Machine Learning for Power Profiles Prediction in Presence of Inter-channel Stimulated Raman Scattering

A. M. Rosa Brusin[(1)], M. Ranjbar Zefreh[(1)], P. Poggiolini[(1)], S. Piciaccia[(2)], F. Forghieri[(2)] and A. Carena[(1)]

[(1)] DET, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, ann.rosabrusin@polito.it
[(2)] Cisco Photonics Italy srl, via Santa Maria Molgora 48/C, 20871 Vimercate (MB), Italy.

Abstract *Two artificial neural network (ANN) models are presented to predict power profiles over C+L–band in presence of inter-channel stimulated Raman scattering (ISRS) and to support non-linear interference (NLI) modeling. High prediction accuracy is obtained with maximum errors $\leq$ 0.1 dB over thousands different partial loads.*

Introduction

Optimization of optical networks requires the knowledge of physical layer behavior: accurate modeling of propagation effects is fundamental for performance assessment of transmission systems. In the past years several analytical models accounting for both amplified spontaneous emission (ASE) noise and non-linear interference (NLI) impairments were proposed[1]–[5].

The ever growing internet data traffic will lead to an increasing capacity demand and a promising solution is to extend the standard C–band toward the L–band, and in a longer term even beyond. In this multi-band scenario, besides the non-linear interaction between channels due to Kerr effect, it becomes relevant the inter-channel stimulated Raman scattering (ISRS), consisting in the power transfer from higher frequency carriers to lower frequency carriers[6]. The resulting effect is a tilt of the power profile over frequencies which depends on the spectral load. Consequently, upgraded analytical models for NLI estimation accounting for ISRS have been proposed[7],[8].

For applications to network control and optimization we need simplified models, capable of real-time predictions. Of particular interest are the closed-form models (CFMs)[9]–[11] which, following some approximations, allow to easily evaluate the generation of non-linear interference (NLI) considering also the ISRS interaction.

In general, a quality of transmission evaluator accounting for both linear and nonlinear effects, considering the impact of ASE and NLI requires a fast and accurate evaluation of power profiles, spectrally and spatially resolved. ISRS models for a large channel count, may not satisfy speed requirements as they are based on the numerical solution of large set of ordinary differential equations (ODEs). Moreover, CFMs are enabled by fitting these power profiles with simplified power profile analytical expressions[11] and this requires further computational effort to get the required parameters for all channels.

In the last years, machine learning (ML) has been successfully applied in different areas of optical communications, with a good success in the design and analysis of Raman amplification[12]–[15]. Based on these recent advances, to overcome the computational complexity problem, we propose two artificial neural network (ANN) models to evaluate power profiles evolution over frequency and distance. A first model is used to predict power profiles, whilst the second one is directly used to predict fitting parameters for the NLI modeling[11].

We perform a comprehensive simulation study considering different spectral loads over C+L–band, showing very high accurate and fast predictions with maximum error always below 0.1 dB.

Scenario and Datasets Generation

We consider a wavelength division multiplexing (WDM) comb of channels over the C+L–band between 185 THz and 196 THz with 220 channels in a 50 GHz grid, carrying 1 mW each when turned ON. The ISRS effect is analyzed over a single span of standard single mode fiber (SSMF) with length $\mathrm{L_{span}}$=100 km and fiber intrinsic attenuation coefficient α=0.21 dB/km. For ANN training and testing, we generate two independent datasets. To reduce the ANN complexity, as proposed in[12], the **training dataset** is generated grouping together 10 adjacent channel slots, resulting in 22 sub-bands of 500 GHz. This simplification is supported by the fact that the power profile does not change significantly over

978-1-6654-3868-1/21/$31.00 ©2021 IEEE

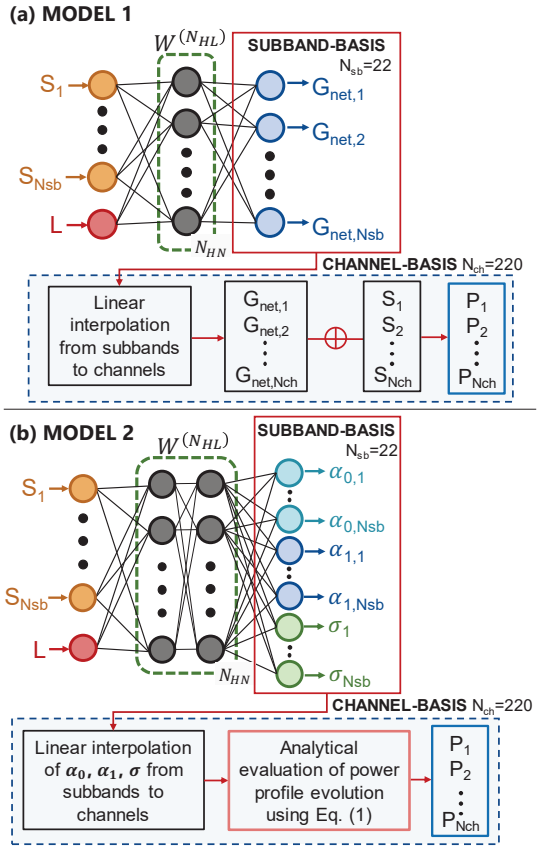

Fig. 1: ANN models, training and testing process for (a) model 1 and (b) model 2.

frequency among adjacent channels, therefore the granularity of 10 channels per sub-band is a good trade-off between accuracy and complexity. To emulate partial loads on sub-bands, three types of datasets are generated based on different power levels per sub-band ($\mathrm{P_{sb}}$): *two* power levels with $\mathrm{P_{sb}}$=[0,10] mW (actually all channels OFF or ON, respectively), *three* power levels with $\mathrm{P_{sb}}$=[0,5,10] mW, and *five* power levels with $\mathrm{P_{sb}}$=[0,2.5,5,7.5,10] mW. To consider a more realistic scenario, the **testing dataset** is generated on a channel-basis with channels being OFF or ON, i.e. $\mathrm{P_{ch}}$=[0,1] mW.

For each type of dataset (training dataset with multi-power levels and testing dataset), we extract $\mathrm{N_{loads}}$=5000 different spectral loads with randomly selected frequency position (and power levels) for sub-bands/channels. The datasets are generated by means of the numerical Raman solver available within the open source library GNPy[16] and power profiles evolution is evaluated over frequency and distance at every kilometer. As the channels are randomly positioned in frequency, their distribution over sub-bands is non-uniform, and this is the reason behind the choice of considering multiple power levels in the generation of the training dataset.

Artificial Neural Network Models

In Fig. 1 the two considered ANN models are shown together with the training and testing process. Inputs and outputs are properly normalized with respect to their mean and standard deviation.

In Fig. 1(a) we have the first ANN (model 1) designed to predict power profiles, which receives at the input the spectral load $\mathbf{S} = [\mathrm{S_1}, \ldots, \mathrm{S_{Nsb}}]$ and the distance L. At the output it has the net gain $\mathbf{G_{net}} = [\mathrm{G_{net,1}}, \ldots, \mathrm{G_{net,Nsb}}]$ defined as the difference in logarithmic units between the output power profiles and the input spectral loads. We set a target maximum prediction error ($\mathrm{E_{MAX}}$) $\leq$ 0.1 dB. To achieve this goal in the first ANN we have to split the span length into four sub-span sections (1-10 km, 11-20 km, 21-50 km, and 51-100 km) and to train four ANNs in parallel, as we were not able to reach the accuracy target with a single ANN working over the entire span. Shorter intervals are in the first kilometers, where the fiber experiences higher ISRS due to higher power levels. For each section we train 20 parallel and independent ANNs (for model averaging) using *random projection* (RP) method with the following hyper-parameters values: $\mathrm{N_{HL}}$=1 hidden layer (HL), $\mathrm{N_{HN}}$=1000 (for 1-10 km, 11-20 km, 21-50 km) and 2000 (for 51-100 km) hidden nodes (HNs), σ=0.1 (standard deviation for weights initialization), λ=10 (regularization parameter) and hyperbolic-tangent as activation function. The training is applied independently on all the three power-levels granularity datasets.

Once trained, the ANNs are used for testing using now the testing dataset generated considering channels, while the ANN works on sub-bands. At the ANN input the power per sub-band is given as the sum of the powers of the ON channels contained in each sub-band. At the ANN output the 22 net gain samples are interpolated over 220 channels that are then applied to the original per-channel spectral load to obtain the resulting power profiles.

Moving to the second ANN (model 2), we need first to describe the fitting parameters introduced in[11] to enable CFM with ISRS. For each channel/sub-band, power profile evolution along distance is assumed to be:

$$P(z) = P(0) \times e^{-2\alpha_0 z + \frac{2\alpha_1}{\sigma}(exp(-\sigma z)-1)}, \qquad (1)$$

where the triplet of coefficients $\{\alpha_0,\alpha_1,\sigma\}$ is numerically determined by fitting the actual power profile. Such fitting procedure is based on a dedicated cost function that weights error along

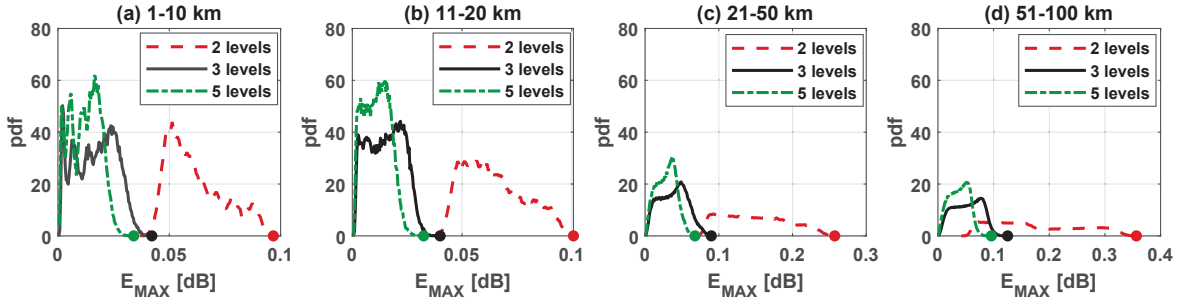

Fig. 2: Probability density function (pdf) of $\mathrm{E_{MAX}}$ of power profiles predicted using **ANN model 1** considering the three power levels granularity for the training dataset for different distance sub-span: (a) 1-10 km, (b) 11-20 km, (c) 21-50 km, and (d) 51-100 km.

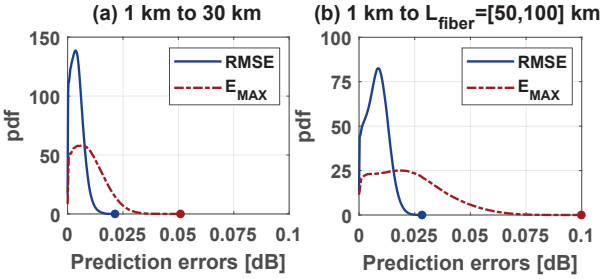

Fig. 3: Probability density functions (pdfs) of RMSE and $\mathrm{E_{MAX}}$ of power profiles obtained using **ANN model 2** considering two distance ranges: (a) 1-30 km and (b) 1-$\mathrm{L_{fiber}}$ km, where $\mathrm{L_{fiber}}$=[50-100] km.

the span: higher weights are assigned when the power is higher, i.e. in the first kilometers of the span, to have a more accurate fitting on the region where most of NLI is generated.

In Fig. 1(b) we have the second ANN (model 2) that receives the same input as model 1, but at the output it has the triplet of vectors $\alpha_{\mathbf{0}} = [\alpha_{0,1}, \ldots, \alpha_{0,\mathrm{Nsb}}]$, $\alpha_{\mathbf{1}} = [\alpha_{1,1}, \ldots, \alpha_{1,\mathrm{Nsb}}]$ and $\sigma = [\sigma_1, \ldots, \sigma_{\mathrm{Nsb}}]$. For ANN model 2 we only consider the training dataset with two power-levels (sub-bands ON or OFF) and a single ANN is trained using *back-propagation* (BP) method with 2 HLs, 20 HNs and logistic sigmoid as activation function. For ANN model 2 testing, we consider only α_0, α_1 and σ for $\mathrm{L_{fiber}}$=[50-100] km. Also in this case linear interpolation is performed to obtain α_0, α_1 and σ on a channel basis.

Simulation Results

The prediction accuracy is evaluated in terms of root-mean-square-error (RMSE) and maximum absolute error $\mathrm{E_{MAX}}$ between the predicted power profiles $\mathrm{P_{pred}}$ and the target profiles $\mathrm{P_{target}}$. Fig. 2 shows the probability density functions (pdfs) of the $\mathrm{E_{MAX}}$ for the four sub-span sections and for the three different power granularity datasets using ANN model 1. In this case $\mathrm{P_{target}}$ is the power profile given by the numerical Raman solver. For sections 1-10 km and 11-20 km (Figs. 2(a),(b)), $\mathrm{E_{MAX}}$ is always below 0.1 dB, therefore there is no need of increasing the granularity of power levels in the training dataset. For section 21-50 km (Fig. 2(c)), with 2 power levels $\mathrm{E_{MAX}}$ reaches values up to 0.26 dB and only with 3 and 5 power levels is below our target value. Finally, for section 51-100 km only with 5 power levels we are able to have all $\mathrm{E_{MAX}}$ below 0.1 dB.

Fig. 3 shows the pdfs of RMSE and $\mathrm{E_{MAX}}$ in case of ANN model 2, for which $\mathrm{P_{target}}$ and $\mathrm{P_{pred}}$ are the power profiles computed inserting in Eq. (1) for each channel the triplet (α_0, α_1 and σ) obtained from fitting or from ANN model 2, respectively. Here, performances are assessed over two ranges of distances: 1-30 km (Fig. 3(a)) where the power is higher and most of NLI generation takes place, and 1-$\mathrm{L_{fiber}}$ (Fig. 3(b)) with $\mathrm{L_{fiber}}$=[50-100] km. In both cases, $\mathrm{E_{MAX}}$ is always below 0.1 dB with very high accuracy in range 1-30 km.

To quantify the advantage of ANNs we compared the time needed to generate the testing dataset (5000 partially loaded power profiles over channels) and to fit α_0, α_1 and σ with respect to the time needed by the ANN to predict same quantities. Numerical solvers ran for $9.5\text{x}10^5$ s and $2.5\text{x}10^4$ s, respectively, while ANN completed the task in few tens of seconds, satisfying the requirements for real-time applications.

Conclusions

We presented two ANN models for power profile prediction in support of NLI modeling over the C+L–band in presence of ISRS, an effect strongly dependent on input spectral loads. Considering a vast set of thousands of partial loads, we showed fast and highly accurate predictions, with $\mathrm{E_{MAX}}$ always below 0.1 dB. Even higher accuracy is achieved in the first 30 km of the fiber span, where most of NLI generation takes place. A significant advantage in terms of computational time is found with respect to numerical approaches .

Acknowledgements

This work was supported by the CISCO Sponsored-Research Agreement (SRA) "OptSys 2022" and by the PhotoNext Center of Politecnico di Torino.

References

[1] M. Secondini and E. Forestieri, "Analytical Fiber-Optic Channel Model in the Presence of Cross-Phase Modulation", *IEEE Photonics Technology Letters*, vol. 24, no. 22, pp. 2016–2019, 2012. DOI: 10.1109/LPT.2012.2217952.

[2] P. Johannisson and M. Karlsson, "Perturbation Analysis of Nonlinear Propagation in a Strongly Dispersive Optical Communication System", *Journal of Lightwave Technology*, vol. 31, no. 8, pp. 1273–1282, 2013. DOI: 10.1109/JLT.2013.2246543.

[3] 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.

[4] 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, 2014. DOI: 10.1109/JLT.2013.2295208.

[5] P. Serena and A. Bononi, "A Time-Domain Extended Gaussian Noise Model", *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1459–1472, 2015. DOI: 10.1109/JLT.2015.2398873.

[6] W. S. Pelouch, "Raman Amplification: An Enabling Technology for Long-Haul Coherent Transmission Systems", *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 6–19, 2016. DOI: 10.1109/JLT.2015.2458771.

[7] M. Cantono, D. Pilori, A. Ferrari, C. Catanese, J. Thouras, J.-L. Augé, and V. Curri, "On the Interplay of Nonlinear Interference Generation With Stimulated Raman Scattering for QoT Estimation", *Journal of Lightwave Technology*, vol. 36, no. 15, pp. 3131–3141, 2018. DOI: 10.1109/JLT.2018.2814840.

[8] D. Semrau, E. Sillekens, R. I. Killey, and P. Bayvel, "A Modulation Format Correction Formula for the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering", *Journal of Lightwave Technology*, vol. 37, no. 19, pp. 5122–5131, 2019. DOI: 10.1109/JLT.2019.2929461.

[9] D. Semrau, R. I. Killey, and P. Bayvel, "A Closed-Form Approximation of the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering", *Journal of Lightwave Technology*, vol. 37, no. 9, pp. 1924–1936, 2019. DOI: 10.1109/JLT.2019.2895237.

[10] M. Ranjbar Zefreh, F. Forghieri, S. Piciaccia, and P. Poggiolini, "Accurate Closed-Form Real-Time EGN Model Formula Leveraging Machine-Learning Over 8500 Thoroughly Randomized Full C-Band Systems", *Journal of Lightwave Technology*, vol. 38, no. 18, pp. 4987–4999, 2020. DOI: 10.1109/JLT.2020.2997395.

[11] M. Ranjbar Zefreh and P. Poggiolini. (). "A Real-Time Closed-Form Model for Nonlinearity Modeling in Ultra-Wide-Band Optical Fiber Links Accounting for Inter-channel Stimulated Raman Scattering and Co-Propagating Raman Amplification", [Online]. Available: https://arxiv.org/abs/2006.03088.

[12] A. M. Rosa Brusin, U. C. de Moura, V. Curri, D. Zibar, and A. Carena, "Introducing Load Aware Neural Networks for Accurate Predictions of Raman Amplifiers", *Journal of Lightwave Technology*, vol. 38, no. 23, pp. 6481–6491, 2020. DOI: 10.1109/JLT.2020.3014810.

[13] M. Ionescu, A. Ghazisaeidi, and J. Renaudier, "Machine Learning Assisted Hybrid EDFA-Raman Amplifier Design for C+L Bands", in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–3. DOI: 10.1109/ECOC48923.2020.9333241.

[14] U. C. de Moura, M. A. Iqbal, M. Kamalian, L. Krzczanowicz, F. Da Ros, A. M. Rosa Brusin, A. Carena, W. Forysiak, S. Turitsyn, and D. Zibar, "Multiband programmable gain raman amplifier", *Journal of Lightwave Technology*,

[15] G. Marcon, A. Galtarossa, L. Palmieri, and M. Santagiustina, "Model-Aware Deep Learning Method for Raman Amplification in Few-Mode Fibers", *Journal of Lightwave Technology*, vol. 39, no. 5, pp. 1371–1380, 2021. DOI: 10.1109/JLT.2020.3034692.

[16] (). "GNPy", [Online]. Available: DOI:%2010.5281/zenodo.3458320,%20%5Cfootnotesize%7B%5Curl%7Bhttps://github.com/Telecominfraproject/oopt-gnpy%7D%7D.