Iterative supervised learning approach using transceiver bit-error-rate measurements for optical line system optimization

An Iterative Supervised Learning Approach using Transceiver Bit-Error-Rate Measurements for Optical Line System Optimization

GIACOMO BORRACCINI[1*], ANDREA D'AMICO[1], STEFANO STRAULLU[2], FEHMIDA USMANI[1,3], ARSALAN AHMAD[3], AND VITTORIO CURRI[1]

[1] Department of Electronics and Telecommunications, Politecnico di Torino, corso Duca degli Abruzzi 24, Torino (TO), 10129, Italy
[2] LINKS Foundation, via Pier Carlo Boggio 61, Torino (TO), 10138, Italy
[3] National University of Sciences & Technology (NUST), Islamabad, Pakistan
[*] Corresponding author: giacomo.borraccini@polito.it

Compiled January 20, 2023

Defining the working points of optical amplifiers is a key factor when managing optical networks, particularly for the quality of transmission (QoT) of deployed connections. However, given the lack of knowledge of physical layer parameters, in many cases operators use these infrastructures sub-optimally. In this work, a methodology is presented that optimizes the QoT of an optical line system (OLS) by setting the working points of the erbium-doped fiber amplifiers (EDFAs), by analysis of simulations that use synthetic data derived from experimental characterization of commercial devices. The procedure is divided into three phases: a physical layer characterization, a design process, and an iterative supervised learning approach. Within the first phase, a novel amplifier physical layer characterization is used, exploiting a simple EDFA model that allows an efficient estimation of the OLS behaviour, knowing only the setting operative ranges of the devices. The results show that the satisfactory outcome produced during the design phase is further improved by the iterative supervised learning approach. The latter approach is implemented for single OLSs between couples of adjacent re-configurable optical add & drop multiplexers (ROADMs), each equipped with a certain set of transceivers (TRXs), enabling the QoT estimation of the specific OLS.

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

Optical system optimization is a fundamental procedure in efficient and cost-effective utilization of network infrastructures, and it has been the focus of numerous research activities in the optical network framework [1–3]. This is mainly due to the increasing interest of operators to sustain the growing traffic demand experienced in the latest years [4, 5]. The level of robustness of an optical network is in some way correlated both to its scalability in managing more connection requests and the capability of the system to prevent or counteract failures. Component aging, fiber cuts or the modification of a patch panel at the beginning of a fiber span are the most common examples of events affecting the system performance. From this perspective, the optimization of the infrastructure depends upon the knowledge of the physical layer in terms of network element features, such as optical amplifiers and fibers that compose the various optical line systems (OLSs) of a network.

Recently, various machine learning (ML) algorithms have shown outstanding results in optical communication and have been used for various applications in this field, in particular for quality of transmission (QoT) estimation improvements [6–8]. Several implementations have been used for QoT predictions and simulations of OLSs. In particular, the proposed solutions have been based on both stand-alone ML frameworks or ML-aided QoT estimator relying on analytical models [9, 10]. All of these solutions require resources and time to acquire large datasets. Furthermore, physical layer parameters used for feeding analytical models are rarely available, presenting a not negligible uncertainty [11]. The authors already tackled this issue providing methodologies that aim to automatically characterize the system, but this is limited to the optical fiber parameters [12–14].

In this work, a methodology is presented that optimizes the QoT of an OLS by setting the working points of the EDFAs. The procedure is divided into three phases: a physical layer characterization, a design process, and an iterative supervised learning approach. In the first phase, a novel amplifier physical layer characterization exploiting a simple EDFA model that allows an efficient estimation of the OLS behaviour is used. The generalized signal-to-noise ratio (GSNR), which includes both the amplified spontaneous emission (ASE) and nonlinear interference (NLI) impairments, has been considered as a fundamental metric for the system QoT [15]. The aim of the developed methodology is to optimize the OLS QoT by automatically characterizing both the optical fibers and the EDFA of an OLS, relying only on the setting operative ranges of the EDFAs, such as the maximum/minimum values of the input parameters and the maximum output power. Furthermore, a ML agent is employed to improve the overall QoT estimation, enabling a more effective optimization of the OLS under investigation through the mitigation of deviations from the model. This proof of concept is given via simulation using synthetic data retrieved from an experimental campaign aimed at characterizing and modelling an EDFA, and an experimentally validated QoT estimation framework [16]. In particular, the reference OLS utilized to test the proposed methodology is created using an accurate ML model of a commercial EDFA based on an experimental dataset.

The proposed iterative supervised learning approach is able to achieve a remarkable improvement with a significantly limited number of the training data samples (i.e. just a few cases of actual QoT performance using four selected channels). The optical network architecture considered in this scenario consists of single OLSs between couples of adjacent re-configurable optical add & drop multiplexers (ROADMs), each equipped with a certain set of transceivers (TRXs), enabling QoT estimation for each specific OLS for a given amplifier configuration setting. In the transmission scenario under investigation, the proposed solution provides a 0.3 dB improvement on the mean GSNR value and a reduction from 0.11 to 0.03 dB in the GSNR standard deviation over all the propagated channels with respect to the amplifiers' configuration optimized after the physical layer characterization, achieving a significant leveling of the QoT over the entire band.

The article is divided as follows: Sec. 2 presents the network architecture and the OLS under investigation. Sec. 3 reports the experimental procedure to characterize a commercial EDFA and then to built the corresponding ML model based on the decoupling between operative settings and gain/noise fluctuations. This EDFA ML model is implied in the simulation to built the reference OLS. Sec. 4 introduces the developed methodology describing the three phases in succession. In Sec. 5 the simulation framework is explained and the obtained results are analyzed and discussed. Sec. 6 reports the conclusion of the work.

2. OPTICAL NETWORK ARCHITECTURE

The OLS has been considered as a sub-system of an open and partially disaggregated optical network, shown in Fig. 1. In this framework, the optical network controller (ONC) is assumed to have direct access to the TRXs and ROADMs for route-wavelength assignment and lightpath deployment purposes. Whereas, the management of the OLSs, identified by the ROADM-to-ROADM physical connections including boosters and pre-amplifiers, is delegated to the optical line controllers (OLCs), which have direct access to the telemetry from the available monitoring devices. Moreover, each OLC autonomously evaluates and sets the optimized working points of the specific OLS amplifiers, providing to the ONC the QoT metrics required for the lightpath computations.

The developed framework is based on a further assumption: each OLS is equipped with a certain set of TRXs, enabling QoT estimation in different portions of the propagating spectrum by means of the bit-error-rate (BER) measurement for a given amplifier configuration setting. This assumption reflects the reality in the case of optical networks with linear topology or in which there is the need to route connections between adjacent nodes.

As a consequence, the reference transmission scenario has been set as a fully-loaded spectrum composed of 70 channels at 32 GBd and 50 GHz fixed spacing centred in the C-band, with an equalized uniform power at the input of the booster (0.1 dBm total power). The optimization target has been defined as the maximum average GSNR over all the channels, achievable with a limited GSNR variation of each channel in order to obtain a minimal QoT complexity.

Regarding the telemetry required for the proposed implementation, each amplification site is equipped with an optical time domain reflectometer (OTDR) and a couple of optical channel monitors (OCMs) at the input and the output of the device in order to reproduce the same conditions presented in [13]. Moreover, the presented iterative approach is based on the QoT feedback retrieved from four fixed channels under test (CUTs), equally spaced within the overall spectrum, which are assumed to be propagated exclusively on the specific OLS under investigation.

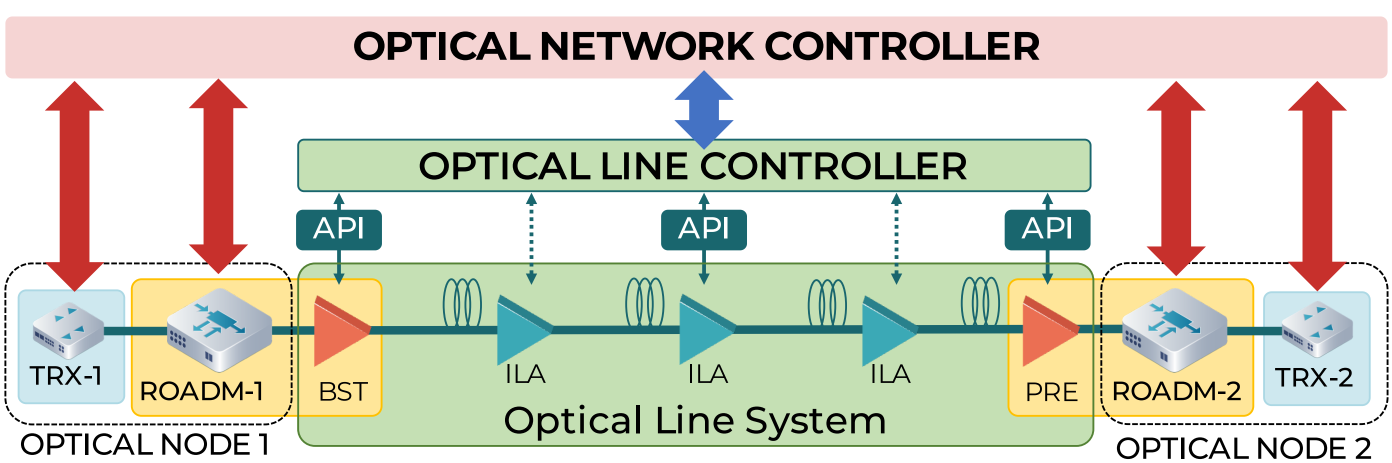

Fig. 1. ROADM-to-ROADM optical line system within an open and partially disaggregated optical network architecture.

3. EDFA MACHINE LEARNING MODEL

First, an EDFA realistic model both in terms of gain and ASE noise has been made by means of a ML technique in order to define the reference OLS of the simulation framework. In particular, a dataset has been collected performing a full spectral load characterization of a commercial EDFA with maximum output power of 20 dBm and maximum gain of 20 dB. A large and valuable research has been done on the EDFA modelling using ML techniques [17–19]. In this work, the ML model is adapted to be complementary with respect to the EDFA simple model proposed in the physical layer characterization.

The experimental setup is depicted in Fig. 2. A commercial wave shaper filter (1000S from Finisar) is programmed to shape the output of an ASE noise source generating a C-band wavelength division multiplexing (WDM) comb centered at 193.5 THz and composed by 38 channels, 100-GHz spaced, with 32 GHz of bandwidth each. The WDM comb is introduced at the EDFA under test input with 9 different overall power values, ranging from -10 up to $+6$ dBm. Working in constant gain mode, the EDFA under test is set in different conditions (i.e. the optical gain and the total tilt settings); the gain parameter, G, ranges from 14 to 20 dB (1 dB step), whereas the total tilt parameter, T, (gain difference between the extreme channels) ranges from -5 to $+5$ dB (1 dB step). For each combination of input power, EDFA gain and tilt values, the WDM comb spectrum is captured both at the EDFA's input and output by means of an optical spectrum analyzer (OSA) and integrated photodiodes. The OSA resolution bandwidth was set to 10 GHz in order to appreciate both the signal peaks and the noise level. All the measurements performed with the EDFA in saturation condition are removed from the dataset as they introduce an ambiguity in the relation between target parameters and the actual EDFA output. Moreover, knowing the specifications of the in-field amplifiers, the saturation condition is generally avoided in real-case scenarios with the design phase.

The gain profiles, $G(f)$, are evaluated from the difference between the input and the output power peak profiles. Thanks to the internal feedback mechanism of the amplifier, the G setting parameter corresponds exactly to the difference between the total input and output power values measured by the integrated photodiodes. The gain ripple profile, $\Delta G(f)$, predicted by the ML model is defined in logarithmic units as:

$$\Delta G(f) = G(f) - \left[G + \frac{T}{B}(f - f_0)\right] \tag{1}$$

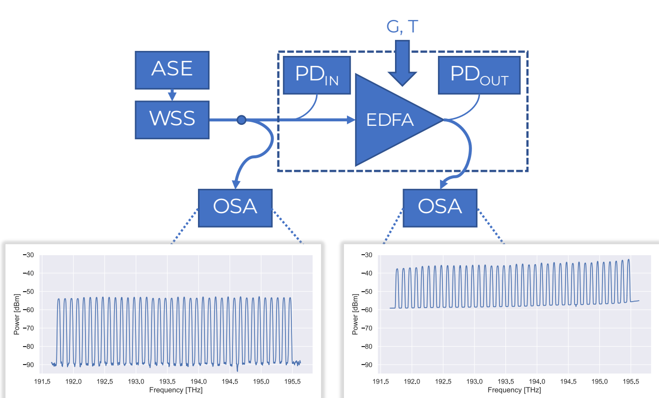

Fig. 2. Experimental setup sketch for the characterization a commercial EDFA and measurement examples.

where B and f_0 are the band on which the tilt is applied and the central frequency of the gain profile, respectively, expressed in Hz. These two parameters are fitted in advance minimizing the root-mean-squared error (RMSE) between all the measured gain profiles and the profiles obtained using this linear expression, $\tilde{G}(f)$, properly considering the corresponding G and T parameters:

$$\mathrm{RMSE}\left(X^p, X^m\right) = \sqrt{\frac{\sum_{i=0}^{N}\left(X_i^p - X_i^m\right)^2}{N}}\,, \tag{2}$$

where X^p and X^m are the arbitrary predicted and the measured profiles, respectively, and N is the total number of frequency samples.

For each amplifier configuration, the evaluated ASE power spectral density (PSD) profile, $\mathrm{PSD}_{\mathrm{ASE}}(f)$, is reported at the amplifier input subtracting the corresponding evaluated gain profile in logarithmic units. Then, the residual noise of the input source is removed from the profile by taking the difference between the two quantities in linear units [20]. The ASE power spectral density ripple, $\Delta\mathrm{PSD}_{\mathrm{ASE}}(f)$, is evaluated in logarithmic units as:

$$\Delta\mathrm{PSD}_{\mathrm{ASE}}(f) = \mathrm{PSD}_{\mathrm{ASE}}(f) - \overline{\mathrm{PSD}_{\mathrm{ASE}}}\,, \tag{3}$$

where $\overline{\mathrm{PSD}_{\mathrm{ASE}}}$ is the profile average, which is stored in a table according to the corresponding amplifier configuration and used in order to perform the overall prediction within amplifier EDFA model.

On the basis of the described dataset, the ML technique exploits two different artificial neural networks (ANNs), to predict the ripples of both the gain and ASE PSD profiles, respectively. To determine the appropriate configurations of the ANN model in order to minimize the complexity, extensive simulations have been performed by changing the ANN parameters, such as number of layers, neurons, epochs, batch size and types of activation function. Each ANN is implemented using the open source library TensorFlow© [21], and consist of one input layer, one hidden layer with 256 neurons, and one output layer. Moreover: a ReLU-based activation function is used for all the neurons to avoid the vanishing gradient problem and the adaptive moment estimation (Adam) optimizer and the RMSE metric are used to optimize and evaluate the model, which is trained on 5000 epochs with a batch size of 64. The input feature space for both gain and ASE PSD ripple estimation includes the amplifier setting parameters, G and T, and the input and output total powers. For a given pattern of features, the predicted label values of the two ripple profiles are related to 38 C-band frequencies fixed in the experimental setup. Then, the two predicted ripples profiles are linearly interpolated over the propagated spectrum frequencies. Moreover, the ASE PSD average is linearly interpolated according to the amplifier settings.

For both the ripples, the dataset consists in 510 samples and the model is validated splitting it in a proportion 90-10% for training and testing, respectively. The goodness of the prediction is estimated in terms of RMSE, maximum absolute error (MAE) and error considering the test dataset over the entire spectrum of each sample (Fig. 3). The MAE is expressed by the following definition:

$$\mathrm{MAE}(X^p, X^m) = \max\left(\left|X_i^p - X_i^m\right|\right) \quad \forall\, 1 \le i \le N\,. \tag{4}$$

Observing the statistics related to the gain ripple prediction, the average value for both the RMSE and the MAE is below

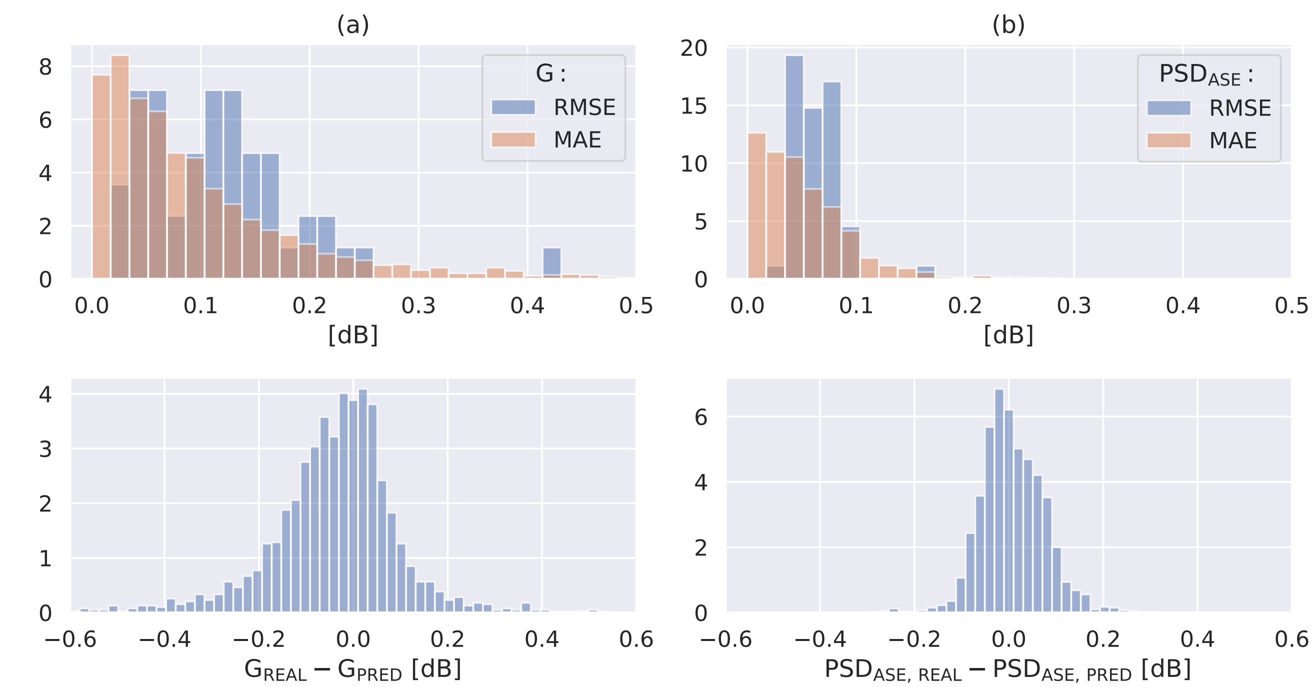

Fig. 3. EDFA ML model testing results: RMSE and MAE distributions, and error distribution between the measured profile and the predicted profile for the gain (a-column) and the ASE PSD (b-column).

0.2 dB. The maximum error values are 1.4 dB and −0.9 dB, both recorded in rarely used amplifier configurations in which the tilt orientation enhances the inter-channel Raman scattering. Similarly, considering the ASE PSD ripple predictions, the average value for both the RMSE and the MAE is below 0.1 dB. For both the gain and ASE PSD ripple profiles, the error distributions are concentrated into a dense zone of values around zero mean. Some examples of the ripples produced by the ML EDFA model are shown in Fig. 4.

4. METHODOLOGY

From the point of view of cognitive optical networks [22], the proposed methodology aims to optimize the working points of an OLS maximizing and leveling the QoT for all the modulated channels composing the WDM comb, in order to guarantee the same performance for whatever traffic connection allocated between a couple of nodes of the optical infrastructure passing through the specific OLS. This procedure involves the use of a set of TRXs, transmitting as many channels equally spaced in frequency, properly characterized in back-to-back (B2B) configuration in order to retrieve the SNR estimation on the basis of the measured BER [23]. In the following, the GSNR derived from the in-field measured BER is referred to this conversion using the B2B TRX characterization of all these CUTs. Furthermore, the adopted terminology is based on the use of three key words. The word *model* refers to the physical layer model built while performing the proposed methodology, which emulates the behavior of a given system. Then, the word *reference* means the system against which the model is compared. It can be a real system when referring to an experiment or a virtual object

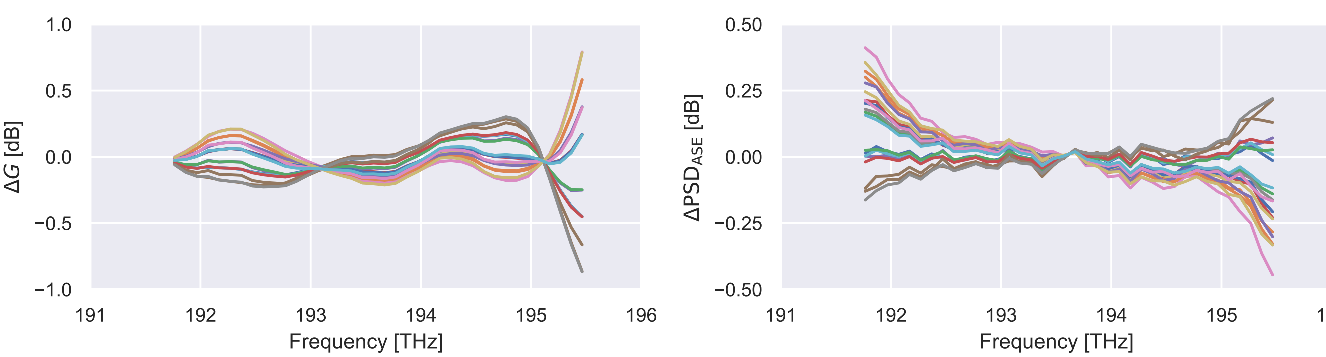

Fig. 4. Examples of $\Delta G(f)$ and $\Delta \mathrm{PSD}_{\mathrm{ASE}}(f)$ profiles produced by the ML EDFA model, fixing the total input power at -4.0 dBm and randomizing the values of gain and tilt with the corresponding operative ranges.

in a simulation. Finally, the word *prediction* concerns the result obtained by the ML agent's elaboration in the third step of the methodology.

The algorithmic process that brings the OLS to be operative is articulated into three phases, which will be described in detail in the following. The block diagrams of each of them are depicted in Fig. 5. In the first step, generally called *physical layer characterization*, the in-field telemetry apparatus is exploited to probe the status of the system in a specific working condition and the measurements are processed in order to retrieve the equivalent physical layer model parameters of the device under test. In this work, only the novel EDFA physical layer characterization is presented, assuming that the fiber characterization, which is independent from the installed EDFA characteristics, is performed as preliminary step [12]. This procedure brings to the full virtualization of each fiber span of an OLS, providing the estimation of the fiber length, L_S, the loss coefficient function, $\alpha(f)$, the Raman efficiency curve, $C_R(\Delta f)$, and the lumped losses, $l(z)$.

Due to the problem complexity, the optimization algorithm adopted within all the steps of the procedure is the covariance matrix adaptation evolution strategy (CMA-ES) [24], a stochastic optimization algorithm based on an evolutionary strategy.

Regarding the adopted physical layer model for QoT estimation, in the built software framework, GNPy open source Python library [25] is used to emulate the optical propagation using two main classes: the optical fiber and the propagating WDM comb.

A. EDFA Physical Layer Characterization

A simple physical layer EDFA model is considered, describing the gain profile, $G(f)$, in logarithmic as:

$$G(f) = G + \frac{T}{B}(f - f_0) \ , \qquad \textbf{(5)}$$

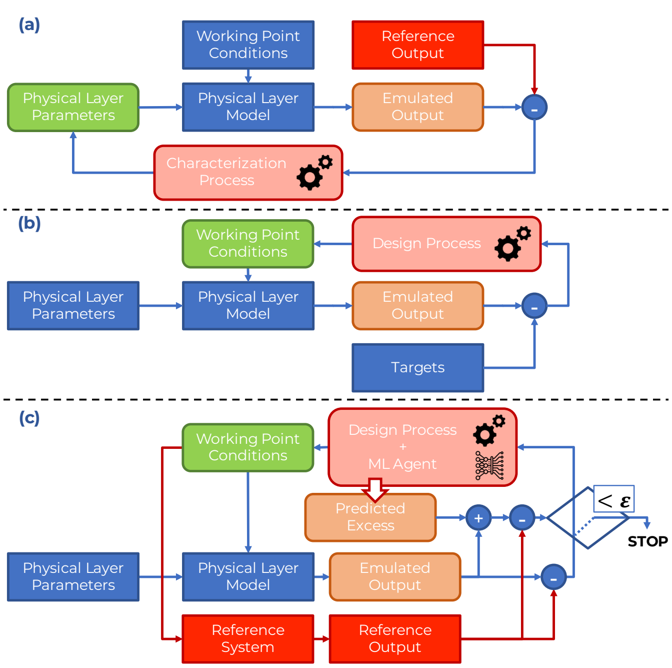

Fig. 5. Methodology conceptual block schemes: (a) physical layer characterization, (b) working point optimization relying on the retrieved physical layer model, (c) working point refinement using an iterative supervised learning approach.

where all the parameters are graphically represented in Fig. 6, and the introduced ASE noise profile, $P_{\text{ASE}}(f)$, in linear units as:

$$P_{\text{ASE}}(f) = h\, f\, B_n\, \left[G(f) - 1\right]\, \text{NF} \ , \qquad \textbf{(6)}$$

where h is the Planck constant, B_n is the noise bandwidth and NF is the noise figure, constant and fixed for all the gain values. The amplifier settings coincide with the parameters G and T, while B, f_0 and NF represent the physical layer parameters of the EDFA to be probed. In particular, B and f_0 parameters define the slope of the tilt and where it is applied within the amplification band. The NF parameter is used to extract the average level of ASE noise introduced by the EDFA in a specific working condition.

All the EDFAs within the OLS are set in a known working condition. Exploiting the telemetry, it is assumed that the system is set in transparency mode, choosing the gain value to restore the required output power and the tilt values to balance the flatness of the channel power profile. With this configuration, the input spectrum is propagated through the reference OLS, collecting at the output GSNR for the complete set of available TRXs, $\text{GSNR}(f)^r$.

The characterization process aims to jointly retrieve the three EDFA model physical layer parameters for all the EDFAs within the OLS model observing the emulated, $\text{GSNR}(f)^e$, and the reference, $\text{GSNR}(f)^r$, GSNR profiles under the defined working conditions, as stated by the following objective function:

$$\min_{\forall\, B,\, f_0,\, \text{NF}} \text{RMSE}\left(\text{GSNR}(f)^e, \text{GSNR}(f)^r\right) \ , \qquad \textbf{(7)}$$

where f are referred to the frequencies of the CUTs. The optimization problem presents a dimension of $3 \times \text{N}_{\text{EDFA}}$, where N_{EDFA} is the number of EDFAs in the OLS.

The main advantage of this novel EDFA characterization process is that it allows to fully abstract the OLS only knowing the setting operative ranges, without performing any in-laboratory experimental characterization for these lumped components.

B. Working Point Design Phase

At this point, the OLS full abstraction has been achieved, obtaining an equivalent representation of all the fibers and EDFAs. On the basis of the previous characterization, the abstracted OLS can be exploited in order to optimize the EDFAs' working point fixing the retrieved physical layer parameters and modifying the amplifiers settings within the optimization process (Fig. 5-b).

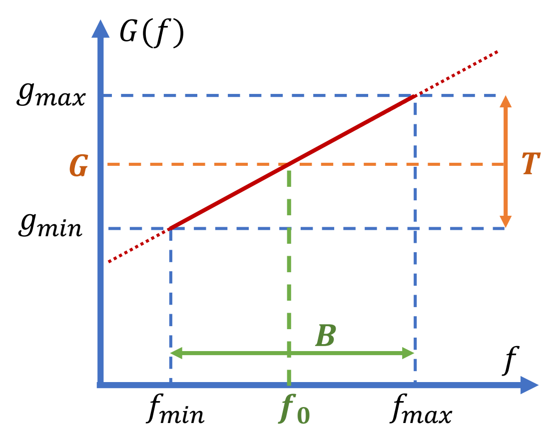

Fig. 6. Qualitative representation of the EDFA simple model gain profile.

Trusting the performed characterization accuracy, the design process manipulates the OLS model in order to evaluate which is the amplifier setting configuration that produces the highest and most homogeneous QoT at the output [13]. Assuming two setting parameters for a single EDFA (G and T), the problem dimension amount to $2 \times \mathrm{N}_{\mathrm{EDFA}}$. The objective function adopted to achieve this purpose is:

$$\max_{\forall\, G,T} \overline{\mathrm{GSNR}} - \sigma_{\mathrm{GSNR}} \;, \tag{8}$$

where $\overline{\mathrm{GSNR}}$ and σ_{GSNR} are the mean and the standard deviation of the emulated output GSNR profile, respectively. This problem formulation addresses the choice of the amplifier working point observing the overall QoT estimated at the line output, aiming to achieve the maximum average value for all the channels composing the spectrum with the minimum profile dispersion.

C. Iterative Supervised Learning Refinement

The result achieved following the procedure until this point could be sufficient to make the OLS ready to start with standard transmission operations. Ignoring the uncertainties related to the knowledge of the physical model of the fibers, the TRXs and the ROADMs penalties, the relevant aspect of the adopted EDFA model is that it does not consider any ripple in the gain and ASE noise profile and any dependency of the noise figure with respect to the set gain. In order to properly mitigate the EDFA gain and noise ripple inaccuracies, the methodology ends with a refinement phase of the EDFAs' working point using an iterative supervised learning approach (Fig. 5-c, Alg. 1).

Algorithm 1. Amplifier working point refinement using the iterative supervised learning approach (pseudo-code)

```
procedure
    Retrieving GSNR(f)^r
    Evaluating MAE(GSNR(f)^p, GSNR(f)^r)
    if MAE(GSNR(f)^p, GSNR(f)^r) > ε then
        Dataset initialization
        while MAE(GSNR(f)^p, GSNR(f)^r) > ε do
            ML agent training
            Design process including the ML agent
            Retrieving GSNR(f)^r
            Evaluating MAE(GSNR(f)^p, GSNR(f)^r)
            Evaluating a new sample: GSNR(f)^e − GSNR(f)^r
            Dataset update
```

A ML agent that maps the difference between the emulated and the reference output GSNR profiles, ΔGSNR, according to the corresponding complete EDFA configuration is created by means of a neural network. An initial dataset is collected both emulating the system behaviour through the OLS model and obtaining the reference GSNR profile using configurations similar to the optimal working point defined in the previous step. The design framework is integrated with a single additional ANN and inserted within a larger loop in which the problem dimension and the objective function remain unvaried but the evaluated GSNR profile is expressed as:

$$\mathrm{GSNR}(f)^p = \mathrm{GSNR}(f)^e + \Delta\mathrm{GSNR}(f) \;, \tag{9}$$

where $\mathrm{GSNR}(f)^p$ is the overall predicted GSNR profile, $\mathrm{GSNR}(f)^e$ is the GSNR profile emulated by the model and

Table 1. Results of the EDFA Physical Layer Characterization

	B [THz]	f_0 [THz]	NF [dB]
BST	5.49	193.52	4.2
ILA-1	3.50	193.45	4.2
ILA-2	3.50	193.82	4.2
ILA-3	3.68	193.30	4.2
ILA-4	5.49	193.86	4.2
ILA-5	5.49	192.08	4.2
PRE	4.72	192.17	4.3

$\Delta\mathrm{GSNR}(f)$ is the GSNR excess predicted by the ANN for each CUTs. At each iteration, the ANN of the ML agent is trained using the current dataset, the OLS is set according to the actual optimal configuration and the dataset is updated by adding the residue between the emulated and the reference GSNR profiles. The refinement loop ends when the difference between the predicted and the reference GSNR profiles is below a given tolerance, ε:

$$\mathrm{MAE}(\mathrm{GSNR}(f)^p, \mathrm{GSNR}(f)^r) < \varepsilon \;. \tag{10}$$

5. SIMULATION FRAMEWORK & RESULTS

In order to test the proposed methodology, a reference OLS is created using the EDFA ML model produced from the previously described experimental characterization on a commercial EDFA (Sec. 3) and a set of 6 standard single mode fibers characterized following the physical layer characterization described in [12], composing 6 EDFA-fiber span with nominal length of 65 km and an additional EDFA at the end as pre-amplifier. Each fiber span presents a loss coefficient function varying the frequency around 0.19 dB/km, a Raman efficiency scale factor that ranges from 0.38 to 0.44 1/W/km, input and output connector losses ranging from 0.1 to 2.5 dB and the total span loss is varied from 14 to 22 dB. The dispersion, D, is assumed to be 16.7 ps/nm/km for all the fibers. The reference OLS represents in simulation what the behavior of a real system would be and it shares with the OLS model built following the proposed methodology the same fiber physical layer abstractions but differs for the adopted EDFA model. In fact, as described in Sec. 4, the OLS model relies on the use of the EDFA simple model. In this simulation analysis, the mitigation on the QoT produced by the iterative supervised learning approach is limited to effects of the EDFAs' gain ripple and ASE noise.

Regarding the EDFA physical layer characterization, the bounds for the band, the central frequency and the noise figure are 3.5−5.5 THz, 192−194 THz and 4.2−6 dB, respectively. The OLS is set in transparency mode at a total power level of 19 dBm. The characterization process result is reported in Tab 1. On the basis of the produced models, the working point of the EDFAs has been optimized in the design phase posing as bounds for the gain and tilt parameters the ranges 12−20 dB and -5−+5 dB, respectively. The refinement process starts with the initial dataset creation consisting of 5 samples. The latter are obtained through the evaluation of amplifier configurations similar to the optimal one, randomizing each parameter in a range ±0.5 dB. The initial size of the dataset has been voluntarily fixed at a small value

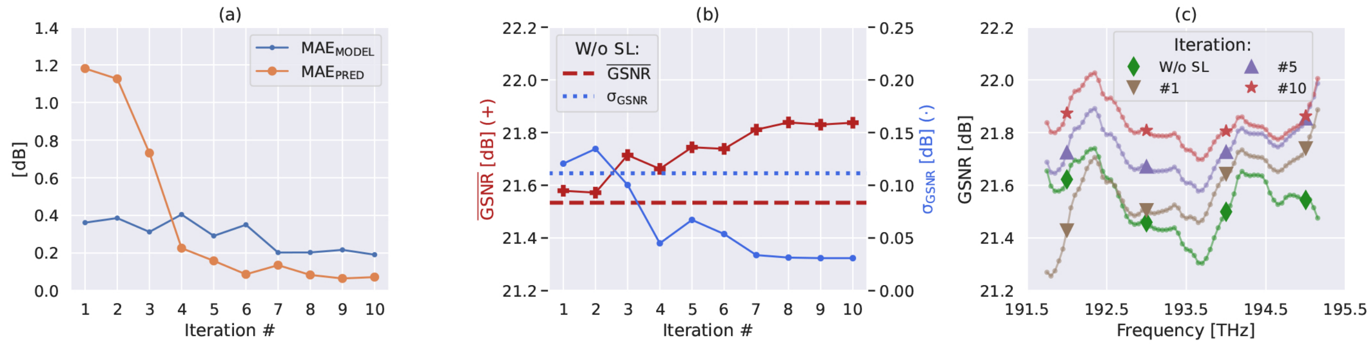

Fig. 7. Simulation results using the refinement with the iterative supervised learning (SL) approach: (a) MAE evolution comparison between the GSNR profiles obtained with the physical layer model only and the combination of the model with the ANN prediction, (b) GSNR aggregated metrics evolution, mean and standard deviation, (c) GSNR profiles at a specified iteration produced with the reference OLS.

Table 2. Results of the Design and Refinement Optimizations

	Design		Refinement	
	G [dB]	T [dB]	G [dB]	T [dB]
BST	16.2	-0.1	17.4	-3.7
ILA-1	14.6	-0.4	13.3	-2.6
ILA-2	14.2	-2.9	16.0	-0.6
ILA-3	16.2	-1.8	15.0	-0.2
ILA-4	15.5	1.8	16.3	-0.2
ILA-5	17.4	2.8	16.0	0.8
PRE	12.1	-2.3	17.6	0.9

in order to stress the convergence of the process. The proposed iterative supervised learning approach is based on the use of a dedicated ANN having the amplifiers gain and tilt parameters as features and the residual GSNR between the reference and the emulated GSNR, $\Delta\text{GSNR}(f)$, as labels. The ANN comprises a single hidden layer with 256 number of neurons (ReLU activation function). An optimization strategy based on Adam is used to update the weights with a batch size of 64.

Fixing $\varepsilon = 0.1$ dB, the refinement process ends after 10 iterations, having the consecutive last 3 iterations below the tolerance. The results, including the comparison with the design outcome, are graphically reported in Fig. 7. The final amplifier working point configuration of both processes are reported in Tab. 2. After the 4-th iteration the ML agent enables the outperforming of the accuracy prediction obtained with the only OLS model (Fig. 7-a). It is remarkable how the latter provides an estimation accuracy below 0.5 dB, proving the goodness of the EDFA characterization. The aggregated metrics of GSNR mean and standard deviation reported in Fig. 7-b show the clear improvement trend, refining the mean of 0.3 dB and bringing the standard deviation from 0.11 dB to 0.03 dB. The performance difference between the configurations obtained through the design process and the refinement is represented in Fig. 7-c in terms of GSNR profile, showing how the reference OLS improves the QoT of all the CUTs.

6. CONCLUSION

In this work, a methodology based on an iterative supervised learning approach for the optimization of the amplifiers' working point of an OLS in a context of cognitive optical networks is presented, proving the effectiveness of each step by means of a simulation campaign. It relies on a novel EDFA physical layer characterization that exploits a simple amplifier model, efficiently reproducing the device behaviour using only device setting parameters. The simulation results obtained through the iterative supervised learning approach in the described highly controlled environment are referred to a mitigation of the QoT limited to the effect of the EDFAs' gain ripple and ASE noise. It is expected that the application of this methodology to a real case will have a much greater impact on performance improvement due to deviations from the model of other components that directly affect the QoT, such as fiber spans, ROADMs and TRXs. In addition, the proposed methodology allows to efficiently cope with the optimization of an optical line system even after a fiber cut, thanks to the adaptability of the procedure.

As further investigations, the authors will stress this approach exasperating the physical layer scenario and analysing the behaviour of the system in presence of strong/soft failures. The final goal is to apply it in a real experimental setup.

REFERENCES

1. B. Mukherjee, "Wdm optical communication networks: progress and challenges," IEEE J. on Sel. Areas communications **18**, 1810–1824 (2000).
2. V. Curri, A. Carena, A. Arduino, G. Bosco, P. Poggiolini, A. Nespola, and F. Forghieri, "Design strategies and merit of system parameters for uniform uncompensated links supporting nyquist-wdm transmission," J. Light. Technol. **33**, 3921–3932 (2015).
3. Y. Pointurier, "Design of low-margin optical networks," J. Opt. Commun. Netw. **9**, A9–A17 (2017).
4. Cisco, "Cisco annual internet report (2018–2023) white paper," (2020).
5. A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez *et al.*, "The lockdown effect: Implications of the covid-19 pandemic on internet traffic," in *Proceedings of the ACM internet measurement conference,* (2020), pp. 1–18.
6. F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," IEEE Commun. Surv. & Tutorials **21**, 1383–1408 (2018).

7. Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," J. Opt. Commun. Netw. **13**, B60–B71 (2021).
8. R. Ayassi, A. Triki, N. Crespi, R. Minerva, and M. Laye, "Survey on the use of machine learning for quality of transmission estimation in optical transport networks," J. Light. Technol. **40**, 5803–5815 (2022).
9. A. D'Amico, S. Straullu, A. Nespola, I. Khan, E. London, E. Virgillito, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Using machine learning in an open optical line system controller," J. Opt. Commun. Netw. **12**, C1–C11 (2020).
10. A. D'Amico, S. Straullu, G. Borraccini, E. London, S. Bottacchi, S. Piciaccia, A. Tanzi, A. Nespola, G. Galimberti, S. Swail, and V. Curri, "Enhancing lightpath qot computation with machine learning in partially disaggregated optical networks," IEEE Open J. Commun. Soc. **2**, 564–574 (2021).
11. F. Usmani, I. Khan, M. U. Masood, A. Ahmad, M. Shahzad, and V. Curri, "Transfer learning aided qot computation in network operating with the 400zr standard," in *2022 International Conference on Optical Network Design and Modeling (ONDM),* (2022), pp. 1–6.
12. G. Borraccini, S. Straullu, A. D'Amico, A. Nespola, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Autonomous physical layer characterization in cognitive optical line systems," in *2021 Optical Fiber Communications Conference and Exhibition (OFC),* (IEEE, 2021), pp. 1–3.
13. G. Borraccini, A. D'Amico, S. Straullu, A. Nespola, S. Piciaccia, A. Tanzi, G. Galimberti, S. Bottacchi, S. Swail, and V. Curri, "Cognitive and autonomous qot-driven optical line controller," J. Opt. Commun. Netw. **13**, E23–E31 (2021).
14. G. Borraccini, S. Staullu, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Cognitive raman amplifier control using an evolutionary optimization strategy," IEEE Photonics Technol. Lett. **34**, 223–226 (2022).
15. V. Curri, "Gnpy model of the physical layer for open and disaggregated optical networking," J. Opt. Commun. Netw. **14**, C92–C104 (2022).
16. G. Borraccini, S. Straullu, A. Giorgetti, R. D'Ingillo, D. Scano, A. D'Amico, E. Virgillito, A. Nespola, N. Sambo, F. Cugini *et al.*, "Qot-driven optical control and data plane in multi-vendor disaggregated networks," in *Optical Fiber Communication Conference,* (Optica Publishing Group, 2022), pp. M4F–5.
17. Y. You, Z. Jiang, and C. Janz, "Machine learning-based edfa gain model," in *2018 European Conference on Optical Communication (ECOC),* (IEEE, 2018), pp. 1–3.
18. A. Mahajan, K. Christodoulopoulos, R. Martínez, S. Spadaro, and R. Muñoz, "Modeling edfa gain ripple and filter penalties with machine learning for accurate qot estimation," J. Light. Technol. **38**, 2616–2629 (2020).
19. J. Yu, S. Zhu, C. L. Gutterman, G. Zussman, and D. C. Kilper, "Machine-learning-based edfa gain estimation," J. Opt. Commun. Netw. **13**, B83–B91 (2021).
20. D. M. Baney, P. Gallion, and R. S. Tucker, "Theory and measurement techniques for the noise figure of optical amplifiers," Opt. fiber technology **6**, 122–154 (2000).
21. https://www.tensorflow.org/.
22. W. Wei, C. Wang, and J. Yu, "Cognitive optical networks: key drivers, enabling techniques, and adaptive bandwidth services," IEEE Commun. magazine **50**, 106–113 (2012).
23. G. Borraccini, S. Straullu, A. Ferrari, E. Virgillito, S. Bottacchi, S. Swail, S. Piciaccia, G. Galimberti, G. Grammel, and V. Curri, "Using qot-e for open line controlling and modulation format deployment: an experimental proof of concept," in *2020 European Conference on Optical Communications (ECOC),* (IEEE, 2020), pp. 1–4.
24. N. Hansen, Y. Akimoto, and P. Baudis, "CMA-ES/pycma on Github," Zenodo, DOI:10.5281/zenodo.2559634 (2019).
25. A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrát, G. Grammel, G. Galimberti, and V. Curri, "Gnpy: an open source application for physical layer aware open optical networks," J. Opt. Commun. Netw. **12**, C31–C40 (2020).