Modeling of the Impact of Nonlinear Propagation Effects in Uncompensated Optical Coherent Transmission Links

A. Carena, *Member, IEEE*, V. Curri, *Member, IEEE*, G. Bosco, *Member, IEEE*, P. Poggiolini, *Member, IEEE*, and F. Forghieri, *Member, IEEE*

***Abstract*—We address perturbative models for the impact of nonlinear propagation in uncompensated links. We concentrate on a recently-proposed model which splits up the signal into spectral components and then resorts to a four-wave-mixing-like approach to assess the generation of nonlinear interference due to the beating of the signal spectral components. We put its founding assumptions on firmer ground and we provide a detailed derivation for its main analytical results. We then carry out an extensive simulative validation by addressing an ample and significant set of formats encompassing PM-BPSK, PM-QPSK, PM-8QAM, and PM-16QAM, all operating at 32 GBaud. We compare the model prediction of maximum system reach and optimum launch power versus simulation results, for all four formats, three different kinds of fibers (PSCF, SMF, and NZDSF) and for several values of WDM channel spacing, ranging from 50 GHz down to the symbol-rate. We found that, throughout all tests, the model delivers accurate predictions, potentially making it an effective general-purpose system design tool for coherent uncompensated transmission systems.**

***Index Terms*—Coherent systems, DWDM, GN model, nonlinear effects, PM-QAM, PM-BPSK, polarization-multiplexed quadrature phase-shift keying (PM-QPSK), Uncompensated transmission.**

I. Introduction

THE LAST few years have been characterized by steady progress in coherent detection systems, achieving ever increasing performance records. Polarization-multiplexed quadrature phase shift keying (PM-QPSK) has been a prominent format in this quest for better performance [1], [2] but lately polarization-multiplexing quadrature-amplitude-modulation with 16 symbols (PM-16QAM) has been attracting very substantial interest too [3], as well as other higher cardinality formats [4], [5].

An important common aspect of all recent records is that they have been achieved over uncompensated links. This circumstance has strongly confirmed earlier predictions, based both on simulative and ad hoc experimental investigations [6]–[8], which showed better performance with uncompensated transmission (UT) versus dispersion-managed transmission (DMT). The current consensus is that green-field installations, as well as major overhauling and refurbishing of existing links, should adopt UT. In this new scenario, there is obvious interest in achieving a better understanding of UT theoretical limitations and in obtaining accurate and effective performance prediction tools.

Long-haul transmission systems are mainly limited by two distinct phenomena: amplified spontaneous emission (ASE) noise accumulation and the generation of nonlinear interference (NLI) due to the Kerr effect in the fiber. ASE noise accumulation is well understood. As for NLI, several approximated models of nonlinear propagation have been proposed over the years. However, in the context of DMT, the goal of describing NLI in a comprehensive, simple and accurate analytical fashion has proved elusive. In contrast, in the context of UT, the prospects for achieving such goal appear more favourable. This is because UT alters the properties of signal propagation in quite a substantial way with respect to DMT. Specifically, there are two key aspects of UT signal propagation that appear to be of great impact.

First, due to the large values of accumulated dispersion, the four electric field components of each transmitted WDM channel appear to quickly take on identical, statistically-independent, zero-mean Gaussian distributions, as they propagate along the link [9]. This phenomenon tends to turn the signal into a noise-like source of nonlinear disturbance, thus amenable to easier statistical manipulations.

In addition, the other remarkable aspect of UT, discussed and simulatively tested in [9], and recently experimentally confirmed in [10], is that after digital signal processing (DSP) the statistical distribution of each of the received constellation points appears to be Gaussian as well, with independent components, *even in the absence of ASE noise in the link*. It seems, in other words, that the effect of NLI could be modeled as excess additive Gaussian noise, at least for low-to-moderate nonlinearity. If so, NLI noise could then simply be added to ASE noise and its impact on system performance could be assessed through a modified signal-to-noise-ratio.

A key question is whether typical transmission systems indeed operate in a low-to-moderate nonlinearity regime. This seems to be the case: it was recently shown that *at maximum reach* the variance of the disturbance due to nonlinearity

Manuscript received July 14, 2011; revised December 23, 2011; accepted February 10, 2012. Date of publication February 27, 2012; date of current version April 04, 2012. This work was supported by CISCO Systems within a SRA contract. It was also supported by the EURO-FOS project, a Network of Excellence (NoE) funded by the European Commission through the 7th ICT-Framework Programme.

A. Carena, V. Curri, G. Bosco, and P. Poggiolini are with Politecnico di Torino, Dipartimento di Elettronica e Telecomunicazioni, C.so Duca Degli Abruzzi 24, 10129 Torino, Italy (e-mail: andrea.carena@polito.it).

F. Forghieri is with Cisco Photonics Italy s.r.l., via Philips 12, 20900 Monza, Italy (e-mail: fforghie@cisco.com).

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

Digital Object Identifier 10.1109/JLT.2012.2189198

amounts to only half that due to ASE noise [11], [12]. The fact that NLI cannot be too large in practical systems also suggests that a perturbative approach could be accurate enough to estimate it.

Over the years, a few nonlinear propagation models based on perturbative approaches have been proposed. Focusing on those that can be applied to UT, they can be broadly classified into two categories: Volterra-series (VS) models and four-wave-mixing-like (FWM-like) models.

The VS models are based on the truncated Volterra-series solution of the nonlinear Schroedinger equation (NLSE) originally proposed in [13]. They can be either expressed in frequency-domain VS [14] or time-domain VS [15].

The FWM-like models are based on ideally slicing up the signal spectrum into spectral components, whose nonlinear beating during propagation is then analytically expressed in a fashion similar to the classical formulas of FWM. FWM-like models have been recently proposed to deal with OFDM coherent systems, where the frequency slicing is quite naturally carried out by looking at the many electrical subcarriers of each WDM channel [16]–[18]. However, the FWM-like approach can also be extended to encompass single-carrier per channel transmission, through proper signal modeling. The first conventional-transmission FWM-like model was actually proposed in 1993 [19] and so it predates all others, including all VS models. It was later reproposed in [20]. More recently, another FWM-like approach has been presented in [21].

Interestingly, many analytical results stemming from both the VS or the FWM-like models are similar, though not identical, showing that once a perturbative approach is undertaken, then the end results tend to agree to a good extent to other perturbative approaches, independently of the derivation procedure.

Though promising, these models have so far undergone only limited validation. If accurate, they could represent quite effective tools for the analysis and design of high-performance coherent systems and therefore a comprehensive validation effort appears to be of substantial interest. Indeed, a main goal of this paper is to undertake such validation campaign through computer simulations. Another main goal is to put on firmer grounds the FWM-like approach for conventional transmission, especially regarding signal modeling.

Specifically, we concentrate on the FWM-like model summarized in [21], which we call "GN model" to stress the 'Gaussian-Noise' assumption for both signal modeling and NLI disturbance on which it is based.

We first address signal modeling in depth and show that it can be accurately carried out by means of spectrally-shaped periodic Gaussian noise with independent spectral components. We also provide a detailed derivation of the NLI power spectral density, which was not supplied in [21].

We then carry out a broad simulative validation effort of such model. This effort encompasses four different formats, namely polarization-multiplexed binary-phase-shift-keying (PM-BPSK), PM-QPSK, PM-8QAM, and PM-16QAM. We elect to operate at a fixed symbol rate while also spanning several values of channel spacing and three different types of fiber: nonzero dispersion-shifted fiber (NZDSF), standard single-mode fiber (SMF) and pure-silica-core fiber (PSCF). In all, we address a realistic, significant and very wide range of systems, over distances spanning between a few hundreds of km, to almost 20 000 km. Our results show the GN model to yield very accurate predictions in all link scenarios and for all modulation formats, confirming its capability to rather precisely predict NLI in UT systems.

The paper is structured as follows. In Section II we provide the details regarding the assumptions and derivation of the GN model. The bulk of signal modeling part and model derivation were transferred to appendices for the readers's convenience. In Section III, we introduce the test system layouts and simulation details. In Section IV, we compare the model results with those obtained through simulations. In Section V we provide comments and conclusions. Several appendices provide specific details regarding various aspects of the model.

II. The GN Model and Its Derivation

We start out by recalling a few fundamental results on the relationship among BER, the electrical and the optical signal-to-noise-ratio (SNR and OSNR). We then introduce modified SNR and OSNR in order to take NLI noise into account. We then proceed towards assessing the NLI noise variance by first properly modeling the transmitted signal and then applying FWM-like formulas to such signal.

A. BER, SNR, and OSNR

Assuming linear propagation, additive Gaussian ASE noise and neglecting polarization-dependent loss (PDL), the BER of any coherent system exploiting QPSK/QAM modulation, including PM systems, can be expressed as a function of the signal-to-noise ratio (SNR) evaluated over the constellation scattering diagram at the decision stage of the Rx, after DSP. Formally

$$\mathrm{BER} = \Psi(\mathrm{SNR}). \tag{1}$$

The function Ψ depends on the modulation format. As an example, for PM-QPSK, it is

$$\mathrm{BER} = \frac{1}{2} \cdot \mathrm{erfc}\left(\sqrt{\mathrm{SNR}/2}\right) \tag{2}$$

For the other formats addressed in this paper, see Appendix A. The SNR is found as

$$\mathrm{SNR} = \frac{\overline{A^2}}{\sigma^2_{\mathrm{ASE}}} \tag{3}$$

where $\overline{A^2}$ is the average of the squared distance of the (noiseless) signal constellation points from the origin and σ^2_{ASE} is the ASE noise variance about each point.

The value of $\overline{A^2}$ depends on the actually transmitted waveforms and on the overall baseband scalar transfer function of the coherent Rx, $H_{\mathrm{Rx}}(f)$, including the adaptive equalizer if present. Assuming that the signal at the Tx does not suffer from either ISI or linear crosstalk among channels and assuming that $H_{\mathrm{Rx}}(f)$ is shaped so that *matched filtering* occurs, it can be shown that

$$\overline{A^2} = P_{\mathrm{Rx}} \cdot R_s^{-1} \int_{-\infty}^{\infty} |H_{\mathrm{Rx}}(f)|^2 \, df \tag{4}$$

where P_{Rx} is the received optical power per channel and R_s is the symbol rate. In the following we will assume matched filtering. This assumption is well justified due to the fact that the DSP adaptive equalizers present in coherent Rx's tend to make $H_{\mathrm{Rx}}(f)$ converge to matched filtering.

As for σ^2_{ASE}, it is found as

$$\sigma^2_{\mathrm{ASE}} = \int_{-\infty}^{\infty} G_{\mathrm{ASE}} \cdot |H_{\mathrm{Rx}}(f)|^2 \, df \tag{5}$$

where G_{ASE} is the unilateral power spectral density (PSD) of dual-polarization ASE noise impinging onto the Rx, down-converted to baseband.

We then proceed to invoke one of the fundamental assumptions on which the GN model, as well as essentially all other models [13]–[20], are based: the effect of NLI on WDM signals in UT systems can be modeled as additive Gaussian noise, statistically independent of ASE noise. See Appendix C for a discussion of this aspect. The consequence of such assumption is that ASE and NLI noise contributions simply add up in variance.

Therefore, the system BER still depends on SNR through (1), but the definition of SNR needs to be modified to include NLI noise

$$\mathrm{SNR}_{\mathrm{NL}} = \frac{\overline{A^2}}{\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLI}}} \tag{6}$$

where, similarly to (5)

$$\sigma^2_{\mathrm{NLI}} = \int_{-\infty}^{\infty} G_{\mathrm{NLI}}(f) |H_{\mathrm{Rx}}(f)|^2 \, df \tag{7}$$

The quantity $G_{\mathrm{NLI}}(f)$ is the unilateral PSD of NLI, down-converted to baseband.

Using (5) and (4) a straightforward relationship can be established between SNR and the widely used optical two-polarization signal-to-noise ratio (OSNR):

$$\mathrm{OSNR} = \frac{R_s}{B_n} \mathrm{SNR} \tag{8}$$

where B_n is the OSNR bandwidth. Note that this formula is valid only for matched Rx filtering, otherwise the relationship between OSNR and SNR is more complex. Again, throughout this paper we assume matched filtering so (8) holds as is.

Combining (6) and (8) we can also define a nonlinear equivalent OSNR, as follows:

$$\mathrm{OSNR}_{\mathrm{NL}} = \frac{R_s}{B_n} \mathrm{SNR}_{\mathrm{NL}}. \tag{9}$$

Throughout this paper, when using OSNR or $\mathrm{OSNR}_{\mathrm{NL}}$, we will always imply $B_n = 12.48$ GHz (0.1 nm).

According to (7), $G_{\mathrm{NLI}}(f)$ is the fundamental quantity that needs to be assessed to characterize the nonlinear behavior of the system. In the following, we present the main steps of its derivation, assuming for convenience single polarization. The final result will however be extended to dual polarization.

For notational clarity, we assume an odd number of channels N_{ch} in the WDM comb and we assume that frequency f is properly shifted so that $f = 0$ represents the center frequency of the center channel. We also assume lumped (EDFA) optical amplification. A model extension to encompass Raman amplification is possible, but will not be dealt with here. In addition, we assume that all spans are identical and that each EDFA exactly compensates for the loss of the preceding span.

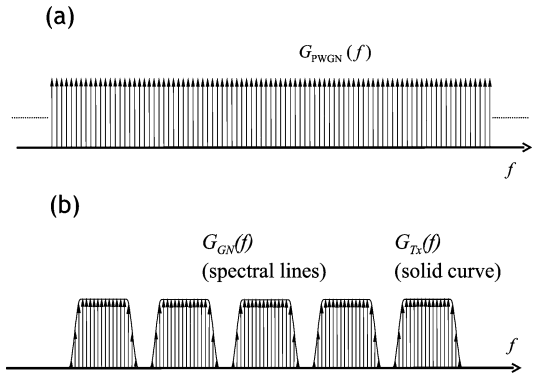

Fig. 1. (a) The average power spectrum of the PWGN noise process of (11); (b) the average power spectrum of the spectrally-shaped periodic Gaussian noise process of (13).

B. Modeling the Transmitted Signal

We start out by making the key assumption that each channel of the comb and, as a consequence, the overall WDM signal, can be modeled as a Gaussian random process. In addition, we need the signal spectrum to be made up of spectral lines, because we want to use a FWM-like approach when assessing NLI. One way to simultaneously satisfy both these requirements consists in resorting to a complex periodic white Gaussian noise (PWGN) process, which we then spectrally shape in a suitable manner.

A PWGN process of period T_0 can be expressed through the Karhunen-Loève formula [22], as

$$E_{\mathrm{PWGN}}(t) = \sqrt{f_0} \sum_{i=-\infty}^{\infty} \xi_i \, e^{j2\pi i f_0 t} \tag{10}$$

where $f_0 = 1/T_0$ and the ξ_k's are identically-distributed statistically-independent complex Gaussian random variables (RV's) which we assume to have unit variance: $\sigma^2_{\xi_n} = 1$. The Fourier transform of such process is

$$E_{\mathrm{PWGN}}(f) = \sqrt{f_0} \sum_{i=-\infty}^{\infty} \xi_i \, \delta(f - i f_0) \tag{11}$$

and its average PSD is

$$G_{\mathrm{PWGN}}(f) = f_0 \sum_{i=-\infty}^{\infty} \delta(f - i f_0) \tag{12}$$

which is depicted in Fig. 1(a).

To obtain the desired transmitted Gaussian-noise (GN) signal model E_{GN} from E_{PWGN}, we spectrally shape the latter so that the resulting 'envelope' of the spectral lines of E_{GN} coincides with the PSD $G_{\mathrm{Tx}}(f)$ of the actually transmitted WDM optical signal

$$E_{\mathrm{GN}}(f) = \sqrt{G_{\mathrm{Tx}}(f)} \sqrt{f_0} \sum_{i=-\infty}^{\infty} \xi_i \, \delta(f - i f_0). \tag{13}$$

In fact the resulting average PSD of $E_{\rm GN}(f)$ is

$$G_{\rm GN}(f) = G_{\rm Tx}(f)\, f_0 \sum_{i=-\infty}^{\infty} \delta(f - if_0) \tag{14}$$

which is depicted in Fig. 1(b).

Note that (13) assumes that each frequency component is statistically independent of all others, since the ξ_k's are inepended RV's. This aspect is discussed in depth in Appendix B. Note also that the assumption of $E_{\rm GN}$ being periodic engenders no actual loss of generality, since the period T_0 can be chosen as large as desired. It can even be made to tend to infinity, as discussed in Appendix D.

C. The NLI Power Spectral Density

The main analytical steps of the derivation of the NLI noise PSD, $G_{\rm NLI}(f)$, are reported in Appendix D. The final expression for the PSD of dual-polarization NLI noise *in a single span* (ss) is

$$G_{\rm NLI,ss}(f) = \frac{16}{27}\gamma^2 \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_{\rm Tx}(f_1)G_{\rm Tx}(f_2)G_{\rm Tx}(f_1+f_2-f) \cdot \left|\frac{1-e^{-2\alpha L_s}e^{j4\pi^2|\beta_2|L_s(f_1-f)(f_2-f)}}{2\alpha - j4\pi^2|\beta_2|(f_1-f)(f_2-f)}\right|^2 df_2\, df_1 \tag{15}$$

where γ is the fiber nonlinearity coefficient, α is the fiber loss parameter, β_2 is the fiber dispersion and L_s is the span length.

Note that the integration range only formally extends to $[-\infty, \infty]$. In practice, it is limited by the optical bandwidth of the WDM signal spectrum $G_{\rm Tx}(f)$. Assuming that $G_{\rm Tx}(f) \neq 0$ for $f \in [-B_o/2, B_o/2]$, where B_o is the overall WDM signal optical bandwidth, then the exact extremes for the integration range are

$$\begin{aligned} f_1 &\in [\max(-B_o/2, f - B_o), \min(B_o/2, f + B_o)] \\ f_2 &\in [\max(-B_o/2, f - f_1 - B_o/2), \\ &\quad \min(B_o/2, f - f_1 + B_o/2)]. \end{aligned} \tag{16}$$

Equation (15) provides the PSD of the NLI noise that is generated in any single span of the link. We then assume that the NLI noise contributions from each span propagate linearly through the rest of the link till they reach the Rx. There, all these NLI contributions add up. This noise sum can be dealt with in two ways.

One way is to assume that the NLI noise generated in any given span can be summed in power, i.e., *incoherently*, with the NLI noise generated in any other span. If so, the PSD of the total NLI noise at the Rx would be, quite simply

$$G_{\rm NLI}(f) = N_s \cdot G_{\rm NLI,ss}(f) \tag{17}$$

where N_s is the number of spans in the link.

Another approach is to assume that the NLI contributions generated in each span are summed at the Rx *coherently*, as shown in Appendix D. The resulting total NLI noise PSD at the Rx is, accordingly

$$G_{\rm NLI}(f) = \frac{16}{27}\gamma^2 \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_{\rm Tx}(f_1)G_{\rm Tx}(f_2)G_{\rm Tx}(f_1+f_2-f) \cdot \left|\frac{1-e^{-2\alpha L_s}e^{j4\pi^2|\beta_2|L_s(f_1-f)(f_2-f)}}{2\alpha - j4\pi^2|\beta_2|(f_1-f)(f_2-f)}\right|^2 \cdot \frac{\sin^2(2N_s\pi^2(f_1-f)(f_2-f)|\beta_2|L_s)}{\sin^2(2\pi^2(f_1-f)(f_2-f)|\beta_2|L_s)}\, df_2\, df_1. \tag{18}$$

Somewhat unexpectedly, given the substantial difference between the resulting integrand functions, the two assumptions deliver results that are typically quite similar. We will come back to this interesting aspect in Section IV.B.

Either (17) or (18) can then be inserted into (7). As a result, to estimate $\sigma^2_{\rm NLI}$ it is typically necessary to carry out a triple numerical integration. Note however that all integrands are everywhere positive. This facilitates the assessment of the convergence of a numerical integration procedure, because such assessment consists in detecting a saturation in the result versus repeated increases in the algorithm nominal accuracy.

Note also that (17), (18) can be integrated analytically when operating with systems at the *Nyquist limit*, that is when each WDM channel has a rectangular spectrum of bandwidth equal to the symbol rate and the channel spacing is equal to the symbol rate as well [21]. We will not deal with this aspect in this paper. For similar results see also [17], [20].

III. Simulated Test System Layout and Back-to-Back Performance

In this section we describe in detail the simulation test set-up that was used to check the model predictions. We employed the commercial optical system simulator OptSim™ to carry out the simulations.

The UT link layout we analyzed is shown in Fig. 2(a). We addressed PM-BPSK, PM-QPSK, PM-8QAM and PM-16QAM. For all modulation formats, we used standard I-Q transmitters based on nested Mach-Zehnder modulators driven by nonreturn-to-zero (NRZ) signals. Each modulated channel was spectrally shaped by a 4th-order super-Gaussian (SG) optical filter, whose −3 dB bandwidth $B_{\rm opt}$ was optimized (see Section III.A). The optical link was composed of $N_{\rm s}$ identical uncompensated spans. Each span was composed of 100 km of transmission fiber followed by an EDFA with noise-figure $F = 5$ dB that completely recovered the span loss $A_{\rm span}$. As transmission fibers, we considered pure-silica-core fiber (PSCF), standard single-mode fiber (SMF) and large-effective-area nonzero-dispersion fiber (NZDSF). The fiber parameters (loss coefficient α, dispersion coefficient D and nonlinear coefficient γ) are reported in Table I.

The Rx structure for each channel is described in Fig. 2(b) and it is the same for the four considered modulation formats. It includes a local oscillator (LO), that was mixed with the in-

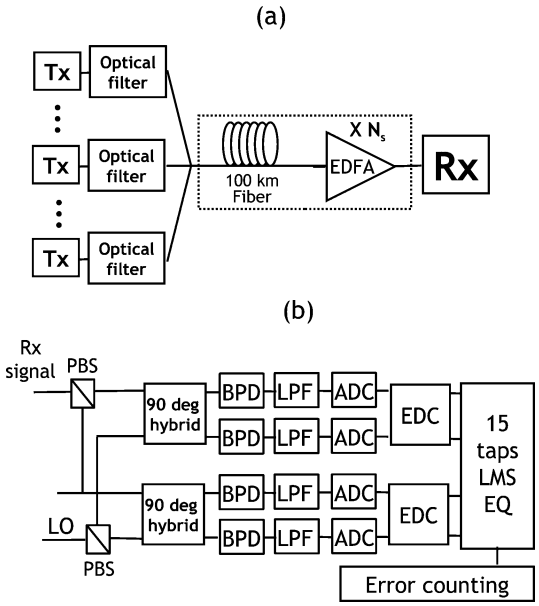

Fig. 2. (a) Layout of the analyzed optical link. (b) Layout of the coherent DSP-based Rx used for the demodulation of a single channel. EDFA: erbium-doped fiber amplifer, N_s: number of spans, LO: local oscillator, PBS: polarization beam splitter, BPD: balanced photo-detector, LPF: low-pass filter, ADC: analog to digital converter, EDC: electronic dispersion compensator, LMS: least-mean square, EQ: FIR-filter-based butterfly equalizer.

TABLE I
PARAMETERS OF THE THREE FIBER TYPES USED IN SIMULATIONS

Fiber	α [dB/km]	D [ps/nm/km]	γ [1/W/km]
PSCF	0.18	20.1	0.9
SMF	0.22	16.7	1.3
NZDSF	0.22	3.8	1.5

coming signal in two 90-degree hybrids, one for each polarization. No optical filtering was present at the Rx: the channel was selected by tuning the LO to the proper frequency. Such tuning was assumed ideal. Lasers were assumed ideal (zero linewidth). Four balanced photo-detectors (BPD) were used to detect the received signal components. The signals were then filtered by a 5th-order low-pass Bessel filter (LPF), with optimized -3 dB bandwidth $B_{\rm Rx}$ (see Section III.A).

After detection and low-pass filtering, each signal was sampled by an analog-to-digital converter (ADC), which was assumed to operate at 2 samples per symbol. We also assumed infinite ADC resolution, so that there was no quantization penalty. After sampling, dispersion was fully and ideally compensated for by the DSP. Then, the four signal components were processed by a multiple-input multiple-output (MIMO) equalizer [23], which consists of four complex 15-tap FIR filters. The equalizer FIR filter coefficients were estimated using the least mean square (LMS) algorithm, first using a training sequence, then operating on the basis of the decided bit string in a decision-directed mode. For each modulation format, a different LMS error function was applied taking into account their own signal constellation and bit mapping. The signals out of the equalizer were used for decision and BERs were evaluated using direct error counting. BER performance was probed on the center channel.

The signaling rate was chosen to be 32 GBaud, for all formats. This symbol-rate corresponds to a net throughput of 25 GBaud plus overheads, encompassing both overhead for forward error correction (FEC) and for network protocols. In our investigation we set the target BER at 10^{-3}, even though the available FEC overhead would have allowed a higher BER threshold, in order to account for a realistic system margin.

Each channel was assembled by encoding onto the constellation points the proper number of pseudo-random bit sequences (PRBSs), each delivering a binary stream at 32 Gb/s. We used 2, 4, 6, and 8 PRBSs for PM-BPSK, PM-QPSK, PM-8QAM and PM-16QAM, respectively. All PRBSs were independent of all others, not only within a single channel, but within the overall WDM comb. The PRBSs were of degree 16 and simulations were carried out transmitting one full sequence, corresponding to a total of 131070 simulated bits for PM-BPSK, 262140 for PM-QPSK, 393210 for PM-8QAM and 524280 for PM-16QAM.

We addressed six different channel spacings, ranging from the standard WDM spacing $\Delta f = 50$ GHz down to the symbol-rate, i.e., $\Delta f = R_S = 32$ GHz. In most tests we simulated WDM propagation with 9 channels but for a selected test we went up to 27 channels corresponding to an overall optical bandwidth equal to 864 GHz.

All link simulations were run with in-line ASE noise. We checked in selected cases whether Rx noise loading produced significant differences, but we always found negligible performance change.

A. Back-to-Back Sensitivity Evaluation

The nonlinear model relies on (3) to estimate the decision-stage SNR in the presence of NLI. However, besides ASE and NLI, other impairments may be present, too. Most prominent, the scattering diagram can also be corrupted by linear inter-channel crosstalk noise and by constellation smearing due to intersymbol interference (ISI).

In our test setups, whether ISI and crosstalk are present essentially depends on the WDM frequency spacing. When $\Delta f = 50$ GHz their impact is small or negligible. When Δf is close to the symbol-rate, channel crosstalk becomes important. It can be reduced by using narrow filtering at the Tx, but this typically causes ISI. As a result, BER estimation based on (1) and (3) is no longer accurate.

We can overcome this problem by adopting a *modified* SNR which adds an extra noise variance $\sigma^2_{\rm XI}$ that accounts for ISI and crosstalk. In back-to-back (btb) we have

$$\mathrm{SNR}_{\rm mod} = \frac{\overline{A^2}}{\sigma^2_{\rm XI} + \sigma^2_{\rm ASE}} \tag{19}$$

where "XI" stands for "Xtalk" and "ISI." If we characterize this extra variance, which is different for each format and frequency spacing, we can then use it to correct the SNR and obtain the exact system BER law in linearity. We also make the assumption that this disturbance is essentially Gaussian-distributed, so that the BER law given by (1) is still valid. For a discussion of this assumption, see Appendix C.

To do so properly, we first point out that the impact of crosstalk and ISI *alone* is clearly independent of the transmitted

power. In other words, the SNR due to these impairments should not depend on the signal strength. This constraint can be met by assuming: $\sigma^2_{\rm XI} = \overline{A^2} \cdot \chi^2_{\rm XI}$. The SNR due to crosstalk and ISI only would then be

$$\mathrm{SNR}_{\rm XI} = \frac{\overline{A^2}}{\sigma^2_{\rm XI}} = \frac{1}{\chi^2_{\rm XI}} \tag{20}$$

which is indeed constant versus signal strength.

We then remark that the level of crosstalk and ISI also depends on the optical transmission filter bandwidth, which should actually be optimized so that the best compromise between ISI (filter too narrow) and crosstalk (filter too wide) is found. This was done by testing the various systems in btb, as explained in the following.

For every considered format and for every frequency spacing, we performed a joint Tx/Rx optimization by varying both the -3 dB bilateral bandwidth $B_{\rm opt}$ of the Tx optical filter, which spectrally shapes each transmitted channel, and the -3 dB bandwidth $B_{\rm Rx}$ of the Rx low-pass electrical filter. The latter has the important function of limiting the aliasing induced by ADC sampling. The simulative optimization was conducted at a target $\mathrm{BER} = 10^{-3}$. For every pair $(B_{\rm opt}, B_{\rm Rx})$ we searched for the OSNR value yielding such target BER. Note that OSNRs were varied using optical noise loading. Remarkably, the performance optimum always fell at approximately $B_{\rm opt} = \Delta f$ and $B_{\rm Rx} = R_S/2$, for all spacings and modulation formats. Such optimum values of filter bandwidths are therefore used throughout the remainder of the paper.

From the btb OSNR values thus found, by means of (8) we calculated the corresponding SNR. Since OSNR includes ASE noise only, the corresponding SNR obtained through (8) is also due to ASE alone, and we call it $\mathrm{SNR}_{\rm ASE}$. Then, we computed the SNR which results from inverting the ideal BER-to-SNR relationship of (1): $\mathrm{SNR}_{\rm BER} = \Psi^{-1}(10^{-3})$. If ASE noise was the only disturbance, we would have $\mathrm{SNR}_{\rm ASE} = \mathrm{SNR}_{\rm BER}$. If instead crosstalk and ISI are present as well, $\mathrm{SNR}_{\rm ASE} > \mathrm{SNR}_{\rm BER}$. Attributing the difference to $\chi^2_{\rm XI}$, the latter simply turns out to be

$$\chi^2_{\rm XI} = \frac{1}{\mathrm{SNR}_{\rm BER}} - \frac{1}{\mathrm{SNR}_{\rm ASE}}. \tag{21}$$

This formula allowed us to obtain the specific $\chi^2_{\rm XI}$ for each modulation format and frequency spacing.

We would like to stress the fact that this characterization procedure is performed in btb and in linearity and has nothing to do with the nonlinear model. It is however necessary because each system configuration has a different btb penalty that needs to be correctly taken into account to later obtain accurate predictions in the nonlinear regime. Thanks to the $\chi^2_{\rm XI}$'s, BER analytical calculations in nonlinearity can then be performed using the overall *modified* SNR

$$\mathrm{SNR}_{\rm NL,mod} = \frac{\overline{A^2}}{\sigma^2_{\rm XI} + \sigma^2_{\rm ASE} + \sigma^2_{\rm NLI}} \tag{22}$$

Fig. 3 shows the btb performance for all modulation formats, in terms of OSNR needed to achieve $\mathrm{BER} = 10^{-3}$, as a function of Δf (solid curves). The dashed horizontal lines are the OSNR values corresponding to the $\mathrm{SNR}_{\rm BER}$ as obtained by inverting

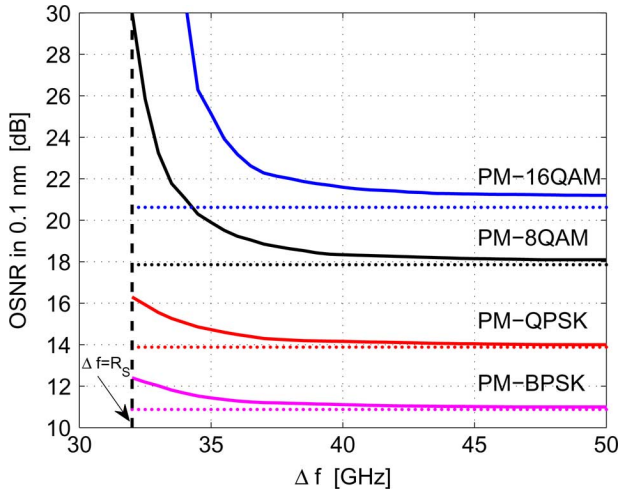

Fig. 3. Solid lines: required OSNR (in 0.1 nm) granting $\mathrm{BER} = 10^{-3}$ on the center channel of the WDM comb, versus Δf. Back-to-back setup with noise loading and optimum filtering. Dotted lines refer to a single-channel ideal Rx using a matched filter.

the theoretical formulas in completely ideal conditions and assuming a matched $H_{\rm Rx}(f)$. As it can be seen, for $\Delta f = 50$ GHz all formats perform almost ideally, with a slight residual penalty for PM-16QAM due to the fact that such format is clearly the most sensitive to signal distortion of all those considered.

While PM-BPSK and PM-QPSK show a limited OSNR penalty when decreasing Δf down to the symbol-rate (32 GHz), PM-8QAM and PM-16QAM are substantially more impacted. For all formats it is however quite clear that the btb characterization of $\chi^2_{\rm XI}$ is indispensable for correct performance assessment at low values of Δf.

IV. Comparison Between Model and Simulation Results

In this section we compare simulative results with model predictions. Throughout Figs. 4–7, noise accumulation was assumed 'incoherent', that is (17) was used to predict NLI noise. This is because the results that we obtained with incoherent noise accumulation turned out to match simulations somewhat better than those using coherent accumulation. Section IV.B is devoted to a discussion of this aspect.

The bulk of our tests was conducted as follows. For each of the four modulation formats, frequency spacing and fiber type, we estimated the system reach $L_{\rm tot}$ as a function of the launched power per channel $P_{\rm Tx}$, both through simulations and using the NLI model together with (1) and (22). As an example, in Fig. 4(a) we show a set of such curves for PM-QPSK over SMF. Each curve refers to a different spacing. In Fig. 4(b) a similar set is shown for PM-8QAM over PSCF. These figures already show what we have consistently found throughout this comparison: the model (solid lines) matches rather closely the simulative results (markers).

Since showing similar plots for all twelve combinations of formats and fibers would be impractical, we organized the results in a different way. The key information in the plots of Fig. 4 is clearly the maximum of each curve, which represents the *maximum achievable system reach*, $L_{\rm max}$. We estimated $L_{\rm max}$ for all

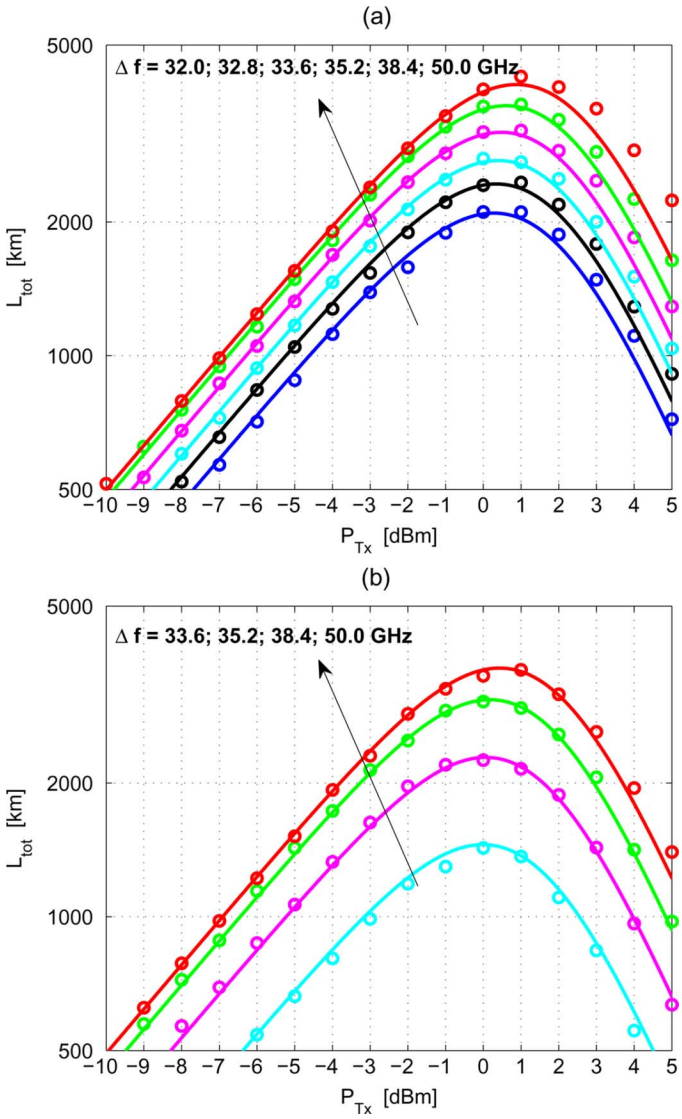

Fig. 4. System reach L_{tot} versus launch power per channel P_{Tx}, 9 channels at 32 GBaud, $\mathrm{BER} = 10^{-3}$. Solid lines: GN model predictions. Markers: simulative results. The different curves correspond to different channel spacings, as listed in the figures. Plot (a) PM-QPSK over SMF; plot (b) PM-8QAM over PSCF.

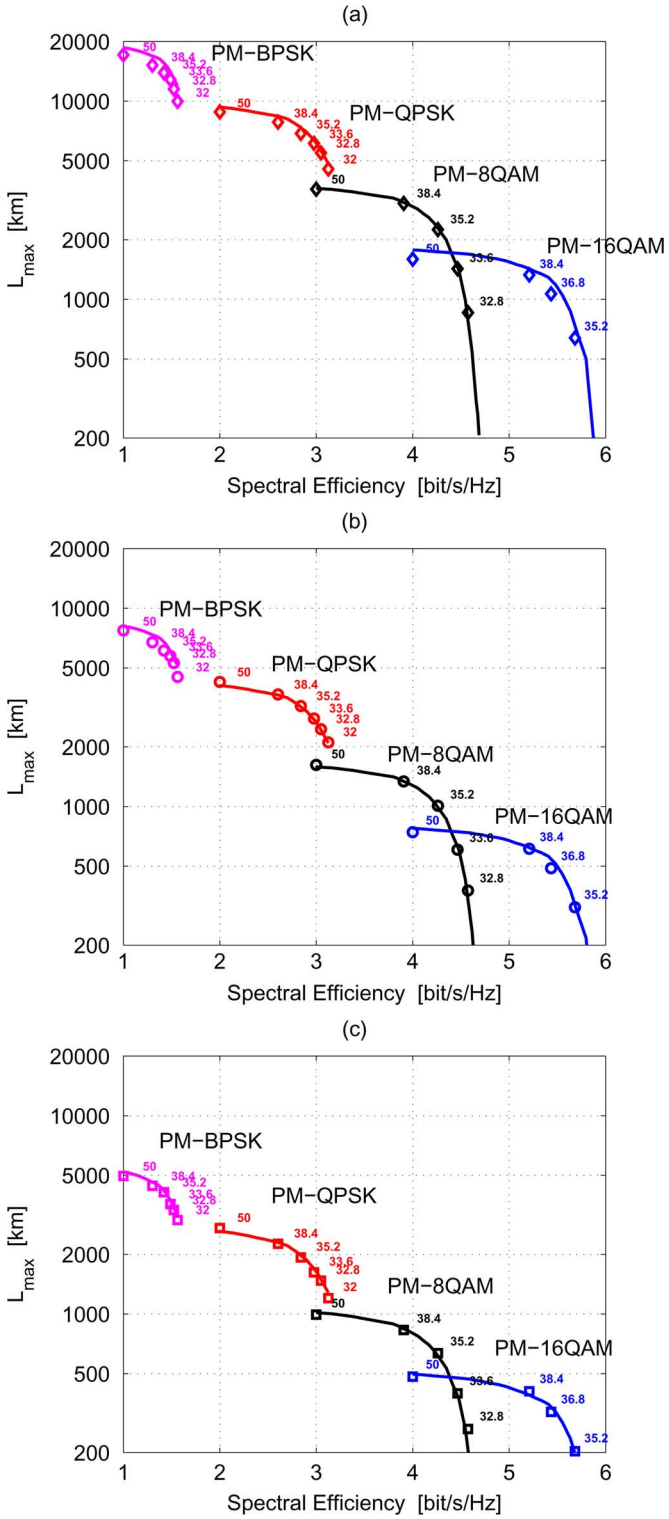

Fig. 5. Maximum achievable system reach L_{max} versus net spectral efficiency SE, in bit/s/Hz, 9 channels at 32 GBaud, $\mathrm{BER} = 10^{-3}$. Solid lines refer to model and markers to simulative results: each colored curve corresponds to a different modulation format as labeled in the figure. Numbers next to markers specify channel spacing used in simulations. (a) PSCF plot. (b) SMF plot. (c) NZDSF plot.

system configuration, both by simulations and using the model. We then gathered all results in three plots, one per fiber type, of L_{max} versus net spectral efficiency (SE), shown in Fig. 5. The net spectral efficiency is defined as SE=$M \cdot R'_s/\Delta f$, where M is the number of bits per symbol carried by each format (2 for PM-BPSK, 4 for PM-QPSK, 6 for PM-8QAM and 8 for PM-16QAM) and $R'_S = 25$ GBaud is the net payload symbol-rate. The latter is found by subtracting from the gross symbol-rate of 32 GBaud the FEC and protocol overhead, which we assume to globally amount to 28%. For each format and for each fiber a different curve was drawn, while varying the channel spacing parameter. In Fig. 5, solid lines are model predictions, whereas markers are simulations. Labels refer to the value of frequency spacing.

The most prominent feature of Fig. 5 is that the correspondence between model predictions and simulations is very good across all modulation formats, fibers and spacings. Notice that the plots range between extremes that are drastically far apart: from a PM-BPSK system reaching 18 000 km over PSCF, at a spectral efficiency of 1.0 bit/s/Hz, all the way to a PM-16QAM

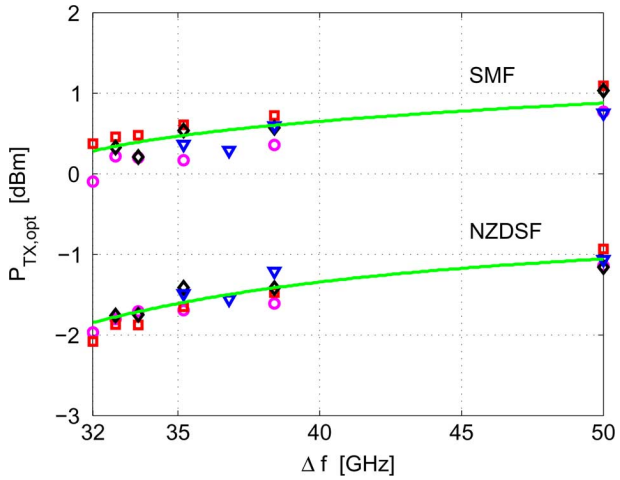

Fig. 6. Optimal launch power $P_{\mathrm{Tx,opt}}$ as a function of channel spacing Δf for two fiber types: SMF and NZDSF. Solid lines: model predictions; markers: simulations (circles: PM-BPSK; squares: PM-QPSK; diamonds: PM-8QAM; triangles: PM-16QAM).

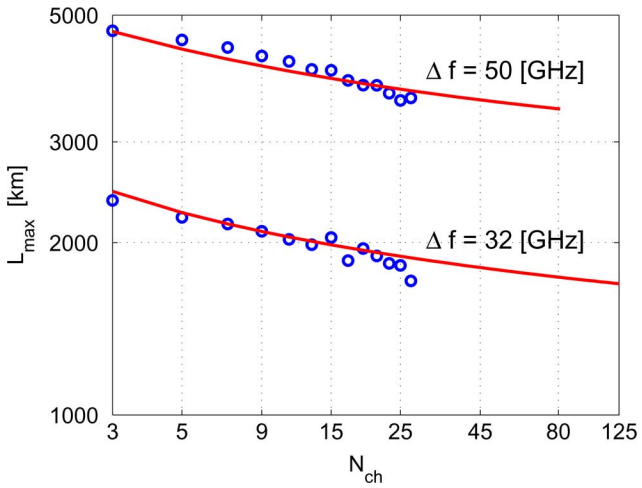

Fig. 7. Maximum system reach $L_{\max}$ versus number of WDM channels considering PM-QPSK over SMF at 32 GBaud for channel spacing equal to 32 GHz and 50 GHz, $\mathrm{BER} = 10^{-3}$. Solid lines: model predictions; markers: simulations.

system achieving 200 km over NZDSF at a spectral efficiency of 5.7 bit/s/Hz. This shows that the model predictive capability is quite broad.

Besides maximum reach, the other key parameter for system analysis or design is the optimum launch power per channel $P_{\mathrm{Tx,opt}}$, i.e., the launch power corresponding to $L_{\max}$. We collected the $P_{\mathrm{Tx,opt}}$'s for all formats, fibers and spacings. We then plotted such $P_{\mathrm{Tx,opt}}$'s versus frequency spacing, for all system configurations over SMF and NZDSF fibers, as shown in Fig. 6. PSCF results are very similar and are not shown for readability of Fig. 6, because they are partially overlapped to SMF data. Different formats are identified with different marker shapes. Solid lines are model predictions. Note that only two curves are present, one per fiber type. This is because the model, somewhat counter-intuitively, predicts that $P_{\mathrm{Tx,opt}}$ does not depend on the modulation format. It should be remarked, though, that same $P_{\mathrm{Tx,opt}}$ in Fig. 6 does not necessarily mean same $L_{\max}$. The latter *does* depend on format and all the other system parameters, as shown in Fig. 5. In any case, the simulative results confirm this interesting behavior.

A. Performance versus Number of Channels

To further extend the range of test scenarios for the model, we decided to increase the number of channels. In so doing, we also tried to address a question that has not been yet conclusively answered in the context of coherent UT systems: how does performance depend on the number of WDM channels? A related question is: how many channels need to be simulated, or actually implemented in an experiment, to ensure that interchannel crosstalk, or NLI, is fully taken into account?

We concentrated on PM-QPSK over SMF at spacings 32 and 50 GHz. We simulatively estimated $L_{\max}$ versus the number of transmitted channels N_{ch}. We increased N_{ch} till it was practically possible due to CPU time constraints. We then computed the model prediction, increasing N_{ch} till the whole C-band (4 THz) was filled up, corresponding to 80 and 125 channels respectively at 50 and 32 GHz. The results are shown in Fig. 7, where simulations are markers and the model is a solid line. The simulations, as far as it was possible to perform them, fully confirm the model accuracy, even in this wide-optical-bandwidth scenario.

In essence, from Fig. 7 it appears that there is no saturation of performance degradation versus the number of channels. In this respect, there seems to be no 'sufficiently large' number of channels ensuring that NLI is fully taken into account. However, Fig. 7 clearly shows that performance degrades quite slowly as N_{ch} goes up.

On the other hand, it is also clear that the performance of a few-channel experiment is not fully representative of a complete 'C' or 'C+L' band system performance. In this respect, the model presented here could be of substantial help: since the model appears to be quite accurate in predicting system performance versus N_{ch}, then the performance of a few-channel system could be extrapolated through the model to reliably predict it for a larger channel count.

B. Coherent versus Incoherent Noise Accumulation

In Fig. 8, we show a plot of maximum reach $L_{\max}$ versus Δf, for PM-QPSK, on SMF and NZDSF, with 9 channels. The incoherent model of (17) delivers the solid lines whereas the coherent model of (18) yields the dashed line. The two predictions, despite the substantially different analytical forms of the PSD equations used to obtain them, are not far apart. Yet, they are not the same and it is important to better investigate this divergence.

To do so, in Fig. 9 we plotted for NZDSF the ratio ρ between P_{NLI} computed with the coherent and with the incoherent model, as a function of N_s

$$\rho = \frac{P_{\mathrm{NLI,coherent}}}{P_{\mathrm{NLI,incoherent}}}. \tag{23}$$

The two P_{NLI} coincide after the first span so $\rho = 1$ for $N_s = 1$. Adding spans, $P_{\mathrm{NLI,coherent}}$ grows slightly more than

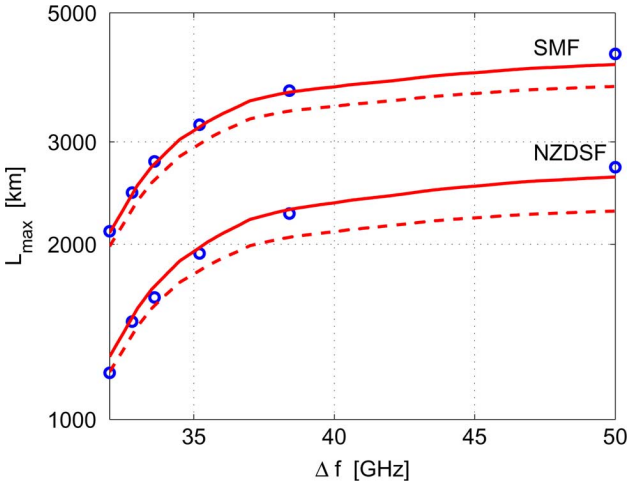

Fig. 8. Maximum system reach $L_{\max}$ versus Δf, for PM-QPSK at 32 GBaud on SMF and NZDSF, 9 channels. Solid lines: incoherent model from (17); dashed lines: coherent model from (18); markers: simulations.

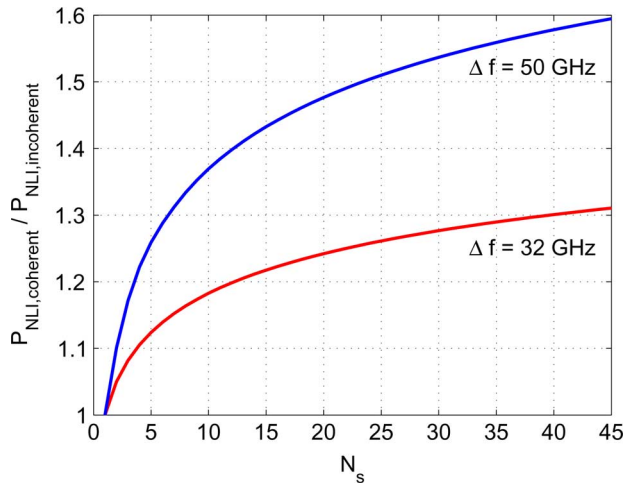

Fig. 9. Ratio between $P_{\rm NLI}$ computed with the coherent model of (18) and $P_{\rm NLI}$ computed with the incoherent model of (17) as a function of the number of spans N_s, for a 9-channel WDM system at 32 GBaud with frequency spacing $\Delta f = 32$ GHz and 50 GHz, on NZDSF fiber. The curves are well fitted by (24) with α =0.08 and 0.16, for 32 and 50 GHz spacing, respectively.

$P_{\rm NLI,incoherent}$, so ρ steadily goes up. The ρ curve turns out to be well fitted by the heuristic law

$$\rho = (1 + \alpha \cdot \log N_s). \tag{24}$$

Since by assumption $P_{\rm NLI,incoherent}$ grows linearly versus N_s, the heuristic law implies that $P_{\rm NLI,coherent}$ must grow approximately according to

$$P_{\rm NLI,coherent} \propto N_s(1 + \alpha \cdot \log N_s) \tag{25}$$

that is, $P_{\rm NLI,coherent}$ grows slightly super-linearly versus N_s.

Note that (17)–(18) are independent of modulation format and therefore α is independent of the modulation format too. On the other hand, we tried changing various system parameters and it turned out that α depends on most of them. As an example, α is smaller when dispersion is higher: with 9 WDM channels, for $\Delta f = 32$ GHz we find $\alpha = 0.08$ over NZDSF while $\alpha = 0.06$ over SMF. Also, α grows with Δf: going up to $\Delta f = 50$ GHz over NZDSF, α grows from 0.08 to 0.16 (see Fig. 9).

Interestingly, α decreases substantially when the number of channels grows, that is, when the overall optical bandwidth B_o grows. For instance, with $\Delta f = 32$ GHz over SMF, going from 9 channels to 125 channels (that is $B_o = 4$ THz, the overall C band) α shrinks from 0.06 to a mere 0.035. Note that for such a low value of α, coherent and incoherent NLI accumulation differ very little and from a practical viewpoint the two models provide the same predictions.

We tried to find out which analytical model matched simulation results better. Our simulations hinted at a better accuracy of the incoherent model (see markers in Fig. 8). On the other hand, many factors can cause small deviations in simulation results and we believe that our evidence is insufficient to conclusively establish which model is closer to reality. Therefore, in our opinion, the issue of coherent versus incoherent noise accumulation is not settled and deserves further investigation.

C. Ultralow Dispersion Scenarios

In this paper we have concentrated on model validation using "typical" fibers. The lowest assumed dispersion value, that of NZDSF, was $\mathrm{D} = 3.8$ ps/nm/km: such value is low, but not pathologically so. Our results appear to confirm that down to this value the model accuracy is good.

We have however preliminary evidence that the model may start being less accurate when $\mathrm{D} < 2$ ps/nm/km. This is probably due to the signal not achieving a Gaussian distribution for a large part of the link length. We consider the investigation of ultra-low dispersion regimes outside of the scope of the present paper but would like to point out this circumstance nonetheless.

D. Experimental Validation

After the submission of the present manuscript, we have carried out a model validation experiment, using ten densely spaced 120 Gb/s PM-QPSK channels, propagating over either PSCF, SMF or NZDSF [24]. The experiment confirmed a good predictive accuracy for the model. Further validation efforts using PM-16QAM are currently underway.

V. Conclusion

In this paper, we have addressed a perturbative FWM-like model for the impact of nonlinear propagation in uncompensated links. Specifically, we concentrated on the GN model preliminarily introduced in [21], pointing out its connections with other perturbative approaches [13]–[20]. We have put the GN model on firmer ground as to the key assumptions regarding signal modeling which permit to derive it, and we have provided the details regarding its analytical derivation.

We have then carried out a thorough simulative validation of the GN model, encompassing four different modulation formats, three different fiber types and several values of WDM channel spacing. The results consistently indicate that the model is rather accurate and dependable in all the addressed systems scenarios.

We point out that the model has no free fitting parameters and that all physical parameter dependencies were the sole consequence of the analytical derivation. This suggests, on *a posteriori* basis, that the assumptions which the model is based on

must actually hold true to a substantial extent. Specifically, the assumption that NLI can be viewed as additive noise, essentially Gaussian, independent of ASE noise, must in fact be quite accurate. Also, the modeling of the transmission signal as zero-mean Gaussian noise, with the same PSD as the actual WDM signal, must be capable of capturing the relevant signal features, as far as the generation of NLI is concerned. Furthermore, it appears that NLI is indeed generated as the mixing of fine spectral components through a FWM-like mechanism, and that this single effect can in fact account for most of the relevant nonlinear propagation impairments in UT systems.

Given the similarities between the GN model [21] and the other perturbative models, pointed out in Section I, this validation effort indirectly validates this class of models as a whole. Minor differences should be investigated to achieve consensus and perhaps even better predictive power.

One intriguing aspect specifically needs further examination: NLI noise accumulation across different spans can be assumed to occur coherently or incoherently. Interestingly, the two assumptions produce results that are typically rather similar, so that from a practical viewpoint this aspect is of limited impact. Nonetheless, it should be clarified, to achieve a deeper understanding of NLI noise generation and accumulation in optical links.

In any case, the GN model, already in its present form, appears to provide an effective tool for the high-level analysis and design of coherent UT system, across a very wide range of system scenarios.

Appendix A
BER Formulas

The BER formulas used in this paper are as follows.

- For PM-BPSK

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\mathrm{SNR}}\right). \tag{26}$$

  This formula is exact.

- For PM-QPSK

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\mathrm{SNR}/2}\right). \tag{27}$$

  This formula is exact, assuming Gray coding.

- For PM-8QAM

$$\mathrm{BER} = \frac{2}{3}\mathrm{erfc}\left(\sqrt{\frac{3}{14}\mathrm{SNR}}\right). \tag{28}$$

  Gray coding is not possible for PM-8QAM. Also, no closed-form BER formulas are available. Assuming the best possible coding, the above formula was found as a very tight best-fit of the actual BER curve. The accuracy is better than ± 0.05 dB of SNR over the range 10^{-2} to 10^{-4}.

- for PM-16QAM

$$\mathrm{BER} = \frac{3}{8}\mathrm{erfc}\left(\sqrt{\frac{1}{10}\mathrm{SNR}}\right). \tag{29}$$

  It is derived from a tight approximation of the exact SER formula, then converted into BER as: $\mathrm{BER} = \mathrm{SER}/4$. Assuming Gray coding, the accuracy is better than ± 0.05 dB of SNR over the range 10^{-2} to 10^{-4}.

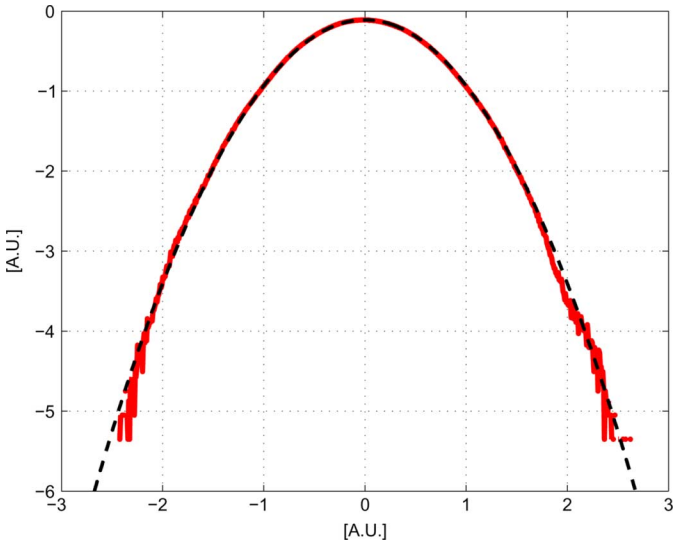

Fig. 10. Vertical bars: Histogram of the samples of one of the four components of the signal after 500 km of SMF, linear propagation, for PM-QPSK transmission at 32 GBaud. Red dashed line: analytical Gaussian distribution. Vertical and horizontal axes: arbitrary units.

Appendix B
Gaussian-Noise Assumption on the Transmitted Signal

In this Appendix we discuss the assumption that the transmitted signal can be modeled as a Gaussian noise process.

Regarding the circumstance that the signal behaves as a Gaussian noise process after propagation across even relatively limited fiber stretches, this can be verified easily by direct simulation. For instance, in Fig. 10 we show the histogram of the time samples of one of the four components of a 32 GBaud PM-QPSK signal propagated across 500 km of SMF, in linearity. The solid line is a Gaussian fit and it is quite evident that the matching is very good. The same conclusion can also be drawn based on the analytical argument that at any point in time the resulting field is a linear combination of contributions coming from many different random data symbols (tens to hundreds, typically). The sum of many random independent contributions converges to Gaussian. It therefore appears that the Gaussianity assumption is reasonable and it approximates the signal features quite effectively.

However, the modeling of such Gaussian process through (13) requires a more detailed discussion. Specifically, we assume that the RVs appearing in (13) are Gaussian, zero-mean, with same variance and independent of one another. The independence assumption, in particular, may be questionable and needs to be discussed.

We do so for a single channel, being evident that if this is proved for a single channel, it is also proved for an overall WDM comb of channels. We will approach this task by comparing the signal model given by (13) with the Fourier transform of an actual transmitted signal, $E_{\mathrm{a}}(f)$. To ease the comparison, we

shall assume that $E_{\rm a}(f)$ too is periodic in time-domain. As discussed in Section II.B this assumption causes no loss of generality, because the signal period can be made arbitrarily large and even tend to infinity. Also, it is customary to use periodic test sequences in laboratory experiments, albeit long ones, without considering it a limitation as to the validity of the experiments.

Recalling that such period is $T_0 = W \cdot T_s$, with W an arbitrary integer, we can write the actual transmitted signal as

$$E_a(t) = \sum_{n=-\infty}^{\infty} q(t - nT_0). \tag{30}$$

The signal $q(t)$ can in turn be written as

$$q(t) = \sum_{w=0}^{W-1} a_w\, s(t - wT_s) \tag{31}$$

where $s(t)$ is the elementary "pulse" carrying a single symbol and a_w is the complex number mapping the bits of the w-th symbol onto a constellation point. We consider the a_w's to be independent RV's. All customary constellations, such as BPSK, QPSK, 16-QAM, 64-QAM, etc., are such that $\mathbf{E}\{a_w\} = 0$, so we make this assumption too.

We then proceed to compute the Fourier transform of the signal $E_{\rm a}(t)$. According to well-known properties of periodic signals [22], it is immediately written as

$$E_{\rm a}(f) = f_0 \sum_{n=-\infty}^{\infty} q(nf_0)\, \delta(f - nf_0) \tag{32}$$

where

$$\begin{aligned} q(f) &= \mathrm{F}\{q(t)\} = \mathrm{F}\left\{\sum_{w=0}^{W-1} a_w\, s(t - wT_s)\right\} \\ &= \sum_{w=0}^{W-1} a_w\, \mathrm{F}\{s(t - wT_s)\} = \sum_{w=0}^{W-1} a_w\, s(f) e^{-j2\pi f w T_s} \end{aligned} \tag{33}$$

and F means Fourier transform. Substituting (33) into (32) we get:

$$\begin{aligned} E_{\rm a}(f) &= f_0 \sum_{n=-\infty}^{\infty} \sum_{w=0}^{W-1} a_w\, s(f)\, e^{-j2\pi w T_s f}\, \delta(f - nf_0) \\ &= s(f) \sum_{n=-\infty}^{\infty} \delta(f - nf_0) \sum_{w=0}^{W-1} f_0\, a_w\, e^{-j\frac{2\pi}{W} wn}. \end{aligned} \tag{34}$$

The above equation can be rewritten as

$$E_{\rm a}(f) = \sqrt{f_0\, |s(f)|^2} \sum_{n=-\infty}^{\infty} \zeta_n\, \delta(f - nf_0) \tag{35}$$

where the ζ_n's are RV's defined as

$$\zeta_n = \sqrt{f_0} \sum_{w=0}^{W-1} a_w e^{-j\frac{2\pi}{W} wn} e^{j\varphi_s(nf_0)} \tag{36}$$

where $\varphi_s(nf_0) = \arg\{s(nf_0)\}$. Note that the ζ_n's are RV's because their definition contains the a_w's, which are RV's too.

Now we can compare the signal model of (13) with the actual signal (35). They are clearly quite similar. In fact, if we could prove that the ζ_n's in (35) have the same statistical features as the ξ_n's in (13), then we could conclude that indeed (13) is an accurate model for the signal.

To do so, we need to show that the ζ_n's are Gaussian statistically independent complex RV's. The Gaussianity of the ζ_n's is immediately proved by the central-limit theorem. In fact, according to (36), each ζ_n arises from the sum of W independent complex RV's of the form

$$\sqrt{f_0} a_w e^{-j2\pi wn/W} e^{j\varphi_s(nf_0)}. \tag{37}$$

Since we can make W arbitrarily large, convergence of the distribution of each ζ_n to a Gaussian distribution can be made arbitrarily accurate.

Regarding the independence of the ζ_n's, we directly calculate the covariance between any two of them

$$\begin{aligned} &\mathbf{E}\{\zeta_n \zeta_m^*\} \\ &\quad = f_0\, \mathbf{E}\left\{\sum_{w_1=0}^{W-1} a_{w_1} e^{-j\frac{2\pi}{W} w_1 n} e^{j\varphi_s(nf_0)} \right. \\ &\qquad \left. \cdot \sum_{w_2=0}^{W-1} a_{w_2}^* e^{j\frac{2\pi}{W} w_2 m} e^{-j\varphi_s(mf_0)} \right\} \\ &\quad = f_0\, e^{j[\varphi_s(nf_0) - \varphi_s(mf_0)]} \\ &\qquad \cdot \sum_{w_1=0}^{W-1} \sum_{w_2=0}^{W-1} \mathbf{E}\left\{a_{w_1} a_{w_2}^*\right\} e^{-j\frac{2\pi}{W} w_1 n} e^{j\frac{2\pi}{W} w_2 m} \\ &\quad = f_0\, \sigma_a^2 e^{j[\varphi_s(nf_0) - \varphi_s(mf_0)]} \\ &\qquad \cdot \sum_{w_1=0}^{W-1} \sum_{w_2=0}^{W-1} \delta_{w_1 w_2} e^{-j\frac{2\pi}{W} w_1 n} e^{j\frac{2\pi}{W} w_2 m} \\ &\quad = f_0\, \sigma_a^2\, e^{j[\varphi_s(nf_0) - \varphi_s(mf_0)]} \sum_{w=0}^{W-1} e^{j\frac{2\pi}{W} w(m-n)}. \end{aligned}$$

The last summation has the following behavior

$$\sum_{w=0}^{W-1} e^{j\frac{2\pi}{W} w(m-n)} = \begin{cases} 0 & \text{if } [m \neq n \text{ and } (m-n) \neq kW] \\ W & \text{if } (m-n) = pW \end{cases}. \tag{38}$$

As a result, the ζ_n's are uncorrelated, and hence independent, unless $(m - n) = pW$, where p is an integer. Physically speaking, this means that there is correlation between those spectral lines in (35) that are $pWf_0 = pR_s$ apart, i.e., that are multiples of the symbol-rate apart.

Therefore, in general, (13) is not an exact model of (35). However, it is in fact exact in one significant case. If we assume to be operating at the Nyquist limit, i.e., with a pulse $s(f)$ whose spectrum is rectangular with bandwidth equal to R_s. In this case, the summation in (35) can be limited to those n's for which: $nf_0 \in [-R_s/2, R_s/2]$. As a result, there are no two nonzero spectral components in (35) that are (pR_s) apart, for any value

of p, including $p = 1$. In other words, there are no two nonzero spectral components in (35) that are correlated.

When the spectrum is nonrectangular, then there is correlation, but at frequency distance $(p\,R_s)$ only. In fact, given that modern systems have Tx spectra whose bandwidth is typically limited to less than $(2\,R_s)$, then each signal frequency component is correlated with none, or at most one other frequency component. This correlation appears mild enough to enable us to use the approximate signal model of (13) in our calculations. The very good accuracy that the GN model appears to have, seems to confirm that the mild correlation present in (35) leads to negligible errors in the estimation of the NLI noise spectrum.

We would also like to point out that in principle it is possible to revise the GN model by incorporating the correlation present in (35). This leads to substantially more complex calculations, that are at present work in progress.

A final comment must be devoted to an intriguing aspect of the signal features. As shown above, the actual Tx signal has the same spectral features as that of a periodic Gaussian noise process to a large extent, and exactly so at the Nyquist limit. Notice that this is already the case in btb: in fact, the above calculations were carried out in btb, without dispersion being considered.

However, *in time-domain*, the actual Tx signal in btb *does not* typically look like Gaussian noise. In fact, the btb transmission signal can be pictured through the customary eye diagram and is not Gaussian-distributed. So, the paradox is that in frequency domain the btb Tx signal appears like a Gaussian noise process whereas in time-domain it does not. Notice however that this seeming paradox ceases to exist as soon as propagation starts. When dispersion comes into play, every frequency component is phase-scrambled. This scrambling destroys the precise phase-matching that produces the orderly signal pulses in time-domain in btb. Once this has occurred, as shown in Fig. 10, the statistical distribution of the signal in time-domain does indeed take on a Gaussian shape. Nonetheless, this aspect is intriguing and is being investigated, even though we expect no repercussion on the GN model results from such investigation.

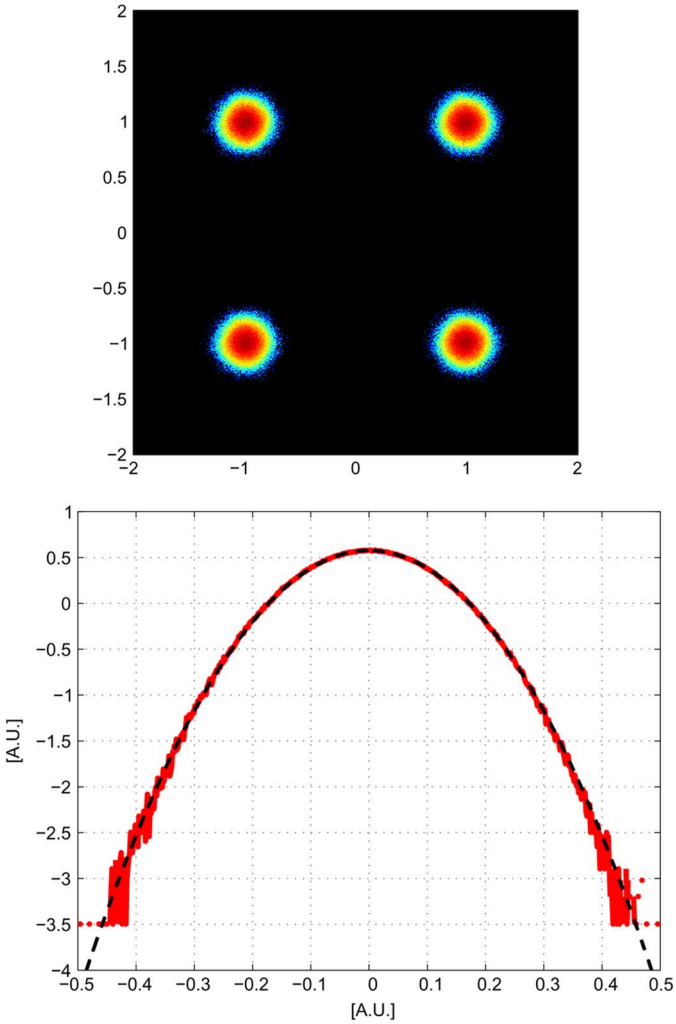

Fig. 11. Top: PM-QPSK back-to-back received constellation in the absence of ASE noise, at a channel spacing equal to the symbol rate (32 GHz). Bottom: histogram of the received signal points about their statistical average, all points together, one quadrature. The other quadrature and polarization have similar distributions. The dashed line is an ideal Gaussian distribution with same variance as the histogram.

APPENDIX C
GAUSSIAN-NOISE ASSUMPTION ON THE RECEIVED SIGNAL

The expression of SNR when both ISI/crosstalk and NLI noise are taken into account is given by (22). By writing this formula and using it in (1) we implicitly assume that these noises and ASE noise are all independent among themselves and are all Gaussian, at least to 'sufficient' degree.

As for independence, this was discussed elsewhere and in addition it appears to be a quite reasonable assumption. Regarding Gaussianity, in the following we address this issue, first for ISI/crosstalk and then for NLI noise.

A. Interchannel Crosstalk and ISI

Fig. 11 shows the received constellation and the histogram of the samples about the constellation points (one quadrature) for PM-QPSK. The simulated symbols were 500 000. The dashed line shows a Gaussian distribution with same variance as the signal samples. Gaussianity is seemingly verified to a very high degree. The same result is found in Fig. 12, concerning PM-16QAM. Not shown, similar results are found for PM-BPSK and PM-8QAM. We tested the assumption of the Gaussianity of ISI/crosstalk by looking at the statistical distribution of such disturbance on the received constellation points in back-to-back, when no ASE is present and there is no chromatic dispersion. We show here the results for PM-QPSK at symbol-rate spacing (32 GHz) and of PM-16QAM at 1.1 times the symbol rate (35.2 GHz). Note that the PM-16QAM spacing is larger because at symbol-rate spacing it does not achieve $\mathrm{BER} = 10^{-3}$. We chose tight spacings because ISI/crosstalk is substantial only in this case, whereas it tends to be negligible for larger ones.

Based on these results, we can conclude that the Gaussian assumption on ISI/crosstalk appears to be verified to a sufficient extent to justify dealing with ISI/crosstalk noise as an additional Gaussian noise contribution.

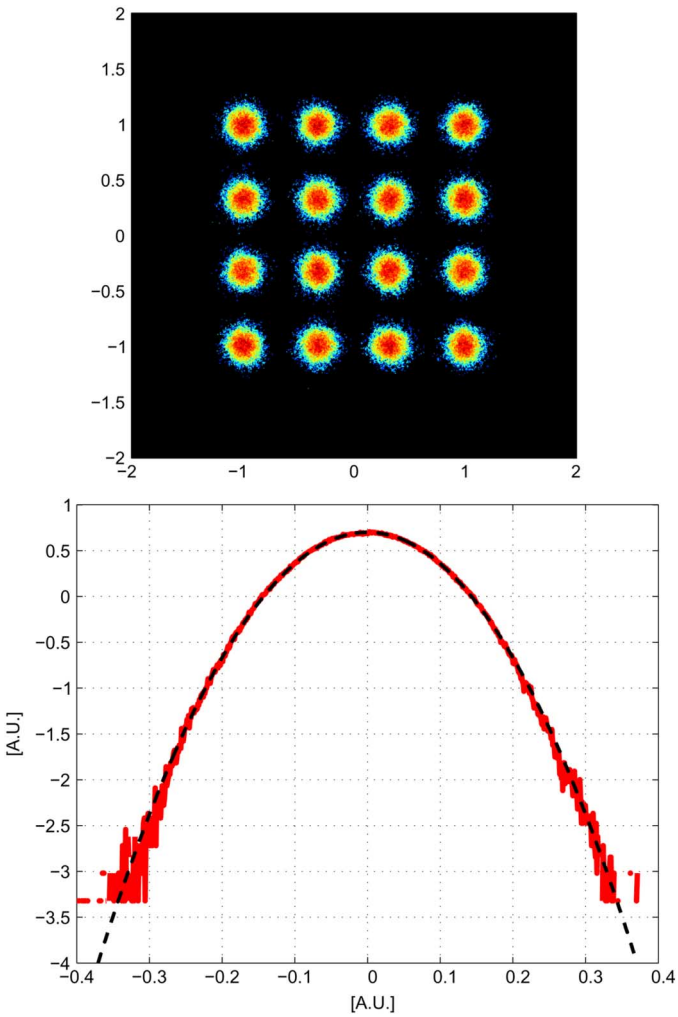

Fig. 12. Top: PM-16QAM back-to-back received constellation in the absence of ASE noise, at a channel spacing equal to 1.1 times the symbol rate (35.2 GHz). Bottom: histogram of the received signal points about their statistical average, all points together, one quadrature. The other quadrature and polarization have similar distributions. The dashed line is an ideal Gaussian distribution with same variance as the histogram.

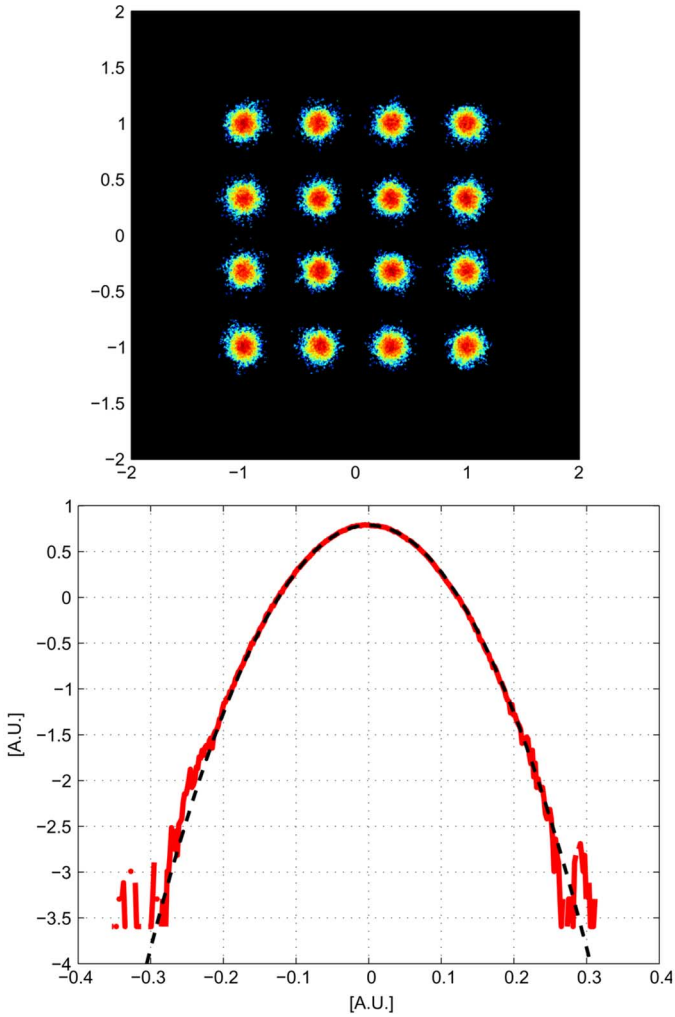

Fig. 13. Top: PM-16QAM received constellation in the absence of ASE noise, at a channel spacing equal to 1.2 times the symbol rate (38.4 GHz), over 600 km of SMF. Launch power is the optimum one from Fig. 6. Bottom: histogram of the received signal points about their statistical average, all points together, one quadrature. The other quadrature and polarization have similar distributions. The dashed line is an ideal Gaussian distribution with same variance as the histogram.

B. NLI Noise

As mentioned in the introduction, the Gaussianity of NLI noise was simulatively shown to be well verified for PM-QPSK in [9]. After this paper was submitted, an experimental test was carried out as well [10], where a 1500 km 80-channel system using 112 Gb/s PM-QPSK was tested. Fiber was SMF. The experiment showed a very high degree of noise Gaussianity about the constellation points, both in a rather linear regime and at a substantially nonlinear regime, even beyond the maximum-reach launch power for the system. These results appear as a strong indication of substantial NLI noise Gaussianity, at least with PM-QPSK.

The question remains whether Gaussianity is also found for higher-order transmission formats. To answer this question we tested the Gaussianity of received noise on higher-order formats at their maximum reach, that is at the distance reported in Fig. 5, corresponding to the top of their reach versus launch-power curves (similar to those shown in Fig. 4).

In Figs. 13 and 14, we show the results for PM-16QAM over SMF (600 km). PM-8QAM had completely equivalent results, which we omit for brevity. The spacing was set at 38.4 GHz, where the impact of ISI and linear crosstalk on the constellation noise is low. In other words, what is shown in the figures contains almost no ISI and crosstalk. The simulated symbols were 500 000. We carried out the test with and without ASE noise. When present, ASE noise was injected along the transmission line (not just added at the receiver). The figures show the received constellation and the histogram of the samples about the constellation points (one quadrature). The dashed lines show a Gaussian distribution with same variance as the signal samples.

Note that, qualitatively, NLI noise alone appears in Fig. 13 to have a relatively small variance. In fact, in [11], we pointed out that at maximum reach the ratio of ASE noise variance versus NLI noise variance is 2, that is 2/3 of the total noise variance is ASE and only 1/3 is NLI. This circumstance was then also discussed in [12] and experimentally confirmed in [10].

In any case, Gaussianity is seemingly verified to a very substantial degree, for both the case when ASE is present or when ASE noise is absent, although, in practice, the truly significant case is that of ASE noise present, the other being presented here

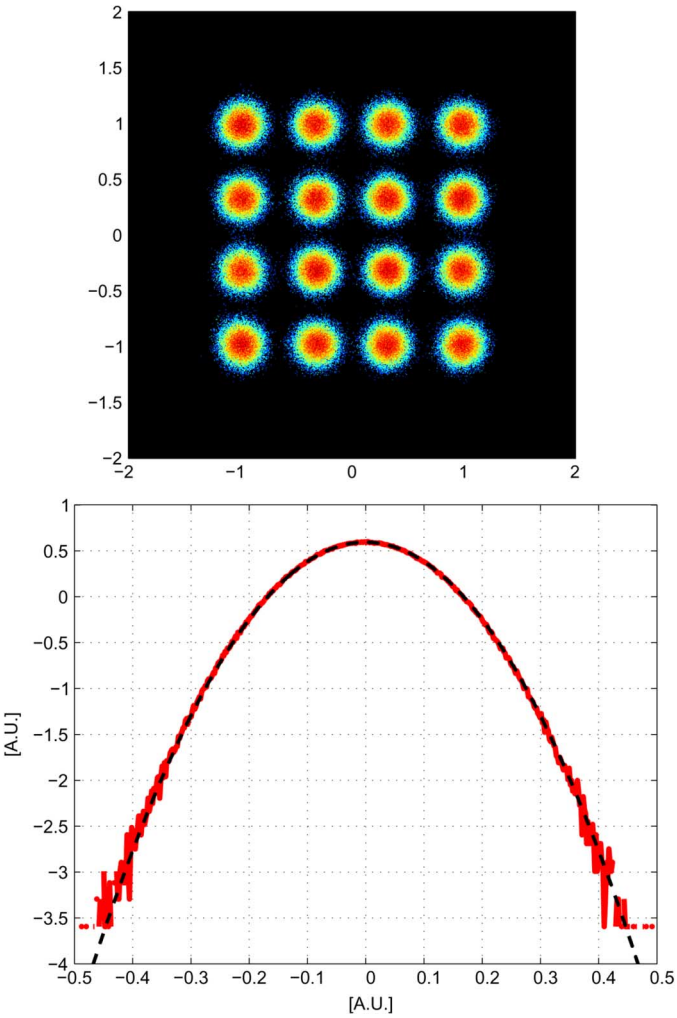

Fig. 14. Same as Fig. 13, except here ASE noise is present and is injected at each amplifier.

for completeness. As mentioned, both cases were simulated at the optimum launch power for the case when ASE is present. We have also conducted similar Gaussianity tests at 1 dB greater launch power than the optimum and found a similar degree of Gaussianity.

Based on the these results, we conclude that the Gaussian assumption on NLI noise appears to be verified to a very substantial extent on higher-order formats as well, at least up to PM-16QAM.

As a final comment, it is possible that the shown distributions would not pass extremely stringent Gaussianity tests, i.e., when looking at the far tails of the distribution. However, modern systems operate at high BERs, and the trend is actually towards increasingly higher ones, even greater than 10^{-2}. At such high BERs, Gaussianity to the extreme tails of the distribution becomes irrelevant in practical terms.

APPENDIX D
THE GN MODEL DERIVATION

The spectral-line modeling of the signal through (13) allows us to carry out NLI noise estimation using the standard FWM formulas [25]. Note, however, that we do not compute FWM in the customary sense of the terminology, i.e., assuming a *single* spectral line per channel. Rather, each channel combines with the others, and with itself, through the nonlinear beating of *many* spectral lines per channel, in fact arbitrarily many, since the longer T_0, the smaller f_0 and the greater the number of spectral lines per channel.

To evaluate such nonlinear FWM-like beating, we use the "undepleted pump" assumption, which corresponds to assuming operation in low-to-moderate nonlinearity. Specifically, we first solve for the *linear* propagation of the signal, which simply yields

$$E_{\mathrm{GN}}(f,z) = E_{\mathrm{GN}}(f) e^{-j2\pi^2\beta_2 f^2 z} e^{-\alpha z} \tag{39}$$

where $E_{\mathrm{GN}}(f)$ is the Tx signal model given by (13). The symbol α is fiber loss, β_2 is fiber dispersion and z is propagation distance into the fiber. We then use $E_{\mathrm{GN}}(f,z)$ to force the generation of the nonlinear FWM-like beating along the fiber.

In this paper we neglect the frequency dependence of β_2. The model can conceivably be extended to account for it, at least through the inclusion of the frequency derivative of β_2, generally indicated as β_3. We leave this aspect for future investigation.

As a result of the nonlinear FWM-like beating calculation, assuming for now single-polarization, we obtain the NLI field generated within a *single span* (ss). Its Fourier transform has the form

$$E_{\mathrm{NLI,ss}}(f) = \sum_{i=-\infty}^{\infty} \mathcal{E}_{\mathrm{NLI,ss},i}\, \delta(f - i f_0) \tag{40}$$

where $\mathcal{E}_{\mathrm{NLI,ss},i}$ is the complex coefficient of the i-th NLI frequency line, centered at (if_0), present in $E_{\mathrm{NLI,ss}}(f)$. These coefficients are given by the following formula:

$$\begin{aligned}\mathcal{E}_{\mathrm{NLI,ss},i} &= -j\gamma f_0^{\frac{3}{2}} e^{-j2\pi^2\beta_2 i^2 f_0^2 N_s L_s} \\ &\cdot \sum_{m,n,k\in \mathrm{A}_i} \xi_m \xi_n^* \xi_k \sqrt{G_{\mathrm{Tx}}(mf_0) G_{\mathrm{Tx}}(nf_0) G_{\mathrm{Tx}}(kf_0)} \\ &\cdot \frac{1 - e^{-2\alpha L_s} e^{j4\pi^2\beta_2 f_0^2 (k-n)(m-n) L_s}}{2\alpha - j4\pi^2\beta_2 f_0^2 (k-n)(m-n)}\end{aligned} \tag{41}$$

where γ is the fiber nonlinearity coefficient and L_s is the span length. The symbol A_i represents the set of all triples (m,n,k) for which $m - n + k = i$, *except* those triples for which $m = n$ or those for which $k = n$. The triples left out of A_i are those generating XPM-like and SPM-like terms. Note once more that these are neither XPM nor SPM in the customary sense of the acronyms. They are only mathematically akin to them. The XPM-like term can be removed because their aggregated impact amounts to a fixed, frequency-independent phase-shift, proportional to the total WDM average signal power, and hence irrelevant. As for the SPM-like terms, their contribution vanishes as $f_0 \to 0$.

The NLI power $P_{\mathrm{NLI,ss},i}$ impinging over each single frequency (if_0) due to a single span is then found by absolute-value squaring of the corresponding coefficient $\mathcal{E}_{\mathrm{NLI,ss},i}$

$$P_{\mathrm{NLI,ss},i} = \mathbf{E}\{|\mathcal{E}_{\mathrm{NLI,ss},i}|^2\} \tag{42}$$

where the symbol $\mathbf{E}$ denotes the statistical expectation operator. Statistical averaging is necessary because the $\mathcal{E}_{\mathrm{NLI,ss},i}$'s are RVs, since they contain the ξ_i's deriving from (13).

Before proceeding further, we address the polarization aspect. So far we have assumed single-polarization. All results can be readily recalculated in dual-polarization by using the Manakov equation [26]. Remarkably, under the assumption that the transmitted field on one polarization is statistically independent of the field transmitted on the other, and that the total transmitted power is evenly split between polarizations, accounting for dual-polarization only generates a constant multiplying factor equal to 8/27. As a result, for dual-polarization we simply have

$$P_{\mathrm{NLI,ss},i} = \frac{8}{27}\mathbf{E}\left\{|\mathcal{E}_{\mathrm{NLI,ss},i}|^2\right\}. \tag{43}$$

Note that we employ the definition of the fiber nonlinearity coefficient as in [26], that is: $\gamma = n_2\, k_0$, where n_2 is the Kerr nonlinearity coefficient and k_0 is the wavenumber. As a consequence, we used the Manakov equation with factor 8/9 multiplying the Kerr nonlinearity term, as shown in [26]. Other authors define $\gamma = 8/9 \cdot n_2\, k_0$ and drop the 8/9 factor from the Manakov equation [18]. According to this latter definition of γ, the factor appearing in (43) should then be 3/8 rather than 8/27. In this paper we adopt the conventions of [26].

Incidentally, it turns out that the contribution to $P_{\mathrm{NLI,ss},i}$ due to same-polarization NLI is 2/3 of the total whereas the contribution of cross-polarization NLI is 1/3. In other words, same-polarization NLI is twice as powerful as cross-polarization NLI. Apart from this factor of scale, the two NLI contributions are exactly alike and their PSD is also identical.

The squaring process in (43) gives rise, for each frequency (if_0), to many cross-products of different NLI contributions. The majority of these cross-products is however eliminated by the statistical averaging, thanks to the independence of the ξ_i's. Eventually, for each frequency (if_0), the total NLI power $P_{\mathrm{NLI,ss},i}$ can be expressed as a double summation over discrete indices of many NLI contributions

$$\begin{aligned} P_{\mathrm{NLI,ss},i} = {} & \frac{16}{27}\gamma^2 f_0^3 e^{-2\alpha L_s} \\ & \cdot \sum_m \sum_k \left| \frac{1 - e^{-2\alpha L_s} e^{j4\pi^2\beta_2 f_0^2 (k-i)(m-i) L_s}}{2\alpha - j4\pi^2\beta_2 f_0^2 (k-i)(m-i)} \right|^2 \\ & \cdot G_{\mathrm{Tx}}(mf_0) G_{\mathrm{Tx}}(kf_0) G_{\mathrm{Tx}}([m+k-i]f_0). \end{aligned} \tag{44}$$

By letting $T_0 \to \infty$, i.e., by letting the periodicity of the PWGN noise representing the signal in (13) go to infinity, the summations in (44) can be converted into integrals and a frequency-continuous version of (44) can be obtained. The result is shown as (15).

A. Coherent NLI Noise Sum

As mentioned in Section II.C, one way of dealing with the sum at the Rx of the single-span NLI noise contributions, each of the form of (15), is to add them up *coherently*. This means that the correct phase relationships accumulated during propagation must be accurately taken into account. The sum therefore cannot be carried out in power, but must be carried out over the NLI *field* contributions propagated to the end of the link.

Therefore, one must first derive the total NLI *field* at the end of the link. Its general form is similar to (40)

$$E_{\mathrm{NLI}}(f) = \sum_{i=-\infty}^{\infty} \mathcal{E}_{\mathrm{NLI},i}\delta(f - if_0) \tag{45}$$

with coefficients $\mathcal{E}_{\mathrm{NLI},i}$. These coefficients are now given by

$$\begin{aligned} \mathcal{E}_{\mathrm{NLI},i} = {} & -j\gamma f_0^{\frac{3}{2}} e^{-j2\pi^2\beta_2 i^2 f_0^2 N_s L_s} \\ & \cdot \sum_{m,n,k\in\mathrm{A}_i} \xi_m \xi_n^* \xi_k \sqrt{G_{\mathrm{Tx}}(mf_0) G_{\mathrm{Tx}}(nf_0) G_{\mathrm{Tx}}(kf_0)} \\ & \cdot \frac{1 - e^{-2\alpha L_s} e^{j4\pi^2\beta_2 f_0^2 (k-n)(m-n) L_s}}{2\alpha - j4\pi^2\beta_2 f_0^2 (k-n)(m-n)} \\ & \times \frac{\sin\left(2\pi^2\beta_2 f_0^2 (k-n)(m-n) N_s L_s\right)}{\sin\left(2\pi^2\beta_2 f_0^2 (k-n)(m-n) L_s\right)} \\ & \cdot e^{-j2\pi^2\beta_2 f_0^2 (k-n)(m-n)(N_s-1)L_s}. \end{aligned} \tag{46}$$

Comparing (41) with (46), the difference resides in the presence in (46) of a $[\sin(N_s\phi)/\sin(\phi)]$ factor, as well as of a following phase shift factor. These two factors account for the coherently interfering sum at the end of the link of the NLI contributions created within the different spans. This coherent beating is influenced by both the accumulated dispersive phase shift along the link of the signal field of (39), which generates the NLI contributions, as well as by the phase shift due to the linear propagation of NLI from the span where it is generated till the Rx. Such interference effect, with similar analytical form, was first pointed out in the context of conventional FWM calculations [25], [27]. It has also been stressed in various FWM-like nonlinear propagation models developed for OFDM systems [16], [17].

Subsequent squaring and statistical averaging, according to (43), followed by letting $T_0 \to \infty$, leads to the result shown as (18).

ACKNOWLEDGMENT

The authors thank RSoft Design Group Inc. for supplying the optical transmission system simulation software OptSim™.

REFERENCES

[1] J.-X. Cai *et al.*, "112 × 112 Gb/s transmission over 9 360 km with channel spacing set to the baud rate (360% spectral efficiency)," in *Proc. ECOC 2010*, Torino, Italy, Sep. 2010, paper PD2.1.

[2] M. Salsi *et al.*, "Transmission of 96 × 100 Gb/s with 23 super-FEC overhead over 11 680 km, using optical spectral engineering," in *Proc. OFC 2011*, Los Angeles, CA, Mar. 2011, paper OMR2.

[3] S. Yamanaka *et al.*, "11 × 171 Gb/s PDM 16-QAM transmission over 1440 km with a spectral efficiency of 6.4 b/s/Hz using high-speed DAC," in *Proc. ECOC 2010*, Torino, Italy, Sep. 2010, paper We.8.C.1.

[4] A. Sano *et al.*, "100 × 120 Gb/s PDM 64-QAM transmission over 160 km using linewidth-tolerant pilotless digital coherent receiver," in *Proc. ECOC 2010*, Torino, Italy, Sep. 2010, paper PD2.4.

[5] D. Qian *et al.*, "101.7 Tb/s (370 × 294-Gb/s) PDM-128QAM-OFDM transmission over 3 × 55-km SSMF using pilot-based phase noise mitigation," in *Proc. OFC 2011*, Los Angeles, CA, Mar. 2011, paper PDPB5.

[6] V. Curri, P. Poggiolini, A. Carena, and F. Forghieri, "Dispersion compensation and mitigation of non-linear effects in 111 Gb/s WDM coherent PM-QPSK systems," *IEEE Photon. Technol. Lett.*, vol. 20, no. 7, pp. 1473–1475, Sep. 1, 2008.

[7] D. van den Borne, V. Sleiffer, M. Alfiad, S. Jansen, and T. Wuth, "POLMUX-QPSK modulation and coherent detection: The challenge of long-haul 100G transmission," in *Proc. of ECOC 2009*, Vienna, Austria, Sep. 2009, paper 3.4.1.

[8] G. Gavioli *et al.*, "NRZ-PM-QPSK 16 × 100 Gb/s transmission over installed fiber with different dispersion maps," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 371–373, Mar. 2010.

[9] A. Carena *et al.*, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," in *Proc. of ECOC 2010*, Torino, Italy, Sep. 2010, paper P4.07.

[10] F. Vacondio, C. Simonneau, L. Lorcy, J. C. Antona, A. Bononi, and S. Bigo, "Experimental characterization of Gaussian-distributed nonlinear distortions," in *Proc. of ECOC 2011*, Geneva, Switzerland, Sep. 2011, paper We.7.B.1.

[11] G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, "Performance prediction for WDM PM-QPSK transmission over uncompensated links," in *Proc. OFC 2011*, Los Angeles, CA, Mar. 2011, paper OThO7.

[12] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed non-linear noise," *Opt. Exp.*, vol. 19, pp. 12781–12788, 2011.

[13] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *J. Lightw. Technol.*, vol. 15, no. 12, pp. 2232–2241, Dec. 1997.

[14] J. Tang, "The channel capacity of a multispan DWDM system employing dispersive nonlinear optical fibers and an ideal coherent optical receiver," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1095–1101, Jul. 2002.

[15] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.

[16] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. 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, pp. 15778–15810, 2008.

[17] 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, 2010.

[18] W. Shieh and X. Chen, "Information spectral efficiency and launch power density limits due to fiber nonlinearity for coherent optical OFDM systems," *IEEE Photon. J.*, vol. 3, no. 2, pp. 158–173, Apr. 2011.

[19] A. Splett, C. Kurzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. ECOC 1993*, 1993, vol. 2, pp. 41–44.

[20] H. Louchet *et al.*, "Analytical model for the performance evaluation of DWDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 15, no. 9, pp. 1219–1221, Sep. 2003.

[21] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of non-linear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 1, 2011.

[22] J. G. Proakis, *Digital Communication*. New York: McGraw-Hill, 1989.

[23] Y. Han and G. Li, "Coherent optical communication using polarization multiple-input-multiple-output," *Opt. Exp.*, vol. 13, pp. 7527–7534, 2005.

[24] E. Torrengo, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, and F. Forghieri, "Experimental validation of an analytical model for nonlinear propagation in uncompensated optical links," *Opt. Exp.*, vol. 19, no. 26, pp. B790–B798, Dec. 2011.

[25] K. Inoue and H. Toba, "Fiber four-wave mixing in multi-amplifier systems with nonuniform chromatic dispersion," *J. Lightw. Technol.*, vol. 13, no. 1, pp. 88–93, Jan. 1995.

[26] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[27] W. Zeiler, F. Di Pasquale, P. Bayvel, and J. E. Midwinter, "Modeling of four-wave mixing and gain peaking in amplified WDM optical communication systems and networks," *J. Lightw. Technol.*, vol. 14, no. 9, pp. 1933–1942, Sep. 1996.

Author biographies not included at authors' request due to space constraints.