Gain Design of Few-Mode Fiber Raman Amplifiers Using an Autoencoder-Based Machine Learning Approach

Gianluca Marcon(1,⋆), Andrea Galtarossa(1), Luca Palmieri(1), Marco Santagiustina(1)

(1)Department of Information Engineering, University of Padova, Italy, (⋆)gianluca.marcon@dei.unipd.it

Abstract *We propose an unsupervised machine learning method based on autoencoders to design the gain profile of few-mode-fiber Raman amplifiers. We test the method for flat and tilted profiles across the C+L optical band, using a few-mode fiber supporting 6 LP mode groups.*

Introduction

Space-division multiplexing (SDM) has emerged as the leading technology for future optical fiber communications, exploiting few-mode fibers (FMFs) and multi-core fibers (MCFs) to surpass the fundamental limit to the capacity of single-mode fibers (SMFs) and avoid a "capacity crunch"[1],[2]. Mode-multiplexed transmissions using FMFs have experimentally demonstrated the feasibility of using distributed Raman amplifiers (DRAs) in single- and multi-span links[3],[4]. Methods to flatten the Raman gain using multiple pumps have been studied extensively for SMFs, and few works have also been published regarding FMFs[4],[5]. Recently, a machine learning (ML) approach has been proposed for SMFs[6] and FMFs[7], using a dataset of thousands of gain curves generated with random pump powers and wavelengths to train a neural networks (NNs) to learn the relationship between the pump parameters and respective gain. We propose a more robust unsupervised learning scheme based on autoencoders that embeds a numerical Raman model in the training process of the NN to design the gain of multi-pump multi-mode Raman amplifiers, without the need to pre-compute a synthetic dataset. The trained NN can be used to predict the correct pump parameters to generate a target gain profile with low computational complexity. We test this method on a FMF supporting 6 groups of LP modes, shaping the gain of 50 wavelengths on the C+L band using 6 counterpropagating pumps. We show results regarding flat and tilted gain profiles, achieving low mode-dependent gain (MDG) and mode-dependent effective noise figure (NF), with values lower than 0.6 and 0.05 dB, respectively. For flat profiles, a flatness value of about 6% of the total gain is achieved. Finally, the robustness of the predictions with respect to the variation of the input signal power are presented, showing negligible variations when the power of each wavelength and spatial channel is below -10 dBm.

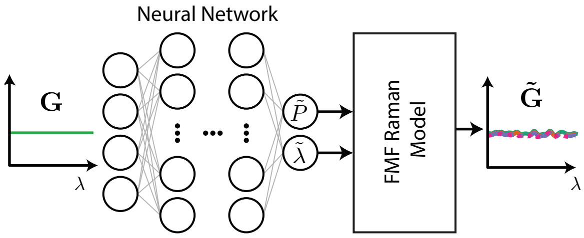

Fig. 1: A diagram of the proposed ML model.

Proposed method

In FMF DRAs, the evolution of the signals, pumps and amplified spontaneous emission (ASE) is governed by the following nonlinear differential equations[5],[8]:

$$\xi_i \frac{dP_i^m}{dz} = -\alpha_i P_i^m + P_i^m \sum_{j\neq i,n} I_{m,n} g_R(\nu_i,\nu_j) P_j^n, \tag{1}$$

$$\begin{aligned}\xi_i \frac{dN_i^m}{dz} =& -\alpha_i N_i^m + N_i^m \sum_{j\neq i,n} I_{m,n} g_R(\nu_i,\nu_j) P_j^n \\ &+ 2h\nu_i B_{ref} \sum_{j\neq i,n} \eta_{i,j} I_{m,n} g_R(\nu_i,\nu_j) P_j^n,\end{aligned} \tag{2}$$

where P_i^m and N_i^m are the signal/pump and ASE power at the ith frequency and mth mode, with $i = 1,\dots,N_s+N_p$, $m = 1,\dots,M$, and N_s, N_p, M the number of signals, pumps and modes, respectively; ξ_i is equal to -1 at the frequency of counterpropagating pumps and $+1$ otherwise; α_i is the attenuation coefficient at the ith frequency; $g_R(\nu_i,\nu_j)$ is the Raman gain coefficient between frequencies ν_i and ν_j. Finally, $I_{m,n}$ are the overlap integrals between mode m and n, defined by

$$I_{m,n} = \frac{\iint_{-\infty}^{+\infty} F_m(x,y)F_n(x,y)dxdy}{\iint_{-\infty}^{+\infty} F_m(x,y)dxdy \iint_{-\infty}^{+\infty} F_n(x,y)dxdy}, \tag{3}$$

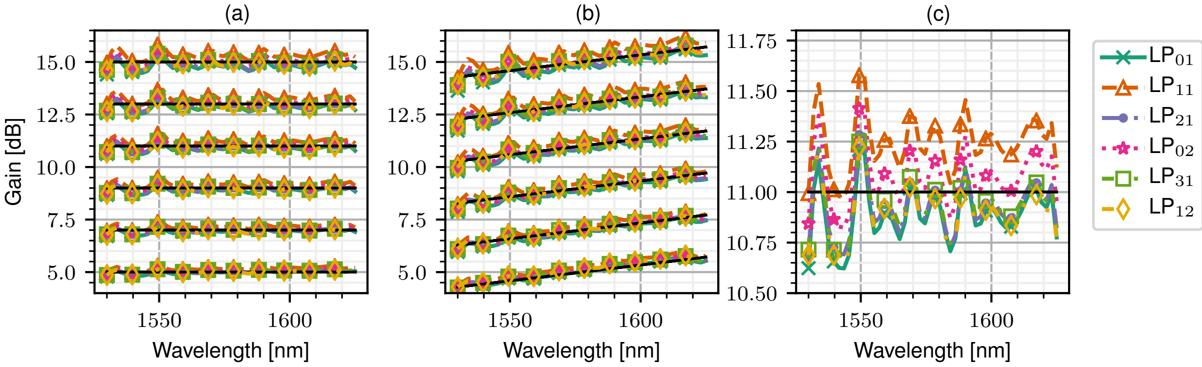

Fig. 2: (a) Predicted gain for the flat case. (b) Predictions for the tilted case. (c) Close-up view of the 11 dB gain profile.

where F_k is the mode intensity profile of the kth mode. The noise source term $\eta_{i,j}$ is defined by $\eta_{i,j} = 1 + [\exp(h|\nu_i - \nu_j|/(k_B T)) - 1]^{-1}$, where h is the Planck's constant, k_B is the Boltzmann's constant, and T is the temperature of the fiber; B_{ref} is the reference noise bandwidth. For signal frequencies $i = 1, \ldots, N_s$ and mode m we can define the on-off gain in a fiber of length L as

$$G_i^m = \frac{P_i^m(L) \text{ with pumps on}}{P_i^m(L) \text{ with pumps off}}. \tag{4}$$

Defining $\mathcal{R}(\lambda_k, P_k^n(L)) = G_i^m$ as the function that computes the on-off gain from the pump wavelengths, λ_k, and powers $P_k^n(L)$, we can define an autoencoder-like model whose diagram is reported in Fig. 1:

$$\tilde{\lambda}_k, \tilde{P}_k^n(L) = \mathcal{N}(G_i^m) \tag{5}$$

$$\tilde{G}_i^m = \mathcal{R}(\tilde{\lambda}_k, \tilde{P}_k^n(L)), \tag{6}$$

where $\mathcal{N}(\cdot)$ is a NN. This model is then trained to minimize a cost function $\mathcal{C}(G_i^m, \tilde{G}_i^m)$ between its input and output gain profiles, so that $\mathcal{N}(\cdot) \approx \mathcal{R}^{-1}(\cdot)$. The NN parameters are iteratively adjusted using backpropagation and gradient-based algorithms such as Adam, using batches of gain profiles as input of the NN in each iteration. Once trained, $\mathcal{N}(\cdot)$ can be used to instantly obtain the pump parameters required for a given gain. For the practical case of flat and tilted curves, we can directly use ideal profiles with average gain level g and tilt t sampled from a training region of interest $\mathcal{T} = [g_{min}, g_{max}] \times [t_{min}, t_{max}]$, removing the need to generate a representative dataset. For the case of counterpropagating pumps, the training process is sped up considerably by using the NN to predict the pump powers at $z = 0$, and obtain the corresponding gain and pump powers at $z = L$ by solving an initial value problem (IVP). This requires a single numerical integration of Eq. (1), avoiding the complex shooting algorithms that are typically used in counterpumping schemes[9]. This advantage comes at the cost of solving an IVP also in the prediction phase, when we interrogate the trained NN for a target gain. However, this approach is still significantly faster than using classical gain flattening methods.

Tab. 1: Values of $I_{m,n}$ as defined in Eq. (3) (units of 10^9 m^{-2}).

$I_{m,n}$	LP_{01}	LP_{11}	LP_{21}	LP_{02}	LP_{31}	LP_{12}
LP_{01}	4.42	2.97	2.14	3.56	1.58	2.84
LP_{11}	2.97	4.76	2.87	1.83	2.44	2.84
LP_{21}	2.14	2.87	4.48	1.69	2.8	1.43
LP_{02}	3.56	1.83	1.69	5.12	1.73	2.48
LP_{31}	1.58	2.44	2.8	1.73	4.25	1.3
LP_{12}	2.84	2.84	1.43	2.48	1.3	4.2

Results

The considered fiber is a $L = 70$ km long step-index fiber supporting 6 groups of LP modes, and whose computed overlap integrals are reported in Tab. 1. The fiber attenuation coefficient as a function of the wavelength is computed as $\alpha(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2$, with $\alpha_0 = 5.788$ dB km^{-1}, $\alpha_1 = -7.1246 \times 10^{-3}$ dB km^{-1} nm^{-1}, $\alpha_2 = 2.268 \times 10^{-6}$ dB km^{-1} nm^{-2}. A standard Raman response function is used[10], with a peak Raman gain coefficient $g_R = 7 \times 10^{-14}$ W^{-1} m. The NN is a feed-forward NN with 5 fully-connected layers of 1000 hidden neurons each, rectified linear unit activations, and a sigmoid function $\sigma(x) = 1/(1+e^{-x})$ applied to the output. The NN is trained using the root-mean-square error cost function and the Adam algorithm. The input signals consist in 50 wavelengths on the C+L band, with an input power of -20 dBm per channel; 6 counterpropagating pumps are used. The training algorithm is run for 1000 iterations, using batches of 1024 curves with average gain and tilt uniformly sampled from the intervals from 5 to 15 dB, and from -0.015 to 0.015 dB/nm, respectively. In Fig. 2 (a) and (b) we show the excellent results in terms of mean error and flatness, for the case of flat and tilted gain

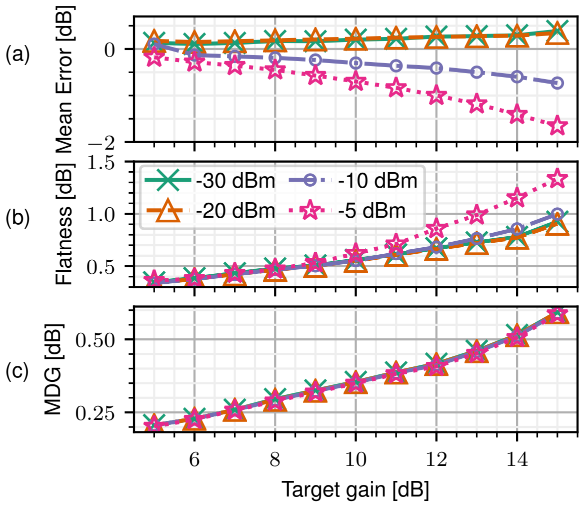

Fig. 3: Mean error (a), flatness (b) and MDG (c) for flat target gains, varying the signal input power.

profiles, respectively. The target tilt is set to 0.015 dB/nm. In Fig. 2 (c) we report a close-up view of the predicted gain curve for the target gain level of 11 dB, showing that the LP_{11} mode is the most amplified, followed by the LP_{02}, and then the rest of the modes with MDG close to zero. Next, we test the robustness of the predictions by evaluating flatness, MDG, and mean error with respect to flat target profiles, varying the input power of the transmitted signals. Mean error and flatness are computed as their maximum value among the different modes. The mean error is reported in Fig. 3 (a), showing that increasing the signal power leads to under-amplification at higher gains. From Fig. 3 (b), similar comments can be made about flatness, which rises to over 1 dB when using -5 dBm per channel. For the other cases, the gain flatness is approximately 6% of the total gain. Varying the input signal power has no significant effect on the MDG, which remains equal to about 3.5% of the target gain level (Fig. 3 (b)). The output ASE spectrum is computed solving Eq. (2) with the predicted pump parameters and used to determine the effective NF at the ith frequency and mth mode[11]:

$$\mathrm{NF}_{\mathrm{eff}}(i, m) = \frac{P_i^m(0) N_i^m(L)}{h\nu_i B_{ref} P_i^m(L)\mathcal{L}(\nu_i)}, \tag{7}$$

where $\mathcal{L}(\nu_i)$ are the total link losses at frequency ν_i. In Fig. 4 we report the maximum $\mathrm{NF}_{\mathrm{eff}}$ among the modes as a function of the achieved gain, for each signal wavelength, using a reference noise bandwidth of 0.5 nm. Recalling the definition of $\mathrm{NF}_{\mathrm{eff}}$[11], Fig. 4 shows that this DRA always outperforms any quantum-limited lumped amplifier. The $\mathrm{NF}_{\mathrm{eff}}$ difference among modes is then computed varying the power of the input signals, using the pump parameters predicted by the model trained

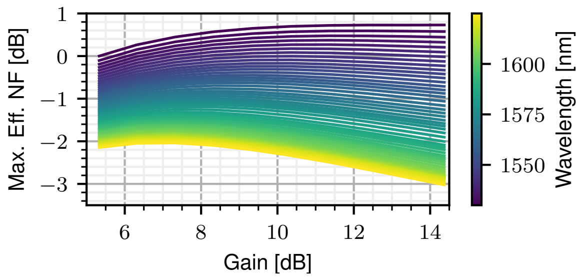

Fig. 4: Maximum effective noise figure among modes as a function of the gain, for every signal wavelength.

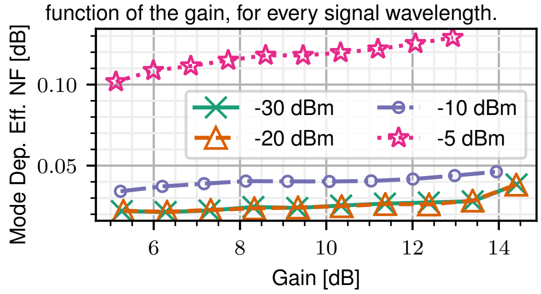

Fig. 5: The mode-dependent effective noise figure as a function of the Raman gain, for various signal power levels.

with -20 dBm per channel. These results are reported in Fig. 5, showing that for low input signal power $\mathrm{NF}_{\mathrm{eff}}$ is mode-equalized, with the total variation remaining below 0.05 dB for every gain level, while rising to about 0.1 dB when -5 dBm per channel are injected in the fiber.

Conclusions

We presented a robust machine learning method based on autoencoders to design the gain of distributed Raman amplifiers in few-mode fibers. A numerical Raman amplification model is embedded in the training of a neural network to learn the mapping between gain profile and the pump parameters that generate it, without using a precomputed dataset. This approach is tested on a fiber supporting 6 LP groups, trasmitting 50 wavelength channels on the C+L band and using 6 counterpropagating pumps: good results have been obtained both for the flat and tilted gain case, showing flatness and mode-dependent gain values of about 6% and 3.5% of the target gain. The noise figure of the amplifier was evaluated, revealing a maximum variation among the modes always lower than 0.05 dB when the input signal power is lower than -10 dBm.

Acknowledgements

Authors acknowledge financial support from the Italian Ministry for Education, University and Research (MIUR, "Departments of Excellence"–law 232/2016, PRIN 2017–project Fiber Infrastructure for Research on Space-division multiplexed Transmission (FIRST)) and the University of Padova (BIRD 2019–project MACFIBER).

References

[1] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres", *Nature Photonics*, vol. 7, no. 5, pp. 354–362, May 2013.

[2] A. Chralyvy, "The coming capacity crunch", in *European Conference on Optical Communication (ECOC)*, Vienna, Austria, Sep. 2009, plenary talk.

[3] M. Esmaeelpour, R. Ryf, N. K. Fontaine, H. Chen, A. H. Gnauck, R.-J. Essiambre, J. Toulouse, Y. Sun, and R. Lingle, "Transmission Over 1050-km Few-Mode Fiber Based on Bidirectional Distributed Raman Amplification", *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1864–1871, Apr. 2016.

[4] J. Li, L. Wang, J. Du, S. Jiang, L. Ma, C. Cai, L. Zhu, A. Wang, M.-J. Li, H. Chen, J. Wang, and Z. He, "Experimental demonstration of a few-mode Raman amplifier with a flat gain covering 1530–1605 nm", *Optics Letters*, vol. 43, no. 18, p. 4530, Sep. 2018.

[5] J. Zhou, "An analytical approach for gain optimization in multimode fiber Raman amplifiers", *Optics Express*, vol. 22, no. 18, p. 21 393, Sep. 2014.

[6] D. Zibar, A. M. Rosa Brusin, U. C. de Moura, F. Da Ros, V. Curri, and A. Carena, "Inverse System Design Using Machine Learning: The Raman Amplifier Case", *Journal of Lightwave Technology*, vol. 38, no. 4, pp. 736–753, Feb. 2020.

[7] Y. Chen, J. Du, Y. Huang, K. Xu, and Z. He, "Intelligent gain flattening in wavelength and space domain for FMF Raman amplification by machine learning based inverse design", *Optics Express*, vol. 28, no. 8, pp. 11 911–11 920, Apr. 2020.

[8] W. Wang, J. Zhao, Z. Yang, C. Li, Z. Wang, L. Yu, and R. Mi, "Amplified Spontaneous Emission and Rayleigh Scattering in Few-Mode Fiber Raman Amplifiers", *IEEE Photonics Technology Letters*, vol. 29, no. 14, pp. 1159–1162, Jul. 2017.

[9] H. M. Jiang and K. Xie, "Efficient and robust shooting algorithm for numerical design of bidirectionally pumped Raman fiber amplifiers", *Journal of the Optical Society of America B*, vol. 29, no. 1, p. 8, Jan. 2012.

[10] D. Hollenbeck and C. D. Cantrell, "Multiple-vibrational-mode model for fiber-optic Raman gain spectrum and response function", *Journal of the Optical Society of America B*, vol. 19, no. 12, p. 2886, Dec. 2002.

[11] J. Bromage, "Raman Amplification for Fiber Communications Systems", *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 79–93, Jan. 2004.