Modeling nonlinearity in coherent transmissions with dominant intrachannel-four-wave-mixing

A. Bononi,[1,*] P. Serena,[1] N. Rossi,[1] E. Grellier,[2] and F. Vacondio[2]

[1]*Dip. Ing. informazione, Università di Parma, Italy*
[2]*Alcatel-Lucent Bell Labs, Villarceaux, France*
[*] *alberto.bononi@unipr.it*

Abstract: By extending a well-established time-domain perturbation approach to dual-polarization propagation, we provide an analytical framework to predict the nonlinear interference (NLI) variance, i.e., the variance induced by nonlinearity on the sampled field, and the nonlinear threshold (NLT) in coherent transmissions with dominant intrachannel-four-wave-mixing (IFWM). Such a framework applies to non dispersion managed (NDM) very long-haul coherent optical systems at nowadays typical baudrates of tens of Gigabaud, as well as to dispersion-managed (DM) systems at even higher baudrates, whenever IFWM is not removed by nonlinear equalization and is thus the dominant nonlinearity. The NLI variance formula has two fitting parameters which can be calibrated from simulations. From the NLI variance formula, analytical expressions of the NLT for both DM and NDM systems are derived and checked against recent NLT Monte-Carlo simulations.

OCIS codes: (060.1660) Coherent communications, (060.4370) Nonlinear optics, fibers.

References and links

1. A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. T. Taiba, and F. Forghieri, "Statistical Characterization of PM-QPSK Signals after Propagation in Uncompensated Fiber Links," Proc. ECOC'10, paper P4.07 (2010).
2. P. Ramantanis and Y. Frignac, "Pattern-dependent nonlinear impairments on QPSK signals in dispersion-managed optical transmission systems," Proc. ECOC'10, paper Mo.1.C.4 (2010).
3. G. Bosco, A. Carena, R. Cigliutti, V. Curri, P. Poggiolini, and F. Forghieri, "Performance Prediction for WDM PM-QPSK Transmission over Uncompensated Links," Proc. OFC'11, paper OThO7 (2011).
4. 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. **23**, 742–744 (2011).
5. E. Grellier and A. Bononi, "Quality Parameter for Coherent Transmissions with Gaussian-distributed Nonlinear Noise," Opt. Express **19**, 12781–12788 (2011).
6. F. Vacondio, C. Simonneau, L. Lorcy, J.-C. Antona, A. Bononi, and S. Bigo, "Experimental characterization of Gaussian-distributed nonlinear distortions," Proc. ECOC'11, paper We.7.B.1 (2011).
7. F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, and S. Bigo, "On nonlinear distortions of coherent systems," Opt. Express (to be published).
8. 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," Proc. ECOC'11, paper We.7.B.2 (2011).
9. A. Bononi, N. Rossi, and P. Serena, "Transmission Limitations due to Fiber Nonlinearity," Proc. OFC'11, paper OWO7 (2011).
10. E. Ip and J. M. Kahn, "Compensation of Dispersion and Nonlinear Impairments Using Digital Backpropagation," J. Lightwave Technol. **26**, 3416–3425 (2008).

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

11. D. S. Millar, S. Makovejs, C. Behrens, S. Hellerbrand, R. I. Killey, P. Bayvel, and S. Savory, "Mitigation of Fiber Nonlinearity Using a Digital Coherent Receiver," IEEE J. Sel. Top. Quantum Electron. **16**, 1217–1226 (2010).
12. K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra Series Approach for Optimizing Fiber-Optic Communications Systems Design," J. Lightwave Technol. **16**, 2046–2055 (1998).
13. Y. Gao, F. Zhang, L. Dou, Z. Chen, and A. Xu, "Intra-channel nonlinearities mitigation in pseudo-linear coherent QPSK transmission systems via nonlinear electrical equalizer," Opt. Commun. **282**, 2421–2425 (2009).
14. A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of Intrachannel Nonlinear Effects in Highly Dispersed Optical Pulse Transmission," IEEE Photon. Technol. Lett. **12**, 392–394 (2000).
15. A. Mecozzi, C. B. Clausen, and M. Shtaif, "System Impact of Intra-Channel Nonlinear Effects in Highly Dispersed Optical Pulse Transmission," IEEE Photon. Technol. Lett. **12**, 1633–1635 (2000).
16. X. Wei, "Power-weighted dispersion distribution function for characterizing nonlinear properties of long-haul optical transmission links," Opt. Lett. **31**, 2544–2546 (2006).
17. A. Bononi, P. Serena, and M. Bertolini, "Unified Analysis of Weakly-Nonlinear Dispersion-Managed Optical Transmission Systems from Perturbative Approach," C. R. Phys. **9**, 947–962 (2008).
18. A. Bononi, P. Serena, and N. Rossi, "Modeling Nonlinearity in Coherent Transmissions with Dominant Interpulse-Four-Wave-Mixing," Proc. ECOC'11, paper We.7.B.4 (2011).
19. For interleaved RZ (iRZ) we would need two different support pulses for each polarization, so here iRZ is excluded.
20. A. Bononi, P. Serena, and A. Orlandini, "A Unified Design Framework for Single-Channel Dispersion-Managed Terrestrial Systems," J. Lightwave Technol. **26**, 3617–3631 (2008).
21. J. C. Antona and S. Bigo, "Physical design and performance estimation of heterogeneous optical transmission systems," C. R. Phys. **9**, 963–984 (2008).
22. G. A. Korn and T. A. Korn, *Mathematical Handbook for Scientists and Engineers* (Dover, 2000).
23. C. R. Menyuk and B. Marks, "Interaction of Polarization Mode Dispersion and Nonlinearity in Optical Fiber Transmission Systems," J. Lightwave Technol. **24**, 2806–2826 (2006).

1. Introduction

It has recently been shown that, in high bit-rate coherent optical links with no dispersion management (NDM), the nonlinear interference (NLI) is a zero-mean signal-independent additive circular complex-Gaussian noise already after a few spans [1, 2]. Based on such a key observation, a nonlinear Gaussian model for NDM coherent communications has been proposed [3–5] and experimentally validated [6–8]. In wavelength division multiplexing (WDM), the nonlinear noise comes both from intrachannel nonlinearity and from interchannel nonlinearity. As we increase propagation distance, intrachannel nonlinearity eventually becomes dominant at baudrates in the range of tens of Gigabaud, such as those typically envisaged for modern coherent optical communications [9], unless some form of nonlinear equalization is employed [10–13]. However up to date such nonlinear equalization techniques at long distance and high baudrate have unmanageable hardware implementation complexity.

Building on a well-established time-domain perturbation approach [14–17], in this paper we extend the study of the nonlinear Gaussian model to the regime in which single-channel intrachannel four wave mixing (IFWM) is the dominant nonlinearity. Such a regime applies to both NDM and dispersion-managed (DM) long links at sufficiently large baudrates.

The paper, which is an extended version of [18], is organized as follows. Section 2 introduces the basics of the nonlinear Gaussian model, in which a single parameter, the NLI parameter a_{NL}, completely determines nonlinearity. Section 3 proves that, within the applicability of the nonlinear Gaussian model, we can analytically derive the nonlinear threshold (NLT) at a target bit error rate (BER). Section 4 derives explicit analytical expressions of a_{NL} for dual-polarization modulation both for NDM and for DM links under the assumption that IFWM is dominant. Section 5 finally compares the analytical NLT to the NLT derived by time-consuming Monte-Carlo simulations in [9]. The five appendices contain all the analytical derivations needed to support the results summarized in the main body of the paper. Table 1 provides a list of the main symbols used in the paper.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

Table 1. List of Main Symbols Used in the Paper

Symbol	*Unit*	*Definition*	*Symbol*	*Unit*	*Definition*
P	[mW]	average signal power	$\hat{P}_{NLT}$	[mW]	power at max SNR S_0
N_A	[mW]	ASE power	$\hat{P}_1$	[mW]	power at 1 dB SNR penalty
N_{NL}	[mW]	NLI power	Φ_{NL}	[rad]	nonlinear phase
a_{NL}	[mW^{-2}]	NLI parameter	$G(z)$		power gain from 0 to z
N		number of spans	$\eta(t_1t_2)$		time-kernel
z_A	[km]	span length	ξ		normalized dispersion
L	[km]	total link length	τ_M		time-kernel effective duration
α	[1/km]	attenuation coefficient	d_f		degeneracy factor
γ	[1/W/km]	nonlinear coefficient	$\mathscr{S}$		map strength = $-\frac{\beta_2}{\alpha}R^2$
β_2	[ps^2/km]	dispersion coefficient	η_p		polarization fitting factor
R	[Gbaud]	symbol rate	μ		τ_M fitting factor

2. Nonlinear Gaussian Model

Consider a single-channel long-haul optical link with dual polarization (DP) coherent reception. Assume that both the amplified spontaneous emission (ASE) and the NLI are independent additive complex-Gaussian noises. After coherent reception with polarization demultiplexing and ideal linear electrical equalization, followed by matched filtering and ideal carrier estimation, the 2-dimensional sampled received complex field vector at sampling time t is: $\boldsymbol{r}(t) = \sqrt{P}\boldsymbol{U}(t) + \boldsymbol{n}_L(t) + \boldsymbol{n}_{NL}(t)$, where P [W] is the average signal power, $\boldsymbol{U}$ the normalized 2x1 complex signal vector, $\boldsymbol{n}_L$ the ASE, and $\boldsymbol{n}_{NL}$ the NLI vectors. The electrical signal-noise ratio (SNR) at the decision gate is

$$S = \frac{P}{N_A + N_{NL}} \tag{1}$$

where: $N_A = E[|\boldsymbol{n}_L|^2] = \beta N$ is the ASE power from both polarizations, where N is the number of spans and $\beta = h\nu F(\mathscr{G}-1)$ depends on the in-line amplifiers noise figure F and gain $\mathscr{G} = \mathrm{e}^{\alpha z_A}$, being α the fiber loss coefficient and z_A the span length; $N_{NL} = E[|\boldsymbol{n}_{NL}|^2] = a_{NL}P^3$ is the NLI power obtained from a first-order regular perturbation [3, 5]. The final relationship between Q-factor and SNR then depends on the modulation format [5, 6, 8].

The main goal of this paper is to provide an approximate analytical expression of the NLI coefficient a_{NL}, valid for dominant IFWM, in any DM or NDM link. Such an expression will be used to analytically cross-validate recent simulation results on nonlinear threshold in DP coherent transmissions [9].

3. Nonlinear Threshold

We define the *constrained* NLT at reference bit error rate BER_0 (i.e., at its corresponding format-dependent SNR S_0) as the transmitted power $\hat{P}_{NLT}$ yielding the maximum of the "bell-curve" S versus P, where the maximum value is *constrained* to S_0. Maximization of Eq. (1) with ASE noise adjusted such that the top value is $S = S_0$ yields [3]

$$\hat{P}_{NLT} = \frac{1}{(3S_0a_{NL})^{1/2}} \tag{2}$$

and depends only on S_0 and a_{NL}. It has been shown that the model [Eq. (1)], at the top S value, yields an SNR penalty with respect to linear propagation of 1.76 dB [3, 4]. Appendix 1 reviews such results and extends them to prove that the 1dB NLT $\hat{P}_1$, i.e., the transmitted power needed

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

to achieve S_0 with 1 dB of SNR penalty, is ~ 1.05 dB smaller than $\hat{P}_{NLT}$. $\hat{P}_1$ corresponds to the NLT simulated in [9] that we wish to double-check with our theory.

4. Nonlinear Interference coefficient

We now describe a procedure to derive closed-form analytical expressions of the NLI coefficient a_{NL}. Generalizing the work in [16, 17] to DP, Appendix 2 shows that, in absence of polarization mode-dispersion (PMD), the NLI field vector can be obtained from a first-order regular perturbation (RP1) as:

$$\boldsymbol{n}_{NL}(t) = -j\sqrt{P}\Phi_{NL}(P)\iint_{-\infty}^{\infty}\eta(t_1t_2)\boldsymbol{U}(t+t_1)\boldsymbol{U}^{\dagger}(t+t_1+t_2)\boldsymbol{U}(t+t_2)dt_1dt_2 \tag{3}$$

where: the nonlinear phase is $\Phi_{NL}(P) \triangleq PL < \gamma G >$ with L the total link length, $\gamma(s)$ 8/9 times the fiber nonlinear coefficient, $G(s)$ the power gain at coordinate s, and the average $< \gamma G >$ is defined in Eq. (23); $\eta(t_1t_2)$ is the time-kernel (where time is normalized to the symbol time $T = 1/R$, and R is the baudrate), which is the 2D-inverse Fourier transform of the frequency-kernel $\tilde{\eta}(\omega_1\omega_2)$ given explicitly in Appendix 2; $\boldsymbol{U}(t)$ is the received unit-power-normalized desired signal field, and † stands for transpose conjugate.

For a linear digital modulation we have $\boldsymbol{U}(t) = \sum_{k=-\infty}^{\infty}\boldsymbol{s}_k p(t-k)$, where $\boldsymbol{s}_k = [X_k, Y_k]^T$ is the vector of constellation symbols on polarizations X and Y at time k (^{T} stands for transpose), and $p(t)$ is the real, scalar common supporting pulse [19]. We are interested in the NLI at the sampling time of interest, say $t = 0$. We assume a Nyquist pulse $p(0) = 1$ and $p(k) = 0$ at any other integer k. Thus assuming uncorrelated symbols we have $E[|\boldsymbol{U}(0)|^2] = E[|\boldsymbol{s}_k|^2] = 1$. When the time-kernel is much broader than the symbol time, a regime we call IFWM dominated, then we can approximate the supporting pulse with a delta function in each field term in the double integral of Eq. (3), and the NLI term simplifies to $\boldsymbol{n}_{NL}(0) = \boldsymbol{c}_{NL}P^{3/2}$, with [16, 17]

$$\boldsymbol{c}_{NL} = -jL < \gamma G > \sum_{m,n,l}\boldsymbol{s}_m\boldsymbol{s}_l^{\dagger}\boldsymbol{s}_n\eta\left((m-l)(n-l)\right) \tag{4}$$

where the sum accounts for IFWM terms, i.e., runs over all integers m, n, l such that $m+n-l = t = 0$, with $m \neq l$, $n \neq l$. The above expression does not apply to intrachannel cross-phase modulation (IXPM) ($m = l$ or $n = l$) or pure self-phase modulation (SPM) ($m = n = 0$); both such terms however tend to give a negligible contribution to the overall NL power with respect to IFWM as the time-kernel gets broader and broader, i.e., when IFWM is dominant. The NL power is $P_{NL} \triangleq E[|\boldsymbol{n}_{NL}(0)|^2] = E[|\boldsymbol{c}_{NL}|^2]P^3$, hence we recognize that $a_{NL} \equiv E[|\boldsymbol{c}_{NL}|^2]$, where the expectation is taken over the random symbols. Appendix 3 shows that, for any DP constellation with $E[\boldsymbol{s}_k] = 0$ and $E[|\boldsymbol{s}_k|^2] = 1$, we get when IFWM is dominant

$$a_{NL} = \eta_p 8(L < \gamma G >)^2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|\eta(mn)|^2 \tag{5}$$

where $\eta_p = \frac{3}{8}$ for DP, and $\eta_p = 1$ for SP transmission.

Expression (5) is simple, yet it requires the explicit evaluation of the time-kernel, which is analytically known only for lossless links [16] or for a single lossy span of infinite length [20]. For practical lossy links of interest the time-kernel evaluation is a challenging numerical problem, and in this paper we seek an alternative procedure able to avoid its direct numerical computation. The idea is the following. We first approximate the double sum as

$$A_{lim} \triangleq \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|\eta(mn)|^2 \cong \int_1^{\infty}\int_1^{\infty}|\eta(t_1t_2)|^2dt_1dt_2 = \int_1^{\infty}\ln(\tau)|\eta(\tau)|^2d\tau. \tag{6}$$

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

Since the time-kernel magnitude decreases for increasing τ and eventually vanishes after an effective time duration τ_M, we may then upper-bound the double sum as

$$A_{lim} \cong \int_1^{\tau_M} \ln(\tau)|\eta(\tau)|^2 d\tau \leq \ln(\tau_M) \int_1^{\tau_M} |\eta(\tau)|^2 d\tau \leq \ln(\tau_M) \int_0^{\infty} |\eta(\tau)|^2 d\tau. \quad (7)$$

What we need are expressions of both the kernel duration τ_M and of the above integral of the kernel magnitude that do not need the explicit time-kernel evaluation. We may choose $\tau_M \triangleq \mu\tau_{rms}$ for some positive multiplier μ of the rms width $\tau_{rms}^2 = \int_{-\infty}^{\infty} \tau^2 |\eta(\tau)|^2 \mathrm{d}\tau / \int_{-\infty}^{\infty} |\eta(u)|^2 \mathrm{d}u$. If the parameter μ is chosen too large such that $\mu\tau_{rms}$ exceeds the actual time-kernel duration, then we just make the upper-bound of Eq. (7) looser. We will discuss the choice of parameter μ in the results Section. Now, for every optical link, both with and without dispersion management, a physically meaningful function is the *power-weighted dispersion distribution* (PWDD) $J(c)$, representing signal power versus cumulated dispersion c [16]. Appendix 4 shows that

$$DEN \triangleq \int_{-\infty}^{\infty} |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} J^2(c) \frac{dc}{2\pi} \quad (8)$$

$$NUM \triangleq \int_{-\infty}^{\infty} \tau^2 |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} c^2 [J(c) + cJ'(c)]^2 \frac{dc}{2\pi}. \quad (9)$$

Thus a_{NL} in Eq. (5) can be upper-bounded by the following expression depending solely on integrals of $J(c)$, which are easy to evaluate for practical links:

$$a_{NL} \leq \eta_p 8(L < \gamma G >)^2 \frac{DEN}{2} \ln(\mu \sqrt{\frac{NUM}{DEN}}). \quad (10)$$

The derivation in Appendix 3 clearly shows that this bound is valid for *any* zero-mean DP modulation format with independent polarization tributaries, at a given baudrate. In other terms, in the IFWM dominated regime both constant amplitude formats such as quadrature phase shift keying (QPSK), and variable amplitude formats such as quadrature amplitude modulation (QAM) at the same average power P do generate the same nonlinear power $a_{NL}P^3$. Appendix 5 derives closed-form expressions of the a_{NL} upper-bound of Eq. (10) for several links of interest. For instance, for NDM links we obtain for $N \gtrsim 5$:

$$a_{NL} \leq \eta_P (\frac{\gamma}{\alpha})^2 \frac{N}{\pi |\mathscr{S}|} \ln(\frac{4\mu}{\sqrt{5}} (\alpha z_A N)^2 |\mathscr{S}|) \quad (11)$$

with span length z_A, and fiber "strength" $\mathscr{S} \triangleq -\frac{\beta_2}{\alpha} R^2$ [20]. Note the similarity of this expression with that of a Nyquist-WDM NDM system derived in [4] using a frequency-domain approach. The major difference is the $N\log(kN)$ scaling law in the IFWM-dominated regime, as opposed to the simpler N scaling when presumably cross-nonlinearities dominate. The $N\log(kN)$ scaling law is instead confirmed by the amplitude variance results of Mecozzi *et al.* ([15], Eq. (4)), which are based on the same RP1 time-domain approach and large-strength assumption as in this paper, although dealing with Gaussian shaped return-to-zero on-off keying modulation.

5. Results

Figure 1 (left) shows a plot of the a_{NL} formula (10) versus number of spans N (solid), and numerically simulated values (symbols), for a single-channel 28 Gbaud DP-QPSK coherent transmission over single mode fiber (SMF, $\beta_2 = -21\, ps^2/km$, $\alpha = 0.2$ dB/km, $\gamma = 1.26\, W^{-1}km^{-1}$) for an Nx100 km link, both NDM and DM with 30 ps/nm (DM30) of residual dispersion per span (RDPS) and no pre-compensation. For the NDM link we used the formula (11), while for

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

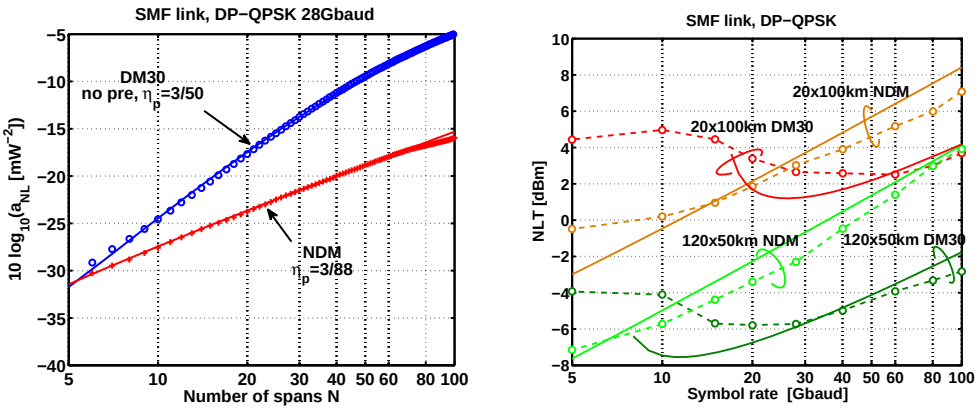

Fig. 1. **(Left)** a_{NL} [mW^{-2}] versus spans N from Eq. (10) (solid) and simulations (symbols). DP-QPSK on Nx100 km SMF links, R=28 Gbaud. **(Right)** 1dB NLT vs. symbol rate R for: theory $\hat{P}_1 = \hat{P}_{NLT} - 1.05$ dBm, with $\hat{P}_{NLT}$ as in Eq. (2) (solid lines); simulations from [9] (symbols). DM30 = DM with 30 ps/nm RDPS.

the DM case the formula (10) and Eqs. (35)-(36). In the theoretical curves we used the value $\mu = 6$, which roughly matches the actual duration $\tau_M \cong 6\tau_{rms}$ of the time-kernel and gives the best fit of the shape of a_{NL} versus N for both links, although for NDM links the dependence of a_{NL} on μ is rather weak. Instead of the theoretical DP value $\eta_p = 3/8$, a smaller fitting factor $\eta_p = 3/50$ was used for DM, and $\eta_p = 3/88$ for NDM, in order to compensate for the upper-bounding in Eq. (7). We appreciate the match of theory and simulation, as well as the announced $N\log(kN)$ scaling law in the NDM case. The perceived NDM slope over a 50 span range is ~ 1.25 dB/dB as in [1], although restricting the range to the first 15 spans gives ~ 1.35 dB/dB, as we experimentally verified in a companion study [6, 7]. NLI grows faster in the DM case: a_{NL} has an initial slope of ~ 2 dB/dB and then bends at larger N.

Figure 1 (right) shows the 1dB NLT at $BER_0 = 10^{-3}$ versus baudrate for a DP-QPSK format for both NDM, and a DM30 link with straight-line rule (SLR) pre-compensation [20], both at 20x100 km and at 120x50 km distance. Symbols refer to single-channel simulations taken from [9], solid lines to the formula $\hat{P}_1 = \hat{P}_{NLT} - 1.05$ [dBm] using Eq. (2) and the same (η_p, μ) fitting factors as in Fig. 1 (left). While for DM links theory only captures the general trend versus R with obviously major discrepancies at lower R where IFWM is not dominant, the match in NDM links (optimized at 28 Gbaud through the fitting factors η_p as in Fig. 1 (left)) seems more reasonable. Notwithstanding the numerical discrepancies observed in Fig. 1 (left), which may be large for practical design purposes, the analytical NLT curves are of great theoretical importance, as they provide a first model able to confirm the general trends of NLT observed in simulations, and quickly predict the NLT qualitative trends as we vary the main system parameters.

Of course, one may play with the two fitting parameters to improve the prediction of a_{NL} (and thus NLT) versus symbol rate. Focusing for instance on the NDM link, Fig. 2 (left) shows a_{NL} versus R for a 20x100 km link. Symbols represent simulations, while the red line the theoretical a_{NL} of Eq. (11) with the same $(\eta_p, \mu) = (\frac{3}{88}, 6)$ parameters as in Fig. 1. We can decrease the gap to simulations by using the "optimized" parameters $(\eta_p, \mu) = (\frac{3}{31}, 0.02)$ as shown by the magenta line, i.e., by pretending the time-kernel duration is smaller than its actual value. However this comes at the price of a reduced accuracy of the a_{NL} versus spans N as shown in Fig. 2 (right).

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

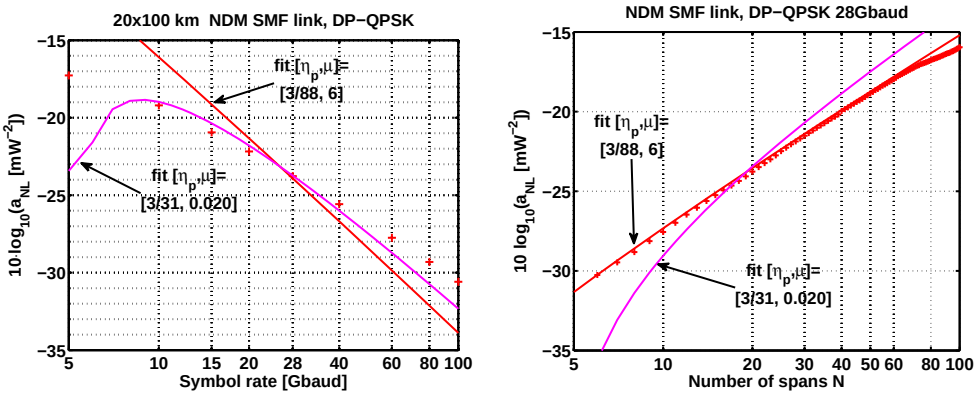

Fig. 2. **(Left)** a_{NL} versus symbol rate R for DP-QPSK 20x100 km NDM link. Symbols: simulations. Red line: theory (11) with $(\eta_p,\mu)=(\frac{3}{88},6)$ as in Fig. 1. Magenta line: theory (10) with optimized $(\eta_p,\mu)=(\frac{3}{31},0.02)$. **(Right)** a_{NL} versus spans N for 28 Gbaud DP-QPSK NDM link. Symbols: simulations. Red and Magenta lines: theory.

We verified that the main reason of the inability of the model to correctly predict the shape of a_{NL} versus symbol rate R (hence NLT versus R) over the wide range shown Figs. 1 and 2 stems from the key approximation [16]

$$\iint_{-\infty}^{\infty}\eta(t_1t_2)p(t_1-m)p(t_2-n)p(t_1+t_2-l)dt_1dt_2\simeq\eta\left((m-l)(n-l)\right)$$

used to derive Eq. (4), which requires shorter and shorter pulses $p(t)$ as $(m-l)$ and $(n-l)$ grow.

6. Conclusions

We provided a time-domain model of NLI in IFWM dominated links, which reasonably models NDM links, as well as high baudrate DM links. Such a model provides a quick qualitative tool to compare transmission link parameters in terms of their impact on received SNR. The model has two fitting parameters which may be optimized to best fit simulations, although it has difficulties in reproducing the correct behavior of a_{NL} versus symbol rate over the wide range shown in Figs. 1-2. More work is needed on this issue to improve its accuracy.

Appendix 1: NLT at fixed distance N and fixed SNR

In [3,5] it is shown that the power that maximizes the SNR, called the unconstrained NLT, is obtained when *ASE power is twice the nonlinear noise* power. Hence explicitly the unconstrained NLT is

$$P_{NLT}=\left(\frac{N_A}{2a_{NL}}\right)^{\frac{1}{3}} \tag{12}$$

and the corresponding maximum SNR value at NLT is

$$S_{NLT}\equiv\frac{P_{NLT}}{\frac{3}{2}N_A}=\left(3^3a_{NL}(\frac{N_A}{2})^2\right)^{-\frac{1}{3}} \tag{13}$$

with an SNR penalty with respect to linear propagation of $SP=10\text{Log}\frac{3}{2}\sim 1.76$ dB.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

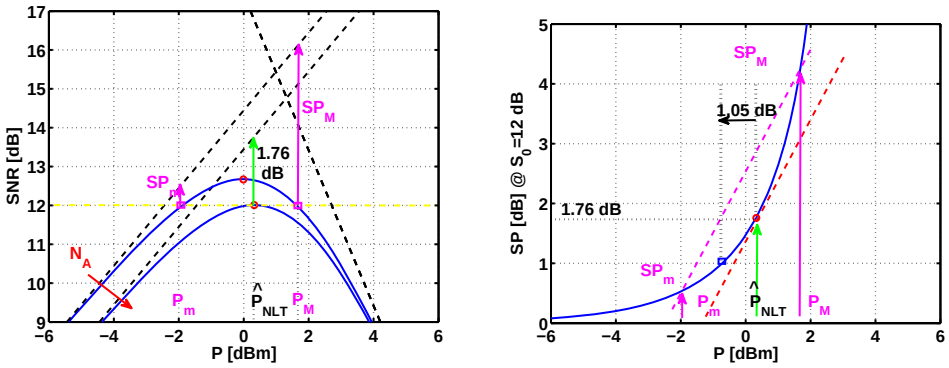

Fig. 3. (Left) example of SNR [dB] vs. P [dBm] "Bell" curve and its roots at $S_0 = 12$ dB, along with graphical definition of constrained NLT $\hat{P}_{NLT}$ at 1.76 dB of penalty; (Right) SNR penalty $SP = S^{dB} - S_0^{dB}$ at reference S_0 vs. P [dBm]. It is shown in Eq. (18) that the constrained NLT at 1dB of penalty is below $\hat{P}_{NLT}$ by 1.05 dB.

Figure 3 (left) shows an example of the "bell curve" S versus P, where a reference SNR S_0 was fixed, and the smallest and largest intersections of the SNR vs. P curve with the line at level S_0 occur at power P_m and P_M, respectively. The corresponding penalties are marked as SP_m and SP_M in the figure. The two intersections coincide at a specific value of ASE noise $\hat{N}_A$, and the corresponding power value $\hat{P}_{NLT}$ is called the NLT *at* S_0, or the *constrained NLT* [21]. Clearly, $SP = 1.76$ dB also at the constrained NLT. At $N_A > \hat{N}_A$ no intersections are found, i.e., the target SNR S_0 is unachievable. Figure 3 (right) reports the sensitivity penalty values SP_m and SP_M at their respective powers P_m and P_M as we vary the ASE noise over all achievable values $N_A \leq \hat{N}_A$. The graph in Fig. 3 (right) is routinely used in system design [21]. Since both N and S_0 are fixed, we stress that the SP vs P points are actually obtained by using varying amounts of ASE noise. For each N_A, the two corresponding (P_m, SP_m) and (P_M, SP_M) points are found at the intersection of the SP curve with the unit slope straight line $S_L^{dB} = P^{dB} - N_A^{dB}$, as shown in Fig. 3 (right).

Objective of this Appendix is to provide explicit expressions of SP_m, SP_M, $\hat{P}_{NLT}$, and the NLT $\hat{P}_1$ at $SP = 1$ dB.

i) Expressions of SP_m, SP_M and $\hat{P}_{NLT}$ at $N_A \leq \hat{N}_A$

Inverting Eq. (13) at $S_{NLT} = S_0$ we get

$$\hat{N}_A = \frac{2}{(3S_0)^{3/2} a_{NL}^{1/2}}. \tag{14}$$

From Eq. (1), P_m and P_M are seen to solve the cubic equation $P^3 - \frac{1}{S_0 a_{NL}} P + \frac{N_A}{a_{NL}} = 0$. Cardan's solutions ([22], p. 23) of the cubic equation $y^3 + py + q = 0$ are discriminated by the value of the discriminant $Q = (\frac{p}{3})^3 + (\frac{q}{2})^2$. When $Q < 0$ the cubic has 3 real roots, which can be expressed in trigonometric form as

$$y_1 = 2\sqrt{-\frac{p}{3}}\cos(\frac{\alpha}{3}) \qquad y_{2,3} = -2\sqrt{-\frac{p}{3}}\cos(\frac{\alpha}{3} \pm \frac{\pi}{3}) \tag{15}$$

with $\alpha = \arccos(-\sqrt{\frac{(q/2)^2}{-(\frac{p}{3})^3}})$. In our case $\frac{p}{3} = -\frac{1}{3S_0 a_{NL}}$, $\frac{q}{2} = \frac{N_A}{2a_{NL}}$, so $Q = -\frac{1}{(3S_0 a_{NL})^3} + \frac{(N_A/2)^2}{a_{NL}^2}$, and using Eq. (14) $Q = -\frac{(\hat{N}_A/2)^2}{a_{NL}^2} + \frac{(N_A/2)^2}{a_{NL}^2} < 0$ for all $N_A \leq \hat{N}_A$. In such a case, $\alpha =$

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

$\arccos(-N_A/\hat{N}_A)$, with $90^o < \alpha \leq 180^o$, and thus $30^o < \frac{\alpha}{3} \leq 60^o$, so that $-\cos(\frac{\alpha}{3}+60^o)$>0, i.e. y_1 is the largest positive solution, while y_2 the smallest positive solution corresponds to the + sign in Eq. (15): $P_M = \frac{2}{\sqrt{3S_0 a_{NL}}}\cos(\frac{\alpha}{3})$ and $P_m = \frac{2}{\sqrt{3S_0 a_{NL}}}\cos(\frac{2\pi-\alpha}{3})$. Using Eq. (14), we rewrite the solutions explicitly as:

$$P_M = 3S_0\hat{N}_A\cos(\frac{\arccos(-N_A/\hat{N}_A)}{3}) \qquad P_m = 3S_0\hat{N}_A\cos(\frac{2\pi-\arccos(-N_A/\hat{N}_A)}{3}). \tag{16}$$

From Eq. (1), the sensitivity penalty is the ratio of the linear SNR P/N_A and the nonlinear SNR S_0: $SP_{m,M} = \frac{P_{m,M}/N_A}{S_0}$, hence finally the sought SP values are

$$SP_M = 3\frac{\hat{N}_A}{N_A}\cos(\frac{\arccos(-N_A/\hat{N}_A)}{3}) \qquad SP_m = 3\frac{\hat{N}_A}{N_A}\cos(\frac{2\pi-\arccos(-N_A/\hat{N}_A)}{3}). \tag{17}$$

As a check, when $N_A = \hat{N}_A$ the angle $\alpha = \pi$, $\cos(\pi/3) = 1/2$ and we obtain the known value $SP_{m,M} = \frac{3}{2}$, and the NLT explicit value is $\hat{P}_{NLT} = \frac{3}{2}S_0\hat{N}_A = \frac{1}{(3S_0 a_{NL})^{1/2}}$, which can more directly be obtained by substituting Eq. (14) into Eq. (12).

ii) Expression of $\hat{P}_1$

We are now ready to answer the following question: at which power $\hat{P}_1$ does the SP w.r.t. S_0 reach a value of 1 dB? From Eq. (17), letting $SP_m = 10^{0.1} \cong 1.26$ and $x = N_A/\hat{N}_A$, we look for the solution of equation $10^{0.1} = \frac{3}{x}\cos(\frac{2\pi-\arccos(-x)}{3})$, which is $x_1 \cong 0.936$. Hence

$$\frac{\hat{P}_{NLT}}{\hat{P}_1} = \frac{\frac{3}{2}S_0\hat{N}_A}{3S_0\hat{N}_A\cos(\frac{2\pi-\arccos(-x_1)}{3})} \cong 1.273 \tag{18}$$

which means $\hat{P}_1$ is $10\log_{10}(1.273) \cong 1.05$ dB below the NLT at 1.76 dB penalty. This result is also sketched in Fig. 3 (right).

Appendix 2: Regular Perturbation Solution

We will generalize here the scalar case time-domain analysis presented in [16, 17], which extends previous analytical work based on Gaussian supporting pulses [14, 15]. The propagation equation of a DP single channel in the retarded normalized time frame t (physical time normalized to the symbol interval $T = 1/R$), can be described in absence of PMD by the Manakov-Nonlinear Schroedinger equation (M-NLSE) in engineering notation as ([23], Eq. (73))

$$\frac{\partial \boldsymbol{A}(z,t)}{\partial z} = \frac{g(z)}{2}\boldsymbol{A}(z,t) + \frac{j\beta_2(z)R^2}{2}\frac{\partial^2\boldsymbol{A}(z,t)}{\partial t^2} - j\gamma(z)|\boldsymbol{A}|^2\boldsymbol{A}(z,t) \tag{19}$$

where $\boldsymbol{A} = [A_x, A_y]^T$ is the signal field envelope on the two polarizations (in $\sqrt{W}$), $\gamma(z)$ is 8/9 times the nonlinear coefficient, $g(z)$ is the net gain/attenuation coefficient per unit length, $\beta_2(z)$ is the dispersion coefficient, and such parameters are $z-$varying functions, with span k ending at coordinate kz_A, $k = 1,\ldots,N$. The function $G(z) = e^{\int_0^z g(s)ds}$ is the power gain from 0 to z. Since the nonlinear term may be also written as $|\boldsymbol{A}|^2\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{A}^\dagger\boldsymbol{A})$, by taking the Fourier transform of Eq. (19) we get:

$$\begin{aligned}\frac{\partial\tilde{\boldsymbol{A}}(z,\omega)}{\partial z} =& \frac{g(z)-j\omega^2\beta_2(z)R^2}{2}\tilde{\boldsymbol{A}}(z,\omega) - \\ & j\gamma(z)\iint_{-\infty}^{\infty}\tilde{\boldsymbol{A}}(z,\omega+\omega_1)\tilde{\boldsymbol{A}}^\dagger(z,\omega+\omega_1+\omega_2)\tilde{\boldsymbol{A}}(z,\omega+\omega_2)\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\end{aligned} \tag{20}$$

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

where for any function $g(t)$ we define its Fourier transform (with engineering sign) as $\tilde{g}(\omega) = \int_{-\infty}^{\infty} g(t)\mathrm{e}^{-j\omega t}dt$, where $\omega = 2\pi f$, and f is the frequency normalized to the baud rate. The input modulated field $\tilde{\boldsymbol{A}}_0(\omega)$ may be pre-chirped to give $\tilde{\boldsymbol{A}}(0,\omega) \equiv \tilde{\boldsymbol{A}}_0(\omega)\mathrm{e}^{j\frac{\omega^2}{2}\xi_{pre}}$, where $\xi_{pre} \triangleq -L_{pre}\beta_{pre}R^2$ is the normalized cumulated dispersion in the pre-compensation fiber of dispersion coefficient β_{pre} and length L_{pre}. Now make the change of variable

$$\tilde{\boldsymbol{A}}(z,\omega) = \sqrt{P_0}e^{\frac{\ln G(z)+jC(z)\omega^2}{2}}\tilde{\boldsymbol{U}}(z,\omega) \tag{21}$$

where P_0 is a reference normalizing power, and $C(z) \triangleq \xi_{pre} - R^2\int_0^z \beta_2(s)$ṣ is the *normalized cumulated dispersion* up to z. Differentiating Eq. (21) and substituting into Eq. (20) one gets:

$$\begin{aligned}\frac{\partial\tilde{\boldsymbol{U}}(z,\omega)}{\partial z} =& -j\gamma(z)P_0G(z)\iint_{-\infty}^{\infty}\mathrm{e}^{-jC(z)\omega_1\omega_2}\tilde{\boldsymbol{U}}(z,\omega+\omega_1)\cdot \\ &\cdot\tilde{\boldsymbol{U}}^{\dagger}(z,\omega+\omega_1+\omega_2)\tilde{\boldsymbol{U}}(z,\omega+\omega_2)\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}.\end{aligned} \tag{22}$$

Now define the nonlinear phase (referred to nominal power P_0) as $\Phi_{NL}(P_0) \triangleq P_0L < \gamma G >$, where $L = Nz_A$ is the total link length, and

$$< \gamma G > \triangleq \frac{1}{L}\int_0^L \gamma(s)G(s)ds. \tag{23}$$

Using such definitions, multiply and divide Eq. (22) by $L < \gamma G >$, thus finally obtaining the Manakov dispersion-managed NLSE (M-DM-NLSE) in the form:

$$\begin{aligned}\frac{\partial\tilde{\boldsymbol{U}}(z,\omega)}{\partial z} =& -j\Phi_{NL}(P_0)\iint_{-\infty}^{\infty}\frac{\gamma(z)G(z)\mathrm{e}^{-jC(z)\omega_1\omega_2}}{L<\gamma G>}\tilde{\boldsymbol{U}}(z,\omega+\omega_1)\cdot \\ &\cdot\tilde{\boldsymbol{U}}^{\dagger}(z,\omega+\omega_1+\omega_2)\tilde{\boldsymbol{U}}(z,\omega+\omega_2)\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}.\end{aligned} \tag{24}$$

If the field terms in the integrand in Eq. (24) are approximated as z-independent, then Eq. (24) can be integrated on the link $[0,L]$ to yield the first-order regular perturbation (RP1) solution:

$$\begin{aligned}\tilde{\boldsymbol{U}}(L,\omega) =& \tilde{\boldsymbol{U}}(0,\omega) - j\Phi_{NL}(P_0)\iint_{-\infty}^{\infty}\frac{\int_0^L\gamma(s)G(s)\mathrm{e}^{-jC(s)\omega_1\omega_2}ds}{L<\gamma G>}\cdot \\ &\cdot\tilde{\boldsymbol{U}}(0,\omega+\omega_1)\tilde{\boldsymbol{U}}^{\dagger}(0,\omega+\omega_1+\omega_2)\tilde{\boldsymbol{U}}(0,\omega+\omega_2)\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\end{aligned}$$

with $\tilde{\boldsymbol{U}}(0,\omega) \equiv \tilde{\boldsymbol{A}}_0(\omega)/\sqrt{P_0}$ the initial condition. Define now the (scalar) *frequency-kernel* as

$$\tilde{\eta}(w) \triangleq \frac{\int_0^L\gamma(s)G(s)\mathrm{e}^{-jC(s)w}ds}{\int_0^L\gamma(s)G(s)ds} \tag{25}$$

so that the RP1 solution writes as $\tilde{\boldsymbol{U}}(L,\omega) = \tilde{\boldsymbol{U}}(0,\omega) + \tilde{\boldsymbol{U}}_{NL}(\omega)$, with

$$\tilde{\boldsymbol{U}}_{NL}(\omega) \triangleq -j\Phi_{NL}(P_0)\iint_{-\infty}^{\infty}\tilde{\eta}(\omega_1\omega_2)\tilde{\boldsymbol{U}}(0,\omega+\omega_1)\tilde{\boldsymbol{U}}^{\dagger}(0,\omega+\omega_1+\omega_2)\tilde{\boldsymbol{U}}(0,\omega+\omega_2)\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}. \tag{26}$$

If one adds at the receiver a post-compensating fiber with accumulated normalized dispersion ξ_{post}, one finally has the RP1 field at the receiver as:

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

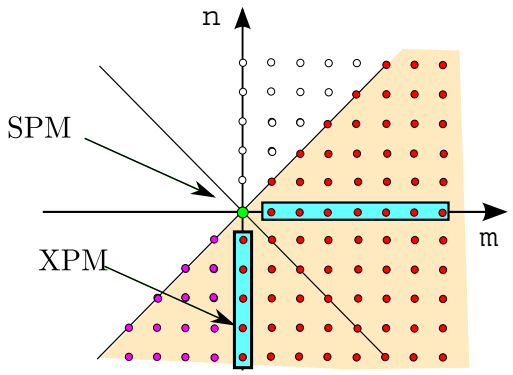

Fig. 4. Colored dots represent IFWM points on (m,n) plane when infinitely many precursors and postcursors are taken into account. They all have degeneracy factor $d_f = 2$, except those on the $m = n$ line (partially degenerate) which have degeneracy $d_f = 1$. Points on axes (IXPM) and (0,0) point (pure SPM) should not be included in IFWM count.

$$\tilde{\boldsymbol{r}}(\omega) = \sqrt{P_0} e^{\frac{\ln G(L) + j\xi_{tot}\omega^2}{2}} [\tilde{\boldsymbol{U}}(0,\omega) + \tilde{\boldsymbol{U}}_{NL}(\omega)] \tag{27}$$

where $\xi_{tot} \triangleq C(L) + \xi_{post}$. For a "power-transparent" line $G(L) = 1$, and typically for coherent systems $\xi_{tot} = 0$, so that after chromatic dispersion compensation at the receiver we have $\tilde{\boldsymbol{r}}(\omega) = \sqrt{P_0}[\tilde{\boldsymbol{U}}(0,\omega) + \tilde{\boldsymbol{U}}_{NL}(\omega)]$, and in the time domain $\boldsymbol{r}(t) = \sqrt{P_0}\boldsymbol{U}(0,t) + \boldsymbol{n}_{NL}(t)$ where $\boldsymbol{n}_{NL}(t)$ is the inverse Fourier transform of $\sqrt{P_0}\tilde{\boldsymbol{U}}_{NL}(\omega)$ and thus has the expression reported in Eq. (3). Note that the reference power P_0 can be freely chosen to simplify the analysis.

Appendix 3: Power of NL term

We need to evaluate $a_{NL} = E[|\boldsymbol{c}_{NL}|^2]$, where we rewrite $\boldsymbol{c}_{NL}$ in Eq. (4) as:

$$\begin{aligned} \boldsymbol{c}_{NL} &= -jL < \gamma G > \sum_{m,n} \boldsymbol{s}_m \boldsymbol{s}^{\dagger}_{m+n} \boldsymbol{s}_n \eta\,(mn) \\ &= -jL < \gamma G > \sum_{m,n} \begin{pmatrix} X_m (X_n X^*_{m+n} + Y_n Y^*_{m+n}) \\ Y_m (X_n X^*_{m+n} + Y_n Y^*_{m+n}) \end{pmatrix} \eta\,(nm) \end{aligned} \tag{28}$$

and the summation runs over all signed non-zero m,n integer pairs, which we visualize as points on the (m,n) plane. Each point corresponds to a pair of RVs, one per polarization, as given by the big parenthesis in Eq. (28). When we swap $m \leftrightarrow n$ the constituent random variables (RV) $X_n X_m X^*_{n+m}$ and $Y_n Y_m Y^*_{n+m}$ (Type I) remain unchanged: they represent the same RV, which in the double summation in Eq. (28) must be counted $d_f = 2$ times, and the double summation for them then runs on half the (m,n) plane, i.e., for instance on the pairs (m,n) below and on the bisectrix $m = n$ (actually the points at which $m = n$ have degeneracy $d_f = 1$, but we will disregard this subtlety for very broad time-kernels, and use $d_f = 2$ even for them), except the axes $m = 0$ and $n = 0$ which collect the IXPM and pure SPM terms. The situation for Type I RVs is summarized in Fig. 4. On the contrary, when we swap $m \leftrightarrow n$, the RVs $X_m Y_n Y^*_{m+n}$ and $Y_m X_n X^*_{m+n}$ (Type II) do change into new RVs.

We will restrict the analysis to common DP coherent modulation formats, for which $E[\boldsymbol{s}_k] = 0$. Moreover, we choose $E[|\boldsymbol{s}_k|^2] = 1$ to be the unit power of the normalized constellation symbols $\boldsymbol{s}_k = [X_k, Y_k]^T$. Symbols are assumed to be uncorrelated in time. Tributary symbols X_k, Y_k are also zero-mean uncorrelated and have the same power $E[|X_k|^2] = E[|Y_k|^2] = 1/2$. Then the

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

RV $\boldsymbol{S}_{mn} \triangleq \boldsymbol{s}_m \boldsymbol{s}_{m+n}^{\dagger} \boldsymbol{s}_n$ is zero-mean, and so is $\boldsymbol{c}_{NL}$. Therefore

$$a_{NL} = Var[\boldsymbol{c}_{NL}] = (L < \gamma G >)^2 \sum_{m,n} Var[\boldsymbol{S}_{mn}] |\eta\,(mn)\,|^2$$

since the RVs adding up to build $\boldsymbol{c}_{NL}$ are uncorrelated. Taking into account the degeneracy factor d_f we thus get

$$\frac{a_{NL}}{(L < \gamma G >)^2} = 2 \left(\sum_{\substack{(m,n):\\ m \le n}} \frac{1}{8} d_f^2 |\eta\,(mn)\,|^2 + \sum_{(m,n)} \frac{1}{8} |\eta\,(mn)\,|^2 \right)$$

where the first sum is on Type I RVs and accounts for the "self-polarization" variance, while the second sum is on Type II RVs and accounts for the variance due to cross-polarization crosstalk between X and Y. The factor 2 accounts for the contribution to NL variance from the two polarizations, and $1/8 = (E[|X|^2])^3 = E[|X|^2](E[|Y|^2])^2$ is the variance of both Types of RVs. Using the fact that the magnitude square of the kernel $|\eta\,(mn)|^2$ is the same on the 4 quadrants of the (m,n) plane, the above further simplifies to

$$\frac{a_{NL}^{DP}}{(L < \gamma G >)^2} = 2 \left(\frac{2}{8} d_f^2 + \frac{4}{8} \right) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\eta\,(mn)\,|^2 \tag{29}$$

where we added the superscript DP for clarity. The per-component a_{NL}^{pc} in DP (such that $a_{NL}^{pc} P_0^3$ is the NL variance on each component and $P_0 = P/2$ is the per-component power) is obtained using $E[|X|^2] = 1$, $E[|Y|^2] = 1$ (i.e., normalizing the M-DM-NLSE of Eq. (24) to P_0):

$$\frac{a_{NL}^{pc}}{(L < \gamma G >)^2} = \left(2 d_f^2 + 4\right) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\eta\,(mn)\,|^2. \tag{30}$$

The result for SP is obtained by using $E[|X|^2] = 1$ and keeping only the Type I RV, hence

$$\frac{a_{NL}^{SP}}{(L < \gamma G >)^2} = \left(2 d_f^2\right) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\eta\,(mn)\,|^2.$$

We clearly see from Eqs. (29) and (30) that variance coming from Type I RVs (the one present also in SP transmission) is twice that due to cross-polarization, hence $a_{NL}^{pc} = \frac{3}{2} a_{NL}^{SP}$, and therefore $a_{NL}^{DP} = \frac{3}{8} a_{NL}^{SP}$. Figure 5 shows both a_{NL}^{SP} estimated from SP transmission, and a_{NL}^{pc} from DP transmission of a single QPSK modulated channel in a 20x100 km single-mode fiber (SMF) NDM link. We observe the convergence of the gap to the value $3/2$ predicted by theory already at 28 Gbaud. When convergence is reached, we are in the "IFWM dominated regime". Note that a_{NL}^{pc} is 6 dB larger than a_{NL}^{DP}, as confirmed by the simulated a_{NL} in Fig. 2 (left).

Appendix 4

In this Appendix we prove formulas (8)-(9).

We start by recalling two important results that can easily be derived from [16]:

1) the PWDD $J(c)$ is the inverse 1D Fourier transform of the frequency kernel considered as a function of the single variable $w = \omega_1 \omega_2$: $J(c) = \mathscr{F}^{-1}[\tilde{\eta}(w)] \equiv \int_{-\infty}^{\infty} \tilde{\eta}(w) \mathrm{e}^{jwc} \frac{dw}{2\pi}$. Since $\tilde{\eta}(w)$ is Hermitian, as per Eq. (25), then $J(c)$ is real.

2) the time-kernel $\eta(\tau)$ seen as a function of the single variable $\tau = t_1 t_2$ can be obtained as the following inverse 1D Fourier transform: $\eta(\tau) = \mathscr{F}^{-1}\left[\frac{1}{|\omega|} J(\frac{1}{\omega})\right]$.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

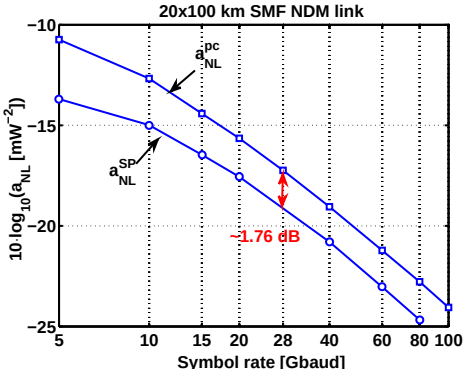

Fig. 5. a_{NL} versus baudrate in 20x100 km SMF NDM coherent link with SP- and DP-QPSK single channel transmission. a_{NL}^{pc} is the per-component NLI coefficient in DP, with $a_{NL}^{pc} = 4a_{NL}^{DP}$. The figure shows convergence of a_{NL}^{pc} to the value $\frac{3}{2}a_{NL}^{SP}$ theoretically predicted when IFWM is dominant.

We thus can prove the following two results:

A)

$$\int_{-\infty}^{\infty} |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} J^2(c)\frac{dc}{2\pi} \tag{31}$$

Proof: from Parseval's theorem for Fourier pairs we have

$$\int_{-\infty}^{\infty} |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} \left|\frac{1}{|\omega|}J(\frac{1}{\omega})\right|^2 \frac{d\omega}{2\pi} = \int_{-\infty}^{\infty} J^2(\frac{1}{\omega})\frac{d\omega}{\omega^2}\frac{1}{2\pi}$$

and after the change of variable $c = 1/\omega$ we finally get Eq. (31). Since the Fourier transform of $\eta(\tau)$ is real, then $\eta(-\tau) = \eta(\tau)^*$, and thus $|\eta(\tau)|^2$ is even. Hence $\int_0^{\infty} |\eta(\tau)|^2 d\tau = \frac{1}{2}\int_{-\infty}^{\infty} J^2(c)\frac{dc}{2\pi}$.

B)

$$\int_{-\infty}^{\infty} \tau^2 |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} c^2 [J(c) + cJ'(c)]^2 \frac{dc}{2\pi}. \tag{32}$$

Proof: we know that time function $\eta(\tau)$ has real Fourier transform $V(\omega) = \frac{1}{|\omega|}J(\frac{1}{\omega})$, since $J(c)$ is real. Hence $\tau\eta(\tau)$ has transform $j\frac{dV(\omega)}{d\omega}$ and by Parseval's theorem then $\int_{-\infty}^{\infty} \tau^2 |\eta(\tau)|^2 d\tau = \int_{-\infty}^{\infty} (\frac{dV(\omega)}{d\omega})^2 \frac{d\omega}{2\pi}$.

Now, for $\omega > 0$, $\frac{dV(\omega)}{d\omega} = \frac{d}{d\omega}(\frac{1}{\omega}J(\frac{1}{\omega}))$ and by the change $c = 1/\omega$: $\frac{dV(\omega)}{d\omega} = \frac{d}{dc}(cJ(c)) \cdot \frac{dc}{d\omega} = [J(c) + cJ'(c)](\frac{-1}{\omega^2})$, where $J'(c) \triangleq \frac{d}{dc}J(c)$. Similarly, for $\omega < 0$, $\frac{dV(\omega)}{d\omega} = -[J(c) + cJ'(c)](\frac{-1}{\omega^2})$. Hence in general $\frac{dV(\omega)}{d\omega} = [J(\frac{1}{\omega}) + \frac{1}{\omega}J'(\frac{1}{\omega})](\frac{-\mathrm{sgn}(\omega)}{\omega^2})$ so $(\frac{dV(\omega)}{d\omega})^2 = [J(\frac{1}{\omega}) + \frac{1}{\omega}J'(\frac{1}{\omega})]^2 \frac{1}{\omega^4}$, and thus by the change $c = 1/\omega$ we get Eq. (32).

Appendix 5

We compute in this Appendix the closed-form expression of the upper-bound Eq. (10) for NDM and DM links. In these calculations, the term $NUM = \int_{-\infty}^{\infty} c^2 [J(c) + cJ'(c)]^2 \frac{dc}{2\pi}$ is the most critical: if the PWDD $J(c)$ has a discontinuity at $c_0 \neq 0$, then $J'(c)$ has a term $\delta(c - c_0)$ which causes NUM (and thus the kernel width $\tau_{rms} = \sqrt{NUM/DEN}$) to diverge to infinity. In such a case, which occurs in most links of interest, the double sum A_{lim} in Eq. (6) still converges, and we verified that the value $\tau'_{rms} = \sqrt{NUM'/DEN}$ obtained by neglecting the Dirac deltas in $J'(c)$ still provides a meaningful time scale for measuring the width of the time-kernel.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

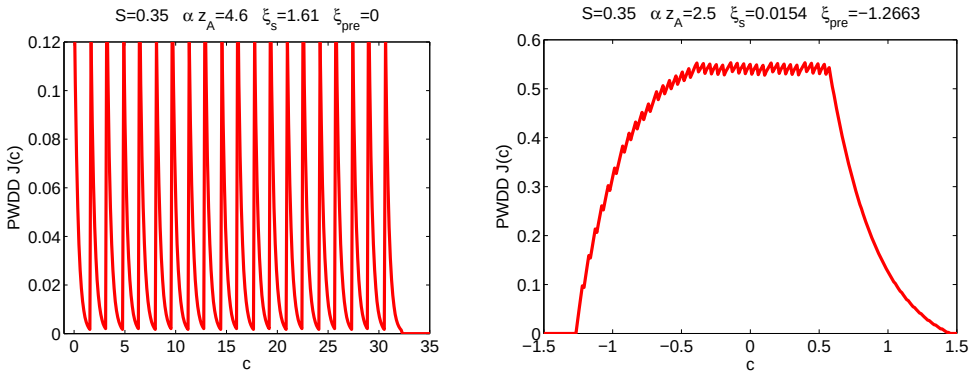

Fig. 6. PWDD versus normalized cumulated dispersion on SMF fiber link ($D = 17$ ps/nm/km, $\alpha = 0.2$ dB/km) and transmission at $R = 28$ Gbaud, corresponding to strength $\mathscr{S} = 0.35$, for **(Left)** $N = 20$ span NDM system with span length $z_A = 100$ km **(Right)** $N = 120$ span DM system at 30 ps/nm RDPS and span length $z_A = 50$ km.

NDM link without pre-compensation

Consider an NDM link composed of N identical spans of length $z_A \gg \frac{1}{\alpha}$, with fiber dispersion parameter β_2 $[ps^2/km]$. The fiber strength is defined as $\mathscr{S} = -\frac{\beta_2}{\alpha}R^2$ [20]. Assuming $\mathscr{S} > 0$, the PWDD is [17]:

$$J_{NDM}(c) = \frac{1}{N}\sum_{k=0}^{N-1} J_s(c - k\xi_s) \tag{33}$$

where $J_s(c) = \frac{1}{\mathscr{S}}e^{-\frac{c}{\mathscr{S}}}U(c)$, with $U(c)$ the unit step function, and the normalized cumulated dispersion *per span* is $\xi_s = -\beta_2 z_A R^2 = \alpha z_A \mathscr{S}$. For instance, Fig. 6 (left) shows $J_{NDM}(c)$ for a 20x100 km SMF NDM system such as the one whose NLT is reported in Fig. 1 (right). Dropping the subscript NDM for brevity, we get $J'(c) = \frac{1}{N}\sum_{k=0}^{N-1} J'_s(c - k\xi_s)$ and $(J(c) + cJ'(c))^2 = \frac{1}{N^2}\sum_{k=0}^{N-1}(J_s(c-k\xi_s) + cJ'_s(c-k\xi_s))^2$ because the PWDDs of the various spans practically do not overlap. Hence

$$NUM = \frac{\sum_{k=0}^{N-1}\int_0^\infty c^2\left(J_s(c-k\xi_s) + cJ'_s(c-k\xi_s)\right)^2 dc}{2\pi N^2} = \frac{2\Sigma_4\xi_s^4 + 2\Sigma_2\xi_s^2\mathscr{S}^2 + 2\Sigma_1\xi_s\mathscr{S}^3 + \mathscr{S}^4}{8\pi N^2\mathscr{S}^3}$$

where $\Sigma_1 = \sum_{k=0}^{N-1} k = \frac{(N-1)N}{2}$, $\Sigma_2 = \sum_{k=0}^{N-1} k^2 = \frac{(N-1)N(2N-1)}{6}$, and $\Sigma_4 = \sum_{k=0}^{N-1} k^4 = \frac{(N-1)N(2N-1)(3N^2-3N-1)}{30}$. Similarly,

$$DEN = \frac{\sum_{k=0}^{N-1}\int_0^\infty \left(J_s(c-k\xi_s)\right)^2 dc}{2\pi N^2} = \frac{1}{4\pi N\mathscr{S}}.$$

For large N, we get $\Sigma_k \cong \frac{N^{k+1}}{k+1}$, hence

$$\tau_{rms}^{NDM} \cong \left(\frac{2\frac{N^4}{5}\xi_s^4 + 2\frac{N^2}{3}\xi_s^2\mathscr{S}^2 + 2\frac{N}{2}\xi_s\mathscr{S}^3}{2\mathscr{S}^2}\right)^{\frac{1}{2}} = \frac{1}{\sqrt{5}}(\alpha z_A N)^2\mathscr{S}$$

and the time-kernel width is seen to scale with N^2. Note that we ignored the delta terms in $J'(c)$.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012

DM link with small RDPS

Assuming RDPS$\ll D/\alpha$ (D is fiber dispersion) we can neglect the ripples in the PWDD [16,17] and use the smooth approximation (we assume $\mathscr{S}>0$)

$$J_{DM}(c)=\begin{cases}\frac{1-\mathrm{e}^{-\frac{c-\xi_{pre}}{\mathscr{S}}}}{\xi_{in}} & \text{if } \xi_{pre}<c<\xi_{pre}+\xi_{in}\\ \frac{\mathrm{e}^{-\frac{c-\xi_{pre}-\xi_{in}}{\mathscr{S}}}-\mathrm{e}^{-\frac{c-\xi_{pre}}{\mathscr{S}}}}{\xi_{in}} & \text{if } c>\xi_{pre}+\xi_{in}\end{cases} \tag{34}$$

with normalized pre-compensation $\xi_{pre}=|\beta'_{pre}|R^2=\frac{D_{pre}}{D/\alpha}\mathscr{S}$ (where β'_{pre} [ps^2] and D_{pre} [ps/nm] are the pre-compensation parameters) and total in-line dispersion $\xi_{in}=N\xi_s$, with normalized per-span residual dispersion $\xi_s=|\beta_s|R^2=\frac{RDPS}{D/\alpha}\mathscr{S}$ (where β_s [ps^2] and $RDPS$ [ps/nm] are the residual dispersion parameters per span). For instance, Fig. 6 (right) shows the true $J(c)$ for a 120x50 km SMF DM system with 30 ps/nm RDPS and SLR pre-compensation, such as the one whose NLT is reported in Fig. 1 (right); $J_{NDM}(c)$ in Eq. (34) represents the smooth average of the true $J(c)$. Dropping for brevity the subscript DM:

$$(J+cJ')^2=\begin{cases}\frac{\left[1-\mathrm{e}^{-\frac{c-\xi_{pre}}{\mathscr{S}}}(1-\frac{c}{\mathscr{S}})\right]^2}{\xi_{in}^2} & \text{if } \xi_{pre}<c<\xi_{pre}+\xi_{in}\\ \frac{(1-\frac{c}{\mathscr{S}})^2\mathrm{e}^{-\frac{2(c-\xi_{pre})}{\mathscr{S}}}(1-\mathrm{e}^{-\frac{\xi_{in}}{\mathscr{S}}})^2}{\xi_{in}^2} & \text{if } c>\xi_{pre}+\xi_{in}\end{cases}$$

and thus $NUM=\frac{1}{\xi_{in}^2}\{\int_{\xi_{pre}}^{\xi_{pre}+\xi_{in}}c^2\left[1-\mathrm{e}^{-\frac{c-\xi_{pre}}{\mathscr{S}}}(1-\frac{c}{\mathscr{S}})\right]^2\frac{dc}{2\pi}+\left(1-\mathrm{e}^{-\frac{\xi_{in}}{\mathscr{S}}}\right)^2\int_{\xi_{pre}+\xi_{in}}^{\infty}c^2(1-\frac{c}{\mathscr{S}})^2\mathrm{e}^{-\frac{2(c-\xi_{pre})}{\mathscr{S}}}\frac{dc}{2\pi}\}$. Long calculations lead to

$$\begin{aligned}NUM=\frac{1}{2\pi\xi_{in}^2 6\mathscr{S}}\{&\\ &-\mathrm{e}^{-\frac{\xi_s}{\mathscr{S}}}[6(\xi_{in}+\xi_{pre})^4+12(\xi_{in}+\xi_{pre})^3\mathscr{S}+30(\xi_{in}+\xi_{pre})^2\mathscr{S}^2+54(\xi_{in}+\xi_{pre})\mathscr{S}^3+\\ &+51\mathscr{S}^4]+[3\xi_{in}^4+6\xi_{pre}^4+12\xi_{pre}^3\mathscr{S}+30\xi_{pre}^2\mathscr{S}^2+54\xi_{pre}\mathscr{S}^3+51\mathscr{S}^4+2\xi_{in}^3(6\xi_{pre}+\mathscr{S})+\\ &+3\xi_{in}(2\xi_{pre}+\mathscr{S})(2\xi_{pre}^2+\mathscr{S}^2)+3\xi_{in}^2(6\xi_{pre}^2+2\xi_{pre}\mathscr{S}+\mathscr{S}^2)]\}.\end{aligned} \tag{35}$$

Similarly, $DEN=\frac{1}{2\pi\xi_{in}^2}\{\int_{\xi_{pre}}^{\xi_{pre}+\xi_{in}}\left[1-\mathrm{e}^{-\frac{c-\xi_{pre}}{\mathscr{S}}}\right]^2dc+\left(1-\mathrm{e}^{-\frac{\xi_{in}}{\mathscr{S}}}\right)^2\int_{\xi_{pre}+\xi_{in}}^{\infty}\mathrm{e}^{-\frac{2(c-\xi_{pre})}{\mathscr{S}}}dc\}$, leading to

$$DEN=\frac{\xi_{in}-\mathscr{S}(1-\mathrm{e}^{-\frac{\xi_{in}}{\mathscr{S}}})}{2\pi\xi_{in}^2} \tag{36}$$

independently of ξ_{pre}. Note that DEN is always ≥ 0.

#155635 - $15.00 USD Received 29 Sep 2011; revised 7 Dec 2011; accepted 8 Dec 2011; published 21 Mar 2012