A Time-Domain Extended Gaussian Noise Model

Paolo Serena, *Member, IEEE*, and Alberto Bononi, *Senior Member, IEEE*

(Invited Paper)

***Abstract*—From first-order perturbation theory, we derive the autocorrelation function of the nonlinear interference in coherent optical links. We show that the fundamental assumptions of the Gaussian noise (GN) model regarding stationary Gaussian statistics of the transmitted signal can be removed for a more complete model accounting for the fine details of the cyclostationary modulation format. We first give an intuitive presentation of the theory, and then provide a formal mathematical treatment based on symbol cumulants and discuss its key assumptions and limitations. The proposed model includes dual polarization effects, wavelength-division multiplexing and the nonlinear signal to amplified spontaneous emission noise interaction along the line, thus neglecting only the impact of four-wave mixing.**

***Index Terms*—Cumulants, GN model, nonlinear interference.**

I. Introduction

THE search of an effective model yielding the statistical properties of the nonlinear interference (NLI) in coherent links is an endeavor that stimulated lots of research in recent years [1]–[9]. The key observation of Gaussian statistics of the detected signal in dispersion uncompensated (DU) links [10] spurred simplified analyses of the coherent receiver performance [11], [12], re-aligning bit error rate (BER) estimation of practical DU optical coherent systems to the additive Gaussian noise (GN) paradigm of digital communications textbooks [13]. The problem thus became one of finding reliable models of the NLI variance or, more generally, of the NLI power spectral density (PSD) function.

Among the several approaches proposed in the literature, the most successful was the GN model [1], [3] which essentially assumed Gaussian statistics of the signal at any point along the optical link and even at its input, an assumption motivated by the strong inter-symbol interference (ISI) induced by the large fiber dispersion cumulated along typical DU links. The GN model expresses the NLI as an additive perturbation and has numerical complexity dramatically reduced compared with experiments or typical split step Fourier (SSF) simulations. Such an advantage comes at the expense of a reduced accuracy, especially for short DU links or legacy dispersion managed (DM) links where the Gaussian signal assumption is arguable [14]. The problem is partially relaxed for long DU links, e.g., close to the reach of practical coherent systems as the ones analyzed in this manuscript, thanks to the larger dispersion cumulated in the link.

Manuscript received October 6, 2014; revised December 23, 2014 and January 28, 2015; accepted January 29, 2015. Date of publication February 1, 2015; date of current version March 4, 2015.

The authors are with the Department of Information Engineering, Università degli Studi di Parma, 43124 Parma, Italy (e-mail: paolo.serena@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 10.1109/JLT.2015.2398873

Such GN model accuracy problems motivated the search of extended GN models able to cope with the actual signal modulation. The first of such models was proposed by Mecozzi and Essiambre [4], where the variance of the NLI was evaluated in the time domain in cross-phase modulation (XPM) dominated links. The XPM-model in [4] was later exploited in [15] by working in the frequency domain to investigate some limitations of the GN model. Recently, Carena *et al.* proposed a frequency-domain enhanced GN (EGN) model that includes all the NLI components, namely, self-phase modulation (SPM), XPM, cross polarization modulation (XPolM) and four-wave mixing (FWM) [9].

Unlike [9], in this work, which is an extended version of [16], we took a time-domain approach to extend the XPM perturbation-model in [4] to include all NLI components but FWM for polarization division multiplexed (PDM) transmissions, by following the approach proposed in [17]. Our model is not just an alternative derivation of the EGN in [9], but aims to provide very general results that fully exploit the cyclostationarity of the input digital signals. Moreover, our proposed EGN leads to an efficient, fast Fourier transform (FFT) based algorithm that evaluates the entire auto-correlation function of the received NLI, and not just its variance.

The modulation-format aware EGN model is more accurate than the GN model, as shown both in [9], [15], [18] and in this work. In this paper, for the first time, we show that the EGN can be used not just for DU optical links, but with reasonable accuracy even for DM optical links, where the GN model grossly fails. Moreover, we show that with the proposed time-domain EGN model the nonlinear signal/amplified spontaneous emission noise (ASE) interaction along the optical line can be easily accounted for, thus finessing the problems encountered by the frequency domain approach [19].

The paper is organized as follows: Section II introduces a simplified description of the perturbative model underlying all the GN and EGN models and provides the key ideas of the derivation of our proposed EGN model, while Section IV provides its rigorous mathematical derivation. In Section III we show numerical results supporting the theory. In Section V we comment on the limits of the EGN model and their connection with the underlying model assumptions. Finally in Section VI we draw our main conclusions. Theorems supporting the theory can be found in the appendices.

II. Intuitive Approach

In this section we give an intuitive description of the statistical properties of the NLI. Then in Section IV we will provide their rigorous derivation.

Many studies showed that in most practical cases the highly dispersive nonlinear channel with coherent detection can be rea-

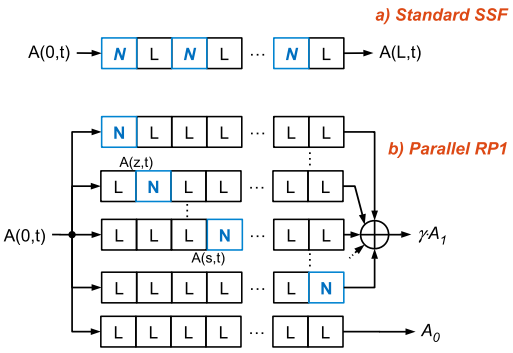

Fig. 1. SSF and RP1 algorithm. N: block implementing Kerr effect [17]. L: block implementing linear effects.

sonably modeled as an additive GN channel, so that performance takes well-known simple textbook expressions [6], [10]–[12]. NLI was thus introduced as a way to identify the additive impairment brought by the nonlinear effect and distinguish it from ASE noise. An additive interference naturally leads to perturbation theory, which in most of the proposed models of nonlinear optical links takes the form of a first order regular perturbation (RP1) [20], where NLI scales linearly with γ [2]–[7], [9], [21]–[23]. If we look at the nonlinear Schrödinger equation from the SSF perspective as an (infinite) concatenation of linear/nonlinear steps, it should be clear that the only way to get an RP1 output linearly related to γ is to account for all possible crossings of a *single* nonlinear block along the line, as depicted by Fig. 1 [23].

From Fig. 1 we also realize that, in highly dispersive links, in many branches of the RP1 diagram the signal cumulates a strong dispersion before entering the nonlinear block. This observation may justify the stationary Gaussian assumption for the input signal, since a strong dispersion causes a large ISI which both induces first-order Gaussian statistics for the signal field and apparently attenuates its symbol rate cyclic properties. However, it also highlights the weakness of such an assumption, since along paths with an early presence of the nonlinearity (top branches in Fig. 1(b)) the dispersive effect is limited [14].

How do we evaluate the auto-correlation function of the received NLI from the above diagram? Since NLI is a summation of contributions, the task translates into finding the cross-correlation between the signals outgoing any two generic paths. Moving backwards on such paths from the output, relating the cross-correlation to the corresponding one of the signals outgoing the nonlinear blocks is a matter of pure linear system theory. Similar reasoning can be applied between the transmitter and the input of the nonlinear blocks, where things are even simpler, since a linearly modulated digital signal just changes the supporting pulse after filtering. Thus, given the digital signal $A(z,t)=\sum_k a_k p(z,t-kT)$ with zero mean independent symbols a_k, the correlation between $A(z,t)$ and $A(s,t)$ is simply:

$$\begin{aligned} R(t,\tau) &= E\left[A(z,t+\tau)A^*(s,t)\right] \\ &= \sum_{k,n} E[a_k a_n^*]p(z,t+\tau-kT)p^*(s,t-nT) \\ &= \sum_{k} E[|a_k|^2]p(z,t+\tau-kT)p^*(s,t-kT) \end{aligned} \quad (1)$$

since only identical indexes yield a non-zero average value. Moreover, $A(z,t)$ is cyclostationary, hence one usually removes the dependence on t by averaging (1) over one symbol time [30], thus obtaining $R(\tau)\propto p(z,\tau)\otimes p^*(s,-\tau)$ (see Appendix C), with $\otimes$ indicating time-convolution. Let us now evaluate the cross-correlation between $V(z,t)$ and $V(s,t)$, with $V(z,t)\triangleq|A(z,t)|^2A(z,t)$. We now have to deal in principle with six summations, which can be reduced to less for the same reasoning as in (1). In fact, a symbol participates to a non-zero average value only if it is taken in absolute value. This leaves us with just three kinds of summations: one involving $\{|a_k|^2|a_n|^2|a_l|^2\}$ with[1] $k\neq n\neq l$, one with $\{|a_k|^4|a_n|^2\}$ with $k\neq n$ and one for $\{|a_k|^6\}$. We are thus partitioning the problem into subproblems. Let us look at the first partition, $\{|a_k|^2|a_n|^2|a_l|^2\}$. Such a partition generates three summations in k, n, l, each identical to (1), whose product is thus proportional to $|R(\tau)|^2R(\tau)$. This way, the only detail of the modulation format that matters here is the power $E[|a_k|^2]$. Hence, at fixed signal power, this contribution is format-independent.

The other contributions instead involve forth and sixth-order moments of the modulation format, and at fixed signal power they are therefore format-dependent. Let us concentrate on the contribution given by the partition created by $\{|a_k|^6\}$. Its contribution to the cross-correlation is much like (1), thus with just one summation justifying all equal indexes with $|p(z,t)|^2p(z,t)$ in place of $p(z,t)$. After averaging, we thus get a term proportional to $(|p(z,\tau)|^2p(z,\tau))\otimes(|p(s,-\tau)|^2p^*(s,-\tau))$. Similar reasoning can be applied to the partition created by $\{|a_k|^4|a_n|^2\}$. We thus expect the correlation between $V(z,t)$ and $V(s,t)$ to be a combination of convolutions between nonlinear functions of the pulse, thus of similar complexity to (1).

At this point the question is: which is the most important term? Even neglecting attenuation, the pulse $p(z,t)$ broadens due to dispersion, and its height asymptotically decreases with distance z at rate $1/\sqrt{z}$ to preserve energy. Hence contributions fade away at increasing z and s, with one exception: $Q(z,s,\tau)\triangleq p(z,\tau)\otimes p^*(s,-\tau)$, due to $\{|a_k|^2|a_n|^2|a_l|^2\}$ only, is distance independent for $s\to z$. As we noted before, this term is modulation format independent: not surprisingly, it is the only one accounted for by the GN model. Its double integration along s and z that creates the NLI correlation function is expected to be dominated by the strip $s\approx z$, not just because of the above fading, but also for the oscillating behavior of $Q(z,s,\tau)$ for $s\neq z$. It thus contributes to the overall correlation by an almost linear growth with spans. The only term that can compete with this one is the one related to $F_4(z,s,\tau)\triangleq|p(z,\tau)|^2\otimes|p(s,-\tau)|^2$, due to $\{|a_k|^4|a_n|^2\}$, which is indeed always positive at any coordinate, such that its fading in z,s is eventually balanced by the double integration. However, attenuation breaks such a balance, thus making the contribution of $F_4(z,s,\tau)$ to grow more slowly over the initial spans than the GN contribution. In conclusion, in all practical lossy systems, the GN contribution is the most important.

While this section introduced the reader to the key ideas, the exact derivation of the NLI auto-correlation in the RP1 model, as well as a check of the claims made here, is deferred to Section

[1]The fact that indexes are different may seem a minor point. Indeed, it implies a non trivial combinatorial problem that can be solved by introducing the concept of cumulant, as we will see in Appendix B.

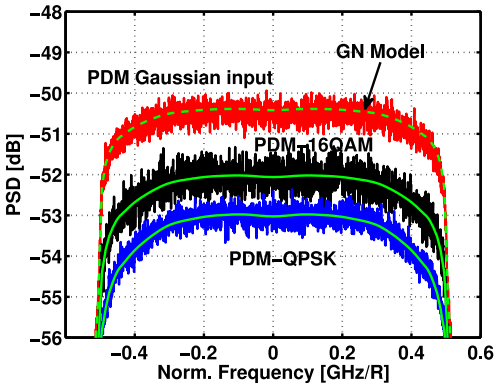

Fig. 2. NLI PSD from MC simulations (noisy curves) the GN model (dashed smooth lines) and our EGN model (solid smooth lines). 15 WDM channels, 32 Gbaud, spacing 37.5 GHz, 5×100 km SMF DU link.

IV. The next section anticipates numerical checks using our EGN model.

III. Numerical Checks

We tested the proposed EGN by comparing its predicted NLI PSD against predictions of both of the GN model and of SSF Monte–Carlo (MC) simulations. We applied the model to both PDM-quadrature phase shift keying (QPSK) and PDM-16 quadrature amplitude modulation (QAM), while to double-check the GN Model we transmitted a Gaussian process with the same PSD as the modulated cases. The optical link was either DU or DM with residual dispersion per span $D_{\rm in} = 30$ ps/nm (DM30). Transmission spans were composed of single mode fibers (SMF; Attenuation $\alpha = 0.2$ dB/km; dispersion $D = 17$ ps/nm/km; nonlinear coefficient $\gamma = 1.3$ W^{-1} km^{-1}; length: 100 km) with end-span amplification. Fiber propagation was modeled by the Manakov equation without polarization mode dispersion [24]. Before transmission, in the DM30 case we used a pre-compensating fiber equal to $-D/\alpha - D_{\rm in}(N_{\rm s}-1)/2$, with $N_{\rm s}$ the number of spans.

As a sanity check, we first estimated the PSD of the received NLI by SSF MC simulations in the DU case. We estimated the NLI PSD by the periodogram method averaged over 100 runs of 4096 symbols each. Symbol rate was $R = 32$ Gbaud, for a total of 15 channels of power -4 dBm each, with channel spacing $\Delta f = 37.5$ GHz. In this setup we shaped the pulses as $\mathrm{sinc}(tR)$, and then modulated the laser sources with a random state of polarization from channel to channel.

If $s(t)$ and $s_{\rm tx}(t)$ are the signals after the receiver optical matched filter with/without propagation along the link, respectively, and in absence of ASE, in MC simulations we estimated the NLI field as $s_{\rm NLI} = s(t)e^{-j\langle\phi\rangle} - s_{\rm tx}(t)$, where $\langle\phi\rangle$ is the average nonlinear phase induced by the Kerr effect [17], [23]. The NLI PSD for a 5 span DU system is reported in Fig. 2. We note that both the GN model and our EGN model PSD predictions (smooth lines) well match the average PSD obtained from MC simulations (noisy curves). It is also worth noting that the GN model always yields a larger NLI PSD compared with the EGN PSD that accounts for the details of the digital input. 16 QAM is closer to the GN PSD thanks to its larger constellation that is "more similar" to a circular Gaussian distribution than

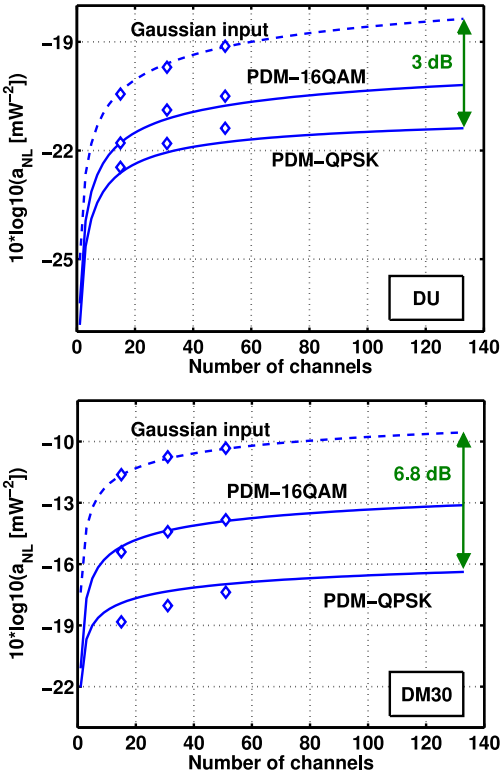

Fig. 3. Normalized NLI variance ($\sigma^2_{\rm NLI} \triangleq a_{\rm NL}P^3$, Ppower) versus number of channels. 10×100 km SMF link at 32 Gbaud, spacing 37.5 GHz. Solid lines: EGN. Dashed lines: GN model. Symbols: MC SSF simulations. DU: dispersion uncompensated. DM30: dispersion managed with residual 30 ps/nm/span and pre-compensation.

the QPSK. The area under the PSD is the NLI variance $\sigma^2_{\rm NLI}$, which can be used to evaluate the signal to noise ratio (SNR). When the RP1 holds, such a variance scales with the cube of the average power P [11], [12], and it is customary to provide the normalized NLI coefficient $a_{\rm NL}$, defined by $\sigma^2_{\rm NLI} \triangleq a_{\rm NL}P^3$. In our 5×100 km SMF DU link we estimated by SSF MC simulations $a^{\rm dB}_{\rm NL} \triangleq 10\log_{10}(a_{\rm NL}\,[\mathrm{mW}^{-2}]) = -23.5$, -25.1, -26.3 dB for the Gaussian, QAM and QPSK modulated cases, respectively. Our EGN model provided the same values, within an error around 0.1 dB. In the QPSK case we also tried to estimate $a_{\rm NL}$ by extracting NLI as in [26] and [37], as a way to capture phase noise, and obtained a negligible difference with the previous measurement (close to 0.1 dB). Such an error can be larger in a DU network scenario with different channels experiencing different light-paths [26].

In a second set of simulations we checked how $a_{\rm NL}$ scales with the number of channels. The link was now 10×100 km SMF, either DU or DM30, and thus close to the reach of the 16 QAM case (see next Fig. 4). Results are summarized in Fig. 3, again by using MC simulations (symbols), our EGN model (solid) and the GN model (dashed). MC simulations used a maximum of 16384 symbols per channel, and were limited to a maximum of 51 channels for feasible execution times. The good agreement between MC simulations and our EGN is an indication of the EGN accuracy. However, the NLI PSD with our EGN can be evaluated for many more channels than an MC simulation, here up to 133 channels to fill the entire bandwidth

(5 THz) of erbium doped fiber amplifiers (EDFA). From the figure we make the following observations: 1) GN gap to EGN increases with increasing number of channels[2] : for instance, in the DU case with PDM-QPSK the gap at five channels is ∼2 dB, while at 133 channels it is ∼3 dB; 2) while the GN model curve well fits a logarithmic scaling with the number of channels, the scaling of the EGN curve appears to be slower; 3) $a_{\rm NL}$ for EGN 16QAM is closer to the GN prediction due to a more Gaussian-like constellation; 4) as expected, GN model error is significant in the DM30 case; 5) the EGN model works with reasonable accuracy even in the DM case, where assumption iii) of Section IV may be questionable, thus yielding a non-circularly distributed NLI with nonlinear phase noise.

It is worth noting that in DM most of the error comes from single channel (in Fig. 3 we read 4.6 dB between GN and QPSK curves at one channel, while the gap is 6.8 dB at 133 channels), an indication that may be useful for numerical modeling since it suggests to evaluate cross-channel effects with the GN model, thus restricting the computationally-heavier EGN model to SPM effects. This issue will be investigated in future studies.

Having checked the accuracy of the NLI variance predicted by the GN/EGN model, we move to analyze a more significant quality of transmission parameter, i.e., the power yielding a Q-factor of 8.5 dB (BER of 4×10^{-3}). At a given number of spans there are two powers achieving the target BER, one in the ASE noise dominated regime, one in the NLI dominated regime. We estimated these powers in [27] through MC simulations by counting at least 400 errors and averaging over ten different random realizations of input states of polarization and symbol patterns. ASE was loaded at the receiver since for such systems the nonlinear signal-noise interaction is negligible [19], [27] at BER around 10^{-3} with an equivalent noise figure per amplifier of 6 dB. Pulses were sinc shaped, modulated with a symbol rate of 10, 32, or 80 Gbaud at fixed bandwidth efficiency $\Delta f/R = 0.85$ (e.g., Δf =37.5 GHz at R =32 Gbaud). We fixed the total aggregate bandwidth[3] to 562.5 GHz (e.g., 15 channels at 32 Gbaud). The number of symbols was set to capture the maximum walk-off of the outer channel, with a minimum of 4096.

At the receiver the central channel was filtered by a matched filter, sampled at two samples per symbol, and processed by a digital signal processing including trained least squares equalization with 15 taps and blind phase estimation with 27 taps. Symbols were differentially encoded.

Fig. 4 shows the BER contour levels versus both power and number of spans for an $N_{\rm s} \times 100$ km SMF link either DU (top row) and DM30 (bottom row). Symbols denote MC simulations, solid lines the analytical contours in [28, eq. (8)] where the NLI variance from our EGN model was used, while dashed lines indicate the same analytical contours where the GN model was used. With GN and EGN models we first evaluated $a_{\rm NL}$, then converted it to BER as in [1] using the true back-to-back performance of our system.

In the DU link, we observe that by including modulation dependent NLI terms in the RP1 model the match with SSF simulations is excellent, with a minor error at 10 Gbaud where FWM starts to play a role. The GN model error is almost 1 dB on the NLI asymptote. This value is smaller than the error on variance observed in Fig. 3 because the power $P_{\rm M}$ representing the NLI asymptote, where ASE noise is negligible, is [28] $P_{\rm M} = 1/\sqrt{S_0 a_{\rm NL}}$, S_0 being the SNR yielding Q-factor = 8.5 dB. The presence of the square root halves variance errors in a decibel scale. Regarding the reach, it was theoretically proved that an error of $\Delta\alpha$ [dB] in $a_{\rm NL}$ translates into an error of $\Delta N_0 = -\frac{\Delta\alpha}{3+\epsilon}$ [dB] in reach, $0 < \epsilon < 1$ being the noise accumulation factor [28]. Thus the dB error on reach is at least 1/3 smaller than the dB error on $a_{\rm NL}$. For these reasons, reach estimations are well tolerant to NLI modeling errors, so that the GN model, despite its non negligible inaccuracy in predicting variance, works quite well in predicting the reach [29]. Note the good match of the EGN contours with MC simulations also at one span for $R >$ 10 Gbaud, an indication that the BER is well approximated, for the systems analyzed in Fig. 4, by a function of the SNR only.

The bottom row of Fig. 4 refers to the DM30 case. Here the NLI Gaussian assumption may be questionable, even at the receiver side. Not surprisingly, the GN model error is most of the times unacceptable, e.g., 50% in reach prediction at 32 Gbaud PDM-QPSK. The good match between GN and MC simulations at 10 Gbaud and 16 QAM is just a coincidence. Our EGN RP1 model including modulation dependent NLI shows instead good accuracy, except at smaller symbol rates where the circularity of the received NLI statistics fails. For QAM, modulation dependent NLI and GN curves are closer, because of the faster convergence to Gaussian-like first order statistics of the propagating field. However, it is worth mentioning that the GN model always provides a lower bound on maximum reach.

IV. Rigorous Approach

In this section we derive the NLI auto-correlation function by a rigorous approach. We adopt the following assumptions:

i) a linearly modulated digital signal:

$$A(0,t) = \sum_k a_k p(0, t - kT) \tag{2}$$

with a_k zero mean, independent and identically distributed (i.i.d.) digital symbols with fourfold rotational symmetry (see Appendix B), and $p(0,t) \in \mathbb{C}$ the supporting pulse at coordinate $z = 0$ and time t;

ii) the received signal can be approximated by a RP1 solution;

iii) NLI is statistically independent of the symbol of interest;

iv) FWM can be neglected.

Please note that binary-PSK (BPSK) does not satisfy assumption i), as discussed in Appendix B, while assumption iv) is actually not necessary, but it simplifies a lot the analysis. GN model uses hypotheses i)–iii) as well, with a main difference in i), where $A(0,t)$ is substituted with a stationary Gaussian process [1], [17], [22].

A. RP of the Electric Field

We start by deriving the RP1 solution from the nonlinear Schrödinger equation using the operator notation (α: attenuation; $\beta_{2,3}$: second/third order dispersion; γ: nonlinear coeffi-

[2]See also the discussion at point IV of Section V.

[3]Similar curves but at a fixed number of channels and at $\Delta f/R = 0.56$ can be found in [16].

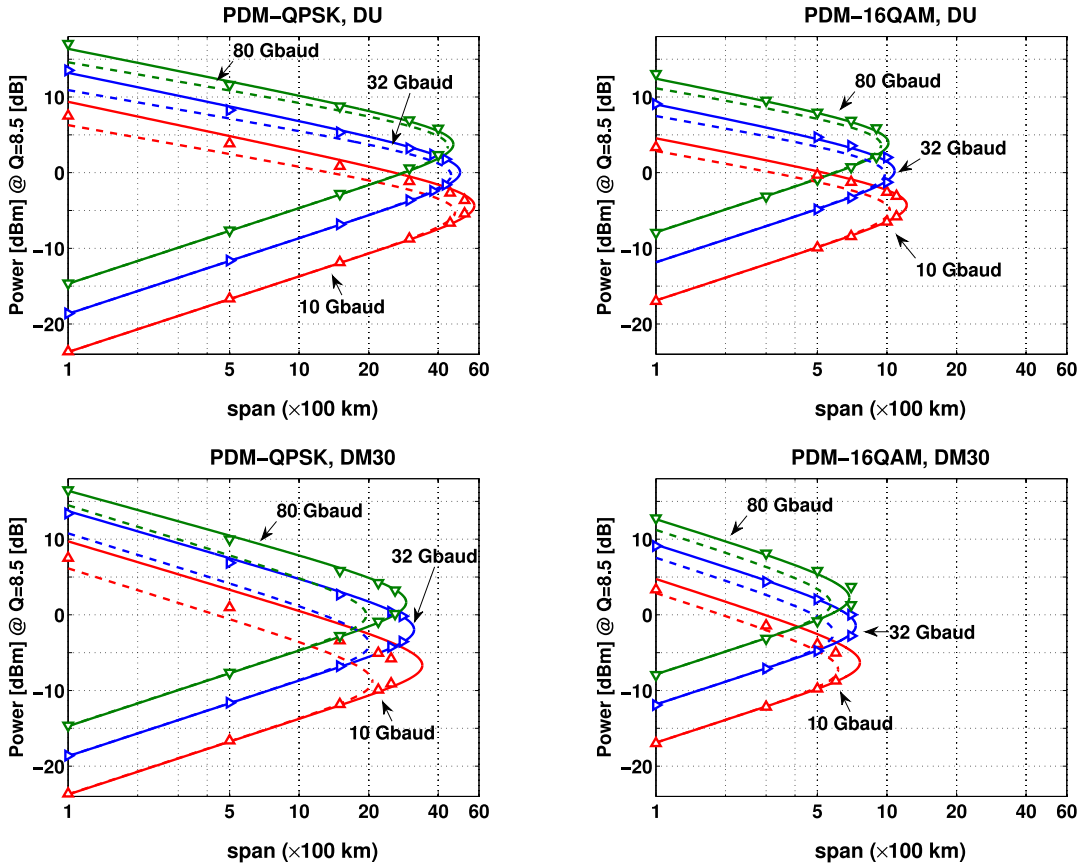

Fig. 4. Power yielding Q-factor = 8.5 dB versus number of spans. Solid lines: our EGN model. Dashed lines: GN Model. Symbols: MC Split-step simulations. Amplifiers noise figure F = 6 dB, span loss = 20 dB.

cient):

$$\frac{\partial A}{\partial z} = (\mathcal{L} - j\gamma\mathcal{N})\, A \tag{3}$$

where $\mathcal{L} \triangleq -\frac{\alpha}{2} + j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3}{\partial t^3}$ and $\mathcal{N}A \equiv \mathcal{N}(A) \triangleq |A|^2 A$. Using z-independent parameters to simplify notation[4], using the integrating factor $A(z,t) = e^{\mathcal{L}z}U(z,t)$ we can rewrite (3) in integral form as [4]:

$$U(z,t) = U(0,t) - j\gamma \int_0^z e^{-\mathcal{L}\xi}\mathcal{N}\left(e^{\mathcal{L}\xi}U\right)\mathrm{d}\xi. \tag{4}$$

Despite the elegant form expressed by $A = e^{\mathcal{L}z}U$, its practical meaning is a convolution with a linear impulse response $h_{0z}(t)$ accounting for all linear effects from coordinate 0 to z, i.e., $e^{\mathcal{L}z}U \equiv h_{0z}(t) \otimes U(0,t)$ [17]. In the general case, the Fourier transform of $h_{0z}(t)$ is:

$$\tilde{h}_{sz}(f) \triangleq \mathscr{F}\{h_{sz}(t)\} = \sqrt{G(s,z)}e^{-j\frac{\int_s^z \beta_2(x)\mathrm{d}x}{2}\omega^2 - j\frac{\int_s^z \beta_3(x)\mathrm{d}x}{6}\omega^3}$$

with $\omega \triangleq 2\pi f$, and $G(s,z)$ the cumulated gain from coordinate s to z.

Since in real conditions the right hand side of (4) is bounded, a brute force solution of (4) is obtained by the Picard recursion that iteratively solves (4) by substituting U_{k+1} and U_k in place of $U(z,t)$ in the left/right hand side, respectively. If the nonlinear effect is small, the RP1 idea is to stop such an iteration at first term, thus yielding $U(z,t) \simeq U_0(z,t) + U_1(z,t)$ with the *unperturbed* solution

$$U_0(z,t) \triangleq h_{0z}(t) \otimes U(0,t) = \sum_k a_k p(z, t-kT) \tag{5}$$

which defines $p(z,t) \triangleq h_{0z}(t) \otimes p(0,t)$, and *first order perturbation*

$$U_1(z,t) = -j\gamma \int_0^z h_{\xi 0}(t) \otimes \mathcal{N}\Big(h_{0\xi}(t) \otimes U(0,t)\Big)\mathrm{d}\xi$$

where $h_{\xi 0}(t) = h^*_{0\xi}(-t)/G(0,\xi)$. The perturbed solution can thus be read as an infinite summation of echoes induced by nonlinearity. To complete our model of the optical system, we must introduce in the discussion the receiver, placed at coordinate L. The receiver performs correlation detection and phase recovery, while frequency recovery can be overlooked since we are implicitly assuming that all local oscillators are ideal. Correlation detection is made by filtering the detected signal with the receiver filter $h_{\mathrm{r}}(t)$. $h_{\mathrm{r}}(t)$ is designed (normally with the help of a fractionally spaced equalizer at the receiver) such that the unperturbed signal after it, $I_0(t) = h_{\mathrm{r}}(t) \otimes U_0(L,t)$, has supporting pulse close to satisfying both the Nyquist condition of zero ISI and the matched filter condition. Similarly, the perturb-

[4]In the general case one should substitute $e^{\mathcal{L}z}$ with $e^{\int_0^z \mathcal{L}(\xi)\mathrm{d}\xi}$ and include $\gamma(z)$ within $\mathcal{N}$.

bation $I_1(t)$ is:

$$I_1(t) = -j\gamma \int_0^L h(\xi,t) \otimes V(\xi,t)\mathrm{d}\xi \tag{6}$$

where:

$$h(\xi,t) \triangleq h_{\mathrm{r}}(t) \otimes h_{\xi 0}(t) \tag{7}$$

and $V(z,t) = |A(z,t)|^2 A(z,t)$ explicitly is:

$$V(z,t) = \sum_{k,n,l} a_k a_n^* a_l p(z,t-kT)p^*(z,t-nT)p(z,t-lT). \tag{8}$$

Carrier phase estimation (CPE) is an important issue since the nonlinear Kerr effect induces phase rotation. We assume a long window CPE which essentially recovers the average phase rotation. This way, the digital CPE can be performed as well in the analog domain and corresponds to solving (3) with $\left(\mathcal{N} - \overline{P}\right) A$ in place of $\mathcal{N}A$, with $\overline{P}$ a proper power. The corresponding new RP solution is called enhanced RP (eRP). It was shown [17] that for highly dispersive systems the value of $\overline{P}$ is twice the average power in single polarization links.

B. Correlation Function of the Perturbation

Since all symbols are unknown by assumption, $I_1(t)$ is wide-sense cyclostationary as it is a function of the cyclostationary signal $A(0,t)$ in (2). Consequently, the auto-correlation function of $I_1(t)$ is:

$$R_L(t,\tau) = E\left[I_1(t+\tau)I_1^*(t)\right] = R_L(t+nT,\tau), \quad n \in \mathbb{Z}.$$

The periodicity in t can be efficiently described by a Fourier series, yielding:

$$R_L(t,\tau) = \sum_k R_L^{(k)}(\tau) e^{j2\pi k \frac{t}{T}}$$

$R_L^{(k)}(\tau)$ is called the cyclic auto-correlation function at *cycle frequency* k/T. By analogy, its Fourier transform $\tilde{S}_L^{(k)}(f) = \mathscr{F}\left\{R_L^{(k)}(\tau)\right\}$ is called *cyclic spectrum* [30]. For a stationary signal only the $k=0$ component survives, thus making $R_L(t,\tau)$ a function of τ only.

Our target is relating $R_L(t,\tau)$ to the transmitted signal properties. Let $\tilde{S}_{zs}^{(k)}(f)$ the cyclic spectrum of the Fourier coefficients of $R_{zs}(t,\tau) = E\left[V(z,t+\tau)V^*(s,t)\right]$. We start by relating $\tilde{S}_L^{(k)}(f)$ to $\tilde{S}_{zs}^{(k)}(f)$. From result (22) in Appendix A, and the fact that RP1 implies additive contributions, we have:

$$\tilde{S}_L^{(k)}(f) = \int_0^L \int_0^L \tilde{h}(z,f)\,\tilde{h}^*\left(s, f-\frac{k}{T}\right)\tilde{S}_{zs}^{(k)}(f)\mathrm{dzds} \tag{9}$$

where $\tilde{h}(z,f)$ is the Fourier transform of $h(z,t)$ defined in (7). Looking at (7) we understand that for practical filters $h_{\mathrm{r}}(t)$ of bandwidth close to $R/2$, the most important cycle frequency is the zeroth one. In particular, it is exactly the only one differing from zero when $h_{\mathrm{r}}(t)$ is matched to a sinc pulse $p(0,t) = \mathrm{sinc}(t/T)$. Hence, treating $V(z,t)$ as a stationary signal is a reasonable approximation. In general, this assumption does not mean that only the stationary component of $U(0,t)$ matters, since we still have to play with a nonlinear function. Similar considerations can be made about the unperturbed signals $A_0(z,t)$ and $A_0(s,t)$ of which $V(z,t)$ and $V(s,t)$ are nonlinear functions. Thus the cyclic cross correlation $R_{zs}^0(t,\tau) = E\left[A_0(z,t+\tau)A_0^*(s,t)\right]$ has cyclic spectrum $\tilde{S}_{zs}^{0,(k)}(\omega) \triangleq \mathscr{F}\left\{R_{zs}^{0,(k)}(\tau)\right\}$ whose relation with the cyclic spectrum $\tilde{S}_{00}^{(k)}(f)$ of $U(0,t)$ is[5] :

$$\tilde{S}_{zs}^{0,(k)}(f) = \tilde{h}_{0z}(f)\tilde{h}_{0s}^*\left(f-\frac{k}{T}\right)\tilde{S}_{00}^{(k)}(f).$$

At this step, the only missing part is relating the cyclic correlation of $R_{zs}(t,\tau)$ to $R_{zs}^0(t,\tau)$, i.e., introducing the nonlinearity in the formulation. Before proceeding, it is interesting to note that in the GN model each signal is assumed to be stationary [17], thus only cycle frequency 0 exists in any cyclic correlation. Moreover, thanks to the Gaussian assumption for signal statistics, the relation between $R_{zs}(t,\tau)$ and $R_{zs}^0(t,\tau)$ takes the following simple and elegant expression under the eRP assumption [17]:

$$R_{zs}(\tau) = 2\mathcal{N}\left(R_{zs}^0(\tau)\right) = 2\left|R_{zs}^0(\tau)\right|^2 R_{zs}^0(\tau).$$

This relation does not hold for generic linearly modulated digital signals. Nevertheless, it is possible to evaluate $R_{zs}(\tau)$ proceeding as in [4]. To simplify notation we introduce the following definitions:

$$c_k \triangleq p(z,t+\tau-kT), \qquad g_k \triangleq p(s,t-kT) \tag{10}$$

which allow to significantly compact the notation even if they do not highlight that c_k, g_k are functions of time and distance. Calling $W_k \triangleq a_k c_k$ and $Y_k \triangleq a_k g_k$, the function $R_{zs}(t,\tau)$ can be written as:

$$R_{zs}(t,\tau) = \sum_{\substack{k,n,l \\ p,m,q}} E\left[W_k W_n^* W_l Y_p^* Y_m Y_q^*\right]. \tag{11}$$

This summation can be evaluated by exploiting the properties of cumulants as discussed in Appendix B. As a result, we get $R_{zs}(t,\tau)$ as a combination of periodic summations, consistently with its cyclostationary nature. As previously discussed, the most important contribution comes from the zeroth cycle frequency, as if the signals were stationary. Thus, using the results of Appendix C in (28) we have:

$$\begin{aligned} R_{zs}^{(0)}(\tau) = {} & \frac{\mu_2^3}{T^3} 2\left|Q(z,s,\tau)\right|^2 Q(z,s,\tau) + \\ & \frac{\mu_2\kappa_{2;2}}{T^2}\left\{4F_4(z,s,\tau) + Q_4(z,s,\tau)\right\}Q(z,s,\tau) + \\ & \frac{\kappa_{3;3}}{T} Q_6(z,s,\tau) \end{aligned} \tag{12}$$

[5] In particular $\tilde{S}_{00}^{(0)}(f)$ is the average PSD of $U(0,t)$.

with $\kappa_{i;i}$ the i-th cumulant (see Appendix B), $\mu_2 \equiv \kappa_{1;1}$, and:

$$\begin{aligned} Q(z,s,\tau) &\triangleq p(z,\tau) \otimes p^*(s,-\tau) \\ Q_4(z,s,\tau) &\triangleq p^2(z,\tau) \otimes (p^*(s,-\tau))^2 \\ F_4(z,s,\tau) &\triangleq |p(z,\tau)|^2 \otimes |p(s,-\tau)|^2 \\ Q_6(z,s,\tau) &\triangleq \left\{|p(z,\tau)|^2 p(z,\tau)\right\} \\ &\quad \otimes \left\{|p(s,-\tau)|^2 p^*(s,-\tau)\right\}. \end{aligned} \tag{13}$$

For all the other cycle frequencies, $R_{zs}^{(k)}(\tau)$ can be found by substituting the generic convolution $x(\tau) \otimes y(\tau)$ in (13) with $x(\tau) \otimes (y(\tau)e^{j2\pi k \frac{t}{T}})$, as suggested by (29) in Appendix C.

Equation (12) is the key ingredient for the evaluation of $R_L^{(0)}(\tau)$. For numerical purposes the algorithm is the following: 1) set a numerical grid for coordinates z and s in the integrals in (9); 2) For each $(z,\ s)$ evaluate all functions in (13) by numerical convolution. The efficiency comes from the fact that the grid in $(z,\ s)$ can be cleverly setup by adaptive quadrature routines, while the convolutions can be implemented using FFTs. The algorithm returns the full auto-correlation function; if one is just interested in the NLI variance, a fast algorithm working in the frequency domain can be found in [18].

It is worth noting that the GN model can be derived using the same theory, but with $\kappa_{2;2} = \kappa_{3;3} = 0$ [4], [15], thus leaving only the first term in (12). In this case (9) for $k = 0$ becomes $\tilde{S}_L^{(0)}(f) \triangleq 2\tilde{S}_{\rm GN}(f)$, with:

$$\begin{aligned} \tilde{S}_{\rm GN}(f) \triangleq \frac{\mu_2^3}{T^3} \int_0^L \int_0^L \tilde{h}(z,f)\,\tilde{h}^*(s,f) \\ \times \mathscr{F}\left\{|Q(z,s,\tau)|^2 Q(z,s,\tau)\right\} \mathrm{d}z\mathrm{d}s. \end{aligned} \tag{14}$$

In this special case the double integration can be efficiently reduced to a single integration, as shown in [17].

In Section II, following [14], we intuitively ascribed the main reason of the failure of the GN model to the failure of the Gaussian assumption along paths with an early presence of the nonlinear block in the diagram of Fig. 1. Carena *et al.* [14] motivated this claim by showing that the magnitude of the EGN correction terms can be made arbitrarily small at fixed distance by using a pre-compensating fiber of sufficiently large dispersion before transmission, with the same sign as the transmission fiber. This observation translates into both a significantly wrong GN prediction of the NLI variance brought by the paths with an early presence of nonlinearity and a wrong cross-correlation of such paths with all the others. It is worth noting that while one is tempted to neglect the contribution of the cross-correlations since they vanish at increasing coordinates z, s, still their double integration keeps increasing with distance L. As a result, an asymptotic offset of GN and EGN estimations of $a_{\rm NL}$ remains even in presence of pre-compensation, as first observed in [18].

As an application of our model, we now provide analytical support to these observations in the simplified case of a single channel transmission with sinc-like pulses propagating in a lossless system, a peculiar situation where the role of pre-compensation is magnified. Since the dominant correction term to the GN model is by far F_4 [18], we concentrate on it. In presence of a large pre-compensation the supporting pulse $p(z,t)$ can be asymptotically described by the method of the stationary phase [4]. This way, the power of $p(z,t)$ is of rectangular shape in the time domain. Therefore, according to (13), F_4 is a trapezoid, essentially flat over the effective duration of $Q(z,s,\tau)$. Hence, assuming $F_4(z,s,\tau) \simeq F_4(z,s,0)$, the double integral of $\tilde{S}_L^{(0)}(f)$ involving only F_4 can be closed yielding a flat PSD over the signal bandwidth of value:

$$\tilde{S}_{F_4} \sim \frac{L + z_{\rm pre} \log\left(\frac{z_{\rm pre}}{L+z_{\rm pre}}\right)}{|\beta_2|\, R^2 \pi},\ z_{\rm pre} \gg 1 \tag{15}$$

where L is the system length and $z_{\rm pre}$ is the length of the pre-compensating fiber, whose dispersion is β_2. This formula is asymptotically exact for $z_{\rm pre} \to \infty$. At fixed L, the limit of (15) for $z_{\rm pre} \to \infty$ is 0, thus the exact convergence of the EGN to the GN model, in agreement with [14]; nevertheless, at fixed $z_{\rm pre}$ the limit for $L \to \infty$ is ∞, thus the asymptotic mismatch between EGN and GN models, in agreement with [18]. Hence (15) simply explains the conundrum of the apparently contrasting conclusions in [14] and [18].

We numerically observed that in WDM transmission the qualitative conclusions are essentially the same. Equation (15) can be evaluated even in presence of lossy fibers, but with a more involved expression that uses exponential integral functions. However, the qualitative conclusions remain the same.

Fig. 5 provides further supporting evidence to the above remarks, by plotting $a_{\rm NL}$ versus number of spans $N_{\rm s}$ for a 133 channel PDM-QPSK system at 32 Gbaud with spacing 37.5 GHz over an $N_{\rm s} \times 100$ km SMF DU link. The figure shows $a_{\rm NL}$ evaluated by selectively turning on each contribution in (13), as indicated in the graph. We note that F_4 provides the largest correction to the GN model prediction. Reason is that Q_4 and Q_6, unlike F_4, are complex terms whose oscillations average out their net contribution after double integration in (9)[6] . GN and F_4 contributions show to scale with an average slope of 1.04 dB/dB and 0.88 dB/dB, respectively, in the range 10 to 100 spans, so

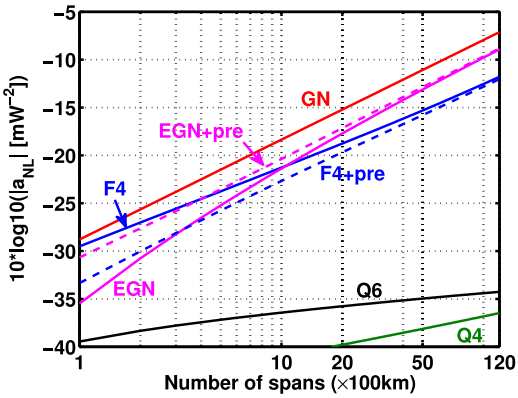

Fig. 5. Contributions to $a_{\rm NL}$ versus number of spans N_s. 133 PDM-QPSK channels ($R = 32$ Gbaud, $\Delta f = 37.5$ GHz) transmitted into a $N_{\rm s} \times 100$ km SMF DU link. F_4 and Q_4 contributions are plotted in absolute value, since they are actually negative.

[6]For this link, we verified that F_4 dominates $Q_{4,6}$ even in single channel transmission.

that the overall EGN model slope is 1.15 dB/dB. Such a number is thus larger than the value 1 predicted by the GN model for large WDM systems, and is consistent with early experimental observations [31]. Such numbers indicate that the GN part is expected to dominate F_4 at increasing distance; however, F_4 is negligible only in unrealistically long systems. Insertion of a pre-compensating fiber of 8500 ps/nm modifies the F_4 terms as shown in dashed line. As discussed previously, it relaxes the role of F_4 in short systems, but it has asymptotically no impact at long distance.

C. Extension to WDM

The generalization of (5) after z km to a WDM field is:

$$U_0(z,t) = \sum_i \sum_k a_{ki} p_i(z, t - kT) \tag{16}$$

where a_{ki} indicates the kth symbol of the channel centered at angular frequency ω_i, whose supporting pulse is p_i. In particular, in absence of third order dispersion $p_i(z,t)$ takes the expression:

$$p_i(z,t) \triangleq e^{-j\frac{\beta_2}{2}\omega_i^2 z} p(z, t - w_i(z)) e^{j\omega_i t} \tag{17}$$

where $w_i(z) \triangleq \omega_i \int_0^z \beta_2(\xi)\mathrm{d}\xi$ is the walk-off. Eq. (17), compared to the (arbitrary) reference frequency 0, experienced i) a frequency modulation, ii) a z-dependent group delay (walk-off) and iii) a z-dependent phase shift. Since (16) has the same form as (5), one can proceed as for the single channel case, by generalizing c_k and g_k to:

$$c_{ki} \triangleq a_{ki} p(z, t + \tau - w_i(z) - kT)$$
$$g_{ki} \triangleq a_{ki} p(s, t - w_i(s) - kT).$$

If FWM can be safely neglected, the dispersion-induced phase shift in (17) has no effect, so that the signal in (8) after the nonlinear block in (8) and for channel at $\omega_0 = 0$ changes into:

$$V(s,t) = \sum_{p,m,q} \Big(\underbrace{g_{p0} g^*_{m0} g_{q0}}_{\mathrm{SPM}} + \underbrace{2 \sum_{i\neq 0} g_{pi} g^*_{mi} g_{q0}}_{\mathrm{XPM}} \Big) \tag{18}$$

in which we identify intra-channel SPM and XPM-induced terms. The XPM term is independent of SPM and can be treated with the same techniques as SPM. Calculations are simpler since the general sixth order moment of XPM is:

$$E\left[c_{ki} c^*_{ni} c_{l0} g^*_{pb} g_{mb} g^*_{q0}\right] = E\left[c_{ki} c^*_{ni} g^*_{pb} g_{mb}\right] E\left[c_{l0} g^*_{q0}\right]$$

hence one has to deal with fourth and second order moments only [4]. As a result, by introducing the differential walk-off $w_{sz,i} \triangleq w_i(s) - w_i(z)$, the XPM contribution is:

$$R^{(0)}_{zs}(\tau)\Big|_{\mathrm{XPM}} = 4\frac{\mu_2^3}{T^3} \sum_{i\neq 0} |Q(z,s,\tau + w_{sz,i})|^2 Q(z,s,\tau)$$
$$+\ 4\frac{\mu_2 \kappa_{2;2}}{T^2} \sum_{i\neq 0} F_4(z,s,\tau + w_{sz,i}) Q(z,s,\tau). \tag{19}$$

D. Extension to Dual Polarization

In single polarization the nonlinear operator is $\mathcal{N}(A) = |A|^2 A$. Using the Manakov equation in dual polarization with electric field $\vec{A} = [A_x;\, A_y]$, such an operator generalizes to $\mathcal{N}(\vec{A}) = \frac{8}{9}\vec{A}^\dagger \vec{A} \vec{A}$, where $\dagger$ indicates transpose conjugate. In WDM, accounting just for SPM and XPM, with $\vec{A}_i = [A_{ix};\, A_{iy}]\, e^{j\omega_i t}$ the nonlinear operator acting on channel 0 is [24]:

$$\mathcal{N}_0 = \frac{8}{9}\left\{ ||\vec{A}_0||^2 + \sum_{i\neq 0} \left(\vec{A}_i \vec{A}_i^\dagger + \vec{A}_i^\dagger \vec{A}_i \right) \right\} \vec{A}_0.$$

In particular, the nonlinearity acting on polarization x is:

$$\mathcal{N}_{0x} = \underbrace{|A_{0x}|^2 A_{0x}}_{①} + \underbrace{|A_{0y}|^2 A_{0x}}_{②}$$
$$+ \sum_{i\neq 0} \Big[\underbrace{2|A_{ix}|^2 A_{0x}}_{③} + \underbrace{|A_{iy}|^2 A_{0x}}_{④} + \underbrace{A_{ix} A^*_{iy} A_{0y}}_{⑤} \Big].$$

Let us discuss each term separately:

① Phase modulation (PM) term identical to SPM discussed in Section IV.

② PM term, manifesting as XPM of y over x polarization. Hence, it is XPM with zero walk-off and no degeneracy factor 2. It contributes to the overall correlation like (19), with a factor $1/4$ to remove the missing degeneracy.

③ PM term identical to XPM discussed in Section IV-C.

④ PM term but without degeneracy factor 2. It contributes to the overall correlation like (19) with a factor $1/4$ to remove the missing degeneracy.

⑤ XPolM term containing the mixing of three independent fields. This way, it contributes to the correlation like (19), but just as if $\kappa_{2;2} = 0$.

Please note that the overall effect created by ③ and ④ induces both a PM common to both polarizations, hence XPM, as well a differential PM, hence XPolM. Once averaged over a uniformly distributed state of polarization of channel i, the XPM is weighted by a factor 3/2 [25].

Similarly to (14), by introducing the definitions:

$$\tilde{S}_{\mathrm{XGN,i}}(f) = \frac{\mu_2^3}{T^3} \int_0^L \int_0^L \tilde{h}(z,f)\, \tilde{h}^*(s,f) \times \mathscr{F}\left\{ |Q(z,s,\tau + w_{sz,i})|^2 Q(z,s,\tau) \right\} \mathrm{dzds}$$

$$\tilde{S}_{\mathrm{F4,i}}(f) = \frac{\kappa_{2;2}\mu_2}{T^2} \int_0^L \int_0^L \tilde{h}(z,f)\, \tilde{h}^*(s,f) \times \mathscr{F}\left\{ F_4(z,s,\tau + w_{sz,i}) Q(z,s,\tau) \right\} \mathrm{dzds}$$

$$\tilde{S}_{\mathrm{Q4}}(f) = \frac{\kappa_{2;2}\mu_2}{T^2} \int_0^L \int_0^L \tilde{h}(z,f)\, \tilde{h}^*(s,f) \times \mathscr{F}\left\{ Q_4(z,s,\tau) Q(z,s,\tau) \right\} \mathrm{dzds}$$

$$\tilde{S}_{\mathrm{Q6}}(f) = \frac{\kappa_{3;3}}{T} \int_0^L \int_0^L \tilde{h}(z,f)\, \tilde{h}^*(s,f) \times \mathscr{F}\left\{ Q_6(z,s,\tau) \right\} \mathrm{dzds}$$

we can discriminate between single/dual polarization and single/multi channel according to Table I. For numerical purposes, note that the summations in i can be easily closed, thus generating phased array factors [7].

TABLE I
PSD OF NLI

Case	NLI PSD
Single-polarization	$\tilde{S}_x(f) = \left(\frac{8}{9}\right)^2 \left\{ 2\tilde{S}_{\mathrm{SGN}}(f) + 4\sum_{i\neq 0} \tilde{S}_{\mathrm{XGN}}(f) + 4\sum_i \tilde{S}_{\mathrm{F4,i}}(f) + \tilde{S}_{\mathrm{Q4}}(f) + \tilde{S}_{\mathrm{Q6}}(f) \right\}$
Dual-polarization	$\tilde{S}_{x+y}(f) = 2\left(\frac{1}{2}\right)^3 \left(\frac{8}{9}\right)^2 \left\{ 3\tilde{S}_{SGN}(f) + 6\sum_{i\neq 0} \tilde{S}_{\mathrm{XGN}}(f) + 5\sum_i \tilde{S}_{\mathrm{F4,i}}(f) + \tilde{S}_{\mathrm{Q4}}(f) + \tilde{S}_{\mathrm{Q6}}(f) \right\}$
	$= \frac{16}{27}\left(\tilde{S}_{\mathrm{SGN}}(f) + 2\sum_{i\neq 0} \tilde{S}_{\mathrm{XGN}}(f) \right) + \frac{80}{81}\sum_i \tilde{S}_{\mathrm{F4,i}}(f) + \frac{16}{81}\tilde{S}_{\mathrm{Q4}}(f) + \frac{16}{81}\tilde{S}_{\mathrm{Q6}}(f)$

Regarding the dual-polarization case we note that: factor $\left(\frac{1}{2}\right)^3$ comes after substituting symbols a_k with $a_k/\sqrt{2}$ so that power P is the total power (*x*+*y*) in any case; the first factor 2 comes from the assumption $\tilde{S}_{x+y}(f) = \tilde{S}_x(f) + \tilde{S}_y(f) = 2\tilde{S}_x(f)$; the factors $3, 5, 6$ come from ②, ④and ⑤, respectively. The $\sum_i \tilde{S}_{\mathrm{F4,i}}(f)$ is over all channels, thus even the reference one at $\omega_i = 0$.

E. Extension to ASE Noise

In [19] it was shown that nonlinear signal/ASE interaction along DU links plays a role at large BERs. In this section we show how to include ASE noise into our EGN model in the special case of Nyquist pulses, i.e., $p(0,t) = \mathrm{sinc}\left(\frac{t}{T}\right)$. Let $w(t)$ be white GN of one-sided PSD $N_{0\mathrm{A}}$ introduced by an EDFA, it can be written as:

$$w(t) = \sum_\ell w^{(\ell)}(t) e^{j2\pi\ell\frac{t}{T}}$$

with $w^{(\ell)}(t)$ independent[7], zero-mean, Gaussian stationary processes with rectangular PSD over the range $-\frac{R}{2} + \ell R < f < \frac{R}{2} + \ell R$:

$$\begin{aligned}\tilde{S}_\ell(f) &= N_{0\mathrm{A}} \prod\left(\frac{f-\ell R}{R}\right) \overset{\mathscr{F}^{-1}}{\longleftrightarrow} R_\ell(\tau) \\ &= N_{0\mathrm{A}} R \cdot p(t) e^{j2\pi\ell\frac{t}{T}}. \end{aligned} \tag{20}$$

Each $w^{(\ell)}(t)$ is a stationary band-limited Gaussian process, hence it can be written as [32, p. 599]:

$$w^{(\ell)}(t) = \sum_k w^{(\ell)}(kT) p(t-kT) \tag{21}$$

where equality is in the mean square sense. From (20) we infer that random variables (RV) $w^{(\ell)}(kT)$ and $w^{(\ell)}(nT)$ are independent if $k \neq n$. We are thus expressing ASE as a linear digital modulated signal with Gaussian distributed symbols. This way, for an optical link with EDFAs spaced by z_{A} km, the unperturbed solution after z km is a generalization of (16)[8] :

$$A_0(z,t) = \sum_\ell \sum_k \left[a_k^{(\ell)} + \sum_{m=1}^{\left\lfloor \frac{z}{z_{\mathrm{A}}} \right\rfloor} w_m^{(\ell)}(kT) \right] p_\ell(z, t-kT)$$

where subscript m indicates the amplifier number while p_ℓ was introduced in (17) and here has to be used with a sinc pulse. This expression is formally similar to (16) but with noisy symbols. Note that $a_k^{(\ell)}$ may be zero if channel ℓ is off, hence the summation over ℓ is confined to the EDFA bandwidth.

We can thus apply the same theory as in the previous section, with different statistics for the symbols. Since cumulants of independent RVs add up, it turns out that cumulants $\kappa_{3;3}$ and $\kappa_{2;2}$ are not affected by ASE noise. ASE remains just in the GN-model contribution to the overall NLI, i.e., cumulant $\kappa_{1;1}$, which becomes distance dependent and changes into:

$$\kappa_{1;1} = \mu_2 + N_{0\mathrm{A}} R \cdot \left\lfloor \frac{\min(z,s)}{z_{\mathrm{A}}} \right\rfloor = \mu_2 \left(1 + \frac{N_{\mathrm{s}} \left\lfloor \frac{\min(z,s)}{z_{\mathrm{A}}} \right\rfloor}{\mathcal{E}_s/N_0} \right)$$

where N_{s} corresponds to the number of EDFA, while $\mathcal{E}_s/N_0 \triangleq \mu_2/(N_{0\mathrm{A}} N_{\mathrm{s}})$ is the received SNR over the bandwidth R in absence of nonlinear effects. For instance, for a QPSK modulation using matched filtered detection, $\mathcal{E}_s/N_0$ coincides with the Q-factor.

The algorithm discussed at the end of Section IV-B can thus be applied as well in presence of ASE with the new $\kappa_{1;1}$ cumulant at each coordinate z and s. As a preliminary test, we investigated the change in a_{NL} due to the presence of ASE noise in a single channel PDM-QPSK transmission at $P = 0$ dBm. Results are reported in Fig. 6 for $R = 32$ Gbaud and a 10×100 km DU SMF link. ASE noise was loaded at each EDFA along the line over a bandwidth of 637.5 GHz to speed up SSF simulations. The same value was used even in the EGN model. We note a good agreement between SSF simulations and EGN theory at practical $\mathcal{E}_s/N_0 > 2$ dB. Note that even the GN model (dashed line) predicts an enhancement of NLI variance at small SNRs.

V. DISCUSSION ON THE HYPOTHESES

Investigating the range of validity of the EGN assumptions is a big challenge due to the large amount of parameters affecting performance. Here we just provide some general indications.

A. RP1 Approximation

It is a good approximation for a large class of optical links. We verified in [33] for different fiber dispersions that in absence of ASE noise RP1 overestimates the true SNR by less than 1

[7] $w(t)$ is stationary if and only if $w^{(\ell)}$ are uncorrelated, otherwise it is cyclostationary [30].

[8] Note that adding an arbitrary phase shift and/or a time delay to (21) does not change its statistics.

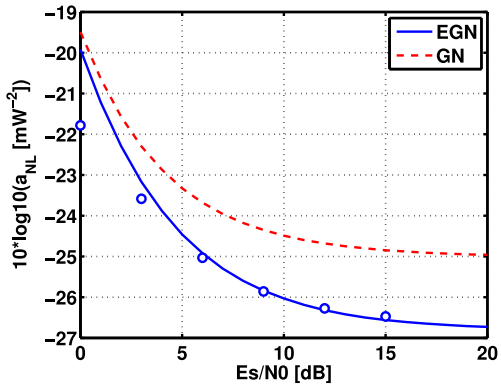

Fig. 6. a_{NL} versus received $\mathcal{E}_s/N_0$ evaluated using EGN, GN and SSF simulations (symbols). Single channel PDM-QPSK transmission, noisy EDFA along the line with ASE filling a bandwidth of 637.5 GHz.

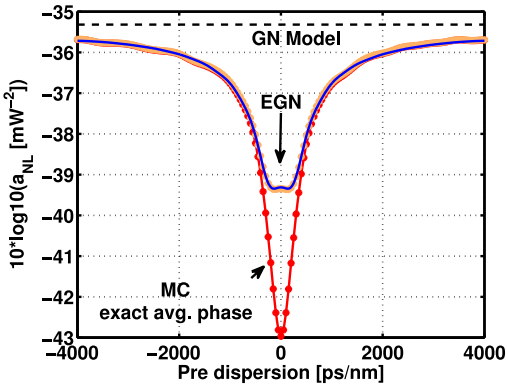

Fig. 7. a_{NL} in a single dispersion-less span preceded by a linear pre-compensating fiber and fully compensated at the receiver (corresponds to a generic branch in Fig. 1). Single channel PDM-QPSK. While the GN curve is independent of the value of the pre-dispersion, the EGN does depend on it. However, EGN fails at small values of pre-dispersion because it is based on the asymptotic value of the average nonlinear phase in presence of strong dispersion.

TABLE II
$a_{\mathrm{NL}}^{\mathrm{MC}} - a_{\mathrm{NL}}^{\mathrm{EGN}}$ [dB]

	$\rho = 0$	$\rho = 0.4$	$\rho = 0.7$
PDM-QPSK	<0.1	0.14	0.5
PDM-16QAM	<0.1	0.11	0.38

Optical link as in Fig. 2.

dB at SNR larger than 4 dB.[9] In presence of ASE this gap is even smaller, particularly around the best power at which the contribution of ASE noise is twice the one of NLI. We also tested, even if not reported here, that logarithmic perturbation [8] is never more accurate than RP in the DU and DM systems analyzed in this paper, however with a Q-factor discrepancy smaller than 1 dB.

B. Cyclostationarity

Assumption i) in Section IV of a linearly modulated digital signal implies cyclostationarity for the received NLI. This way, we have to deal with cyclic spectra whose zeroth cyclic frequency is the well known stationary component. As we noted after (12), higher order cyclic frequencies are proportional to convolution of band-limited signals shifted in frequency by multiples of R, hence they are expected to be of small impact. To give a feeling, assume a system working at $R = 28$ Gbaud and spacing 50 GHz. For a roll-off ρ the channel bandwidth is $R(1+\rho)$, hence ρ can be varied between 0 and 0.78 to avoid spectrum overlap. For the 5-span optical link considered in Fig. 2 we estimated the errors reported in Table II between MC and the EGN model of Table I.

For practical roll-off $\rho < 0.5$ the error is negligible.

C. Linearly Modulated Digital Signal

In assumption i) in Section IV we assumed independent symbols with at least fourfold rotational symmetry. Notable violations of this assumption are BPSK [34], where the rotational symmetry is twofold, and polarization-switched QPSK and all its set-partitioning variants, for which symbols in *X* and *Y* are dependent. Further extending the EGN to account for such modulation formats is possible even if quite complex. However, preliminary simulations in a 5 span DU link indicate that the deviations in a_{NL} from the prediction given by the EGN model as described in this paper are confined to 0.4 dB.

D. Average Nonlinear phase

To simplify the model, the main EGN result summarized in Table I was derived using a CPE that just recovered the large dispersion Kerr induced nonlinear phase (see Corollary 12) due to a constant signal of power equal to $2P$ in single polarization and $\frac{3}{2}P$ in dual polarization, respectively, as discussed in [17]. An intuitive explanation of the factor two in single polarization is related to intra-channel and inter-channel XPM, which both dominate the contribution of SPM when a large dispersion is present. In [17] we showed that many practical DU links belong to this category. However, when dispersion is small, which can be the case especially with DM systems, the true average nonlinear phase is smaller. In principle this effect can be accounted for by including the terms removed in Corollary 12. However, a small average nonlinear phase is an indication of a small dispersive regime, where the non-circularity of the NLI distribution is the main issue.

To get a feeling of the error introduced by the use of a wrong average nonlinear phase in our EGN, we simulated the simple single-span system preceded by linear pre-compensation that was recently adopted to analyze special network scenarios [26], [35]. We investigated the extreme case of transmission fiber with zero dispersion, thus corresponding to an eRP of just one linear block followed by a nonlinear block. This amounts to investigating just one path of the RP diagram in Fig. 1. Fig 7 shows a_{NL} for the single channel case when varying the amount of pre-dispersion. a_{NL} was estimated by MC simulations in a reference system rotated by either the average nonlinear phase or the theoretical one expected by the model according to Corollary 12. Keeping the power small enough, -4 dBm in this case,

[9] Recall that for an optimal memoryless detection, SNR and Q-factor coincide for PDM-QPSK in absence of differential decoding.

and measuring $a_{\rm NL}$ as in Section III following the eRP definition, despite the absence of dispersion the EGN model should work correctly since all the other hypotheses of Section IV are satisfied. Fig. 7 indicates that this is the case only by using the (false) theoretical average phase. Using the true estimated average phase rotation induces a significant error for pre-dispersion smaller than $\sim$400 ps/nm in absolute value. At larger values of pre-dispersions, indeed the ISI is large enough to let the average nonlinear phase rotation approach the theoretical large-dispersion value. Asymptotically, the dispersion is so strong that even the Gaussian assumption for the input field holds.

Even if not reported here, we numerically verified that the mismatch between the reference average nonlinear phase used by the EGN and the true nonlinear phase is the main reason of the small gap between EGN and SSF results in Fig. 3.

E. FWM

In our EGN derivation we neglected FWM. To get a feeling of the impact of FWM we refer to the system of Fig. 2, thus close to a Nyquist WDM. By removing FWM in the MC we got an error on $a_{\rm NL}$ of 0.3 dB, an indication that with $R = 32$ Gbaud and channel spacing 37.5 GHz FWM in a DU system can be safely neglected. In our SMF-based system FWM starts to play a role at 10 Gbaud at bandwidth efficiency 0.85, where we separately verified that it is almost the entire cause of the small gap between SSF and EGN nonlinear asymptotes of the QPSK-DU curves in Fig. 4. This is not the case of the DM30 curve, where, by using the nonlinearity decoupling method used in [27], we estimated that FWM is responsible of about 30% of the gap between EGN and SSF asymptotes. We ascribe the remaining part of the gap to the breakdown of the noise circularity. FWM does play a role either at small dispersions or at the Nyquist WDM, as shown in [9].

F. Independence on Symbol of Interest

Probably this is the weakest point of all EGN models for weakly dispersive systems. In assumption iii) in Section IV we assumed that the NLI variance conditioned on the transmitted symbol of interest (i.e., the actual variance we need for BER computation) is equal to the unconditional variance, i.e., the one computed by the EGN model by averaging on all symbols. The two variances coincide when the NLI field is independent of the symbol of interest. While EGN averaging over all symbols is straightforward since the process that gets averaged is cyclostationary, when we condition on a specific symbol of interest the received NLI is not cyclostationary anymore and computations get much more involved despite the fact that the conditioned symbol has zero cumulant.

This problem may be significant in the weakly dispersive regime where the reduced system memory makes the contribution of the symbol of interest relevant to the total NLI variance. The impact of such a problem was analyzed in [36], where it was show that even with Gaussian distributed symbols, the circularly distributed unconditional NLI distribution becomes non-circular after conditioning.

An implication of assumption iii) is that it cannot account for nonlinear phase noise. A simple proof is by looking at the pseudo-correlation function $E\left[V(z,t+\tau)V(s,t)\right]$, which is always zero under assumption iii). This way, $V(z,t)$ is a proper [13] stochastic process, implying that the real and imaginary components of the NLI have the same variance, contrary to the in-phase/quadrature unbalance naturally produced by nonlinear phase noise [26]. Investigations of the phase noise nature of the NLI appear in [15], [26], [37]. An attempt to include nonlinear phase noise in modeling of coherent systems was proposed in [38].

VI. Conclusion

From a first-order perturbative propagation model, we derived a time-domain extended GN model yielding the time auto-correlation function of the received NLI with linearly-modulated input digital signals. We assumed an additive NLI to the useful signal, obtaining excellent reach estimations both in SMF-based DU links at all considered symbol rates, and in DM links at rates above 10 Gbaud. Compared with the EGN in [9] our model neglects FWM and caution should then be used when dealing with non-zero-dispersion-shifted DU links, especially at low frequency separation, a case, however, where also the assumed circularity of NLI becomes questionable, especially in a network scenario [26]. On the positive side, our EGN is more easily extended to include distributed ASE. Also, double-integral computations in the time domain may more naturally be carried out due to similarity with the FFT based SSF approach. Finally, our time-domain derivation based on cyclostationary properties of the input signals leads to a cross-validation of the SPM and cross-channel NLI theoretical formulas in [9] and to an interpretation of their key coefficients in terms of signal cumulants. Compared with the EGN in [15], our EGN does provide not only the NLI variance, but the whole NLI auto-correlation, i.e., even its spectral properties. Also, our eq. (15) was able to simply explain the apparent contrasting conclusions of the studies in [14] and [18] regarding the convergence/misconvergence of the EGN to the GN model. More work is needed to fully understand the implications of the key model assumption of independence of NLI and received field at the symbol of interest, in order to fully extend the method to DM links at lower symbol rates.

Appendix A
Linear Filtering of Cyclostationary Signals

In this appendix we report a general property of cyclostationary signals passing through linear time invariant (LTI) systems [30]. Let:

$$y_1(t) = h_1(t) \otimes x_1(t)$$
$$y_2(t) = h_2(t) \otimes x_2(t).$$

Assume $x_{1,2}(t)$ are jointly cyclostationary processes of period T such that:

$$R_{x_1x_2}(t,\tau) = E\left[x_1(t+\tau)x_2^*(t)\right] = \sum_n R^{(n)}_{x_1x_2}(\tau)e^{j2\pi n\frac{t}{T}}.$$

Also $y_{1,2}$ are jointly cyclostationary, thus $R_{y_1y_2}(t,\tau)$ can be expanded in Fourier series as well with cyclic correlation $R^{(n)}_{y_1y_2}(\tau)$

at cycle frequency n/T equal to [30]:

$$R^{(n)}_{y_1y_2}(\tau) = \left(h_2^*(-\tau)e^{j2\pi n\frac{\tau}{T}}\right) \otimes h_1(\tau) \otimes R^{(n)}_{x_1x_2}(\tau).$$

In Fourier domain, with $\tilde{S}^{(n)}_{y_1y_2}(f) \triangleq \mathscr{F}\left\{R^{(n)}_{y_1y_2}(\tau)\right\}$ and $\tilde{H}_{1,2}(f) \triangleq \mathscr{F}\{h_{1,2}(\tau)\}$:

$$\tilde{S}^{(n)}_{y_1y_2}(f) = \tilde{H}_1(f)\tilde{H}_2^*\left(f - \frac{n}{T}\right)\tilde{S}^{(n)}_{x_1x_2}(f). \tag{22}$$

For a stationary signal only the $n = 0$ component exists, yielding the well known input-output relation of LTI systems.

Appendix B
Cumulants and Master Theorem

In this appendix we review the main properties of cumulants and derive the master equation (12). More information can be found in [39] and at Wikipedia [40].

Definition 1: The joint cumulant generating function (CGF) of the RVs $X_1, \ldots, X_n$ is defined as:

$$K(s_1, \ldots, s_n) = \log M(s_1, \ldots, s_n) \tag{23}$$

where $M(s_1, \ldots, s_n)$ is the moment generating function (MGF):

$$M(s_1, \ldots, s_n) = E\left[e^{s_1X_1 + \ldots + s_nX_n}\right].$$

□

Definition 2: The joint cumulant of the RVs $X_1, \ldots, X_n$ is defined as:

$$\kappa(X_1, \ldots, X_n) = \left.\frac{\partial^n K}{\partial s_1 \cdots \partial s_n}\right|_{s_1 = \ldots = s_n = 0}. \tag{24}$$

The cumulant of a complex RV Z is indeed defined as [41]:

$$\kappa_{p,q}(Z) = \kappa(\underbrace{Z, \ldots, Z}_{p}, \underbrace{Z^*, \ldots, Z^*}_{q}).$$

□

Cumulants have several interesting properties. For instance they are multilinear:

Proposition 3: For any list of constants $c_1, \ldots, c_n$, cumulants satisfy the property [40]:

$$\kappa(X_1 + Z, \ldots, X_n) = \kappa(X_1, \ldots, X_n) + \kappa(Z, \ldots, X_n)$$
$$\kappa(c_1X_1, \ldots, c_nX_n) = c_1 \cdots c_n \kappa(X_1, \ldots, X_n).$$

□

Unlike moments, cumulants of independent RVs are zero:

Theorem 4: Given a list $X_1, \ldots, X_n$ of RVs, if at least one X_k is mutually independent of the remaining RVs in the list, the joint cumulant is zero.

Proof: Suppose that X_1 is independent of $X_2, \ldots, X_n$. Its MGF can be thus factored out of the overall MGF. The CGF is thus:

$$\begin{aligned} K(s_1, \ldots, s_n) &= \log\left[M(s_1)M(s_2, \ldots, s_n)\right] \\ &= \log M(s_1) + \log M(s_2, \ldots, s_n) \end{aligned}$$

and thus (24) is zero. ■

The most important property of cumulants is their combinatoric meaning, according to this lemma [39], [40]:

Lemma 5: By expanding in Taylor series both sides of (23), using the Faá di Bruno's formula we get:

$$\kappa(X_1, \ldots, X_n) = \sum_{\pi} (|\pi| - 1)!(-1)^{|\pi|-1} \prod_{B \in \pi} E\left[\prod_{i \in B} X_i\right] \tag{25}$$

$$E[X_1 \cdots X_n] = \sum_{\pi} \prod_{B \in \pi} \kappa(X_i : i \in B) \tag{26}$$

where π runs through the list of all partitions of $\{1, \ldots, n\}$, B runs through the list of all blocks in the running partition, and $|\pi|$ is the number of blocks in the running partition.

Example 6: If $n = 3$, the set π contains five partitions: $\{1, 2, 3\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{2, 3\}, \{1\}\}$, $\{\{1\}, \{2\}, \{3\}\}$. For instance, the second partition $\{\{1, 2\}, 3\}$ contains two blocks ($|\pi| = 2$), i.e., $\{1, 2\}$ and $\{3\}$. We thus have:

$$\begin{aligned} \kappa(X_1, X_2, X_3) &= E[X_1X_2X_3] - E[X_1X_2]E[X_3] \\ &\quad - E[X_1X_3]E[X_2] - E[X_2X_3]E[X_1] \\ &\quad + 2E[X_1]E[X_2]E[X_3] \\ E[X_1X_2X_3] &= \kappa(X_1, X_2, X_3) + \kappa(X_1X_2)\kappa(X_3) \\ &\quad + \kappa(X_1X_3)\kappa(X_2) + \kappa(X_2X_3)\kappa(X_1) \\ &\quad + \kappa(X_1)\kappa(X_2)\kappa(X_3). \end{aligned}$$

□

In our case we are mainly interested in searching for cumulants of RVs having rotational symmetry, according to the following definition:

Definition 7: A complex RV Z has n-fold rotational symmetry if the distribution of Z is identical to the distribution of $Ze^{j\frac{2\pi}{n}m}$ for any $m \in \mathbb{Z}$. □

The direct consequence of such a definition is the following property:

Theorem 8: Let Z an n-fold rotational symmetric complex RV. Hence, generalizing [13, p. 499]:

$$E[Z^p] = 0, \qquad p \neq kn,\ k \in \mathbb{Z}.$$

Proof: It is:

$$E[Z^p] = e^{-j\frac{2\pi}{n}kp}E\left[\left(Ze^{j\frac{2\pi}{n}k}\right)^p\right] = e^{-j\frac{2\pi}{n}kp}E[Z^p] \tag{27}$$

where the first equality follows by rewriting Z^p as $e^{-j\frac{2\pi}{n}kp}(Ze^{j\frac{2\pi}{n}k})^p$, and where the second equality follows because the n-fold rotational symmetry of Z implies that Z and $Ze^{-j\frac{2\pi}{n}k}$ have the same law $\forall k \in \mathbb{Z}$. If p is not a multiple of n, the only way to satisfy (27) is by $E[Z^p] = 0$. ■

An interesting implication in our context is given by the following theorem:

Theorem 9: Cumulant $\kappa_{p,q}(Z)$ of a complex RV Z with n-fold rotational symmetry is zero if $p - q \neq kn$ with $k \in \mathbb{Z}$.

Proof. According to Theorem 8, if the cumulant were non zero, there would be at least one partition in (25) whose generic block $Z^m\,(Z^*)^l$ would have $m - l = kn$, k being an integer.

Since this identity must be true for all the blocks forming the partition, even $p-q$ would be a multiple of n, in contradiction with the hypotheses. ■

Example 10: Using (25), the first three complex cumulants of a zero mean complex RV having $\mu_n = E\left[|Z|^n\right]$ are [42]:

$$\begin{aligned}\kappa_{1;1} &= \kappa(Z) = \mu_2^2\\ \kappa_{2;2} &= \kappa(Z,Z,Z^*,Z^*) = \mu_4 - 2\mu_2^2\\ \kappa_{3;3} &= \kappa(Z,Z,Z,Z^*,Z^*,Z^*) = \mu_6 - 9\mu_4\mu_2 + 12\mu_2^3.\end{aligned}$$

In particular, square M-QAM modulation with i.i.d. symbols has:

$$\kappa_{2;2} = -\frac{3}{5}\frac{M+1}{M-1}\mu_2^2, \qquad \kappa_{3;3} = \frac{36}{21}\frac{M^3-1}{(M-1)^3}\mu_2^2$$

while $\kappa_{2;2} = \kappa_{3;3} = 0$ for Gaussian distributed symbols. □

We are now in position to collect all together the previous statements within the following master theorem:

Theorem 11: Assume $a_k,\ k \in \mathbb{Z}$, complex i.i.d. RVs with n-fold rotational symmetry and $n \geq 4$. Let $W_k = c_k a_k$ and $Y_k = g_k a_k$ with $c_k,\ g_k$ constants. Then:

$$\begin{aligned}\sum_{\substack{k,n,l\\p,m,q}} E\left[W_k W_n^* W_l Y_p^* Y_m Y_q^*\right] &= \kappa_{3;3}\sum_k |c_k|^2 c_k |g_k|^2 g_k^*\\ &+ \kappa_{2;2}\kappa_{1;1}\left(2\sum_{k,n}|c_k|^2 c_k g_k^* |g_n|^2 + 2\sum_{k,n}|g_k|^2 g_k^* c_k |c_n|^2\right)\\ &+ \kappa_{2;2}\kappa_{1;1}\left(4\sum_{k,n}|c_k|^2|g_k|^2 c_n g_n^* + \sum_{k,n} c_k^2 g_k^{*2} c_n^* g_n\right)\\ &+ \kappa_{1;1}^3\left(4\sum_{k,n,l}|c_k|^2|g_n|^2 c_l g_l^* + 2\sum_{k,n,l} c_k g_k^* c_n g_n^* c_l g_l^*\right).\end{aligned} \tag{28}$$

Proof. From (26) each average value can be expressed as a sum of cumulants of the RV creating the partitions. Using Theorem 4, in each block of each partition the indexes must be identical. Thus, using Prop. 3 the set partitioning of (26) contains cumulants $\kappa_{p,q}$ with $|p-q| \leq 3$. Theorem 9 ensures that such cumulants are zero if $p \neq q$. Hence, only cumulants $\kappa_{1;1},\ \kappa_{2;2},\ \kappa_{3;3}$ participate to (26), thus completing the proof. ■

Equation (28) is rather complex. A first simplification is provided by the eRP assumption. Hence, if the value of P introduced in Section IV-A is twice the average power, i.e., $P = 2\kappa_{1;1}\sum_n |c_n|^2$, the second line of (11) is zero, according to the following Corollary:

Corollary 12: If $\sum_n |c_n|^2 = \sum_n |g_n|^2$, with $\mathcal{E} \triangleq 2\kappa_{1;1}\sum_n |c_n|^2$ we have (eRP assumption):

$$\begin{aligned}&\sum_{\substack{k,n,l\\p,m,q}} E\left[\left(W_k W_n^* - \mathcal{E}\right) W_l \left(Y_p^* Y_m - \mathcal{E}\right) Y_q^*\right]\\ &= \kappa_{3;3}\sum_k |c_k|^2 c_k |g_k|^2 g_k^* + 2\kappa_{1;1}^3\sum_{k,n,l} c_k g_k^* c_n g_n^* c_l g_l^*\\ &+ \kappa_{2;2}\kappa_{1;1}\left(4\sum_{k,n}|c_k|^2|g_k|^2 c_n g_n^* + \sum_{k,n} c_k^2 g_k^{*2} c_n^* g_n\right)\\ &+ \kappa_{1;1}^3\left(4\sum_{k,n,l}|c_k|^2|g_n|^2 c_l g_l^* + 2\sum_{k,n,l} c_k g_k^* c_n g_n^* c_l g_l^*\right).\end{aligned}$$

Proof. Obtained by straightforward application of the steps in Theorem 11, and by using the identities:

$$\begin{aligned}\sum_{k,n,l,q} E\left[W_k W_n^* W_l Y_q^*\right] &= \sum_k c_k g_k^*\left(\kappa_{2;2}|c_k|^2 + 2\kappa_{1;1}^2\sum_n |c_n|^2\right)\\ \sum_{p,m,l,q} E\left[Y_p Y_m^* Y_q^* W_l\right] &= \sum_k g_k^* c_k\left(\kappa_{2;2}|g_k|^2 + 2\kappa_{1;1}^2\sum_n |g_n|^2\right).\end{aligned}$$

■

All summations in (28) are periodic summations that can be expressed in terms of convolutions by using the results of Appendix C.

Appendix C
Periodic Summations

Consider the periodic summation:

$$s(t,\tau) = \sum_k f(t+\tau-kT)g(t-kT).$$

This function is periodic in t of period T, hence it can be expanded in Fourier series:

$$s(t,\tau) = \sum_\ell s^{(\ell)}(\tau)e^{j2\pi\ell\frac{t}{T}}$$

with:

$$\begin{aligned}s^{(\ell)}(\tau) &= \frac{1}{T}\int_{-T/2}^{T/2}\sum_k f(t+\tau-kT)g(t-kT)e^{-j2\pi\ell\frac{t}{T}}\,\mathrm{dt}\\ &= \frac{1}{T}\int_{-\infty}^{\infty} f(t_1+\tau)g(t_1)e^{-j2\pi\ell\frac{t_1}{T}}\,\mathrm{dt_1}\\ &= \frac{1}{T}f(\tau)\otimes\left(g(-\tau)e^{j2\pi\ell\frac{\tau}{T}}\right).\end{aligned}$$

Hence:

$$s(t,\tau) = \frac{1}{T}\sum_\ell\left[f(\tau)\otimes\left(g(-\tau)e^{j2\pi\ell\frac{\tau}{T}}\right)\right]e^{j2\pi\ell\frac{t}{T}}. \tag{29}$$

In frequency:

$$\mathscr{F}_\tau\left\{s(t,\tau)\right\} = \frac{1}{T}\sum_\ell \tilde{F}(f)\tilde{G}\left(\frac{\ell}{T}-f\right)e^{j2\pi\ell\frac{t}{T}}.$$

ACKNOWLEDGMENT

The authors wish to thank J.-C Antona for his continuous support and invaluable advice, P. Ramantanis and N. Rossi for stimulating discussions and A. Mecozzi for focusing their attention on the behavior in ultralong links.

REFERENCES

[1] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of non-linear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[2] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Exp.*, vol. 18, pp. 19039–19054, Aug. 2010.

[3] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[4] A. Mecozzi and R. J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[5] A. Bononi, P. Serena, N. Rossi, E. Grellier, and F. Vacondio, "Modeling nonlinearity in coherent transmissions with dominant intrachannel-four-wave-mixing," *Opt. Exp.*, vol. 20, no. 7, pp. 7777–7791, Mar. 2012.

[6] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, H. Wymeersch, and P. Andrekson, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[7] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, Pak Cho, R. Noe, I. Shpantzer, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Exp.*, vol. 16, no. 20, pp. 15778–15810, Sep. 2008.

[8] M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[9] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "The EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362, Jun. 2014.

[10] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. Tapia Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," presented at the European Conf. Optical Communication, Torino, Italy, 2010, Paper P4.07.

[11] G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, "Performance prediction for WDM PM-QPSK transmission over uncompensated links," presented at the Optical Fiber Communication Conf., Los Angeles, CA, USA, 2011, Paper OThO7.

[12] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with gaussian-distributed nonlinear noise," *Opt. Exp.*, vol. 19, no. 13, pp. 12781–12788, Jun. 2011.

[13] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[14] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, "Impact of the transmitted signal initial dispersion transient on the accuracy of the GN-model of non-linear propagation", presented at the European Conf. Optical Communication, London, U.K., 2013, Paper Th.1.D.4.

[15] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699, Oct. 2013.

[16] P. Serena, A. Bononi, and N. Rossi, "The impact of the modulation dependent nonlinear interference missed by the gaussian noise model," presented at the European Conf. Optical Communication, Cannes, France, 2014, Paper Mo.4.3.1.

[17] 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.

[18] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in multi-span fiber-optic systems," *Opt. Exp.*, vol. 22, no. 12, pp. 14199–14211, Jun. 2014.

[19] P. Poggiolini, A. Carena, and Y. Jiang, G. Bosco, V. Curri, and F. Forghieri, "Impact of low-OSNR operation on the performance of advanced coherent optical transmission systems," presented at the European Conf. Optical Communication, Cannes, France, 2014, Paper Mo.4.3.2.

[20] D. Zwillinger, *Handbook of Differential Equations*, 3rd ed. New York, NY, USA: Academic, 1997.

[21] A. Mecozzi, "A unified theory of intrachannel nonlinearity in pseudolinear transmission," in *Impact of Nonlinearities on Fiber Optic Communication* , S. Kumar, Ed. New York, NY, USA: Springer-Verlag, 2011, pp. 253–291.

[22] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 31, no. 8, pp. 1273–1282, Apr. 2013.

[23] A. Vannucci, P. Serena, and A. Bononi, "The RP method: A new tool for the iterative solution of the nonlinear Schroedinger equation," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1102–1112, Jul. 2002.

[24] C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *J. Lightw. Technol.*, vol. 24, no. 7, p. 2806–2826, Jul. 2006.

[25] M. Winter, C. A. Bunge, D. Setti, and K. Petermann, "A statistical treatment of cross-polarization modulation in DWDM systems," *J. Lightw. Technol.*, vol. 27, no. 17, pp. 3739–3751, Sep. 2009.

[26] N. Rossi, P. Ramantanis, and J.-C. Antona, "Nonlinear interference noise statistics in unmanaged coherent networks with channels propagating over different lightpaths," presented at the European Conf. Optical Communication, Cannes, France, 2014, Paper Mo.4.3.4.

[27] A. Bononi, N. Rossi, and P. Serena, "Performance dependence on channel baud-rate of coherent single-carrier WDM systems," presented at the European Conf. Optical Communication, London, U.K., 2013, Paper Th.1.D.5.

[28] A. Bononi, N. Rossi, and P. Serena, "On the nonlinear threshold versus distance in long-haul highly-dispersive coherent systems," *Opt. Exp.*, vol. 20, no. 26, pp. B204–B216, Nov. 2012.

[29] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 4, no. 32, pp. 100–127, Feb. 2014.

[30] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2005.

[31] F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, and S. Bigo, "On nonlinear distortions of highly dispersive optical coherent systems," *Opt. Exp.*, vol. 2, no. 20, pp. 1022–1032, Jan. 2012.

[32] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2008.

[33] P. Serena and A. Bononi, "On the accuracy of the gaussian nonlinear model for dispersion-unmanaged coherent Links;" presented at the European Conf. Optical Communication, London, U.K., 2013, Paper Th.1.D.3.

[34] A. Lau, S. Rabbani, and J. M. Kahn, "On the statistics of intrachannel four-wave mixing in phase-modulated optical communication systems," *J. Lightw. Technol.*, vol. 26, no. 14, pp. 2128–2135, Jul. 2008.

[35] E. Seve, P. Ramantanis, J.-C. Antona, E. Grellier, O. Rival, F. Vacondio, and S. Bigo, "Semi-analytical model for the performance estimation of 100Gb/s PDM-QPSK optical transmission systems without inline dispersion compensation and mixed fiber types," presented at the European Conf. Optical Communication, London, U.K., 2013, Paper Th.1.D.2.

[36] P. Serena and A. Bononi, "Perturbative propagation models for coherent systems," presented at the Asia Communications Photonics Conf., Shanghai, China, 2014, Paper AF1E.3.

[37] Y. Jiang, A. Carena, P. Poggiolini, and F. Forghieri, "On the impact of non-linear phase noise on the sssessment of long-haul uncompensated coherent systems performance," presented at the European Conf. Optical Communication, Cannes, France, 2014, Paper P5.12.

[38] Y. Fan, L. Dou, Z. Tao, L. Li, S. Oda, T. Hoshida, and J. C. Rasmussen, "Modulation format dependent phase noise caused by intra-channel nonlinearity," presented at the European Conf. Optical Communication, Amsterdam, The Netherlands, 2012, Paper We.2.C.3.

[39] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.

[40] Wikipedia. (2014, Jun. 3). Cumulant—Wikipedia, the Free Encyclopedia [Online]. Available: http://en.wikipedia.org/w/index.php?title=Cumulant&oldid=610246675

[41] J. Eriksson, E. Ollila, and V. Koivunen, "Essential statistics and tools for complex random variables," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5400–5408, Oct. 2010.

[42] N. Petrochilos and P. Comon, "A zero-cumulant random variable and its applications," *Signal Process.*, vol. 86, pp. 3334–3338, Mar. 2006.

Authors' biographies not available at the time of publication.