 0, 0]$, $[0, 1, 0]$, and $[0, 0, 1]$ represent a linear polarization at 45° , a right-handed circular polarization, and a linear polarization at 0° from the local birefringence axes, respectively; the SOPs $[-1, 0, 0]$, $[0, -1, 0]$, and $[0, 0, -1]$ represent a linear polarization at -45° , a left-handed circular polarization, and a linear polarization at 90° , respectively.

The input CW pump beam power is set to $P_{\text{tot}} = S_{0p} = 1 \text{ W}$, and its SOP is defined by the Stokes vector $\vec{S}_p = [\sqrt{0.5}, \sqrt{0.4}, \sqrt{0.1}]$. We set the input signal power to $P_{s,\text{in}} = 1 \text{ mW}$, whereas the idler is zero at the fiber input, as in typical FWM experiments. We compared the numerical solution of Eq. (8) with the analytical solution of Eq. (9). As the initial condition we employed a set of 10 000 input signal SOPs, whose corresponding Stokes vectors are uniformly distributed over the Poincaré sphere. Figures 1 and 2 illustrate the dependence on sideband detuning of the signal gain g_s and its output degree of polarization (DOP), respectively, for four different values of D_p (namely, $D_p = 0, 0.50 \text{ ps km}^{-1/2}$, $0.75 \text{ ps km}^{-1/2}$, and $D_p = 5 \text{ ps km}^{-1/2}$), and the fiber length $L = 300 \text{ m}$. The signal gain was computed as $g_s = (2L)^{-1} \log[P_{s,\text{out}}/P_{s,\text{in}}]$, where $P_{s,\text{out}}$ is the output signal power. The output DOP was calculated as discussed in [35]. In Figs. 1 and 2, the curves refer to numerical solutions, and the dots to analytical solutions: as can be seen, the first-order term of the Magnus expansion provides an excellent approximation of the exact solution. Figure 1 shows that, as the PMD grows larger, the signal gain g_s is progressively degraded; at the same time, both the peak gain frequency detuning $\Delta\omega_p$ and the cutoff frequency $\Delta\omega_c$ shrink towards zero. In addition, Fig. 2 shows that the signal DOP is maximum for sideband frequencies close to peak gain values; however, the peak DOP rapidly drops from

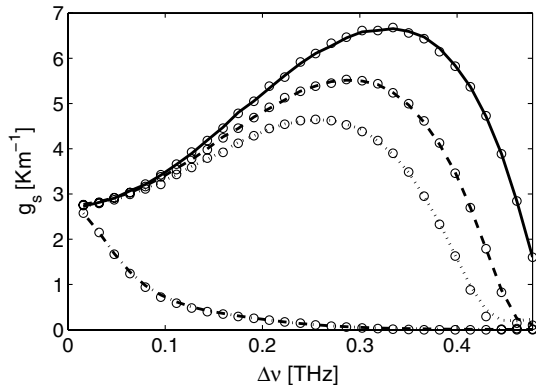


Fig. 1. Dependence of signal gain g_s on its frequency detuning from the pump, with $L = 300$ m. Curves and circles were obtained with z -varying or average M coefficients, respectively. Moreover $D_p = 0$ (solid curve), $D_p = 0.50$ ps km $^{-1/2}$ (dashed curve), $D_p = 0.75$ ps km $^{-1/2}$ (dotted curve), and $D_p = 5$ ps km $^{-1/2}$ (dash-dotted curve).

unity as the PMD strength is increased (i.e., for $D_p \geq 0.5$ ps km $^{-1/2}$). It is interesting to point out that, in contrast with the case of the signal, the output DOP of the idler (not shown here) remains close to unity throughout the entire gain bandwidth. The increased attraction of the idler towards the pump is due to the fact that the idler grows from zero at the fiber input; hence its projection on the growing eigenvector \vec{p} is much larger than for the signal.

In the second example of Fig. 3, we show the signal DOP as a function of the fiber length L , for four different values of the sideband frequency detuning $\Delta\nu = \Delta\omega/(2\pi)$ (i.e., $\Delta\nu = 0.255$ THz, 0.350 THz, 0.365 THz, 0.380 THz); here the PMD value is kept fixed to $D_p = 0.75$ ps km $^{-1/2}$. As can be seen, for $\Delta\nu = 0.255$ THz (which corresponds to the peak gain value), the DOP is monotonically increasing with distance, and it approaches the unit value for $L \geq 500$ m; in this case, $L_d = 2.1$ km. On the other hand, Fig. 3 shows that for other values of the sideband detuning, the output DOP exhibits a damped oscillating behavior and it converges to relatively low values after fiber lengths of the order of 1 km. Correspondingly, L_d decreases from 1.2 km down to 1 km. Indeed, the evolution of the signal polarization as described by Eqs. (3)–(5) in the small-signal limit is determined by two distinct physical effects: namely, parametric gain, which pulls the signal polarization towards the pump, and pump-induced nonlinear Kerr birefringence, which turns the signal SOP around the representative point of the pump SOP on the Poincaré sphere

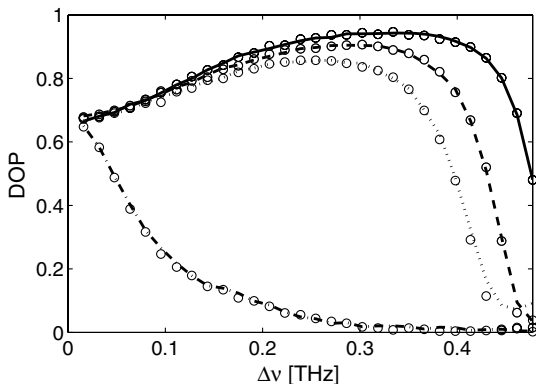


Fig. 2. Same as Fig. 1, but for the signal DOP.

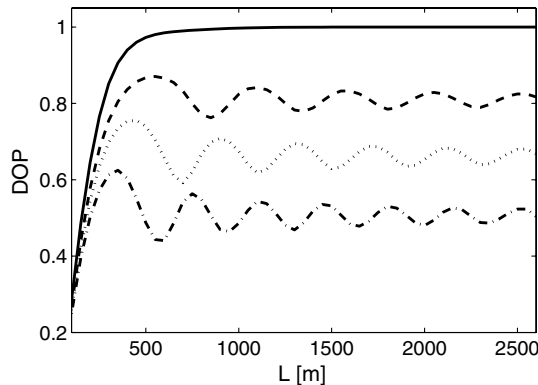


Fig. 3. Signal DOP versus fiber length L with $D_p = 0.75$ ps km $^{-1/2}$, and different values of the sideband detuning frequency: $\Delta\nu = 0.255$ THz (solid curve), $\Delta\nu = 0.350$ THz (dashed curve), $\Delta\nu = 0.365$ THz (dotted curve), and $\Delta\nu = 0.380$ THz (dash-dotted curve).

(see also the discussion in [31]). Thus, parametric gain and nonlinear birefringence lead to a motion of the signal SOP along two orthogonal directions on the Poincaré sphere. Hence, unless the parametric gain is so strong that the signal is immediately attracted towards the pump, spiral trajectories for the signal SOP may result on the sphere, which explains the DOP oscillations that are observed in Fig. 3.

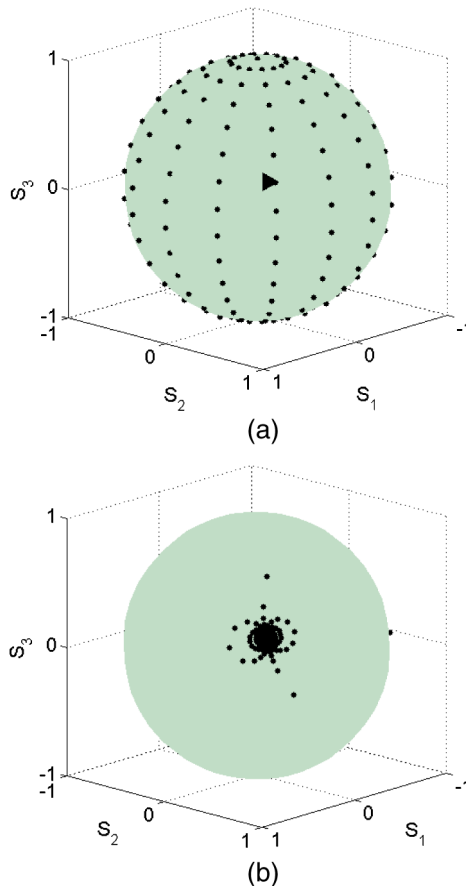


Fig. 4. (Color online) Tips of input (a) and output (b) signal Stokes vectors on the Poincaré sphere for a fiber length $L = 500$ m, $D_p = 0.75$ ps km $^{-1/2}$, and $\Delta\nu = 0.255$ THz. For the sake of clarity, only 225 vectors are represented instead of the 10 000 used in the simulations. Input vectors are distributed uniformly over the sphere. The black triangle represents the input pump Stokes vector.

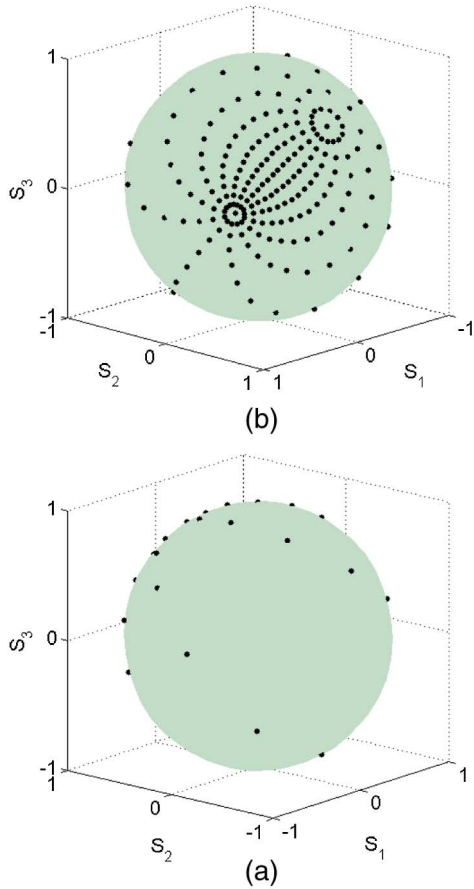


Fig. 5. (Color online) Distribution of the output signal Stokes vectors with $L = 500$ m and $\Delta\nu = 0.350$ THz. Panels (a) and (b) display opposite views of the Poincaré sphere.

It is useful to visualize the effectiveness of polarization attraction by means of parametric gain or FWM by plotting on the Poincaré sphere the end points of the Stokes vectors corresponding to either the input or the output distributions of signal SOPs, corresponding to the results of Figs. 1–3. In Fig. 4, we compare the distribution of input signal SOPs, which uniformly covers the sphere [Fig. 4(a)], to the output signal SOP distribution [Fig. 4(b)] from a fiber of length $L = 500$ m with $D_p = 0.75$ ps km^{-1/2}; here the signal detuning is $\Delta\nu = 0.255$ THz. These parameters correspond to the sideband detuning for peak signal gain (see the dotted curve in Fig. 2). As shown by the solid curve in Fig. 3, the output DOP is as high as 0.97, which means a nearly full attraction towards the input pump Stokes vector \vec{S}_p . On the other hand, in Fig. 5, we show the output distribution of signal SOPs when the sideband detuning is increased up to $\Delta\nu = 0.350$ THz. Figure 3 shows that the output DOP is only 0.73 in this case, which results in a relatively poor polarization attraction. It is important to point out that the polarization attraction (to the pump SOP) behavior that is illustrated in Figs. 4 and 5 does not depend upon the specific input pump SOP that is selected; indeed, the strength of polarization attraction only depends on the pump power level.

5. CONCLUSIONS

In our study we proposed and analyzed a novel type of nonlinear polarizer, exploiting the degenerate FWM process or

parametric optical amplification in a standard telecom fiber with randomly varying birefringence. In the FWM polarizer, the SOP of the amplified signal (or idler) beam is attracted to the SOP of the copropagating, fully polarized pump wave. We have derived the coupled wave equations that describe the propagation of the pump, the idler, and the signal in the presence of weak PMD. Our model substantially extends previous theory of FWM in optical fibers, since it may analytically describe the rate of degradation of FWM efficiency and polarization attraction for low-PMD fibers. Knowing the spatial rate of PMD-induced degradation permits the proper design of practical nonlinear polarizers based on optical parametric amplification in kilometer-long nonlinear-optical fibers. Polarization attraction and control by parametric amplification in fibers is potentially applicable to frequency-conversion and phase-sensitive amplification devices when combined with polarization-sensitive optical processing devices (e.g., a heterodyne receiver). In addition, codirectional parametric repolarizers based on low-PMD telecom fibers may be used for compensating ultrafast input signal SOP fluctuations. Although FWM-based polarizers suffer from output RIN, RIN suppression could be obtained when operating the amplifier in the depleted pump regime, as it occurs with Raman polarizers [36].

APPENDIX A: STOCHASTIC THEORY OF PARAMETRIC AMPLIFICATION

Our goal is to convert the initial equations for the field (1) and (2) with stochastic coefficients into corresponding equations with deterministic coefficients. In other words, we need to find a way to average the initial equations over the ensemble of fibers, which represents all possible realizations of the random fiber birefringence with a given statistics. Since both initial and final equations are nonlinear, our procedure cannot be done exactly and it will require a number of approximations. Thus, the final equations will have a limited range of applicability.

We use the approach first introduced into the fiber-optics theory by Wai and Menyuk in [33]. This approach was formulated for a single beam (or pulse), and lead to the derivation of the celebrated Manakov equation and its generalization in the form of the Manakov-PMD equation. An extension of this theory for the two-beam configuration was undertaken in [8,17] and led to the formulation of the theoretical basis of XPoIM-induced polarization attraction effect in telecom fibers and of Raman polarizers. Here, we need to extend this theory even further by fully taking into account the three interacting beams. Given that all these theories have very much in common, we shall omit many repetitions and where appropriate we simply refer to prior literature for more details.

We start with the transformation of field vectors from the laboratory (x, y) frame into the local reference frame $(1, 2)$, which is defined by the z -dependent orientation of the axis of birefringence: $\Psi_f = M(z)U_f$, where $M(z)$ is the 2×2 rotation matrix defined in Eq. (4) of [8]. Here $\Psi_f = (\psi_{1f}, \psi_{2f})$. All terms except one in the field equations stay immune to this transformation. The only change is the form of the birefringence matrix, which now becomes

$$\overline{\Delta B}(\omega_f) = \begin{pmatrix} b(\omega_f) & \mp \frac{i}{2}g_\theta \\ \pm \frac{i}{2}g_\theta & -b(\omega_f) \end{pmatrix}. \quad (\text{A1})$$

The next transformation: $\Phi_f = T_f(z)\Psi_f$, is aimed at the decoupling of the linear portions of the field equations. This goal is reached if the transformation matrix

$$T_p(z) = \begin{pmatrix} a_1(z) & a_2(z) \\ -a_2^*(z) & a_1^*(z) \end{pmatrix} \quad (\text{A2})$$

obeys the following equation:

$$i \frac{\partial T_p}{\partial z} + \overline{\Delta B}_p \cdot T_p = 0. \quad (\text{A3})$$

Matrices T_s and T_i are defined in a similar way, with $b_{1,2}$ and $c_{1,2}$ elements used instead of $a_{1,2}$. The unitarity of this transformation is preserved by requiring that $|a_1(z)|^2 + |a_2(z)|^2 = |b_1(z)|^2 + |b_2(z)|^2 = |c_1(z)|^2 + |c_2(z)|^2 = 1$. Initial conditions for the elements of the $T_f(z)$ matrices are to be determined from the requirement that $\Phi_f = \Psi_f$ at $z = 0$. Thus, $a_1(0) = b_1(0) = c_1(0) = 1$ and $a_2(0) = b_2(0) = c_2(0) = 0$.

The transformation associated with the $T_f(z)$ matrix brings the equations for three fields in the form

$$i \frac{\partial \Phi_f}{\partial z} + \gamma(N_{spm} + N_{xpm} + N_{ex})_f = 0. \quad (\text{A4})$$

As expected, in this reference frame, the fields are coupled by nonlinearity only through three types of cubic terms: SPoLM terms N_{spm} , XPolM terms N_{xpm} , and energy-exchange terms N_{ex} . The number of these nonlinear terms is very large, and we do not provide here their detailed structure. Instead, we refer to Eqs. (9)–(12) in [8], where the SPoLM and XPolM nonlinear terms are written down explicitly. In our present theory, we have all these terms as well, and in addition get energy-exchange terms in the form of cubic products involving three different fields.

Coefficients prior to these terms are some self- and cross-fourth-order polynomials composed of $a_{1,2}(z)$, $b_{1,2}(z)$, $c_{1,2}(z)$, and their complex conjugates. It is convenient to work with quadratic coefficients u_m and u_m^* ($m = 1 \div 30$). Coefficients u_m with $m = 1 \div 14$ are identical to those introduced immediately below Eq. (12) in [8]. They are divided into self-terms— $u_1 = |a_1|^2 - |a_2|^2$, $u_2 = -(a_1 a_2 + a_1^* a_2^*)$, $u_3 = i(a_1 a_2 - a_1^* a_2^*)$, $u_4 = 2a_1 a_2^*$, $u_5 = a_1^2 - a_2^2$, $u_6 = -i(a_1^2 + a_2^2)$ —and cross-terms— $u_7 = a_1^* b_1 - a_2 b_2^*$, $u_8 = -(b_1 a_2 + b_2^* a_1^*)$, $u_9 = i(b_1 a_2 - a_1^* b_2^*)$, $u_{10} = -i(a_1^* b_1 + a_2 b_2^*)$, $u_{11} = a_1 b_2^* + b_1 a_2^*$, $u_{12} = a_1 b_1 - a_2^* b_2^*$, $u_{13} = -i(a_1 b_1 + a_2^* b_2^*)$, and $u_{14} = i(a_1 b_2^* - a_2^* b_1)$. The three-beam theory additionally brings 16 new coefficients. Coefficients u_m with $m = 15 \div 22$ are the same as u_m with $m = 7 \div 14$ but with $b_{1,2}$ replaced with $c_{1,2}$. Coefficients u_m with $m = 23 \div 30$ are the same as u_m with $m = 7 \div 14$, where $a_{1,2}$ is replaced with $b_{1,2}$, and simultaneously $b_{1,2}$ is replaced with $c_{1,2}$.

Nonlinear coefficients in Eq. (A3) are products of the type $u_m u_n$ or $u_m u_n^*$. They are z -dependent random coefficients, because they depend on the stochastic variable $g_\theta(z)$. We need to find average values of all nonlinear terms, which are of the form, for instance, $u_9^2 \phi_{1s} \phi_{2p} \phi_{2s}^*$. This is the place where the most important approximation comes into play. We assume that the following factorization is valid: $\langle u_9^2 \phi_{1s} \phi_{2p} \phi_{2s}^* \rangle \approx \langle u_9^2 \rangle \langle \phi_{1s} \phi_{2p} \phi_{2s}^* \rangle$. This factorization is justified whenever the spatial evolution of the fields is much slower than the spatial evolution of the nonlinear coefficients, or vice versa, whenever the spatial evolution of the fields is much faster than

the spatial evolution of the nonlinear coefficients. In the context of parametric amplification and in the absence of group-velocity dispersion, the nonlinear evolution of the fields scales with the nonlinear length $L_{NL} = (\gamma P_p)^{-1}$, where P_p is the pump power. In its turn, the z -dependence of $\langle u_m u_n \rangle$ or $\langle u_m u_n^* \rangle$ is governed by two different length scales. On the one hand, we have the relatively short spatial scales that are associated with the correlation length L_c and the beat length L_B , both of which are typically less than 100 m. On the other hand, we have the relatively long spatial scale, which is associated with the PMD diffusion length L_d . For practically interesting situations, we need to provide the following hierarchy of scales: $L_c, L_B \ll L, L_{NL} \ll L_d$. In this range, the factorization approximation is well justified; with this limitation in mind, we may proceed further.

In order to find averages of the type u_m^2 and $|u_m|^2$, it is convenient to group coefficients as $G_1 = \{u_1, u_2, u_3\}$, $G_2 = \{u_4, u_5, u_6\}$, $G_3 = \{u_7, u_8, u_9, u_{10}\}$, $G_4 = \{u_{11}, u_{12}, u_{13}, u_{14}\}$, $G_5 = \{u_{15}, u_{16}, u_{17}, u_{18}\}$, $G_6 = \{u_{19}, u_{20}, u_{21}, u_{22}\}$, $G_7 = \{u_{23}, u_{24}, u_{25}, u_{26}\}$, and $G_8 = \{u_{27}, u_{28}, u_{29}, u_{30}\}$. For each group we were able to formulate a closed system of linear first-order differential equations by using Eq. (A3). For an example of such a system, we may refer to Eq. (13) in [8].

Next we need to know the average values of quadratic forms composed by these coefficients. They can be found from the solutions to the equations of motion for the average of the generic function F . For instance, for $F(u_1, u_2, u_3, \theta)$, we need to solve the equation $\partial_z \langle F \rangle = \langle G(F) \rangle$. The generator G is to be constructed by a procedure described in the appendix of [33]. For a specific example of G , we may refer to Eqs. (A1) and (A7) in [8]. Note also that the average over different realizations of the fiber birefringence can be replaced by a spatial average as

$$\langle f \rangle = \lim_{z \rightarrow \infty} \frac{1}{z} \int_0^z dz' f(z'), \quad (\text{A5})$$

by assuming that the ergodicity hypothesis is valid.

With this procedure at hand, we are able to find the mean values of u_m^2 , with $m = 9, 10, 13, 14, 17, 18, 21, 22, 25, 26, 29, 30$ by solving the equation $\partial_z V_A = M_A V_A$ for the vector $V_A = (\langle S_1^2 \rangle, \langle S_2^2 \rangle, \langle S_3^2 \rangle, \langle S_4^2 \rangle, \langle S_2 S_3 \rangle, \langle S_1 S_4 \rangle)^T$, where $\{S_1, S_2, S_3, S_4\}$ is any of the groups G_i with $i = 4 \div 8$, and with the matrix M_A given by

$$\begin{pmatrix} -2L_c^{-1} & 2L_c^{-1} & 0 & 0 & 0 & 2\Delta^{(-)} \\ 2L_c^{-1} & -2L_c^{-1} & 0 & 0 & -2\Delta^{(+)} & 0 \\ 0 & 0 & 0 & 0 & 2\Delta^{(+)} & 0 \\ 0 & 0 & 0 & 0 & 0 & -2\Delta^{(-)} \\ 0 & \Delta^{(+)} & -\Delta^{(+)} & 0 & -L_c^{-1} & 0 \\ -\Delta^{(-)} & 0 & 0 & \Delta^{(-)} & 0 & -L_c^{-1} \end{pmatrix}. \quad (\text{A6})$$

Here $\Delta^{(+)} = b(\omega_p) + b(\omega_i)$ and $\Delta^{(-)} = b(\omega_p) - b(\omega_i)$. It is a straightforward calculation to get an estimate $\Delta^{(-)} / \Delta^{(+)} \sim \Delta \omega / \omega_p$, where $\Delta \omega = \omega_p - \omega_i = \omega_s - \omega_p$. For typical fiber parameters, the evolution associated with $\Delta^{(+)}$ is very fast, while $\Delta^{(-)}$ defines a much slower spatial scale. Setting $\Delta^{(-)}$ to zero defines the Manakov limit, and brings us back to the formulation of coupled wave equations with constant in z nonlinear coefficients. The difference of $\Delta^{(-)}$ from zero

means the inclusion of effects caused by the PMD. In this case we are dealing with z -dependent nonlinear coefficients.

Next we calculate the averages of the type $|u_m|^2$, with $m = 9, 10, 13, 14, 17, 18, 21, 22, 25, 26, 29, 30$. Thereto we formulate the equation of motion $\partial_z V_B = M_B V_B$ for the vector $V_B = (\langle |S_1|^2 \rangle, \langle |S_2|^2 \rangle, \langle |S_3|^2 \rangle, \langle |S_4|^2 \rangle, \langle S_2 S_3^* \rangle, \langle S_2^* S_3 \rangle, \langle S_1 S_4^* \rangle, \langle S_1^* S_4 \rangle)^T$, where $\{S_1, S_2, S_3, S_4\}$ is any of the groups G_i , with $i = 4 \div 8$, and where the matrix M_B reads as

$$\begin{pmatrix} -2L_c^{-1} & 2L_c^{-1} & 0 & 0 & 0 & 0 & \Delta^{(-)} & \Delta^{(-)} \\ 2L_c^{-1} & -2L_c^{-1} & 0 & 0 & -\Delta^{(+)} & -\Delta^{(+)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta^{(+)} & \Delta^{(+)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\Delta^{(-)} & -\Delta^{(-)} \\ 0 & \Delta^{(+)} & -\Delta^{(+)} & 0 & -L_c^{-1} & 0 & 0 & 0 \\ 0 & \Delta^{(+)} & -\Delta^{(+)} & 0 & 0 & -L_c^{-1} & 0 & 0 \\ -\Delta^{(-)} & 0 & 0 & \Delta^{(-)} & 0 & 0 & -L_c^{-1} & 0 \\ -\Delta^{(-)} & 0 & 0 & \Delta^{(-)} & 0 & 0 & 0 & -L_c^{-1} \end{pmatrix}. \quad (A7)$$

Note that initial conditions for the averages of the type $\langle u_m(z)u_n(z) \rangle$ and $\langle u_m(z)u_n^*(z) \rangle$ can be found from the initial conditions for the coefficients $a_{1,2}$, $b_{1,2}$, and $c_{1,2}$, and by observing that $\langle u_m(0)u_n(0) \rangle = u_m(0)u_n(0)$. Thus we find $u_{1,5,7,12,15,20,25,28}(0) = 1$ and $u_{6,10,13,17,21,26,29}(0) = -i$, while the remaining coefficients are all zero.

Next, we turn to cross-terms like $\langle u_m u_n \rangle$ with $m \neq n$. Many of these terms are zero, mainly because of the imposed zero initial conditions. Nonzero coefficients are $\langle u_{14}u_{22} \rangle$, $\langle u_{14}u_{22}^* \rangle$, $\langle u_{10}u_{18} \rangle$, $\langle u_{10}u_{18}^* \rangle$, $\langle u_6u_{29} \rangle$, $\langle u_6u_{29}^* \rangle$, $\langle u_3u_{25} \rangle$, and $\langle u_3u_{25}^* \rangle$. The first four of these coefficients can be found by solving the equation $\partial_z V_B = M_B V_B$ with the matrix M_B defined as in Eq. (A7), and where the vector V_B is identified with any of the following vectors:

$$\begin{aligned} & (\langle u_{11}u_{19} \rangle, \langle u_{12}u_{20} \rangle, \langle u_{13}u_{21} \rangle, \langle u_{14}u_{22} \rangle, \langle u_{13}u_{20} \rangle, \langle u_{12}u_{21} \rangle, \\ & \langle u_{14}u_{19} \rangle, \langle u_{11}u_{22} \rangle)^T; \\ & (\langle u_{11}u_{19}^* \rangle, \langle u_{12}u_{20}^* \rangle, \langle u_{13}u_{21}^* \rangle, \langle u_{14}u_{22}^* \rangle, \langle u_{13}u_{20}^* \rangle, \langle u_{12}u_{21}^* \rangle, \\ & \langle u_{14}u_{19}^* \rangle, \langle u_{11}u_{22}^* \rangle)^T; \\ & (\langle u_7u_{15} \rangle, \langle u_8u_{16} \rangle, \langle u_9u_{17} \rangle, \langle u_{10}u_{18} \rangle, \langle u_9u_{16} \rangle, \langle u_8u_{17} \rangle, \\ & \langle u_{10}u_{15} \rangle, \langle u_7u_{18} \rangle)^T; \\ & (\langle u_7u_{15}^* \rangle, \langle u_8u_{16}^* \rangle, \langle u_9u_{17}^* \rangle, \langle u_{10}u_{18}^* \rangle, \langle u_9u_{16}^* \rangle, \langle u_8u_{17}^* \rangle, \\ & \langle u_{10}u_{15}^* \rangle, \langle u_7u_{18}^* \rangle)^T. \end{aligned}$$

In turn, the coefficients $\langle u_6u_{29} \rangle$, $\langle u_6u_{29}^* \rangle$, $\langle u_3u_{25} \rangle$, and $\langle u_3u_{25}^* \rangle$ can be found from the equation $\partial_z V_C = M_C V_C$ with the matrix M_C defined as

$$\begin{pmatrix} 0 & 0 & \Delta^{(+)} & \Delta^{(+)} & 0 & 0 \\ -\Delta^{(+)} & -L_c^{-1} & 0 & \Delta^{(+)} & 0 & 0 \\ -\Delta^{(+)} & 0 & -L_c^{-1} & \Delta^{(+)} & 0 & 0 \\ 0 & -\Delta^{(+)} & -\Delta^{(+)} & -2L_c^{-1} & 2L_c^{-1} & 0 \\ 0 & 0 & 0 & 0 & 2L_c^{-1} & -2L_c^{-1} \\ 0 & 0 & 0 & 0 & -\Delta^{(-)} & -L_c^{-1} \end{pmatrix} \quad (A8)$$

when we associate the vector V_C with any of the following vectors:

$$\begin{aligned} & (\langle u_6u_{29} \rangle, \langle u_5u_{29} \rangle, \langle u_6u_{28} \rangle, \langle u_5u_{28} \rangle, \langle u_4u_{27} \rangle, \langle u_4u_{30} \rangle)^T; \\ & (\langle u_6u_{29}^* \rangle, \langle u_5u_{29}^* \rangle, \langle u_6u_{28}^* \rangle, \langle u_5u_{28}^* \rangle, \langle u_4u_{27}^* \rangle, \langle u_4u_{30}^* \rangle)^T; \\ & (\langle u_3u_{25} \rangle, \langle u_3u_{24} \rangle, \langle u_2u_{25} \rangle, \langle u_2u_{24} \rangle, \langle u_1u_{23} \rangle, \langle u_1u_{26} \rangle)^T; \\ & (\langle u_3u_{25}^* \rangle, \langle u_3u_{24}^* \rangle, \langle u_2u_{25}^* \rangle, \langle u_2u_{24}^* \rangle, \langle u_1u_{23}^* \rangle, \langle u_1u_{26}^* \rangle)^T. \end{aligned}$$

APPENDIX B: ANALYTIC ESTIMATION OF THE NONLINEAR COEFFICIENTS

In this appendix, we look for approximate analytical solutions to the linear systems of equations for the vectors V_A , V_B , and V_C . This task is equivalent to finding the eigenvalues and eigenvectors of the matrices $M_{A,B,C}$. Additionally, we need to find the decomposition of the initial vectors $V_{A,B,C}(0)$ in the basis of the corresponding eigenvectors. In this way, we may determine the z -dependence of the nonlinear coefficients.

We shall give a detailed analysis for the M_A matrix, and sketch only briefly the results for the other matrices. We develop a perturbative approach, by assuming that $\Delta^{(-)}$ is much smaller than $\Delta^{(+)}$ and L_c^{-1} . First, setting $\Delta^{(-)}$ to zero, we get a much simpler matrix \tilde{M}_A . The difference $\Delta M_A = M_A - \tilde{M}_A$ is therefore a small correction. The matrix \tilde{M}_A has a doubly degenerate eigenvalue $\tilde{\lambda}_A = 0$ and two corresponding eigenvectors $\tilde{e}_{A1} = (0, 0, 0, 1, 0, 0)^T$ and $\tilde{e}_{A2} = (1, 1, 1, 0, 0, 0)^T$. The other eigenvalues of \tilde{M}_A all have relatively large negative real parts, in the sense that the corresponding eigenvectors vanish with distance very quickly. The spatial scale of this decay is determined by the correlation length L_c and the beat length L_B , both of which are typically less than 100 m. So, the characteristic decay rate is estimated as $L_{\text{transient}} \sim 100$ m. After the transient decay is over, we can write the solution of $\partial_z \tilde{V}_A = \tilde{M}_A \tilde{V}_A$ as $\tilde{V}_A(z) = (\tilde{C}_1 \tilde{e}_{A1} + \tilde{C}_2 \tilde{e}_{A2}) \exp(\tilde{\lambda}_A z) = \tilde{C}_1 \tilde{e}_{A1} + \tilde{C}_2 \tilde{e}_{A2}$, where $\tilde{C}_1 = (V_A(0) \cdot \tilde{e}_{A1}^*) / (\tilde{e}_{A1} \cdot \tilde{e}_{A1}^*)$ and $\tilde{C}_2 = (V_A(0) \cdot \tilde{e}_{A2}^*) / (\tilde{e}_{A2} \cdot \tilde{e}_{A2}^*)$, thanks to the orthogonality of the set of eigenvectors of \tilde{M}_A .

When $\Delta^{(-)}$ is different from zero, the degeneracy is lifted and the doubly degenerate eigenvalue $\tilde{\lambda}_A$ split into two different eigenvalues $\lambda_{A1} = 0$ and $\lambda_{A2} \neq 0$. Let us find λ_{A2} by way of developing the perturbative analysis. First we find the eigenvalue equation for the exact M_A matrix. It is $\det(M_A - \lambda I) = 0$, where I is the unity matrix and

$$\begin{aligned} \det(M_A - \lambda I) &= 32[\Delta^{(-)}]^2[\Delta^{(+)}]^2 L_c^{-1} \lambda \\ &+ (16[\Delta^{(-)}]^2[\Delta^{(+)}]^2 + 12[\Delta^{(-)}]^2 L_c^{-2} \\ &+ 12[\Delta^{(+)}]^2 L_c^{-2}) \lambda^2 \\ &+ (16[\Delta^{(-)}]^2 L_c^{-1} + 16[\Delta^{(+)}]^2 L_c^{-1} + 4L_c^{-3}) \lambda^3 \\ &+ (4[\Delta^{(-)}]^2 + 4[\Delta^{(+)}]^2 + 9L_c^{-2}) \lambda^4 + 6L_c^{-1} \lambda^5 \\ &+ \lambda^6. \end{aligned}$$

Since $\Delta^{(-)}$ is small, we expect that the correction to the unperturbed zero eigenvalue $\tilde{\lambda}_A$ is also small. By keeping in the eigenvalue equation terms no higher than second order in λ , we get after some simplifications the approximated solution $\lambda_{A2} \approx -(8/3)(\Delta^{(-)})^2 L_c = -L_d^{-1}$. The eigenvector corresponding to eigenvalue λ_{A1} (λ_{A2}) is $e_{A1} = \tilde{e}_{A1} + \tilde{e}_{A2}$ (e_{A2}). The perturbed solution is $V_A(z) = C_1 e_{A1} \exp(\lambda_{A1} z) + C_2 e_{A2} \exp(\lambda_{A2} z) = C_1 e_{A1} + C_2 e_{A2} \exp(-z/L_d)$. Here $C_1 = (V_A(0) \cdot e_{A1}^*) / (e_{A1} \cdot e_{A1}^*)$. The exact expression for $C_2 e_{A2}$ is cumbersome; however, under the condition $\Delta^{(-)} \ll \Delta^{(+)}, L_c^{-1}$, we can use the equality

of \tilde{V}_A and V_A in the limit of $\Delta^{(-)} \rightarrow 0$, and write $C_2 e_{A2} = \tilde{C}_1 \tilde{e}_{A1} + \tilde{C}_2 \tilde{e}_{A2} - C_1 e_{A1}$.

Now we can turn to evaluation of the nonlinear coefficients. Let us start with the averages $\langle u_9^2 \rangle$ and $\langle u_{10}^2 \rangle$. Coefficients u_9 and u_{10} belong to the group of coefficients denoted earlier as G_3 . For this group, the vector V_A contains $\langle u_9^2 \rangle$ and $\langle u_{10}^2 \rangle$ as the third and the fourth element, respectively. The initial condition reads as $V_A(0) = (1, 0, 0, -1, 0, i)^T$. Thus, for $L \geq L_{\text{transient}}$, we find $\langle u_9^2 \rangle = (1/3) \exp(-z/L_d)$ and $\langle u_{10}^2 \rangle = -\exp(-z/L_d)$. Similarly, we find $\langle u_{17}^2 \rangle = \langle u_{25}^2 \rangle = \langle u_9^2 \rangle$ and $\langle u_{18}^2 \rangle = \langle u_{26}^2 \rangle = \langle u_{10}^2 \rangle$. With initial conditions $V_A(0) = (0, 1, -1, 0, -i, 0)^T$, we get $\langle u_{13}^2 \rangle = \langle u_{14}^2 \rangle = \langle u_{21}^2 \rangle = \langle u_2^2 \rangle = \langle u_{29}^2 \rangle = \langle u_{30}^2 \rangle = 0$.

The matrix M_B can be considered similarly. Again, in the limit $\Delta^{(-)} \rightarrow 0$, this matrix possesses a doubly degenerate zero eigenvalue $\tilde{\lambda}_B$, while the other eigenvalues have large negative real parts, so that the corresponding eigenvectors vanish after a certain propagating distance, say $L_{\text{transient}}$. Whenever $\Delta^{(-)}$ is different from zero, the degeneracy is lifted and the doubly degenerate eigenvalue $\tilde{\lambda}_B$ is split into $\lambda_{B1} = 0$ and $\lambda_{B2} = -L_d$. Thus, for $L \geq L_{\text{transient}}$, we find $\langle |u_9|^2 \rangle = \langle |u_{17}|^2 \rangle = \langle |u_{25}|^2 \rangle = (1/2) - (1/6) \exp(-z/L_d)$, $\langle |u_{10}|^2 \rangle = \langle |u_{18}|^2 \rangle = \langle |u_{26}|^2 \rangle = (1/2) + (1/2) \exp(-z/L_d)$, $\langle |u_{13}|^2 \rangle = \langle |u_{21}|^2 \rangle = \langle |u_{29}|^2 \rangle = (1/2) + (1/6) \exp(-z/L_d)$, and $\langle |u_{14}|^2 \rangle = \langle |u_{22}|^2 \rangle = \langle |u_{30}|^2 \rangle = (1/2) - (1/2) \exp(-z/L_d)$. By using the same eigenvectors and eigenvalues of matrix M_B , we find also that $\langle u_{14} u_{22} \rangle = 0$, $\langle u_{14} u_{22}^* \rangle = (1/2) - (1/2) \exp(-z/L_d)$, $\langle u_{10} u_{18} \rangle = -\exp(-z/L_d)$, and $\langle u_{10} u_{18}^* \rangle = (1/2) + (1/2) \exp(-z/L_d)$.

Finally, matrix M_C possesses a nondegenerate eigenvalue $\tilde{\lambda}_C = 0$ in the limit $\Delta^{(-)} = 0$, with the other eigenvalues vanishing for $z \geq L_{\text{transient}}$. The perturbative approach yields the correction to the zero eigenvalue: $\lambda_C = -(1/8)L_d^{-1}$. Then, for $L \geq L_{\text{transient}}$ we find $\langle u_6 u_{29} \rangle = 0$, $\langle u_6 u_{29}^* \rangle = (2/3) \exp(-z/L_d)$, $\langle u_3 u_{25} \rangle = \langle u_3 u_{25}^* \rangle = (1/3) \exp(-z/L_d)$.

When all these nonlinear coefficients are substituted in the equations for the field, we arrive to the final result: Eqs. (3)–(5), which represent equations with deterministic coefficients, as desired.

ACKNOWLEDGMENTS

This work was carried out with support from the Italian Ministry of Research and the University (MIUR) through the grant 2008MPSSNX.

REFERENCES

- J. E. Heebner, R. S. Bennink, R. W. Boyd, and R. A. Fisher, "Conversion of unpolarized light to polarized light with greater than 50% efficiency by photorefractive two-beam coupling," *Opt. Lett.* **25**, 257–259 (2000).
- S. Pitois, G. Millot, and S. Wabnitz, "Nonlinear polarization dynamics of counterpropagating waves in an isotropic optical fiber: theory and experiments," *J. Opt. Soc. Am. B* **18**, 432–443 (2001).
- S. Pitois, J. Fatome, and G. Millot, "Polarization attraction using counter-propagating waves in optical fiber at telecommunication wavelengths," *Opt. Express* **16**, 6646–6651 (2008).
- S. Pitois, A. Picozzi, G. Millot, H. R. Jauslin, and M. Haelterman, "Polarization and modal attractors in conservative counterpropagating four-wave interaction," *Europhys. Lett.* **70**, 88–94 (2005).
- J. Fatome, S. Pitois, P. Morin, and G. Millot, "Observation of light-by-light polarization control and stabilization in optical fibre for telecommunication applications," *Opt. Express* **18**, 15311–15317 (2010).
- V. V. Kozlov, K. Turitsyn, and S. Wabnitz, "Nonlinear repolarization in optical fibers: polarization attraction with copropagating beams," *Opt. Lett.* **36**, 4050–4052 (2011).
- P. Morin, J. Fatome, C. Finot, S. Pitois, R. Claveau, and G. Millot, "All-optical nonlinear processing of both polarization state and intensity profile for 40 Gbit/s regeneration applications," *Opt. Express* **19**, 17158–17166 (2011).
- V. V. Kozlov, J. Nuño, and S. Wabnitz, "Theory of lossless polarization attraction in telecommunication fibers," *J. Opt. Soc. Am. B* **28**, 100–108 (2011).
- V. V. Kozlov and S. Wabnitz, "Theoretical study of polarization attraction in high-birefringence and spun fibers," *Opt. Lett.* **35**, 3949–3951 (2010).
- D. Sugny, A. Picozzi, S. Lagrange, and H. R. Jauslin, "Role of singular tori in the dynamics of spatiotemporal nonlinear wave systems," *Phys. Rev. Lett.* **103**, 034102 (2009).
- E. Assémat, S. Lagrange, A. Picozzi, H. R. Jauslin, and D. Sugny, "Complete nonlinear polarization control in an optical fiber system," *Opt. Lett.* **35**, 2025–2027 (2010).
- S. Lagrange, D. Sugny, A. Picozzi, and H. R. Jauslin, "Singular tori as attractors of four-wave-interaction systems," *Phys. Rev. E* **81**, 016202 (2010).
- E. Assémat, D. Dargent, A. Picozzi, H. R. Jauslin, and D. Sugny, "Polarization control in spun and telecommunication optical fibers," *Opt. Lett.* **36**, 4038–4040 (2011).
- E. Assémat, A. Picozzi, H. R. Jauslin, and D. Sugny, "Hamiltonian tools for the analysis of optical polarization control," *J. Opt. Soc. Am. B* **29**, 559–571 (2012).
- M. Martinelli, M. Cirigliano, M. Ferrario, L. Marazzi, and P. Martelli, "Evidence of Raman-induced polarization pulling," *Opt. Express* **17**, 947–955 (2009).
- V. V. Kozlov, J. Nuño, J. D. Ania-Castañón, and S. Wabnitz, "Theory of fiber optic Raman polarizers," *Opt. Lett.* **35**, 3970–3972 (2010).
- V. V. Kozlov, Javier Nuño, Diego Juan Ania-Castañón, and S. Wabnitz, "Theoretical study of optical fiber Raman polarizers with counterpropagating beams," *J. Lightwave Technol.* **29**, 341–347 (2011).
- L. Ursini, M. Santagiustina, and L. Palmieri, "Raman nonlinear polarization pulling in the pump depleted regime in randomly birefringent fibers," *IEEE Photon. Technol. Lett.* **23**, 254–256 (2011).
- F. Chiarello, L. Ursini, L. Palmieri, and M. Santagiustina, "Polarization attraction in counterpropagating fiber Raman amplifiers," *IEEE Photon. Technol. Lett.* **23**, 1457–1459 (2011).
- V. V. Kozlov and S. Wabnitz, "Suppression of relative intensity noise in fiber-optic Raman polarizers," *IEEE Photon. Technol. Lett.* **23**, 1088–1090 (2011).
- S. Sergeev and S. Popov, "Two-section fiber optic Raman polarizer," *IEEE J. Quantum Electron.* **48**, 56–60 (2012).
- V. V. Kozlov, J. Nuño, J.-D. Ania-Castañón, and S. Wabnitz, *Trapping Polarization of Light in Nonlinear Optical Fibers: An Ideal Raman Polarizer* (Springer-Verlag, 2012).
- Nelson J. Muga, Mario F. S. Ferreira, and Armando N. Pinto, "Broadband polarization pulling using Raman amplification," *Opt. Express* **19**, 18707–18712 (2011).
- A. Zadok, E. Zilka, A. Eyal, L. Thevenaz, and M. Tur, "Vector analysis of stimulated Brillouin scattering amplification in standard single-mode fibers," *Opt. Express* **16**, 21692–21707 (2008).
- Q. Lin and G. P. Agrawal, "Vector theory of stimulated Raman scattering and its application to fiber-based Raman amplifiers," *J. Opt. Soc. Am. B* **20**, 1616–1631 (2003).
- V. V. Kozlov and S. Wabnitz, "Silicon Raman polarizer," *Opt. Lett.* **37**, 737–739 (2012).
- C. McKinstrie, H. Kogelnik, R. Jopson, S. Radic, and A. Kanaev, "Four-wave mixing in fibers with random birefringence," *Opt. Express* **12**, 2033–2055 (2004).
- M. Guasoni and S. Wabnitz, "Nonlinear polarizers based on four-wave mixing in high birefringence optical fibers," *J. Opt. Soc. Am. B* **29**, in press (2012).
- Q. Lin and G. P. Agrawal, "Vector theory of four-wave mixing: polarization effects in fiber-optic parametric amplifiers," *J. Opt. Soc. Am. B* **21**, 1216–1224 (2004).
- Q. Lin and G. P. Agrawal, "Effects of polarization-mode dispersion on fiber-based parametric amplification and wavelength conversion," *Opt. Lett.* **29**, 1114–1116 (2004).
- J. F. L. Freitas, C. J. S. de Matos, M. B. Costa e Silva, and A. S. L. Gomes, "Impact of phase modulation and parametric gain on

- signal polarization in an anomalously dispersive optical fiber," *J. Opt. Soc. Am. B* **24**, 1469–1474 (2007).
32. S. Wabnitz, "Broadband parametric amplification in photonic crystal fibers with two zero-dispersion wavelengths," *J. Lightwave Technol.* **24**, 1732–1738 (2006).
 33. P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightwave Technol.* **14**, 148–157 (1996).
 34. W. Magnus, "On the exponential solution of differential equations for a linear operator," *Commun. Pure Appl. Math.* **7**, 649–673 (1954).
 35. V. V. Kozlov, J. Nuño, and S. Wabnitz, "Theory of lossless polarization attraction in telecommunication fibers: erratum," *J. Opt. Soc. Am. B* **29**, 153–154 (2012).
 36. V. V. Kozlov and S. Wabnitz, "Suppression of relative intensity noise in fiber-optic Raman polarizers," *IEEE Photon. Technol. Lett.* **23**, 1088–1090 (2011).