Modeling errors in long-haul optical fiber transmission systems by using instantons and Edgeworth expansion

Miloš Ivković, Ivan Djordjević Member, IEEE, Predrag Rajković and Bane Vasić Senior Member, IEEE

Abstract—In this work we use a new approach to model error events in long-haul optical fiber transmission systems. Existing approaches for obtaining probability density functions (PDFs) rely on numerical simulations or analytical approximations. Numerical simulations make far tails of the PDFs difficult to obtain, while analytical approximations are often inaccurate, as they neglect nonlinear interaction between pulses and noise.

Our approach combines the instanton method from statistical mechanics, to model far tails of the PDFs, with numerical simulations to refine the middle part of the PDFs. We combine the two methods by using an orthogonal polynomial expansion constructed specifically for this problem. We demonstrate the approach on an example of a specific submarine transmission system.

I. INTRODUCTION

Investigation of error statistics in high speed optical fiber communication systems is a fundamental task. The nonlinear nature of the propagation of light, nonlinear inter-symbol interference (ISI) between neighboring pulses coupled with noise makes this task very challenging.

With the first Viterbi equalizers operating at more than 10Gbps already available [1] and other advanced detection algorithms such as the BCJR algorithm being considered for high speed optical transmission, the need for an efficient methodology for approximating the probability density functions for energy of detected signals is even more important.

Due to ISI, it is necessary to study probability density functions (PDFs) of signal samples corresponding to longer bit configurations and not just individual bits. Existing approaches for modelling these PDFs rely either on extensive numerical simulations [2], [3] or on simplified and often inaccurate analytical approximations ([4] and references therein).

Numerical approaches approximate the middle part of distributions well, but far tails of PDFs are very hard to obtain numerically. Unfortunately, the tails of the PDFs are very important, since acceptable bit error-rates in the optical fiber transmission systems are low (\(\sim 10^{-9} - 10^{-12}\)), so error events fall into far tails of distributions.

All the existing analytical approaches neglect non-linear interaction between pulses and noise (which is implicitly incorporated in numerical modelling). This leads to PDFs approximations that are applicable only under very severe restrictions in terms of system speed, distance and types of fiber used (see [5], [6]).

Several recent papers [8], [9], [6], [12] used Karhunen-Loève series expansion (KL) to determine PDFs. In the cases when covariance function is known this method works well [6]. In the case of a general optical fiber communication system covariance between received pulses is not known and the KL expansion approach has several drawbacks. Firstly, a covariance matrix needs to be approximated numerically (in which case Karhunen-Loève expansion is in fact the Principle Component Analysis (PCA)) and numerical calculation of eigen-quantities is often unstable. Secondly, a separate set of eigen-quantities needs to be numerically calculated for each bit configuration (for more details about problems related with this approach see [13], [8], appendix of [6] and references therein).

In this paper, we use a method of optimal fluctuations, or instantons, to model far tails, and numerical simulations to refine the middle part of PDFs. We combine these two approaches by using Edgeworth expansion with orthogonal polynomials specially constructed for this problem.

The method of optimal fluctuations was originally developed in statistical physics [14]. This method allows one to reformulate the stochastic problem of pulse deterioration due to short correlated amplifier noise in terms of the deterministic problem of ”the most damaging” noise configuration. This approach is very similar to numerical saddle point approaches used in [10] and [11], but in contrast to those papers we calculate the saddle point, i.e., “the most damaging” noise configuration, analytically. The details are given in Section II.

Edgeworth expansion is a statistical method for approximating unknown PDFs [15]. An unknown distribution \(w(x)\) can be represented as \(w(x) = u(x) \sum_{i=0}^{\infty} C_i P_i(x)\), where \(u(x)\) is a starting approximate distribution and \(P_i(x), i \in N\) is a family of polynomials orthogonal with respect to the weight \(u(x)\). The coefficients \(C_i\) are derived from numerically (or experimentally) obtained moments of \(w(x)\). In practice it is sufficient to truncate the sum above to a certain (finite) number of terms. In the existing literature this method is almost always used with a Gaussian distribution as \(u(x)\) [15], probably because in this case it involves the widely known Hermite
polynomials as $P_1(x)$ (this special case of Edgeworth expansion is referred to as Gram-Charlier expansion). Recently, use of Gram-Charlier expansion was suggested [4] in the context of optical communication systems.

In this work we use PDFs derived by instanton method as $u(x)$. By using these PDFs, we obtain better asymptotic properties of the approximate distributions, which leads to faster and more accurate approximation of unknown PDFs. This requires derivation of a special family of orthogonal polynomials, which is given in Section III.

In Section IV we apply the proposed method to find PDFs for a system (operating at 40 Gb/s) that belongs to the class of dispersion-managed systems, where pieces of optical fiber with positive and negative dispersion are periodically incorporated. This is a single mode system with return-to-zero (RZ) pulses (similar calculations can be repeated for the NRZ format) and erbium-doped fiber amplifiers (EDFA). Data is encoded by the return-to-zero (RZ) pulses. This system has only one solution for $t, z$.

In this work we use PDFs derived by instanton method (optimal fluctuations [14]). Denoting

$$
\Phi = \left( -\frac{1}{2Nz} \int |y|^2 dt - i\lambda(E - \int_{-T/2}^{T/2} |A_s + y|^2 dt) \right)
$$

we have:

$$
\begin{align*}
\frac{\delta \Phi}{\delta \vartheta} &= 0 \\
\frac{\delta \Phi}{\delta \varphi} &= 0 \\
\frac{\delta \varphi}{\delta \varphi} &= 0 \\
\frac{\delta \varphi}{\delta \lambda} &= 0
\end{align*}
$$

This system of equations gives:

$$
\begin{align*}
&-\frac{1}{2Nz} \varphi^* + i\lambda(A_s^* + \varphi^*) = 0 \\
&-\frac{1}{2Nz} \varphi^* + i\lambda(A_s + \varphi) = 0 \\
&E - \int_{-T/2}^{T/2} |A_s + \varphi|^2 dt = 0
\end{align*}
$$

This system has only one solution for $\varphi$ given by

$$
\varphi = \frac{2i\lambda Nz}{1 - 2i\lambda Nz} A_s.
$$

This is the noise configuration that has the greatest influence on the energy of the received pulse. It is not surprising that it has the same shape as the pulse $A_s$.

Evaluating the expression $\Phi$ at the solution of the system (4), we obtain an estimate of the probability density function:

$$
P(E|s) \approx C \exp \left[ -\frac{1}{2Nz} \left( \sqrt{E - \int_{-T/2}^{T/2} |A_s(t, z)|^2 dt} \right)^2 \right]
$$

where $C$ is normalization constant.

Thus, the asymptotic behavior of PDFs is not Gaussian as often assumed in the existing literature. We note that this can be concluded from the study in [5] and it is explicitly stated.

II. PROBABILITY DENSITY FUNCTIONS OBTAINED BY INSTANTON METHOD

In this section we derive approximate PDFs for energy of the received pulse at the center slot of different bit patterns. This derivation follows the method derived in [20], and it is already presented in [16], but for the sake of completeness we shall repeat it here.

According to the central limit theorem, short correlated amplifier spontaneous emission (ASE) noise at the position $z$ in the moment $t$, $\xi(t, z)$, can be considered to be $\delta$-correlated Gaussian with zero mean (in which noise from EDFA is correlated is short even compared with the duration of the bit slot [17]):

$$
\langle \xi(t_1, z)\xi(t_2, z) \rangle = N \delta(t_1 - t_2)
$$

where $N$ is noise intensity and correlation is denoted by $\langle \rangle$.

Propagation of noise through the fiber is approximated by [19]:

$$
i\varphi_z(t, z) + u'(z) \varphi_z(t, z) = i\xi(t, z)
$$

(1)

where nonlinear interaction of the noise is neglected. The dispersion map function is $u'(z)$ and $\varphi(t, 0) = 0$.

The solution of Eq. (1) is

$$
\varphi(t, z) = \int_0^z \int_R G(t - s, y) \xi(s, y) ds dy
$$

where the Green’s function of the linear operator $i\varphi_z + u'(z) \varphi_z^2$ is given by

$$
G(t, z) = \exp \left( \frac{i\varphi^2}{4[u(z) - u(0)]} \right)
\sqrt{4\pi i[u(z) - u(0)]}
$$

Taking into account that

$$
\lim_{z_1 \rightarrow z_2} \exp \left( \frac{i(t_1 - t_2)^2}{4[u(z_1) - u(z_2)]} \right) = \delta(t_1 - t_2)
$$

we see that:

$$
\lim_{z_1 \rightarrow z_2} < \varphi(t_1, z_1), \varphi(t_2, z_2) > = N \delta(t_1 - t_2)
$$

The probability density function of the energy at the center slot is:

$$
P(E|s) = \delta(E - \int_{-T/2}^{T/2} |A_s(t, z) + \varphi(t, z)|^2 dt) >_\sigma
$$

where $T$ is the size of a time slot and $s$ is a bit configuration surrounding the center bit that gives $A_s(t, z)$ via non-linear interaction of the pulses in the absence of noise. Averaging $<>_\sigma$ is done over all noise configurations.

Taking Eq.(2) into account, Eq.(3) can be rewritten as:

$$
P(E|s) = \int d\lambda \int D\varphi \exp \left( -\frac{1}{2Nz} \int |\varphi|^2 dt \right) \\
\times \exp \left( -i\lambda(E - \int_{-T/2}^{T/2} |A_s(t, z) + \varphi(t, z)|^2 dt) \right)
$$

This integral can be estimated by evaluating it around its saddle point(s) (optimal fluctuations [14]). Denoting

$$
\Phi = \left( -\frac{1}{2Nz} \int |\varphi|^2 dt - i\lambda(E - \int_{-T/2}^{T/2} |A_s + \varphi|^2 dt) \right)
$$

we have:

$$
\begin{align*}
\frac{\delta \Phi}{\delta \varphi} &= 0 \\
\frac{\delta \Phi}{\delta \lambda} &= 0 \\
\frac{\delta \lambda}{\delta \varphi} &= 0
\end{align*}
$$

This system of equations gives:

$$
\begin{align*}
&-\frac{1}{2Nz} \varphi^* + i\lambda(A_s^* + \varphi^*) = 0 \\
&-\frac{1}{2Nz} \varphi^* + i\lambda(A_s + \varphi) = 0 \\
&E - \int_{-T/2}^{T/2} |A_s + \varphi|^2 dt = 0
\end{align*}
$$

This system has only one solution for $\varphi$ given by

$$
\varphi = \frac{2i\lambda Nz}{1 - 2i\lambda Nz} A_s.
$$

This is the noise configuration that has the greatest influence on the energy of the received pulse. It is not surprising that it has the same shape as the pulse $A_s$.

Evaluating the expression $\Phi$ at the solution of the system (4), we obtain an estimate of the probability density function:

$$
P(E|s) \approx C \exp \left[ -\frac{1}{2Nz} \left( \sqrt{E - \int_{-T/2}^{T/2} |A_s(t, z)|^2 dt} \right)^2 \right]
$$

where $C$ is normalization constant.

Thus, the asymptotic behavior of PDFs is not Gaussian as often assumed in the existing literature. We note that this can be concluded from the study in [5] and it is explicitly stated.
in [6]. Similar approximations for PDFs have been derived in different ways, see for example [3], [18].

At the end of this section, we would like to mention that the above method allows any noise statistics as long as it is uncorrelated. The expression \( \exp\left(\frac{1}{2\pi T} \int |\phi|^2 dt\right) \) in \( \Phi \) is a consequence of using Gaussian noise. Noise with different statistics will have a different expression other than \( \exp\left(-\frac{1}{2\pi T} \int |\phi|^2 dt\right) \) reflecting different density. If noise is correlated, expressions for partial derivatives of \( \Phi \) become integral equations that need to be solved numerically. Also, the same calculations can be repeated for the NRZ format, by numerically calculating energy accumulated at the center slot \( \int_{-T/2}^{T/2} |A_n(t, z)|^2 \) for bit configurations in this format.

### III. Edgeworth Expansion

In this section we use Edgeworth expansion [15] to refine PDFs derived in the previous section. An unknown distribution \( w(x) \) can be represented as

\[
w(x) = u(x) \left[ \sum_{i=1}^{\infty} C_i P_i(x) \right]
\]

where \( u(x) \) is a starting approximate distribution and \( P_i(x), n \in N \) is a family of polynomials orthogonal with respect to the weight \( u(x) \). Let \( P_i(x) = \sum_{k=1}^{\infty} a_k x^k \). By multiplying Eq. (6) by \( P_i(x) \) and integrating over the domain of orthogonality (in our case \( x \in (0, +\infty) \) since \( x \) represents energy) we get

\[
\sum_{k=1}^{i} a_k \eta_k = C_i
\]

where \( \eta_k, k \in N \) are moments of the distribution \( w(x) \). We can obtain a finite number of \( j \) of these moments numerically or experimentally, therefore, deriving an approximate PDF \( \tilde{w}_j(x) \) by truncating infinite sum Eq. (6) to \( j \) terms. These approximate PDFs, \( \tilde{w}_j(x), j \in N \) are guaranteed to converge to \( w(x) \) uniformly [15].

#### A. Orthogonal polynomials

In order to use PDFs given in Eq. (5) as starting distributions we need polynomials \( P_n(x; m; p), n \in N \) orthogonal with respect to the weight

\[
w(x; m; p) = e^{-m \sqrt{x-p}}, \quad x \in (0, +\infty) \quad (m, p > 0).
\]

These polynomials can be seen as generalized Laguerre polynomials and to the best of our knowledge have not been studied before. We shall briefly explain how to construct them; their properties will be studied in a separate publication, which is in preparation. The moments of distribution \( w(x; m; p) \) can be written as

\[
\mu_n(m; p) = \frac{2}{m^{n+1}} \int_{-p}^{\infty} (t + p)^{2n+1} e^{-t^2} dt.
\]

We denote

\[
\sigma_n(p) = 2 \int_{-p}^{\infty} (t + p)^n e^{-t^2} dt,
\]

and by partial integration rule, it can be shown that

\[
\sigma_{n+2}(p) = p \sigma_{n+1}(p) + \frac{n+1}{2} \sigma_n(p) \quad (n = 0, 1, \ldots),
\]

\[
\sigma_0(p) = \sqrt{\pi} \left(1 + \text{Erf}(p)\right), \quad \sigma_1(p) = e^{-p^2} + p \sigma_0(p)
\]

where Erf is the “error function” \( \text{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \).

Knowing all the moments \( \mu_n(m; p) \) we can calculate the polynomials by:

\[
P_n(x; m; p) = \frac{1}{\sigma_1^2(p)} \begin{vmatrix} \sigma_1(p) & \sigma_2(p) & \cdots & \sigma_{2n+1}(p) \\ \sigma_2(p) & \sigma_3(p) & \cdots & \sigma_{2n+2}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{2n+1}(p) & \sigma_{2n+3}(p) & \cdots & \sigma_{2n+2n+1}(p) \end{vmatrix} x^n
\]

Coefficients of these polynomials quickly become very cumbersome if expressed in general terms i.e. as functions of \( e^{-p^2} \) and \( \text{Erf}(p) \). However, for a given numerical value of \( p = \sqrt{\int_{-T/2}^{T/2} |A_n(t, z)|^2} \) there are no computational problems.

### IV. Numerical Results

In this section we shall illustrate the developed method. We considered the system in Fig. 1. As mentioned in the introduction, this system consists of periodically distributed sections of fiber with positive \( D_+ \) and negative dispersion \( D_- \) separated by amplifiers (EDFA). One span consists of one section of fiber with positive dispersion, one section of fiber with negative dispersion, and corresponding amplifiers (see Fig. 1).

![Fig. 1. Scheme of the optical transmission system considered](image)

The transmission of a signal through the fiber is modelled by the nonlinear Schrödinger equation (NLSE)

\[
\frac{\partial A}{\partial z} = -\frac{\alpha}{2} A - i \frac{\beta_3}{2} |A|^2 A + \frac{\beta_3}{6} (\frac{\partial^3 A}{\partial t^3}) + i \gamma |A|^2 A
\]

where \( z \) is the propagation distance along the fiber, relative time \( t = t_{\text{real}} - z/v_g \) gives a frame of reference moving at the group velocity \( v_g \). \( A(z, t) \) is the complex field amplitude of the pulse, \( \alpha \) is the attenuation coefficient of the fiber, \( u'(z) \) is the group velocity dispersion (GVD) coefficient, \( \beta_3(z) \) is the second-order GVD, \( \gamma \) is the nonlinearity coefficient giving rise to Kerr effect nonlinearities: self-phase modulation (SPM), intrachannel cross-phase modulation (IXPM) and intrachannel four-wave mixing (IFWM). In short, this calculation takes into account modulation, extinction ratio, realistic models of transmitter, optical filter and electrical filter, crosstalk effects, Kerr nonlinearities, ASE noise, and dispersion effects (GVD and second order GVD). In the system simulator, propagation
of pulses through the system, i.e. solving NLSE, was done numerically by the split-step Fourier method [19].

The parameters of positive dispersion $D_+$ and negative dispersion $D_-$ fibers are given in Table I. Pre-compensation of -330 ps/nm and corresponding post-compensation were also applied. The RZ modulation format has duty cycle of 33%, and the launched power was set to -6dBm. EDFAs with noise figure of 8dB were deployed after every fiber section, the bandwidth of optical filter was set to 3R$_b$ and the bandwidth of electrical filter to 0.65R$_b$, with R$_b$ being the bit rate (40 Gb/s).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$D_+$ fiber</th>
<th>$D_-$ fiber</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion [ps/(nm km)]</td>
<td>20</td>
<td>-40</td>
</tr>
<tr>
<td>Dispersion Slope [ps/(nm$^2$ km)]</td>
<td>0.06</td>
<td>-0.12</td>
</tr>
<tr>
<td>Effective Cross-sectional Area [$\mu$m$^2$]</td>
<td>110</td>
<td>50</td>
</tr>
<tr>
<td>Nonlinear refractive index [m$^2$/W]</td>
<td>$2.2 \times 10^{-20}$</td>
<td>$2.2 \times 10^{-20}$</td>
</tr>
<tr>
<td>Attenuation Coefficient [dB/km]</td>
<td>0.19</td>
<td>0.25</td>
</tr>
<tr>
<td>Length (in one span) [km]</td>
<td>33.4</td>
<td>16.7</td>
</tr>
</tbody>
</table>

We first ran the system simulator in the training mode (without noise) with random a bit sequence (of length $2^{17}$) to obtain values $E_s = \int_{-T/2}^{T/2} |A_s(t,z)|^2$ needed in the probability density functions Eq. (5).

The nonlinear distance of this system is roughly 6000km. As expected, for distances below this number the instanton approximation itself approximates the true PDF well. To illustrate this, in the Fig. 2, we plot both PDFs obtained by instanton method and histograms obtained numerically, for two bit configurations: (i) with zero in the center slot $s = 0110110$ and (ii) with “1” at the center slot $s = 0001000$. Both bit configurations were propagated through 100 spans, that is 5000km.

However, when the propagation distance is longer the instanton approximation is not sufficient. As Fig. 3 and Fig. 4 show, after 300 spans (approximately two and a half nonlinear distances) neglecting of nonlinear interaction between noise and pulses in the instanton approximation makes this approximation overly “optimistic”, i.e., too narrow. After polynomial correction is added, the refined PDFs show negligible differences to numerically obtained histograms. Note that only polynomials up to fourth order are needed to refine the PDF for “0110110”, and only first two polynomials are needed for “0001000”. This is in sharp difference with what is reported in [4], where, for a similar system, more than a dozen of Hermite polynomials are needed to sufficiently improve starting Gaussian distribution (also presented on figures).

![Fig. 2. Probability density functions after 100 spans](image)

![Fig. 3. Comparison of various PDF approximations for energy of bit pattern “0110110” after 300 spans](image)

![Fig. 4. Comparison of various PDF approximations for energy of bit pattern “0001000” after 300 spans](image)

Of course, question arises how many polynomials are needed to approximate an unknown distribution sufficiently well. From the engineering point of view, satisfactory answer...
can be to add higher order polynomials until the moment difference between two consecutive refinements falls under a certain threshold.

V. Final Remarks

We developed an approach for approximating probability density functions that is both practical and accurate. The instanton method gives right the asymptotic behavior for the tails of distributions. Use of the parameterized family of orthogonal polynomials reduces computational cost, making this approach applicable for high speed applications.

Method is also very general (it is not restricted to specific pulse shaping, bit rate, propagation distance) and therefore applicable to a wide range of systems.

Acknowledgments

Authors would like to acknowledge valuable discussions with Ildar Gabitov, Robert Indik and Misha Stepanov.

Los Alamos Report LA-UR-06-4518. This work was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396, and supported by DOE Office of Science Advanced Scientific Computing Research (ASCR) Program in Applied Mathematics Research. It is also funded in part by the NSF under the grant ITR 0325979.

References