Combining Machine Learning and the GN Model for Fast NLI Prediction in Dispersion-Managed Links

Combining Machine Learning and the GN Model for Fast NLI Prediction in Dispersion-Managed Links

Emanuele Virgillito(1*), Rosario Ietro(1), Antonio Napoli(2), Sai Kishore Bhyri(3) Gabriele Galimberti(4), Walid Wakim(5), Vittorio Curri(1)

(1*) DET, Politecnico di Torino, Italy, emanuele.virgillito@polito.it (2) Nokia, Germany, (3) Nokia, India, (4) Nokia, Italy, (5) Nokia, USA

Abstract *Non-linearity buildup in dispersion-managed network sections can be severe due to the small accumulated dispersion causing non-Gaussian signal distribution and correlation between noise contributions. We propose a machine learning-assisted, spatially disaggregated model able to provide a real-time and accurate non-linearity estimation in this scenario. ©2025 The Author(s)*

Introduction

Nowadays, the method to transport 400 Gbps per wavelength and beyond in optical networks consists of dual-polarization (DP) coherent technologies, whose propagation is optimal over uncompensated transmission (UT) links[1].
However, a significant number of legacy dispersion-managed (DM) optical links, employing IMDD channels, are still operated in metro, access, and submarine networks. Notably, the presence of DM-based links in submarine systems remains non-negligible. While coherent-suited UT has become the standard for all newly deployed cables, enabling the routing of coherent channels over existing DM optical cables would significantly increase capacity and maximize operators' return on CAPEX investments. Numerous studies have shown that nonlinear propagation impairments affecting DP signals can be effectively modeled by nonlinear interference (NLI)—a perturbation to the linear propagation of a WDM channel comb[2]–[6]. In most cases, NLI is dominated by self-channel interference (SCI) and cross-channel interference (XCI), caused by the channel under test (CuT) interfering with itself (SCI) and with co-propagating channels (XCI)[7],[8]. The underlying physical phenomena enable a disaggregated analysis of NLI, even in systems comprising mixed fiber types[9],[10], under UT configurations. This methodology allows the isolation of the contribution to the overall nonlinear impact on the CuT from each individual WDM channel and fiber span, achieving both spectral and spatial disaggregation, while an independent effect is caused by the *Gaussianization* of the signal constellation[11] induced by dispersion accumulation, which leads the GN model to overestimate the NLI. However, this effect is small in UT links due to the large accumulated dispersion. In DM links, the low level of accumulated dispersion—resulting from the use of inline dispersion compensation units (DCUs)—triggers correlation among NLI generated by different fiber spans. This introduces spatial memory effects that disrupt the assumption of spatial disaggregation[12],[13]. As optical networking evolves toward open and disaggregated architectures[14],[15], achieving a disaggregated NLI estimation to build the physical layer digital twin[9] has become essential, and the use of the Gaussian Noise (GN) model[16] is recognized as accurate and effective. For DM optical links, accurate yet time-consuming *aggregated* models have been proposed and validated, such as the EGN model[17] and its evolution toward closed-form expressions[18],[19], while a disaggregated approach was first proposed in[20]. To address spatial coherency, we introduced in[21] a spatially disaggregated method that extends the GN model to estimate the coherent accumulation of SCI in DM, mixed-fiber links, and in[22] we validated this approach also for XCI, using the split-step Fourier method (SSFM) as a reference. We demonstrated that the per-span NLI contribution can be effectively evaluated by the GN model, while both the coherency and *Gaussianization* effects can be evaluated separately, paving the way for an innovative and potentially ultra-fast NLI evaluation applicable to any optical link. In these studies, the effect of *Gaussianization* was

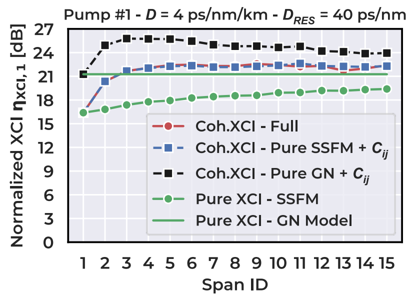

Fig. 1: Example of XCI per-span modeling for a generic interfering channel: *Gaussianization* is represented by the gap between the green curves, while coherency is indicated by the displacement between each green, pure-XCI curve and its corresponding coherent-XCI curve.

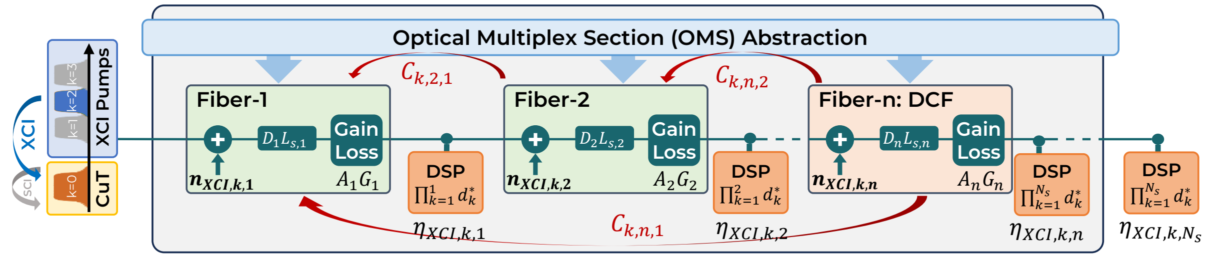

Fig. 2: Block diagram abstraction of a generic DM/UT OMS considered in this paper. XCI noise is modeled as a lumped noise field at the fiber input. ASE noise is not simulated but added in the final GSNR computation.

observed but not explicitly addressed in the modeling, resulting in a worst-case assessment; moreover, a heuristic approach was proposed for the treatment of correlation effects. In this work, we continue along the direction set in[21],[22], and propose the innovative use of machine learning methods to accurately *capture* both the correlation and *Gaussianization* phenomena, combined with the GN model for a per-span, spectrally separated evaluation of NLI. The training dataset was generated using SSFM simulations across scenarios covering a wide region of the feature hyperspace, and the results have shown surprisingly high accuracy on SSFM-simulated testing sets, enabling NLI evaluation within timeframes on the order of milliseconds.

Simulation and modeling coherent NLI

In Fig. 2 is shown the abstraction of the system used for coherent NLI modeling. We will test our model on an actual subsea cable consisting of $N_s = 47$ mixed fiber spans, thus presenting mixed values for all fiber parameters, i.e., attenuation α, dispersion D, and non-linear coefficient γ. All EDFAs, which noise figure is set to $F = 5.5\ dB$, operate in transparency ($A_i G_i = 1$), recovering completely fiber loss. The qualitative dispersion map for the link under analysis is illustrated in Fig. 3. We considered a WDM spectrum composed of $N_{ch} = 15$, identical DP-QPSK channels shaped by raised-root cosine filter with roll-off coefficient $r = 0.0625$, $R_s = 63$ GBaud and spacing between central frequencies $\Delta f = 67$ GHz, consisting so in CuT (probe) with $N_p = 7$ XCI pumps situated on each side.

To compare our modeling results, we use SSFM simulator output as reference, which encompasses all non-linear actual aggregated impairments between channels. In order to estimate

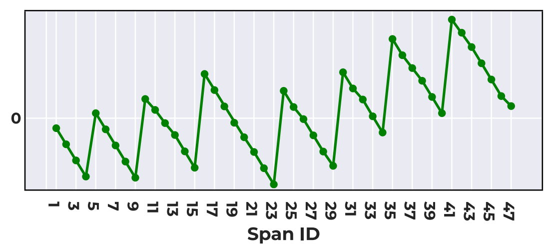

Fig. 3: Dispersion map for the considered link.

the individual contribution of each span for what concerns SCI/XCI only (from now on, we will refer to SCI as XCI_0, to enhance notational brevity), we apply

$$\Delta P_{XCI_{k,i}} = \rho_{k,i}\sigma_{k,i}^2 + 2\sum_{j=1}^{i-1} C_{k,ij} \cdot \sqrt{\rho_{k,i}}\sigma_{k,i} \cdot \sqrt{\rho_{k,j}}\sigma_{k,j} \tag{1}$$

where: $\sigma_{k,i}^2$ is the pure XCI power for pump k introduced at span i computed with GN model, $\rho_{k,i} \in [0,1]$ is the *Gaussianization* coefficient, which accounts for non-Gaussian actual spectrum distribution, and $C_{k,ij} \in [-1,1]$ is the correlation coefficient quantifying coherency accumulation between span i and j due to accumulated dispersion compensation. Both $C_{k,ij}$ and $\rho_{k,i}$ are evaluated with ML methods: this essentially translates into two distinct and independent regression problems where the predicted functions are:

The

Where $D_{acc,(i,j)}$ is the accumulated dispersion between the spans i and j:

$$D_{acc,(i,j)} = \sum_{k=j}^{i-1} D_k L_{s,k}. \tag{2}$$

To predict both $C_{k,ij}$ and $\rho_{k,i}$ values, Random Forest regressor with $N_{estimators} = 100$ underwent extensive training on a wide array of simulation data where different values, for spectral and fiber physical parameters have been considered: to obtain ground truth $C_{k,ij}$, we run SSFM simulations launching a spectrum composed, in XCI case, by pump k and probe. Here, probe power is kept low to avoid residual SCI effect, while in SCI case we launch the CuT alone at its nominal power. Here, we turn on the Kerr effect on all span pair i, j combinations (with $j < i$) per simulation, thus isolating each coherence contribution. To extract the ground truth $\rho_{k,i}$ instead, we run the so described pump and probe simulations by turning on the Kerr effect only in the span i to extract the actual SCI/XCI pure contribution per span $\sigma_{k,i,SSFM}^2$. Finally, we obtain the *Gaussianization* coefficient $\rho_{k,i}$ as $\rho_{k,i} = \frac{\sigma_{k,i,SSFM}^2}{\sigma_{k,i,GN}^2}$. For both models, ML performances can be found at Figures 4 and 5:

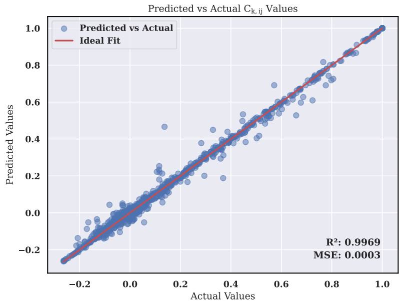

Fig. 4: $C_{k,ij}$ ML metrics

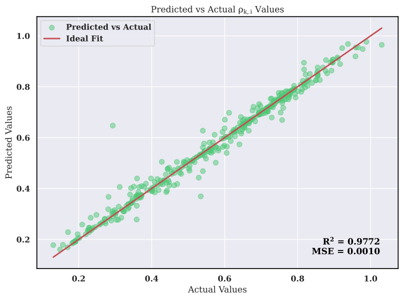

Fig. 5: $\rho_{k,i}$ ML metrics

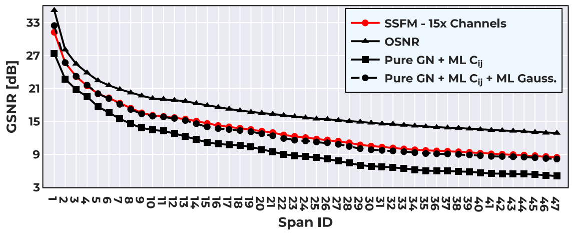

Fig. 6: Test cable OSNR, $\mathrm{GSNR_{SSFM}}$ and GSNR estimation by modeling with and w/o *Gaussianization* correction.

In the actual scenario, each XCI_k power increase presented in Eq. 1 is accumulated according to

$$P_{XCI_{k,i}} = P_{XCI_{k,i-1}} + \Delta P_{XCI_{k,i}} \quad (3)$$

in order to obtain the effective noise power in the span i. Adopting a spectral disaggregation methodology, the total NLI value is derived by aggregating all the calculated XCI components via

$$P_{NLI_i} = P_{SCI_i} + 2\sum_{k=1}^{N_p} P_{XCI_{k,i}} \quad (4)$$

which holds because of the symmetric spectrum with an odd number of channels.

In Fig. 6, we observe the GSNR results, considering contributions from SNR_{NLI} and $OSNR$:

$$\frac{1}{GSNR} = \frac{1}{SNR_{NLI}} + \frac{1}{OSNR} \quad (5)$$

In Fig. 6, the red curve shows the output of the SSFM simulation, which serves as the ground truth. The black curve, with square markers, represents GSNR computed using ML-based estimation for the correlation coefficients, while the *Gaussianization* correction coefficients were not applied (i.e., $\rho_{k,i}$, $\rho_{k,j}$ = 1). In this scenario, the resulting GSNR is conservative compared to the SSFM reference. Conversely, the black curve with circular markers illustrates the model output when ML is used to estimate both the correlation and *Gaussianization* correction coefficients. In this case, the resulting GSNR curve aligns more closely with the SSFM reference than in the previous case. In Fig. 7, we quantify the error on two GSNR curves obtained with and w/o correction for GN model *Gaussianization*: in the case where such correction has not been applied, we have roughly 3.5 dB more conservative results in terms of GSNR at the end of the link, while with *Gaussianization* correction, such gap is slightly below 0.5 dB at the end of the link.

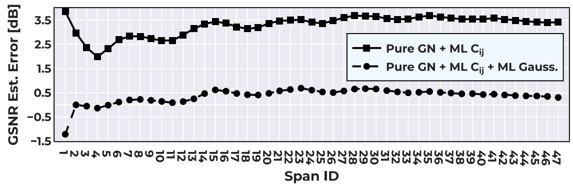

Fig. 7: GSNR estimation error with (circles) and without (squares) *Gaussianization* correction. Error is computed as $\mathrm{GSNR_{SSFM,dB}} - \mathrm{GSNR_{Model,dB}}$

Significant computational power variations exist among the curves in Fig. 6. The SSFM result demands extensive simulation time, often spanning several days and necessitating a high-performance GPU. Without this equipment, obtaining the result would be infeasible. On the contrary, GSNR black curves obtained by the proposed modeling require a computation in the order of milliseconds, which can be performed by any device.

Conclusions

We have proposed a semi-analytical approach to model NLI in DM optical links employing coherent transmission. This approach distinctly accounts for both spatial coherency and *Gaussianization* effects, which are separately estimated using dedicated machine learning models and integrated in semi-analytical, spatial and spectral disaggregated model extending GN model in DM scenario, providing a quick QoT estimation tool avoiding to run time-consuming SSFM simulations.

Acknowledgements

This publication has received funding and support from the EU - Next Generation EU under the Italian NRRP, Mission 4, Component 2, Investment 1.3, CUP E13C22001870001, partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”) and from the EU Horizon 2020 funded project ALLEGRO (GA No. 101092766).

References

[1] V. Curri, P. Poggiolini, A. Carena, and F. Forghieri, “Dispersion compensation and mitigation of nonlinear effects in 111-gb/s wdm coherent pm-qpsk systems”, *IEEE Photonics Technology Letters*, vol. 20, no. 17, pp. 1473–1475, 2008. DOI: 10.1109/LPT.2008.927906.

[2] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, “Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links”, *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1524–1539, 2012. DOI: 10.1109/JLT.2012.2189198.

[3] A. Mecozzi and R.-J. Essiambre, “Nonlinear shannon limit in pseudolinear coherent systems”, *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, 2012. DOI: 10.1109/JLT.2012.2190582.

[4] F. Vacondio, O. Rival, C. Simonneau, *et al.*, “On nonlinear distortions of highly dispersive optical coherent systems”, *Opt. Express*, vol. 20, no. 2, pp. 1022–1032, Jan. 2012. DOI: 10.1364/OE.20.001022. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-20-2-1022.

[5] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “Properties of nonlinear noise in long, dispersion-uncompensated fiber links”, *Opt. Express*, vol. 21, no. 22, pp. 25 685–25 699, Nov. 2013. DOI: 10.1364/OE.21.025685. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-21-22-25685.

[6] D. Semrau, R. I. Killey, and P. Bayvel, “The gaussian noise model in the presence of inter-channel stimulated raman scattering”, *J. Lightwave Technol.*, vol. 36, no. 14, pp. 3046–3055, Jul. 2018. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?URI=jlt-36-14-3046.

[7] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “Pulse Collision Picture of Inter-Channel Nonlinear Interference in Fiber-Optic Communications”, *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 593–607, Jan. 2016.

[8] E. Virgillito, A. Castoldi, S. Straullu, R. Pastorelli, and V. Curri, “QoT Computation for 100G Lightpaths Routed on 10G-loaded Dispersion-Managed Network Segments”, in *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Jun. 2021, pp. 1–6. DOI: 10.1109/ICECCE52056.2021.9514126.

[9] V. Curri, “GNPy model of the physical layer for open and disaggregated optical networking [Invited]”, *Journal of Optical Communications and Networking*, vol. 14, no. 6, pp. C92–C104, Jun. 2022, ISSN: 1943-0639. DOI: 10.1364/JOCN.452868.

[10] E. London, A. D’Amico, E. Virgillito, A. Napoli, and V. Curri, “Modelling non-linear interference in non-periodic and disaggregated optical network segments”, *Optics Continuum*, vol. 1, no. 4, pp. 793–803, Apr. 2022, ISSN: 2770-0208. DOI: 10.1364/OPTCON.453253.

[11] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, “Impact of the transmitted signal initial dispersion transient on the accuracy of the gn-model of non-linear propagation”, in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, 2013, pp. 1–3. DOI: 10.1049/cp.2013.1515.

[12] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, *A Detailed Analytical Derivation of the GN Model of Non-Linear Interference in Coherent Optical Transmission Systems*, Jun. 2014. DOI: 10.48550/arXiv.1209.0394. arXiv: 1209.0394 [physics].

[13] P. Poggiolini and Y. Jiang, “Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems”, *Journal of Lightwave Technology*, vol. 35, no. 3, pp. 458–480, Feb. 2017.

[14] R. Casellas, R. Martínez, R. Vilalta, and R. Muñoz, “Abstraction and Control of Multi-Domain Disaggregated Optical Networks With OpenROADM Device Models”, *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2606–2615, May 2020.

[15] M. Birk, O. Renais, G. Lambert, *et al.*, “The OpenROADM initiative [Invited]”, *Journal of Optical Communications and Networking*, vol. 12, no. 6, pp. C58–C67, Jun. 2020, ISSN: 1943-0639. DOI: 10.1364/JOCN.380723.

[16] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-Model of Fiber Non-Linear Propagation and its Applications”, *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, Feb. 2014, ISSN: 0733-8724, 1558-2213. DOI: 10.1109/JLT.2013.2295208.

[17] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, “Egn model of non-linear fiber propagation”, *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun. 2014. DOI: 10.1364/OE.22.016335. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-22-13-16335.

[18] Y. Jiang, A. Nespola, S. Straullu, *et al.*, “Experimental test of a closed-form egn model over c+l bands”, *Journal of Lightwave Technology*, vol. 43, no. 2, pp. 439–449, 2025. DOI: 10.1109/JLT.2024.3455752.

[19] F. Balasis, D. J. Elson, M. Jarmolovicius, *et al.*, “An extended closed form of the isrs gn model for the zero-dispersion regime”, in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 1435–1438.

[20] E. Seve, P. Ramantanis, J.-C. Antona, *et al.*, “Semi-analytical model for the performance estimation of 100gb/s pdm-qpsk optical transmission systems without inline dispersion compensation and mixed fiber types”, in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, 2013, pp. 1–3. DOI: 10.1049/cp.2013.1513.

[21] E. Virgillito, A. Castoldi, A. D’Amico, *et al.*, “Spatially disaggregated model for self-channel interference in mixed fiber optical network segments”, *Journal of Optical Communications and Networking*, vol. 15, no. 8, pp. C271–C287, Aug. 2023, ISSN: 1943-0639. DOI: 10.1364/JOCN.486713.

[22] E. Virgillito, R. Ietro, S. C. Ramesh, *et al.*, “Spatially disaggregated approach to cross-channel interference in dispersion-managed optical links”, *OFC*, 2025.