Forward Raman Amplifier Optimization Using Machine Learning-aided Physical Modeling

(Article begins on next page)

Forward Raman Amplifier Optimization Using Machine Learning-aided Physical Modeling

Metodi P. Yankov*, Darko Zibar*, Andrea Carena†, Francesco Da Ros*
*Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark,
†Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino - Italy
meya@fotonik.dtu.dk

Abstract: *A differentiable nonlinear interpolation function learns the Raman gain efficiency and enables gradient-descent-based optimization of a Raman amplifier with arbitrary number of pumps. Example is given for unrepeatered links with a remote pumping stage.*
Keywords: *Raman amplifier, Machine learning, Unrepeatered transmission*

I. INTRODUCTION

Raman amplifiers (RAs) present a significant advantage over erbium-doped fiber amplifiers (EDFAs) in terms of noise figure and potential for gain shaping [1]. Broadband amplification using RAs can be achieved by employing multiple Raman pumps at different frequencies. However, such configurations pose a challenge for the optimization of the pumps' frequency and power due to the increased dimensionality of the problem [1]. Machine learning (ML) methods for the optimization are gaining traction [2], especially methods that train inverse system models to predict the required pump power and frequency for a given target gain profile [2] [3] [4]. ML methods provide excellent performance, however, they require a lot of training data to be generated in order to populate the $2N_p$ dimensional space, where N_p is the number of pumps. Furthermore, each training dataset is specific to N_p and the number of wavelength division multiplexed (WDM) channels to be amplified and does not allow re-optimization when pumps are to be added or when the data load changes, unless an expanded training dataset including all the configurations is considered [5].

In this paper, we lift the requirement for training data by optimizing directly using the physical model for stimulated Raman scattering. The model is made differentiable in the pump powers and frequencies by first training a differentiable nonlinear interpolation function for the Raman gain coefficient. The optimization is then completely flexible in terms of 1) number of pumps; 2) number of WDM channels; 3) number of RA stages; 4) fiber length.

II. DIFFERENTIABLE PHYSICAL MODEL OF THE SRS

The SRS is typically described with ordinary differential equations [6], which can be solved numerically *for forward propagating carriers* using the expression:

$$P(n,z) = P(n,z-\Delta_z) - \alpha\Delta_z + \sum_{m=1}^{N} \frac{g_R(\omega_m-\omega_n)}{A_{eff}} L_{eff}(\Delta_z) e^{P(m,\ z-\Delta_z)}, \quad (1)$$

where P(n, z) is the power at frequency index n and distance z, g_R is the Raman gain coefficient for a given offset between the angular frequencies ω_m and ω_n, A_{eff} is the fiber effective area, $L_{eff}(L) = (1- \exp(-2\alpha L))/ (2\alpha)$ is the effective power interaction length [6], α is the fiber attenuation, and Δ_z is the step size, set to 100 m in this paper (justified below).

In (1), no distinction is made between a 'pump' and a 'channel'. Each step in z is differentiable in power, but not in frequency due to g_R which is typically obtained using a look-up table or piece-wise interpolation. In [7], g_R is approximated using linear interpolation g_R^{LIN}, which is sufficient for estimating the SRS between WDM channels, but fails to provide the required accuracy when high-power pumps are added near the maximum efficiency (as will be demonstrated later in the paper). To that end, we train a deep neural network (DNN) to learn a nonlinear interpolation function of g_R^{NL}. The DNN is depicted in Fig. 1a), has 3 layers with 100 nodes per layer and a ReLU activation function, and is trained using gradient descent (GD) with the Adam optimizer and the mean squared error (MSE) cost function. In Fig. 2a), the normalized g_R^{NL}, g_R^{LIN} and the true g_R are given as a function of the frequency offset for standard, single-mode fiber (SSMF). An MSE between g_R and g_R^{NL} of $4.1 \cdot 10^{-5}$ was achieved. The DNN is differentiable in the frequency offset and allows optimization w.r.t. pump frequency by substituting it for g_R in Eq. (1).

The main drawback of this method is that it only allows optimization of forward Raman pumps, since solving (1) with

The authors are supported by Danish National Research Foundation Centre of Excellence SPOC (DNRF123), the Villum YIP, OPTIC-AI, (grant no. 29344), the European Research Council (ERC-CoG FRECOM Grant NFo. 771878) and the Italian Ministry for University and Research (PRIN 2017, project FIRST).

backward-propagating pumps requires iterative algorithms which are not differentiable.

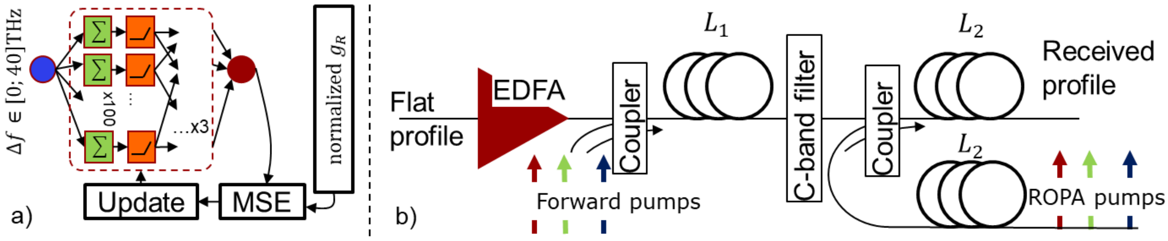

Fig. 1. **a)**: Raman efficiency model and training; **b)**: Considered single-span setup with a ROPA.

III. Considered Setup

Forward pumping is applicable to e.g. remote pumping stages and hybrid EDFA-RA for boosting the launch power in e.g. unrepeatered links [8], which are single-span and typically very long. The multi-pump multi-stage example we consider is given in Fig. 1b). A full C-band load is considered with 40 channels in the [191.6; 195.6] THz range (C-band). An EDFA provides the initial amplification to P_{EDFA}^{out} = 15 dBm (≈ -1 dBm per channel). A flat profile is assumed at the EDFA input at a total power P_{EDFA}^{in} = 0 dBm. The EDFA model considered in this work is trained experimentally [8] and provides a highly non-flat gain (seen in the results section). Flattening the EDFA output using a filter would require wasting power, which is highly undesirable in such power-deficient links. Instead, an additional RA is used to 1) boost the power further; and 2) optimize the transmitter output w.r.t. the power spectral profile at the receiver. An SSMF of L_1 = 210 km follows, at which point a forward pumping Raman remote optical power amplifier (ROPA) is assumed with Raman pump lasers guided from the receiver location by an additional fiber. An L_2 = 40 km of SSMF span is assumed between the ROPA and the receiver. We assume a fiber loss of 0.2 dB/km in the C-band. In the S-band, where the pump lasers would typically be placed, the loss is assumed to be 0.25 dB/km. These numbers are chosen to be in line with SSMF, however, the model in Eq.(1) is completely flexible w.r.t. the individual loss coefficient per channel (e.g. obtained from fiber characterization and for specialized fiber types).

IV. Optimization Strategy

The RAs frequency and power are optimized such that a flat profile is received at a given power level P_{target}^{tot} = -10 dBm (≈ -26 dBm per channel), which means that the RAs need to provide a total gain of 25 dB to the C-band. Some semi-heuristic iterative methods have been developed for the optimization of such systems [8] [10], which become challenged when the number of parameters grows (e.g. due to the increased number of RA pumps). Furthermore, here the full C-band load together with the non-flat launch characteristics pose an additional challenge to the optimization of the RAs. The RAs are initialized uniformly in the frequency range of [198; 220] THz and with equal power per pump $P_i^F = P_{tot}/N_F$, $P_i^R = P_{tot}/N_R$, where P_i^F and P_i^R are the powers of the i−th forward and i−th remote pump, respectively, N_F and N_R are the number of forward and remote pumps, respectively, and P_{tot} = 2.5 W is the total power constraint of the RA, as considered in this work. Gradient descent optimization is then applied w.r.t. both RAs pumps' frequency and power with a target cost of minimizing the MSE between the received profile and the target profile.

In order to apply frequency and power constraints, the MSE is appended with the following cost terms:

$$\begin{aligned} Cost = MSE &+ ReLU\left(F_{min} - \min_i f_i^R\right) + ReLU\left(F_{min} - \min_i f_i^F\right) + ReLU\left(\max_i f_i^R - F_{max}\right) \\ &+ ReLU\left(\max_i f_i^F - F_{max}\right) + ReLU\left(\max_i P_i^R - P_{max}\right) + ReLU\left(\max_i P_i^F - P_{max}\right) \\ &+ ReLU\left(\sum_i P_i^R - P_{tot}\right) + ReLU\left(\sum_i P_i^F - P_{tot}\right), \end{aligned} \tag{2}$$

where P_{max} = 2 W is the power constraint per laser, f_i^F and f_i^R are the frequency and power of the i−th forward and remote pump, respectively, and F_{max} = 220 THz and F_{min} = 198 THz are the frequency constraints (range chosen to support 1st and 2nd order pumping). After initial convergence, pump lasers with frequency offset ≤ 200 GHz are merged together in a single pump with a power equal to the sum of powers (if the P_{max} constraint for the new pump is satisfied) and frequency equal to the average of the frequencies. Optimization then resumes. This ensures the efficient use of lasers.

Different variants can be devised of merging strategy, e.g. always selecting the N pumps with highest powers and pruning the rest, or prune based on minimum resulting deterioration of the MSE. In either case, the designer is free to impose a constraint on the number of pumps available for the RA and/or their frequency.

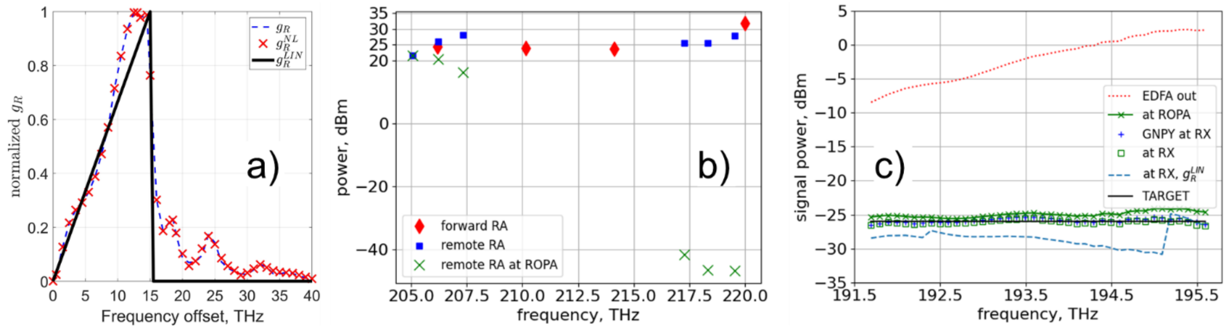

Fig. 2. Simulation results. **a)**: Raman gain coefficient modeled with a linear approximation (solid line) and a DNN (markers), together with the true g_R (dashed line); **b)**: optimized Raman configuration when initialized to 20 remote and 20 forward pumps, converged to 7 remote and 6 forward pumps; **c)**: resulting power profiles for the configuration **b)**, together with the GNPy benchmark.

V. Results

The received power profiles optimized and estimated using the model (1) are compared to 1) a full implementation in GNPy of the setup from Fig. 1 including a Raman solver for the set of ordinary differential equations describing the SRS and using a piece-wise interpolation for the tabulated g_R values [11]; and 2) the model in Eq. (1) which assumes g_R^{LIN}. In Fig. 2b), the optimized forward and remote pumps are given when the optimization algorithm was initialized with 20 pumps each. The model converged to $N_F = 4$ and $N_R = 6$, an MSE of 0.07dB^2 and a max error to target of 0.67 dB. The benchmarks are then given in Fig. 2c). We see excellent correspondence between the model and the GNPy benchmark (maximum error of 0.7dB between the target profile and GNPy validation), which also justifies the chosen step size. In this case, the forward RA provides ≈ 18.07 dB of gain, while the ROPA provides ≈ 6.88 dB with total powers $\sum_i P_i^F =$ 2.20 W and $\sum_i P_i^R =$ 1.52W, respectively. The linear assumption on g_R results in significant discrepancies of ≈ 4.8 dB of maximum error, indicating that it is not feasible to apply this simple model for optimization.

The minor discrepancies between the model in Eq. (1) and GNPy are attributed to the finite step size. At the chosen value of Δ_z= 100 m and the optimization requires ≈ 500 iterations to converge which takes ≈5 min on a standard CPU.

VI. Conclusions

An optimization method was presented for forward Raman amplifiers which is completely flexible in the main system and amplifier parameters. The optimization follows the physical model of the SRS and does not require training data to be generated. An obvious extension is to include the RA and EDFA NFs in the model in order to optimize the received SNR profile instead of power similar to [9].

Acknowledgment

The authors would like to thank Dr. Uiara Celine de Moura for the fruitful discussions in preparing the manuscript.

References

[1] G.C.M. Ferreira, et al. “Optimization of Distributed Raman Amplifiers Using a Hybrid Genetic Algorithm With Geometric Compensation Technique,” IEEE Phot. J., 3(3), 2011.
[2] U.C. de Moura, et al. “Experimental Characterization of Raman Amplifier Optimization Through Inverse System Design,” JLT, 39(4), 2021.
[3] G. Marconi, et al. “Model-Aware Deep Learning Method for Raman Amplification in Few-Mode Fibers,” JLT, 39(5), 2021.
[4] M. Soltani, et al. “Inverse Design of a Raman Amplifier in Frequency and Distance Domains Using Convolutional Neural Networks,” OL, 46(11), 2021.
[5] A.M.R. Brusin, et al. “Introducing load aware neural networks for accurate predictions of Raman amplifiers”, JLT, 38(23) 2020.
[6] I. Roberts, et. al. “Channel Power Optimization of WDM Systems Following Gaussian Noise Nonlinearity Model in Presence of Stimulated Raman Scattering”, JLT, 35(23), 2017.
[7] D. Semrau, et. al. “Modeling and Mitigation of Fiber Nonlinearity in Wideband Optical Signal Transmission [Invited]”, JOCN, 12(6), 2020.
[8] T. Sutili, et al. “Cost-Effective Solution for High-Capacity Unrepeatered Transmission”, OFC 2020 T4I.4.
[9] M. P. Yankov, et. al. “SNR Optimization of Multi-span Fiber Optic Communication Systems Employing EDFAs with Non-flat Gain and Noise Figure”, JLT, 39(21), 2021.
[10] T. Sutili, et. al., ”Unrepeatered Optical Systems Design based on Iterative Channel Pre-Emphasis Optimization,” ICECCE, 2021, pp. 1-6
[11] A. Ferrari, et. al. “GNPy: an Open Source Application for Physical Layer Aware Open Optical Networks”, JOCN, 12(6), 2020.