

Effective pulse dynamics in optical fibers with polarization mode dispersion

Josselin Garnier ^{a,*}, Renaud Marty ^b

^a *Laboratoire de Probabilités et Modèles Aléatoires and Laboratoire Jacques-Louis Lions, Université Paris VII, 2 Place Jussieu, 75251 Paris Cedex 05, France*

^b *Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 4, France*

Received 19 April 2006; accepted 3 May 2006
Available online 16 June 2006

Abstract

This paper investigates pulse propagation in randomly birefringent optical fibers. We consider two polarization-mode dispersion models. Using a separation of scales technique we derive an effective stochastic partial differential equation for the envelope of the field. This equation is driven by three independent Brownian motions, and it depends on the polarization-mode dispersion model through a single effective parameter. This shows that pulse dynamics in randomly birefringent fibers does not depend on the microscopic model. Numerical simulations are in excellent agreement with the theoretical predictions.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Optical fiber; Polarization mode dispersion; Waves in random media

1. Introduction

Optical fibers have been extensively studied because they play an important role in modern communication systems [1]. In particular, polarization mode dispersion (PMD) has attracted the attention of engineers, physicists and, more recently, applied mathematicians. PMD is one of the main limiting effects of high bit rate transmission in optical fiber links. It has its origin in the birefringence [2,3], i.e. the fact that the electric field is a vector field and the index of refraction of the medium depends on the polarization state (i.e. the unit vector pointing in the direction of the electric vector field). For a fixed position in the fiber, there are two orthogonal polarization states which correspond to the maximum and the minimum propagation velocities. These two polarization states are parameterized by an angle with respect to a fixed pair of axes that is called the birefringence angle. The difference between the maximum and the minimum of the index of refraction is the so-called birefringence strength. If the birefringence angle and strength were constant along the fiber, then a pulse polarized along one of the eigenstates would travel at constant velocity. However the birefringence angle is randomly varying which involves coupling between the two polarized modes. The modes travel with different

* Corresponding author. Tel.: +33 1 44 27 86 93; fax: +33 1 44 27 76 50.

E-mail addresses: garnier@math.jussieu.fr (J. Garnier), marty@cict.fr (R. Marty).

velocities, which involves pulse spreading. Random birefringence results from variations of the fiber parameters such as the core radius or geometry. There exist various physical reasons for the fluctuations of the fiber parameters. They may be generated during the drawing process as a result of a drift of the operating parameters. They may also be induced by mechanical distortions on fibers in practical use, such as point-like pressures or twists [4].

In our paper we consider the linear propagation of a pulse in a randomly birefringent optical fiber. The electric field is solution of two coupled Schrödinger equations with random coefficients. Randomness is inherent in PMD. The length scales present in the problem are the fiber length, the beat length (proportional to the inverse of the birefringence strength), and the PMD correlation length. We consider realistic configurations where the fiber length (typically 10^3 km) is larger than the PMD correlation length (typically a few tens of meters) which is itself larger than the beat length (a meter or less) [5]. We study the asymptotic behavior of the field envelope in the framework based on the separation of these scales, using a technique first introduced by Papanicolaou and coworkers [6]. We derive an asymptotic stochastic partial differential equation driven by three independent Brownian motions and parameterized by a single effective PMD parameter. This result can be extended to more general configurations, where other length scales come into the play: the nonlinear length and the length of the dispersion map in the presence of a dispersion-management scheme. The result obtained in this paper should remain true as long as the PMD correlation length is smaller than these new length scales, which holds true in practical configurations. We shall give numerical evidence of this conjecture.

The paper is organized as follows. Section 2 is devoted to the presentation of the first model introduced by Wai and Menyuk [2,3] for linear randomly birefringent optical fibers. In Section 3 we introduce the second model introduced by Wai and Menyuk [2,3]. Section 4 is devoted to the study of the asymptotic dynamics of the electric field in the two PMD models. In Section 5 we present numerical results which are compared with the theoretical asymptotic results.

2. Random birefringence with constant birefringence strength

The evolution of the two polarization modes of the electric field in a randomly birefringent fiber is governed by the coupled Schrödinger equation [2]:

$$i \frac{\partial E}{\partial z} + bK(z)E + ib'K(z) \frac{\partial E}{\partial t} + \frac{d_0}{2} \frac{\partial^2 E}{\partial t^2} = 0, \tag{1}$$

where $E(z, t)$ is the column vector $(E_1(z, t), E_2(z, t))^T$ representing the field vector envelope depending on the position $z \in [0, +\infty)$ and the local time $t \in \mathbb{R}$. b (resp. b') is the birefringence strength (resp. the frequency-derivative of the birefringence strength). The usual beat length is related to b through the identity $L_B = \pi/b$. We denote by $\theta(z)$ the random birefringence angle which depends on the position z . The matrix K is given by

$$K(z) = \cos \theta(z) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \sin \theta(z) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

d_0 is the group velocity dispersion (GVD) parameter. We perform the change of variables $\tilde{E} = M^{-1}(z)N^{-1}(z)E$ where

$$N(z) = \begin{pmatrix} \cos(\theta(z)/2) & -\sin(\theta(z)/2) \\ \sin(\theta(z)/2) & \cos(\theta(z)/2) \end{pmatrix} \quad \text{and} \quad M(z) = \begin{pmatrix} \tilde{v}_1(z) & \tilde{v}_2(z) \\ -\overline{\tilde{v}_2(z)} & \overline{\tilde{v}_1(z)} \end{pmatrix}$$

is the unitary matrix solution of

$$i \frac{dM}{dz} + \left[\frac{i}{2} \frac{d\theta}{dz}(z) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + b \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] M = 0.$$

Here \bar{x} stands for the conjugate of a complex number x . The equation describing the evolution of the new electric field \tilde{E} is:

$$i \frac{\partial \tilde{E}}{\partial z} + \frac{d_0}{2} \frac{\partial^2 \tilde{E}}{\partial t^2} + ib' \tilde{\Sigma}(z) \frac{\partial \tilde{E}}{\partial t} = 0, \tag{2}$$

where $\tilde{\Sigma} = M^{-1}N^{-1}KNM$. Note that NM is unitary so that, for any z, t , we have $\|E(z, t)\|_{\mathbb{C}^2} = \|\tilde{E}(z, t)\|_{\mathbb{C}^2}$. The matrix $\tilde{\Sigma}(z)$ depends on the position $z \in [0, +\infty)$ in the fiber. It is imposed by PMD, and it has the form

$$\tilde{\Sigma}(z) = \begin{pmatrix} |\tilde{v}_1(z)|^2 - |\tilde{v}_2(z)|^2 & 2\overline{\tilde{v}_1(z)}\tilde{v}_2(z) \\ 2\tilde{v}_1(z)\overline{\tilde{v}_2(z)} & |\tilde{v}_2(z)|^2 - |\tilde{v}_1(z)|^2 \end{pmatrix}.$$

In realistic experimental configurations the typical amplitude of $d\theta/dz$ is much smaller than the birefringence strength b [5]. Accordingly we can consider that the derivative $d\theta/dz$ of the orientation angle can be written as $2\varepsilon\alpha$ where α is a stochastic process and ε ($0 < \varepsilon \ll 1$) is a small parameter which characterizes the amplitudes of the fluctuations of θ . The \mathbb{C}^2 -valued process $\tilde{v} := (\tilde{v}_1, \tilde{v}_2)$ is governed by the random differential equation:

$$\frac{d\tilde{v}_1}{dz} = i b \tilde{v}_1 - \varepsilon \alpha(z) \overline{\tilde{v}_2}, \quad \frac{d\tilde{v}_2}{dz} = i b \tilde{v}_2 + \varepsilon \alpha(z) \overline{\tilde{v}_1}.$$

Note that $|\tilde{v}_1|^2 + |\tilde{v}_2|^2 = 1$. We assume that α is centered, bounded a.s., and has a unique invariant probability measure \mathbb{P} , under which it is ergodic and that its infinitesimal generator \mathcal{A} satisfies the Fredholm alternative, i.e. \mathcal{A} admits an inverse on the subspace of functions centered under \mathbb{P} . The symbol \mathbb{E} denotes the expectation with respect to the invariant probability measure \mathbb{P} . Moreover, we assume that the integrals

$$\gamma_c := \int_0^{+\infty} \cos(2bz) \mathbb{E}[\alpha(0)\alpha(z)] dz \quad \text{and} \quad \gamma_s := \int_0^{+\infty} \sin(2bz) \mathbb{E}[\alpha(0)\alpha(z)] dz$$

are well-defined. Note that γ_c is nonnegative by Wiener–Khintchine theorem. We can now study the birefringence state. First, we apply a diffusion-approximation theorem to the process \tilde{v} . Second, we show that the limit process has good ergodic properties.

We consider that the fiber length is much larger than the correlation length of the fluctuations α . Considering that the correlation length is of order one we look for the birefringence state \tilde{v} after a propagation distance of order ε^{-2} . Accordingly we introduce the re-scaled processes $\tilde{v}^\varepsilon(\cdot) := \tilde{v}(\cdot/\varepsilon^2)$ and $\alpha^\varepsilon(\cdot) := \alpha(\cdot/\varepsilon^2)$. We have

$$\frac{d\tilde{v}_1^\varepsilon}{dz} = \frac{i b}{\varepsilon^2} \tilde{v}_1^\varepsilon - \frac{\alpha^\varepsilon(z)}{\varepsilon} \overline{\tilde{v}_2^\varepsilon}, \quad \frac{d\tilde{v}_2^\varepsilon}{dz} = \frac{i b}{\varepsilon^2} \tilde{v}_2^\varepsilon + \frac{\alpha^\varepsilon(z)}{\varepsilon} \overline{\tilde{v}_1^\varepsilon}.$$

We perform the change of variables:

$$v_1^\varepsilon := \tilde{v}_1^\varepsilon \exp\left(-\frac{i b z}{\varepsilon^2}\right), \quad v_2^\varepsilon := \overline{\tilde{v}_2^\varepsilon} \exp\left(\frac{i b z}{\varepsilon^2}\right), \tag{3}$$

so that we get

$$\frac{dv_1^\varepsilon}{dz} = -\frac{\alpha^\varepsilon(z)}{\varepsilon} \exp\left(-\frac{2i b z}{\varepsilon^2}\right) v_2^\varepsilon, \quad \frac{dv_2^\varepsilon}{dz} = \frac{\alpha^\varepsilon(z)}{\varepsilon} \exp\left(\frac{2i b z}{\varepsilon^2}\right) v_1^\varepsilon. \tag{4}$$

We are now in a position to prove the following proposition:

Proposition 1. (i) *As $\varepsilon \rightarrow 0$, the processes v^ε converge in distribution in $\mathcal{C}([0, +\infty), \mathbb{C}^2)$ to the continuous Markov process v solution of the stochastic differential equation*

$$dv(z) = i\sqrt{\gamma_c}(P_1 v(z) \circ dW_1(z) + P_2 v(z) \circ dW_2(z)) + i\gamma_s P_3 v(z) dz, \tag{5}$$

where W_1 and W_2 are two independent standard Brownian motions, P_1, P_2 , and P_3 are the Pauli matrices

$$P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad P_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and \circ stands for the Stratonovich integral. Throughout this section the infinitesimal generator of v will be denoted by \mathcal{G}_v . (ii) *The process v has a unique invariant probability μ which is the uniform measure on the unit sphere \mathbb{S}^3 of $\mathbb{C}^2 \sim \mathbb{R}^4$. Moreover, if $f : \mathbb{S}^3 \rightarrow \mathbb{R}$ is a \mathcal{C}^2 bounded function such that $\int_{\mathbb{S}^3} f d\mu = 0$, then the Poisson equation $\mathcal{G}_v g + f = 0$ has a \mathcal{C}^2 bounded solution g which can be written as*

$$g(v_0) = \int_0^{+\infty} \mathbb{E}[f(v(z)) | v(0) = v_0] dz.$$

Proof. The proof of (i) is an application of a diffusion-approximation theorem [6]. The proof of (ii) is a consequence of Theorem 1.3.6 and Theorem 1.3.15 [7]. Indeed, the process v has a compact state space, thus it remains to prove that its transition probability admits a continuous positive density with respect to the uniform measure on the unit sphere \mathbb{S}^3 of \mathbb{R}^4 . We remark that the Lie bracket of iP_1 and iP_2 is $[iP_1, iP_2] = P_2 P_1 - P_1 P_2 = 2iP_3$. So the Lie algebra generated by iP_1 and iP_2 spans the tangent space to \mathbb{S}^3 at every point. According to the probabilistic Hörmander’s theorem (see [8] volume 2, p. 251) on \mathbb{S}^3 the transition probability of v admits a positive density with respect the uniform measure on \mathbb{S}^3 . Thus, v has a unique invariant probability, and we can readily check that this probability is the uniform measure μ on \mathbb{S}^3 . \square

By Proposition 1, the birefringence model that has been introduced can be simplified by substituting

$$\Sigma(z) := \begin{pmatrix} |v_1(z)|^2 - |v_2(z)|^2 & 2\overline{v_1(z)}v_2(z) \\ 2v_1(z)v_2(z) & |v_2(z)|^2 - |v_1(z)|^2 \end{pmatrix}$$

for the matrix $\tilde{\Sigma}(z)$ in Eq. (2). Thus, we now consider the new coupled Schrödinger equation for the polarized fields in a randomly birefringent fiber:

$$i \frac{\partial \tilde{E}}{\partial z} + \frac{d_0}{2} \frac{\partial^2 \tilde{E}}{\partial t^2} + ib' \Sigma(z) \frac{\partial \tilde{E}}{\partial t} = 0. \tag{6}$$

We define the real-valued random processes η_1, η_2 , and η_3 by

$$\eta_1 + i\eta_2 = 2v_1v_2 \quad \text{and} \quad \eta_3 = |v_1|^2 - |v_2|^2, \tag{7}$$

and η stands for (η_1, η_2, η_3) . Then, Eq. (6) can be written as

$$i \frac{\partial \tilde{E}}{\partial z} + \frac{d_0}{2} \frac{\partial^2 \tilde{E}}{\partial t^2} + ib'(\eta_1(z)P_1 + \eta_2(z)P_2 + \eta_3(z)P_3) \frac{\partial \tilde{E}}{\partial t} = 0. \tag{8}$$

From the definition (7) we write $\eta_1^2 + \eta_2^2 + \eta_3^2 = (|v_1|^2 + |v_2|^2)^2 = 1$ which shows that the process η takes values on the unit sphere \mathbb{S}^2 of \mathbb{R}^3 . We give some other properties of η in the next proposition. We denote by \mathbb{E}_μ the expectation with respect to the process v starting from the invariant distribution μ .

Proposition 2. (i) The process η is a function of v , which is a bounded, ergodic, Markov process with invariant probability measure μ . (ii) For $j = 1, 2, 3$ we have $\mathbb{E}_\mu[\eta_j(z)] = 0$ and $\mathbb{E}_\mu[\eta_j^2(z)] = 1/3$ for any $z \geq 0$. (iii) Denoting $L_c = 1/(4\gamma_c)$, we have, for $j, k = 1, 2, 3$,

$$\int_0^{+\infty} \mathbb{E}_\mu[\eta_j(0)\eta_k(z)] dz = \begin{cases} \frac{L_c}{3} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Note that L_c is the correlation length of the process η , which is the usual PMD correlation length. The hypothesis saying that the derivative of the birefringence angle with respect to distance is smaller than the birefringence strength is thus equivalent to saying that the PMD correlation length is larger than the beat length. This corresponds to practical configurations where the beat length is typically one meter or even less while the PMD correlation length is 10–100 m (see Section 5 for numerical applications).

Proof. (i) The first point follows immediately from Proposition 1 and Eq. (7). (ii) We check that $\mathbb{E}_\mu[\eta_j(0)] = 0$ and we have $\mathbb{E}_\mu[\eta_j(z)] = \mathbb{E}_\mu[\eta_j(0)] = 0$ because μ is the invariant probability of v . (iii) We apply Itô’s formula [7, Theorem 2.3.1] to calculate the infinitesimal generator \mathcal{G}_v of v : if we define $v_{1,r} := \Re(v_1)$, $v_{1,i} := \Im(v_1)$, $v_{2,r} := \Re(v_2)$, and $v_{2,i} := \Im(v_2)$, we have

$$\begin{aligned} \mathcal{G}_v = & (-\gamma_c v_{1,r} - \gamma_s v_{1,i}) \frac{\partial}{\partial v_{1,r}} + (-\gamma_c v_{1,i} + \gamma_s v_{1,r}) \frac{\partial}{\partial v_{1,i}} + (-\gamma_c v_{2,r} + \gamma_s v_{2,i}) \frac{\partial}{\partial v_{2,r}} + (-\gamma_c v_{2,i} - \gamma_s v_{2,r}) \frac{\partial}{\partial v_{2,i}} \\ & + \frac{\gamma_c}{2} (v_{2,r}^2 + v_{2,i}^2) \left(\frac{\partial^2}{\partial v_{1,r}^2} + \frac{\partial^2}{\partial v_{1,i}^2} \right) + \frac{\gamma_c}{2} (v_{1,r}^2 + v_{1,i}^2) \left(\frac{\partial^2}{\partial v_{2,r}^2} + \frac{\partial^2}{\partial v_{2,i}^2} \right) \\ & + \gamma_c (v_{1,i} v_{2,i} - v_{1,r} v_{2,r}) \left(\frac{\partial^2}{\partial v_{1,r} \partial v_{2,r}} - \frac{\partial^2}{\partial v_{1,i} \partial v_{2,i}} \right) + \gamma_c (-v_{1,r} v_{2,i} - v_{1,i} v_{2,r}) \left(\frac{\partial^2}{\partial v_{1,r} \partial v_{2,i}} + \frac{\partial^2}{\partial v_{1,i} \partial v_{2,r}} \right). \end{aligned}$$

We can check that $\mathcal{G}_v \eta_k = -4\gamma_c \eta_k$ for $k = 1, 2, 3$. Then for any k and j we have

$$\frac{d}{dz} \mathbb{E}_\mu[\eta_j(0)\eta_k(z)] = \mathbb{E}_\mu[\eta_j(0)\mathcal{G}_v \eta_k(z)] = -4\gamma_c \mathbb{E}_\mu[\eta_j(0)\eta_k(z)]$$

which can be readily integrated as $\mathbb{E}_\mu[\eta_j(0)\eta_k(z)] = \mathbb{E}_\mu[\eta_j(0)\eta_k(0)] \exp(-4\gamma_c z)$. Finally, the calculations of the expectations $\mathbb{E}_\mu[\eta_j(0)\eta_k(0)]$ and the integral $\int_0^{+\infty} \exp(-4\gamma_c z) dz$ complete the proof. \square

3. Random birefringence with nonconstant birefringence strength

Experimental results [9] have provided evidence that the second model of Wai and Menyuk [2,3] is more appropriate for the modeling of PMD in optical fibers. This section is devoted to the study of this model. As previously, the evolution of the two polarization modes of the electric field is governed by the coupled Schrödinger equation (1) where $bK(z)$ is now given by

$$bK(z) = \tilde{\theta}_1(z) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \tilde{\theta}_2(z) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$\tilde{\theta}_1$ and $\tilde{\theta}_2$ are independent Langevin (also called Ornstein–Uhlenbeck) processes

$$d\tilde{\theta}_1(z) = -c\tilde{\theta}_1(z)dz + \sigma dW_1(z), \quad d\tilde{\theta}_2(z) = -c\tilde{\theta}_2(z)dz + \sigma dW_2(z), \tag{9}$$

W_1 and W_2 are two independent Brownian motions, and c and σ are two positive constants related to the fiber parameters.

Note that $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are Gaussian ergodic Markov processes. They admit a unique invariant probability measure which is the Gaussian distribution with zero-mean and variance $\sigma^2/(2c)$. We denote by $\langle \cdot \rangle$ the expectation with respect to this probability measure. The square average birefringence strength is $b^2 = \langle \tilde{\theta}_1^2 \rangle + \langle \tilde{\theta}_2^2 \rangle$, with $b^2 = \sigma^2/c$. To express c in terms of the experimental fiber parameters, we compute the fiber autocorrelation function

$$\langle \tilde{\theta}_1(0)\tilde{\theta}_1(z) \rangle + \langle \tilde{\theta}_2(0)\tilde{\theta}_2(z) \rangle = b^2 \exp(-cz),$$

which shows that $c = 1/L_c$ where L_c is the PMD correlation length. Note that this model is clearly different from the model presented in Section 2. Indeed, in the first model, the birefringence strength (that is to say the squared root of $\det(bK(z))$) is constant, whereas in the second model the square of the birefringence strength has an exponential probability distribution function (because it is the sum of the squares of two independent Gaussian random variables).

We make the change of variables $\tilde{E} = M^{-1}(z)E$ where

$$M(z) = \begin{pmatrix} v_1(z) & v_2(z) \\ -\bar{v}_2(z) & \bar{v}_1(z) \end{pmatrix}$$

is the unitary matrix solution of $i \frac{dM}{dz} + bK(z)M = 0$. The coefficients v_1 and v_2 satisfy the ordinary differential equations

$$\frac{dv_1}{dz} = i\tilde{\theta}_1 v_1 - i\tilde{\theta}_2 \bar{v}_2, \quad \frac{dv_2}{dz} = i\tilde{\theta}_2 \bar{v}_1 + i\tilde{\theta}_1 v_2.$$

The partial differential equation describing the evolution of the new electric field \tilde{E} is

$$i \frac{\partial \tilde{E}}{\partial z} + \frac{d_0}{2} \frac{\partial^2 \tilde{E}}{\partial t^2} + ib' \Sigma(z) \frac{\partial \tilde{E}}{\partial t} = 0, \tag{10}$$

where $\Sigma = M^{-1} K M$. Note that M is unitary so that, for any z, t , we have $\|E(z, t)\|_{\mathbb{C}^2} = \|\tilde{E}(z, t)\|_{\mathbb{C}^2}$. The matrix $\Sigma(z)$ is given by

$$\Sigma(z) = \eta_1(z) P_1 + \eta_2(z) P_2 + \eta_3(z) P_3,$$

where P_1, P_2 , and P_3 are the Pauli matrices and η_1, η_2 , and η_3 are the real-valued processes given by

$$\begin{aligned} \eta_1 &= \Re(2\bar{v}_1 v_2 \theta_1 + (\bar{v}_1^2 - v_2^2) \theta_2), \\ \eta_2 &= \Im(2\bar{v}_1 v_2 \theta_1 + (\bar{v}_1^2 - v_2^2) \theta_2), \\ \eta_3 &= (|v_1|^2 - |v_2|^2) \tilde{\theta}_1 - 2\Re(v_1 v_2) \tilde{\theta}_2, \end{aligned} \tag{11}$$

and we have introduced the normalized processes $\theta_1 = \tilde{\theta}_1/b$ and $\theta_2 = \tilde{\theta}_2/b$. The random matrix $\Sigma(z)$ is quite different from the corresponding matrix that was obtained with the first model in Section 2. In particular, we have

$$|\eta(z)|^2 = \eta_1^2(z) + \eta_2^2(z) + \eta_3^2(z) = \theta_1^2(z) + \theta_2^2(z)$$

which shows that the process $\eta(z)$ does not take values on the sphere \mathbb{S}^2 of \mathbb{R}^3 , in contrast with $\eta(z)$ defined by (7), but the statistics of $|\eta(z)|^2$ is an exponential distribution with mean 1. However we are going to show that the random matrix $\Sigma(z)$ possesses the same key statistical properties as the ones obtained with the first model. We denote $\theta = (\theta_1, \theta_2)$ and $v = (v_1, v_2)$.

Proposition 3. *The process (θ, v) is $\mathbb{R}^2 \times \mathbb{S}^3$ -valued, where \mathbb{S}^3 is the unit sphere of $\mathbb{C}^2 \sim \mathbb{R}^4$. It is an ergodic Markov process with the infinitesimal generator \mathcal{G} given (12). It has a unique invariant probability measure*

$$\mu(d\theta dv) = \frac{1}{\pi} \exp(-\theta_1^2 + \theta_2^2) \lambda_{\mathbb{S}^3}(dv) d\theta_1 d\theta_2,$$

where $\lambda_{\mathbb{S}^3}$ is the uniform distribution on \mathbb{S}^3 . Moreover, if $f : \mathbb{R}^2 \times \mathbb{S}^3 \rightarrow \mathbb{R}$ is a \mathcal{C}^2 bounded function with compact support such that $\int f d\mu = 0$, then the Poisson equation $\mathcal{G}g + f = 0$ has a \mathcal{C}^2 bounded solution g which can be written as

$$g(\theta_0, v_0) = \int_0^{+\infty} \mathbb{E}[f(\theta(z), v(z)) | \theta(0) = \theta_0, v(0) = v_0] dz.$$

Proof. We have $|v|^2 = |v_1|^2 + |v_2|^2 = 1$, which shows that v takes values in \mathbb{S}^3 . The Ornstein–Uhlenbeck processes θ_1 and θ_2 are Markov, and v is the solution of an ordinary differential equation driven by θ . As a result, the joint process (θ, v) is Markov and its infinitesimal generator is given by

$$\begin{aligned} \mathcal{G} &= -c\theta_1 \frac{\partial}{\partial \theta_1} - c\theta_2 \frac{\partial}{\partial \theta_2} + \frac{c}{2} \frac{\partial^2}{\partial \theta_1^2} + \frac{c}{2} \frac{\partial^2}{\partial \theta_2^2} + ib(\theta_1 v_1 - \theta_2 \bar{v}_2) \frac{\partial}{\partial v_1} - ib(\theta_1 \bar{v}_1 - \theta_2 v_2) \frac{\partial}{\partial \bar{v}_1} \\ &+ ib(\theta_1 v_2 - \theta_2 \bar{v}_1) \frac{\partial}{\partial v_2} - ib(\theta_1 \bar{v}_2 + \theta_2 v_1) \frac{\partial}{\partial \bar{v}_2}. \end{aligned} \tag{12}$$

Here we have written the generator in terms of $(v_1, \bar{v}_1, v_2, \bar{v}_2)$ rather than in terms of the real and imaginary parts of v_1 and v_2 . As usual, if $z = a + ib$ is complex, then $\partial_z = (1/2)(\partial_a - i\partial_b)$ and $\partial_{\bar{z}} = (1/2)(\partial_a + i\partial_b)$. Once the infinitesimal generator is known, we can check that the measure μ is an invariant probability measure. From the classification of Markov processes [7, Theorem 1.3.10], the process is ergodic. The last point is a consequence of [7, Theorem 1.3.15]. \square

We denote by \mathbb{E}_μ the expectation with respect to the process (v, θ) starting from the invariant distribution μ . We can then write the analog of Proposition 2.

Proposition 4. (i) *The process (η_1, η_2, η_3) defined by (11) is a function of (v, θ) , which is an ergodic Markov process with invariant probability measure μ . (ii) For $j = 1, 2, 3$ we have $\mathbb{E}_\mu[\eta_j(z)] = 0$ and $\mathbb{E}_\mu[\eta_j^2(z)] = 1/3$ for any $z \geq 0$. (iii) We have, for $j, k = 1, 2, 3$,*

$$\int_0^{+\infty} \mathbb{E}_\mu[\eta_j(0)\eta_k(z)]dz = \begin{cases} \frac{L_c}{3} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Proof. (i) The first point follows from the definition of v and the properties of the Ornstein–Uhlenbeck processes θ_1 and θ_2 .(ii) The computations of the first and second moments with respect to the invariant measure μ gives the result, as well as the decorrelation property

$$\mathbb{E}_\mu[\eta_j(0)\eta_k(0)] = \begin{cases} \frac{1}{3} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{13}$$

(iii) We apply the infinitesimal generator \mathcal{G} given by (12) on η_j and we get simply $\mathcal{G}\eta_j = -c\eta_j$. As a result we obtain the differential equation

$$\frac{d}{dz} \mathbb{E}_\mu[\eta_j(z)\eta_k(0)] = -c\mathbb{E}_\mu[\eta_j(z)\eta_k(0)].$$

Integrating and using the initial condition (13) completes the proof. \square

4. Asymptotic pulse dynamics

The typical fiber length for transoceanic transmission is 10^3 – 10^4 km and the typical PMD correlation length L_c is of the order of 10–100 m (see Section 5 for numerical applications). Accordingly, we introduce a new small parameter ε so that the correlation length is of order l and the fiber length is of order ε^{-2} . We look for the evolution of the field for macroscopic propagation distances of the order of ε^{-2} . We address initial conditions that will be affected by GVD and PMD over such a propagation distance. In this framework this means that U_0 varies on a time scale of order ε . We show in this section that PMD has a macroscopic effect on such a pulse. Accordingly we consider $U^\varepsilon(z, t) := \tilde{E}(z/\varepsilon^2, t/\varepsilon)$ which is solution of

$$i \frac{\partial U^\varepsilon}{\partial z}(z, t) + \frac{d_0}{2} \frac{\partial^2 U^\varepsilon}{\partial t^2}(z, t) + \frac{ib'}{\varepsilon} \Sigma\left(\frac{z}{\varepsilon^2}\right) \frac{\partial U^\varepsilon}{\partial t}(z, t) = 0, \tag{14}$$

with the initial condition $U^\varepsilon(z = 0, t) = U_0(t)$. Denoting by $\eta^\varepsilon = (\eta_1^\varepsilon, \eta_2^\varepsilon, \eta_3^\varepsilon)$ the process $\eta(\cdot/\varepsilon^2)$, we have

$$\Sigma\left(\frac{z}{\varepsilon^2}\right) = \eta_1^\varepsilon(z)P_1 + \eta_2^\varepsilon(z)P_2 + \eta_3^\varepsilon(z)P_3.$$

The main result of this section is the following one.

Proposition 5. *Assume for simplicity that the Fourier transform of U_0 is a \mathcal{C}^2 compactly supported function. (i) Eq. (14) has a unique solution U^ε a.s. in $\mathcal{C}([0, \infty), L^2(\mathbb{R}, \mathbb{C}^2))$ such that $U^\varepsilon(z = 0, t) = U_0(t)$. Moreover, for all $z \geq 0$,*

$$\|U^\varepsilon(z, \cdot)\|_{L^2(\mathbb{R}, \mathbb{C}^2)} = \|U_0\|_{L^2(\mathbb{R}, \mathbb{C}^2)}. \tag{15}$$

(ii) *Let W_1, W_2 , and W_3 be three independent Brownian motions and $\gamma = 2b^2L_c/3$. The equation*

$$U(z, t) - U_0(t) - \frac{id_0}{2} \int_0^z \frac{\partial^2 U}{\partial t^2}(\zeta, t)d\zeta + \sqrt{\gamma} \sum_{j=1}^3 \int_0^z P_j \frac{\partial U}{\partial t}(\zeta, t) \circ dW_j(\zeta) = 0, \tag{16}$$

has a unique solution U a.s. in $\mathcal{C}([0, \infty), L^2(\mathbb{R}, \mathbb{C}^2))$. Moreover, for all $z \geq 0$,

$$\|U(z, \cdot)\|_{L^2(\mathbb{R}, \mathbb{C}^2)} = \|U_0\|_{L^2(\mathbb{R}, \mathbb{C}^2)}. \tag{17}$$

(iii) *As $\varepsilon \rightarrow 0$, U^ε converge in distribution in $\mathcal{C}([0, \infty), L^2(\mathbb{R}, \mathbb{C}^2))$ to U .*

We can see that PMD has a macroscopic effect on the pulse: in the asymptotic framework $\varepsilon \rightarrow 0$ the pulse dynamics is driven by three real-valued Brownian motions. A fact which is remarkable is that the three driving processes η_j are correlated in both PMD models (this is especially striking for the first model where

$\eta_1^2 + \eta_2^2 + \eta_3^2 = 1$), but they act on the pulse like three independent processes at the macroscopic scale. The effective propagation equation is the same for the first and second PMD models. All that remains from the microscopic PMD model is the value of the effective PMD parameter γ .

The proof of Proposition 5 is given in Appendix A.1. It consists in studying the Fourier transforms \hat{U}^ε and \hat{U} of U^ε and U , applying relative compactness criteria in the space of L^2 -valued processes, and proving approximation-diffusion results based on the perturbed test function method. Note that the stochastic integrals in Eq. (16) are Stratonovich integrals. In the standard Itô form, the equation reads as

$$U(z, t) - U_0(t) - \frac{id_0}{2} \int_0^z \frac{\partial^2 U}{\partial t^2}(\zeta, t) d\zeta - \frac{3\gamma}{2} \int_0^z \frac{\partial^2 U}{\partial t^2}(\zeta, t) d\zeta + \sqrt{\gamma} \sum_{j=1}^3 \int_0^z P_j \frac{\partial U}{\partial t}(\zeta, t) dW_j(\zeta) = 0, \tag{18}$$

which exhibits an effective diffusion term.

5. Numerical simulations

The theory developed in this paper shows that the pulse propagation does not depend on the PMD model in the linear propagation regime. The goal of this section is to illustrate this fact, and to show by numerical simulations that it also holds true in the nonlinear regime.

We first consider linear pulse propagation in presence of PMD and in absence of GVD, as modeled by Eq. (1) with $d_0 = 0$. We consider a Gaussian pulse as an initial condition with an initial r.m.s. width $T_0 = 4$ ps. In order to simulate the random fluctuations of birefringence, we consider the first model studied in Section 2 and assume that the birefringence strength is constant and the birefringence angle is stepwise linear. The process α is consequently stepwise constant. We take $\alpha(z) = W_{N(z)}$ where $(W_n)_{n \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables with Gaussian distribution, zero-mean and variance σ^2 . $(N(z))_{z \geq 0}$ is a Poisson process, that is to say $N(z) = n \in \mathbb{N}$ if $\sum_{k=1}^n \zeta_k \leq z < \sum_{k=1}^{n+1} \zeta_k$ where $(\zeta_k)_{k \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables with exponential distribution and mean h . Accordingly α is a stationary Markov process and $\mathbb{E}[\alpha(0)\alpha(z)] = \sigma^2 \exp(-z/h)$. The parameter γ_c is given by $\gamma_c = \sigma^2 h / (1 + 4b^2 h^2)$. We take standard values for the fiber parameters [10,4]. The birefringence is small in absolute values in communication fibers, with values of the order of $\Delta n / n \sim 10^{-7} - 10^{-5}$ so that $b = \pi \Delta n / \lambda \simeq 0.2 - 20 \text{ m}^{-1}$ ($\lambda = 1.55 \text{ }\mu\text{m}$). The orientation of the birefringence is randomly varying on a length scale of a few centimeters. We take $b = 4000 \text{ km}^{-1}$, $b' = 1 \text{ ps/km}$, $h = 10^{-4} \text{ km}$, and $\sigma = 800 \text{ km}^{-1}$. In these conditions $\gamma_c \simeq 39 \text{ km}^{-1}$, which means that the PMD correlation length is $L_c = 1/(4\gamma_c) \simeq 6.5 \text{ m}$ (which is larger than the beat length $L_B \simeq 0.8 \text{ m}$). These values correspond to measured experimental values [11,12]. Finally $\gamma = 4.3 \times 10^{-3} \text{ ps}^2/\text{km}$. We consider a fiber length of the order of a few $L_0 = T_0^2/\gamma \simeq 400 \text{ km}$ so as to exhibit a noticeable pulse spreading. Note that $L_B \ll L_c \ll L_0$ so that the separation of scales that we have introduced in this paper holds true. The number of samples used to evaluate the averaged values and variances is 10^4 . We have solved the linear Schrödinger equation in the Fourier domain, so that the numerical problem is reduced to the integration of a set of ordinary differential equations for the Fourier components. Comparisons of the numerical results and theoretical formulas given in [13,14] show excellent agreement (Figs. 1 and 2).

We now address the propagation equation in presence of GVD, PMD, and nonlinearity. The evolution of the two polarization modes of the electric field is governed by the coupled nonlinear Schrödinger equations [2]

$$i \frac{\partial E}{\partial z} + bK(z)E + ib'K(z) \frac{\partial E}{\partial t} + \frac{d_0}{2} \frac{\partial^2 E}{\partial t^2} + N(E) = 0, \tag{19}$$

where $N(E)$ stands for the nonlinear terms

$$N(E) = n_2 \begin{pmatrix} (|E_1|^2 + \alpha_2 |E_2|^2) E_1 + \frac{\alpha_2}{2} E_2^2 \overline{E_1} \\ (|E_2|^2 + \alpha_2 |E_1|^2) E_2 + \frac{\alpha_2}{2} E_1^2 \overline{E_2} \end{pmatrix}, \tag{20}$$

and the cross-phase modulation is $\alpha_2 = 2/3$ for linearly birefringent fiber. In absence of high-order PMD $b' = 0$, a direct averaging procedure [3] leads to the result that U^ε converge to U solution of

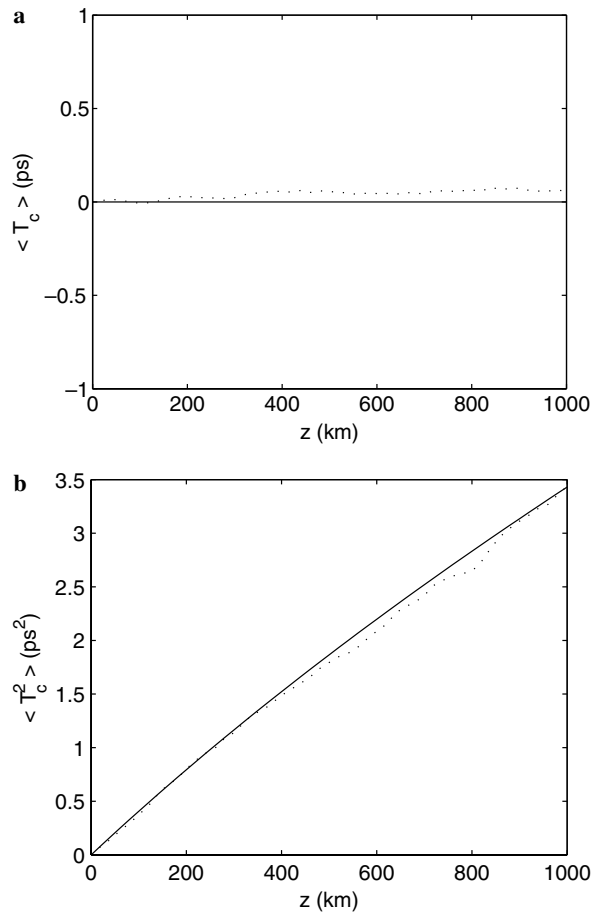


Fig. 1. Expectation (a) and variance (b) of the time displacement $T_c(z) := \int t |E(z,t)|^2 dt / \int |E(z,t)|^2 dt$. Comparison between numerical simulations (dotted lines) and theoretical predictions (solid lines): $\langle T_c(z) \rangle = 0$ and $\langle T_c^2(z) \rangle = (T_0^2/2) [\sqrt{1 + 4\gamma z/T_0} - 1]$.

$$i \frac{\partial U}{\partial z} + \frac{d_0}{2} \frac{\partial^2 U}{\partial t^2} + \frac{8}{9} n_2 |U|^2 U = 0,$$

where the nonlinear term $|U|^2 U$ reads explicitly as $(|U_1|^2 + |U_2|^2) (U_1, U_2)^T$. In absence of high-order PMD $b' = 0$ we thus get the integrable Manakov system that supports soliton solutions [15]. The numerical simulations are performed with a split-step Fourier method. We have carried out numerical simulations of Eq. (19) in presence of high-order PMD $b' \neq 0$ with the two birefringence models introduced in Sections 2 and 3, respectively. Our main motivation to compare the two different PMD models is to show that, even in the presence of nonlinearity, there is no difference in behavior.

The first PMD model is considered with a stepwise constant α . We take $\alpha(z) = W_{[z/h]}$ where $[z]$ stands for integral part of a real number z and $(W_n)_{n \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables with Gaussian distribution, zero-mean and variance σ^2 . Accordingly the parameter γ_c is given by $\gamma_c = \sigma^2 [1 - \cos(2bh)] / (4b^2h)$. Note that this numerical model is very close to the one considered here above, but it involves an expression for γ_c which is quite different. This illustrates the fact that the effective PMD parameter has a complicated dependence on the microscopic PMD model. We take $b = 4000 \text{ km}^{-1}$, $b' = 2 \text{ ps/km}$, $h = 10^{-4} \text{ km}$, and $\sigma = 800 \text{ km}^{-1}$. In these conditions $\gamma_c \simeq 30 \text{ km}^{-1}$, which means that the PMD correlation length is $L_c = 1/(4\gamma_c) \simeq 8 \text{ m}$ and the effective PMD parameter is $\gamma = 2.2 \cdot 10^{-2} \text{ ps}^2/\text{km}$. The initial pulse is a sech pulse $\sqrt{P_0} \text{sech}(t/T_0)$ with $T_0 = 6 \text{ ps}$ and $P_0 = 1/36 \text{ W}$. We carry out two series of simulations, the first one in the linear propagation regime in absence of GVD, the second one in the nonlinear

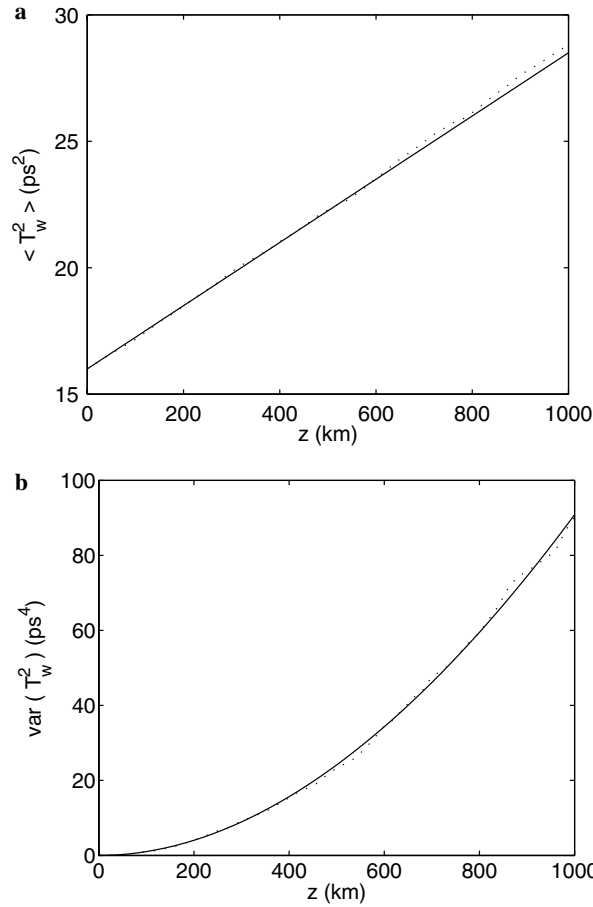


Fig. 2. Expectation (a) and variance (b) of the r.m.s. pulse width $T_w^2(z) := \int t^2 |E(z, t)|^2 dt / \int |E(z, t)|^2 dt$. Comparison between numerical simulations (dotted lines) and theoretical predictions (solid lines): $\langle T_w^2 \rangle = T_0^2 + 3\gamma z$ and $\text{Var}(T_w^2) = T_0(T_0^2 + 4\gamma z)^{3/2} - 6T_0^2\gamma z - T_0^4$.

dispersive propagation regime. More exactly we consider a nonlinear coefficient $n_2 = 9/8 \text{ W}^{-1} \text{ km}^{-1}$ and a GVD parameter $d_0 = 1 \text{ ps}^2/\text{km}$ so that the initial sech pulse is a Manakov soliton in absence of high-order PMD $b' = 0$. The number of samples used to evaluate the averaged values is 200. Once again, we can observe in Fig. 3 that the numerical simulations in the linear regime agree with the theoretical predictions. We can also observe that the broadening of the Manakov soliton is in the first steps similar to that of the corresponding linear pulse, but after a while the departure increases and it eventually turns out that the Manakov soliton is less stable. This may seem surprising and in contradiction with previous results [16–18]. However this can be explained by the fact that the r.m.s. pulse width is plotted in Fig. 3, while the Full width at half maximum (FWHM) was plotted in the previous studies. We plot in Fig. 4b the FWHM for the same set of experiments as in Fig. 3. The results are in perfect agreement with the ones obtained in [17] for instance. We recover the intuitive result that the Manakov soliton is more robust than the linear pulse. The opposite behaviors of the r.m.s. pulse width and the FWHM can be explained by considering the typical pulse profile at the fiber output plotted in Fig. 4a.

- (1) On the one hand, we can see that the central part of the soliton is very robust, because of the self-trapping induced by the nonlinear potential generated by the nonlinear term. The FWHM is only sensitive to the central part of the pulse. As a result the pulse broadening measured in terms of the FWHM is reduced compared to the linear case.
- (2) On the other hand, we can observe that the soliton emits small-amplitude radiation, which escapes the attraction of the nonlinear potential. The phenomenon of dispersion radiation emitted by solitons in the

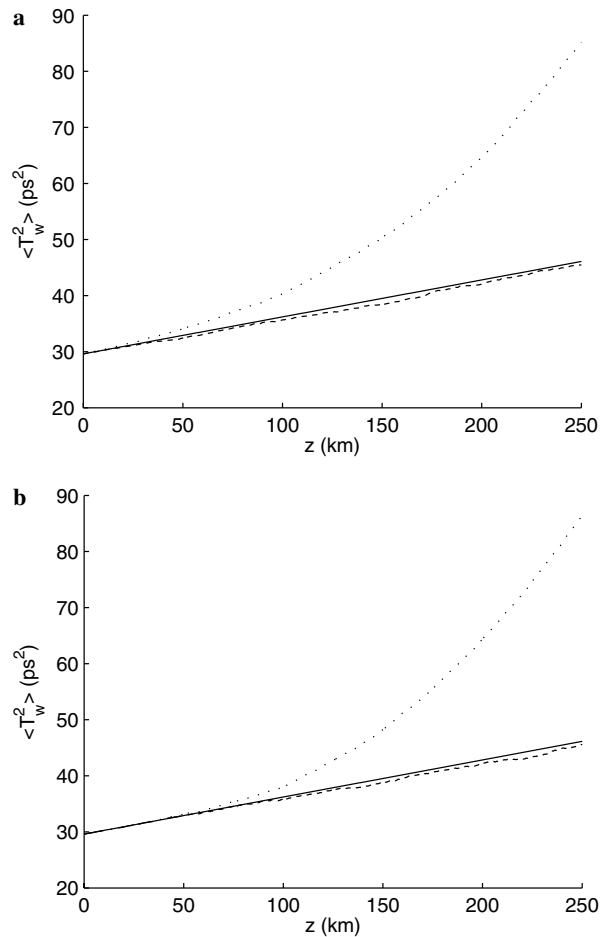


Fig. 3. Mean square pulse width for the first (a) and second (b) PMD model. The solid lines represent the theoretical predictions for the linear propagation regime, the dashed lines represent the numerical simulations for the linear regime, and the dotted lines represent the numerical simulations for the nonlinear dispersive regime.

presence of PMD has been discussed in great detail. It was first mentioned in the work of Mollenauer et al. [19], and then in more detail in [20,21]. The FWHM is not sensitive to the small-amplitude radiation, but the r.m.s. pulse width is very sensitive to it. The radiation spreading actually involves the fast increase of the r.m.s. pulse width exhibited in Fig. 3.

The second PMD model is also considered. We choose $b = 4000 \text{ km}^{-1}$, $L_c = 25 \text{ m}$, and we take $b' = 1.15 \text{ ps/km}$ so that the effective PMD parameter is $\gamma = 2.2 \cdot 10^{-2} \text{ ps}^2/\text{km}$, as for the first model. We use an exact updating formula to integrate Eq. (9) [22], where $\sigma = b/\sqrt{L_c}$ and $c = 1/L_c$. The results are plotted in Fig. 3b. By comparing qualitatively and quantitatively Fig. 3b and Fig. 3a, we can see that the results do not depend on the choice of the microscopic PMD model. This observation is in perfect agreement with the theoretical predictions because both models give rise to the same value of the effective PMD parameter γ . These series of simulations confirm the main theoretical prediction that the pulse dynamics, in linear and nonlinear media, depends on a unique effective PMD parameter. We have carried out our numerical simulations with very different PMD models, but we have taken care to consider models involving the same values for the effective PMD parameters. The numerical results have shown that the pulse dynamics with the different models are the same.

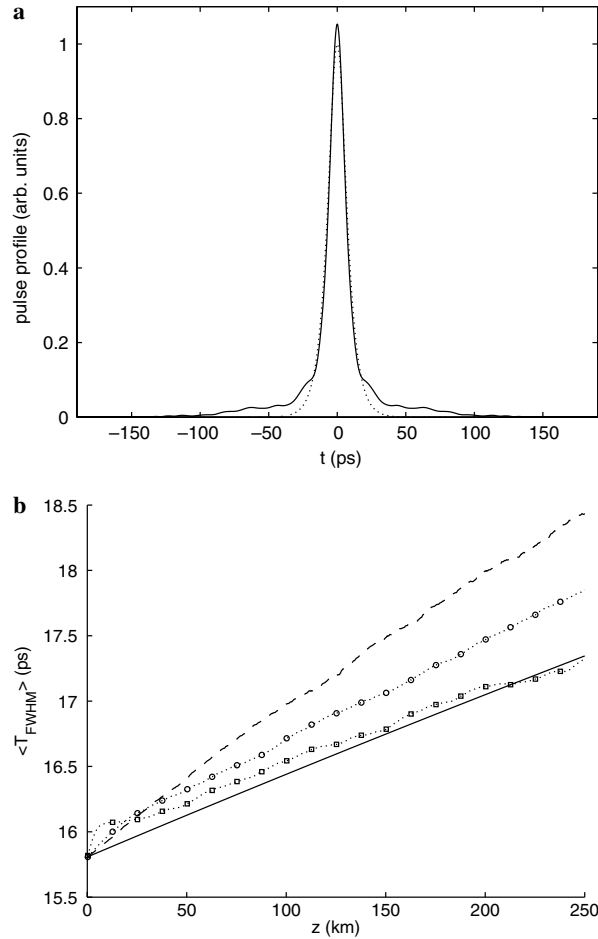


Fig. 4. (a) Input (dotted) and output (solid) pulse profiles in the nonlinear regime. The input profile is a Manakov soliton $U_0(t) = \sqrt{P_0} \text{sech}(t/T_0)$ with $T_0 = 6$ ps and $P_0 = 1/36$ W. We can observe that the main part of the soliton is well preserved, but small-amplitude radiation is emitted and spreads out far from the soliton. (b) FWHM in the linear case in absence of GVD (dashed), in the nonlinear case with $d_0 = 1$ ps²/km, $n_2 = 9/8$ W⁻¹ km⁻¹ (dotted with circles) and with $d_0 = 4$ ps²/km, $n_2 = 9/2$ W⁻¹ km⁻¹ (dotted with squares). The enhanced robustness of the soliton compared to the linear pulse is obvious. For (d_0, n_2) large enough, the soliton spreads out according to the relation predicted by Matsumoto et al. [16] $T_{FWHM}(z)/T_{FWHM}(z = 0) = \sqrt{1 + 4\gamma z / (3T_0^2)}$ (solid line).

6. Conclusion

In this paper we show that the effective pulse propagation equation in randomly birefringent fibers reads as a stochastic partial differential equation driven by three independent Brownian motions. This result holds true when the beat length is smaller than the PMD correlation length which is itself smaller than the other length scales present in the problem (dispersion length, nonlinear length, dispersion map period, fiber length). The striking feature is that this stochastic equation depends on PMD through a single effective parameter whatever the PMD model. The PMD effective parameter has a complicated dependence on the microscopic PMD model, however our explicit calculations carried out in the linear propagation regime shows that it can be estimated very simply by an experimental measure of the pulse broadening in the linear propagation regime. Accordingly, PMD simulations can be carried out with an arbitrary microscopic model without loss of generality. The only requirement is that the simulated PMD gives rise to the same effective PMD parameter as the experimental value.

Appendix A

A.1. Proof of Proposition 5

Proof of (i): We consider the equation

$$i \frac{\partial X^\varepsilon}{\partial z}(z, \theta) - \theta^2 \frac{d_0}{2} X^\varepsilon(z, \theta) + \frac{b'\theta}{\varepsilon} \tilde{\Sigma}\left(\frac{z}{\varepsilon^2}\right) X^\varepsilon(z, \theta) = 0 \tag{A.1}$$

with the initial condition $X^\varepsilon(z = 0, \theta) = X_0(\theta)$. If θ is fixed, Eq. (A.1) is an ordinary differential equation; by Cauchy–Lipschitz Theorem, it admits a unique solution $X^\varepsilon(\cdot, \theta)$ a.s. in $\mathcal{C}([0, Z], \mathbb{C}^2)$. A calculation of $d\|X^\varepsilon(z, \cdot)\|_{\mathbb{C}^2}^2/dz$ shows that

$$\|X^\varepsilon(z, \cdot)\|_{L^2(\mathbb{R}, \mathbb{C}^2)} = \|X_0\|_{L^2(\mathbb{R}, \mathbb{C}^2)}. \tag{A.2}$$

We define

$$U^\varepsilon(z, t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{it\theta} X^\varepsilon(z, \theta) d\theta,$$

and we derive the existence and uniqueness of the solution of (14). By Parseval’s formula and Eq. (A.2), we have (15). By (A.2) and classical theorems about regularity of solutions of ordinary differential equation with respect to initial conditions [23], we get that $X^\varepsilon(z, \cdot)$ is a \mathcal{C}^2 compactly support function in θ .

Proof of (ii): We consider the equation

$$X(z, \theta) - X_0(\theta) + i\theta^2 \frac{d_0}{2} \int_0^z X(\zeta, \theta) d\zeta - i\theta\sqrt{\gamma} \sum_{j=1}^3 \int_0^z P_j X(\zeta, \theta) \circ dW_j(\zeta) = 0. \tag{A.3}$$

If θ is fixed, (A.3) is a stochastic differential equation with linear coefficients; by the usual existence and uniqueness theorem [7, Theorem 3.4.7], it admits a unique solution $X(\cdot, \theta)$ a.s. in $\mathcal{C}([0, Z], \mathbb{C}^2)$ such that $X(0, \theta) = X_0(\theta)$. Applying Itô’s formula to $\|X(z, \cdot)\|_{\mathbb{C}^2}^2$ shows that

$$\|X(z, \cdot)\|_{L^2(\mathbb{R}, \mathbb{C}^2)} = \|X_0\|_{L^2(\mathbb{R}, \mathbb{C}^2)}. \tag{A.4}$$

We define

$$U(z, t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{it\theta} X(z, \theta) d\theta,$$

and we derive the existence and uniqueness of the solution of (16). By Parseval’s formula and (A.4), we have (17).

Proof of (iii): We consider the Fourier transforms of U^ε and U :

$$\widehat{U}^\varepsilon(\cdot, \theta) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-it\theta} U^\varepsilon(\cdot, t) dt \quad \text{and} \quad \widehat{U}(\cdot, \theta) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-it\theta} U(\cdot, t) dt.$$

We suppose that the support of \widehat{U}_0 belongs to $[-A, A]$ where $A \in \mathbb{N}^*$. To prove (iii), we show that $(\widehat{U}^\varepsilon)_\varepsilon$ converge in distribution in $\mathcal{C}([0, Z], L^2([-A, A], \mathbb{C}^2))$ to \widehat{U} as $\varepsilon \rightarrow 0$.

First, we prove the convergence of the finite-dimensional distributions. We fix n frequencies $\theta_1, \theta_2, \dots, \theta_n$ and we apply Theorem 6 to the process $(\widehat{U}^\varepsilon(\cdot, \theta_1), \widehat{U}^\varepsilon(\cdot, \theta_2), \dots, \widehat{U}^\varepsilon(\cdot, \theta_n))$. The hypotheses of the theorem are readily fulfilled by the first model, in view of Propositions 1 and 2. For the second model, we must invoke Propositions 3 and 4, and we also note that we can interpret the process (θ, v) as having compact support, at the expense of considering $(\text{atan}(\theta_1), \text{atan}(\theta_2))$ instead of (θ_1, θ_2) .

Second, we prove the tightness of $(\widehat{U}^\varepsilon)_\varepsilon$ in $\mathcal{C}([0, Z], L^2([-A, A], \mathbb{C}^2))$. Because the weak limit belongs to $\mathcal{C}([0, Z], L^2([-A, A], \mathbb{C}^2))$, it is sufficient to show that $(\widehat{U}^\varepsilon)_\varepsilon$ is tight in $\mathcal{D}([0, Z], L^2([-A, A], \mathbb{C}^2))$ equipped with the Skorohod topology. We recall a definition and two standard tightness criteria [24]:

Definition A.1. Let $(E, \|\cdot\|)$ be a normed space and $(X^\varepsilon)_{\varepsilon>0} = ((X^\varepsilon(z))_{z \in [0, Z]})_{\varepsilon>0}$ be a sequence of processes with paths in $\mathcal{D}([0, Z], E)$. We say that $(X^\varepsilon)_{\varepsilon>0}$ satisfies the Aldous property if for all $M > 0$, $\eta > 0$ and $\lambda > 0$, there exists $\delta > 0$ such that

$$\limsup_{\varepsilon \rightarrow 0} \sup_{\tau \in \mathcal{T}_M} \sup_{0 < \delta' < \delta} \mathbb{P}(\|X^\varepsilon(\tau + \delta') - X^\varepsilon(\tau)\| > \lambda) < \eta,$$

where \mathcal{T}_M is the space of all the stopping times bounded by M .

Lemma A.1. Let $(E, \|\cdot\|)$ be a normed space and $(X^\varepsilon)_{\varepsilon>0} = ((X^\varepsilon(z))_{z \in [0, Z]})_{\varepsilon>0}$ be a sequence of processes with paths in $\mathcal{D}([0, Z], E)$. If for all z in a dense subset of $[0, Z]$ $(X^\varepsilon(z))_{0 < \varepsilon \leq 1}$ is tight in E and satisfies the Aldous property, then $(X^\varepsilon)_{0 < \varepsilon \leq 1}$ is tight in $\mathcal{D}([0, Z], E)$.

Lemma A.2. Let H be a Hilbert space and $(H_k)_{k \in \mathbb{N}}$ be an increasing sequence of finite-dimensional subspace of H such that, for any h in H , $\lim_{k \rightarrow \infty} \pi_{H_k} h = h$, where π_{H_k} stands for the projection onto H_k . Let $(Y^\varepsilon)_{\varepsilon>0}$ be a sequence of random variables in H . The family $(Y^\varepsilon)_{\varepsilon>0}$ is tight if and only if for all $\eta > 0$ and $\lambda > 0$, there exist $\rho_\eta > 0$ and an integer $k_{\eta, \lambda}$ such that

$$\sup_{\varepsilon \in (0, 1]} \mathbb{P}(\|Y^\varepsilon\| \geq \rho_\eta) \leq \eta \quad \text{and} \quad \sup_{\varepsilon \in (0, 1]} \mathbb{P}(d(Y^\varepsilon, H_{k_{\eta, \lambda}}) > \lambda) \leq \eta.$$

In view of **Lemma A.1**, the proof of the tightness of $(\widehat{U}^\varepsilon)_\varepsilon$ consists of two steps: the tightness of $(\widehat{U}^\varepsilon(z, \cdot))_\varepsilon$ in $L^2([-A, A], \mathbb{C}^2)$ for all z , and the Aldous property.

Step 1: $(\widehat{U}^\varepsilon(z, \cdot))_\varepsilon$ is tight in $L^2([-A, A], \mathbb{C}^2)$ FOR ALL z . We apply **Lemma A.2** with $H = L^2([-A, A], \mathbb{C}^2)$ and H_n the set of all simple functions $h = (h_1, h_2)$ such that

$$h_j(\theta) = \sum_{k=-A2^n}^{A2^n-1} \alpha_k \mathbb{1}_{\theta \in [k/2^n, (k+1)/2^n)},$$

with $\alpha_k \in \mathbb{C}$. The following criteria insure that $(\widehat{U}^\varepsilon(z, \cdot))_\varepsilon$ fulfills the hypotheses of **Lemma A.2**:

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{E}_\mu \left[\sup_{|\theta_1 - \theta_2| \leq \delta, |\theta_1| \leq A, |\theta_2| \leq A} \|\widehat{U}^\varepsilon(z, \theta_1) - \widehat{U}^\varepsilon(z, \theta_2)\|_{\mathbb{C}^2}^2 \right] = 0, \tag{A.5}$$

$$\limsup_{\varepsilon \rightarrow 0} \sup_{|\theta| \leq A} \mathbb{E}_\mu [\|\widehat{U}^\varepsilon(z, \theta)\|_{\mathbb{C}^2}^2] < \infty. \tag{A.6}$$

Let us introduce $f(\widehat{U}) := \|\widehat{U}\|_{\mathbb{C}^2}^2$, \mathcal{L}_θ the generator of $\widehat{U}(\cdot, \theta)$, $\mathcal{L}_\theta^\varepsilon$ the generator of $(\widehat{U}^\varepsilon(\cdot, \theta), v(\cdot/\varepsilon^2))$, and f_θ^ε the perturbed test functions of f which are built according to the procedure described in the proof of **Theorem 6**. The martingale property ensures that

$$\mathbb{E}_\mu[f_\theta^\varepsilon(\widehat{U}^\varepsilon(z, \theta), v(z/\varepsilon^2))] = \mathbb{E}_\mu[f_\theta^\varepsilon(\widehat{U}_0(\theta), v(0))] + \int_0^z \mathbb{E}_\mu[\mathcal{L}_\theta^\varepsilon f_\theta^\varepsilon(\widehat{U}^\varepsilon(\zeta, \theta), v(\zeta/\varepsilon^2))] d\zeta,$$

then, using the perturbed test function method (see the proof of **Theorem 6**), we have

$$\mathbb{E}_\mu[f(\widehat{U}^\varepsilon(z, \theta))] = \mathbb{E}_\mu[f(\widehat{U}_0(\theta))] + \int_0^z \mathbb{E}_\mu[\mathcal{L}_\theta f(\widehat{U}^\varepsilon(\zeta, \theta))] d\zeta + R_\theta^\varepsilon,$$

with

$$|R_\theta^\varepsilon| \leq K_\theta \varepsilon \int_0^z \mathbb{E}_\mu[f(\widehat{U}^\varepsilon(\zeta, \theta))] d\zeta,$$

where K_θ is such that $|f_\theta^\varepsilon - f| \leq K_\theta \varepsilon f$ and $|\mathcal{L}_\theta^\varepsilon f_\theta^\varepsilon - \mathcal{L}_\theta f| \leq K_\theta \varepsilon f$. We note that K_θ can be chosen to be uniformly bounded with respect to $\theta \in [-A, A]$. Moreover there exists a constant $C > 0$ such that $\mathcal{L}_\theta f \leq C\theta^2 f$, so,

$$\mathbb{E}_\mu[f(\widehat{U}^\varepsilon(z, \theta))] \leq \mathbb{E}_\mu[f(\widehat{U}_0(\theta))] + (C\theta^2 + K_\theta \varepsilon) \int_0^z \mathbb{E}_\mu[f(\widehat{U}^\varepsilon(\zeta, \theta))] d\zeta.$$

By Gronwall’s lemma, we have

$$\mathbb{E}_\mu[f(\widehat{U}^\varepsilon(z, \theta))] \leq \mathbb{E}_\mu[f(\widehat{U}_0(\theta))] \exp(C\theta^2 z + K_\theta \varepsilon).$$

To get (A.6), we take the sup on $|\theta| \leq A$ and we make $\varepsilon \rightarrow 0$. Let us show (A.5). We have for all $\theta_1, \theta_2 \in [-A, A]$,

$$\|\widehat{U}^\varepsilon(z, \theta_1) - \widehat{U}^\varepsilon(z, \theta_2)\|_{\mathbb{C}^2} \leq |\theta_1 - \theta_2| \sup_{|\theta| \leq A} \left\| \frac{\partial \widehat{U}^\varepsilon}{\partial \theta}(z, \theta) \right\|_{\mathbb{C}^2}.$$

So,

$$\mathbb{E}_\mu \left[\sup_{|\theta_1 - \theta_2| \leq \delta, |\theta_1| \leq A, |\theta_2| \leq A} \|\widehat{U}^\varepsilon(z, \theta_1) - \widehat{U}^\varepsilon(z, \theta_2)\|_{\mathbb{C}^2}^2 \right] \leq \delta^2 \mathbb{E}_\mu \left[\sup_{|\theta| \leq A} \left\| \frac{\partial \widehat{U}^\varepsilon}{\partial \theta}(z, \theta) \right\|_{\mathbb{C}^2}^2 \right].$$

From Sobolev’s imbedding $H^1([-A, A]) \hookrightarrow L^\infty([-A, A])$ there exists a constant $C > 0$ such that

$$\begin{aligned} \mathbb{E}_\mu \left[\sup_{|\theta| \leq A} \left\| \frac{\partial \widehat{U}^\varepsilon}{\partial \theta}(z, \theta) \right\|_{\mathbb{C}^2}^2 \right] &\leq C \int_{-A}^A \mathbb{E}_\mu \left[\left\| \frac{\partial \widehat{U}^\varepsilon}{\partial \theta}(z, \theta) \right\|_{\mathbb{C}^2}^2 + \left\| \frac{\partial^2 \widehat{U}^\varepsilon}{\partial \theta^2}(z, \theta) \right\|_{\mathbb{C}^2}^2 \right] d\theta \\ &\leq 2AC \sup_{|\theta| \leq A} \mathbb{E}_\mu \left[\left\| \frac{\partial \widehat{U}^\varepsilon}{\partial \theta}(z, \theta) \right\|_{\mathbb{C}^2}^2 + \left\| \frac{\partial^2 \widehat{U}^\varepsilon}{\partial \theta^2}(z, \theta) \right\|_{\mathbb{C}^2}^2 \right]. \end{aligned}$$

Then, to get (A.5), it is sufficient to show (A.6) for $\partial \widehat{U}^\varepsilon / \partial \theta$ and $\partial^2 \widehat{U}^\varepsilon / \partial \theta^2$: we use again the perturbed test function method for $(\widehat{U}^\varepsilon(\cdot, \theta), Y^\varepsilon(\cdot, \theta), Z^\varepsilon(\cdot, \theta))$ with $Y^\varepsilon := \partial \widehat{U}^\varepsilon / \partial \theta$ and $Z^\varepsilon := \partial^2 \widehat{U}^\varepsilon / \partial \theta^2$ for a fixed θ .

Step 2: The Aldous property.

The following criterion insures that $(\widehat{U}^\varepsilon)_\varepsilon$ verifies the Aldous property:

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \sup_{|\theta| \leq A} \sup_{\tau \in \mathcal{T}_M} \sup_{0 < \delta' < \delta} \mathbb{E}_\mu [1 \wedge \|\widehat{U}^\varepsilon(\tau + \delta', \theta) - \widehat{U}^\varepsilon(\tau, \theta)\|_{\mathbb{C}^2}^2] = 0,$$

where \mathcal{T}_M is the set of all stopping times bounded by M . For $k = 1, 2, 3, 4$, we introduce the function f_k from \mathbb{R}^4 to \mathbb{R} defined by $f_k(s_1, s_2, s_3, s_4) = s_k$. We build the perturbed test functions $f_{k,\theta}^\varepsilon$ according to the procedure described in the proof of Theorem 6. The process

$$M_{f_{k,\theta}^\varepsilon}(z) := f_{k,\theta}^\varepsilon(\widehat{U}^\varepsilon(z, \theta), v(z/\varepsilon^2)) - f_{k,\theta}^\varepsilon(\widehat{U}^\varepsilon(0, \theta), v(0)) - \int_0^z \mathcal{L}_\theta^\varepsilon f_{k,\theta}^\varepsilon(\widehat{U}^\varepsilon(w, \theta), v(w/\varepsilon^2)) dw$$

is a martingale with quadratic variation

$$\langle M_{f_{k,\theta}^\varepsilon} \rangle(z) = \int_0^z H_{f_{k,\theta}^\varepsilon}(\widehat{U}^\varepsilon(w, \theta), \tilde{v}(w/\varepsilon^2)) dw,$$

where $H_{f_{k,\theta}^\varepsilon} = \mathcal{L}_\theta^\varepsilon (f_{k,\theta}^\varepsilon)^2 - 2f_{k,\theta}^\varepsilon \mathcal{L}_\theta^\varepsilon f_{k,\theta}^\varepsilon$. Because of (A.8), $M_{f_{k,\theta}^\varepsilon}$ can be written as

$$M_{f_{k,\theta}^\varepsilon}(z) = f_k(\widehat{U}^\varepsilon(z, \theta)) - f_k(\widehat{U}_0(\theta)) - \int_0^z \mathcal{L}_\theta f_k(\widehat{U}^\varepsilon(w, \theta)) dw + \mathcal{O}(\varepsilon),$$

where $\mathcal{O}(\varepsilon)$ is uniform in θ, k , and z in a compact set. Then for all $\tau \in \mathcal{T}_M$ and $\delta' \in (0, \delta)$:

$$\begin{aligned} \left(f_k(\widehat{U}^\varepsilon(\tau + \delta', \theta)) - f_k(\widehat{U}^\varepsilon(\tau, \theta)) \right)^2 &\leq 4 \left(M_{f_{k,\theta}^\varepsilon}(\tau + \delta') - M_{f_{k,\theta}^\varepsilon}(\tau) \right)^2 \\ &\quad + 4 \left(\int_\tau^{\tau + \delta'} \mathcal{L}_\theta f_k(\widehat{U}^\varepsilon(w, \theta)) dw \right)^2 + \mathcal{O}(\varepsilon). \end{aligned} \tag{A.7}$$

Using

$$\mathbb{E}_\mu \left[\left(M_{f_{k,\theta}^\varepsilon}(\tau + \delta') - M_{f_{k,\theta}^\varepsilon}(\tau) \right)^2 \right] = \int_\tau^{\tau + \delta'} \mathbb{E}_\mu [H_{f_{k,\theta}^\varepsilon}(\widehat{U}^\varepsilon(w, \theta), v(w/\varepsilon^2))] dw,$$

taking the expectation in (A.7), summing over k , taking the sup over δ', τ and θ , and letting $\delta \rightarrow 0$ then $\varepsilon \rightarrow 0$, establish the desired result.

A.2. Diffusion-approximation theorem

In this appendix, we state and prove a diffusion-approximation theorem that is applied in this paper. We consider the following problem: we aim at proving the weak convergence of the \mathbb{R}^d -valued process X^ε defined by the system

$$\frac{dX^\varepsilon}{dz}(z) = F(X^\varepsilon(z)) + \frac{1}{\varepsilon}G(X^\varepsilon(z), m^\varepsilon(z))$$

starting from $X^\varepsilon(0) = X_0 \in \mathbb{R}^d$. We denote $m^\varepsilon(z) = m(z/\varepsilon^2)$ and we assume that:

- (i) The Markov process m is bounded a.s., stationary, ergodic with generator \mathcal{M} , satisfying the Fredholm alternative.
- (ii) $F(x)$ and $G(x, m)$ are \mathbb{R}^d -valued functions. $G(x, m)$ satisfies the centering condition: for any x , $\mathbb{E}[G(x, m(0))] = 0$ where \mathbb{E} denotes the expectation with respect to the invariant probability measure of m . Instead of technical sharp conditions, we assume that all partial derivatives in x of F and G are bounded.

Note that nor the result, neither the proof is original [6,25]. However the expression of the perturbed test function that we derive in the proof is necessary for the technical estimates used in the proof of Proposition 5.

Theorem 6. As $\varepsilon \rightarrow 0$, the processes X^ε converge in distribution in $\mathcal{C}([0, Z], \mathbb{R}^d)$ to the diffusion Markov process whose infinitesimal generator is given by

$$\mathcal{L}f(x) = F(x) \cdot \nabla_x f(x) + \int_0^{+\infty} \mathbb{E}[G(x, m(0)) \cdot \nabla_x(G(x, m(z)) \cdot \nabla_x f(x))] dz.$$

Proof. We use the perturbed test function method. The process $(X^\varepsilon, m^\varepsilon)$ is a Markov process with generator \mathcal{L}^ε given by:

$$\mathcal{L}^\varepsilon = F(x) \cdot \nabla_x + \frac{1}{\varepsilon}G(x, m) \cdot \nabla_x + \frac{1}{\varepsilon^2}\mathcal{M}.$$

According to Kushner [25 Chapter III, Theorems 2 and 4], the most important fact to show is: for any bounded \mathcal{C}^∞ function $f(x)$, there exists a family of functions $(f^\varepsilon(x, m))_\varepsilon$ such that for any compact set K

$$\sup_{x \in K, m} |f^\varepsilon(x, m) - f(x)| \xrightarrow{\varepsilon \rightarrow 0} 0 \quad \text{and} \quad \sup_{x \in K, m} |\mathcal{L}^\varepsilon f^\varepsilon(x, m) - \mathcal{L}f(x)| \xrightarrow{\varepsilon \rightarrow 0} 0. \tag{A.8}$$

Let f be a bounded, \mathcal{C}^∞ function. Define

$$f^\varepsilon(x, m) = f(x) + \varepsilon f_1(x, m) + \varepsilon^2 f_2(x, m),$$

where f_1 and f_2 will be chosen so that (A.8) will be satisfied. Applying \mathcal{L}^ε to f^ε , one gets:

$$\begin{aligned} \mathcal{L}^\varepsilon f^\varepsilon &= \frac{1}{\varepsilon}G(x, m) \cdot \nabla_x f + \frac{1}{\varepsilon}\mathcal{M}f_1 + F(x) \cdot \nabla_x f + G(x, m) \cdot \nabla_x f_1 + \mathcal{M}f_2 + \varepsilon F(x) \cdot \nabla_x f_1 + \varepsilon G(x, m) \cdot \nabla_x f_2 \\ &\quad + \varepsilon^2 F(x) \cdot \nabla_x f_2. \end{aligned}$$

Defining $f_1 := -\mathcal{M}^{-1}(G(x, m) \cdot \nabla_x f)$, which is well-defined because of the Fredholm alternative and the centering condition, the $\mathcal{O}(1/\varepsilon)$ -term in \mathcal{L}^ε becomes 0. An explicit expression of f_1 is

$$f_1(x, m) = \int_0^{+\infty} \mathbb{E}[G(x, m(z)) \cdot \nabla_x f(x) | m(0) = m] dz.$$

Our aim is now to get an expression with a $\mathcal{O}(1)$ -term independent of m . Therefore we define:

$$f_2 := -\mathcal{M}^{-1}(G(x, m) \cdot \nabla_x f_1 - \mathbb{E}[G(x, m(0)) \cdot \nabla_x f_1(m(0))]).$$

It thus remains

$$\mathbb{E}[G(x, m(0)) \cdot \nabla_x f_1(m(0))] = \int_0^{+\infty} \mathbb{E}[G(x, m(0)) \cdot \nabla_x (G(x, m(z)) \cdot \nabla_x f(x))] dz,$$

which proves (A.8). \square

References

- [1] G.P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed., Academic Press, San Diego, 2001.
- [2] P.K.A. Wai, C.R. Menyuk, Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly birefringence, *J. Lightwave Technol.* 14 (1996) 148–157.
- [3] P.K.A. Wai, C.R. Menyuk, Polarization decorrelation in optical fibers with randomly varying birefringence, *Opt. Lett.* 19 (1994) 1517–1519.
- [4] S.C. Rashleigh, Origins and control of polarization effects in single-mode fibers, *J. Lightwave Technol.* 1 (1983) 312–331.
- [5] C.R. Menyuk, P.K.A. Wai, Polarization evolution and dispersion in fibers with spatially varying birefringence, *J. Opt. Soc. Am. B* 11 (1994) 1288–1296.
- [6] G. Papanicolaou, D.W. Stroock, S.R.S. Varadhan, Martingale approach to some limit theorem, in: D. Ruelle (Ed.), *Statistical Mechanics and Dynamical systems. Duke Turbulence Conf., Duke Univ. Math. Series III, Part IV, 1976*, pp. 1–120.
- [7] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge university press, Cambridge, 1990.
- [8] L.C.G. Rogers, D. Williams, *Diffusions, Markov processes and martingales*, Wiley, Chichester, 1987.
- [9] A. Galtarossa, L. Palmieri, M. Schiano, T. Tambosso, Statistical characterization of fiber random birefringence, *Opt. Lett.* 25 (2000) 1322–1324.
- [10] A. Simon, R. Ulrich, Evolution of polarization along a single-mode fiber, *Appl. Phys. Lett.* 31 (1977) 517–520.
- [11] M.C. de Lignie, H.G.J. Nagel, M.O. van Deventer, Large polarization mode dispersion in fiber optic cables, *J. Lightwave Technol.* 12 (1994) 1325–1329.
- [12] A. Galtarossa, G. Gianello, C.G. Someda, M. Schiano, In-field comparison among polarization mode dispersion measurement techniques, *J. Lightwave Technol.* 14 (1996) 42–49.
- [13] M. Karlsson, Polarization mode dispersion-induced pulse broadening in optical fibers, *Opt. Lett.* 23 (1998) 688–690.
- [14] J. Garnier, J. Fatome, G. Le Meur, Statistical analysis of pulse propagation driven by polarization-mode dispersion, *J. Opt. Soc. Am. B* 19 (2002) 1968–1977.
- [15] S.V. Manakov, On the theory of two-dimensional stationary self focussing of electromagnetic waves. *Zh. Eksp. Teor. Fiz.* 65 (1973) 505–516 [*Sov. Phys. JETP* 38 (1974) 248].
- [16] M. Matsumoto, Y. Akagi, A. Hasegawa, Propagation of solitons in fibers with randomly varying birefringence: effects of soliton transmission control, *J. Lightwave Technol.* 15 (1997) 584–589.
- [17] C. Xie, M. Karlsson, P.A. Andrekson, Soliton robustness to the Polarization-Mode Dispersion in optical fibers, *IEEE Photon. Technol. Lett.* 12 (2000) 801–803.
- [18] A. Hasegawa, Effect of polarization mode dispersion in optical soliton transmission in fibers, *Physica D* 188 (2004) 241–246.
- [19] L.F. Mollenauer, K. Smith, J.P. Gordon, C.R. Menyuk, Resistance of solitons to the effects of polarization dispersion, *Opt. Lett.* 14 (1989) 1219–1221.
- [20] T.I. Lakoba, D.J. Kaup, Perturbation theory for the Manakov soliton and its applications to pulse propagation in randomly birefringent fibers, *Phys. Rev. E* 64 (1997) 6147–6165.
- [21] S.M. Baker, J.N. Elgin, H.J. Harvey, Soliton jitter in birefringent fiber, *Opt. Commun.* 165 (1999) 27–32.
- [22] D.T. Gillespie, Exact numerical simulation of the Ornstein–Uhlenbeck process and its integral, *Phys. Rev. E* 54 (1996) 20842091.
- [23] H. Amann, *Ordinary differential equations, an introduction to nonlinear analysis*, Walter de Gruyter, Berlin, 1990.
- [24] M. Métivier, Convergence faible et principe d’invariance pour des martingales à valeurs dans des espaces de Sobolev, internal report CMAP-Ecole Polytechnique 106 (1984).
- [25] H.J. Kushner, *Approximation and weak convergence methods for random processes*, MIT Press, Cambridge, 1984.