PDL-impaired coherent optical communication systems: BER vs information-theory limit

Pierre Delesques,1,2,* Philippe Ciblat,2 Guillerm Froc,1 Yves Jaouën,2 and Cédric Ware2

1Mitsubishi Electric R&D Centre Europe, Rennes, France
2Institut Mines-Télécom/Télécom ParisTech; CNRS LTCI, Paris, France
*p.delesques@fr.merce.mee.com

Abstract: In coherent optical communications, PDL is one of the major linear impairments that can seriously impact systems performance. From an information-theoretic point of view, said PDL has to be analyzed through the so-called outage probability. After deriving this outage probability in closed-form expression for various statistical PDL models, we check these models’ accuracy numerically against a phenomenological one. Finally, we compare the outage probability to the BER of a simulated transmission system using polarization-time and LDPC coding schemes in simulation, especially in terms of SNR gaps and penalties.

© 2012 Optical Society of America

OCIS codes: (060.1660) Coherent Communications; (060.4510) Optical Communications.

References and links
1. Introduction

Coherent optical communication is a key enabling technique to satisfy the high data rate requirements in future core networks. Thanks to high-speed electronic devices, coherent detection can now be implemented at very high bit rates (e.g. [1, 2]), enabling the use of digital signal processing techniques against the impairments that the signal undergoes through its physical propagation in fiber and other optical devices.

These impairments can be categorized into two classes: linear and nonlinear. The major linear impairments are chromatic dispersion (CD), polarization mode dispersion (PMD), and polarization dependent loss (PDL). Against CD and PMD, the equalization principle is very effective in a single-carrier context (see [3–5]); better, in multi-carrier systems, they have virtually no impact on performance, since the matrices corresponding to CD and PMD are unitary in an inherently spectral-domain context. Furthermore, under the standard low-power regime assumption, corresponding to the operating points of deployed systems, linear impairments are predominant, allowing us to neglect nonlinear effects in this paper; for more details about the impact of nonlinear effects in the presence of PDL, the reader may refer to [6].

Therefore, this paper will focus on PDL as the main limiting linear impairment. Our goal is twofold: first, to determine the impact of PDL from an information-theoretic point of view; and, second, to compare the resulting fundamental limit to simulated transmission systems performance.

In optical systems, PDL comes from the small anisotropy induced by non-ideal physical elements such as fiber slicing and optical components [7]. These elements are spread along the optical link and the contribution of each element can change unpredictably over time. Consequently, PDL may be modeled via a random process; its impact would then be most naturally studied using the outage probability [8]. Although several statistical models have been proposed in the literature [9–12], and some work done towards numerical evaluation of the outage probability [13], we have not found it as closed-form expressions so far. Section 2 will be devoted to deriving such expressions for various PDL models, which will then be compared numerically to a phenomenological model in section 3.

We shall next focus on simulated PDL-impaired transmission systems, whose BER performance can be compared to a fundamental limit governed by the outage probability. We have selected coding schemes using both Polarization-Time (PT) codes and Low-Density Parity Check (LDPC) codes. PT codes are space-time codes applied to polarization-multiplexed systems, and
have been widely advocated for PDL mitigation: spreading the information symbols over time and both polarizations [14–17] partly counteracts the uneven energy repartition amongst the polarizations. An LDPC code complements them to further improve BER performance, which is numerically evaluated in section 4 and compared to the corresponding outage probability.

2. Outage probability derivations

Assuming a polarization division multiplexing (PDM) based coherent optical communications scheme only disturbed by the linear impairments, the bivariate received signal, denoted by $Y(\omega)$ at pulsation $\omega$, can be written as follows

$$Y(\omega) = H(\omega)X(\omega) + N(\omega)$$

where $X(\omega)$ is the transmitted signal, and $N(\omega)$ an additive white Gaussian noise (AWGN) with variance $N_0$ per real dimension [8]. The channel transfer matrix $H(\omega)$ corresponds to the concatenation of the linear impairments, namely CD, PMD, and PDL.

We remind that the instantaneous channel capacity associated with one channel realization $H(\omega)$, denoted by $C(H)$, takes the following shape

$$C(H) = \frac{1}{2\pi} \int_{-\pi B}^{\pi B} \log_2 \left( \det \left[ I_2 + \rho H(\omega)H(\omega)^H \right] \right) d\omega$$

when the input $X(\omega)$ is Gaussian-distributed [8]. The superscript $(.)^H$ stands for the transpose-conjugate operator, $I_2$ is the 2-by-2 identity matrix, $B$ is the transmitted signal bandwidth, and $\rho = E_s/2N_0$ (i.e. the ratio between the energy per received symbol and the noise energy per polarization).

As $H(\omega)$ can be written as the concatenation of transfer matrices related to the different impairments (CD, PMD, PDL) [16] and as the transfer matrices of CD an PMD are unitary, the instantaneous capacity can be simplified as follows

$$C(H) = \frac{1}{2\pi} \int_{-\pi B}^{\pi B} \log_2 \left( \det \left[ I_2 + \rho H_{PDL}(\omega)H_{PDL}(\omega)^H \right] \right) d\omega$$

where $H_{PDL}(\omega)$ represents the transfer matrix induced by the PDL.

As defined in [8], the outage probability is the probability that the instantaneous channel capacity $C(H)$ (i.e. for one channel realization) is below a given transmission rate $r$ (expressed in bits/s)

$$P_o = \text{Prob}\{C(H) < r\}.$$  \hspace{1cm} (2)

Our objective is thus to exhibit closed-form expressions for $P_o$ when various PDL statistical models are considered.

Although PDL is generated by numerous elements spread along the fiber, it is usual to model the PDL through an unique “equivalent” element [7, 18]. The validity of this assertion will be tested in section 3; in the meantime, we shall admit this assertion and consider

$$H_{PDL} = R_\alpha D_\gamma R_\beta$$  \hspace{1cm} (3)

where

$$D_\gamma = \begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & \sqrt{1+\gamma} \end{bmatrix},$$

and where $R_\alpha$ and $R_\beta$ are two random 2-by-2 rotation matrices of angles $\alpha$ and $\beta$ respectively. The term $\gamma$ corresponds to the mismatch between both polarizations and belongs to the interval
− 1, 1[. We also define the so-called PDL coefficient as
\[
\Gamma = 10 \log_{10} \left( \frac{1 + \gamma}{1 - \gamma} \right)
\] (4)
(which is thus inherently expressed in dB).

Under this hypothesis, the capacity can be readily calculated: first, \( H_{\text{PDL}} \) does not depend on \( \omega \), implicitly assuming that the signal is narrowband enough; the integral over the bandwidth is thus trivial; then, by replacing \( \mathbf{I}_2 \) with \( \mathbf{R}_\alpha \mathbf{R}_\alpha^T \) and using the commutativity of the determinant, Eq. (1) becomes:
\[
C(\mathbf{H}) = B \log_2 (\det (\mathbf{I}_2 + \rho \mathbf{D}_\gamma^2)) = B \log_2 \left( (1 + \rho)^2 - \rho^2 \gamma^2 \right)
\]
and the outage probability given in Eq. (2) becomes:
\[
P_o = \text{Prob} \{ \log_2 \left( (1 + \rho)^2 - \rho^2 \gamma^2 \right) < R \}
\]
where \( R = r/B \) is the spectral efficiency.

As a consequence, the outage probability only depends on the probability density function (pdf) of \( \gamma \) in the following way
\[
P_o = \text{Prob} \{ \gamma^2 > f(R, \rho) \} \text{ with } f(R, \rho) = 1 - \frac{1}{\rho^2}(2^R - 1 - 2\rho)
\] (5)

2.1. PDL statistical models

Several works have been devoted to the modeling of \( \gamma \) (or, equivalently, of \( \Gamma \)), most of which are listed below:

- **\( \gamma \) constant** [16]: this is the simplest one and has been introduced for a rough analysis of the influence of the PT codes on the overall system performance. The constant value is denoted by \( \gamma = \mu_c \).

- **\( \gamma \) Gaussian** [7, 9]: this model has been roughly validated, for low values of the PDL, against a concatenation of many elementary PDL slices. The probability density function is
\[
p_{\gamma}^{(g)}(x) = \frac{1}{\sqrt{2\pi \sigma_g^2}} \exp \left[ -\frac{(x - \mu_g)^2}{2\sigma_g^2} \right]
\]
with \( \mu_g = \mathbb{E}[\gamma] \) and \( \sigma_g^2 \) the variance.

This model has many unfortunate drawbacks: it is quite realistic only in a few situations; the link between \( \mu_g, \sigma_g^2 \) and the mean PDL (given by \( \mathbb{E}[\Gamma] \)) is intractable; and \( \gamma \) may be out of the interval \(-1, 1\[, \) which does not have a physical meaning. In our simulation setup, we have fixed \( \mu_g \) so that \( 10\log_{10}(1 + \mu_g)/(1 - \mu_g) \) is equal to \( \mathbb{E}[\Gamma] \), then we have built a lookup table to find the relationship between \( \sigma_g^2 \) and \( \mathbb{E}[\Gamma] \).

- **\( \gamma \) truncated Gaussian**: this model, introduced in this paper, simply ensures that the previous Gaussian model keeps \( \gamma \) in the interval \(-1, 1\[. \) Therefore the Gaussian distribution is truncated as follows
\[
p_{\gamma}^{(t)}(x) = \begin{cases} 
\frac{K}{\sqrt{2\pi \sigma_g^2}} \exp \left[ -\frac{(x - \mu_g)^2}{2\sigma_g^2} \right], & |x| < 1 \\
0 & \text{otherwise}
\end{cases}
\]
where

\[
K = \frac{1}{1 - Q\left(1 + \frac{\mu_g}{\sigma_g}\right) - Q\left(1 - \frac{\mu_g}{\sigma_g}\right)}
\]

is a normalization coefficient ensuring that \( p_{\Gamma}^{(\gamma)} \) is a pdf. Function \( x \mapsto Q(x) \) stands for the Gaussian tail function.

Then we have chosen \( \mu_g = \mu_g \) where \( \mu_g \) is given as above, and \( \sigma_g \) is obtained thanks to a similarly-built lookup table.

• \( \Gamma \) (positive) Gaussian [10]: by construction, if \( \Gamma \) is defined as the ratio (in dB) between the maximum power over the minimum power on the eigenvectors, then \( \Gamma \) is positive. After some practical measurements, [10] has observed that a Gaussian distribution (actually its right part on the x-axis) fits roughly well with their empirical histogram of \( \mathbb{E}[\Gamma] \). Therefore, they have suggested the following probability density function for \( \Gamma \)

\[
p_{\Gamma}^{(\gamma)}(x) = \begin{cases} 
\frac{2}{\sqrt{2\pi}\sigma_{\gamma}} \exp\left[-\frac{x^2}{2\sigma_{\gamma}^2}\right], & x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

After some algebraic manipulations, we show that \( \sigma_{\gamma}^2 \) has to satisfy \( \sigma_{\gamma}^2 = \left(\frac{\sqrt{2}}{\pi}\right) \mathbb{E}[\Gamma] \).

We note that an alternative choice, of also including the negative part of the Gaussian distribution, can be encountered in the literature: in [19], \( \Gamma \) is modeled as a Gaussian-distributed random variable with a zero mean and a variance equal to \( \mathbb{E}[\Gamma] \). The choice of those values is not justified, and seems somewhat arbitrary. Therefore, in the following, this alternate Gaussian model is not pursued.

• \( \Gamma \) Rayleigh [11]: this statistical model for \( \Gamma \) has been obtained by concatenating “small” PDL elements without taking into account the birefringence. Then the probability density function is given by the Rayleigh distribution as follows

\[
p_{\Gamma}^{(r)}(x) = \begin{cases} 
\frac{x}{\sigma_r^3} \exp\left[-\frac{x^2}{2\sigma_r^2}\right], & x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

where \( \sigma_r = \sqrt{\frac{2}{\pi}} \cdot \mathbb{E}[\Gamma] \).

• \( \Gamma \) Maxwellian [12]: the most sophisticated model, well-supported by mathematical calculations and taking the birefringence into account, considers that \( \Gamma \) is distributed according to a Maxwell-Boltzmann distribution given by

\[
p_{\Gamma}^{(m)}(x) = \begin{cases} 
\frac{\sqrt{2}}{\pi} \frac{x^3}{\sigma_m^4} \exp\left[-\frac{x^2}{2\sigma_m^2}\right], & x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

where \( \sigma_m = \sqrt{\frac{\pi}{8}} \cdot \mathbb{E}[\Gamma] \).

Based on Eq. (5), we can find closed-form expressions for all the above-mentioned PDL models. For brevity’s sake, we will hereafter only sketch the derivations for the two most realistic models: on the one hand, when \( \gamma \) is truncated Gaussian, and on the other hand when \( \Gamma \) is Maxwellian. The results for the other models are reported (without derivation details) in Table 1. Part of this work was presented in [20].
2.2. Derivations for the γ truncated Gaussian model

Let us start with the two trivial cases of $f(R, \rho) > 1$ and $f(R, \rho) < 0$, respectively equivalent to $\rho > (2^R - 1)/2$ and $\rho < \sqrt{2^R} - 1$. In the first case, since $\gamma$ lies in the interval $[-1,1]$, it is impossible to have $\gamma^2 > f(R, \rho) > 1$; given Eq. (5), the outage probability vanishes. Conversely, if $f(R, \rho)$ is negative, $\gamma^2$ is always greater, and the outage probability is 1. To summarize:

$$\begin{cases} 
\text{If } \rho > \rho_a, & \text{with } \rho_a = \frac{1}{2}(2^R - 1), \text{ then } f(R, \rho) > 1, \text{ so } P_o = 0 \\
\text{If } \rho < \rho_c, & \text{with } \rho_c = \sqrt{2^R} - 1, \text{ then } f(R, \rho) < 0, \text{ so } P_o = 1
\end{cases} \quad (6)$$

The remaining case corresponds to $\rho \in [\rho_c, \rho_a]$, which is equivalent to $f(R, \rho) \in [0,1]$, which allows us to decompose the outage probability as

$$P_o = \int_{-1}^{\sqrt{f(R, \rho)}} p_\gamma^{(f)}(x)dx + \int_{\sqrt{f(R, \rho)}}^{1} p_\gamma^{(f)}(x)dx$$

leading to the following result

$$P_o = K \left[ Q\left(\sqrt{f(R, \rho)} + \frac{\mu_g}{\sigma_g}\right) + Q\left(\frac{\sqrt{f(R, \rho)} - \mu_g}{\sigma_g}\right) - Q\left(1 + \frac{\mu_g}{\sigma_g}\right) - Q\left(1 - \frac{\mu_g}{\sigma_g}\right) \right]$$

We have thus obtained a closed-form expression of the outage probability in the $\gamma$ truncated Gaussian case and the final result is reported in Table 1.

2.3. Derivations for the $\Gamma$ Maxwellian model

Here we have to replace the inequality on $\gamma$ with an equivalent inequality on $\Gamma$. However, we are still assuming that $\gamma \in [-1,1]$ (implicit in the definition of $\Gamma$), which means that Eq. (6) still holds. This leaves us to focus on the case $\rho \in [\rho_c, \rho_a]$, that is, $f(R, \rho) \in [0,1]$. Moreover, $\Gamma$ being positive under the Maxwellian model, so is $\gamma$ according to Eq. (4). We can thus rewrite Eq. (5) by replacing inequality $\gamma^2 > f(R, \rho)$ with $\gamma > \sqrt{f(R, \rho)}$.

From there, remarking that Eq. (4) can also be written as $\Gamma = \frac{20}{\log(10)} \text{artanh}(\gamma)$, and that the function $\text{artanh}$ is strictly increasing, we simply obtain

$$P_o = \text{Prob} \{ \Gamma > T(R, \rho) \} = \int_{T(R, \rho)}^{\infty} p_\Gamma^{(m)}(x)dx$$

with

$$T(R, \rho) = \frac{20}{\log(10)} \text{artanh} \left(\sqrt{f(R, \rho)}\right).$$

The cumulative density function (cdf) of the Maxwell-Boltzmann distribution being well-known, we obtain the following closed-form expression for the outage probability:

$$P_o = 2Q\left(\frac{T(R, \rho)}{\sigma_m}\right) + \sqrt{\frac{2}{\pi}} T(R, \rho) e^{-\frac{T(R, \rho)^2}{2\sigma_m^2}}.$$
Outage probability

Consider the total received power in the frequency domain (by considering the power on both polarizations), given by

\[
\text{PDL} = \text{mean of } \gamma
\]

Let us consider a transfer matrix \(H\), as in [12], and its variance is

\[
\sigma_{\gamma}^2 = \frac{1}{N} \sum_{i=1}^{N} \left( R_{\gamma_i} \right) \left( D_{\gamma_i} \right) \left( B_{\phi_i} \right)
\]

where \(R_{\gamma_i}\) and \(D_{\gamma_i}\) are 2-by-2 rotation and PDL matrices as above, and \(B_{\phi}\) is a 2-by-2 birefringence diagonal matrix with diagonal \([\exp(-i\phi), \exp(i\phi)]\).

We still assume that we can neglect the variation of PDL over the bandwidth of the signal. We consider that the rotation angles \(\gamma_i\) and the phases \(\phi_i\) are independent and uniformly distributed in \([0, 2\pi]\), and that \(\gamma_i\) is identically distributed according to a truncated Gaussian distribution. The mean of \(\gamma_i\) is set to \(\sqrt{9/(\pi N)} \log(1 + \log(10)/30)\), and its variance is chosen such that the target mean PDL \(\text{E[\gamma]}\) is reached.

This matrix cannot be used directly in place of \(H_{\text{PDL}}\) in Eq. (1), because it does not conserve power: consider the total received power in the frequency domain (by considering the power on both polarizations), given by

\[
P_T = \text{E}[\text{Y}^H(Y)]
\]

Assuming independent symbols in \(X(\omega)\),

<table>
<thead>
<tr>
<th>Model</th>
<th>Outage probability</th>
</tr>
</thead>
</table>
| \(\gamma\) constant         | \[
1 \quad \text{if } \rho < \frac{\sqrt{1-(\mu_1^2)(\mu_2^2)-1}}{\mu_2^2} \\
0 \quad \text{otherwise}
\] |
| \(\gamma\) Gaussian          | \[
Q \left( \frac{\sqrt{T(R, \rho)+\mu_x}}{\sigma_x} \right) + Q \left( \frac{\sqrt{T(R, \rho)-\mu_x}}{\sigma_x} \right) \\
\text{if } \rho < \rho_l \\
K \left[ Q \left( \frac{\sqrt{T(R, \rho)+\mu_x}}{\sigma_x} \right) + Q \left( \frac{\sqrt{T(R, \rho)-\mu_x}}{\sigma_x} \right) \right] - Q \left( \frac{\mu_x}{\sigma_x} \right) - Q \left( \frac{-\mu_x}{\sigma_x} \right) \\
\text{if } \rho_l < \rho < \rho_u \\
0 \quad \text{otherwise}
\] |
| \(\gamma\) truncated Gaussian| \[
2Q \left( \frac{T(R, \rho)}{\sigma_y} \right) \\
\text{if } \rho < \rho_l \\
\exp \left[ -\frac{T(R, \rho)}{2\sigma_y^2} \right] \\
\text{if } \rho_l < \rho < \rho_u \\
0 \quad \text{otherwise}
\] |
| \(\Gamma\) positive Gaussian  | \[
2Q \left( \frac{T(R, \rho)}{\sigma_y} \right) + \sqrt{\frac{2}{\pi\sigma_y}} T(R, \rho) \exp \left[ -\frac{T(R, \rho)^2}{2\sigma_y^2} \right] \\
\text{if } \rho < \rho_l \\
\text{if } \rho_l < \rho < \rho_u \\
0 \quad \text{otherwise}
\] |

Table 1. Outage probability closed-form expressions for different PDL models

3. Comparison of statistical and phenomenological PDL models

As noted in section 2, statistical models based on Eq. (3) do not reflect the way in which the actual physical effect appears. Therefore, we have numerically evaluated the outage probability in a phenomenological PDL model inspired by [9, 12] in order to estimate the validity of the statistical models in this respect.

3.1. Phenomenological PDL model

Let us consider a transfer matrix \(H_{\text{PDL}}\) that consists in the concatenation of \(N\) elementary PDL slices:

\[
\hat{H}_{\text{PDL}} = \prod_{i=1}^{N} \left( R_{\alpha_i} D_{\gamma_i} B_{\phi_i} \right)
\]
and a noiseless, PDL-only case, we obtain

\[ P_y = \text{Tr}(H_{\text{PDL}} H_{\text{PDL}}^H) P_x \]

where \( \text{Tr} \) is the trace operator, and \( P_x \) is the power transmitted on each polarization.

According to Eq. (3), the theoretical model leads to \( \text{Tr}(H_{\text{PDL}} H_{\text{PDL}}^H) = 2 \) for any PDL value \( \gamma \). In contrast, \( \text{Tr}(\tilde{H}_{\text{PDL}} \tilde{H}_{\text{PDL}}^H) \) is not constant and even depends on the realization of the elementary PDL slices, as can be observed in Fig. 1 where the histogram of \( \text{Tr}(\tilde{H}_{\text{PDL}} \tilde{H}_{\text{PDL}}^H) \) has been displayed for two values of mean PDL and with \( N = 100 \).

This is not surprising since PDL, as a physical phenomenon, is lossy. However, Eq. (1) assumes that all power loss and noise considerations be taken into account through parameter \( \rho \), and \( H_{\text{PDL}} \) must only reflect the power mismatch between the different polarizations.

Therefore, we suggest to normalize the matrix \( \tilde{H}_{\text{PDL}} \):

\[ H_{\text{PDL}} = \tilde{H}_{\text{PDL}} \cdot \frac{\sqrt{2}}{\| \tilde{H}_{\text{PDL}} \|_F} \]

where \( \| A \|_F^2 = \text{Tr}(A A^H) \) for any square matrix \( A \) (Frobenius norm).

On a related note, [21] points out that PDL has different effects on signal and noise in systems with multiple amplified spans, because each amplifier adds noise that has not propagated in the spans preceding said amplifier. Consequently, the noise variance is time-varying inducing a Signal-to-Noise Ratio (SNR) compression, and the noise is correlated. In our case, however, the SNR compression entailed by this phenomenon is also assumed to be accounted for in \( \rho \). In addition, we have observed that the noise correlation induced by the PDL has a very small impact on the outage probability and thus can be neglected without loss of generality.

There remains to define a value of \( \Gamma \) for the resulting matrix, which can be done by analogy to the theoretical model: given that \( H_{\text{PDL}} \) as defined in Eq. (3) has eigenvalues \( \sqrt{1 + \gamma} \) and \( \sqrt{1 - \gamma} \), we note that Eq. (4) defines \( \Gamma \) as the ratio of these eigenvalues (the matrix’s so-called condition number), squared and expressed in dB. Likewise, we propose to identify the mean PDL for the phenomenological model as the square of the condition number of \( H_{\text{PDL}} \).

The outage probability of this phenomenological model could only be evaluated numerically, as done in the next section.

3.2. Numerical illustrations for the outage probability

In Fig. 2, we plot the theoretical outage probabilities for the PDL models considered, and the empirical one for the phenomenological model versus \( E_b/N_0 \) (equal to \( 2\rho/R \)) for a mean \( \Gamma \).
equal to 3 dB (left) and 6 dB (right). The spectral efficiency is $R = 4 \text{ bits/s/Hz}$, and the number of slices in the phenomenological model is $N = 100$.

As expected, the $\gamma$ Gaussian model becomes useless when the mean PDL becomes too high. Except for the $\gamma$ Gaussian model and the $\gamma$ constant model, all the curves drop to zero when $E_b/N_0$ reaches 5.74 dB, as predicted by Eq. (6). We remark that the $\gamma$ constant model is quite far away from the other models and thus is clearly inadequate. Finally, the $\Gamma$ Maxwellian model is the closest one to the phenomenological model and so is the most relevant (as already asserted in [12]), and will be next analyzed more deeply.

Before going further, we introduce the so-called SNR penalty (on the outage probability) as follows: the SNR penalty is defined as the ratio between the value of $E_b/N_0$ needed to yield a given outage probability over the $E_b/N_0$ needed for the same outage probability in a PDL-free transmission. In Fig. 3, we plot the SNR penalty for the Maxwellian model versus the mean PDL $\mathbb{E}[\Gamma]$ for different outage probability targets. On the left, $R = 4 \text{ bits/s/Hz}$; on the right, $R = 20 \text{ bits/s/Hz}$.

We notice that for standard PDL values, around 3 dB, the penalty is roughly between 1 and 3 dB. The exact value depends slightly on the outage probability target and strongly on the spectral efficiency $R$. Nevertheless, the SNR penalty is upper-bounded whatever the mean PDL.
and the outage probability target. This upper bound can be found using Eq. (6): first, in a PDL-
free situation ($\gamma = 0$), the SNR required to achieve any outage probability below 1 corresponds
to $\rho = \rho_t$, since $P_o$ drops from 1 to 0 at this point; then, however high the PDL is, the same
outage probability can be matched at $\rho < \rho_u$, since $P_o$ always drops to 0 at $\rho = \rho_u$. The SNR
penalty is thus always below $\rho_u/\rho_t = (2^R - 1)/(2(\sqrt{2^R} - 1))$, which is confirmed by Fig. 3.

4. Performance of simulated communication systems

Having determined the outage probability in the presence of PDL now raises the question of
how well simulated communication systems fare against this fundamental limit, as the out-
age probability is known to be a lower-bound of the BER [8]. In this section, we perform the
comparison using a simulated PDL-impaired system using PT and LDPC codes.

4.1. System setup

In our simulation setup, we have only considered the linear impairments (CD, PMD, PDL).
As CD and PMD can be completely removed by implementing a well-designed Orthogonal
Frequency Division Multiplexing (OFDM) transmission scheme, an OFDM scheme is thus
carried out.

Consequently, our communication scheme boils down to a PDM system disturbed only by
PDL, and is well described by

$$Y(\omega) = H_{PDL}X(\omega) + N(\omega).$$

where all these terms have been defined in Section 2.

As described in Fig. 4, we thus have considered a PDM-based OFDM transmission scheme
assuming that the channel transfer matrix $H_{PDL}$ on each subcarrier is perfectly known at the
receiver side.

![Fig. 4. Architecture of the simulated system.](image)

At the transmitter side, the bit-stream is first coded using a forward error correction (FEC)
code. The specific code chosen is an LDPC code based on the quasi-cyclic progressive edge
growth algorithm (QC-PEG): LDPC codes are very powerful [22] and QC-PEG is particu-
larly well adapted to coherent optical communications [19]. The codeword then passes through
an interleaver in order to scramble the bits with respect to time and polarization, increasing
the robustness to burst errors. Next, the interleaved codewords are modulated using Quadra-
ture Phase-Shift Keying (QPSK). The modulated symbols are further precoded thanks to a
PT code, which spreads the information over time and polarization, significantly improving
the performance of PDL-disturbed systems [14–17]. Finally, the PT symbol is spread over the
OFDM subcarriers; this last spreading does not affect performance since the PDL channel is not frequency-selective.

The channel is assumed to be invariant over the duration of a LDPC codeword. This assumption is not restrictive since PDL (as well as PMD) is a slow time-varying phenomenon.

At the receiver side, after the OFDM demodulator, we have implemented two decoding schemes:

- A “Hard” decoding approach for which hard decisions are taken at each step: specifically, a hard Maximum-Likelihood (ML) decoder for the PT code is implemented, then a hard decision is provided to the symbol demodulator, and so on.

- A “Soft” decoding approach for which the hard decision is taken only on the last step and soft decisions are propagated along the decoding process. In this case, the ML PT decoder and demodulator are done jointly and yield the so-called Log-Likelihood Ratio (LLR) for each received bit; then the LLR is the input for the LDPC decoder, which in turn is implemented via the Sum-Product Algorithm [22].

According to coding theory, it is well known that the soft approach is much more efficient than the hard one at the expense of complexity. The following configurations have been implemented:

- QPSK only (no FEC, no PT),
- PT only (Silver or Golden codes),
- LDPC only (with either hard or soft decision),
- PT+LDPC with hard decoding,
- PT+LDPC with soft decoding.

For each configuration, we sought a useful data rate of at least 100 Gbit/s over a standard 50-GHz ITU channel. Allowing for some filtering, we considered an effective bandwidth of 40 GHz. The combination of QPSK and PDM has a spectral efficiency of 4 bits/s/Hz, thus a useful data rate of 160 Gbit/s, in the no-FEC case; this is valid whether or not PT coding is used, as the PT codes considered (Silver and Golden) are full-rate and have no impact on the spectral efficiency. In the case with FEC, the LDPC we use has a coding rate \( r_{LDPC} = \frac{3}{4} \), which yields a spectral efficiency equal to 3 bits/s/Hz, and the useful data rate becomes 120 Gbit/s.

4.2. Performance of simulated systems vs outage probability

The performance of the simulated systems introduced above will now be compared to the fundamental limit yielded by the outage probability. Given the results of section 3.2, we will only consider the \( \Gamma \) Maxwellian and phenomenological PDL models. The spectral efficiency \( R \) is taken to be \( R = 4 \) bits/s/Hz (without FEC) or \( R = 3 \) bits/s/Hz (with LDPC). The number of PDL elements in the phenomenological model is set to \( N = 100 \).

In Fig. 5, we plot the outage probabilities and BER versus SNR for the Maxwellian model (on the top) and for the phenomenological one (on the bottom) with a mean PDL equal to a standard value of 3 dB. First of all, we remark that the best configuration (PT+LDPC with soft decoding) is quite close to the fundamental limit, namely, around 1.5 dB. Consequently, even though optimizing PT and LDPC codes for PDL mitigation could be of some interest, the existing ones already yield remarkable performance. In addition, we observe that the gains offered by the PT and LDPC codes are cumulative, and thus both coding techniques have to be employed for combating the PDL efficiently. Also, unlike in wireless communications, the
Silver PT code is as good as the Golden PT code, as already noticed in [15, 17]. Finally, soft decoding dramatically improves the performance compared to hard decoding.

Fig. 5. Outage probability and BER with or without FEC for Maxwellian model (top) and phenomenological model (bottom). Mean $\Gamma \sim 3$ dB

To explore other values of the mean PDL, we use the so-called SNR gap, defined for a given outage probability as: the ratio between the SNR required for the considered configuration to yield a BER equal to the target outage probability, over the SNR needed to ensure the same outage probability. In other words, the SNR gap corresponds to the “distance” in SNR between the performance of the considered simulated configuration and the fundamental limit.

In Fig. 6, we plot the SNR gap versus the mean PDL for the Maxwellian model (on the left) and for the phenomenological one (on the right), with an outage probability target fixed to $10^{-7}$ in both cases.

We first remark that the LDPC code is not robust to high mean PDL whereas the PT code enables us to keep the SNR gap small enough even for high mean PDL. So the PT code is much more efficient for mitigating the PDL, but concatenating both coding techniques (especially when soft decoding is considered) is very relevant since the SNR gap is less than 6 dB for
very high mean PDL whatever the PDL model. For usual mean PDL values (around 3 dB), we observe that the minimum SNR gap is only 1.5 dB.

Finally, in Fig. 7, we plot the SNR Penalty versus the mean PDL for the Maxwellian model (on the left) and for the phenomenological one (on the right). The SNR penalty is defined here, for each simulated configuration, as the ratio between the value of $E_b/N_0$ needed to achieve a given BER over that needed for the same BER in a PDL-free transmission.

We can see that, for a mean PDL of 3 dB, the SNR penalty is quite small (around a few dB) as soon as PT codes are involved; overall, the SNR penalty with PT coding is on the same order as that of the outage probability. In contrast, when only LDPC codes are used, the SNR penalty increases rapidly and becomes huge. Therefore, the use of PT codes is necessary, and concatenating them with an LDPC code further improves performance, bringing it fairly close to the fundamental limit.

5. Conclusion

We have studied the fundamental performance limit for coherent optical communications in the presence of PDL, which is one of the major impairments in such systems. This limit, represented by the outage probability, was derived in closed-form expressions for some PDL models;
we then used BER simulations to evaluate how far from said limit simulated systems could be. We have compared the outage probability expressions under the usual PDL models with the numerical evaluation of that associated with a phenomenological model. We have observed that the $\Gamma$-Maxwellian model, already known to be the best-justified physically, also fits quite well with the phenomenological model from an outage probability standpoint. Other models are inaccurate to various degrees, against which their simplicity must be weighed. Additionally, we have found a bound on SNR above which PDL can in theory be completely mitigated.

The outage probability has also been compared numerically to the BER offered by simulated systems using QPSK modulation and several coding schemes: LDPC and/or PT Codes, or no coding. These simulations have shown that, with the codes we used, PT codes are more robust to PDL than LDPC, and that the concatenation of a PT code with an LDPC is even better. Indeed, we have found that for a mean PDL of 3 dB, a simulated system based on PT and LDPC codes (with soft decoding) performs with an SNR gap of only 1.5 dB to the fundamental limit.

While more powerful coding schemes could undoubtedly edge closer to the limit, the modest 1.5 dB gain that can be hoped for must be balanced against the computational complexity, which is already the major drawback of LDPC soft decoding, especially at the very high bit rates of modern optical communication systems. Nevertheless, other options for LDPC codes could be investigated. The question also remains of whether this result holds with higher-order modulation formats.

However, the performance of PT codes hints that an interesting way for further study would be whether improved PT codes might alleviate the need for powerful LDPC codes, allowing near-limit performance with less-expensive codes.