# Theory of lossless polarization attraction in telecommunication fibers

**Victor V. Kozlov,[1,2,*] Javier Nuño,[3] and Stefan Wabnitz[1]**

[1]*Department of Information Engineering, Universitá di Brescia, Via Branze 38, 25123 Brescia, Italy*
[2]*Department of Physics, Saint-Petersburg State University, Petrodvoretz, St. Petersburg 198504, Russia*
[3]*Instituto de Optica, Consejo Superior de Investigaciones Cientificas, 28006 Madrid, Spain*
**Corresponding author: victor.kozlov@email.com*

Received August 27, 2010; accepted October 18, 2010;
posted October 21, 2010 (Doc. ID 134145); published December 14, 2010

In this work, polarization attraction is meant to be the conservative nonlinear effect that transforms any arbitrary input state of polarization (SOP) of an intense optical signal beam fed to a nonlinear medium into approximately one and the same SOP at the output, provided that the medium is driven by a relatively stronger counterpropagating pump beam. Essentially, the combination of the nonlinear medium and the pump beam serves as a lossless polarizer for the signal beam. The degree of polarization of the outcoming signal beam can be close to 100% (90% in our present simulations). With an eye toward the development of such lossless polarizers for fiber optics applications, we theoretically study the polarization attraction effect in the optical fibers that are used in telecommunication links; i.e., randomly birefringent fibers. A generic model for the fiber-based lossless polarizers is derived, and a statistical scheme for the quantification of their performance is proposed. 

*OCIS codes:* 230.5440, 060.4370, 230.1150, 230.4320.

## 1. INTRODUCTION

The most popular and robust optical element used to exercise control over the polarization state of a light beam is a linear polarizer. However, in many real-world applications, e.g., in telecommunications, such polarizers have limited use. The reason is that a linear polarizer transforms input beams with an essentially random distribution of polarizations into beams with a well-defined deterministic state of polarization (SOP) by wasting the orthogonal component. The overall 50% loss of energy inherent to this method can, in principle, be accepted. A much more serious problem is that, in the presence of signal polarization fluctuations and as a result of polarization-dependent loss, outgoing beams acquire large intensity fluctuations. Thus, the interconnection with optical devices that postprocess these strongly fluctuating signals becomes problematic, especially if these devices are nonlinear. A need for polarization control methods free from polarization-dependent losses comes to the forefront.

In this paper, we theoretically analyze such a method, in which the polarization state of the signal beam is tightly controlled by means of the nonlinear Kerr interaction with a counterpropagating pump or control beam with the same or similar frequency in a span of randomly birefringent fiber. In particular, we identify regimes in which an initial random distribution of polarizations for the signal beam is transformed at the output end of the fiber into a single well-defined deterministic SOP, as imposed by the SOP of the control beam. The beam evolution inside the fiber is unitary and, therefore, free from polarization-dependent losses, as desired. Sometimes such devices are called lossless polarizers or polarization funnels.

The possibility of achieving nonlinear or all-optical polarization control is rooted in the soliton theory, namely in the conservative structures, such as the polarization domain wall solitons that were earlier reported [1–6]. In that framework, one aims to demonstrate that polarization solitons may act as strong attractors, so that these soliton structures may describe the ultimate result of the evolution of a beam with an arbitrary initial SOP, which is the property that is expected from lossless polarizers. A large variety of soliton structures and quite general conclusions about their evolution in a nonlinear medium are available, thanks to the Hamiltonian nature of the problem and the power of the technique of the inverse scattering transform. The important theoretical assumption here is that the medium is not bounded, i.e., of infinite extent, whereas relatively little is known about solitons in bounded media.

Our interest here is exactly in describing beam propagation in a bounded medium with its associated appropriate boundary conditions for the beams at both boundaries. The method of the inverse scattering transform is well suited for Cauchy problems, but meets with difficulties when applied to boundary-value problems [7]. With the goal of checking the stability of standing-wave polarization solitons (which are known from the solution of the same problem but for an infinite medium) in the presence of boundaries, we solved the problem numerically [8]. We found that the soliton that is initially formed inside the medium and expected to stand still nevertheless moves toward the nearest boundary, collides with it, and disappears. Thus, such solitons are characterized by a finite lifetime in the bounded media and no longer serve as attractors. Instead of the solitons, we identified a different type of so-called polarization attractors. They are found to be strong attractors, representing the unique distribution of the SOPs of both beams inside the medium, which is the ultimate result of their evolution from arbitrary initial conditions. It is these polarization attractors that play a key role in the present study.

 

More information on polarization attractors can be found in [8].

The described stability analysis implies that the conclusions about the polarization dynamics that are derived from soliton theory for an infinite medium may lead to misleading findings. Thus, here we prefer to draw our conclusions entirely on the basis of numerical integration results, thus excluding any possible bias from soliton-theory intuition. Our numerical results may serve as a guide for the development of an analytic theory. As a matter of fact, some analytical results on lossless polarizers based on the four-wave mixing interaction of counterpropagating beams in an isotropic fiber have been reported in [9–12].

As stated in the beginning of this section, the development of lossless polarizers is a practice-driven problem, whose ultimate goal is the experimental verification and the technological development of simple and reliable devices. The first lossless polarizer was based on photorefractive two-beam coupling (experimentally demonstrated in [13]). The response time of nonlinearity in photorefractive materials is rather slow, which makes it difficult for them to be used in devices for high-speed telecommunications links. From this viewpoint, lossless polarizers based on the virtually instantaneous Kerr nonlinearity of optical fibers appear more suitable for telecommunications application needs. The experimental study of such devices was pioneered in Dijon: the first results were reported in [14] and followed by [15]. However, the need to use isotropic fibers (to avoid any deterministic or random birefringence-related effects) practically limited the length of the fiber to no more than a few meters. Obtaining a significant nonlinear interaction within such a short distance required beam powers that are prohibitively high from the viewpoint of telecommunications applications.

Practically relevant results have emerged only recently, when a lossless polarizer consisting of a 20 km randomly weakly birefringent fiber pumped by an incoherent counterpropagating control beam from the output end was demonstrated in [16]. The signal beam power was as low as 300 mW. The transition from the impractical isotropic fibers to the real-world randomly birefringent fibers was a major step, indicating that a delicate coherent four-wave mixing process, which is the necessary prerequisite for the polarization control, is present even in the randomly birefringent fibers. A proper theoretical description of these experiments is the main goal of this study. Here, we formulate our theoretical approach in such a form that it can be also used in future studies on nonlinear polarization control, in particular in different types of optical fibers and other nonlinear media. To provide an immediate example of this approach, we limit our present study to one model of the lossless polarizer, which is based on the nonlinear interaction of counterpropagating frequency-degenerate (or close in frequency) beams in randomly birefringent fibers.

So far, we have presented a brief overview of progress in the development of lossless polarizers, and we did not mention nonconservative schemes for achieving polarization control. The latter are based on the stimulated Raman or Brillouin scattering effects [17–19]. Here, the amplified probe beam acquires the SOP that is dictated by the pump beam because these scattering processes dominantly amplify a particular polarization mode. Essentially, the SOP of the probe beam is simply newly generated from the portion of the beam that has the polarization aligned with the SOP of the amplified mode, while the remaining part can be ignored as contributing very little to the amplified outcoming beam. This physically more trivial dynamics is very different from the intelligent transformation of the SOP of the entire beam, which is inherent to lossless polarizers.

## 2. MODEL

The model is based on four-wave equations for two counterpropagating beams. Two out of these four equations describe the evolution of the two amplitudes of the forward-propagating beam (signal) associated with its two orthogonal polarizations. Similarly, the other pair of equations describe the two amplitudes of the backward-propagating beam (control). All four equations are coupled by the nonlinear polarization. The medium response is typical for transparent weakly anisotropic dielectrics with inversion symmetry, for which the dominant nonlinear contribution comes from the Kerr effect. The slowly varying approximation that is usual for nonlinear optics is applied.

In addition to the separation of the fastest field evolution, which occurs on the scale of the optical wavelength, another important length scale in our problem is the fiber beat length, which is typically of the order of meters or tens of meters for most silica fibers. The beat length is the characteristic runaway distance of the phases in the orthogonal polarization components of each beam. Most nonlinear cross-polarization terms average to zero over the distance of a few kilometers, the overall length scale that is of interest to us here. However, some terms are not averaged to zero and provide the effect of polarization attraction. In this way, we arrive at the model of a deterministically linearly birefringent fiber, which is realistic only for distances less than 100 m, the scale of the correlation length $L_c$ of random birefringence variations in a typical fiber. On this scale, the fiber randomly changes the orientation of the axis of anisotropy. By properly taking into account this stochastic behavior, one can formulate the model of a randomly linearly birefringent fiber as an extension of the one-beam theory in [20] into the two-beam theory that is derived in detail in Section 3. The randomly birefringent fiber is the prototype of most telecommunications fibers.

Besides the two important length scales characterizing the fiber: the correlation length $L_c$ and the beat length $L_B$, the most critical for us is the length scale that is associated with the difference beat length $L_B'$. Namely, $L_B' = [L_B^{-1}(\omega_s) - L_B^{-1}(\omega_p)]^{-1}$, where $\omega_p(\omega_s)$ is the pump (signal) carrier frequency. The necessary prerequisite for the lossless nonlinear polarizer to be efficient is the fulfillment of the condition that $L_B'$ is much longer than the nonlinear length $L_{\mathrm{NL}}$ or the total fiber length $L$. Otherwise, the mutual polarizations of the beams are rapidly depolarized on the Poincaré sphere, thus damping the correlation effect that is induced by the nonlinear interaction of the two beams. From the mathematical point of view, we thus keep in the propagation equations four-wave mixing terms, which have exponential factors $\exp(z/L_B')$. In the frame of the local birefringence axes (see Section 3 for details), the evolution of the fields appears as slow motion with the characteristic length scale $L_{\mathrm{NL}}$. In this way, we arrive at the equations that can be conveniently formulated in Stokes space:

$$\partial_{\xi}\mathbf{S}^{+} = \mathbf{S}^{+} \times J_s\mathbf{S}^{+} + \mathbf{S}^{+} \times J_x\mathbf{S}^{-}, \tag{1}$$

$$\partial_{\eta}\mathbf{S}^{-} = \mathbf{S}^{-} \times J_s\mathbf{S}^{-} + \mathbf{S}^{-} \times J_x\mathbf{S}^{+}. \tag{2}$$

As a matter of fact, these equations represent in a unified form self- and cross-polarization interactions that arise in the different cases of isotropic, spun, and randomly birefringent fibers. In Section 3, we provide all details on the derivation of these equations for the model of randomly birefringent fibers. The only difference between the different fiber models is in the numerical values of the elements of the self- and cross-polarization matrices, $J_s$ and $J_x$, which have very simple forms in all three cases. We believe that Eqs. (1) and (2) are rather generic and arise in the consideration of any medium exhibiting Kerr nonlinearity. Only $J_s$ and $J_x$ are different in each specific situation. Therefore, these equations provide a simple mathematical model for studies of polarization attraction in a Kerr medium. Note that the difference in the structure of the $J_s$ and $J_x$ matrices sensitively affect the effect of polarization attraction, which may not necessarily be present at all in some instances. Even the opposite is true: namely, finding such matrices $J_s$ and $J_x$ that can produce the effect of polarization attraction for at least a single SOP of the pump beam is already a sort of success. Fortunately, in addition to the already known results for isotropic fibers, we have also observed the effect of polarization attraction in deterministically birefringent, randomly birefringent, and spun fibers.

Equations (1) and (2) govern the evolution of three-dimensional vectors $\mathbf{S}^{+} = (S_1^{+}, S_2^{+}, S_3^{+})$ and $\mathbf{S}^{-} = (S_1^{-}, S_2^{-}, S_3^{-})$ along the space coordinate $z$ in the course of time $t$. Here, wave variables are $\xi = (ct + z)/2$ and $\eta = (ct - z)/2$ with $c$ as the group speed of light in the medium. $\times$ denotes vector product. Vector $\mathbf{S}^{+}(z,t)$ governs the evolution of the SOP of the forward beam on the Poincaré sphere. Vector $\mathbf{S}^{-}(z,t)$ does the same for the backward beam. Thus, we prefer to work with Stokes vectors rather than the original amplitudes of the fields. The reason for this preference is that we are interested in the dynamics of the SOPs of the beams, which is naturally and conveniently visualized as trajectories on the Poincaré sphere.

In the case of isotropic fibers $J_s = \frac{1}{3}\gamma_{ss}\mathrm{diag}(-2,0,-2)$, $J_x = \frac{2}{3}\gamma_{ps}\mathrm{diag}(-2,0,-2)$, where $\gamma_{ss}$ and $\gamma_{ps}$ are the self-phase modulation (SPM) and cross-phase modulation (XPM) coefficients, each equal to the Kerr coefficient $\gamma$.

In the case of spun fibers (see [21]), $J_s = \frac{1}{3}\gamma_{ss}\mathrm{diag}(0, 0, 2\sin^2\phi - \cos^2\phi)$, $J_x = \frac{2}{3}\gamma_{ps}\mathrm{diag}(\cos^2\phi, -\cos^2\phi, -2\cos^2\phi)$, where $\phi$ is the artificial ellipticity that is induced in the fiber by spinning it during the drawing process: $\phi = \mathrm{atan}(\tau/\Delta\beta)$. Here, $\tau$ is the spinning period and $\Delta\beta$ is the value of linear birefringence. In the limit of zero spinning period ($\phi = 0$), we recover the case of deterministically birefringent fibers. The deterministic nature of the model implies that the fiber is not longer than the correlation length (<100 m).

In the case of randomly birefringent fibers, we obtain (see Section 3) $J_s = \gamma_{ss}\mathrm{diag}(0,0,0)$ and $J_x = \gamma_{ps}\frac{8}{9}\mathrm{diag}(-1,1,-1)$. It is this pair of tensors that we use throughout the rest of the paper.

## 3. NONLINEAR INTERACTION OF TWO BEAMS IN RANDOMLY BIREFRINGENT FIBERS

We start from the equation of motion for the pump field, written for the two-component field vector $U_p = (u_{px}, u_{py})^T$, where $u_{px}$ and $u_{py}$ are the amplitudes of the normal polarization modes $\mathbf{e}_x$ and $\mathbf{e}_y$ of the fiber: $\mathbf{U}_p = u_{px}\mathbf{e}_x + u_{py}\mathbf{e}_y$. This equation is derived under the (as usual for nonlinear optics) unidirectional and slowly varying approximations (see, for instance, [22]) and reads

$$\begin{aligned}\pm i\frac{\partial U_p}{\partial z} + i\beta'(\omega_p)\frac{\partial U_p}{\partial t} + \Delta B(\omega_p)U_p \\ + \gamma_{pp}\left[\frac{2}{3}(U_p^* \cdot U_p)U_p + \frac{1}{3}(U_p \cdot U_p)U_p^*\right] \\ + \frac{2}{3}\gamma_{ps}[(U_s^* \cdot U_s)U_p + (U_s \cdot U_p)U_s^* \\ + (U_p \cdot U_s^*)U_s] = 0.\end{aligned} \tag{3}$$

Upper + (lower −) correspond to the copropagating (counterpropagating) configuration. A similar equation (with indices $p$ and $s$ interchanged) arises for the signal beam, which is characterized by the field vector $U_s$.

The terms that are responsible for the Raman interaction are not included because, for beams of similar frequencies, which are of interest to us here, Raman gain is negligible. Here, $\gamma_{pp}$ and $\gamma_{ps}$ are self- and cross-modulation terms, whose values depend on frequency and, therefore, in principle, are different for the signal and pump beams. However, since the signal and pump beams are close in frequency to each other, we can safely approximate $\gamma_{pp}, \gamma_{ss}, \gamma_{ps}$, and $\gamma_{sp}$ by the same coefficient $\gamma$, which is the usual Kerr coefficient of the nonlinear fiber. $\beta'(\omega_{p,s})$ is the inverse group velocity of the pump/signal beam. $\Delta B(\omega_{p,s})$ is the birefringence tensor. For a linearly birefringent fiber, it takes the form $\Delta B(\omega_{p,s}) = \Delta\beta(\omega_{p,s})(\cos\theta\sigma_3 + \sin\theta\sigma_1)$, where $\Delta\beta(\omega_{p,s})$ is the value of birefringence at frequency $\omega_{p,s}$ and $\theta$ is the angle of orientation of the axis of the birefringence with respect to the reference frame defined by polarization modes $\mathbf{e}_x$ and $\mathbf{e}_y$. $\sigma_3$ and $\sigma_1$ are the usual Pauli matrices.

The orientation angle $\theta$ is randomly varying in fibers used for telecommunications applications, thus explaining the term randomly birefringent fibers that is applied to them. In principle, the magnitude of the birefringence $\Delta\beta$ also varies stochastically. However, as noticed in [20], the two approaches, one in which $\theta$ is the only stochastic variable, and the second, where both $\theta$ and $\Delta\beta$ are stochastic variables, produce nearly identical results. Thus, here we shall develop our theory by assuming the single stochastic variable $\theta$. Our theory is a generalization of the one-beam theory by Wai and Menyuk [20] to the case of two beams interacting via the Kerr nonlinearity.

The angle $\theta$ 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 at which $\theta$ changes randomly. Its value rarely exceeds 100 m in typical fibers.

We start with the transformation of field vectors into the local reference frame, which is defined by the $z$-dependent orientation of the axis of birefringence: $\Psi_{p,s} = M(z)U_{p,s}$, where

$$M(z)=\begin{pmatrix}\cos\frac{\theta}{2} & \sin\frac{\theta}{2}\\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2}\end{pmatrix}. \tag{4}$$

Here, $\Psi_p=(\Psi_{p1},\Psi_{p2})^T$ and $\Psi_s=(\Psi_{s1},\Psi_{s2})^T$. All terms except one in the master Eq. (3) stay immune to this transformation. The only change is the form of the birefringent matrix, which now becomes

$$\bar{\Delta B}(\omega_{p,s})=\begin{pmatrix}\Delta\beta(\omega_{p,s}) & \mp\frac{i}{2}g_\theta\\ \pm\frac{i}{2}g_\theta & -\Delta\beta(\omega_{p,s})\end{pmatrix}. \tag{5}$$

This new birefringent matrix defines two length scales, the correlation length $L_c$ and the beat length $L_B=2\pi/\Delta\beta$. Both scales are of the order of 100 m or less and, therefore, much shorter than the nonlinear length $L_{\mathrm{NL}}$ for powers of the beams of the order of 1 W and less. Therefore, it is reasonable to separate these length scales by making yet another transformation: $\Psi_{p,s}=T_{p,s}(z)\Phi_{p,s}$, where

$$T_p(z)=\begin{pmatrix}a_1 & a_2\\ -a_2^* & a_1^*\end{pmatrix},\qquad T_s(z)=\begin{pmatrix}b_1 & b_2\\ -b_2^* & b_1^*\end{pmatrix}, \tag{6}$$

$\Phi_p=(\Phi_{p1},\Phi_{p2})^T$, and $\Phi_s=(\Phi_{s1},\Phi_{s2})^T$. Here, $|a_1|^2+|a_2|^2=1$ and $|b_1|^2+|b_2|^2=1$, while matrices $T_p$ and $T_s$ obey the equation

$$\pm i\frac{\partial}{\partial z}T_{p,s}+\overline{\Delta B}(\omega_{p,s})T_{p,s}=0. \tag{7}$$

It is worth noting that this separation does not involve any approximations yet. In principle, the nonlinear length can be even shorter than $L_B$. The only limitation of our theory is the requirement that the least length between the total length of the fiber $L$ and the nonlinear length $L_{\mathrm{NL}}$ is substantially longer than the correlation length $\min(L_{\mathrm{NL}},L)\gg L_c$. In other respects, the theory we are developing is exact.

There is another length scale inherent to the problem. It is the difference beat length $L_B'=[L_B^{-1}(\omega_p)-L_B^{-1}(\omega_s)]^{-1}$. This scale is of particular importance to our study. Indeed, the effect of polarization attraction only takes place whenever $L\ll L_B'$. To satisfy this inequality, we need to choose $\omega_p$ close to $\omega_s$, although so far our theory is valid for any value of $L_B'$. However, for getting simple analytical results relevant to the problem of polarization attraction, we set $L_B'\to\infty$ on a later stage. The full-scale theory, valid for any value of $L_B'$ and also accounting for the Raman interaction, will be reported elsewhere.

With all transformations listed earlier, from Eq. (3) we get

$$\pm i\frac{\partial\Phi_p}{\partial z}+i\beta'(\omega_p)\frac{\partial\Phi_p}{\partial t}+\frac{1}{3}\gamma[2(\Phi_p^*\cdot\Phi_p)\Phi_p+\hat{N}_{spm}]+\frac{2}{3}\gamma[(\Phi_s^*\cdot\Phi_s)\Phi_p+\hat{N}_{xpm}]=0. \tag{8}$$

Here, $\hat{N}_{spm}=(N_{s1},N_{s2})^T$ and $\hat{N}_{xpm}=(N_{x1},N_{x2})^T$, with

$$N_{s1}=|\Phi_{p1}|^2\Phi_{p1}+u_3^2(2|\Phi_{p2}|^2-|\Phi_{p1}|^2)\Phi_{p1}-u_3u_6^*(2|\Phi_{p1}|^2-|\Phi_{p2}|^2)\Phi_{p2}-u_3u_6\Phi_{p1}^2\Phi_{p2}^*-u_6^{*2}\Phi_{p2}^2\Phi_{p1}^*, \tag{9}$$

$$N_{s2}=|\Phi_{p2}|^2\Phi_{p2}+u_3^2(2|\Phi_{p1}|^2-|\Phi_{p2}|^2)\Phi_{p2}+u_3u_6(2|\Phi_{p2}|^2-|\Phi_{p1}|^2)\Phi_{p1}+u_3u_6^*\Phi_{p2}^2\Phi_{p1}^*-u_6^2\Phi_{p1}^2\Phi_{p2}^*, \tag{10}$$

$$\begin{aligned}N_{x1}=&(|u_{13}|^2+|u_{10}|^2)|\Phi_{s1}|^2\Phi_{p1}-u_9u_{13}\Phi_{s1}\Phi_{p1}\Phi_{s2}^*\\&+u_{10}^*u_{14}^*\Phi_{s1}^*\Phi_{p1}\Phi_{s2}+(u_{10}u_{14}^*-u_9u_{13}^*)|\Phi_{s1}|^2\Phi_{p2}\\&+u_9^2\Phi_{s1}\Phi_{p2}\Phi_{s2}^*+u_{14}^{*2}\Phi_{s1}^*\Phi_{p2}\Phi_{s2}-u_9^*u_{13}^*\Phi_{s2}\Phi_{p1}\Phi_{s1}^*\\&+u_{14}u_{10}\Phi_{s2}^*\Phi_{p1}\Phi_{s1}+(|u_9|^2+|u_{14}|^2)|\Phi_{s2}|^2\Phi_{p1}\\&-u_{13}^{*2}\Phi_{s2}\Phi_{p2}\Phi_{s1}^*-u_{10}^2\Phi_{s2}^*\Phi_{p2}\Phi_{s1}\\&\times(u_9u_{13}^*-u_{10}u_{14}^*)|\Phi_{s2}|^2\Phi_{p2},\end{aligned} \tag{11}$$

$$\begin{aligned}N_{x2}=&(|u_{13}|^2+|u_{10}|^2)|\Phi_{s2}|^2\Phi_{p2}+u_9^*u_{13}^*\Phi_{s2}\Phi_{p2}\Phi_{s1}^*\\&-u_{10}u_{14}\Phi_{s2}^*\Phi_{p2}\Phi_{s1}+(-u_{10}^*u_{14}+u_9^*u_{13}^*)|\Phi_{s2}|^2\Phi_{p1}\\&+u_9^{*2}\Phi_{s2}\Phi_{p1}\Phi_{s1}^*+u_{14}^2\Phi_{s2}^*\Phi_{p1}\Phi_{s1}+u_9u_{13}\Phi_{s1}\Phi_{p2}\Phi_{s2}^*\\&-u_{14}^*u_{10}\Phi_{s1}^*\Phi_{p2}\Phi_{s2}+(|u_9|^2+|u_{14}|^2)|\Phi_{s1}|^2\Phi_{p2}\\&-u_{13}^2\Phi_{s1}\Phi_{p1}\Phi_{s2}^*-u_{10}^{*2}\Phi_{s1}^*\Phi_{p1}\Phi_{s2}\\&\times(-u_9^*u_{13}+u_{10}^*u_{14})|\Phi_{s1}|^2\Phi_{p1}.\end{aligned} \tag{12}$$

The coefficients $u_i$ with $i=1,\ldots,14$ are the quadratic forms of the coefficients $a_1$, $a_2$, $b_1$, and $b_2$ and their complex conjugates. They are divided into self-terms [$u_1=|a_1|^2-|a_2|^2$, $u_2=-(a_1a_2+a_1^*a_2^*)$, $u_3=i(a_1a_2-a_1^*a_2^*)$, $u_4=2a_1a_2^*$, $u_5=a_1^2-a_2^{*2}$, and $u_6=-i(a_1^2+a_2^{*2})$] and cross-terms [$u_7=a_1^*b_1-a_2b_2^*$, $u_8=-(b_1a_2+b_2^*a_1^*)$, $u_9=i(b_1a_2-a_1^*b_2^*)$, $u_{10}=-i(a_1^*b_1+a_2b_2^*)$, $u_{11}=a_1b_2^*+b_1a_2^*$, $u_{12}=a_1b_1-a_2^*b_2^*$, $u_{13}=-i(a_1b_1+a_2^*b_2^*)$, and $u_{14}=i(a_1b_2^*-a_2^*b_1)$].

First, we consider self-terms. These are the same as in the theory by Wai and Menyuk, but we analyze them here for completeness. They are grouped as $\{u_1,u_2,u_3\}$, $\{\mathrm{Re}(u_4),\mathrm{Re}(u_5),\mathrm{Re}(u_6)\}$, and $\{\mathrm{Im}(u_4),\mathrm{Im}(u_5),\mathrm{Im}(u_6)\}$, and each group obeys the same equations of motion as $\{S_1,S_2,S_3\}$:

$$\begin{pmatrix}S_1\\S_2\\S_3\end{pmatrix}=\begin{pmatrix}S_2\\-S_1\\0\end{pmatrix}g_\theta+\begin{pmatrix}0\\\mp2\Delta\beta(\omega_p)S_3\\\pm2\Delta\beta(\omega_p)S_2\end{pmatrix}. \tag{13}$$

We need to know the average values of quadratic forms composed of these coefficients. They can be found from solutions to the equations of motion for the average of the generic function $\psi(S_1,S_2,S_3,\theta)$:$\partial_z\langle\psi\rangle=\langle G(\psi)\rangle$, where generator $G$ has the form

$$G=\frac{1}{2L_c}\left[S_2^2\frac{\partial^2}{\partial S_1^2}+S_1^2\frac{\partial^2}{\partial S_2^2}+\frac{\partial^2}{\partial\theta^2}-2S_1S_2\frac{\partial^2}{\partial S_1}\partial S_2+2S_2\frac{\partial^2}{\partial\theta\partial S_1}-2S_1\frac{\partial^2}{\partial\theta\partial S_2}\right]\mp2\Delta\beta(\omega_p)S_3\frac{\partial}{\partial S_2}\pm2\Delta\beta(\omega_p)S_2\frac{\partial}{\partial S_3}. \tag{14}$$

In this way, we generate a closed system of equations:

$$\partial_z\langle S_1^2\rangle=-2L_c^{-1}(\langle S_1^2\rangle-\langle S_2^2\rangle), \tag{15}$$

$$\partial_z\langle S_2^2\rangle=2L_c^{-1}(\langle S_1^2\rangle-\langle S_2^2\rangle)\mp4\Delta\beta(\omega_p)\langle S_2S_3\rangle, \tag{16}$$

$$\partial_z\langle S_3^2\rangle=\pm4\Delta\beta(\omega_p)\langle S_2S_3\rangle, \tag{17}$$

$$\partial_z\langle S_2S_3\rangle=-L_c^{-1}\langle S_2S_3\rangle\pm2\Delta\beta(\omega_p)(\langle S_2^2\rangle-\langle S_3^2\rangle). \tag{18}$$

Equations (15)–(18) are to be supplemented with the initial conditions $\{u_1,u_2,u_3\}=\{1,0,0\}$, $\{u_4,u_5,u_6\}=\{0,1,-i\}$.

It can be shown that coefficients $u_3$ and $u_6$ are not correlated, so that the average of their product is the product of their averages, which is found to be exactly zero (see [20] for details). However, separately, $u_3^2$ and $u_6^{*2}$ are not zero and are $z$-dependent. The characteristic length of their evolution toward their corresponding steady states does not exceed 100 m for the range $1<L_B<100\,(\mathrm{m})$ and $0.1<L_c<100\,(\mathrm{m})$, which is typical for telecommunications fibers. Thus, we can safely assume that $u_3^2$ and $\mathrm{Re}(u_6^{*2})$ take their steady-state values: 1/3 and 0, respectively. In this way, we find the SPM tensor in the form $J_s=\gamma_{ss}\mathrm{diag}(0,0,0)$.

Note that on the way to the result we replaced the products of coefficients in Eqs. (8)–(11) by their average values. This approximation works well only when the correlation length $L_c$ is much shorter than $\min(L_{\mathrm{NL}},L)$. Another approximation that we used was the application of the ergodic theorem $\langle g\rangle=\lim_{z\to\infty}z^{-1}\int_0^z \mathrm{d}z' g(z')$. This theorem allowed us to replace the spatial average of a function $g(z)$ by an ensemble average, again assuming that the fiber length is much longer than the correlation length.

Cross-terms are analyzed along similar lines. Groups of coefficients $\{\mathrm{Re}(u_7),\mathrm{Re}(u_8),\mathrm{Re}(u_9),\mathrm{Re}(u_{10})\}$, $\{\mathrm{Im}(u_7),\mathrm{Im}(u_8),\mathrm{Im}(u_9),\mathrm{Im}(u_{10})\}$, $\{\mathrm{Re}(u_{11}),\mathrm{Re}(u_{12}),\mathrm{Re}(u_{13}),\mathrm{Re}(u_{14})\}$, and $\{\mathrm{Im}(u_{11}),\mathrm{Im}(u_{12}),\mathrm{Im}(u_{13}),\mathrm{Im}(u_{14})\}$ obey the same equations of motion as $\{S_1,S_2,S_3,S_4\}$:

$$\begin{pmatrix}S_1\\S_2\\S_3\\S_4\end{pmatrix}=\begin{pmatrix}S_2\\-S_1\\0\\0\end{pmatrix}g_\theta+\begin{pmatrix}\Delta_\pm^{(-)}S_4\\-\Delta_\pm^{(+)}S_3\\\Delta_\pm^{(+)}S_2\\-\Delta_\pm^{(-)}S_1\end{pmatrix},\tag{19}$$

where $\Delta_\pm^{(-)}\equiv[\pm\Delta\beta(\omega_p)-\Delta\beta(\omega_s)]$ and $\Delta_\pm^{(+)}\equiv[\pm\Delta\beta(\omega_p)+\Delta\beta(\omega_s)]$ In order to evaluate the average values of the products of these coefficients, we apply the generic equation $\partial_z\langle\psi\rangle=\langle G(\psi)\rangle$, which is valid for an arbitrary function $\psi(S_1,S_2,S_3,S_4,\theta)$. Here, the generator $G(\psi)$ is given by

$$\begin{aligned}G=\frac{1}{L_c}\Bigg[&S_2^2\frac{\partial^2}{\partial S_1^2}+S_1^2\frac{\partial^2}{\partial S_2^2}+\frac{\partial^2}{\partial\theta^2}-2S_1S_2\frac{\partial^2}{\partial S_1\partial S_2}+2S_2\frac{\partial^2}{\partial\theta\partial S_1}\\&-2S_1\frac{\partial^2}{\partial\theta\partial S_2}-S_1\frac{\partial}{\partial S_1}-S_2\frac{\partial}{\partial S_2}\Bigg]-\Delta_\pm^{(-)}S_1\frac{\partial}{\partial S_4}+\Delta_\pm^{(-)}S_4\frac{\partial}{\partial S_1}\\&-\Delta_\pm^{(+)}S_3\frac{\partial}{\partial S_2}+\Delta_\pm^{(+)}S_2\frac{\partial}{\partial S_3},\end{aligned}\tag{20}$$

and in this way we generate the equations of motion

$$\partial_z\langle S_1^2\rangle=-2L_c^{-1}(\langle S_1^2\rangle-\langle S_2^2\rangle)+2\Delta_\pm^{(-)}\langle S_1S_4\rangle,\tag{21}$$

$$\partial_z\langle S_2^2\rangle=2L_c^{-1}(\langle S_1^2\rangle-\langle S_2^2\rangle)-2\Delta_\pm^{(+)}\langle S_2S_3\rangle,\tag{22}$$

$$\partial_z\langle S_3^2\rangle=2\Delta_\pm^{(+)}\langle S_2S_3\rangle,\tag{23}$$

$$\partial_z\langle S_4^2\rangle=-2\Delta_\pm^{(-)}\langle S_1S_4\rangle,\tag{24}$$

$$\partial_z\langle S_2S_3\rangle=-L_c^{-1}\langle S_2S_3\rangle+\Delta_\pm^{(+)}(\langle S_2^2\rangle-\langle S_3^2\rangle),\tag{25}$$

$$\partial_z\langle S_1S_4\rangle=-L_c^{-1}\langle S_1S_4\rangle+\Delta_\pm^{(-)}(\langle S_4^2\rangle-\langle S_1^2\rangle).\tag{26}$$

The initial conditions are $\{u_7,u_8,u_9,u_{10}\}=\{1,0,0,-i\}$, $\{u_{11},u_{12},u_{13},u_{14}\}=\{0,1,-i,0\}$. The average values of cross-products like $u_iu_j$ vanish and the structure of the XPM term $\hat{N}_{xpm}$ is greatly simplified. When Eq. (8) and the similar equation for the signal beam are written as evolution equations for Stokes parameters (1) and (2), the XPM tensor takes the simple form $J_x=\frac{1}{3}\gamma\mathrm{diag}(\langle\mathrm{Re}(u_{10}^2+u_{13}^2-u_9^2-u_{14}^2)\rangle,\langle\mathrm{Re}(u_{10}^2+u_{14}^2-u_9^2-u_{13}^2)\rangle,\langle|u_9|^2+|u_{14}|^2-|u_{13}|^2-|u_{10}|^2\rangle)$.

We numerically computed the coefficients of interest and, for the case when $[\Delta\beta(\omega_p)-\Delta\beta(\omega_s)]=0$, we observed a transient evolution of the elements of the XPM tensor toward their steady-state values. For the range of lengths $1<L_B<100\,(\mathrm{m})$, $0.1<L_c<100\,(\mathrm{m})$ the transient length did not exceed 1 km. Neglecting this transient stage we assumed steady-state values in the XPM tensor. For the counterpropagating case, they are $J_x=\gamma_{ps}\frac{8}{9}\mathrm{diag}(-1,1,-1)$. For comparison, for the copropagating case, the XPM matrix is $J_x=\frac{8}{9}\gamma_{ps}\mathrm{diag}(-1,-1,-1)$. Note that, considering steady state, $z$-independent XPM tensor components are not necessary when numerically solving the Stokes parameters (1) and (2) as we do in Sections 4 and 5. Therefore, our theory involving the spatially varying XPM tensor is applicable to more general situations involving relatively short fibers, where the transient terms may play a role. However, we verified that the contribution of these terms is negligible in the situations of interest in this work. Moreover, the simple form of the XPM tensor is significant since it may lead to analytical solutions. With this, the modeling of the randomly birefringent fiber is completed.

Note that throughout the paper components of the Stokes vectors are defined as $S_1^\pm=\Phi_{p1}^{\pm*}\Phi_{p2}^\pm+\Phi_{p1}^\pm\Phi_{p2}^{\pm*}$, $S_2^\pm=i(\Phi_{p1}^{\pm*}\Phi_{p2}^\pm+\Phi_{p1}^\pm\Phi_{p2}^{\pm*})$, $S_3^\pm=|\Phi_{p1}|^2-|\Phi_{p2}|^2$.

## 4. GENERAL PROPERTIES OF THE POLARIZATION ATTRACTION

In summary, our model describes the situation when two counterpropagating light beams nonlinearly interact inside a randomly birefringent fiber via the Kerr effect. The temporal variations of the beams are supposed to be relatively slow, so that the second- and higher-order chromatic dispersions can be neglected over the entire length $L$ of the fiber.

For each beam we define the zeroth Stokes parameters $S_0^+$ and $S_0^-$ according to the equation $S_0^\pm=\sqrt{(S_1^\pm)^2+(S_2^\pm)^2+(S_3^\pm)^2}$. These parameters represent the powers of forward and backward beams. As we are dealing with a conservative problem, the sum of the powers of both beams is a conserved quantity. Moreover, the equations of motion (1) and (2) imply that the power of each beam is separately a conserved quantity: $S_0^+(z-ct)=S_0^+(z=0,t)$ and $S_0^-(z+ct)=S_0^-(z=L,t)$ for all $z$. Note that for stationary (i.e., independent of time) boundary conditions and uniform (i.e., independent of $z$) initial conditions, which are the only conditions of interest to us here, the powers of the beams do not depend either on $t$ or $z$, so we

skip the explicit indication of the dependence of $S_0^{\pm}$ on $t$ and $z$. Throughout this work, the initial and boundary conditions are connected in the following manner:

$$S_i^+(z=0,t)=S_i^+(z,t=0), \tag{27}$$

$$S_i^-(z=L,t)=S_i^-(z,t=0), \tag{28}$$

with $i=1,2,3$. From now on we only specify the boundary conditions, while for the initial conditions, we refer to the previous relationships (27) and (28). As we shall see, our numerical simulations demonstrate the convergence of the distributions of the Stokes vectors $\mathbf{S}^+$ and $\mathbf{S}^-$ inside the medium to a stationary polarization attractor, whose particular shape only depends on the boundary conditions. In other words, the initial conditions have no influence whatsoever on the shape of the final attractor. Nevertheless, initial conditions do affect the time of convergence (which we denote as $T_{tr}$) to the stationary polarization attractor. Note that the polarization attractor is not necessarily stationary in time.

Let us also introduce the nonlinear length $L_{\mathrm{NL}}=(\gamma S_0^+)^{-1}$, which has the meaning of the characteristic length of nonlinear beam evolution inside the fiber. In its turn, the characteristic time is simply $T_{\mathrm{NL}}=L_{\mathrm{NL}}/c$. In our simulations, we use $L_{\mathrm{NL}}$ as the unit for measuring distance in the fiber medium in a reference frame with the origin ($z=0$) at the left boundary, where we set the boundary conditions for the forward beam. At the right boundary ($z=L$), we set the boundary conditions for the backward beam. Temporal units have no explicit use in representations of our results because we are interested in the long-term solutions for the Stokes parameters, i.e., polarization attractors. The duration $T_{tr}$ of the transient processes depends on the length of the medium, the particular choice of the boundary conditions, and the initial conditions, as discussed earlier. We found that $T_{tr}<10000T_{\mathrm{NL}}$ for a rather general choice of parameters. Therefore, all plots presented in this work were generated for integration times $T=10000T_{\mathrm{NL}}$.

First, let us comment on our choice of the fiber medium length. Clearly, to allow the beams to interact in full strength, this length must exceed $L_{\mathrm{NL}}$. In all simulations, we used a fiber sample of total length $L=10L_{\mathrm{NL}}$.

Polarization attraction is an effect that can be characterized as follows: a random (or uniform) distribution of initial states of polarization of the signal converges upon propagation through the medium into a different, very tightly localized statistical distribution of polarizations, with its own mean and root mean square deviation. The characterization of such transformation between input and output polarization distributions is one of the main goals of this study. From this description, it is clear that we are interested in studying the statistics of scrambled beams, thus dealing with their ensemble average. A different approach is to consider a single input signal beam with its SOP randomly varying in time, and then compute the statistics of the output signal SOP on the basis of time averaging. In a forthcoming publication, we shall demonstrate that the lossless polarizer is an ergodic system (i.e., the ensemble and time averages coincide) when the input signal SOP time fluctuations are slower than the characteristic response time $T_{\mathrm{NL}}$ of the polarizer. The existence of such an equivalence permits us to anticipate that the results reported here indeed represent the universal statistical steady-state properties of the polarizer.

Currently, no analytical methods are available to characterize the polarization attraction dynamics of the system of equations (1) and (2) coupled with their initial and boundary conditions [Eqs. (27) and (28)]. Therefore, we may only rely on their numerical analysis and the method of trial and error. For instance, not necessarily all polarizations of the pump beam produce a desirable polarization attraction effect. As we show, some pump polarizations perform better than others.

From the practical viewpoint, it is already an obvious success if one finds at least one SOP of the pump beam that serves as an effective lossless polarizer. Then, methods to convert an output polarized signal beam into any other desirable SOP are known. Driven by these practical considerations, we see our goal in finding and characterizing the polarization attraction effect for a limited set of SOPs of the pump beam.

## 5. RESULTS

We performed the integration of Eqs. (1) and (2) based on the numerical method proposed in [23]. We chose to work with a set of boundary conditions for the SOP of the pump beam that are represented on the Poincaré sphere by points in the vicinity of its six poles, namely $(\pm1,0,0)$, $(0,\pm1,0)$, and $(0,0,\pm1)$. For each value of the pump polarization at $z=L$, we applied at $z=0$ an ensemble of $N=110$ signal beam SOPs with uniform distribution over the Poincaré sphere and followed their evolution into a distribution of output SOPs at $z=L$. From

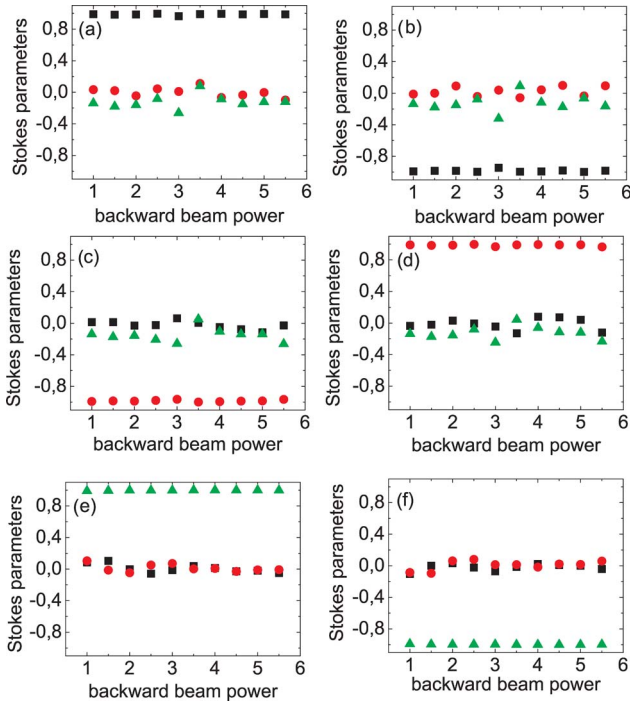


Fig. 1. (Color online) Components of the (normalized) mean Stokes vector of the output signal beam as a function of the relative power of the pump beam: $S_1^+$ (black squares), $S_2^+$ (red circles), and $S_3^+$ (green triangles) for six input SOPs of the pump beam located near the poles of the Poincaré sphere: (a) (–0.99, 0.01, 0.14), (b) (0.99,0.01,0.14), (c) (0.01, –0.99, 0.14), (d) (0.01,0.99,0.14), (e) (0.01, 0.01, –0.9999), (f) (0.01,0.01,0.9999). The Stokes parameters of the signal and pump beam are normalized with respect to $S_0^+(z,t)$ and $S_0^-(z,t)$, respectively.

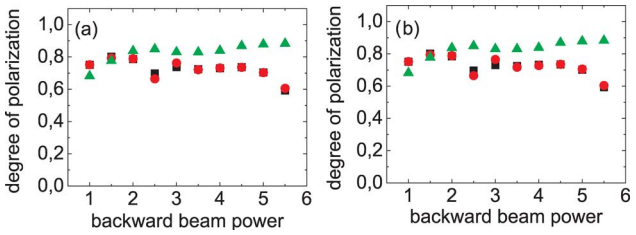


Fig. 2. (Color online) DOP $D$ of the output signal beam as a function of the relative pump beam power for six input SOPs of the pump beam: (a) $(-0.99, 0.01, 0.14)$ (black squares), $(0.01, -0.99, 0.14)$ (red circles), $(0.01, 0.01, -0.9999)$) (green triangles); (b) (0.99,0.01,0.14) (black squares), (0.01,0.99,0.14) (red circles), (0.01,0.01,0.9999) (green triangles).

these $N = 110$ output signal polarizations we deduced the direction of the mean Stokes vector and characterized the spread of the corresponding distribution by means of the degree of polarization (DOP) $D$. The details of this statistical analysis can be found in Appendix A. The DOP may span values ranging from 0 (totally depolarized light) to 1 (perfectly polarized light). Our goal is to achieve an output $D$ as close to unity as possible. Numerical runs were performed for pump beam powers in the range $1–5.5S_0^+$.

Figures 1 and 2 show the main properties of lossless polarization attraction in randomly birefringent fibers. The coordinates of the mean Stokes vector of the output signal beam distribution as a function of the pump beam power are shown in Fig. 1. These results show two important features: (1) the output SOP of the signal is almost independent of the relative pump power and (2) the signal SOP changes when the SOP of the pump beam is varied. Yet even more important are the results that are provided by examining the variations of the output signal DOP, shown in Fig. 2. As can be seen, $D$ changes in a nonmonotonic fashion as the relative pump power grows larger, reaching a maximum value of $D \simeq 0.9$ whenever the SOP of the pump is aligned along the $S_3$ axis. Note that whenever the pump power drops below $S_0^+$ (not shown), the signal DOP quickly degrades, reaching zero for vanishing pump power.

Figure 3 illustrates the width and randomness of the spread of the output signal Stokes vector distribution on the Poincaré

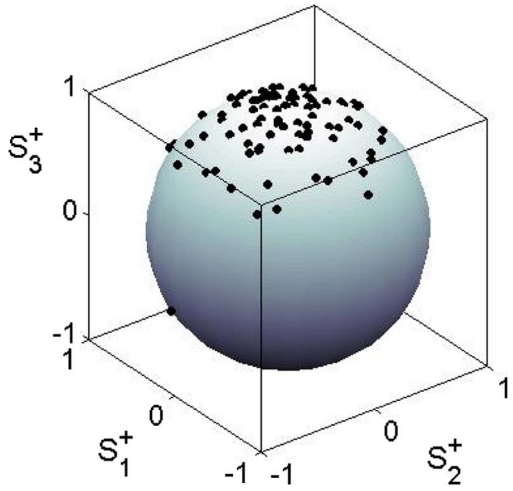


Fig. 3. (Color online) $N = 110$ points on the Poincaré sphere illustrating the output SOPs of the signal beam corresponding to an input distribution of $N = 110$ points uniformly distributed over the Poincaré sphere. Here, the input SOP of the pump beam is $(0.01, 0.01, -0.9999)$. The pump power is $S_0^- = 5.5S_0^+$.

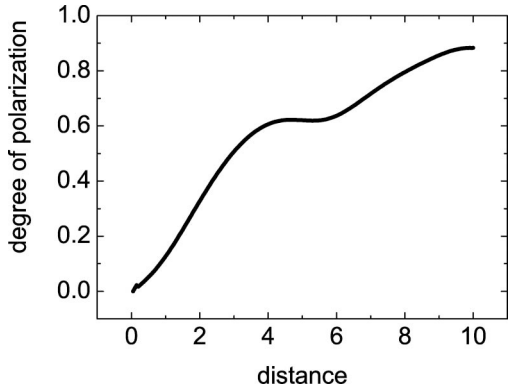


Fig. 4. DOP of the signal as the function of the distance propagated inside the fiber. The initial SOP of the pump is (0.01,0.01,0.9999).

spheres. To plot Fig. 3, we used the data corresponding to the case exhibiting the strongest attraction. It is difficult to say whether the highest recorded value of $D \simeq 0.9$ is large or small because the tolerance to the residual polarization uncertainty sensitively depends on the particular application. The plot of Fig. 4 illustrates the spatial dynamics of the DOP $D$ as the signal beam traverses through the fiber medium. This figure demonstrates the progressive repolarization of the signal toward a well-defined SOP.

On the other hand, the rule of how the output SOP of the signal depends on the input SOP of the pump can be readily deduced from Fig. 1. In all cases shown by this figure, it can be seen that the signal is attracted toward virtually the same or orthogonal SOP of the pump. The fact that this rule is very simple is important in practical applications.

So far we considered the SOPs of the pump that are located near each of the six poles of the Poincaré sphere. We also performed simulations for the additional four pump polarization states: $(1/\sqrt{3})(1, -1, 1)$, $(1/\sqrt{3})(-1, -1, 1)$, $(1/\sqrt{3})(-1, 1, 1)$, and $(1/\sqrt{3})(1, 1, -1)$. We found that polarization attraction is still present and relatively strong in these cases; nevertheless, the DOP did not exceed the value of $D \simeq 0.83$. Thus, our results indicate that polarization attraction appears to be a general property of the cross-polarization interaction among counterpropagating waves in a randomly birefringent fiber. However, the degree of signal repolarization that can be achieved in a given fiber span is a sensitive function of the pump beam power and its SOP.

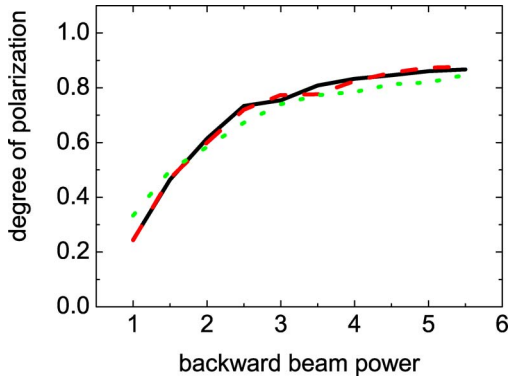


Fig. 5. (Color online) DOP of the pump at $z = 0$ as a function of the relative pump power. The initial SOP of the pump is $(-0.99, 0.01, 0.14)$ (black solid curve), $(0.01, -0.99, 0.14)$ (red dashed curve), (0.01,0.01, 0.9999) (green dotted curve).

We also recorded the SOP of the pump beam after it traversed through the medium and interacted with the signal beam. Figure 5 shows the DOP of the pump at its output end as a function of the relative pump power and for different pump polarizations at $z = L$. As can be seen, for relatively low powers, the pump beam becomes nearly depolarized at its output end $z = 0$. This indicates the presence of a sort of conservation of order in our system—the signal beam becomes polarized at the expense of the depolarization of the pump beam. The results of Fig. 5 show that, in practice, a nonlinear randomly birefringent fiber may also be used as an efficient polarization scrambler. Indeed, scrambling of the output pump beam in isotropic fibers was theoretically and experimentally demonstrated in [15].

It is important to point out that the cross-polarization modulation tensor $J_x$ acquires the first two nonvanishing terms on the diagonal, which are responsible for the polarization attraction effect, only under the assumption that the frequency difference $\Delta\omega = \omega_p - \omega_s$ is small enough to fulfill $L'_B \ll L$. This inequality can be written as a condition for the difference between signal and pump wavelengths $\lambda_s - \lambda_p \ll [2\pi/\Delta\beta(\lambda_p)[(\lambda_p/L)$. For $L = 10\,\mathrm{km}$, $\lambda_p = 1.55\,\mu\mathrm{m}$, and $\Delta\beta(\lambda_p) = 1\,\mathrm{m}^{-1}$, we get the estimate $\lambda_s - \lambda_p \ll 1\,\mathrm{nm}$. For low-birefringent fibers, the magnitude $\Delta\beta(\lambda_p)$ can be up to two orders of magnitude smaller and, therefore, pump and signal wavelengths can be substantially farther apart (although not too far, otherwise the Raman effect may harm the performance of the polarizer).

## 6. CONCLUSIONS

In previous work, the effect of polarization attraction was thoroughly studied only for the case of isotropic fibers, for which the required powers of the beams were orders of magnitude higher than the typical power levels in telecommunications-related devices. Therefore, the necessity to move toward more practical implementations of this effect became evident. With the goal of developing practical lossless polarizers for telecom applications, we demonstrated the effect of polarization attraction in randomly birefringent fibers. As a matter of fact, an initial bold step forward in using randomly birefringent fibers for polarization attraction was experimentally carried out in [16]. Our theory can be viewed as the theoretical support for these experimental results. It opens the way to study the effect of polarization attraction in more detail, and it may help to predict some new interesting nonlinear dynamics. Clearly, the formulation of the equations of motion for such a practically relevant problem sets the basis for its subsequent formal study. Indeed, an interesting analytical insight in the polarization attraction phenomenon has been recently provided in [10–12] for the case of isotropic fibers.

The model based on the system of equations (1) and (2) may also form the basis for an adequate description of polarization attraction in other nonlinear materials that so far have not been considered for this purpose. The only requirement is finding the specific self- and cross-polarization tensors that are relevant for the Kerr-like medium of interest.

## APPENDIX A: CALCULATION OF THE STATISTICS OF THE STOKES VECTORS

As explained earlier, $N = 110$ signal beams with different SOPs uniformly distributed over the Poincaré sphere are launched into the fiber, one at a time. For all of these $N = 110$ realizations, the pump beam with one and the same SOP was launched from the opposite end of the fiber. For each realization, we measure $S_1^+(L)$, $S_2^+(L)$, and $S_3^+(L)$. At the end of the simulations we calculate the mean values:

$$\langle S_i^+(L)\rangle = \frac{1}{N}\sum_{j=1}^{N}[S_i^+(L)]_j, \tag{A1}$$

where $N = 110$ and $i = 1, 2, 3$. In this form, these averages do not define the mean direction of the Stokes vector on the Poincaré sphere simply because the sum of squares of these means does not yield the square of the power $(S_0^+)^2$. The length of the vector satisfies the inequality $\sqrt{\langle S_1^+(L)\rangle + \langle S_2^+(L)\rangle + \langle S_3^+(L)\rangle} \le S_0^+$ and can be even zero.

However, the information contained in the three means is sufficient to restore the direction of the Stokes vector and, moreover, to quantify the DOP of the outcoming signal light. The two angles $\theta_0$ and $\phi_0$ (also called circular means), which determine the direction of the Stokes vector on the Poincaré sphere, are defined as

$$\theta_0 = \arccos\left(\frac{\langle S_3\rangle}{\sqrt{\langle S_1^+\rangle^2 + \langle S_2^+\rangle^2 + \langle S_3^+\rangle^2}}\right), \tag{A2}$$

$$\phi_0 = \arctan 2(\langle S_2^+\rangle, \langle S_1^+\rangle). \tag{A3}$$

Here,

$$\mathrm{arc\,tan}\,2(x,y) = \begin{cases} \arctan(y/x), & x > 0, \\ \pi + \arctan(y/x), & y \ge 0, x < 0 \\ -\pi + \arctan(y/x), & y < 0, x < 0. \end{cases} \tag{A4}$$

Finally, the Cartesian coordinates of the Stokes vector are restored as

$$\bar{S}_1 = S_0^+ \sin\theta_0 \cos\phi_0, \tag{A5}$$

$$\bar{S}_2 = S_0^+ \sin\theta_0 \sin\phi_0, \tag{A6}$$

$$\bar{S}_3 = S_0^+ \cos\theta_0. \tag{A7}$$

It is these coordinates that are displayed in Fig. 1.

The DOP is defined as $D = (\langle S_1^+\rangle^2 + \langle S_2^+\rangle^2 + \langle S_3^+\rangle^2)^{1/2}/S_0^+$. It is this DOP that is displayed in Fig. 2.

## ACKNOWLEDGMENTS

We thank J. Fatome for sending us a preprint of [16] prior to its publication. This work was carried out with support from the Italian Ministry of the University and Research (MIUR) through the PRIN 2008 project "Nonlinear cross-polarization interactions in photonic devices and systems" and the 2009 Italy–Spain bilateral initiative "Nonlinear Optical Systems and Devices" (HI2008-0075). JN acknowledges funding from the Spanish Ministry of Science and Innovation through an FPI fellowship (TEC2008-05791).

## REFERENCES


1. V. E. Zakharov and A. V. Mikhailov, "Polarization domains in nonlinear optics," Pis'ma Zh. Eksp. Teor. Fiz., **45**, 279–282 (1987) [JETP Lett. **45**, 349–352 (1987)].
2. S. Pitois, G. Millot, and S. Wabnitz, "Polarization domain wall solitons with counterpropagating laser beams," Phys. Rev. Lett. **81**, 1409–1412 (1998).
3. S. Wabnitz, "Chiral polarization solitons in elliptically birefringent spun optical fibers," Opt. Lett. **34**, 908–910 (2009).
4. S. Wabnitz, "Cross polarization modulation domain wall solitons for WDM signals in birefringent optical fibers," IEEE Photonics Technol. Lett. **21**, 875–877 (2009).
5. A. V. Mikhailov and S. Wabnitz, "Polarization dynamics of counterpropagating beams in optical fibers," Opt. Lett. **15**, 1055–1057 (1990).
6. S. Wabnitz and B. Daino, "Polarization domains and instabilities in nonlinear optical fibers," Phys. Lett. A **182**, 289–293 (1993).
7. A. Degasperis, S. V. Manakov, and P. M. Santini, "On the initial-boundary value problems for soliton equations," JETP Lett. **74**, 481–485 (2001).
8. V. V. Kozlov and S. Wabnitz, "Instability of optical solitons in the boundary value problem for a medium of finite extension," Lett. Math. Phys., in press (2010).
9. 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).
10. 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).
11. 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 (2010).
12. 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).
13. 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).
14. 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).
15. 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).
16. 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).
17. M. Martinelli, M. Cirigliano, M. Ferrario, L. Marazzi, and P. Martelli, "Evidence of Raman-induced polarization pulling," Opt. Express **17**, 947–955 (2009).
18. A. Zadok, E. Zilka, A. Eyal, L. Thèvenaz, and M. Tur, "Vector analysis of stimulated Brillouin scattering amplification in standard single-mode fibers," Opt. Express **16**, 21692–21707 (2008).
19. J. Fatome, S. Pitois, and G. Millot, "Experimental evidence of Brillouin-induced polarization wheeling in highly birefringent optical fibers," Opt. Express **17**, 12612–12618 (2009).
20. P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," J. Lightwave Technol. **14**, 148 (1996).
21. V. V. Kozlov and S. Wabnitz, "Theoretical study of polarization attraction in high-birefringence and spun fibers," Opt. Lett., **35**, 3949–3951 (2010).
22. C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," J. Lightwave Technol. **24**, 2806–2826 (2006).
23. C. Martijn de Sterke, K. R. Jackson, and B. D. Robert, "Nonlinear coupled-mode equations on a finite interval: a numerical procedure," J. Opt. Soc. Am. B **8**, 403–412 (1991).