

PMD-Induced Fluctuations of Bit-Error Rate in Optical Fiber Systems

Vladimir Chernyak, Michael Chertkov, Ildar Gabitov, Igor Kolokolov, and Vladimir Lebedev

Invited Paper

Abstract—This paper presents a method that allows evaluating the performance of an optical fiber system where bit errors result from a complex interplay of spontaneous noise generated in optical amplifiers and birefringent disorder of the transmission fiber. We demonstrate that in the presence of temporal fluctuations of birefringence characteristics, the bit-error rate (BER) itself is insufficient for characterizing system performance. Adequate characterization requires introducing the probability distribution function (PDF) of the BER obtained by averaging over many realizations of birefringent disorder. Our theoretical analysis shows that this PDF has an extended tail indicating the importance of anomalously large values of BER. We present the results of comprehensive analysis of the following issues: 1) The dependence of the PDF tail shape on detection details, such as filtering and regular temporal shift adjustment; 2) the changes in the PDF of BER that occur when the first- or higher order polarization mode dispersion (PMD) compensation techniques are applied; 3) an alternative PMD compensation method capable of providing more efficient suppression of extreme outages.

Index Terms—Bit-error rate (BER), optical fiber telecommunication systems, polarization mode dispersion (PMD) compensation, probability distribution function (PDF).

I. INTRODUCTION

POLARIZATION mode dispersion (PMD) constitutes one of the main limiting factors for reliable optical fiber system performance at transmission rates of 40 Gb/s or higher. PMD causes broadening of initially compact pulses in a data stream that eventually leads to bit-pattern corruption [1]–[4]. This effect can be characterized in terms of the PMD vector [5]–[8], [37]. It has been also recognized that the PMD vector does

not provide a complete description of the PMD phenomenon and some more sophisticated approaches that take into account “higher order” PMD effects, have been recently discussed in the literature [9]–[12].

It is well known from experiment that birefringence in optical fiber systems is slowly, but substantially changing with time under the influence of fluctuations in environmental conditions (stresses, temperature, etc.), see, e.g., [13], [14]. Thus, dynamical PMD compensation becomes a major issue in modern high-speed optical-fiber telecommunication technology [15], [16].

Development of new techniques capable of first- [15], [19], [20] and higher order [20], [21] PMD compensation has raised a question of how to evaluate the compensation success (or failure). Traditionally, the statistics of the PMD vectors of first [5], [6], [37], [8] and higher orders [9]–[11] are considered as a measure for any particular compensation method’s performance. However, these objects are only indirectly related to fiber system reliability.

In this paper, that develops the ideas briefly described in a series of recent publications [22]–[25], we clearly demonstrate that PMD effects should be considered together with impairments due to amplifier (and other types of) noise. Indeed, the system performance for a given realization of birefringent disorder is characterized by a certain value of BER, i.e., probability of detecting an error, which is nonzero because of the noise. However, the value of the bit-error rate (BER) is varying together with the temporal variations of the birefringent disorder. The characteristic time scale of such variations is much longer than the times related to signal transmission, however, it can substantially exceed the overall system operation time. Therefore, evaluating system performance should be based on the analysis of fluctuations in the value of BER. We show that fluctuations of BER caused by variations of the birefringent disorder are substantial. Large fluctuations of BER originate from the very different nature (temporal correlations) of the amplified spontaneous emission (ASE) noise compared to that of the birefringent disorder. Birefringent disorder is practically frozen (i.e., it does not vary at least on the time scales related to the optical signal propagation). Optical noise originating from ASE constitutes an impairment of a different nature: the amplifier noise is short correlated on the time scale of the signal width.

We demonstrate that the probability of extreme outages (i.e., such situations or, stated differently, realizations of birefringent disorder when the BER substantially exceeds its typical value)

Manuscript received June 1, 2003; revised September 12, 2003. This work was supported by LDRD ER on “Statistical Physics of Fiber Optics Communications” at Los Alamos National Laboratory and by a personal grant as well as under Grant 03-02-16 147a from the Russian Foundation for Basic Research (IK).

V. Chernyak is with Corning Inc., SP-DV-02-8, Corning, NY 14831 USA (e-mail: ChernyakV@corning.com).

M. Chertkov is with the Theoretical Division, Los Alamos National Laboratory (LANL), Los Alamos, NM 87545 USA.

I. Gabitov is with the Theoretical Division, Los Alamos National Laboratory (LANL), Los Alamos, NM 87545 USA, and with the Department of Mathematics, University of Arizona, Tucson, AZ 85721 USA, and with the Landau Institute for Theoretical Physics, Moscow, 117334, Russia.

I. Kolokolov and V. Lebedev are with the Theoretical Division, Los Alamos National Laboratory (LANL), Los Alamos, NM 87545 USA, and with the Landau Institute for Theoretical Physics, Moscow, 117334, Russia.

Digital Object Identifier 10.1109/JLT.2004.825237

is much larger than one could expect from naive estimates based on singling out effects of either of these two impairments. This phenomenon is a consequence of a complex interplay between these two different impairments. It may not be rationalized in terms of just an average value of BER, or statistics of any PMD vectors of different orders. Complete description of this phenomena requires studying the probability distribution function (PDF) of BER, and specifically its tail.

The emergence of an extremely extended (algebraic or algebraic-like) tail in the PDF of BER constitutes the major result of the paper. The result is general, i.e. it applies to a whole variety of transmission regimes (linear, nonlinear, and quasi-linear), various signal modulation formats [return-to-zero (RZ), nonreturn-to-zero (NRZ), differential phase-shift keying (DPSK)-RZ, differential quadrature phase-shift keying (DQPSK)-RZ, etc.], and detection techniques (optical and electrical filtering, decision threshold choice, etc.). However, for the sake of simplicity and clarity and also to be specific, we consider the following situation of major practical interest in optical fiber communications: 1) the modulation format is RZ (on-off keying); 2) transmission is linear, i.e. Kerr, Raman, and other nonlinear terms are not taken into account. Besides (and less importantly), the other two model assumptions (which do not restrict the generality of the model independent results reported in the paper) are that: a) the electrical filter (window) is represented by a symmetric step function; b) the optical filter is of Lorentzian shape.

After brief technical introduction into the problem given in Section II, our theoretical analysis starts in Section III with evaluating the signal BER due to the amplifier noise for a given realization of birefringent disorder. We next study the PDF of BER, where the statistics are collected over different fibers or, equivalently, over the birefringence states of a given fiber at different times. At the second step, we focus on evaluating the probability of anomalously large BER. This general scheme will be applied to four situations of interest. We start with the basic (no compensation) case in Section IV-A and compare it with the case of the simplest compensation scheme known as “setting the clock” in Section IV-B, and also with the cases of first- and higher order PMD compensations schemes in Sections V-A and V-B, respectively. Finally, in Section V-C, we discuss a compensation scheme referred to as (quasi)-periodic that appears to be more efficient in reducing the extreme outages compared to the traditional high-order compensation scheme with the same number of compensating degrees of freedom. Section VI is reserved for discussions and conclusions.

II. TECHNICAL INTRODUCTION

In this section, we introduce the basic relations and terminology that describe data transmission (signal propagation) in an optical fiber system. The goal here is to set the problem in formal terms, introduce the objects of interest, and also make some preliminary evaluations.

A. Amplifier Noise and Birefringent Disorder

The envelope of the optical field propagating in a given channel in the linear regime (i.e., at relatively low optical

power), which is subject to PMD distortion and amplifier noise, satisfies the following equation [26]–[28]:

$$\partial_z \Psi - i\hat{\Delta}(z)\Psi - \hat{m}(z)\partial_t \Psi - id(z)\partial_t^2 \Psi = -\gamma(z)\Psi + g(z)\Psi + \xi(z, t). \quad (1)$$

Here z , t , ξ , as well as d , γ , and g are the position along the fiber, retarded time (i.e., time associated with the reference frame moving with group velocity of a chosen frequency channel), the amplifier noise, and the chromatic dispersion, attenuation, and gain coefficients, respectively. (We assume that neither gain nor damping are polarization dependent, leaving the more general problem for future investigation.) The envelope Ψ is a two-component complex field, the two components represent two states of the optical signal polarization. Our approach allows us to treat discrete (erbium) and distributed (Raman) amplification schemes within the same framework. The birefringent disorder is characterized by two random 2×2 traceless matrix fields related to the zero-, $\hat{\Delta}$, and first-, \hat{m} , orders in the frequency expansion with respect to the deviation from the channel carrier frequency ω_0 . Birefringence that affects the light polarization is practically frozen (t -independent) on all propagation-related time scales.

The matrix $\hat{\Delta}$ as well as the attenuation and gain coefficients γ and g can be excluded from consideration by the following transformation $\Psi = A\hat{V}\tilde{\Psi}$, $\xi = A\hat{V}\tilde{\xi}$ and $\hat{m} = \hat{V}\tilde{m}\hat{V}^{-1}$. Here

$$A = \exp \left\{ \int_0^z dz' [g(z') - \gamma(z')] \right\}$$

is a z -dependent number, and the unitary matrix

$$\hat{V}(z) = T \exp \left[i \int_0^z dz' \hat{\Delta}(z') \right]$$

is the ordered exponential defined as a formal solution of the equation $\partial_z \hat{V} = i\hat{\Delta}\hat{V}$ with $\hat{V}(0) = \hat{I}$. We assume that the gain coefficient is properly chosen to perfectly compensate for damping, so that $A(Z) = 1$, where Z is the total system length. The renormalized quantity $\tilde{\Psi}$ satisfies the equation

$$\partial_z \tilde{\Psi} - \tilde{m}(z)\partial_t \tilde{\Psi} - id(z)\partial_t^2 \tilde{\Psi} = \tilde{\xi}(z, t). \quad (2)$$

The solution of (2) can be represented as

$$\tilde{\Psi} \equiv \varphi + \phi$$

$$\varphi = \hat{W}(z)\Psi_0(t) \quad (3)$$

$$\phi = \int_0^z dz' \hat{W}(z)\hat{W}^{-1}(z')\tilde{\xi}(z', t) \quad (4)$$

$$\hat{W}(z) = \exp \left[i \int_0^z dz' d(z')\partial_t^2 \right] \hat{U}(z) \quad (5)$$

$$\hat{U}(z) = T \exp \left[\int_0^z dz' \hat{m}(z')\partial_t \right] \quad (6)$$

where $\Psi_0(t)$ stands for the initial pulse shape and $T \exp$ denotes the so-called ordered exponential operator

$$T \exp \left[\int_0^z dz' \hat{a}(z') \right] \equiv \hat{V}(z)$$

that can be formally described as the solution of the operator equation $\partial_z \hat{V}(z) = \hat{a}(z)\hat{V}(z)$ with the initial condition $\hat{V}(0) = \hat{I}$ (note that when the operators $a(z)$ commute for all z the ordered exponential coincides with the usual one). Solving the operator

equation iteratively leads to a very useful representation of the ordered exponential in a form of a functional infinite series

$$\begin{aligned} T \exp \left[\int_0^z dz' \hat{a}(z') \right] \\ = \sum_{n=0}^{\infty} \int_0^z dz_1 \int_0^{z_1} dz_2 \cdots \int_0^{z_{n-1}} dz_n \hat{a}(z_1) \hat{a}(z_2) \cdots \hat{a}(z_n) \end{aligned}$$

that will be used in this manuscript for performing perturbative computations.

We consider a situation when the pulse propagation distance substantially exceeds the inter-amplifier separation (the system consists of a large number of spans). The additive noise ξ , generated by optical amplifiers, is zero on average. The statistics of ξ are Gaussian with spectral properties determined solely by the steady-state features of amplifiers (gain and noise figure) [30]. The noise correlation time is much shorter than the pulse temporal width, and, therefore, ξ can be treated as δ -correlated in time. Equations (4) and (5) imply that the noise contribution to the output signal ϕ is a zero mean Gaussian field characterized by the following pair correlation function:

$$\langle \phi_\alpha(Z, t_1) \phi_\beta^*(Z, t_2) \rangle = D_\xi Z \delta_{\alpha\beta} \delta(t_1 - t_2) \quad (7)$$

with the product $D_\xi Z$ being the ASE spectral density accumulated along the system. The coefficient D_ξ is introduced into (7) to reveal the linear growth of the ASE factor with the total line length Z [30]. Provided the noise ξ is short-correlated in space (that is, correct for both erbium and Raman amplifiers), the factor D_ξ in (7) is statistically independent of both $d(z)$ and $\hat{m}(z)$, as immediately follows from (4)–(6).

The matrix of birefringence \hat{m} can be parameterized using a three-component real field h_j , $\hat{m} = \sum h_j \hat{\sigma}_j$, with $\hat{\sigma}_j$ being a set of three Pauli matrices. The field \mathbf{h} is zero on average and short-correlated in z . Therefore, in accordance with the central limit theorem (see, e.g., [31]) the integral $\mathbf{H} = \int_0^Z dz \mathbf{h}(z)$ has Gaussian statistics (with zero average) characterized by the pair correlation function

$$\langle H_i H_j \rangle = D_m Z \delta_{ij} \quad (8)$$

where the average is taken over the birefringent disorder realizations (corresponding to different fibers or, equivalently, states of birefringence in a single fiber at different times). The isotropy of the pair correlation function (8) is guaranteed by the above transformation $\hat{m} = \hat{V} \hat{m} \hat{V}^{-1}$ since the presence of \hat{V} leads to fast rotations of the vector \mathbf{h} along z . In the case of weak birefringent disorder \mathbf{H} represents the PMD vector. Thus, $D_m = k^2/12$, with k being the so-called PMD coefficient that is usually measured in units of picoseconds per square root of a kilometer (ps/ $\sqrt{\text{km}}$) and has the following meaning. In a system of length Z short enough so that effects of PMD are typically weak, $k\sqrt{Z}$ represents a typical time splitting between the two principle polarization components of a pulse accumulated along the system. The factor of 12 is obtained in the following way: $k\sqrt{Z}$ is twice the typical value of the differential group delay (DGD) vector resulting in $k^2 Z$ to be four times the typical (defined as the average) value of its square, the latter being naturally given by $3D_m Z$. As we will see later, some observables contain the field

\mathbf{h} in a more sophisticated form than just the integral \mathbf{H} . Statistical properties of these more sophisticated objects can be established by using the relation

$$\langle h_i(z_1) h_j(z_2) \rangle = D_m \delta_{ij} \delta(z_1 - z_2) \quad (9)$$

instead of (8).

B. BER as a Functional of Birefringent Disorder

We consider the RZ modulation format when the pulses are well separated in time. The signal detection at the line output, $z = Z$, corresponds to measuring the output pulse intensity I

$$I = \int dt G(t) |\mathcal{K}\Psi(Z, t)|^2 \quad (10)$$

where $G(t)$ is a convolution of the electrical (current) filter function with the sampling window function. The linear operator \mathcal{K} in (12) stands for an optical filter and a variety of engineering “tricks” applied to the output signal $\Psi(Z, t)$. These tricks consist of the optical filter \mathcal{K}_f , and the compensation \mathcal{K}_c parts, respectively, assuming the compensation is applied first followed by filtering, i.e.,

$$\mathcal{K} = \mathcal{K}_f \mathcal{K}_c. \quad (11)$$

We can replace Ψ by $\tilde{\Psi}$ in (10) since $A(Z) = 1$ and \hat{V} is a unitary matrix. Upon substituting the representation (3) into (10) we obtain

$$I = \int dt G(t) |\mathcal{K}\varphi(Z, t) + \mathcal{K}\phi(Z, t)|^2. \quad (12)$$

Compensating options, coded in specific form(s) of the operator \mathcal{K}_c , are discussed in Sections II-C and II-D. Filtering options, formalized by specific choices of the function $G(t)$ and operator \mathcal{K}_f are described in Appendix I, where we also discuss the specific form of the initial pulse Ψ_0 used for the modeling analysis.

Ideally, I takes two distinct values corresponding to the bits “0” and “1,” respectively. However, the impairments enforce deviations of I from the ideal values. The output signal (bit of information) is identified by introducing a threshold (decision level) I_d and declaring that the signal encodes “1” if $I > I_d$ and “0” otherwise. Sometimes, the information is lost, i.e., an initial “1” is detected as “0” at the output or *vice versa*. BER is the probability to detect a false event measured by counting many pulses coming through a fiber with a given realization of the PMD (birefringent) disorder $h_j(z)$. For successful system performance BER should be extremely small, i.e., typically both impairments (noise and disorder driven) can cause only a small distortion of a pulse or, stated differently, the optical signal-to-noise ratio (OSNR) and the ratio of the squared pulsewidth to the mean squared value of the PMD vector are both large. OSNR can be estimated as $I_0/(D_\xi Z)$ where, according to (10)

$$I_0 = \int dt G(t) |\mathcal{K}_f \Psi_0(t)|^2$$

is the intensity of the unperturbed signal, Ψ_0 being the input signal normalized to one. Therefore, the two small parameters of our theory are represented by

$$\frac{D_\xi Z}{I_0} \ll 1 \quad (13)$$

$$\frac{D_m Z}{b^2} \ll 1 \quad (14)$$

b being the pulsewidth, and the condition (14) is assumed to hold for all cases considered in the paper except for the one discussed in Section V-C.¹

We distinguish events associated with the $1 \rightarrow 0$ transition (loss of the signal), and with the $0 \rightarrow 1$ transition (false pulse detection) and designate the corresponding probabilities as $B_{1 \rightarrow 0}$ and $B_{0 \rightarrow 1}$. These two objects are defined as

$$B_{1 \rightarrow 0} = \int_0^{I_d} dI \mathcal{P}_1(I), \quad B_{0 \rightarrow 1} = \int_{I_d}^{\infty} dI \mathcal{P}_0(I) \quad (15)$$

and according to (12), the PDF \mathcal{P} of the output signal intensity which can be written as

$$\begin{aligned} \mathcal{P}_{0;1} &= \left\langle \delta \left(I - \int dt G(t) |\mathcal{K}_f \varphi(Z, t) + \mathcal{K}_f \phi(Z, t)|^2 \right) \right\rangle_{\phi} \\ \langle A\{\phi\} \rangle_{\phi} &\equiv \int \mathcal{D}\phi A\{\phi\} \bar{\mathcal{P}}\{\phi\} \\ \bar{\mathcal{P}}\{\phi\} &= \bar{N}^{-1} \exp \left(-\frac{1}{D_{\xi} Z} \int_{-\infty}^{\infty} dt |\phi(t)|^2 \right) \end{aligned} \quad (16)$$

where $\bar{\mathcal{P}}$ is the PDF of ϕ , N is normalization constant, and $A\{\phi\}$ is an arbitrary functional of ϕ . Also the subscript 0 in the definition of $\mathcal{P}_{0;1}$ corresponds to the “zero” input bit $\varphi = 0$, while the subscript 1 corresponds to the “one” input bit $\varphi = \Psi_0$. In (16), averaging is performed over the noise statistics. Experimentally, such an average is measured by collecting the statistics over many pulses propagating along a fiber with the same birefringent disorder realization $\mathbf{h}(z)$ since different pulses experience different realizations of the noise the latter being stochastic in nature. Formally, this constitutes averaging an observable that can be represented by any functional $A\{\phi\}$ over all possible realizations of the noise ϕ with the probability distribution $\bar{\mathcal{P}}\{\phi\}$, whereas a specific form of $\bar{\mathcal{P}}\{\phi\}$ also given in (16) corresponds to the situation under consideration, i.e., Gaussian statistics of the noise fully determined by the pair correlation function of (7). Since realizations of noise are represented by functions $\phi(t)$ of time that represent the field $\phi(z, t)$ at $z = Z$. (16) constitutes a path (functional)-integral representation for $\mathcal{P}_{0;1}$ (see, e.g., [32], [33] for an introduction to path-integral techniques). This, however, does not constitute a major conceptual problem since path integrals can (and strictly speaking should) be considered as finite-dimensional integrals where functions are represented by sets of values at a large but finite set of points t_0, \dots, t_n .

An important difference between $B_{1 \rightarrow 0}$ and $B_{0 \rightarrow 1}$, defined by (15), is in the strong dependence of the first case and independence of the second one on the \mathbf{h} realization (i.e., the birefringence profile along the fiber). This difference stems from the fact that $\varphi = 0$ in the case of $B_{0 \rightarrow 1}$ whereas $B_{1 \rightarrow 0}$ is a non-trivial functional of \mathbf{h} , as well as from statistical independence of ϕ and \mathbf{h} . One concludes that even though $B_{1 \rightarrow 0}$ and $B_{0 \rightarrow 1}$ are of the same order in the absence of birefringence, anomalously large values of BER (which is the focus of this paper) originate solely from the “ $1 \rightarrow 0$ ” transitions. Therefore, in what follows

¹Section V-C is devoted to a special case where without compensation, the condition (14) is essentially violated, however, a weaker condition, $D_m Z / b^2 \ll 1/N$ still holds. It is shown then that by using a new quasi-periodic compensation strategy of N th order one can still get an operable system, i.e., a system with typical BER essentially smaller than one.

we concentrate primarily on the analysis of $B_{1 \rightarrow 0}$ thus dropping the $1 \rightarrow 0$ subscript ($B_{1 \rightarrow 0} = B$) to simplify notation.

The PDF $\mathcal{P}_1(I)$ has a maximum near $I = I_0$ and decays quickly as I departs from I_0 . The tail of the transition probability at $|I - I_0| \gg D_{\xi} Z$ is exponential (see Appendix II for details). This implies that the integration in the right-hand side (RHS) of (15) is actually concentrated near $I = I_d$ thus yielding the following estimate for BER, B :

$$\ln B \approx \ln \mathcal{P}_1(I_d). \quad (17)$$

C. “Setting the Clock” Compensation

An essential part of the signal loss can be compensated using a simple procedure, in the fiber-optics jargon usually called “setting the clock.” This procedure accounts for adjusting the overall time shift which is a functional of the birefringent disorder. (We are not discussing here an important engineering problem of how to make this dynamical adjustment, simply assuming that a device capable of doing this operation does exist.) Formally, the “setting the clock” procedure can be described by the following modification of (12):

$$I = \int dt G(t + t_{cl}) |\mathcal{K}_f \varphi(Z, t) + \mathcal{K}_f \phi(Z, t)|^2 \quad (18)$$

or returning to the notation of (11), (18) corresponds to the following form of the compensation operator: $K_c = \exp(-t_{cl} \partial_t)$. As discussed in Section IV-B, the one-parameter flexibility one gains through t_{cl} can be used to minimize system outage. The important question to be addressed is: What is dependence of the “optimal” shift on the birefringent disorder?

D. PMD Compensator

Effects of PMD can be reduced by using a device usually called a PMD compensator (PMDC). Any optical PMDC consists of two parts: a compensating (optical) part responsible for the compensation itself, and a measuring part that extracts (measures) relevant information on the transmission fiber birefringence. We start by considering the optical part of the compensator that usually consists of a set of relatively short elements. Each element includes a piece of polarization-maintaining fiber (this is a fiber characterized by uniform, i.e., position-independent, birefringence vector) usually surrounded by two polarization controllers that allow rotation of the polarization state [15]. This implies that the optical part of a PMDC (hereafter referred to as a PMDC itself when it does not lead to confusion) is characterized by its transfer function that can be parameterized by a finite number of parameters (degrees of freedom). Additionally, one would naturally distinguish between i) describing a compensator in terms of available transfer functions (the subject of this subsection), and ii) compensating strategy, i.e., a prescription of how to fix the compensating degrees of freedom based on the measured data. The compensating strategy part of the problem is discussed in Sections V-A, V-B, and V-C.

The so-called, first-order PMDC corresponds to $\mathcal{K}_c = \mathcal{K}_1$

$$\mathcal{K}_1(\mathbf{M}) = \exp(-M_j \hat{\sigma}_j \partial_t) \quad (19)$$

with $j = 1, 2, 3$. Such a form of the compensating operator \mathcal{K}_c offers richer adjustment options compared to the “setting the

clock” compensation as it actually contains three compensating degrees of freedom, i.e., the three components of the compensating vector \mathbf{M} , instead of one.² Note also, that the transfer matrix \hat{U} of the transmission fiber is defined as an ordered exponential (6), whereas the compensating operator \mathcal{K}_1 is defined in terms of the usual exponential (19). This important difference stems from the fact that the birefringence profile along the transmission fiber $\mathbf{h}(z)$ is a random function of z , while the birefringence of the compensating part is flat, as it is accurately controlled to be z' -independent, with z' being the position marker along the polarization maintaining piece.

A compensation strategy that allows for more compensating options (more degrees of freedom) is potentially better. Thus, a compensator, hereafter referred to as an N th-order PMD compensator consists of N concatenated PMD compensators of the first order [15]. Each of the N compensators is characterized by its own three-component compensating vector \mathbf{M}_n , where $n = 1, \dots, N$, so that the compensating operator generalizes that of (19)

$$\mathcal{K}_c = \mathcal{K}_1(\mathbf{M}_1)\mathcal{K}_1(\mathbf{M}_2)\cdots\mathcal{K}_1(\mathbf{M}_N) \quad (20)$$

and the set of \mathbf{M}_n vectors introduces $3N$ compensating degrees of freedom that are at our disposal for outage optimization.

Once the set of compensation options, described by (19), (20) is fixed, the next task, addressed in Sections V-A, V-B, and V-C is about how to use the compensating degrees of freedom offered by the compensators (19) or (20) to minimize the effects of the system outages. Or rephrasing the question in more formal terms: What are the optimal values of the $3N$ compensating degrees of freedom \mathbf{M}_n that correspond to a given realization $\mathbf{h}(z)$ of the transmission fiber birefringence profile?

III. AMPLIFIER NOISE AVERAGING

In this short section, we present only the basic results, while all derivations can be found in the Appendixes.

Since the OSNR is large, the expression for the transition probability (16) allows for an asymptotic saddle-point evaluation. The details of an analytical calculation, resulting in an algebraic system of equations that implicitly relate the saddle-point value of the transition probability to the inhomogeneous part of the measured signal are given in Appendix II.

²A three-parameter compensator with the transfer function given by (19) can be implemented by surrounding a polarization-sensitive delay line with two polarization controllers (PC). The transfer function of such a device has the form $\mathcal{K}_c = U_c^\dagger \mathcal{K}_d(t_d) U_c$ where U_c^\dagger and U_c are frequency independent of the PC located after and before the delay line, respectively, and $\mathcal{K}_d(t_d) = \exp(-t_d \hat{\sigma}_3 \partial_t)$ is the delay-line transfer function with t_d being the relative time delay. If the time delay t_d is controlled, such a compensator provides the transfer function of (19) since with a proper choice of U_c provided by the PC we can obtain any vector \mathbf{M} with $|\mathbf{M}| = t_d$. A possible implementation of a delay line with adjustable t_d involves a free-space optical-mechanical device that achieves a relative time delay by varying the relative value of the optical paths for two polarization states. Such a device naturally offers three-parametric compensation, as it is capable of dynamically generating any value of the three-component vector \mathbf{M} . A less expensive and maybe more practical option is to use a piece of polarization-maintaining fiber instead of a complicated free-space optical-mechanical device. In this case, only two parameters (the components of \mathbf{M}) are dynamically adjusted with the value of $|\mathbf{M}| = t_d$ being determined by the polarization-maintaining fiber length. The dynamical adjustment is only two-parametric, as it originates from changes in the polarization controllers orientation, while the birefringence of the polarization-maintaining fiber is fixed.

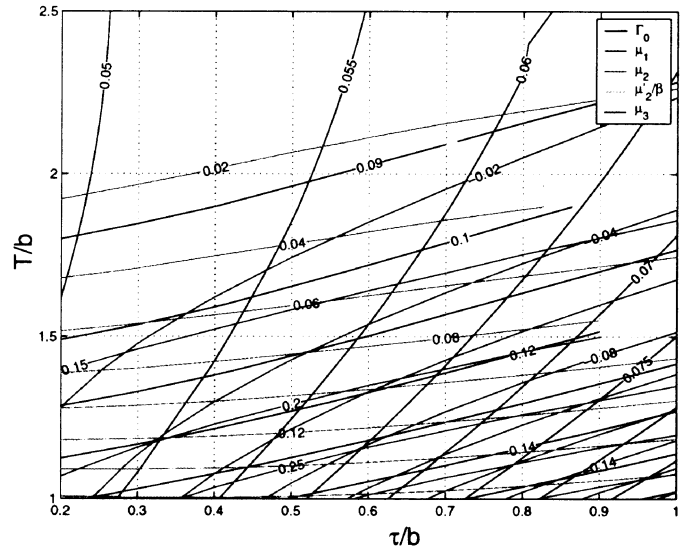


Fig. 1. Dependence of the dimensionless coefficients $\Gamma_0 = -D_\xi Z \ln B_0/I_0$, μ_1 , μ_2 , μ'_2/β , and μ_3 , entering (21), (26), (28), (33), and (34), on the electric filter width T , and the optical filter width τ (both measured in units of the pulsewidth b) for the model introduced in Appendix I. Details of calculations resulted in the dependencies shown in the figure are explained in Appendix II.

The bottom line of these calculations, accounting for averaging with respect to the stochastic noise (i.e., many pulses) in (15), (16) is the saddle-point [i.e. asymptotic, applied whenever the condition (13) holds] expression for the loss probability that is obtained (see Appendixes I and II for the details) by first representing the δ -function in (16) as a Fourier transform of an imaginary exponent [see, e.g., (46)] which yields \mathcal{P}_1 as an integral over $d\lambda D\phi$, with the integrand in a form $\exp[-S_{\text{eff}}(\lambda, \phi, \mathbf{h})]$. Note that S_{eff} depends parametrically on \mathbf{h} through φ [see (3), (6), and (16)]. Evaluating the path integral (16) using the saddle-point approximation (see, e.g., [32], [33], for the general description of the method) and making use of (17) we obtain with the exponential accuracy

$$B = B_0 \exp\left(\frac{\Gamma\{\mathbf{h}\}}{D_\xi Z}\right), \quad B_0 \equiv \exp\left(-\frac{\Gamma_0}{I_0}\right) \quad (21)$$

where $(\Gamma_0 - \Gamma\{\mathbf{h}\})I_0/(D_\xi Z)$ is the saddle-point value of the action S_{eff} . In particular, B_0 corresponds to zero PMD, $\mathbf{h} = 0$, value of B and Γ is a dimensionless quantity with a smooth dependence on \mathbf{h} . By definition, Γ tends to 0 with $\mathbf{h} \rightarrow 0$. The quantity Γ_0 , that determines a typical value of BER, is a dimensionless parameter of order unity. The dependence of Γ_0 on the electric filter width T and the optical filter width τ calculated numerically for the model introduced in Appendix I is displayed in Fig. 1.

The dependence of Γ on the birefringence profile $\mathbf{h}(z)$ is the key subject of the analysis presented in Sections IV and V.

IV. PDF OF BER AND EXTREME OUTAGES

This section constitutes the core of the paper. The bottom line here is that fluctuations of BER from one realization of birefringence to another are strong. To demonstrate that, we study the extended (toward larger values of BER, $B \gg B_0$) tail of the PDF (histogram) of BER, B .

The exponential form of the BER dependence on Γ (21) suggests that one can get an essential enhancement of BER at the expense of a moderate change in the integral birefringence vector, when $\Gamma \gg D_\xi Z I_0^{-1}$. Since the OSNR $I_0/[D_\xi Z]$ is large, the condition is consistent with $\Gamma \ll 1$ (i.e., the regime described by $1 \gg \Gamma_0 \gg D_\xi Z/I_0$ is possible). In other words, one expects that in this “perturbative” region, where the ordered exponential (6) can be approximated by the leading terms in its expansion in the series in \mathbf{h} , BER can be substantially enhanced in comparison with its typical $\mathbf{h} = 0$ value, i.e., $B \gg B_0$. This expectation is indeed confirmed, and detailed in a quantitative way, through our consideration of this section, where we discuss the PDF tail corresponding to the “perturbative” region. We have also studied a universal remote tail of the PDF of BER corresponding to huge fluctuations of the disorder when the signal is almost destroyed by the PMD fluctuations. This remote tail is discussed in Appendix IV.

One concludes that the “perturbative” tail of the PDF of BER can be computed by finding the leading terms in the expansion of Γ in \mathbf{h} , i.e., the key question is how Γ scales with \mathbf{h} at small values of \mathbf{h} . We demonstrate later that $\Gamma = O(h^k)$, where the power k depends on the compensation scheme applied.

Our final result is formulated in terms of the PDF of BER defined as

$$S(B) = \left\langle \delta \left[B - B_0 \exp \left(\frac{\Gamma I_0}{D_\xi Z} \right) \right] \right\rangle_{\mathbf{h}}. \quad (22)$$

Averaging in (22) is performed over the statistics of \mathbf{h} (see Appendix III).

In Section IV-A, we consider the bare case (no compensation applied). Note, that the PDF of BER is not only an interesting object to study but also the key object that characterizes the probability of the system outage due to PMD defined as

$$\mathcal{O} = \int_{B_*}^1 dB S(B) \quad (23)$$

where B_* is the tolerance value for BER outage, $B_* \gg B_0$ with B_0 being the zero-disorder value of BER. This explains why suppression of the PDF tail is the prime target of a compensation strategy. It underlies the analysis of “setting the clock” as well as first- and two different higher orders PMD compensation strategies described in Sections IV-B, V-A, V-B, and V-C, respectively. Since $B_* \gg B_0$, the outage is determined by the tail of the PDF that decays fast enough so that

$$\ln \mathcal{O} \sim \ln S(B_*). \quad (24)$$

A. Bare Case

In this subsection, we consider the bare case, assuming that the output signal is real, i.e., that the chirp part of the input signal (45) was adjusted to compensate for the $d(z)$ -induced dispersive part of \hat{W} , defined in (5). Following the procedure explained in detail in Appendix II, we find the main contribution to Γ to be first order in the field \mathbf{h}

$$\Gamma_{\text{bare}} \approx \frac{\mu_1}{b} H_3, \quad H_i \equiv \int_0^Z dz h_i(z) \quad (25)$$

where μ_1 is a dimensionless coefficient of order one that depends on the signal shape and the detection procedure and it is

also assumed that the initial ($z = 0$) polarization of the pulse is (10). For our simple model, the coefficient μ_1 is displayed in Fig. 1 as a function of T/b and τ/b , where T is the electric filter width, τ is the optical filter width, and b is the pulsewidth. (See Appendix I for explicit definitions of T , τ , and b .) Substituting (25), (21) into (22) and averaging over disorder according to (73) one derives

$$\ln S_{\text{bare}}(B) \approx -\frac{D_\xi^2 Z b^2}{2D_m \mu_1^2 I_0^2} \ln^2 \left(\frac{B}{B_0} \right). \quad (26)$$

Note also, that for a given OSNR, the larger Z is, the more extended the tail is.

B. “Setting the Clock” Compensation

The degree of freedom associated with the “setting the clock” transformation (18) allows us to reduce the effect of anomalous fluctuations of birefringence and thus to reduce the extended tail (26) to a certain extent. It is clear from the standpoint of the tail reduction task that the weaker the dependence of Γ on \mathbf{h} is, the better. Therefore, with the single degree of freedom offered by (18) the optimal value of t_{c1} corresponds to cancellation of the first (linear) term in the expansion of Γ in \mathbf{h} . As shown in Appendix II-C, such an optimal value of t_{c1} corresponds to $t_{c1} = H_3$, with the main term in the expansion of Γ in \mathbf{h}

$$\Gamma_{\text{clock}} \approx \frac{\mu_2}{b^2} \mathbf{H}_\perp^2, \quad \mathbf{H}_\perp^2 \equiv H_1^2 + H_2^2 \quad (27)$$

μ_2 being a dimensionless coefficient of order one and, as before, the output signal is assumed to be real. For our simple model, μ_2 is displayed in Fig. 1 as a function of dimensionless electrical T/b and optical τ/b filter widths. Substituting the expressions (27), (21) into (22) and averaging over disorder according to (74) we arrive at

$$S_{\text{clock}}(B) \sim \frac{B_0^\alpha}{B^{1+\alpha}}, \quad \alpha = \frac{D_\xi b^2}{2\mu_2 D_m I_0}. \quad (28)$$

Note, that the “setting the clock” result (28) shows a steeper decay compared to the bare case one (26) [due to the additional small factor $D_\xi Z/I_0$ in (26)] which is a natural consequence of the compensation procedure applied. We reiterate that the outage is determined by (24).

V. PMD COMPENSATION

As already explained in Section II-D, the key ingredient of any PMD compensation strategy is finding the “optimal” relation between the compensating degrees of freedom (e.g., \mathbf{M} in the case of first-order compensation scheme) and the birefringence profile $\mathbf{h}(z)$ in the transmission fiber.

The standard PMD compensation strategy, discussed in the literature, boils down to compensating for as many terms as possible in the expansion of the received signal I given by (12) in the series in \mathbf{h} [9]–[11], since the more terms in the expansion are compensated the smaller is the PMD-induced signal distortion, and the number of compensated terms in the expansion is usually referred to as the compensation order. In view of (3), (6), and (12) this boils down to compensating the terms in the expansion of $\mathcal{K}_c \hat{U} - \hat{1}$ in powers of \mathbf{h} . For example, in the first-order compensation case (19) the choice $\mathbf{M} = \mathbf{H} = \int_0^Z dz' \mathbf{h}(z')$ guarantees that the expansion of $\mathcal{K}_1 \hat{U} - \hat{1}$ in the series in \mathbf{h} starts with the $O(h^2)$ terms.

Even though the standard PMD compensation criterion was not initially designed to minimize the outage probability (23), we will see below (Sections V-A and V-B) that it efficiently reduces the outage in the case when even without compensation typical realizations of disorder still cause reasonably small values of BER (i.e., when the condition (14) is satisfied). This is not surprising, since this is exactly the combined operator $\mathcal{K}_c \hat{U}$ that enters the value of Γ according to (3), (5), (6), and (18). Thus, the weaker the dependence of $\mathcal{K}_c \hat{U}$ on \mathbf{h} is at $h \rightarrow 0$, the weaker the dependence of Γ on \mathbf{h} is, and, therefore, a more substantial reduction of the outage probability is achieved.

However, as argued in Section V-C, the standard N th-order compensation strategy is not the optimal one when N is large enough. The standard PMD compensation is especially bad in the case when the condition (14) fails, however, a weaker condition, with $1/N$ replacing one in the RHS of (14) still holds. Therefore, in Section V-C, we present an alternative N th-order compensation strategy that outperforms the standard N th-order compensation strategy in the case of a relatively large N , and is even capable of restoring transmission in the aforementioned case of a really bad system, when the standard N th-order compensation applied under the same condition would fail.

A. First-Order PMD Compensation

One deduces from (12), (19) that the output intensity depends on the birefringent disorder via the factor $\mathcal{K}_1 \hat{U}$. Replacing \mathbf{M} with its optimal value \mathbf{H} , expanding $\mathcal{K}_1 \hat{U}$ in \mathbf{h} followed by substituting the result into (12), and evaluating B according to analysis of Appendix II leads to

$$\Gamma_1 = \frac{\mu'_2 Y_c}{b^2} + \frac{\mu_3 Y_r}{b^3} + O\left(\frac{H^4}{b^4}\right) \quad (29)$$

$$Y_c \equiv \int_0^Z dz' \int_0^{z'} dz [h_1(z')h_2(z) - h_2(z')h_1(z)] \quad (30)$$

$$Y_r \equiv \int_0^Z dz_1 \int_0^{z_1} dz_2 \int_0^{z_2} dz_3 \left\{ 2h_3(z_1)\mathcal{H}(z_2, z_3) - h_3(z_2)\mathcal{H}(z_1, z_3) - h_3(z_3)\mathcal{H}(z_1, z_2) \right\} \quad (31)$$

$$\mathcal{H}(z_1, z_2) \equiv h_1(z_1)h_1(z_2) + h_2(z_1)h_2(z_2) \quad (32)$$

where in (29) we present the general expression for the first two terms of the expansion of Γ in a series in H/b . In the general position case, the dimensional coefficient μ'_2 is nonzero, the μ_3 term is subleading and can be neglected. If, however, a degeneracy leads to μ'_2 -term cancellation, the μ_3 term becomes the principal contribution.

Thus, for the Gaussian initial pulse shape, described in Appendix I, the case of nonzero μ'_2 and the degenerate case of zero μ'_2 correspond to the complex (uncompensated chirp) and real (compensated chirp) output signal, respectively. (In the first case, the dimensionless coefficient μ'_2 is related to the output signal chirp, produced by initial signal chirp and/or the nonzero integral chromatic dispersion $\eta = \int_0^Z dz d(z)$. Thus, if the initial chirp is small, $\beta_{\text{in}} \ll 1$, the output signal chirp becomes $\beta = \beta_{\text{in}} + \eta$.) The dependence of the dimensionless coefficients μ'_2/β and μ_3 on T/b and τ/b , found numerically for the model

introduced in Appendix I using the saddle-point equations of Appendix II, are displayed in Fig. 1. Substituting (29) into (21), (22) and averaging over disorder according to (85) and (91), respectively, yields the PDF tail in a form

$$\mathcal{S}_{1;c}(B) \sim \frac{B_0^\gamma}{B^{1+\gamma}}, \quad \gamma = \frac{\pi D_\xi b^2}{2|\mu'_2| D_m I_0} \quad (33)$$

$$\ln \mathcal{S}_{1;r}(B) \approx -4.2 \frac{b^2}{D_m Z} \left(\frac{D_\xi Z}{\mu_3 I_0} \ln \frac{B}{B_0} \right)^{\frac{2}{3}} \quad (34)$$

where (33) and (34) correspond to the nonzero μ'_2 and the degenerate zero μ'_2 cases, (33) and (34) are valid when $\ln(B/B_0) \gg |\mu'_2| D_m I_0 / [D_\xi b^2]$ and $D_\xi Z \ln(B/B_0) \gg \mu_3 (D_m Z)^{3/2} I_0 / b^3$, respectively, and the outage is given by (24).

Note, that although (28) and (33), which correspond to the ‘‘setting the clock’’ compensation and first-order compensation (with uncompensated chirp), respectively, look similar the exponents α and γ that determine the tails of the PDF in these two cases are of different nature: they originate from different bilinear combinations of $\mathbf{h}(z)$ [compare (27) with (29) and (30)].

B. Standard High-Order Compensation

The fiber system performance can be improved even further. First, special filtering efforts can enforce the output pulse symmetry under the $t \rightarrow -t$ transformation (theoretically this can be achieved by a proper choice of the optical filter transfer function, however, practical implementation still remains an open question). Then the $O(H^3)$ contribution to Γ will be also canceled and (29) will be replaced by $\Gamma = O(H^4/b^4)$. Second, and more important, instead of first-order compensation \mathcal{K}_1 one can use a higher order compensation \mathcal{K}_N , described by (20) with the N parameters \mathbf{M}_n , where $n = 1, \dots, N$, chosen in such a way that the first N terms of the operator $\mathcal{K}_c \hat{U} - 1$ expansion in the series in H/b would cancel out, so that $\Gamma = O(H^{N+1}/b^{N+1})$. (Note that in counting the degree of cancellation we do not assume that the system possesses any kind of degeneracy, e.g., one of the kind that led to Γ_1 reduction from being $O(H^2/b^2)$ to $O(H^3/b^3)$.) In this case, the logarithm of the PDF tail of Γ can be estimated by $-(b^2/D_m Z)(\Gamma_N)^{2/(N+1)}$. This results in the following expression for the tail of the PDF of B :

$$\ln \mathcal{S}_N(B) \sim -\frac{b^2}{D_m Z} \left(\frac{D_\xi Z}{I_0} \ln \frac{B}{B_0} \right)^{\frac{2}{N+1}} \quad (35)$$

valid for $D_\xi Z \ln(B/B_0) \gg (D_m Z/b^2)^{(N+1)/2} I_0$. Equation (35) generalizes (26), (33), (34), correspondent to $N = 0, 1, 2$, respectively. We conclude that, as anticipated, the compensation does suppress the PDF tail. The corresponding estimate for the outage probability defined by (23), gives $\ln \mathcal{O} \sim -[(I_0)^{-1} D_\xi Z \ln(B_*/B_0)]^{2/(N+1)} b^2 / (D_m Z)$.

C. Periodic and Quasi-Periodic PMD Compensation: Or How to Resurrect PMD-Damaged Fiber Line

The main purpose of this subsection is introducing new compensation strategies that substantially outperform a standard PMD compensation scheme (discussed in Sections II-D and V-B) with the same number of degrees of freedom N in the case of relatively large N .

Assume that an optical line can be divided into N segments, each of length $l = Z/N$, and apply first-order compensation at

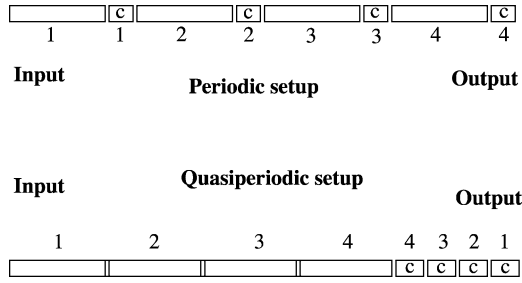


Fig. 2. Cartoon scheme of fiber-line elements installation corresponding to periodic and quasiperiodic compensation strategies.

the end of each segment (as schematically shown in the upper panel of Fig. 2, with “c” denoting the compensating elements). The noise-independent part of the compensated signal for the “periodic compensation” strategy is determined by

$$\mathcal{K}_c \varphi = \exp(i\eta \partial_t^2) \mathcal{K}_{1N} \hat{U}_N \dots \mathcal{K}_{11} \hat{U}_1 \Psi_0(t) \quad (36)$$

$$\hat{U}_n = T \exp \left[\int_{(n-1)l}^{nl} dz h_j(z) \hat{\sigma}_j \partial_t \right] \quad (37)$$

$$\mathcal{K}_{1n} = \exp \left[- \int_{(n-1)l}^{nl} dz h_j(z) \hat{\sigma}_j \partial_t \right] \quad (38)$$

where $\Psi_0(t)$ is the input signal profile, $\eta = \int_0^Z dz d(z)$ is the integral chromatic dispersion, and the ordered product on the RHS of (36) is taken over all the N segments. $T \exp$ is the standard notation for the so-called ordered exponential. The exponential factor \mathcal{K}_{1n} represents the first-order compensation at the end of the n th segment.

Such a “periodic” compensation is not particularly convenient since it requires installation of compensating elements at multiple places along the system. However, one can naturally modify this scheme and have the same compensating elements inserted subsequently but all at once at the fiber output as shown in the lower panel of Fig. 2. If the disorder profile h_j is known (technically such endpoint measurements are possible through the anti-Stokes refraction technique [34]) one can have an endpoint, but multiple, compensation as $\mathcal{K}_c = \prod \mathcal{K}_{1n}$, leading to the following “quasi-periodic” modification of (36):

$$\mathcal{K}_c \varphi = \exp(i\eta \partial_t^2) \mathcal{K}_{11} \dots \mathcal{K}_{1N} \hat{U}_N \dots \hat{U}_1 \Psi_0(t) \quad (39)$$

where \mathcal{K}_{1n} and U_n are defined by (37) and (38). The idea that stands behind the “quasi-periodic” compensation is obvious: We construct (in the compensating part) the best possible approximation (with the given number of the compensating degrees of freedom) for the inverse of the ordered exponential $\hat{U} = \hat{U}_N \dots \hat{U}_1$.

Note that the (quasi)-periodic compensation does not influence the statistics of the noise-dependent part of the signal, i.e., $\mathcal{K}_c \phi$ has the same correlation function (7) as ϕ . Therefore, one arrives at the same expression $\ln(B/B_0) = \Gamma I_0 / (D_\xi Z)$ with a new \mathbf{h} -dependent factor Γ . Furthermore, in the region of our main interest, Γ can be analyzed perturbatively, just as before.

Expanding the factors in (36) up to the second order and making use of (7), (12), and derivations of Appendix II one derives

$$\Gamma_p \approx \frac{\mu'_2}{b^2} \sum_{n=1}^N \int_{a_n}^{nl} dz \int_{a_n}^z dz' [h_1(z)h_2(z') - h_2(z)h_1(z')] \quad (40)$$

where $a_n = (n-1)l$. Here, μ'_2 is the same dimensionless coefficient that has been already discussed in Section V-A, and is also displayed in Fig. 1 for the Gaussian chirped model described in Appendix I. As follows from (39), the same expression (40) is obtained in the second order for the quasi-periodic case. Substituting (40) into the expression for B and evaluating the PDF of B , with the Gaussian statistics of \mathbf{h} described by (9), leads to the following expression for the tail of the PDF of BER:

$$\mathcal{S}_N(B) \sim \frac{B_0^\nu}{B^{1+\nu}}, \quad \nu = \frac{N\pi D_\xi b^2}{2|\mu'_2| D_m I_0}. \quad (41)$$

Equation (41) holds for $\ln(B/B_0) \gg \mu'_2 D_m I_0 / [D_\xi b^2]$. The exponent ν in (41) contains an additional factor N compared to α that explains a steeper tail of $\mathcal{S}(B)$ for (quasi)-periodic compensation compared to the first-order endpoint compensation. It is also instructive to compare the outage probability for the periodic case with the case of higher order endpoint compensation described by (35). One finds that for higher order compensation, i.e., when $N > \mu'_2 \ln(B_*/B_0) I_0 / (D_\xi Z)$, the (quasi)-periodic scheme becomes more efficient compared to the straight N th-order compensation scheme.

Note, that an important computational step that leads to (41) rests on evaluating Γ perturbatively in \mathbf{h} . Besides, in the periodic case, Γ is a direct sum of the individual segment contributions Γ_n , and the perturbative treatment applies separately to each Γ_n , requiring the weakness of the PMD effect at each segment only, i.e., $D_m Z / N \ll b^2$. Therefore, one concludes that even an optical line with not really operable (without compensation) characteristics ($D_m Z$ is of the order or larger than b^2) can still be used for transmission if N is sufficiently large. Moreover, this observation on the applicability of (41) also extends to the quasi-periodic case, in the sense that (41) provides an upper bound for the PDF of BER. This implies that the quasi-periodic arrangement can be superior with respect to the periodic one due to an additional, oscillatory with \mathbf{h} , suppression of Γ_n in the quasi-periodic case versus periodic. This suppression is especially important for segments strongly separated from their compensating counter-segments.

To conclude, in this subsection, we have proposed (quasi)-periodic compensation scheme that appears to be a strong alternative to the standard higher order compensation strategies. The efficiency of the scheme has been demonstrated and even though the technical implementation of this procedure requires expensive equipment, the reduction in the probability of extreme outages can result in an essential overall benefit.

VI. DISCUSSION AND CONCLUSION

In this paper, we evaluated the outage probability that characterizes the reliability of an optical fiber communication system with well-separated time scales related to two different noise mechanisms. The two major impairments which contribute

to the outage are represented by spontaneous emission noise generated in optical amplifiers and birefringent disorder of the fiber. The latter originates from temperature and stress variations in the fiber system. The outage—substantial deviation of BER from its mean value—is characterized in terms of the PDF of the BER. The BER represents the probability of a transmission error (it is found by averaging over many bits of information) for a given realization of birefringent disorder. The outage probability is expressed in terms of the far tail of the PDF of BER (23). This formulation, together with the fact that the object of interest is described in terms of rare events makes numerical or experimental studies extremely difficult (if practical at all).

The proposed method to estimate outage probability is based on first averaging over the amplifier noise, followed by averaging over birefringent disorder. These averaging procedures are very different in nature. The spontaneous noise is short-correlated in time, while birefringent disorder is frozen, i.e., it does not change on the time scale associated with a given pulse transmission throughout the entire system. In an efficient communication system the OSNR is large, thus bit errors, due to fluctuations of the optical field that are large compared to their typical values, occur rarely. The saddle-point (optimal fluctuation) method, developed within the functional integral approach, becomes an adequate (and currently the only) tool for BER evaluation. The BER is a functional of birefringent disorder. As is seen from the general expression (21), even relatively weak variations in the disorder can generate a strong change (additional orders of magnitude) in the value of BER.

It is impossible to compensate for amplifier spontaneous emission noise, whereas effects of birefringent disorder are curable, at least to a certain extent. Since the outage probability characterizes system performance, compensation scheme performance should be compared based on this measure. As illustration, we briefly discuss a relevant example of how efficient various compensation techniques can be. Typical values of the parameters introduced in this paper for real fiber links are: $\Gamma_0 = 0.06$, $\mu_1 = 0.06$, $\mu_2 = 0.12$, $\mu'_2 = 0.15$, and $\mu_3 = 0.35$. Typical bit-error probability is $B_0 = 10^{-12}$ and the value of $I_0/[D_\xi Z] \approx 460$. We also assume that the PMD coefficient $k = \sqrt{12D_m}$ is 0.2 ps/ $\sqrt{\text{km}}$, the pulsewidth is $b = 25$ ps, and the system length is $Z = 2500$ km, i.e., $D_m Z/b^2 \approx 0.013$. Then for these particular set of parameters, the outage probability corresponding to $B_* = 10^{-10}$ is $\mathcal{O} \approx 0.35$ if no compensation is applied, see (26). One derives $\mathcal{O} \approx 0.04$, $\mathcal{O} \approx 4 \cdot 10^{-4}$, and $\mathcal{O} \approx 2 \cdot 10^{-13}$ for (27), (33), and (34), describing the cases of the “setting the clock,” and two cases of the first-order compensations considered in Section V-A, respectively.

Experimental and numerical verification of these results is of considerable importance. Some moderate but significant progress has been made in this direction. Numerical observations corresponding to the “setting the clock” case and consistent with our results (28) are available. Thus, [35, Fig. 2(a)] replotted in log-log variables shows a relation between $\ln S$ and $\ln B$ close to the linear one given by (28). In addition, our major result, the emergence of an extremely extended tail in the PDF of BER that is algebraic or algebraic like (i.e.,

which is much more extended than any log-normal, Gaussian or even exponential expectation would offer), is consistent with experimental measurements on an artificial PMD-modulator system reported in the same volume [36].

We conclude by emphasizing the generality of the theoretical approach developed in the manuscript. Although we restricted our quantitative analysis to the linear model of optical propagation and the RZ modulation format, the fact that *even minor variations of birefringence can lead to major variations of BER* constitutes a general and key feature of the approach. This leads to relatively straightforward generalizations that allow us to consider any other modulation format, account for interchannel nonlinear interactions, and even export the approach to the nonlinear (soliton) transmission regime. Those cases are currently under study and the results will be published elsewhere.

APPENDIX I

OPTICAL AND ELECTRICAL FILTERS AND INITIAL SHAPE OF THE PULSE

Our approach is general. However, for illustrative purposes, we choose to stick to a simple model of signal coding (form of initial pulse) and decoding (particular shape of optical and electrical filters). We also consider the simplest choice of the decision level value $I_d = I_0/2$. These modeling assumptions allow us to get quantitative results, i.e., results valid not only parametrically [and here the major focus will be on the two-dimensionless parameters entering the left-hand side (LHS) of (13), (14)] but also up to numbers, the coefficients, generally being dependent on the model.

Let us formulate our illustrative model. We assume a Lorentzian shape of the optical filter

$$\mathcal{K}_f \Psi = \int_0^\infty dt' \exp\left(-\frac{t'}{\tau}\right) \frac{\Psi(t-t')}{\tau} \quad (42)$$

where τ is the optical filter temporal width. Then, taking into account that the statistics of ϕ are insensitive to the birefringence field \mathbf{h} , one gets from (7) that the statistics of the inhomogeneous contribution are governed by the PDF \mathcal{P}

$$\mathcal{P}(\tilde{\phi}) = N^{-1} \exp\left\{-\frac{1}{D_\xi Z} \int dt \left[|\tilde{\phi}|^2 + \tau^2 |\partial_t \tilde{\phi}|^2\right]\right\} \quad (43)$$

where $\tilde{\phi} \equiv \mathcal{K}_f \phi$, and N is a normalization factor. Equation (43) defines the measure of averaging with respect to the noise: one should integrate over realizations of ϕ with the weight (43). The electrical (window) filter is chosen as

$$G(t) \propto \theta(|t| - T) = \begin{cases} 1, & |t| < T \\ 0, & |t| > T \end{cases} \quad (44)$$

where T stands for the electric filter temporal width. We assume that the initial signal encoding “1” is Gaussian

$$\Psi_0(t) \propto \exp\left[-\frac{t^2}{2b^2}(1 - 2i\beta_{\text{in}})\right] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (45)$$

where b is the pulsewidth and β_{in} stands for the initial pulse chirp.

APPENDIX II
SADDLE-POINT EVALUATION OF BER
(FOR GIVEN REALIZATION OF DISORDER)

The transition probability (16) and the measure of averaging over ϕ , described by (43), are the two expressions that constitute the starting point for the calculations presented in this appendix. Substituting (43) into (16) and introducing an additional Fourier transform leads to

$$\begin{aligned} \mathcal{P}(I_{\text{out}}|I_{\text{in}}; \{\mathbf{h}\}; Z) &= N^{-1} \int \frac{d\lambda}{2\pi} \int \mathcal{D}\tilde{\phi}(t) \\ &\times \exp \left\{ i\lambda \left(I_{\text{out}} - \int dt G(t) |\mathcal{K}\varphi + \tilde{\phi}|^2 \right) \right. \\ &\quad \left. - \frac{1}{D_\xi Z} \int dt \left[|\tilde{\phi}|^2 + \tau^2 |\partial_t \tilde{\phi}|^2 \right] \right\} \quad (46) \end{aligned}$$

where dependencies on I_{out} and \mathbf{h} are contained solely in the $\mathcal{K}\varphi(Z, t)$ field. The condition (13) locates the path-integral over $\tilde{\phi}$ and the usual integral over λ to the saddle point, i.e., the integrals are dominated by a single configuration called the saddle point. The saddle-point evaluation of (46) results in

$$\tau^2 \partial_t^2 \tilde{\phi}_{\text{sp}} - \tilde{\phi}_{\text{sp}} - uG(t)\tilde{\phi}_{\text{sp}} = uG(t)\mathcal{K}\varphi \quad (47)$$

$$I_{\text{out}} = \int dt G(t) |\mathcal{K}\varphi + \tilde{\phi}_{\text{sp}}|^2 \quad (48)$$

$$\begin{aligned} \ln[\mathcal{P}_{\text{sp}}(I_{\text{out}})] &\approx -\frac{1}{D_\xi Z} \int dt \left(|\tilde{\phi}_{\text{sp}}|^2 + \tau^2 |\partial_t \tilde{\phi}_{\text{sp}}|^2 \right) \\ &= \frac{u}{D_\xi Z} \left[I_{\text{out}} - \int dt G(t) (\mathcal{K}\varphi + \tilde{\phi}_{\text{sp}}) \mathcal{K}\varphi^* \right] \quad (49) \end{aligned}$$

where $u \equiv i\lambda D_\xi Z$ is a real number, and (47)–(49) thus constitute the system of integrodifferential-algebraic equations. Subscript “sp” stands for saddle-point and we will skip it in all following formulas to shorten the notation.

Note that (47)–(49) are not singular in the $\tau \rightarrow 0$ limit, which should, however, be treated with caution since the corrections to the leading saddle-point approximation (accounting for the so-called determinant that accounts for Gaussian fluctuations around the saddle point) are actually singular in this limit. The condition that the fluctuations do not destroy the saddle-point result reads $\tau \gg D_\xi Z b / I_0$. Thus, in the interesting asymptotic range, $b \gg \tau \gg D_\xi Z b / I_0$, (47)–(49) transform to

$$\tilde{\phi} = -\frac{uG}{1+uG} \mathcal{K}\varphi \quad (50)$$

$$I_{\text{out}} = \int dt \frac{G(t) |\mathcal{K}\varphi|^2}{(1+uG)^2} \quad (51)$$

$$\ln[\mathcal{P}(I_{\text{out}})] = -\int dt \left(\frac{uG}{1+uG} \right)^2 \frac{|\mathcal{K}\varphi|^2}{D_\xi Z}. \quad (52)$$

For a step-function shape of the electrical (window) filter (44), one gets an explicit expression for u in terms of φ and I_{out} from (51) which, substituted into (52), gives

$$\ln[\mathcal{P}(I_{\text{out}})] = -\frac{\left(\sqrt{\int dt G(t) |\mathcal{K}\varphi|^2} - \sqrt{I_{\text{out}}} \right)^2}{D_\xi Z}. \quad (53)$$

In the general $\tau \sim b$ case, one does not get an explicit analog of (53), however, the general system of (47)–(49) can still be es-

entially simplified for the step-function shape of the electrical (window) filter (44). The solution of (47) that satisfies a zero condition at $t = \pm\infty$ is

$$\tilde{\phi}(t) = \begin{cases} \phi_a(t) + \mathbf{B}_+ \exp\left[-\frac{t\vartheta}{\tau}\right] + \mathbf{B}_- \exp\left[\frac{t\vartheta}{\tau}\right], & |t| < T \\ \mathbf{A}_+ \exp\left[-\frac{t}{\tau}\right], & t > T \\ \mathbf{A}_- \exp\left[\frac{t}{\tau}\right], & t < -T \end{cases} \quad (54)$$

where $\vartheta = \sqrt{1+u}$, and $\phi_a(t)$ is the solution of the auxiliary problem described by modifying (47) with $G(t)$ on the LHS of (47) replaced by one. Thus

$$\phi_a(t) = -\frac{u}{2\tau\vartheta} \int_{-T}^T dt' |\mathcal{K}\varphi|(t') \exp\left(-\frac{|t-t'|\vartheta}{\tau}\right). \quad (55)$$

Four coefficients in (54) are fixed by four obvious conditions imposing continuity for both $\tilde{\phi}(t)$ and $\partial_t \tilde{\phi}(t)$ at $t = \pm T$. Taking into account that $\phi'_a(\pm T) = \mp\vartheta \phi_a(\pm T)/\tau$, one derives

$$\begin{aligned} \mathbf{B}_\pm &\equiv (\vartheta - 1) \\ &\times \frac{(1+\vartheta)e^{\frac{T\vartheta}{\tau}} \zeta_a(\mp T) + (\vartheta-1)e^{-\frac{T\vartheta}{\tau}} \zeta_a(\pm T)}{(1+\vartheta)^2 e^{\frac{2T\vartheta}{\tau}} - (\vartheta-1)^2 e^{-\frac{2T\vartheta}{\tau}}} \quad (56) \end{aligned}$$

$$\begin{aligned} \mathbf{A}_\pm &\equiv 2e^{\frac{T(1+\vartheta)}{\tau}} \vartheta \\ &\times \frac{(1+\vartheta)e^{\frac{T\vartheta}{\tau}} \phi_a(\pm T) + (\vartheta-1)e^{-\frac{T\vartheta}{\tau}} \phi_a(\mp T)}{(1+\vartheta)^2 e^{\frac{2T\vartheta}{\tau}} - (\vartheta-1)^2 e^{-\frac{2T\vartheta}{\tau}}} \quad (57) \end{aligned}$$

that results in the following explicit expression for $\tilde{\phi}$ in terms of $\mathcal{K}\varphi$:

$$\begin{aligned} \tilde{\phi}(t) &= -\frac{u}{2\tau\vartheta} \int_{-T}^T dt' [\mathcal{K}\varphi](t') \\ &\times \left\{ \exp\left(-\frac{|t-t'|\vartheta}{\tau}\right) \right. \\ &\quad \left. + \frac{\vartheta-1}{(1+\vartheta)^2 e^{\frac{2T\vartheta}{\tau}} - (\vartheta-1)^2 e^{-\frac{2T\vartheta}{\tau}}} \right. \\ &\quad \times \left(\exp\left[-\frac{t'+T}{\tau}\vartheta\right] \right. \\ &\quad \times \left[(1+\vartheta)e^{\frac{\vartheta(T-t)}{\tau}} + (\vartheta-1)e^{\frac{\vartheta(t-T)}{\tau}} \right] \\ &\quad \left. + \exp\left[-\frac{T-t'}{\tau}\vartheta\right] \left[(1+\vartheta)e^{\frac{\vartheta(T+t)}{\tau}} \right. \right. \\ &\quad \left. \left. + (\vartheta-1)e^{-\frac{\vartheta(t+T)}{\tau}} \right] \right) \left. \right\} \quad (58) \end{aligned}$$

where only the $|t| < T$ part of the $\varphi(t)$ function is shown. Therefore, (58), substituted into the following two equations derived from (48), (17):

$$\frac{I_0}{2} = \int_{-T}^T dt |\mathcal{K}\varphi + \tilde{\phi}|^2 \quad (59)$$

$$\ln[B] \approx \frac{u}{D_\xi Z} \left[\frac{I_0}{2} - \int_{-T}^T dt (\mathcal{K}\varphi + \tilde{\phi}) \mathcal{K}\varphi^* \right] \quad (60)$$

gives a complete description of saddle-point approximation result. In (59) and (60) we have also assumed that $I_d = I_0/2$.

Note, that even though the explicit solution of (47), presented in (58) is found, one still cannot solve (59), i.e., one cannot express u explicitly in terms of $\varphi(t)$. Therefore, the remainder of this appendix is devoted to numerical solution of (48). Step by step, we analyze the situations corresponding to a variety of cases studied in Section IV. In the remainder of this appendix, we describe an exact recipe for numerical calculation of the dimensionless coefficients that lead to the results shown in Fig. 1.

A. Γ_0 -Calculation

First, one studies the zero PMD, $\mathbf{h} = 0$, case. Replacing $\mathcal{K}\varphi$ in (58), (59), and (60) by $\mathcal{K}_f\Psi_0$, where \mathcal{K}_f and Ψ_0 are defined by (42) and (45), respectively. Then, $u = u_0$ and $\phi = \phi_0$ are the numerical solution of (58), (59) for fixed T/b and τ/b . Finally, from (21), (60)

$$\Gamma_0 = u_0 \left[-\frac{1}{2} + \frac{\int_{-T}^T dt (\mathcal{K}_f\Psi_0 + \tilde{\phi}_0) \mathcal{K}_f\Psi_0^*}{\int_{-T}^T dt |\mathcal{K}_f\Psi_0|^2} \right] \quad (61)$$

where we have used $I_0 = \int_{-T}^T dt |\mathcal{K}_f\Psi_0|^2$. Note that the initial chirp β_{in} does not enter Γ_0 .

B. μ_1 -Calculation

The idea is to solve (58) and (59) perturbatively with respect to \mathbf{h} . Since in this case, $\mathcal{K}\varphi \rightarrow (1 + H_3\partial_t)\mathcal{K}_f\Psi_0$ (also with $\beta_{\text{in}} = 0$), one presents u and ϕ in the forms $u = u_0 + H_3u_1/b + O(h^2)$ and $\tilde{\phi} = \tilde{\phi}_0 + H_3\tilde{\phi}_1/b + O(h^2)$. Then, for

$$D_\xi Z \ln[B/B_0]/I_0 \rightarrow \Gamma_1 = \mu_1 H_3/b$$

one derives from (58)–(60)

$$\tilde{\phi}_1 = \tilde{\phi} \{ \mathcal{K}\varphi \rightarrow \mathcal{K}_f\partial_t\Psi_0, u \rightarrow u_0 \} \quad (62)$$

$$u_1 = - \frac{\int_{-T}^T dt (\tilde{\phi}_0 + \mathcal{K}_f\Psi_0) (\tilde{\phi}_1 + \mathcal{K}_f\partial_t\Psi_0)}{\int_{-T}^T dt (\tilde{\phi}_0 + \mathcal{K}_f\Psi_0) \frac{\partial\tilde{\phi}_0}{\partial u} \Big|_{u \rightarrow u_0}} \quad (63)$$

$$\mu_1 = - \frac{u_1\Gamma_0}{u_0} - \frac{u_0}{\int_{-T}^T dt |\mathcal{K}_f\Psi_0|^2} \times \int_{-T}^T dt \left[\mathcal{K}_f\partial_t\Psi_0 \times (\mathcal{K}_f\Psi_0 + \tilde{\phi}_0) + \mathcal{K}_f\Psi_0 (\tilde{\phi}_1 + \mathcal{K}_f\partial_t\Psi_0) \right]. \quad (64)$$

C. μ_2 -Calculation

A similar perturbative strategy will work in the zero chirp “setting the clock” compensation case. The idea of the compensation is to adjust the time shift t_{cl} in such a way that the first terms of Γ ’s expansion in \mathbf{h} vanish. Formally, pure “setting the clock” compensation means: $\mathcal{K}\varphi(t) \rightarrow \hat{W}\mathcal{K}_f\Psi_0(t - t_{\text{cl}})$. Substituting this expression into (58)–(60), assuming that $t_{\text{cl}} = O(h)$, and making respective variations with respect to h , about $u = u_0$, of (59) one derives that $u = u_0 + (H_3 - t_{\text{cl}})u_1/b + O(h^2)$, $\tilde{\phi} = \tilde{\phi}_0 + (H_3 - t_{\text{cl}})\tilde{\phi}_1/b + O(h^2)$, and finally $\Gamma = \mu_1(H_3 - t_{\text{cl}})/b + O(h^2)$. Therefore, one concludes that to cancel the $O(h)$ -terms in Γ one needs to set the clock according to: $t_{\text{cl}} = H_3$.

Then, modifying the perturbative scheme from the previous subsection, according to $\mathcal{K}\varphi \rightarrow (1 + H_\perp^2\partial_t^2)\mathcal{K}_f\Psi_0$ (also with $\beta_{\text{in}} = 0$), $u = u_0 + H_\perp^2u_2/b + O(h^3)$ and $\tilde{\phi} = \tilde{\phi}_0 + H_\perp^2\tilde{\phi}_2/b + O(h^3)$, one arrives at expressions equivalent to (62)–(64) with all “1”-subscripts replaced by “2” and ∂_t replaced by ∂_t^2 , respectively.

D. μ'_2 -Calculation

Here we discuss the case when first-order PMD compensation is applied while the output signal has small but nonzero chirp. Perturbative calculations, outlined in Subsection II-B of this appendix, are applicable with the following modifications: $\mathcal{K}\varphi \rightarrow (1 + Y_c\partial_t^2)\mathcal{K}_f\Psi_0$, where $\Psi_0 = \Psi_{0;R} + i\mu\Psi_{0;I}$, $\Psi_{0;R} = C_g \exp(-t^2/[2b^2])$, $\Psi_{0;I} = C_g \exp(-t^2/[2b^2])t^2/b^2$, and Y_c is defined in (30).

We are looking for perturbative solution of (58)–(60) in the form: $u = u_0 + Y_c u'_2/b^2 + O(h^3)$ and $\tilde{\phi} = \tilde{\phi}_0 + Y_c \tilde{\phi}_2/b^2 + O(h^3)$, where, however, both $\tilde{\phi}_0$ and $\tilde{\phi}_2$ are complex: $\tilde{\phi}_k = \tilde{\phi}_{k;R} + i\beta\tilde{\phi}_{k;I}$, where $k = 0, 2$. Then, the analogs of (62)–(64) are

$$\tilde{\phi}_{k;P} = \tilde{\phi} \{ \mathcal{K}\varphi \rightarrow \mathcal{K}_f\partial_t^2\Psi_{k;P}, u \rightarrow u_0 \} \quad (65)$$

where $k = 0, 2$ and $P = I, R$

$$u'_2 = \frac{\int_{-T}^T dt (\tilde{\phi}_{0;R} + \mathcal{K}_f\Psi_{0;R}) (\tilde{\phi}_{2;I} + \mathcal{K}_f\partial_t^2\Psi_{0;I})}{\int_{-T}^T dt (\tilde{\phi}_{0;R} + \mathcal{K}_f\Psi_{0;R}) \frac{\partial\tilde{\phi}_{0;R}}{\partial u} \Big|_{u \rightarrow u_0}} - \frac{\int_{-T}^T dt (\tilde{\phi}_{0;I} + \mathcal{K}_f\Psi_{0;I}) (\tilde{\phi}_{2;R} + \mathcal{K}_f\partial_t^2\Psi_{0;R})}{\int_{-T}^T dt (\tilde{\phi}_{0;R} + \mathcal{K}_f\Psi_{0;R}) \frac{\partial\tilde{\phi}_{0;R}}{\partial u} \Big|_{u \rightarrow u_0}} \quad (66)$$

$$\mu'_2 = - \frac{u'_2\Gamma_0}{u_0} + \frac{u_0}{\int_{-T}^T dt |\mathcal{K}_f\Psi_0|^2} \times \int_{-T}^T dt \left[\mathcal{K}_f\partial_t^2\Psi_{0;I} \times (\mathcal{K}_f\Psi_{0;R} + \tilde{\phi}_{0;R}) - \mathcal{K}_f\partial_t^2\Psi_{0;R} (\tilde{\phi}_{0;I} + \mathcal{K}_f\Psi_{0;I}) \right]. \quad (67)$$

E. μ_3 -Calculation

Perturbative calculations for this case (first-order PMD compensation and no chirp, i.e., all the functions are real) are absolutely analogous to those explained in Subsection II-B of this appendix. Then, with $\mathcal{K}\varphi \rightarrow (1 + Y_r\partial_t^3)\mathcal{K}_f\Psi_0$ (also with $\beta_{\text{in}} = 0$), where Y_r is defined in (31), and $u = u_0 + Y_r u_3/b^3 + O(h^4)$, $\tilde{\phi} = \tilde{\phi}_0 + Y_r \tilde{\phi}_3/b^3 + O(h^4)$, and $\Gamma = \mu_3 Y_r/b^3 + O(h^4)$, one derives from (58)–(60)

$$\tilde{\phi}_3 = \tilde{\phi} \{ \mathcal{K}\varphi \rightarrow \mathcal{K}_f\partial_t^3\Psi_0, u \rightarrow u_0 \} \quad (68)$$

$$u_3 = - \frac{4 \int_{-T}^T dt (\tilde{\phi}_0 + \mathcal{K}_f\Psi_0) (\tilde{\phi}_3 + \mathcal{K}_f\partial_t^3\Psi_0)}{3 \int_{-T}^T dt (\tilde{\phi}_0 + \mathcal{K}_f\Psi_0) \frac{\partial\tilde{\phi}_0}{\partial u} \Big|_{u \rightarrow u_0}} \quad (69)$$

$$\mu_3 = - \frac{u_3\Gamma_0}{u_0} - \frac{4u_0}{3 \int_{-T}^T dt |\mathcal{K}_f\Psi_0|^2} \times \int_{-T}^T dt \left[\mathcal{K}_f\partial_t^3\Psi_0 \times (\mathcal{K}_f\Psi_0 + \tilde{\phi}_0) + \mathcal{K}_f\Psi_0 (\tilde{\phi}_3 + \mathcal{K}_f\partial_t^3\Psi_0) \right]. \quad (70)$$

APPENDIX III
h-AVERAGING

A formal definition of the ***h***-averaging, i.e., averaging with respect to the statistics of the birefringence pseudovector, reads

$$\langle A\{\mathbf{h}\} \rangle_h \equiv \int \mathcal{D}\mathbf{h}(z) \mathcal{P}_h A\{\mathbf{h}\} \quad (71)$$

$$\mathcal{P}_h\{\mathbf{h}\} = N_m \exp \left[-\frac{\int_0^Z dz \mathbf{h}^2(z)}{2D_m} \right] \quad (72)$$

where $A\{\mathbf{h}\}$ is an arbitrary functional of ***h***, integration in the RHS of (71) is functional (path-integral), and N_m is a normalization coefficient enforcing $\langle 1 \rangle_h = 1$. Obviously, (71) and (72) are consistent with (9).

Our goal is to derive the PDFs for four auxiliary objects: H_3 , H_\perp^2 , Y_c , and Y_r defined in (25), (27), (30), and (31), respectively, starting from (71) and (72).

While the PDFs of H_3 and H_\perp^2 follow directly from (71), (72)

$$\mathcal{P}_1(H_3) \sim \exp \left[-\frac{H_3^2}{2D_m Z} \right] \quad (73)$$

$$\mathcal{P}_2(H_\perp^2) \sim \exp \left[-\frac{H_\perp^2}{2D_m Z} \right] \quad (74)$$

calculating the PDFs for the other two objects is less straightforward.

A. Statistics of Y_c

The PDF of Y_c can be recast as

$$\mathcal{P}_3(Y_c) = \int \frac{d\lambda}{2\pi} \exp(i\lambda Y) \Xi(i\lambda) \quad (75)$$

$$\Xi(s) = \int \mathcal{D}h \exp \left\{ -\frac{1}{2D_m} \int_0^Z dz (h_1^2 + h_2^2) - s \int_0^Z dz \int_0^z dz' [h_1(z)h_2(z') - h_2(z)h_1(z')] \right\}. \quad (76)$$

Differentiating (76) with respect to s one gets

$$\partial_s \ln \Xi = - \int_0^Z dz \int_0^z dz' G_-(z, z') \quad (77)$$

$$G_-(z, z') \equiv G_{12}(z, z') - G_{21}(z, z') \quad (78)$$

$$G_{ij}(z, z') \equiv \langle h_i(z)h_j(z') \rangle \quad (79)$$

where averaging in the definition of the correlation functions G_{ij} is performed using the Gaussian measure defined by (76). Using the G -function definition and making a set of Gaussian integral transformations (integration by parts) one derives

$$\frac{G_{ij}(z, z_1)}{D_m} + s \int_0^Z \text{sign}(z - z') |\varepsilon_{ik}| G_{kj}(z', z_1) = \delta_{ij} \delta(z - z') \quad (80)$$

where ε_{ik} is the antisymmetric $2 \rightarrow 2$ tensor and summation over k is assumed in the LHS of (80). Equation (80) transforms into the following equation for G_- :

$$\frac{\partial_z G_-(z, z_1)}{2D^2 s} + s \int_0^z dz' \text{sign}(z - z') G_-(z', z_1) = -2\delta(z - z_2). \quad (81)$$

Equation (81) has to be supplemented by the boundary conditions

$$G_-(Z, z_1) = -G_-(0, z_1), \quad \partial_z G_-(Z, z_1) = -\partial_z G_-(0, z_1). \quad (82)$$

The solution of (81) that satisfies (82) has the form

$$G_-(z, z_1) = \frac{\sin\left(\frac{\pi}{2} + Ds(2(z - z_1) + Z\text{sign}[z_1 - z])\right)}{\sin\left(\frac{\pi}{2} + DsZ\text{sign}[z_1 - z]\right)} \text{sign}[z_1 - z]. \quad (83)$$

Substituting (83) into (77) and performing integrations over z , z' explicitly, one derives $\partial_s \ln \Xi = D_m \tan(D_m s Z)$, resulting in

$$\Xi(s) = \frac{1}{\cos(D_m s Z)}. \quad (84)$$

Combining (84), (75), and integrating over λ , one arrives at

$$\mathcal{P}_3(Y_c) = \frac{1}{2D_m Z} \cosh^{-1} \left(\frac{\pi Y_c}{2D_m Z} \right). \quad (85)$$

B. Statistics of Y_r

The PDF of Y_r , defined in (31), can be written as

$$\mathcal{P}_4(Y_r) = \int \frac{d\lambda}{2\pi} \exp(-i\lambda Y_r) \int \mathcal{D}\rho \exp \left\{ -\frac{1}{2D_m^3 Z^3} \times \int_0^1 d\tau \left[\left(\frac{d\rho}{d\zeta} \right)^2 + \lambda^2 (3\rho^2 - 2\rho_1 \rho)^2 \right] \right\} \quad (86)$$

where $\zeta = z/Z$ and $\rho_1 = \rho(\zeta = 1)$. Calculating the integral over λ one arrives at

$$\mathcal{P}_4(Y_r) = \int \frac{\mathcal{D}\rho}{(2\pi V)^{\frac{1}{2}}} \times \exp \left\{ -\frac{1}{2D_m^3 Z^3} \times \left[\int_0^1 d\zeta \left(\frac{d\rho}{d\zeta} \right)^2 + \frac{\tilde{Y}^2}{V} \right] \right\} \quad (87)$$

$$\tilde{Y} = D_m^3 Z^3 Y_r$$

$$V = \int_0^1 d\zeta (3\rho^2 - 2\rho_1 \rho)^2. \quad (88)$$

The condition (14) allows saddle-point calculation of the path-integral on the RHS of (87). The saddle-point equations support the conservation of $E = \text{const}$ where

$$E = \dot{\rho}^2 + \frac{\tilde{Y}^2}{V^2} (3\rho^2 - 2\rho_1 \rho)^2, \quad \ln \mathcal{P}_4 \approx -\frac{E}{2D_m^3 Z^3} \quad (89)$$

and ρ corresponds to projection of $\boldsymbol{\rho}$ on $\boldsymbol{\rho}_1$ (the other component of $\boldsymbol{\rho}$ orthogonal to $\boldsymbol{\rho}_1$ is zero). One finds that (89) can be recast in a form

$$\dot{\rho}^2 + (3\rho^2 - 2\eta\rho)^2 = \frac{\kappa^2 \eta^4}{9}, \quad E = \frac{V^2 \kappa^2 \eta^4}{\tilde{Y}^2 9} \quad (90)$$

where $\eta = \varrho(1)$ and $\rho = (V/\tilde{Y})\varrho$. Note, that $3 < \kappa < \infty$. Substituting (90) for E into (89) for $\ln \mathcal{P}$ and expressing V via ϱ from (88) leads to

$$\ln \mathcal{P}_4 \approx -\frac{SY_r^{\frac{2}{3}}}{D_m Z} \quad (91)$$

$$S = \frac{\kappa^2 \eta^2}{18A^{\frac{2}{3}}} \quad (92)$$

$$A = \eta \int_0^1 d\zeta \left[3 \left(\frac{\varrho}{\eta} \right)^2 - \frac{2\varrho}{\eta} \right]^2.$$

It is convenient to introduce a new parameterization

$$\varrho = \frac{\eta}{3} (1 + \sqrt{1 + \kappa} \sin \varphi) \quad (93)$$

that leads to

$$\eta = \left| \int_{\varphi_0}^{\varphi_n} \frac{d\varphi}{\sqrt{\frac{2\kappa}{(1+\kappa) - \cos^2 \varphi}}} \right| \quad (94)$$

$$A = \frac{1}{9} \left| \int_{\varphi_0}^{\varphi_n} d\varphi \frac{[\kappa - (1 + \kappa) \cos^2 \varphi]^2}{\sqrt{\frac{2\kappa}{(1+\kappa) - \cos^2 \varphi}}} \right| \quad (95)$$

$$\sin \varphi_n = \frac{2}{\sqrt{1 + \kappa}}. \quad (96)$$

One finds (numerically) that $\varphi_2 = \pi - \arcsin(2/\sqrt{1 + \kappa})$ corresponds to the solution with the lowest possible value of S , $S_m \approx 4.185$, achieved at $\kappa \approx 3.145$.

APPENDIX IV REMOTE TAIL OF THE PDF OF BER

In this appendix, we study a universal remote tail of $\mathcal{S}(B)$ corresponding to huge fluctuations of birefringent disorder when the signal is almost destroyed by the fluctuations. Thus, in the parametric range

$$\max \{ (D_m Z)^2, (D_\xi Z)^2 \} \ll D_\xi Z \ln B \ll 1 \quad (97)$$

the BER is formed by configurations of $\mathbf{h}(z)$, where the normalized intensity of the signal, without the noise $\mathcal{I}\{\mathbf{h}\} = I\{\mathbf{h}, \boldsymbol{\xi} = 0\}/I_d$ [see also (12)] satisfies inequality $D_\xi z \ll \mathcal{I} - 1 \ll 1$. In this case, the expression (21) for BER is replaced by

$$B = \exp \left[-C \frac{(\mathcal{I}\{\mathbf{h}\} - 1)^2}{D_\xi Z} \right] \quad (98)$$

where $C \sim 1$. (Strictly speaking, this coefficient also depends on \mathbf{h} , however, for configurations \mathbf{h} that correspond to the leading contribution into $\mathcal{S}(B)$ at (97), this dependence is weak, and thus can be neglected.) The PDF of $\mathcal{I} = \mathcal{I}\{\mathbf{h}\}$, as follows from (9), adopts the following form:

$$\mathcal{P}(\mathcal{I}) \propto \exp \left(-\frac{F(\mathcal{I})}{D_m Z} \right) \quad (99)$$

where the function $F(\mathcal{I})$ does not depend of $D_m Z$, and for $\mathcal{I} \sim 1$, it is $O(1)$. Comparing (98), (99) and also making use of $\mathcal{I} - 1 \ll 1$, we obtain the following universal asymptotics for the PDF of BER, valid on the interval (97):

$$\ln \mathcal{S}(B) \approx -\frac{C_1}{D_m Z} + C_2 \sqrt{\frac{D_\xi}{D_m^2 Z}} \ln \frac{1}{B} \quad (100)$$

where $C_1 = F(1)$ and $C_2 \sim |F'(1)|$ are constants of the order one. The principal factor in (100) is $\exp[-C_1/(D_m z)]$ supplemented by a relatively weak dependence on B .

ACKNOWLEDGMENT

The authors wish to thank D. Chowdhury, M. Hempstead, F. Kueppers, and P. Masmyshev for valuable discussions. They are also thankful to Referee A for useful remarks and suggested corrections.

REFERENCES

- [1] S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single-mode fibers," *Opt. Lett.*, vol. 3, pp. 60–63, 1978.
- [2] S. Machida, I. Sakai, and T. Kimura, "Polarization conservation in single-mode fibers," *Electron. Lett.*, vol. 17, p. 494, 1981.
- [3] N. S. Bergano, C. D. Poole, and R. E. Wagner, "Investigation of polarization dispersion in long lengths of single-mode fiber using multilongitudinal mode lasers," *J. Lightwave Technol.*, vol. LT-5, pp. 1618–1622, Nov. 1987.
- [4] D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurements of the group velocity difference between the principal states of polarization on a low-birefringence terrestrial fiber cable," *Opt. Lett.*, vol. 12, pp. 844–846, 1987.
- [5] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, pp. 1029–1030, 1986.
- [6] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 13, pp. 687–689, 1988.
- [7] C. D. Poole, N. S. Bergano, R. E. Wagner, and H. J. Schulte, "Polarization dispersion and principal states in a 147-km undersea lightwave cable," *J. Lightwave Technol.*, vol. 6, pp. 1185–1190, July 1988.
- [8] C. D. Poole, J. H. Winters, and J. A. Nagel, *Opt. Lett.*, vol. 16, pp. 372–374, 1991.
- [9] H. Bülow, "System outage probability due to first- and second-order PMD," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 696–698, May 1998.
- [10] H. Kogelnik, L. E. Nelson, J. P. Gordon, and R. M. Jopson, "Jones matrix for second-order polarization mode dispersion," *Opt. Lett.*, vol. 25, pp. 19–21, 2000.
- [11] A. Eyal, Y. Li, W. K. Marshall, A. Yariv, and M. Tur, "Statistical determination of the length dependence of high-order polarization mode dispersion," *Opt. Lett.*, vol. 25, pp. 875–877, 2000.
- [12] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 310–312, Mar. 2002.
- [13] C. D. Poole and J. A. Nagel, *Optical Fiber Telecommunications*, I. P. Kaminov and T. L. Koch, Eds. San Diego, CA: Academic, 1997, vol. IIIA, p. 114.
- [14] R. M. Jopson, L. E. Nelson, G. J. Pendlock, and A. H. Gnauck, "Polarization mode dispersion impairment in return to zero and nonreturn-to-zero systems," in *Tech. Dig. Optical Fiber Communication Conf., OFC 1999*, 1999, paper WE3.
- [15] F. Heismann, D. Fishman, and D. Wilson, "Automatic compensation of first-order polarization-mode dispersion in a 10 Gb/s transmission system," in *Proc. ECOC 1998*, 1998, pp. 529–530.
- [16] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Nat. Acad. Science USA*, vol. 97, pp. 4541–4550, 2000.
- [17] N. Gisin, B. Gisin, J. P. Von der Weid, and R. Passy, "How accurately can one measure a statistical quantity like polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 1671–1673, Dec. 1996.
- [18] L. E. Nelson, R. M. Jopson, H. Kogelnik, and J. P. Gordon, "Measurement of polarization mode dispersion vectors using the polarization-dependent signal delay method," *Optics Express*, vol. 6, pp. 158–167, 2000.
- [19] T. Ono, S. Yamazaki, H. Shimizu, and H. Emura, "Polarization control method for suppressing polarization mode dispersion influence in optical-transmission systems," *J. Lightwave Technol.*, vol. 12, pp. 891–898, May 1994.
- [20] L. Moller and H. Kogelnik, "PMD emulator restricted to first and second order PMD generation," in *Proc. ECOC 1999*, 1999, pp. 64–65.

- [21] H. Bülow, F. Buchali, W. Baumert, R. Ballentin, and T. Wehren, "PMD mitigation at 10 Gbit/s using linear and nonlinear integrated electronic equaliser circuits," *Electron. Lett.*, vol. 36, pp. 163–164, 2000.
- [22] V. Chernyak, M. Chertkov, I. Kolokolov, and V. Lebedev, "Probability of anomalously large bit-error-rate in long haul optical transmission," *Phys. Rev. Lett. E*, vol. 68, pp. 066619–066622, 2003, submitted for publication.
- [23] —, "Extreme outages due to polarization mode dispersion," *Opt. Lett.*, vol. 28, pp. 2159–2161.
- [24] —, "Compensation for extreme outages caused by polarization mode dispersion and amplifier noise," *Opt. Express*, vol. 11, p. 1607, 2003.
- [25] —, "Periodic and quasiperiodic compensation strategies of extreme outages caused by polarization mode dispersion and amplifier noise," *Pis'ma v ZhETF*, vol. 78, no. 4, pp. 234–237, 2003.
- [26] R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Appl. Opt.*, vol. 18, pp. 2241–2251, 1979.
- [27] I. P. Kaminow, "Polarization in optical fibers," *IEEE J. Quant. Electron.*, vol. QE-17, pp. 15–22, Jan. 1981.
- [28] G. P. Agrawal, *Nonlinear Fiber Optics*. New York: Academic, 1989.
- [29] A. A. Abrikosov, L. P. Gorkov, and A. I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*. New York: Dover, 1977.
- [30] E. Desurvire, *Erbium-Doped Fiber Amplifiers*. New York: Wiley, 1994.
- [31] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York: Wiley, 1957.
- [32] P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill Higher Education, 1965.
- [33] V. N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics*. Norwell, MA: Kluwer, 2001.
- [34] B. Huttner, B. Gisin, and N. Gisin, "Distributed PMD measurement with a polarization-OTDR in optical fibers," *J. Lightwave Technol.*, vol. 17, pp. 1843–1848, Oct. 1999.
- [35] C. Xie, H. Sunnerud, M. Karlsson, and P. A. Andrekson, "Polarization-mode dispersion-induced outages in soliton transmission systems," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 1079–1081, Oct. 2001.
- [36] U. Neukirch, M. Hempstead, G. Piech, Y. Mauro, M. Mlejnek, M. J. Soulliere, M. Webb, D. Pikula, R. Hoyt, M. Anderegg, M. Dailey, F. Wang, and C. Drewnowski, "Time resolved performance analysis of a second-order PMD compensator," *J. Lightwave Technol.*, vol. 22, pp. 0000–0000, Apr. 2003.
- [37] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 14, pp. 523–525, 1989.

Vladimir Chernyak, photograph and biography not available at the time of publication.

Michael Chertkov, photograph and biography not available at the time of publication.

Ildar Gabitov, photograph and biography not available at the time of publication.

Igor Kolokolov, photograph and biography not available at the time of publication.

Vladimir Lebedev, photograph and biography not available at the time of publication.