POLITECNICO DI TORINO
Repository ISTITUZIONALE

Flexible Raman Amplifier Optimization Based on Machine Learning-Aided Physical Stimulated Raman Scattering Model

(Article begins on next page)

28 April 2024

Flexible Raman Amplifier Optimization Based on Machine Learning-aided Physical Stimulated Raman Scattering Model

Metodi Plamenov Yankov, *Member, IEEE*, Francesco Da Ros, *Senior Member, OSA, Senior Member, IEEE*, Uiara Celine de Moura, *Member, OSA*, Andrea Carena, *Senior Member, OSA, Senior Member, IEEE*, and Darko Zibar

Abstract—The problem of Raman amplifier optimization is studied. A differentiable interpolation function is obtained for the Raman gain coefficient using machine learning (ML), which allows for the gradient descent optimization of forward-propagating Raman pumps. Both the frequency and power of an arbitrary number of pumps in a forward pumping configuration are then optimized for an arbitrary data channel load and span length. The forward propagation model is combined with an experimentally-trained ML model of a backward-pumping Raman amplifier to jointly optimize the frequency and power of the forward amplifier's pumps and the powers of the backward amplifier's pumps. The joint forward and backward amplifier optimization is demonstrated for an unrepeatered transmission of 250 km. A gain flatness of < 1 dB over 4 THz is achieved. The optimized amplifiers are validated using a numerical simulator.

Index Terms—Stimulated Raman scattering, Raman amplifier, frequency and power optimization, machine learning.

I. INTRODUCTION

Raman amplifiers (RAs) present a significant advantage over erbium doped fiber amplifiers (EDFAs) in terms of noise figure and the potential for providing arbitrary gain profiles in a controlled way [1]. They are also a viable option for increasing the transmission bandwidth beyond the conventional C+L band and support S-band amplification [2], [3]. Gain shaping over frequency is important for achieving uniformity of the quality of transmission for all channels in a wavelength division multiplexing (WDM) system. Gain nonuniformity is even more pronounced in ultra wide band systems due to different responses and losses of the optical components in different bands. Furthermore, in unrepeatered link scenarios, RAs are critical for providing amplification from the receiver end, as well as remote pumping [4], [5], [6]. 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 [7]. This problem is exacerbated when higher order pumping is employed, as is often the case for unrepeatered systems. Genetic algorithms were proposed in [7], [8], which fall in

the heuristic optimization methods category, and are prone to finding local optimums. State of the art approaches for optimization are also based on heuristics [4], [9], [10]. Machine learning (ML) methods for the optimization are also gaining traction [11], [12], [13], [14], especially methods which train combinations of forward and inverse system models to predict the required pump power and frequency for a given target gain profile [11], [15], [16], [17], [18]. 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 [19].

This paper is an extension of our previous paper [20], where a completely flexible optimization method was proposed for forward-only pumping. In this paper, the forward propagation model from [20] is combined with an ML-based model for the gain of a backward-pumping RA with a fixed number and frequency of the pumps in order to build a complete and differentiable model for a link with joint backward and forward pumping. The optimization is then completely flexible in terms of 1) number of pumps, their frequency and power for the forward RA; 2) power for the backward RA; and 3) fiber length (to a reasonable extent discussed later).

The paper is summarized as follows. In Section II, the physical stimulated Raman scattering (SRS) model and its forward propagation differentiable approximation are presented, including the ML-based model for the Raman gain coefficient. We also introduce the ML-based model for the gain of backward-pumping RA. In Section III, the unrepeatered systems under consideration are explained. In Section IV, the optimization strategies are discussed. In Section V, the optimized gain profiles are presented and analyzed. Section VI concludes the work and provides an outlook.

II. SRS MODELING

The SRS effect in optical fibers results in transfer of power from carriers at high frequencies to carriers at low frequencies. The efficiency of this transfer depends on their respective carriers' frequency offset and their powers. It also depends on the type of fiber through its effective area. This effect

Metodi Plamenov Yankov, Francesco Da Ros, and Darko Zibar are with the Department of Electrical and Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark. Uiara Celine de Moura was with the Department of Electrical and Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark, and is now with NKT Photonics, 3460 Birkerød, Denmark. Andrea Carena is with the Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino - Italy. e-mail: meya@fotonik.dtu.dk

allows for SRS-based amplification of communication signals. However, it also results in inter-carrier power transfer. At high powers, it results in undesired non-uniformity of the received power levels that will consequently define unbalanced level of optical signal to noise ratio (OSNR) for the different channels in the WDM system. The effect is exacerbated when wideband systems beyond the standard C-band are deployed for data transmission, e.g. C+L [21], or even S+C+L, as proposed in recent studies [2].

The SRS with powers expressed in the linear domain can be described with a set of ordinary differential equations [21]

$$\frac{\partial P_n(z)}{\partial z} = -2\alpha_n P_n(z)$$
$$+ \sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{eff}} P_n(z) P_m(z) \quad (1)$$

where $m, n \in [1; N]$ are the indexes of the carriers co-propagating through the fiber, $P_n(z)$ 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, α_n is the fiber loss at the $n-$th frequency and A_{eff} is the fiber effective area. In (1), no distinction is made between a 'pump' and a data-carrying 'channel'. Commonly, the efficiency of the power transfer, i.e. $g_R(\omega)$, is characterized and tabulated for the more popular types of fiber, e.g. standard, single mode fiber (SSMF) [22].

Since there is no closed form analytical solution for Eq. (1), finite difference numerical solutions are typically applied. Choosing a reference direction of propagation allows for the following expression to be formulated for the power evolution in the log domain after a small propagation distance Δ_z [21]

$$P_n(z) = P_n(z - \Delta_z) - \alpha_n \Delta_z$$
$$+ \sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{eff}} L_{eff}(\Delta_z) e^{P_m(z-\Delta_z)}, \quad (2)$$

where, $L_{eff}(L) = \frac{1-\exp(-2\cdot\alpha_n\cdot L)}{2\cdot\alpha_n}$ is the effective power interaction length. When all carriers propagate in the chosen reference direction, the solution (2) converges to the true solution for (1) for vanishing discretization step. However, the joint operation of forward and backward propagating carriers give rise to a non-trivial boundary value problem. In that case, iterative solvers must be employed for propagation in both directions of the fiber in order to arrive at an accurate solution, e.g. through a shooting algorithm [23]. The forward-only solution is relevant for cases of inter-carrier SRS combined with a forward-only Raman pumping amplification. However, in the presence of backward pumping, the iterative solution needs to be applied.

Optimization of any system requires that a cost function is formulated and minimized. In general, non-heuristic (e.g. gradient-based or analytical) optimization requires that the cost function is differentiable w.r.t. the optimization parameters. In the case of RAs, that means an SRS model is needed in a form that is differentiable in the Raman pumps' power and frequency. First, the iterative nature of the solver for the general case of forward+backward propagating carriers
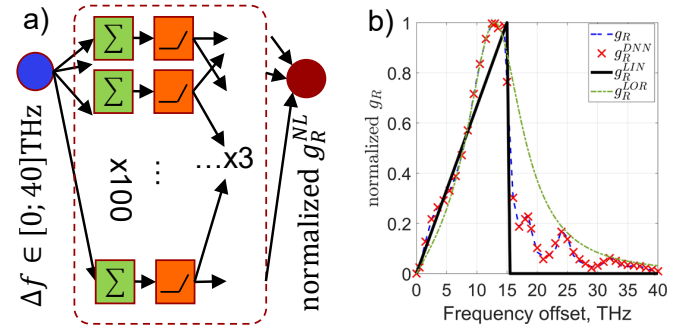

Fig. 1. Neural network modeling of the normalized Raman gain coefficient $g_R(\omega_1 - \omega_2)$. **a):** NN topology for modeling the coefficient: 3 layers, 100 nodes per layer, ReLU activation function; **b):** Modeling accuracy of the NN output g_R^{DNN} compared to the true coefficient g_R, the linear fit g_R^{LIN} applied in [26] and the Lorentzian fit g_R^{LOR} from [25].

results in non-differentiability in itself. Then, in (2), each step in z is differentiable in power, but not in frequency due its dependence on g_R which is typically described using a look-up table, piece-wise interpolation or numerical integration [24][25]. These challenges are addressed in the following.

A. Differentiable forward model

In [26], g_R is approximated using a linear interpolation g_R^{LIN}, which is differentiable until the discontinuity point, chosen for example to coincide with the frequency offset with maximum Raman efficiency. This is sufficient for estimating the SRS between WDM channels, but fails to provide the required accuracy when high-power pumps are added near and beyond the maximum efficiency. Alternatively, a popular method to approximate the Raman gain coefficient is to fit a Lorentzian distribution g_R^{LOR} [25]. Similarly, this fit results in a poor accuracy outside of the peak of the gain coefficient (as will be demonstrated later in the paper). To that end, we train a deep neural network (DNN) to learn a nonlinear interpolation function g_R^{DNN}. The DNN is depicted in Fig. 1a), it has 3 layers with 100 nodes per layer and a ReLU activation function and it is trained using gradient descent (GD) with the Adam optimizer and the mean squared error (MSE) cost function. In Fig. 1b), the normalized g_R^{DNN}, g_R^{LIN}, g_R^{LOR} 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^{DNN} of $4.1 \cdot 10^{-5}$ was achieved. The DNN is differentiable in the frequency offset and allows optimization w.r.t. the pump frequency by substituting it for g_R in (2), while at the same time significantly improves on the accuracy w.r.t. g_R^{LOR}.

B. Differentiable backward model

The method for creating a differentiable model which includes backward pumping is detailed in [11]. For a given fiber length L_{span}, number of backward pumps N_{BCKW}, frequencies of the backward pumps $\mathcal{F}_{BCKW} = \{f_i^B, i = 1 \ldots N_{BCKW}\}$, powers of the backward pumps $\mathcal{P}_{BCKW} = \{P_i^B, i = 1 \ldots N_{BCKW}\}$ and channel load

$\{P_1(0), P_2(0), \ldots, P_N(0)\}$, ML can be applied to learn the on-off gain of the amplifier defined as

$$G(\omega_n) = P_n(L_{span}, \mathcal{P}_{BCKW}, \mathcal{F}_{BCKW})$$
$$- (P_n(0) - \alpha_n L_{span}), \quad (3)$$

where $P_n(L_{span}, \mathcal{P}_{BCKW}, \mathcal{F}_{BCKW})$ is the output power of the fiber ($z = L_{span}$) at the desired frequency index n. The dependence on the backward pumps' frequencies and powers is also explicitly stressed. In order to train a model that predicts this mapping, a training data set is built based on either an iterative solver for the SRS, e.g. implementing a shooting algorithm, or experimentally. In either case, the training dataset is built by first sweeping the required free parameters, which can be any subset of $\{\mathcal{P}_{BCKW} \cup \mathcal{F}_{BCKW}\}$. Then, recording $P_n(L_{span}, \mathcal{P}_{BCKW}, \mathcal{F}_{BCKW})$ for all n. This method is very powerful, but suffers several important drawbacks. The main one is the dimensionality of the problem. Every time a pump is added to the dataset, due to the curse of dimensionality, the number of gain profiles that is needed for accurate representation of the parameter space grows exponentially. Furthermore, the model so constructed cannot generalize to predicting the gain profile for an arbitrary channel load when the amplifier is operated in saturation (high-input signal power leading to pump depletion) [19]. Nevertheless, it allows for an accurate representation of the gain profile of a backward pumping RA to be built, and provides sufficient accuracy in the un-saturated (pump undepleted) regime (see Section III-B). As in typical ML problems, the approximate gain profile $\{G(\omega_n), n = 1 \ldots N\}$ is predicted by an NN. The NN is fully differentiable, and so it allows gradient descent through it in order to optimize the pump configuration w.r.t. a desired gain profile.

In this paper, we follow the process presented in [11] where the NN model is trained using experimentally generated dataset. A 100 km of SSMF is used for transmission. The channel load is assumed fixed to 40 channels with 100 GHz spacing for a total input power of approx. 3 dBm. The pumps' frequencies are also fixed to $[206.1, 207.5, 209.0, 210.6]$ THz, respectively, and the pumps' powers are swept on a regular grid in the range (including boundary points) $[0; 21]$ dBm in order to record the on-off gain profile. Under these conditions, the amplifier is operated outside the saturation regime and pump depletion is negligible. The NN then predicts the gain in dB for an arbitrary pump power configuration by essentially nonlinearly interpolating between the training gain profiles. The NN topology was coarsly optimized to 2 hidden layers of size 256 and 128, respectively with a ReLU activation function. The NN is trained to minimize the MSE

$$MSE = \mathrm{E}_{n,i} \left[|G_{meas}(\omega_n, i) - G_{model}(\omega_n, i)|^2 \right], \quad (4)$$

where $G_{meas}(\omega_n, i)$ and $G_{model}(\omega_n, i)$ are the measured and modeled gains at frequency index n for pump profile i. The maximum absolute error (MAE), defined as

$$MAE = \max_{n,i} \left[|G_{meas}(\omega_n, i) - G_{model}(\omega_n, i)| \right], \quad (5)$$

is also measured. The model achieved an MSE of 0.0329 dB^2 and a MAE of 1.0814 dB, similarly to [11]. For comparison, a
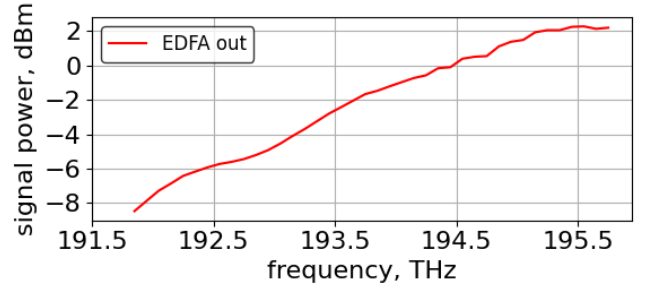

Fig. 2. Power profile launched into the fiber at the transmitter side.

NN with the same topology was applied to a synthetic dataset generated using the GNPy software [27]. The achieved MSE and MAE for that dataset were 0.0033 dB^2 and 0.8892 dB, respectively. The slightly worse modeling performance of the experimental dataset is mainly attributed to the measurement uncertaintity in the experimental data.

III. SYSTEMS UNDER CONSIDERATION

The RA optimization is tested on an unrepeatered link. Due to the long distance, such links require a very high transmit power. Here, forward RA can be applied to aid the transmitter erbium doped fiber amplifier (EDFA) and push some of the gain further into the fiber. This allows to keep a high signal power for longer compared to an EDFA-only setup. The RA then has a dual function: 1) to boost the power; 2) to shape the combined gain profile of the EDFA and the RA. Flattening the gain profile of the EDFA using standard methods, e.g. filtering is undesirable in such links due to the waste of power. In this work, without loss of generality, we assume an EDFA with a non-flat gain profile [28], resulting in a power profile launched into the fiber given in Fig. 2 at a total power of 15 dBm. The total link distance we consider is 250 km. Similar to the backward model training stage, we consider a channel load of $N = 40$ channels on a 100 GHz ITU grid.

A. Forward RA only with remote Raman pumping stage

The forward-only RA optimization is exemplified in Fig. 3. In order to demonstrate the versatility of the model, we consider a Raman-based remote optically pumped amplifier (ROPA), where Raman pumps are guided from the receiver to the ROPA using an independent fiber, coupled into the transmission fiber and provide extra amplification for the last section of the span. In this case, both fibers L_1 and L_2, including the fiber leading the remote pumps to the pumping location are modeled using (2), which is differentiable in the RAs frequency and power if g_R^{LOR} or g_R^{DNN} are applied. This system is motivated by [29], where similar configuration is applied with forward RA and ROPA using a high-order Raman amplification.

B. Forward & backward pumping

The joint forward and backward RA optimization is exemplified in Fig. 4. The receiver directly pumps in the backward
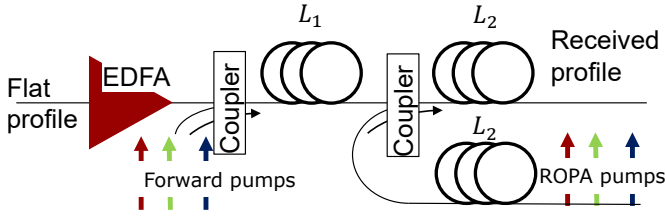
Fig. 3. The unrepeatered transmission setup with a remote pumping stage. Both the transmitter and the remote RA are pumping in the forward direction.
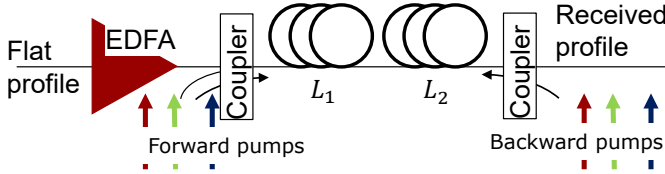

Fig. 4. The unrepeatered transmission setup with both forward and backward pumping RA.
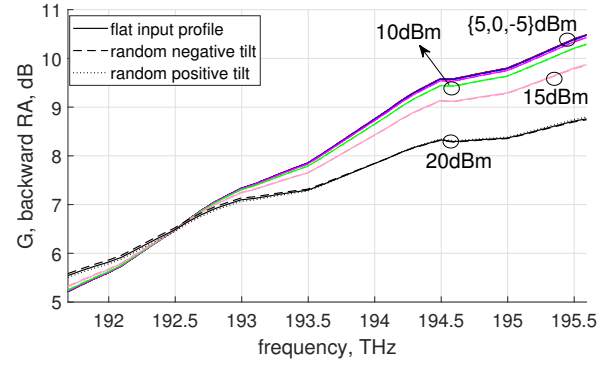

Fig. 5. Depletion study for the backward RA. Slight dependence on the input channel load is seen for high total power (20 dBm). Below a total power of the load channels of 10 dBm, the gain of the RA is independent of both the input power profile and the total power. Total power of the load profiles indicated with annotations.

direction. In this case, the distance is generally long enough and thus the signal power entering the backward-pumping section is weak enough to ensure that the pumps of the backward RA do not suffer from depletion. The un-saturated on-off gain then provides an accurate gain profile of the backward RA, and the ML model discussed in Section II-B can be applied. In order to confirm this assumption, the difference in gain profiles for the backward RA is examined in Fig. 5. The pump powers of all four backward pumps are fixed at their maximum of 21 dBm. Different input load profiles at different total load powers (indicated with circles in the figure) are studied: a flat load profile, a randomly generated positively tilted load profile with an excursion of 5 dB and a randomly generated negatively tilted load profile with an excursion of 5 dB at total power input to a 100 km fiber of $[-5; 20]$ dBm. Below total powers of 10 dBm, the gain profile is constant and independent of both the input load profile and the total load power. This allows for the on-off gain model to be applied in unrepeatered systems. For such systems, after 100+ km of transmission, the total power will typically be well below the threshold of 10 dBm even if we apply a forward RA pumping in the first section of the span. The unrepeatered fiber span of length L_1+L_2 is modeled using two connected models of length L_1 and L_2. The first section is modeled using (2), and the second section is modeled using the ML model from Section II-B. The concatenation of these models is completely differentiable in the forward RA's pump frequencies and powers, as well as in the backward RA's pump powers (pump frequencies are assumed fixed to the values presented in Section II-B).

IV. OPTIMIZATION METHODS

For the system with ROPA, the following cost function is adopted for the optimization of the RAs

$$
\begin{aligned}
\mathcal{L}(\mathcal{P}_{RMT}, \mathcal{P}_{FRW}, \mathcal{F}_{RMT}, \mathcal{F}_{FRW}) = \quad (6)\\
\mathrm{E}_n \left[|P_{modeled,n}(L_{span}) - P_{target,n}(L_{span})|^2 \right] + \\
ReLU \left(F_{min} - \min \left(\mathcal{F}_{RMT} \cup \mathcal{F}_{FRW} \right) \right) + \\
ReLU \left(\max \left(\mathcal{F}_{RMT} \cup \mathcal{F}_{FRW} \right) - F_{max} \right) + \\
ReLU \left(\max \left(\mathcal{P}_{RMT} \cup \mathcal{P}_{FRW} \right) - P_{max} \right) + \\
ReLU \left(\sum_j P_j^R - P_{tot} \right) + ReLU \left(\sum_i P_i^F - P_{tot} \right),
\end{aligned}
$$

where $\mathcal{P}_{RMT} = \{P_i^R, i = 1 \ldots N_{RMT}\}$, $\mathcal{P}_{FRW} = \{P_i^F, i = 1 \ldots N_{FRW}\}$, $\mathcal{F}_{RMT} = \{f_i^R, i = 1 \ldots N_{RMT}\}$ and $\mathcal{F}_{FRW} = \{f_i^F, i = 1 \ldots N_{FRW}\}$ are the powers of the remote RA, powers of the forward RA, frequencies of the remote RA and frequencies of the forward RA, respectively. The first term in (6) is the MSE over frequency between the modeled power $P_{modeled,n}(L_{span})$ and the target power $P_{target,n}(L_{span})$, respectively. Then, P_{max} is the power constraint per laser. The ReLU function is defined as $ReLU(x) = \max[0, x]$ and ensures the penalization of solutions outside of the power and frequency constraints of the pump lasers. Here, $F_{max} = 220$ THz and $F_{min} = 198$ THz are the frequency constraints with a range chosen to support 1st and 2nd order pumping for C-band operation. The total power constraint per amplifier is $P_{tot} = 2.5$ W.

For the system with backward pumping, the following version of the cost function is applied for optimization

$$
\begin{aligned}
\mathcal{L}(\mathcal{P}_{BCKW}, \mathcal{P}_{FRW}, \mathcal{F}_{FRW}) = \quad (7)\\
\mathrm{E}_n \left[|P_{modeled,n}(L_{span}) - P_{target,n}(L_{span})|^2 \right] + \\
ReLU \left(F_{min} - \min \mathcal{F}_{FRW} \right) + \\
ReLU \left(\max \mathcal{F}_{FRW} - F_{max} \right) + \\
ReLU \left(\max \left(\mathcal{P}_{BCKW} \cup \mathcal{P}_{FRW} \right) - P_{max} \right) + \\
ReLU \left(\sum_j P_j^B - P_{tot} \right) + ReLU \left(\sum_i P_i^F - P_{tot} \right),
\end{aligned}
$$

After initial convergence, pump lasers with a frequency separation ≤ 200 GHz are merged together in a single pump with a power equal the sum of powers and frequency equal the average of the frequencies. The iterative optimization process is then resumed. This ensures a practical and efficient use of lasers limiting the number of pumps needed. Different variants can be devised of merging strategy, e.g. always selecting the N pumps with highest powers and pruning the rest, or pruning 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. We note that there may be numerous solutions which provide similar result in terms of optimization, which can be obtained with different initializations and merging strategies.

V. RESULTS

The optimization processes outlined in Section IV are run for the systems under consideration. Without loss of generality, a flat received power profile is targeted. The cases of g_R^{DNN} and g_R^{LOR} are considered for optimization. The results of both are benchmarked against a full implementation in GNPy including a Raman solver for the set of ordinary differential equations describing the SRS (given in Eq.(1)) and using a piece-wise interpolation for the tabulated g_R values [27]. The EDFA output power for both systems is assumed 15 dBm, and results in the power profile in Fig. 2 when its input is a flat profile. The target total received power is chosen to be -10 dBm, or ≈ -26 dBm per channel, which is reasonable for such systems [29].

A. System with remote RA pumping stage

For the system given in Fig. 3, the optimization algorithm was initialized with 8 pumps in each of the forward and the remote amplifier, and $P_{max} = 2$ W. In Fig. 6, the optimized forward and remote pumps are given for the case of g_R^{DNN}. The model converged to $N_{FRW} = 5$ and $N_{RMT} = 4$, an MSE of 0.08dB^2 and an MAE of 0.87 dB. The high-frequency pumps arrive at the coupling stage at very low power, clearly indicating their role as second-order pumping for the low-frequency pumps. The latter then provide gain to the C-band.

The benchmarks are then given in Fig. 7 for g_R^{DNN}. We see excellent correspondence between the model and the GNPy benchmark. The MAE between the target profile and GNPy validation is 1.06dB, and the MAE between the GNPy and the differentiable model prediction is 0.19 dB, which also justifies the chosen step size (see Section V-C). In this case, the forward RA provides ≈ 18.4 dB of total gain, while the ROPA provides ≈ 6.6 dB of total gain with total powers $\sum_i P_i^F = 2.29$ W and $\sum_i P_i^R = 2.5$ W, respectively. In Fig. 8, the GNPy benchmark is given when optimization is performed with the Lorentzian approximation g_R^{LOR}. During training, this model results in MSE and MAE which are slightly higher than the DNN case. Even more importantly, we see significant discrepancies of ≈ 2.45 dB of MAE w.r.t. the GNPy benchmark, indicating that this simple model is not sufficiently accurate for optimization.
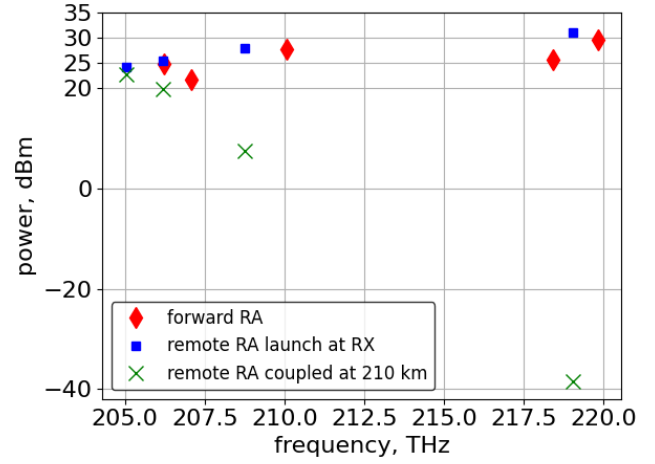

Fig. 6. Optimized Raman configuration using g_R^{DNN} when initialized to 8 remote and 8 forward pumps, converged to 4 remote and 5 forward pumps.
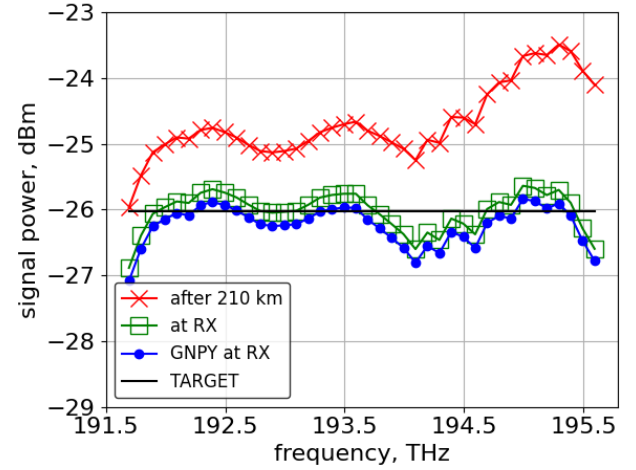

Fig. 7. Resulting power profiles at the receiver when the RAs are optimized using g_R^{DNN}, together with the GNPy benchmarks, for the system with remote pumping.

B. System with backward pumping

For this system, the optimization algorithm was initialized with 8 pumps in the forward amplifier with $P_{max} = 1$ W. For the backward amplifier, $P_{max} = 0.14$ W, which is chosen to be within the operating region of the lasers used to generate the data for the ML model.

In Fig. 9, the optimized forward and backward pumps are given for g_R^{DNN}. The model converged to $N_{FRW} = 4$, an MSE of 0.14dB^2 and an MAE of 0.63 dB. Due to the power constraint per pump, the two pumps at ≈ 215 THz could not be merged. The benchmarks are then given in Fig. 10. A maximum error of 0.65dB between the target profile and GNPy validation of the proposed optimized profile is seen in this case. The total power at the end of the first fiber section is ≈ 0.8 dBm, well below the threshold shown in Fig. 5 for depletion to occur. This assumption is further validated by the good correspondence to the GNPy benchmark at the receiver,
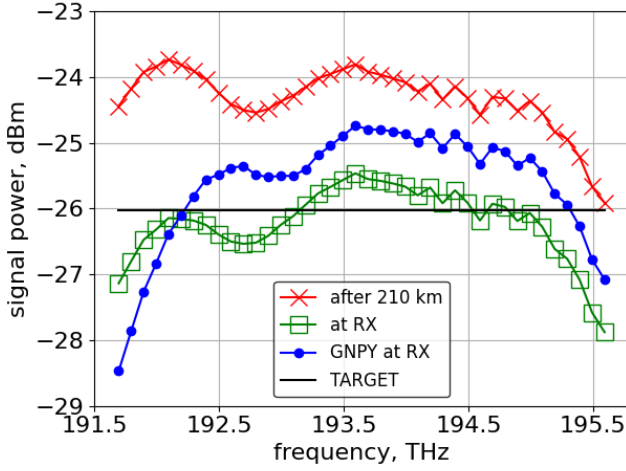
Fig. 8. Resulting power profiles at the receiver when the RAs are optimized using g_R^{LOR}, together with the GNPy benchmarks, for the system with remote pumping.
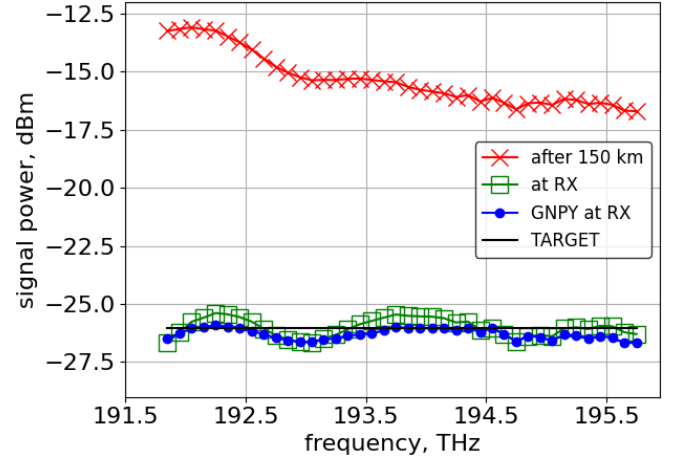

Fig. 10. Resulting power profiles at the receiver when the RAs are optimized using g_R^{DNN}, together with the GNPy benchmarks, for the system with backward pumping.
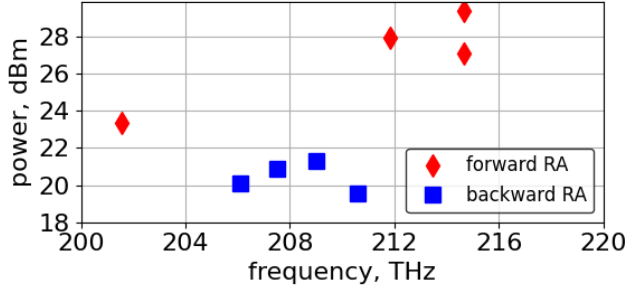

Fig. 9. Optimized Raman pumps' configuration using g_R^{DNN} for the system with backward pumping when initialized to 8 forward pumps, converged to 4 forward pumps
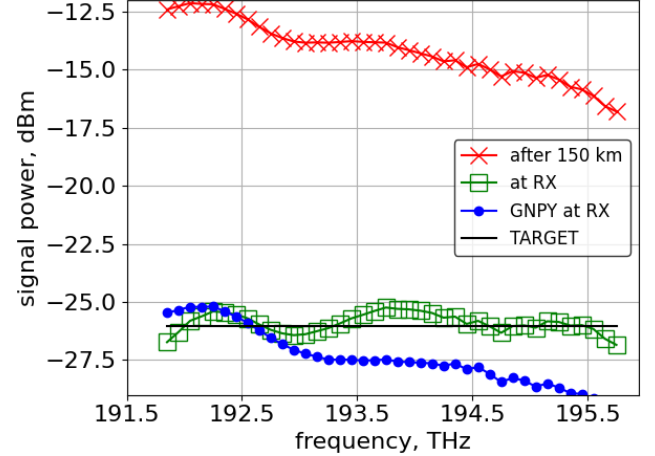

Fig. 11. Resulting power profiles at the receiver when the RAs are optimized using g_R^{LOR}, together with the GNPy benchmarks, for the system with backward pumping.

which does not employ the two virtual sections but considers the fiber as a single span. In this case, the forward RA provides ≈ 15.8 dB of gain, while the backward RA provides ≈ 9.2 dB with total powers $\sum_i P_i^F = 2.21$ W and $\sum_i P_i^B = 0.45$ W, respectively. In Fig. 11, the GNPy benchmark is given when optimization is performed with the Lorentzian approximation g_R^{LOR}. Once again, significant discrepancies of ≈ 3.6 dB of MAE w.r.t. the GNPy benchmark can be observed, even though flatness was achieved during the training stage.

C. Final remarks

The remaining discrepancies between the model from (2) and the GNPy benchmark can potentially be further reduced by decreasing the step size at the expense of slower simulation time. At the chosen step size of $\Delta_z = 100$m, the optimization requires ≈ 500 iteration to converge which takes ≈ 5 minutes on a standard CPU. The discrepancies between the backward model and GNPy are most likely due to remaining mis-calibration between the fiber parameters and non-modeled coupling losses, non-flat fiber loss over the C-band, the un-certainties in the measurements used to obtain the ML model, and similar.

It should be noted that the forward-only pumping model is completely differentiable in distance z. This means that it supports shaping the RA gain over distance, which may be beneficial for systems, where power symmetry across the span is desired. Examples here include optical phase conjugation systems for nonlinearity compensation [30].

Optimization of remote erbium doped fiber-based amplifiers is also of interest for future work, and potentially can be approached with the proposed methods. However, an EDFA model which is differentiable w.r.t. the pump frequency and power would be required, which to our knowledge is currently unavailable.

VI. CONCLUSION

Gradient-based optimization was proposed for the frequency and power of the pumps in Raman amplifiers. In the forward pumping case, a fully differentiable amplification model was

proposed relying on machine learning (ML)-based nonlinear interpolation of the Raman gain coefficient. In the backward pumping case, a differentiable model is obtained using ML to train a NN representing the on-off gain of the amplifier.

For the forward pumping case, this method is completely flexible in the main system and amplifier parameters and it allows for quick optimization of various system configurations, as exemplified with two different unrepeated systems. 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 noise figure in the model in order to optimize the received SNR profile instead of power similar to [31].

ACKNOWLEDGMENT

REFERENCES

[1] C. Headley and G. P. Agrawal, *Raman Amplification in Fiber Optical Communication Systems*. Elsevier, 2005.

[2] B. J. Puttnam, R. S. Luis, G. Rademacher, M. Mendez-Astudilio, Y. Awaji, and H. Furukawa, "S, C and extended L-Band transmission with doped fiber and distributed Raman amplification," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, 2021, pp. 1–3.

[3] U. C. De Moura, M. A. Iqbal, M. Kamalian, L. Krzczanowicz, F. Da Ros, A. M. R. Brusin, A. Carena, W. Forysiak, S. Turitsyn, and D. Zibar, "Multi–band programmable gain Raman amplifier," *Journal of Lightwave Technology*, vol. 39, no. 2, pp. 429–438, 2020.

[4] T. Sutili, P. F. P. Neto, F. D. Simoes, S. M. Rossi, G. J. Suzigan, and R. C. Figueiredo, "Unrepeatered optical systems design based on iterative channel pre-emphasis optimization," in *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2021, pp. 1–6.

[5] T. Sutili, P. F. P. Neto, F. D. Simoes, G. J. Suzigan, and R. C. Figueiredo, "Cost-effective solution for high-capacity unrepeatered transmission," in *2020 Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, pp. 1–3.

[6] J. C. S. S. Januário, A. Chiuchiarelli, S. M. Rossi, J. H. Cruz, J. D. Reis, C. Mornatta, A. Festa, and A. C. Bordonalli, "System design for high-capacity unrepeatered optical transmission," *Journal of Lightwave Technology*, vol. 37, no. 4, pp. 1246–1253, 2019.

[7] G. C. M. Ferreira, S. P. N. Cani, M. J. Pontes, and M. E. V. Segatto, "Optimization of distributed Raman amplifiers using a hybrid genetic algorithm with geometric compensation technique," *IEEE Photonics Journal*, vol. 3, no. 3, pp. 390–399, 2011.

[8] B. Neto, A. L. J. Teixeira, N. Wada, and A. P. S, "Efficient use of hybrid genetic algorithms in the gain optimization of distributed Raman amplifiers," *Optics Express*, vol. 15, no. 26, pp. 17 520–17 528, 2007.

[9] G. Borraccini, S. Staullu, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Cognitive Raman amplifier control using an evolutionary optimization strategy," *IEEE Photonics Technology Letters*, vol. 34, no. 4, pp. 223–226, 2022.

[10] F. L. Della Lucia, J. H. C. Júnior, T. Sutili, and R. C. Figueiredo, "Natural computing algorithms for optimization of high-order distributed Raman amplifiers," in *2019 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*. IEEE, 2019, pp. 1–3.

[11] U. C. de Moura, F. Da Ros, A. M. R. Brusin, A. Carena, and D. Zibar, "Experimental characterization of Raman amplifier optimization through inverse system design," *Journal of Lightwave Technology*, vol. 39, no. 4, pp. 1162–1170, 2021.

[12] M. Soltani, F. Da Ros, A. Carena, and D. Zibar, "Spectral and spatial power evolution design with machine learning-enabled Raman amplification," *Journal of Lightwave Technology*, 2022.

[13] J. Chen and H. Jiang, "Optimal design of gain-flattened Raman fiber amplifiers using a hybrid approach combining randomized neural networks and differential evolution algorithm," *IEEE Photonics Journal*, vol. 10, no. 2, pp. 1–15, 2018.

[14] J. Zhou, J. Chen, X. Li, G. Wu, Y. Wang, and W. Jiang, "Robust, compact, and flexible neural model for a fiber Raman amplifier," *Journal of Lightwave Technology*, vol. 24, no. 6, pp. 2362–2367, 2006.

[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.

[16] M. Soltani, F. Da Ros, A. Carena, and D. Zibar, "Inverse design of a Raman amplifier in frequency and distance domains using convolutional neural networks," *Opt. Lett.*, vol. 46, no. 11, pp. 2650–2653, Jun 2021.

[17] C. Mineto, L. S. Schanner, L. G. M. Riveros, J. H. Da Cruz, F. D. Simões, T. Sutili, F. B. F. Lopes, R. C. Figueiredo, and E. Conforti, "Performance of artificial-intelligence-based modelling for distributed Raman amplification," in *2021 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*. IEEE, 2021, pp. 1–3.

[18] 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, 2020.

[19] A. M. R. 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.

[20] M. Yankov, D. Zibar, A. Carena, and F. Da Ros, "Forward Raman amplifier optimization using machine learning-aided physical modeling," in *accepted, Optoelectronics and Communications Conference (OECC)*, 2022.

[21] I. Roberts, J. M. Kahn, J. Harley, and D. W. Boertjes, "Channel power optimization of WDM systems following gaussian noise nonlinearity model in presence of stimulated Raman scattering," *Journal of Lightwave Technology*, vol. 35, no. 23, pp. 5237–5249, 2017.

[22] G. P. Agrawal, *Fiber Optic Communication Systems*. John Wiley & Sons, Ltd, 2010, ch. 2, pp. 24–78.

[23] X. Liu and B. Lee, "Effective shooting algorithm and its application to fiber amplifiers," *Optics Express*, vol. 11, no. 12, pp. 1452–1461, 2003.

[24] Q. Lin and G. P. Agrawal, "Raman response function for silica fibers," *Optics Letters*, vol. 31, no. 21, pp. 3086–3088, 2006.

[25] 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, pp. 2886–2892, 2002.

[26] D. Semrau, E. Sillekens, P. Bayvel, and R. I. Killey, "Modeling and mitigation of fiber nonlinearity in wideband optical signal transmission [invited]," *Journal of Optical Communications and Networking*, vol. 12, no. 6, pp. C68–C76, 2020.

[27] A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrat, G. Grammel, G. Galimberti, and V. Curri, "GNPy: an open source application for physical layer aware open optical networks," *Journal of Optical Communications and Networking*, vol. 12, no. 6, pp. C31–C40, 2020.

[28] M. P. Yankov, U. C. de Moura, and F. Da Ros, "Power evolution modeling and optimization of fiber optic communication systems with EDFA repeaters," *Journal of Lightwave Technology*, vol. 39, no. 10, pp. 3154–3161, 2021.

[29] H. Bissessur, C. Bastide, S. Etienne, and S. Dupont, "24 Tb/s unrepeatered C-band transmission of real-time processed 200 Gb/s PDM-16-QAM over 349 km," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE, 2017, pp. 1–3.

[30] P. Minzioni, "Nonlinearity compensation in a fiber-optic link by optical phase conjugation," *Fiber and Integrated Optics*, vol. 28, no. 3, pp. 179–209, 2009.

[31] M. P. Yankov, P. M. Kaminski, H. E. Hansen, and F. Da Ros, "SNR optimization of multi-span fiber optic communication systems employing EDFAs with non-flat gain and noise figure," *Journal of Lightwave Technology*, vol. 39, no. 21, pp. 6824–6832, 2021.