The Enhanced Gaussian Noise Model Extended to Polarization-Dependent Loss

The Enhanced Gaussian Noise Model extended to Polarization-dependent Loss

Paolo Serena, *Member, IEEE,* Chiara Lasagni, *Student Member, IEEE,* and Alberto Bononi, *Senior Member, IEEE*

***Abstract*—We show how to extend the enhanced Gaussian noise (EGN) model to account for polarization-dependent loss (PDL) of optical devices placed along a fiber-optic link. We provide a comprehensive theory highlighting the relationships between the time, frequency, and polarization domains in the presence of fiber nonlinear Kerr effect and amplified spontaneous emission. We double-check the new model with split-step Fourier method (SSFM) simulations showing very good accuracy. The model can be efficiently exploited to estimate low values of outage probabilities induced by PDL with computational times orders of magnitude faster than the SSFM, thus opening new opportunities in the design of optical communication links.**

***Index Terms*—Polarization-dependent loss (PDL), Gaussian-noise (GN) model, enhanced-GN (EGN) model.**

I. Introduction

POLARIZATION-dependent loss (PDL) expresses the dependence of the loss of an optical device on the state of polarization of the input electromagnetic field [1]. PDL induces crosstalk between the polarization tributaries and an unequal loss of energy, which are particularly detrimental in polarization-division multiplexing (PDM) transmissions. Contrary to chromatic dispersion and polarization-mode dispersion (PMD), the non-unitary PDL distortion cannot be removed without a penalty even by the best detector [2], [3].

Although typical optical fibers show negligible PDL, PDL may be relevant in optical devices, such as the Erbium-doped fiber amplifiers (EDFA) and inside the wavelength selective switches (WSS) of reconfigurable optical add/drop multiplexers (ROADM) [4].

The axes of maximum/minimum PDL fluctuate randomly over times much longer than the coded data block duration, thus making the optical channel stochastic and non-ergodic. Hence the analysis with PDL should not focus on the average performance, such as the average signal-to-noise ratio (SNR), but rather on the statistics of the SNR. Of particular concern is the outage probability, i.e., the probability that the SNR falls below a given threshold. Because of the random fluctuations, the problem of estimating such a probability is particularly challenging, especially in numerical simulations where low outage values call for many time-consuming simulations, but also in experiments where collecting many observations may require a huge amount of resources to save and post-process the results.

Such difficulties stimulated the development of theoretical models for quick estimation of the PDL effects. Most of the literature focused on the interplay between PDL and amplified spontaneous emission (ASE) noise in the linear regime. Remarkable results have been provided by Gisin [5], who found the statistics of the resulting PDL after concatenation of many devices, and by Mecozzi and Shtaif [6] that investigated the asymptotic properties of PDL showing its Maxwellian statistics when expressed in dB. The implications of the interplay PDL-ASE on the SNR have been investigated by the same authors in [7] and by Shtaif in [8]. The implications of PDL on the channel capacity has been investigated by Nafta *et al.* in [3]. A quaternion approach to analytically investigate PDL has been proposed by Karlsson and Petersson in [9].

The interplay between PDL and the fiber nonlinear Kerr effect received much less analytical attention, and most of the literature focused on numerical/experimental investigations [10]–[17]. Such investigations showed contrasting results, since PDL showed limited interaction with the Kerr effect in [14] while a non-negligible interaction has been pointed out, for instance, in [12], [15].

In modern optical communication systems, it is customary to analyze the performance of the link by employing perturbative models because of their simplicity. Among the available models in the literature, particular attention has been captured by the Gaussian noise (GN) model [18] and its advanced version, the enhanced Gaussian noise (EGN) model [19], [20], also referred to as nonlinear interference noise (NLIN) model [21]. Such models showed excellent accuracy in a wide range of optical links, with savings in computational time of more than an order of magnitude compared with traditional models, such as the split-step Fourier method (SSFM).

We extended the scalar theory of the GN model by including polarization effects in [22], and first included PDL in the GN model framework in [23] for a quick estimation of the probability density function (PDF) of the nonlinear interference (NLI).

In this work, besides providing a novel mathematical formalism to cope with PDL in the GN model, we show how to account for PDL even in the EGN model. The general theory will be double-checked against SSFM simulations.

The resulting PDL-EGN theory extends the standard EGN model when fixed PDL elements are inserted along the link,

Manuscript received xxxxxxxxx xx, xxxx; accepted xxxxxxxxx xx, xxxx. Date of publication xxxxxxxxx xx, xxxx; date of current version xxxxxxxxx xx, xxxx. This work was supported by Nokia Bell-labs, Villarceaux, France. *(Corresponding author: Paolo Serena.)*

The authors are with the Department of Ingegneria e Architettura, Università di Parma, Parma 43124, Italy (e-mail: paolo.serena@unipr.it, chiara.lasagni@unipr.it, alberto.bononi@unipr.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier xx.xxxx/JLT.xxxx.xxxxxxx.

thus providing the NLI variance for a given PDL sequence. The advantages of using the new model as well as some interesting implications, such as the scaling of the outage probability with power, will be discussed.

The paper is organized as follows: in Section II we show the main theory, based on some results provided in the Appendices; in Section III we validate the model. Finally, in Section IV we draw our main conclusions.

II. PDL-EXTENDED EGN MODEL

We adopt the following bra-ket notation

$$|\tilde{A}(\omega)\rangle \triangleq \begin{bmatrix} \tilde{A}_x(\omega) \\ \tilde{A}_y(\omega) \end{bmatrix}, \quad \langle\tilde{A}(\omega)| \triangleq \begin{bmatrix} \tilde{A}_x^*(\omega), & \tilde{A}_y^*(\omega) \end{bmatrix}$$

where $\tilde{A}_{x,y}(\omega)$ indicate the Fourier transform of the two polarization tributaries of the transmitted electric field, with ω the angular frequency.

Under a first-order perturbative approximation, the received electric field $|\tilde{A}_{\mathrm{R}}(\omega)\rangle$ can be related to the transmitted one $|\tilde{A}(\omega)\rangle$ by [24]:

$$|\tilde{A}_{\mathrm{R}}(\omega)\rangle \approx \mathbf{T}(z,\omega)\left(|\tilde{A}(\omega)\rangle + |\tilde{w}(\omega)\rangle + |\tilde{n}(\omega)\rangle\right) \tag{1}$$

with $|\tilde{w}(\omega)\rangle$ and $|\tilde{n}(\omega)\rangle$ the ASE noise and the signal NLI, respectively, and $\mathbf{T}(z,\omega)$ a 2×2 matrix accounting for all linear impairments from input to coordinate z. Such a matrix can be separated into a scalar and a polarization-dependent contribution:

$$\begin{aligned} \mathbf{T}(z,\omega) &= e^{\vartheta(z,\omega)}\mathbf{U}(z) \\ \vartheta(z,\omega) &\triangleq -\int_0^z \left(\frac{\alpha(\xi)}{2} + j\beta(\xi,\omega)\right)\mathrm{d}\xi \end{aligned} \tag{2}$$

with α the fiber attenuation and β the imaginary part of the propagation constant. In this work, we focus on a matrix $\mathbf{U}$ accounting only for a frequency-independent PDL accumulated up to coordinate z. We assume lumped PDL (e.g., WSS and EDFA) at coordinates $z_p: \ p=0,\dots,N-1$, with $z_0=0$ and N the total number of PDL blocks. Each PDL element is followed by a flat constant-gain amplifier recovering the average power loss of PDL[1], such that the matrix $\mathbf{U}$ depends on the kth device at coordinate $z_k \le z$ with PDL matrix $\mathbf{M}_k$ by [1]:

$$\begin{aligned} \mathbf{U}(z) &= \mathbf{M}_p\mathbf{M}_{p-1}\cdots\mathbf{M}_0, \qquad z_p < z < z_{p+1} \\ \mathbf{M}_k &\triangleq \mathbf{W}_k^{\dagger}\begin{bmatrix} \sqrt{1+\Gamma_k} & 0 \\ 0 & \sqrt{1-\Gamma_k} \end{bmatrix}\mathbf{W}_k \end{aligned} \tag{3}$$

where $\dagger$ indicates transpose-conjugate, the $\mathbf{W}_k$ are matrices uniformly distributed in the set of the 2×2 unitary random matrices (Haar matrices), while Γ_k defines the PDL ρ_k by $\rho_k \triangleq (1+\Gamma_k)/(1-\Gamma_k)$. The PDL is usually expressed in dB by $10\log_{10}(\rho_k)$.

In the case of many identically distributed PDL elements, the resulting PDL of the link, expressed in dB, follows a Maxwellian distribution [6] with an average value scaling with $\sqrt{N}$.

Matrix $\mathbf{W}_k$ is statistically independent of matrix $\mathbf{W}_n$, with $k \neq n$. Such matrices are slowly varying in time compared to the symbol timing, hence while each polarization tributary preserves its average power while crossing the generic PDL element, each PDM data-block experiences a power imbalance between polarization tributaries. In absence of noise, such an imbalance can be removed at the receiver side by performing linear equalization, for instance by the zero-forcing equalizer $\mathbf{T}^{-1}(z,\omega)$. However, in the general case, PDL remains both in the ASE and the NLI. Fig. 1 sketches the idea for ASE. In this work, we assume $\mathbf{W}_k$ a random variable, thus time-independent.

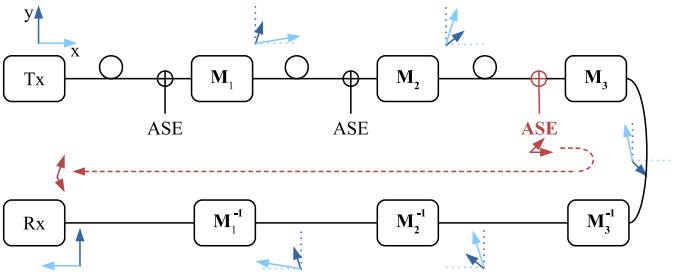

Fig. 1. Sketch of the effects experienced by the signal $|A\rangle$ and the ASE $|w\rangle$ along propagation. The zero-forcing equalization assumption is equivalent to a round-trip propagation of the signal, thus returning equal to itself at reception. However, ASE follows an incomplete round-trip, resulting in PDL.

After linear equalization, the received ASE $|\tilde{w}(\omega)\rangle$ is related to the ASE $|\tilde{w}_m(\omega)\rangle$ emitted by in-line amplifiers at coordinate z_m by:

$$|\tilde{w}(\omega)\rangle = \sum_{m=1}^{M} \mathbf{M}_0^{-1}\mathbf{M}_1^{-1}\cdots\mathbf{M}_{m-1}^{-1}\,|\tilde{w}_m(\omega)\rangle \ . \tag{4}$$

with M the number of amplifiers. For instance, concerning Fig. 1 referred to an ideal source, i.e., $\mathbf{M}_0=\mathbf{I}$ with $\mathbf{I}$ the identity matrix, we have:

$$\begin{aligned} |\tilde{w}(\omega)\rangle &= |\tilde{w}_1(\omega)\rangle + \mathbf{M}_1^{-1}\,|\tilde{w}_2(\omega)\rangle + \mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\,|\tilde{w}_3(\omega)\rangle \\ &= |\tilde{w}_1(\omega)\rangle + \mathbf{U}(z_1)^{-1}\,|\tilde{w}_2(\omega)\rangle + \mathbf{U}(z_2)^{-1}\,|\tilde{w}_3(\omega)\rangle \ . \end{aligned} \tag{5}$$

In Appendix A we show that the NLI takes the expression:

$$\begin{aligned} |\tilde{n}(\omega)\rangle = -j\iint_{-\infty}^{\infty}\sum_{p=0}^{N-1}\eta_p(\omega,\omega_1,\omega_2)\times \\ \langle\tilde{A}(\omega+\omega_1+\omega_2)|\mathbf{P}(z_p)|\tilde{A}(\omega+\omega_2)\rangle\,|\tilde{A}(\omega+\omega_1)\rangle\,\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi} \end{aligned} \tag{6}$$

with η_p the kernel of the optical fiber in the segment $[z_p, \ z_{p+1}]$, and $\mathbf{P}(z_p) \triangleq \mathbf{U}^{\dagger}(z_p)\mathbf{U}(z_p)$. Please note that, without PDL, $\mathbf{P}=\mathbf{I}$. Eq. (6) clearly shows the four-wave mixing (FWM) process generating the NLI.

Both the NLI and the ASE accumulate linearly with the propagation distance, both being additive under the model assumptions. Moreover, by comparing (4) and (6), we observe that after zero-forcing equalization they both depend on the PDL accumulated *before* their generation. However,

[1]This is typical of terrestrial EDFAs that rely on photodiodes and feedback algorithms that can be calibrated to control the output power (after removing the expected, calibrated, ASE contribution).

such a dependence follows completely different relationships. In particular, the quadratic dependence on the entries of matrix $\mathbf{U}(z_p)$ in the NLI is expected to induce larger random fluctuations of the SNR compared to the linear dependence in the ASE case.

We now introduce a discrete-time channel model relating the transmitted/received data symbol [24]. We assume $|A(t)\rangle$ is a wavelength division multiplex (WDM) of linearly modulated digital signals:

$$|A(t)\rangle = \sum_{n=-\infty}^{\infty} \sum_{h=1}^{\#\text{ channels}} \sum_{k\in(x,y)} a_{nhk} p_h(t-nT_h) e^{j\Omega_h t} |k\rangle \quad (7)$$

where a_{nhk} is the digital symbol (e.g., quadrature amplitude modulation, QAM) at time n, WDM channel-index h, and polarization k; $p_h(t-nT_h)$ is the supporting pulse at the nth symbol time of duration T_h and modulated at carrier frequency Ω_h. We find it useful to compact the notation by calling $a_{\mathbf{n}}$ the generic (scalar) information symbol per *(time, frequency, polarization)* channel use, where the vector $\mathbf{n}$ should be read as:

$$\mathbf{n} = \underset{\text{time}}{[n_1,} \ \underset{\text{frequency}}{n_2,} \ \overset{\text{polarization}}{n_3}] \ . \quad (8)$$

We will refer to $a_{\mathbf{n}}$ as an *atom* of the source information. This way, by adopting signal-space representation concepts, eq. (7) in the Fourier domain can be written as:

$$|\tilde{A}(\omega)\rangle = \sum_{\mathbf{n}} a_{\mathbf{n}} |\tilde{G}_{\mathbf{n}}(\omega)\rangle \quad (9)$$

where $\sum_{\mathbf{n}}$ stands for all summations in (7) and the basis functions are:

$$|\tilde{G}_{\mathbf{n}}(\omega)\rangle \triangleq \tilde{p}_{n_2}(\omega-\Omega_{n_2}) e^{-j(\omega-\Omega_{n_2})n_1 T_{n_2}} |n_3\rangle \ .$$

We assume the detector performs demodulation, matched filtering, zero-forcing equalization, sampling, and average carrier phase recovery. In our framework, demodulation, matched filtering, and sampling correspond to the inner product $\int_{-\infty}^{\infty} \langle \tilde{G}_{\mathbf{i}}(\omega)|\tilde{A}(\omega)\rangle \frac{\mathrm{d}\omega}{2\pi}$. In particular, such operation results in $a_{\mathbf{i}}$ in absence of impairments when using orthonormal basis functions, such as root-raised cosine pulses with non-overlapping spectrum among channels:

$$\sum_{\mathbf{n}} a_{\mathbf{n}} \int_{-\infty}^{\infty} \langle \tilde{G}_{\mathbf{i}}(\omega)|\tilde{G}_{\mathbf{n}}(\omega)\rangle \frac{\mathrm{d}\omega}{2\pi} = a_{\mathbf{i}} \ . \quad (10)$$

By following similar steps as [24], [25], from (1) we get the following discrete-time channel model relating the transmitted atom $a_{\mathbf{i}}$ to the received one $u_{\mathbf{i}}$:

$$u_{\mathbf{i}} = a_{\mathbf{i}} + w_{\mathbf{i}} + n_{\mathbf{i}}$$

where $w_{\mathbf{i}}$ and $n_{\mathbf{i}}$ are the sampled ASE and NLI, respectively:

$$\begin{aligned} w_{\mathbf{i}} &= \int_{-\infty}^{\infty} \langle \tilde{G}_{\mathbf{i}}(\omega)|\tilde{w}(\omega)\rangle \frac{\mathrm{d}\omega}{2\pi} \\ n_{\mathbf{i}} &= -j \sum_{\mathbf{k},\mathbf{m},\mathbf{n}} a^*_{\mathbf{k}} a_{\mathbf{m}} a_{\mathbf{n}} \mathcal{X}_{\mathbf{kmni}} \ . \end{aligned} \quad (11)$$

$\mathcal{X}_{\mathbf{kmni}}$ is a tensor weighting the four-atom mixing (FAM) at the symbol level:

$$\begin{aligned} \mathcal{X}_{\mathbf{kmni}} &= \sum_{p=0}^{N-1} \iiint_{-\infty}^{\infty} \eta_p(\omega,\omega_1,\omega_2) \\ &\quad \times \langle \tilde{G}_{\mathbf{k}}(\omega+\omega_1+\omega_2)|\mathbf{P}(z_p)|\tilde{G}_{\mathbf{m}}(\omega+\omega_2)\rangle \\ &\quad \times \langle \tilde{G}_{\mathbf{i}}(\omega)|\tilde{G}_{\mathbf{n}}(\omega+\omega_1)\rangle \frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}\frac{\mathrm{d}\omega}{2\pi} \\ &= \sum_{p=0}^{N-1} P_{k_3 m_3}(z_p) \delta_{i_3 n_3} \mathcal{S}_{\mathbf{kmni}}(z_p) \end{aligned} \quad (12)$$

where the δ indicates Kronecker's delta, and in the final identity of (12) we expanded the tensor in terms of the tensor $\mathcal{S}_{\mathbf{kmni}}$ weighting the FWM interaction at the scalar level:

$$\begin{aligned} &\mathcal{S}_{\mathbf{kmni}}(z_p) \triangleq \iiint_{-\infty}^{\infty} \eta_p(\omega,\omega_1,\omega_2) \times \\ &\tilde{G}^*_{\mathbf{k}}(\omega+\omega_1+\omega_2)\tilde{G}_{\mathbf{m}}(\omega+\omega_2)\tilde{G}^*_{\mathbf{i}}(\omega)\tilde{G}_{\mathbf{n}}(\omega+\omega_1)\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}\frac{\mathrm{d}\omega}{2\pi} \end{aligned} \quad (13)$$

where $\tilde{G}_{\mathbf{n}}(\omega)$ is defined in implicit form by $|\tilde{G}_{\mathbf{n}}(\omega)\rangle \triangleq \tilde{G}_{\mathbf{n}}(\omega)|n_3\rangle$. It is worth noting that in the special, yet relevant, case of a homogeneous link in absence of PDL, the summation $\sum_p$ in (12) can be closed with some advantage for numerical purposes and simplicity.

We now evaluate the covariance of ASE and NLI atoms when acting alone.

A. ASE variance

The PDL impact on ASE has been investigated in several papers in the literature [9], [8], [10], [1], whose main results we now rephrase in our notation.

By definition, the variance of the ASE atom $\mathbf{i}$ is $\sigma^2_{\text{ASE}} = \mathbb{E}[w_{\mathbf{i}} w^*_{\mathbf{i}}]$, with $\mathbb{E}$ indicating expectation. We evaluate it focusing on a link with independent and identically distributed ASE sources, hence with [26, p. 418]:

$$\mathbb{E}\left[|\tilde{w}_m(\omega)\rangle \langle \tilde{w}_p(\mu)|\right] = \frac{N_0}{2}\delta(\omega-\mu)\delta_{mp}\mathbf{I} \quad (14)$$

where the two δ indicate Dirac/Kronecker's delta, respectively, while N_0 is the one-sided, dual-polarization, power spectral density (PSD) of ASE per amplifier. N_0 is related to the noise figure F and the gain G by $N_0 = h\nu FG$, with h Planck's constant and ν carrier frequency.

Let B be the noise equivalent bandwidth of the receiver. By using (4) and the orthogonality property (10) in (14) we have for the generic $i_3 \in (x,y)$:

$$\sigma^2_{\text{ASE},i_3} = \frac{N_0 B}{2} \sum_{p=1}^{M} P^{-1}_{i_3 i_3}(z_p) \quad (15)$$

with M the number of optical amplifiers in the link. Please note that the matrices $\mathbf{P}(z_p)$ and hence their elements $P_{i_3 j_3}(z_p)$ are not independent but related by the concatenation rule of PDL [9].

With PDL, and in absence of joint FEC and equalization, it is more interesting to deal with the SNR per polarization,

because of the asymmetrical behavior of noise power. With reference to the generic polarization i:

$$\mathrm{SNR}^{i}_{\mathrm{ASE}} = \frac{S_i}{\sigma^2_{\mathrm{ASE},i}} = \frac{\mathrm{SNR}^{i}_{\mathrm{ASE}}(\mathrm{PDL{=}0})}{\frac{1}{M}\sum_{p=1}^{M} P^{-1}_{ii}(z_p)}, \qquad i \in (x,y) \quad (16)$$

where S_i is the signal power on polarization i and $\mathrm{SNR}^{i}_{\mathrm{ASE}}(\mathrm{PDL{=}0}) = S_i/\left(\frac{N_0 MB}{2}\right)$. The denominator of (16), equal to the span-average of P^{-1}_{ii}, is the random PDL loss/gain per polarization.

B. NLI variance

The FWM process underpinning the NLI is formally identical to the scalar case, with just a different weighting tensor. However, some symmetries cease to hold, hence the master theorem at the heart of the EGN model (see Appendix B) must be properly generalized by taking care of such a novelty. We observe the following symmetries in indexing:

$$\begin{aligned} &\mathcal{X}_{\mathbf{kmni}} = \mathcal{X}^{*}_{\mathbf{mkin}}, \qquad \text{(always)} \\ &\mathcal{X}_{\mathbf{kmni}} = \mathcal{X}_{\mathbf{knmi}}, \qquad \text{(no PDL)} \ . \end{aligned} \quad (17)$$

The breakdown of the last symmetry induces a small modification in the master theorem as detailed in Appendix B.

With such ingredients, after carrier phase estimation the variance of the NLI atom $n'_{\mathbf{i}}$, $\sigma^2_{\mathrm{NLI}} = \mathbb{E}[n'_{\mathbf{i}} n'^{*}_{\mathbf{i}}]$, can be found by plugging (12) into the master theorem (28). Such a variance can be split into the GN, fourth-order noise (FON) [24], and higher-order noise (HON) [19]–[21] contributions:

$$\sigma^2_{\mathrm{NLI}} = \sigma^2_{\mathrm{GN}} + \underbrace{\sigma^2_{\mathrm{FON}} + \sigma^2_{\mathrm{HON}}}_{\text{EGN correction}} \ .$$

We now analyze the contributions for independent and identically distributed data symbols. The result depends on the statistical cumulants of the symbols [20], which are related to the main moments $\mu_n \triangleq \mathbb{E}[|a_{\mathbf{k}}|^n]$ by:

$$\begin{aligned} \kappa_1 &= \mu_2 \\ \kappa_2 &= \mu_4 - 2\mu_2^2 \\ \kappa_3 &= \mu_6 - 9\mu_4\mu_2 + 12\mu_2^3 \ . \end{aligned}$$

1) GN term: As outlined in Appendix B, the GN contribution to the variance of polarization $i_3 \in (x,y)$ is:

$$\begin{aligned} \sigma^2_{\mathrm{GN}} &= \kappa_1^3 \sum_{\mathbf{k},\mathbf{m},\mathbf{n}} \mathcal{X}_{\mathbf{kmni}} \left(\mathcal{X}^{*}_{\mathbf{kmni}} + \mathcal{X}^{*}_{\mathbf{knmi}}\right) \\ &= \kappa_1^3 \sum_{\mathbf{k},\mathbf{m},\mathbf{n}} \delta_{i_3 n_3} \sum_{p,\ell=0}^{N-1} \mathcal{S}_{\mathbf{kmni}}(z_p) \mathcal{S}^{*}_{\mathbf{kmni}}(z_\ell) \\ &\times \left(P_{k_3 m_3}(z_p) P^{*}_{k_3 m_3}(z_\ell) + \delta_{i_3 m_3} P_{k_3 m_3}(z_p) P^{*}_{k_3 n_3}(z_\ell)\right) \end{aligned}$$

where we used (12). Such a result can be easily generalized to the spatial-covariance matrix. We introduce the 2×2 GN covariance matrix between the polarizations, $\mathbf{K}_{\mathrm{GN}} : K^{(\mathrm{GN})}_{i_3 j_3} = \mathbb{E}[n'_{\mathbf{i}} n'^{*}_{\mathbf{j}}]$, $i_{1,2} = j_{1,2}$, $(i_3, j_3) \in (x,y)$, which takes the elegant form [23]:

$$\begin{aligned} \mathbf{K}_{\mathrm{GN}} &\triangleq \begin{bmatrix} \mathrm{var}\left(\mathrm{NLI}^{(\mathrm{GN})}_x\right) & \mathrm{cov}\left(\mathrm{NLI}^{(\mathrm{GN})}_x, \mathrm{NLI}^{(\mathrm{GN})}_y\right) \\ \mathrm{cov}\left(\mathrm{NLI}^{(\mathrm{GN})}_x, \mathrm{NLI}^{(\mathrm{GN})}_y\right) & \mathrm{var}\left(\mathrm{NLI}^{(\mathrm{GN})}_y\right) \end{bmatrix} \\ &= \sum_{p,\ell=0}^{N-1} \rho_{\mathrm{GN}}(p,\ell) \Big(\mathrm{Tr}\left[\mathbf{P}(z_p)\mathbf{P}^{\dagger}(z_\ell)\right] \mathbf{I} + \mathbf{P}(z_p)\mathbf{P}^{\dagger}(z_\ell) \Big) \end{aligned} \quad (18)$$

where $\rho_{\mathrm{GN}}(p,\ell) \triangleq \sum \mathcal{S}_{\mathbf{kmni}}(z_p) \mathcal{S}^{*}_{\mathbf{kmni}}(z_\ell)$, with the summations limited to the temporal and frequency indexes, is the *scalar* cross-correlation between the NLI accumulated in trunk p and trunk ℓ, while Tr indicates the trace of a matrix. It is worth noting that in absence of PDL we have [24], [21]:

$$\begin{aligned} \sigma^2_{\mathrm{GN}}(\text{no PDL}) = \sum_{p,\ell=0}^{N-1} \rho_{\mathrm{GN}}(p,\ell) = \sum_{h,r,s=1}^{\#\text{ channels}} \iiint_{-\infty}^{\infty} |\eta(\omega,\omega_1,\omega_2)|^2 \\ \times \left|\tilde{P}^{*}_i(\omega-\Omega_i)\right|^2 \left|\tilde{P}^{*}_h(\omega+\omega_1+\omega_2-\Omega_h)\right|^2 \\ \times \left|\tilde{P}_s(\omega+\omega_2-\Omega_s)\right|^2 \left|\tilde{P}_r(\omega+\omega_1-\Omega_r)\right|^2 \frac{\mathrm{d}\omega}{2\pi}\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi} \ . \end{aligned} \quad (19)$$

where the term η is the fiber-kernel of the entire link, see Appendix A. We also observe that, since without PDL $\mathbf{P}(z_p) = \mathbf{I}$ for each p, it is $\mathbf{K}_{\mathrm{GN}} = 3\mathbf{I}$ [20].

The right-hand side in (19) is well known in the literature, see, e.g., [24]. The main contribution of this work is that eq. (18) generalizes the scalar result (19) to the case with PDL. We note that now we need to know all trunk cross-correlations, while in the scalar case such information was not required. However, the matrix of elements $\rho_{\mathrm{GN}}(p,\ell)$ is a Toeplitz matrix, hence it can be calculated in a short simulation (preload), usually of the order of seconds, for instance with the algorithm [21]. Once ρ_{GN} is available, the PDL statistics can be evaluated very quickly by computing the matrix in (18).

The SNR of the generic polarization i finally is [18]:

$$\mathrm{SNR}^{i}_{\mathrm{GN}} = \frac{1}{K^{(\mathrm{GN})}_{ii} S_i^2}, \qquad i \in (x,y) \quad (20)$$

which, contrary to ASE, cannot be expressed in closed-form in terms of the SNR without PDL.

2) FON terms: The FON variance of polarization $i_3 \in (x,y)$ is derived in Appendix B, here repeated for convenience:

$$\sigma^2_{\mathrm{FON}} = \kappa_2 \kappa_1 \sum_{\mathbf{k},\mathbf{n}} \left(|\mathcal{X}_{\mathbf{kkni}} + \mathcal{X}_{\mathbf{knki}}|^2 + |\mathcal{X}_{\mathbf{nkki}}|^2 \right) \ . \quad (21)$$

In Appendix B we label the first absolute value in (21) by F4 and the second by Q4. With similar steps done for the GN counterpart, we introduce the FON covariance matrix $\mathbf{K}_{\mathrm{FON}}$ between spatial coordinates. After inserting (12) in (21) we obtain:

$$\mathbf{K}_{\mathrm{FON}} = \sum_{p,\ell=0}^{N-1} \Big(\rho_{\mathrm{F4}}(p,\ell)\mathbf{F}(z_p,z_\ell) + \rho_{\mathrm{Q4}}(p,\ell)\mathbf{Q}(z_p,z_\ell) \Big)$$

where $\rho_{\mathrm{F4}}(p,\ell) \triangleq \sum \mathcal{S}_{\mathbf{kkni}}(z_p) \mathcal{S}^{*}_{\mathbf{kkni}}(z_\ell)$ and $\rho_{\mathrm{Q4}}(p,\ell) \triangleq \sum \mathcal{S}_{\mathbf{nkki}}(z_p) \mathcal{S}^{*}_{\mathbf{nkki}}(z_\ell)$, with the summations limited to the

temporal and frequency indexes, are the scalar cross-correlations between the two kinds of FON (see Appendix B) accumulated in trunk p and trunk ℓ. They can be evaluated with the scalar EGN model [21], [19], [20].

Matrices $\mathbf{F}$ and $\mathbf{Q}$ are the novelty introduced by PDL. They have the following entries:

$$\begin{aligned}
F_{11} &= 4P_{11}(z_p)P_{11}^*(z_\ell) + P_{22}(z_p)P_{22}^*(z_\ell) + P_{12}(z_p)P_{12}^*(z_\ell)\\
F_{12} &= P_{22}(z_p)P_{21}^*(z_\ell) + P_{12}(z_p)P_{11}^*(z_\ell)\\
F_{21} &= P_{11}(z_p)P_{12}^*(z_\ell) + P_{21}(z_p)P_{22}^*(z_\ell)\\
F_{22} &= 4P_{22}(z_p)P_{22}^*(z_\ell) + P_{11}(z_p)P_{11}^*(z_\ell) + P_{21}(z_p)P_{21}^*(z_\ell)
\end{aligned}$$

while matrix $\mathbf{Q}$:

$$\begin{aligned}
Q_{11} &= P_{11}(z_p)P_{11}^*(z_\ell) + P_{21}(z_p)P_{21}^*(z_\ell)\\
Q_{12} &= Q_{21} = 0\\
Q_{22} &= P_{22}(z_p)P_{22}^*(z_\ell) + P_{12}(z_p)P_{12}^*(z_\ell)\ .
\end{aligned}$$

Please note that in absence of PDL we have $\mathbf{F} = 5\mathbf{I}$ and $\mathbf{Q} = \mathbf{I}$, as expected [20]. It is worth noting that in dispersive optical links the dominant FON term is usually the one related to the F4 term, for the same reasons observed without PDL [20].

3) HON term: The HON is a sixth order term, labeled by Q6 in Appendix B. Such a term occurs only when all the six atoms joining the NLI covariance are identical. We have:

$$\sigma^2_{\text{HON}} = \kappa_3 \sum_{\mathbf{n}} |\mathcal{X}_{\mathbf{nnni}}|^2\ .$$

We introduce a HON covariance matrix $\mathbf{K}_{\text{HON}}$ between spatial coordinates obtaining:

$$\mathbf{K}_{\text{HON}} = \sum_{p,\ell=0}^{N-1} \rho_{\text{Q6}}(p,\ell)\mathbf{H}(z_p, z_\ell)$$

with ρ_{Q6} the scalar HON cross-correlation between trunk p and trunk ℓ, while matrix $\mathbf{H}$ has entries:

$$\begin{aligned}
H_{ii} &= P_{ii}(z_p)P_{ii}^*(z_\ell), \quad i = 1,2\\
H_{ij} &= 0, \quad i \neq j\ .
\end{aligned}$$

Please note that in absence of PDL we have $\mathbf{H} = \mathbf{I}$. In most of the optical links, the HON term is dominated by the FON term.

Finally, the SNR associated to the NLI is:

$$\text{SNR}^i_{\text{NLI}} = \frac{1}{\left(K^{(\text{GN})}_{ii} + K^{(\text{FON})}_{ii} + K^{(\text{HON})}_{ii}\right) S_i^2}, \quad i \in (x,y)$$

while the overall SNR, by neglecting ASE-NLI interaction, follows the usual concatenation rule:

$$\frac{1}{\text{SNR}^i} = \frac{1}{\text{SNR}^i_{\text{ASE}}} + \frac{1}{\text{SNR}^i_{\text{NLI}}}\ . \tag{22}$$

III. Numerical validation

We checked the proposed model against SSFM based simulations. Common parameters to all simulations are the pulse types, i.e., root-raised cosine pulses with roll-off 0.01 sent at 49 Gbaud, the channel spacing, 50 GHz, and the optical fibers, i.e., single-mode fibers (SMF) having length 100 km, dispersion $D = 17$ ps/nm/km, attenuation $\alpha = 0.2$ dB/km, nonlinear coefficient $\gamma = 1.26$ 1/W/km without PMD. The channel under test (CUT) is in any case the central channel of the WDM.

Since the inclusion of PDL in the NLI is the main novelty of this work, in a first test we focused on the above system without ASE. The WDM was made of 11×50 GHz channels, PDM-modulated with lasers having state of polarization (SOP) randomly chosen on the Poincaré sphere. We used Gaussian distributed symbols, i.e., the capacity-achieving modulation format for the additive white Gaussian noise (AWGN) channel. In this case, the EGN model degenerates into the GN model [23].

PDL was included at each amplifier with a value of 0.5 dB. The residual dispersion per span was either 30 ps/nm (dispersion-managed, DM30) or absent (dispersion-uncompensated, DU). In any case, full dispersion compensation was implemented at the receiver input after propagation over 10 or 20 spans.

The link was simulated by the SSFM and compared with the prediction of the PDL-EGN model. In the PDL-EGN matched filtering, zero-forcing PDL equalization and carrier phase estimator (CPE) are implicit in the model. On the other hand, in the SSFM case we implemented matched filtering followed by a 1-tap zero-forcing equalizer, and by a CPE recovering the average phase induced by the fibers.

We estimated the PDF of the received SNR by Monte Carlo simulations over the PDL seeds, both with SSFM runs and with the PDL-EGN model. In the SSFM case, we used 1000 different random PDL realizations. For each realization, we varied the random state of polarization of the channel lasers as well. The SSFM symmetric-step was chosen according to the local-error criterion with a first step accumulating a FWM phase of 20 rad [27]. The number of symbols was 65536, sufficiently high to capture the largest walk-off among channels and to have a negligible error from the Monte Carlo estimation [15]. In the PDL-EGN, besides the Monte Carlo iterations over the PDL seeds, we numerically solved the frequency integrals by another Monte Carlo sampling.

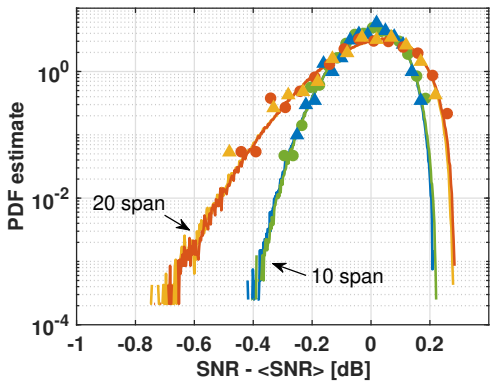

Fig. 2. Estimate of the PDF of the SNR per polarization by SSFM simulations (triangles for DM30, circles for DU) and the corresponding (almost overlapped) PDL-GN PDFs (solid lines).

Fig. 2 depicts the estimate of the PDF of the SNR per

polarization by SSFM (symbols) and the PDL-GN model (eq. (20), lines). We investigated two links lengths, 10 and 20 spans. In order to compare the two links, we plotted the PDFs versus the SNR offset from its mean. For each link, the mean SNR of SSFM simulations was within 0.1 dB of that from the PDL-GN model.

We observe an excellent fit at both distances, thus confirming the validity of the proposed model. For the sake of completeness, we also repeated the DU curve at 10 spans with a relevant fiber PMD of 0.32 ps/$\sqrt{\text{km}}$ [4]. Even if not reported here to avoid confusion, the curve was slightly narrower but almost overlapped to the one without PMD, thus suggesting that the SOP decorrelation induced by fiber PMD is masked by the decorrelation induced by channel walk-off in such a link. A similar result has also been observed in the experiment of [13]. In particular, we note that dispersion management does not affect the PDF shape, while it strongly impacts the average value because of strongly different correlations of the NLI among spans.

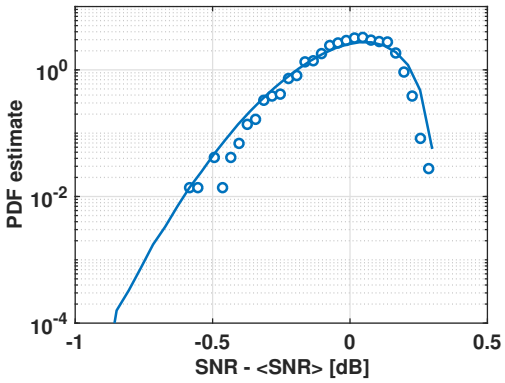

Fig. 3. PDF estimate of the SNR per polarization by the PDL-EGN (solid line) and SSFM (symbols). 32 × 100 km SMF link with star-8QAM- modulation format, with different PDL between EDFA and ROADM.

In a second test, we investigated a more realistic distribution of PDL along the optical link. We thus focused on a 3200 km network scenario where the CUT, besides being added and dropped by ROADMs, crosses ROADMs placed every 4 spans. Within each ROADM working in bypass mode, two WSS were crossed. The PDL was 0.1 dB within EDFAs and 0.4 dB within the WSS, respectively [16]. Fig. 4 sketches the link.

The modulation format was PDM-star 8QAM, for a total of 21 channels. In this setup we included ASE, with a noise figure of 5 dB per amplifier. It is worth noting that with ideal equalization the last drop impacts equally ASE, NLI, and signal, hence with no implications on the statistics.

The PDF estimated by the PDL-EGN and by SSFM at power 1 dBm per channel maximizing the SNR is reported in Fig. 3. Again, we observe a good match between the two models. Even if not shown in the figure, we observed a bias of 0.1 dB by the PDL-EGN model.

Having tested the validity of the model, we applied it to estimate the outage probability, here defined as the probability that the SNR falls below the threshold of 10.56 dB,

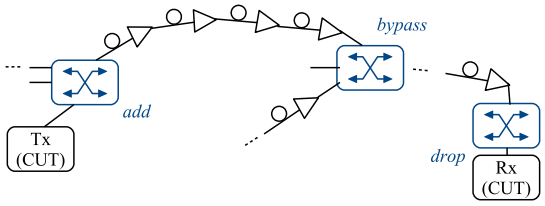

Fig. 4. Network under investigation. 3200 km.

corresponding to a Q-factor of 6.5 dB. At such a threshold the achievable information rate of a symbol-by-symbol star-8QAM detector in AWGN is 2.84 bits/symbol. Since the loss to the nominal 3 bits/symbol carried out by the modulation format is 0.16 bits/symbol, the transmission is expected to be feasible with realistic forward-error correcting (FEC) codes. Moreover, since the Shannon capacity at such threshold is 3.63 bits/symbol, the modulation format is a good candidate for transmission over 3200 km.

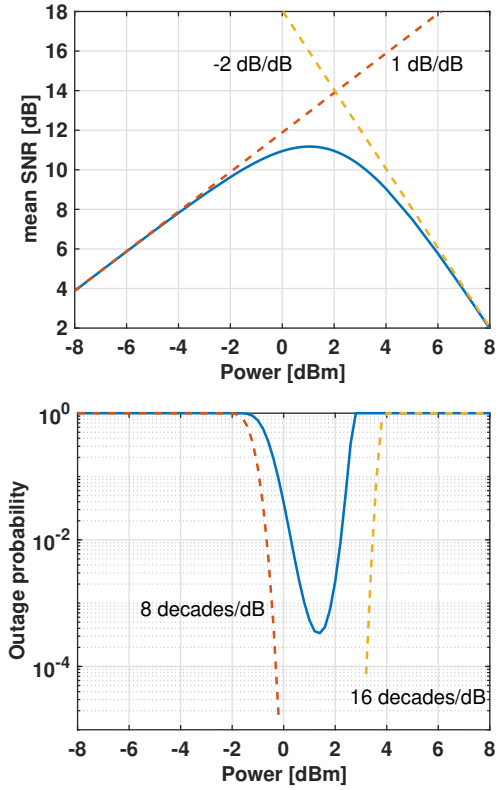

Fig. 5. Top: mean SNR per polarization vs. power. Bottom: outage probability @ Q-factor = 6.5 dB. All curves with the proposed model with only ASE or NLI (dashed lines) or with both ASE and NLI (solid lines). Star-8QAM 32 × 100 km SMF link with different PDL between EDFA and ROADM, see text.

Fig. 5(bottom) reports the outage probability versus launched power per channel, and for reference Fig. 5(top) also shows the mean SNR per polarization versus the same power for the sake of comparison. Several interesting observations can be drawn from Fig. 5. First, we observe that the minimum outage probability is $3 \cdot 10^{-4}$, a non-negligible value indicating the importance of including PDL in link design. Second, the best power for the mean SNR does not coincide with the best

power for the outage probability, with a gap of 0.4 dB. This is strictly related to the nonlinear relation between the outage probability and the SNR. Third, it is interesting to compare the slope of the asymptotes in the ASE-dominated regime (linear regime) and the NLI-dominated regime (nonlinear regime). In the mean SNR curve, we observe a slope of +1 dB/dB in the linear regime and -2 dB/dB in the nonlinear regime. A factor 2 in absolute terms between the two slopes is well known and related to the scaling properties of the ASE and the NLI variance [28]. Quite surprisingly, we still observe a factor 2 between the slopes of the asymptotes on the outage probability graphs. Such an observation can be very useful for quickly scaling the outage probability with power.

In Fig. 6 we compare the individual contributions of ASE and NLI to SNR statistics. We set the power to 2 dBm, i.e., where the two effects yield the same average variance. Symbols indicate SSFM simulations. We observe that, in this setup, the PDF of the ASE-only case is slightly larger than that of the NLI-only case, although the average values are similar. Most important, the ASE-only and the NLI-only PDFs have different shapes, hence the NLI cannot be treated as an equivalent extra-ASE distributed along the link.

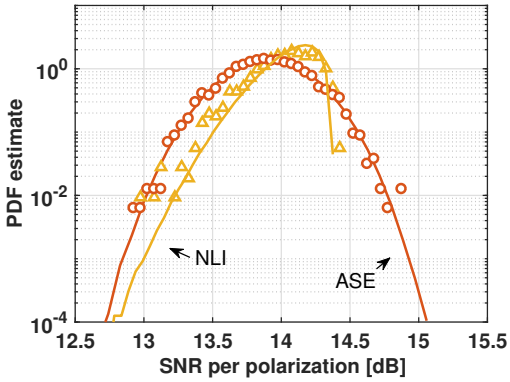

Fig. 6. Individual contribution of ASE and NLI to the PDF by using the proposed PDL-EGN model (solid line) and SSFM simulations (circles: ASE, triangles: NLI). Power 2 dBm.

It is worth noting that if instead of plotting the per-polarization SNR PDF, we plot the PDF of the overall PDM SNR, as usually done in the literature, we obtain a sharply different behavior, as depicted in Fig. 7 with the PDL-EGN model. The PDM SNR, or simply SNR, is defined as $\text{SNR} = (S_x + S_y)/(\sigma_x^2 + \sigma_y^2)$, with $\sigma_{x,y}^2$ the variance of the noise under investigation. The reason for the differences is related to the antithetic impact of PDL on the x and y ASE variances, which is not manifested by SPM- and XPM-like contributions that operate through a common scalar nonlinear phase on both polarizations [15]. Hence, treating NLI as much as ASE, i.e., as an AWGN term that adds incoherently along the link, is inappropriate in the presence of PDL.

It is interesting to compare the computational times of the PDL-EGN model and the SSFM. As a reference, an SSFM simulation, with step set-up as in [27], took 1 day to run 125 PDL seeds on a cluster using INTEL XEON E5- 2683v4 2.1GHz 32 cores central processing units (CPU) with 128 GB of RAM and NVIDIA Tesla P100 graphics processing unit (GPU). The same seeds have been simulated in a fraction of second with the PDL-EGN, plus an overhead for the computation of the span cross-correlations of the order of seconds for the PDL-GN and few minutes for the PDL-EGN. Not surprisingly, with the EGN we were able to simulate 10^6 PDL seeds, while with SSFM only 10^3.

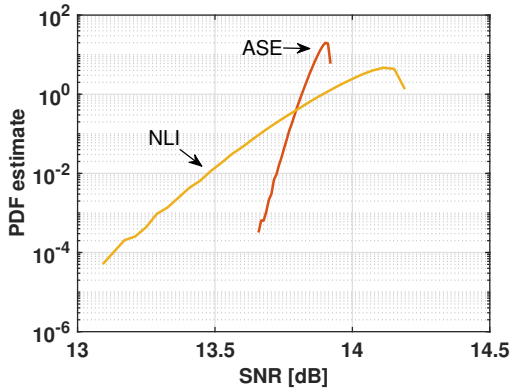

Fig. 7. Estimate of the PDF of the PDM SNR usually adopted in the literature by the proposed PDL-EGN model. Same setup of Fig. 6.

As a final comparison, we tested the impact of the modulation format. The link is the same as Fig. 6 with both ASE and NLI. The resulting PDF is shown in Fig. 8 for different modulation formats. We observe that the GN model underestimates the average value of the other modulation formats [18], with a mismatch up to 0.77 dB w.r.t. the QPSK curve. However, the differences between PDF shapes are smaller. For instance, the lowest SNR that we estimated in the GN case is 0.65 dB smaller than the average value, while for QPSK, by using the same random seeds, we estimated a smaller value by 0.8 dB.

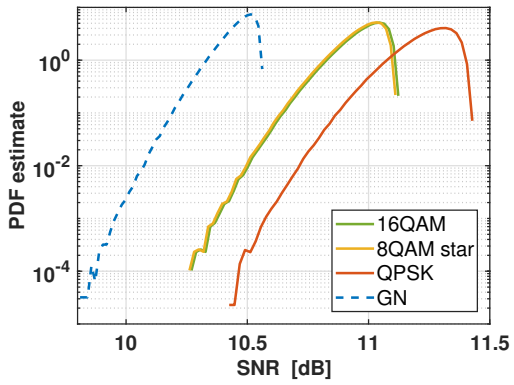

Fig. 8. Same setup of Fig. 6 with ASE and NLI, for different modulation formats.

IV. Conclusions

We extended the EGN model to include PDL. The extension forced us to rework the whole EGN theory, yielding a formally compact master equation based on tensors able to span the time, frequency and polarization axes. With the new model,

it is possible to estimate the statistics of the SNR in the nonlinear regime by exploring very rare events. Our results show that the NLI interplay with PDL follows a different behavior than the one experienced by ASE and thus needs a proper description. Besides this aspect, probably the main advantage of the novel model relies on its extremely fast computational time compared to standard algorithms, such as the SSFM. For instance, we investigated 1000 different optical links with random PDL in a fraction of second after a pre-processing of the order of seconds to minutes. Such times are inaccessible to SSFM simulations. The model can now be used to include a frequency-independent PDL in the design of optical links without allocating empirical margins as usually done. This way, it is possible to perform the design optimizing the system outage probability instead of the average performance.

APPENDIX A
NONLINEAR INTERFERENCE

In the absence of polarization-mode dispersion, the optical propagation within an optical fiber is well described by the Manakov equation. A first-order perturbative solution of the Manakov equation is [24], [18], [21], [20]:

$$|A(z,t)\rangle \simeq e^{\mathcal{L}z}\,|A(0,t)\rangle + \int_0^z e^{\mathcal{L}(z-\xi)}\mathcal{N}\left(e^{\mathcal{L}\xi}\,|A(0,t)\rangle\right)\mathrm{d}\xi \ . \tag{23}$$

where the operator $\mathcal{L}$ accounts for linear effects while $\mathcal{N}$ for nonlinear effects, respectively. $\mathcal{L}$ is best defined in the frequency domain by its Fourier transform $\mathscr{F}\left\{e^{\mathcal{L}z}\right\} \triangleq e^{\vartheta(z,\omega)}$, with ϑ accounting for attenuation and dispersion, see (2). $\mathcal{N}$ is best described in the time domain t, taking the following expression in absence of the Raman effect:

$$\mathcal{N} \triangleq -j\gamma\frac{8}{9}\,\langle A(z,t)|A(z,t)\rangle\,|A(z,t)\rangle \ .$$

The integral in (23) defines the NLI whose physical interpretation is simple: the NLI is additive along distance, and its generic contribution generated at coordinate ξ depends on the unperturbed signal up to that coordinate and experiences only linear effects up to the end [20]. If we factor out $e^{\mathcal{L}z}$ and apply zero-forcing equalization, i.e., concatenation with the inverse $e^{-\mathcal{L}z}$, the received signal $|A_{\mathrm{R}}\rangle$ is:

$$|A_{\mathrm{R}}\rangle \simeq |A(0,t)\rangle + \int_0^z e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}\,|A(0,t)\rangle\right)\mathrm{d}\xi \ . \tag{24}$$

which further suggests interpreting the NLI as an infinite summation of echoes.

The Fourier transform $|\tilde{n}(\omega)\rangle$ of the NLI in (24) is:

$$|\tilde{n}(\omega)\rangle = -j\iint_{-\infty}^{\infty}\eta(\omega,\omega_1,\omega_2)\times \langle\tilde{A}(\omega+\omega_1+\omega_2)|\tilde{A}(\omega+\omega_2)\rangle\,|\tilde{A}(\omega+\omega_1)\rangle\,\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}$$

with the kernel of the optical link in $(0,z)$ given by:

$$\eta(\omega,\omega_1,\omega_2) \triangleq \frac{8}{9}\gamma\int_0^z e^{\vartheta(\xi,\omega+\omega_1)+\vartheta(\xi,\omega+\omega_2)+\vartheta^*(\xi,\omega+\omega_1+\omega_2)-\vartheta(\xi,\omega)}\mathrm{d}\xi \ . \tag{25}$$

It is worth noting that for optical fibers with constant parameters the integral in ξ can be given in closed form [18].

Inserting a frequency-independent PDL in the model corresponds to applying the substitutions $e^{\mathcal{L}\xi} \to e^{\mathcal{L}\xi}\mathbf{U}(z)$ and $e^{-\mathcal{L}\xi} \to e^{-\mathcal{L}\xi}\mathbf{U}^{-1}(z)$, with $\mathbf{U}$ accounting for the cumulative PDL up to coordinate z, as per (2). With lumped PDL the matrix $\mathbf{U}$ is a staircase function in z, hence, it is convenient breaking the integral in z into a summation of integrals between consecutive PDL elements, i.e., $\int_0^z \to \sum_p \int_{z_{p-1}}^{z_p}$. With such substitutions we finally get:

$$|\tilde{n}(\omega)\rangle = -j\iint_{-\infty}^{\infty}\sum_{p=0}^{N-1}\eta_p(\omega,\omega_1,\omega_2)\times \langle\tilde{A}(\omega+\omega_1+\omega_2)|\mathbf{P}(z_p)|\tilde{A}(\omega+\omega_2)\rangle\,|\tilde{A}(\omega+\omega_1)\rangle\,\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi} \tag{26}$$

where $\mathbf{P}(z_p) \triangleq \mathbf{U}^{\dagger}(z_p)\mathbf{U}(z_p)$ and η_p can be evaluated as per (25) but integrated over (z_p, z_{p+1}). If the source does not have PDL, we simply set $\mathbf{U}(z_0) = \mathbf{I}$ at $z_0 = 0$.

APPENDIX B
MASTER THEOREM

The key ingredient to perform a statistical analysis is the variance of the NLI atom, i.e., $\mathbb{E}[|n_{\mathbf{i}}|^2]$:

$$\mathbb{E}\left[n_{\mathbf{i}}n_{\mathbf{i}}^*\right] = \sum\mathbb{E}\left[a_{\mathbf{k}}^*a_{\mathbf{m}}a_{\mathbf{n}}a_{\mathbf{l}}a_{\mathbf{j}}^*a_{\mathbf{o}}^*\right]\mathcal{X}_{\mathbf{kmni}}\mathcal{X}_{\mathbf{ljoi}}^* \ .$$

However, part of the NLI is compensated by digital signal-processing at the receiver. At the simplest level, a basic CPE removes the average phase φ, which, in the perturbative framework, corresponds to work with the following NLI:

$$n_{\mathbf{i}}' = n_{\mathbf{i}} + j\varphi a_{\mathbf{i}} \ . \tag{27}$$

The real target is thus $\mathbb{E}[|n_{\mathbf{i}}'|^2]$, which is given by the following:

Theorem 1: Assume all the $a_{\mathbf{k}}$ to be complex zero-mean independent random variables with n-fold rotational symmetry and $n \geq 4$. Then:

$$\begin{aligned}\mathbb{E}\left[n_{\mathbf{i}}'n_{\mathbf{i}}'^*\right] = &\sum_{\mathbf{n}}\kappa_3^{(\mathbf{n})}\left|\mathcal{X}_{\mathbf{nnni}}\right|^2 \\ &+\sum_{\mathbf{k},\mathbf{n}}\kappa_2^{(\mathbf{k})}\kappa_1^{(\mathbf{n})}\left(\left|\mathcal{X}_{\mathbf{kkni}}+\mathcal{X}_{\mathbf{knki}}\right|^2+\left|\mathcal{X}_{\mathbf{nkki}}\right|^2\right) \\ &+\sum_{\mathbf{k},\mathbf{m},\mathbf{n}}\kappa_1^{(\mathbf{k})}\kappa_1^{(\mathbf{n})}\kappa_1^{(\mathbf{m})}\mathcal{X}_{\mathbf{kmni}}\left(\mathcal{X}_{\mathbf{kmni}}^*+\mathcal{X}_{\mathbf{knmi}}^*\right)\end{aligned} \tag{28}$$

with $\kappa_n^{(\mathbf{k})}$ the n-th order cumulant of data symbols at channel use at time k_1, frequency k_2, and polarization k_3. □

Such a theorem is a generalization of the one provided in [20]. It is worth noting that the main difference is the broken degeneracy between the inner indexes of the tensor. For instance, the last term $\mathcal{X}_{\mathbf{kmni}}^*+\mathcal{X}_{\mathbf{knmi}}^*$ is equal to $2\mathcal{X}_{\mathbf{kmni}}^*$ in the scalar case according to (17). The basic intuition behind (28) is the following. Since the Kerr nonlinearity is cubic, the product $n_{\mathbf{i}}n_{\mathbf{i}}^*$ depends on the product of six atoms. Only combinations with an equal number of conjugate/non-conjugate pairs are non-zero, as depicted in Fig. 9. Each combination is weighted

by the corresponding symbol-cumulant [20]. Please note that terms labeled with - in Fig. 9 correspond to the terms removed by the CPE in (27) [22].

The indexing of the valid combinations yields the terms in (28). Hence, for instance, the PDL-GN model reduces to the last triple summation in (28).

$\mathbf{k}^*$	$\mathbf{m}$	$\mathbf{n}$	$\mathbf{l}$	$\mathbf{j}^*$	$\mathbf{o}^*$	
●	●	●	●	●	●	Q6
●	●	○	●	●	○	F4
●	●	○	●	○	●	F4
●	○	●	●	●	○	F4
●	○	●	●	○	●	F4
●	●	●	○	●	○	-
●	●	●	○	○	●	-
○	●	○	●	●	●	-
○	○	●	●	●	●	-
○	●	●	○	●	●	Q4
●	●	○	△	△	○	-
●	●	○	△	○	△	-
●	○	●	△	△	○	-
●	○	●	△	○	△	-
●	△	○	●	△	○	GN
●	△	○	●	○	△	GN

Fig. 9. Valid combinations yielding non-zero $\mathbb{E}\left[a^*_{\mathbf{k}}a_{\mathbf{m}}a_{\mathbf{n}}a_{\mathbf{l}}a^*_{\mathbf{j}}a^*_{\mathbf{o}}\right]$. For instance, the second row indicates the combination $(\mathbf{k}=\mathbf{m}=\mathbf{l}=\mathbf{j}) \neq (\mathbf{n}=\mathbf{o})$. The labels indicate sixth-order noise (Q6), two types of fourth-order noise (F4 and Q4) and second-order noise, usually called GN contribution. A (-) indicates a phase contribution that is removed by the CPE.

Although the master theorem reduces the number of summations, the final result (28) still depends on an infinite summation over the discrete-time index. However, in the special case of sinc pulses, such a summation can be dropped, with significant simplifications, as already observed in [24], [21], thanks to the Poisson summation formula:

$$\sum_{k=-\infty}^{\infty} e^{jk\omega T} = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{2\pi k}{T}\right) \tag{29}$$

and the finite bandwidth of the pulses that can interact with only one Dirac's delta in such a summation. Instead of entering the fine mathematical details of the proof, the next theorem provides a short-rule for the simplifications:

Theorem 2: With sinc pulses $p_h(t) = \mathrm{sinc}(t/T)$, $\forall h$, all summations in (28) over *temporal* indexes can be dropped. For each drop, an integral in the corresponding tensor-product $\mathcal{X}_{\mathbf{kmni}}\mathcal{X}^*_{\mathbf{ljoi}}$ can be dropped as well. The dropped integral can be identified by solving a linear system obtained by equating the arguments ψ of the $|\tilde{G}_{\mathbf{k}}(\psi)\rangle$ involved in the product owing to the same atom. □

Such a result is best explained by an example. Let us focus on the particular GN-term $\mathcal{G} = \mathcal{X}_{\mathbf{kmni}}\mathcal{X}^*_{\mathbf{kmni}}$ of (28) in the scalar case. This term can be explicitly expanded following (13) as:

$$\begin{aligned}\mathcal{G} = \sum_{t,r=0}^{N-1}\iiiint_{-\infty}^{\infty} &\eta_t(\omega,\omega_1,\omega_2)\eta^*_r(\mu,\mu_1,\mu_2)\\ &\times\tilde{G}^*_{\mathbf{k}}(\omega+\omega_1+\omega_2)\tilde{G}_{\mathbf{m}}(\omega+\omega_2)\tilde{G}^*_{\mathbf{i}}(\omega)\tilde{G}_{\mathbf{n}}(\omega+\omega_1)\\ &\times\tilde{G}_{\mathbf{k}}(\mu+\mu_1+\mu_2)\tilde{G}^*_{\mathbf{m}}(\mu+\mu_2)\tilde{G}_{\mathbf{i}}(\mu)\tilde{G}^*_{\mathbf{n}}(\mu+\mu_1)\\ &\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}\frac{\mathrm{d}\omega}{2\pi}\frac{\mathrm{d}\mu_1}{2\pi}\frac{\mathrm{d}\mu_2}{2\pi}\frac{\mathrm{d}\mu}{2\pi}\,.\end{aligned}$$

Equating the arguments of equal atoms yields the following linear system:

$$\begin{cases}\mathbf{k} & \to \omega+\omega_1+\omega_2 = \mu_1+\mu_2+\mu\\ \mathbf{m} & \to \omega+\omega_2 = \mu+\mu_2\\ \mathbf{n} & \to \omega+\omega_1 = \mu+\mu_1\end{cases}$$

whose solution is $\omega_1 = \mu_1,\ \omega_2 = \mu_2,\ \omega = \mu$. We can thus drop three integrals, for instance the ones with (μ,μ_1,μ_2), by using the previous substitution.

REFERENCES

[1] J. N. Damask, *Polarization Optics in Telecommunications*. New York, NY, USA: Springer, 2005.

[2] A. Dumenil, E. Awwad, and C. Méasson, "Polarization Dependent Loss: Fundamental Limits and How to Approach Them," in Proc. *Advanced Photonics Congress*, New Orleans, LA, USA, 2017, Paper SpM4F.1.

[3] A. Nafta, E. Meron, and M. Shtaif, "Capacity limitations in fiber-optic communication systems as a result of polarization-dependent loss," *Opt. Lett.*, vol. 34, no. 23, pp. 3613–3615, Dec. 2009.

[4] L. E. Nelson et al., "Statistics of polarization dependent loss in an installed long-haul WDM system," *Opt. Express*, vol. 19, no. 7, pp. 6790–6796, Mar. 2011.

[5] N. Gisin, "Statistics of polarization dependent losses," *Opt. Commun.*, vol. 114, nos. 5–6, pp. 399–405, Feb. 1995.

[6] A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 14, no. 3, pp. 313–315, Mar. 2002.

[7] A. Mecozzi and M. Shtaif, "Signal to noise ratio degradation caused by polarization dependent loss and the effect of dynamic gain equalization," *J. Lightw. Technol.*, vol. 22, no. 8, pp. 1856–1871, Aug. 2004.

[8] M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," *Opt. Express*, vol. 16, no. 18, pp. 873–879, Sep. 2008.

[9] M. Karlsson and M. Petersson, "Quaternion Approach to PMD and PDL Phenomena in Optical Fiber Systems," *J. Lightw. Technol*, vol. 22, no. 4, pp. 1137–1146, Apr. 2004.

[10] Z. Tao, L. Dou, T. Hoshida, and J. C. Rasmussen, "A fast method to simulate the PDL impact on dual-polarization coherent systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 24, pp. 1882–1884, Dec. 2009.

[11] C. Xie, "Polarization-dependent loss induced penalties in PDM-QPSK coherent optical communication systems," in Proc. *Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2010, Paper OWE6.

[12] O. Vassilieva, S. Oda, T. Hoshida, J. C. Rasmussen, and M. Sekiya, "Experimental investigation of the statistics of the interplay between nonlinear and PDL effects in coherent polarization multiplexed systems," in Proc. *Opt. Fiber Commun. Conf.*, Ananheim, CA, USA, 2013, Paper OM3B.6.

[13] I. F. J. Ruiz, A. Ghazisaeidi, E. Awwad, P. Tran, and G. Charlet, "Polarization effects in nonlinearity compensated links," in Proc. *Eur. Conf. Optical Commun.*, Düsseldorf, Germany, 2016, Paper We.1.3.3.

[14] H.-M. Chin et al., "Probabilistic Design of Optical Transmission Systems," *J. Lightw. Technol*, vol. 35, no. 4, pp. 931–940, Feb. 2017.

[15] N. Rossi, S. Musetti, P. Ramantanis, and S. Almonacil, "The Impact of Kerr Nonlinearity on the SNR Variability Induced by Polarization Dependent Loss," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5048–5055, Oct. 2019.

[16] A. I. A. El-Rahman and J. C. Cartledge, "Implications of Distributed Link Polarization-Dependent Loss on Bitwise Achievable Information Rates for Probabilistically Shaped and Uniform DP 64-QAM," *J. Lightw. Technol.*, vol. 37, no. 4, pp. 1187–1194, Feb. 2019.

[17] J. C. Cartledge and A. I. A. El-Rahman, “Implications of Polarization Dependent Loss on the Performance of Probabilistically Shaped 64-QAM,” in *Proc. Eur. Conf. Opt. Commun.*, Dublin, Ireland, 2019, Paper Th.2.B.5.
[18] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-model of fiber nonlinear propagation and its applications,” *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.
[19] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, “The EGN model of non-linear fiber propagation,” *Opt. Express*, vol. 22, no. 13, pp. 16335–16362, Jun. 2014.
[20] P. Serena and A. Bononi, "A Time-Domain Extended Gaussian Noise Model," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1459–1472, Apr. 2015.
[21] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “Accumulation of nonlinear interference noise in multi-span fiber-optic systems,” *Opt. Express*, vol. 22, no. 12, pp. 14199–14211, Jun. 2014.
[22] P. Serena and A. Bononi, “An Alternative Approach to the Gaussian Noise Model and its System Implications,” *J. Lightw. Technol.*, vol. 31, no. 22, pp. 3489–3499, Nov. 2013.
[23] P. Serena *et al.*, “The Gaussian noise model extended to polarization dependent loss and its application to outage probability estimation,” in *Proc. Eur. Conf. Opt. Commun.*, Rome, Italy, 2018, Paper Tu4G.4.
[24] A. Mecozzi and R. Essiambre, “Nonlinear Shannon limit in pseudolinear coherent systems,” *J. Lightw. Technol.*, vol. 30, no. 12, pp 2011-2024, Jun. 2012.
[25] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “Properties of nonlinear noise in long , dispersion-uncompensated fiber links,” *Opt. Express*, vol. 21, no. 22, pp. 25685–25699, Oct. 2013.
[26] A. Papoulis, *Probability, random variables, and stochastic processes*, 3rd ed. New York, NY, USA: McGraw-Hill, 1991.
[27] S. Musetti, P. Serena, and A. Bononi, “On the accuracy of split-step Fourier simulations for wideband nonlinear optical communications,” *J. Lightw. Technol.*, vol. 36, no. 23, pp. 5669–5677, Dec. 2018.
[28] E. Grellier, and A. Bononi, “Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise,” *Opt. Express*, vol. 19, no. 13, pp. 12781–12788, Jun. 2011.