Accurate Non-Linearity Fully-Closed-Form Formula based on the GN/EGN Model and Large-Data-Set Fitting

Accurate Non-Linearity Fully-Closed-Form Formula based on the GN/EGN Model and Large-Data-Set Fitting

P. Poggiolini[1], M. Ranjbar Zefreh[1], G. Bosco[1], F. Forghieri[2], S. Piciaccia[2]
(1) DET, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129 Torino, Italy, pierluigi.poggiolini@polito.it
(2) Cisco Photonics Italy srl, via Santa Maria Molgora 48/C, 20871 Vimercate (MB), Italy

Abstract: We tested the accuracy of a fully-closed-form GN/EGN formula over 1,700 different fully-loaded systems. We improved it greatly through a correction that leverages the large data-set, providing an effective tool for real-time physical-layer-aware network management.
OCIS codes: (060.0060) Fiber optics and optical communications; (060.1660) Coherent communications;

1. Introduction

Physical-layer-aware control and optimization is of primary importance for the efficient operation of ultra-high-capacity optical networks. A key enabler is the availability of a sufficiently accurate analytical model of fiber non-linear propagation (or of NLI, Non-Linear-Interference). However, the computational effort of such model must be low enough to allow real-time use. One possible strategy is resorting to an *approximate closed-form formula* derived from known NLI models. Several NLI models have been proposed over the years, such as time-domain [1][2], GN [3], EGN [2][4][5], as well as [6][7], and still others.

Here we concentrate on the GN/EGN model class. Approximate closed-forms of the "incoherent" version of the GN model (the iGN model [3]) have been derived in the past. We considered Eqs.(41)-(43) [3], which can handle arbitrary WDM combs and arbitrary non-uniform span links. We upgraded them to make them capable of dealing with fully-loaded C and (C+L)-band systems, by supporting frequency-dependent loss, gain and dispersion, as well as inter-channel SRS (Stimulated Raman Scattering) [8]. We point out that that a similar approach towards upgrading [3] has been recently undertaken by another group too, whose interesting results are reported in [9].

We believe that extensive validation of such formulas across a *very wide range* of highly randomized WDM systems is needed for them to be considered a reliable physical-layer-awareness tool. The focus of this paper is on such validation, of which we show preliminary results addressing a total of over nearly 1700 randomized C-Band fully-populated WDM systems, of increasingly challenging features. We started out with 400 different WDM systems over SMF. We then moved on to a larger test of 1000 WDM systems where the fiber of each span was randomly selected among SMF and two different low-dispersion NZDSFs. Finally, an additional 280 systems set was considered which included frequency-dependent dispersion too. In these tests, reasonably good accuracy was found. Considerable accuracy improvement could however be obtained by performing a simple polynomial best-fit correction obtained by leveraging the large accumulated test data-sets. After correction, the OSNR *estimation error variance* of the closed-form formula vs. the OSNR predicted by accurate numerical integration of the EGN model was *less than 0.06 dB* across all test sets mentioned above. This is ongoing research and we are currently targeting extensive further testing of the closed-form formula [8], also with the added elements of frequency-dependent loss and inter-channel SRS over (C+L) band.

2. The approximate closed-form iGN model formula

All quantities identified by a capital G are power spectral densities. The general structure of the considered links is shown in Fig.1. Each span is composed of a single fiber type, with the customary parameters as indicated in Fig.1, right. At the end of the span there maybe any combination of amplifiers, filters, and VOAs, whose aggregate transfer function is $\Gamma^{(n_s)}(f)$. The WDM comb is fully arbitrary and can change span by span, with the only exception of the CUT. The quantity $G_{\rm NLI}^{\rm end}(f_{\rm CUT})$ is the PSD of NLI at the frequency of the CUT. Given the incoherent NLI accumulation

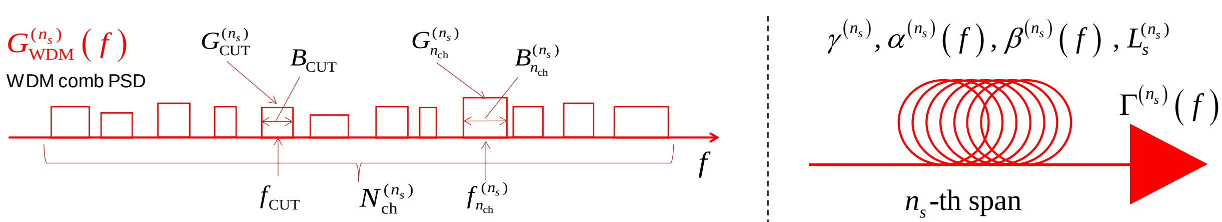

Fig.1 Left: the WDM comb power spectrum $G_{\rm WDM}^{(n_s)}(f)$ at the n_s-th span along the link. Right: the n_s-th span fiber parameters and the lumped elements transfer function $\Gamma^{(n_s)}(f)$.

assumption, $G_{\text{NLI}}^{\text{end}}(f_{\text{CUT}})$ is the sum of the PSD of NLI generated in each single span, $G_{\text{NLI}}^{(n_s)}(f_{\text{CUT}})$, as shown in Eq.(1). Each span contribution is then linearly propagated till the end of the link (the product symbol in Eq.(1)). Eq. (2) provides $G_{\text{NLI}}^{(n_s)}(f_{\text{CUT}})$. In the iGN model it contains double-integrals in frequency whose approximate closed-from solution is shown as Eq. (3) for cross-channel interference (XCI) and as Eq. (4) for self-channel interference (SCI). In the equations, n_s is the considered span. n_{ch} is the channel number, ranging from 1 to $N_{\text{ch}}^{(n_s)}$ for the n_s-th span.

$$G_{\text{NLI}}^{\text{end}}\left(f_{\text{CUT}}\right)=\sum_{n_s=1}^{N_s} G_{\text{NLI}}^{(n_s)}\left(f_{\text{CUT}}\right) p\left(n_s, R_{\text{CUT}}\right)\cdot \prod_{k_s=n_s+1}^{N_s} \Gamma^{(k_s)}\left(f_{\text{CUT}}\right) e^{-2\cdot\alpha^{(k_s)}\left(f_{\text{CUT}}\right)\cdot L_s^{(k_s)}} \qquad \text{Eq. (1)}$$

$$G_{\text{NLI}}^{(n_s)}\left(f_{\text{CUT}}\right)=\frac{16}{27}\left(\gamma^{(n_s)}\right)^2\cdot\Gamma^{(n_s)}\cdot e^{-2\cdot\alpha^{(n_s)}\cdot L_s^{(n_s)}}\cdot G_{\text{CUT}}^{(n_s)}\cdot\left[\sum_{n_{\text{ch}}=1, n_{\text{ch}}\neq n_{\text{CUT}}}^{N_{\text{ch}}} 2\left[G_{n_{\text{ch}}}^{(n_s)}\right]^2 I_{n_{\text{ch}}}^{(n_s)}+\left[G_{\text{CUT}}^{(n_s)}\right]^2 I_{\text{CUT}}^{(n_s)}\right] \qquad \text{Eq. (2)}$$

$$I_{n_{\text{ch}}}^{(n_s)}=\frac{\operatorname{asinh}\left(\frac{\pi^2}{2}\left|\frac{\beta_{2\text{eff},n_{\text{ch}}}^{(n_s)}}{\alpha^{(n_s)}\left(f_{n_{\text{ch}}}^{(n_s)}\right)}\right|\left[f_{n_{\text{ch}}}^{(n_s)}-f_{\text{CUT}}+\frac{B_{n_{\text{ch}}}^{(n_s)}}{2}\right]B_{\text{CUT}}\right)}{8\pi\left|\beta_{2\text{eff},n_{\text{ch}}}^{(n_s)}\right|\cdot\alpha^{(n_s)}\left(f_{n_{\text{ch}}}^{(n_s)}\right)}-\frac{\operatorname{asinh}\left(\frac{\pi^2}{2}\left|\frac{\beta_{2\text{eff},n_{\text{ch}}}^{(n_s)}}{\alpha^{(n_s)}\left(f_{n_{\text{ch}}}^{(n_s)}\right)}\right|\left[f_{n_{\text{ch}}}^{(n_s)}-f_{\text{CUT}}-\frac{B_{n_{\text{ch}}}^{(n_s)}}{2}\right]B_{\text{CUT}}\right)}{8\pi\left|\beta_{2\text{eff},n_{\text{ch}}}^{(n_s)}\right|\cdot\alpha^{(n_s)}\left(f_{n_{\text{ch}}}^{(n_s)}\right)} \qquad \text{Eq. (3)}$$

$$\beta_{2\text{eff},n_{\text{ch}}}^{(n_s)}=\beta_2^{(n_s)}+\pi\beta_3^{(n_s)}\left[f_{n_{\text{ch}}}^{(n_s)}+f_{\text{CUT}}-2f_c^{(n_s)}\right] \qquad \text{Eq. (5)}$$

$$I_{\text{CUT}}^{(n_s)}=\frac{\operatorname{asinh}\left(\frac{\pi^2}{2}\left|\frac{\beta_{2\text{eff},n_{\text{CUT}}^{(n_s)}}^{(n_s)}}{2\alpha^{(n_s)}\left(f_{\text{CUT}}\right)}\right|B_{\text{CUT}}^2\right)}{4\pi\left|\beta_{2\text{eff},n_{\text{CUT}}^{(n_s)}}^{(n_s)}\right|\alpha^{(n_s)}\left(f_{\text{CUT}}\right)} \qquad \text{Eq. (4)}$$

$$\beta_{2\text{eff},n_{\text{CUT}}^{(n_s)}}^{(n_s)}=\beta_2^{(n_s)}+\pi\beta_3^{(n_s)}\left[2f_{\text{CUT}}-2f_c^{(n_s)}\right] \qquad \text{Eq. (6)}$$

Frequency-dependent dispersion can be modeled through the β_3 coefficient. The inclusion of β_3 in the GN model equations was shown in [10] Eq.(C2). We started off from that result but, to close the integrals, the assumption that dispersion was frequency-dependent but *piecewise constant* had to be made. This approach is similar to what is done in [9]. In other words, dispersion is different at each channel frequency, but it is assumed constant over the bandwidth of each channel. The same assumption was made for the loss coefficient $\alpha(f)$. When SRS is included as well, a more complex but still fully closed-form expression is found [8]. Testing with SRS using [8] is work in progress. Note that a SRS closed-form GN-model extension has been proposed in [9] too, showing very interesting results.

3. Large-Data-Set Testing and Fitting

As benchmark we used the EGN model. Using split-step simulations would have not allowed us to accumulate data-sets of adequate size, i.e., in the thousands of cases. The EGN model has repeatedly proved very accurate in several in-depth investigations, such as [11],[12]. We are however planning to double-check a subset of cases through split-step simulations in the near future. In the following, we describe the 1000 systems test-set, first. It is as follows.

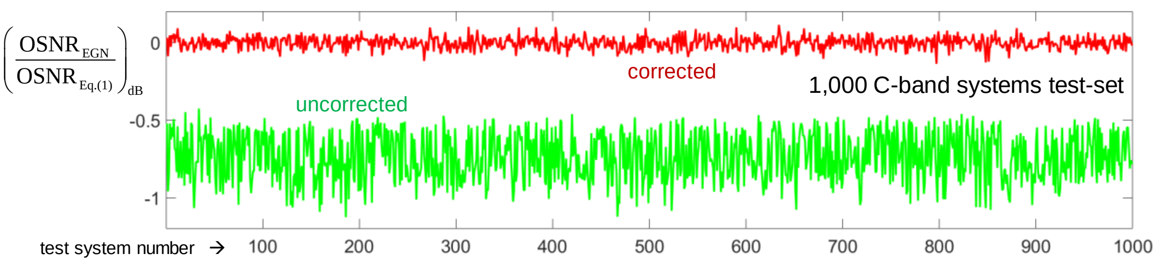

Fig.2: OSNR at max reach, ratio between EGN model estimate and closed-form estimate Eqs.(1)-(6), in dB.

The C band is fully populated by channels; the symbol rate of each channel (including the CUT) is randomly chosen among 32, 64 and 96 GBaud, with spacing 43.75, 87.5 and 125 GHz respectively. The format of each channel is randomly chosen among QPSK, 8, 16, 32, 64, 128 and 256 QAM (all PM), except for the CUT which is is PM-64QAM and it is at the center of the comb. The fiber type is randomly chosen for each span. Fiber data is [α_{dB} (1/km), γ 1/(W km), β_2 (ps^2/km)] equal to [0.21, 1.3, -21], [0.22, 1.77, -4.5], [0.22, 1.35, -2.5] for SMF, NZDSF1 and NZDSF2, respectively. The length of each span is uniformly distributed between 80 and 120 km. The nominal launch power is optimized according to the LOGO strategy [3] Eq.(82), but then each channel launch power (except for the CUT) is randomly altered between 70% to 130% of the optimum power, to mimic real-system power imbalances. For each of these generated systems, first the max-reach is found using the EGN model and then the total OSNR (ASE+NLI) is estimated at that max reach, both with the EGN model and with Eqs.(1)-(6). EDFA NF was 6 dB. The results are shown in Fig. 1. The abscissa is a test-system number, whereas the ordinate is the ratio in dB between the two OSNRs. The green curve shows a mean error of -0.74 dB with a standard deviation σ of 0.155 dB. This first result was quite encouraging, especially because of the small σ. Further investigation showed the error to be strongly related to the *accumulated dispersion and the symbol rate of the CUT*. We therefore introduced in Eq.(1) a factor $p(n_s, R_{\mathrm{CUT}})$ where R_{CUT} is the CUT symbol rate. The correction, fitted over the 1000 test cases, was:

$$p(n_s, 32) = -4.22 \cdot 10^{-5} \cdot x^2 + 8.18 \cdot 10^{-3} \cdot x + 0.361 \qquad p(n_s, 64) = -5.21 \cdot 10^{-5} \cdot x^2 + 8.93 \cdot 10^{-3} \cdot x + 0.373$$

$$p(n_s, 96) = -5.72 \cdot 10^{-5} \cdot x^2 + 9.24 \cdot 10^{-3} \cdot x + 0.391$$

where $x = \mathrm{sqrt}\left(\beta_{2,\mathrm{acc}}^{(n_s)}(f_{\mathrm{CUT}})\right)$ and $\beta_{2,\mathrm{acc}}^{(n_s)}(f_{\mathrm{CUT}})$ is the accumulated dispersion at span n_s at the CUT frequency, in ps^2. With the correction, the red curve in Fig.1 shows an almost complete cancelation of the error vs. the EGN model (mean 0.003 dB, $\sigma = 0.04$ dB) across all 1000 systems. This is remarkable, considering that the system variety in the test-set was quite substantial, with max-reaches ranging from 1 to 8 spans, very diversified WDM combs and links made up of SMF only, NZDSF only or any random mix thereof. We used the strategy over a 400-system test-set similar to the above, where the CUT was either 16QAM or 32QAM (rather than 64QAM) and fiber type was SMF only. We achieved equally good results.

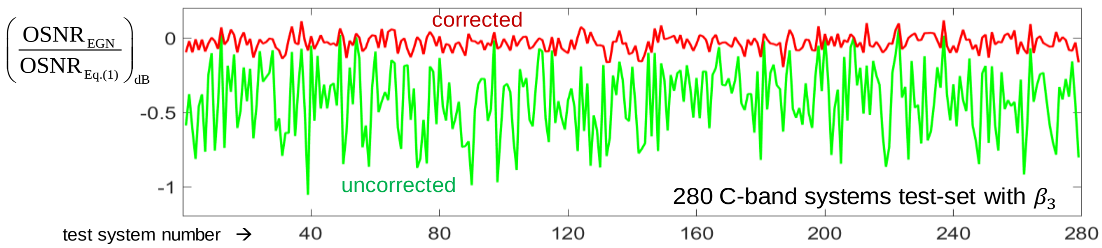

Fig.3: OSNR at max reach, ratio between EGN model estimate and closed-form estimate Eqs.(1)-(6), in dB.

Finally, we performed a 280-systems study, similar to the 1000 test-set, but including fiber dispersion slope as well, with β_3 (ps^3/km): SMF, 0.1452; NZDSF1, 0.1463; NZDSF2, 0.1206. Fig.3 shows the uncorrected (mean -0.42 dB, $\sigma = 0.24$ dB) and corrected results (mean 0.03 dB, $\sigma = 0.06$ dB). The test-set is limited and so these results should be considered preliminary, but the error and especially σ reduction appears to confirm that the method is quite effective. We must point out that *the corrections shown here work well within their respective test-sets*. However, we believe, given the very positive preliminary results shown here, that the approach can be generalized. We are now working on generating a much larger test-set covering all CUT formats and addressing a much wider range of systems, with SRS. Our goal is to find a *simple common correction law* that essentially covers all possible systems of practical interest.

4. Comments and Conclusion

We tested the accuracy of the closed-form iGN formula Eqs.(1)-(6) over 1700 different fully-loaded C-band systems. The formula showed reasonably good accuracy, but we improved it greatly through a correction obtained by leveraging the large result data-sets. Remarkably, a simple best-fit polynomial was sufficient, no complex machine learning algorithms were needed. It enabled us to essentially obtain *EGN-model accuracy at comparatively negligible complexity*. We believe the closed-from formula with the proposed correction strategy can be generalized and thus potentially provide an effective tool for real-time physical-layer-aware management of fully-loaded optical networks.

Acknowledgements: this work was supported by Cisco Systems through an SRA contract and by the PhotoNext Center of Politecnico di Torino.

5. References

[1] A. Mecozzi and R.-J. Essiambre, ‘Nonlinear Shannon limit in pseudolinear coherent systems,’ J. Lightw. Technol., vol. 30, June 2012.

[2] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, ‘Properties of nonlinear noise in long, dispersion-uncompensated fiber links,’ Optics Express, vol. 21, no. 22, pp. 25685-25699, Nov. 2013.

[3] P. Poggiolini, et al., ‘The GN model of fiber non-linear propagation and its applications,’ J. of Lightw. Technol., vol. 32, p. 694, 2014.

[4] A. Carena, et al.,‘EGN model of non-linear fiber propagation,’ Optics Express, vol. 22, no. 13, pp. 16335-16362, June 2014.

[5] P. Serena, A. Bononi, ’A Time-Domain Extended Gaussian Noise Model,’ J. Lightw. Technol., vol. 33, no. 7, pp. 1459-1472, Apr. 2015.

[6] M. Secondini and E. Forestieri, “Analytical fiber-optic channel model in the presence of cross-phase modulation,” IEEE Photon. Technol. Lett., vol. 24, no. 22, pp. 2016–2019, Nov. 2012.

[7] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, ‘Pulse collision picture of inter-channel nonlinear interference noise in fiber-optic communications,’ J. Lightw. Technol., vol. 34, no. 2, pp. 593–607, Jan. 2016.

[8] P.Poggiolini ‘A generalized GN-model closed-form formula,’ submitted to www.arXiv.org Sept. 2018, submission number 2408610.

[9] D. Semrau, R. I. Killey, P. Bayvel, ‘A Closed-Form Approximation of the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering’, www.arXiv.org, paper: arXiv:1808.07940, Aug. 2018.

[10] P. Poggiolini, Y. Jiang,, A. Carena, F. Forghieri ‘Analytical Modeling of the Impact of Fiber Non-Linear Propagation on Coherent Systems and Networks. Chapter 7 in: *Enabling Technologies for High Spectral-efficiency Coherent Optical Communication Networks*. p. 247-310, Wiley, ISBN: 978-111907828-9, doi: 10.1002/9781119078289.

[11] P. Poggiolini and Y. Jiang, ‘Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems’ J. Lightw. Technol., vol. 35, no. 3, pp. 458–480, Feb. 2017.

[12] P. Poggiolini et al, “Non-Linearity Modeling for Gaussian-Constellation Systems at Ultra-High Symbol Rates” ECOC 2018, paper Tu4G.3.