



**HAL**  
open science

## Dynamic power implementation with reinforcement learning for excursion-free amplified optical systems

Maria José Freire Hermelo, Dipankar Sengupta, Antoine Lavignotte, Christine Tremblay, Catherine Lepers

### ► To cite this version:

Maria José Freire Hermelo, Dipankar Sengupta, Antoine Lavignotte, Christine Tremblay, Catherine Lepers. Dynamic power implementation with reinforcement learning for excursion-free amplified optical systems. 24th International Conference on Optical Network Design and Modelling, May 2020, Barcelona (Spain), Spain. hal-04312757

**HAL Id: hal-04312757**

**<https://hal.science/hal-04312757>**

Submitted on 28 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Dynamic Power Predistortion Implementation with Reinforcement Learning for Excursion-Free Amplified Optical Systems

Maria Freire Hermelo  
SAMOVAR, Télécom SudParis, CNRS  
Institut Polytechnique de Paris  
Évry, France

Dipankar Sengupta  
Network Technology Lab  
École de technologie supérieure  
Montreal, Canada  
dipankar.sengupta@etsmtl.ca

Antoine Lavignotte  
Télécom SudParis  
Institut Polytechnique de Paris  
Évry, France  
antoine.lavignotte@telecom-sudparis.eu

Christine Tremblay  
Network Technology Lab  
École de technologie supérieure  
Montreal, Canada  
christine.tremblay@etsmtl.ca

Catherine Lepers  
SAMOVAR, Télécom SudParis, CNRS  
Institut Polytechnique de Paris  
Évry, France  
catherine.lepers@telecom-sudparis.eu

**Abstract**—In optical networks, wavelength dependent gain in optical amplifiers, resulting in power excursion at their output, is still a challenge to address. Current proposed solutions are based on specific wavelength assignment, which can go against the wavelength continuity constraint in transparent optical networks. In this paper, we propose a dynamic power predistortion implementation using reinforcement learning to counter this restriction. We show that a dynamic adjustment of input optical power per each channel permits to reduce the mean power excursion at the output of the optical line system.

**Index Terms**—power predistortion, optical power excursion, reinforcement learning, optical amplifier

## I. INTRODUCTION

In optical networks, dealing with the wavelength dependent gain characteristics of optical amplifiers is still an issue. Producing undesired optical mean power excursions, its effect in wavelength division multiplexing (WDM) systems can be significant. As a result, nonlinearities and Optical Signal to Noise Ratio (OSNR) degradation are observed [1].

Gain equalization techniques mainly developed for static channel load conditions, become unsatisfactory in more dynamic conditions where added and dropped channels change the amplifier operating point [1].

New approaches to reduce power excursions have been investigated based on the knowledge of the behaviour of an amplifier or cascade of amplifiers obtained through amplifier modelling, experimental characterization or data collection [2]. These approaches have been used during wavelength assignment process making wavelength recommendations for add or drop channels [3]. Control of the amplifier operating point has also been demonstrated to be efficient for reducing power excursions [4]. Distributing the signal power on two different wavelengths has also been shown to be efficient to

cancel power discrepancies [5]. Signal power predistortion at the transmitter is generally computationally complex and time consuming. More recently, a machine learning approach to preadjust the input optical channel power in a single step has been proposed for reducing the power discrepancy through a defragmentation process [6].

In this paper, we propose, for the first time to our knowledge, a dynamic power predistortion implementation using reinforcement learning (RL) for self-control of channel power excursions in Erbium-doped fiber amplifiers (EDFAs) through small power adjustment steps. Mean optical power excursion reduction results from the adjustment of the input optical power on a per each channel basis. Small-step channel power adjustments allow for smooth power transitions and reduced impact on the operating conditions of the line system.

The paper is organized as follows. In section II, the simulation model for the double-stage optical amplifier developed with VPItransmissionMaker<sup>TM</sup> (hereafter called VPI) is presented together with the 80-channel experimental setup used for model validation [7]. Section III describes the proposed RL algorithm and shows how the RL-based power predistortion method can be used for reducing maximum absolute channel power excursions to 0.5 dB. Some conclusions are drawn in Section IV.

## II. AMPLIFIER MODEL AND DATASET GENERATION

In this section, the double-stage amplifier VPI model is first described and validated with measurements performed on an experimental testbed, then used for generating the dataset used for training the RL model. First of all, a VPI model of a double-stage amplifier was built. For adjusting the parameters and validating the model, a double-stage EDFA was characterized (Fig. 1). In the experiment, two lasers are

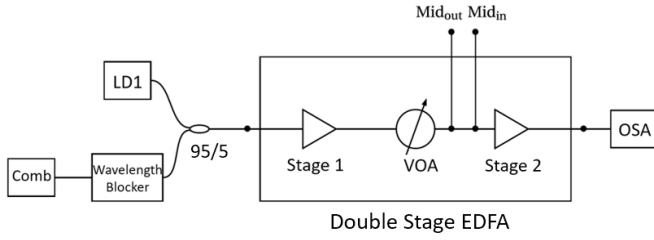


Fig. 1: Double stage EDFA characterization setup.

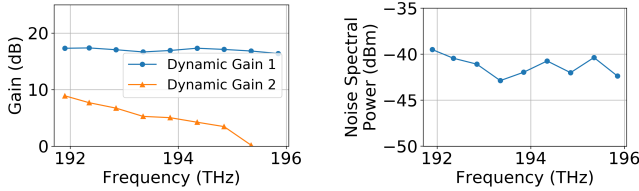


Fig. 2: Experimental dynamic gain (left) and noise spectral power (right).

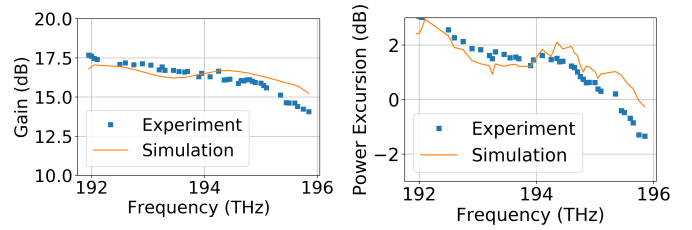
combined at the input of the EDFA operating in current control mode, the first one acting as a gain saturating laser at a fixed wavelength (LD1) and the second one as a tuneable probe laser (Comb). By monitoring the output signal with an optical spectrum analyzer (OSA) at two different saturating power levels, the dynamic gain and noise parameters are calculated (Fig. 2) [8].

Then, the gain saturating laser LD1 was removed from the setup in order to test the amplifier operating in automatic gain control (AGC) mode at 15-dB gain under different working conditions. Three different wavelength loads are considered on a 50-GHz ITU-T grid: 80 channels, 10 channels and 40 channels randomly selected. An equivalent amplifier model was developed based on the black box model in VPI, making use of the previously measured dynamic gain and noise characteristics [8]. Due to noise related issues, an adjustment of the locked gain in the simulation model was required to match the experimental gain.

The resulting model permits to calculate channel gain and power excursion. Power excursion results from the deviations of the output power with respect to the desired mean average power for an ideal flat 15-dB gain amplifier. The results for 40 random channels case show a good agreement between experiment and simulation, justifying the use of the VPI model for generating the synthetic dataset used for the RL model (Fig. 3).

The next step was to generate a dataset using the simulation setup which includes a loop comprising 75 km of standard single mode fiber (SSMF) and a 15-dB gain inline EDFA (Fig. 4). After transmission in 10 loops, output power and power excursions are derived for different channel loads, frequency allocations and input channel power.

A dataset containing 2000 instances is generated for training the RL model composed of 80 neural networks (one per



(a) 40 ch.-Gain

(b) 40 ch.-Power excursion

Fig. 3: Gain and power excursion for 40 random channels.

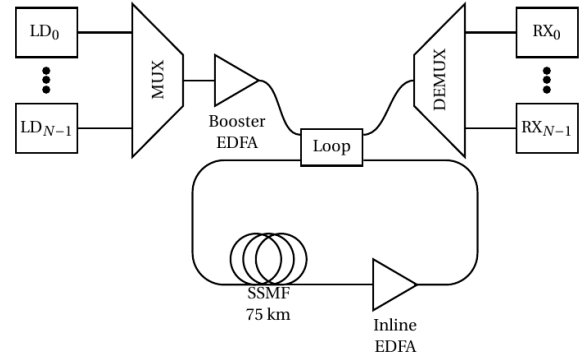


Fig. 4: VPI simulation setup.

channel) and developed using Keras running on TensorFlow Python library [9].

### III. REINFORCEMENT LEARNING

RL algorithms learn from the interaction of an agent with the environment. Here, the environment is the optical line system whereas the agent is the power predistortion module. Overall, for a specific power excursion at the line system output, the power predistortion module takes an action changing the input channel powers in small steps of 0.1 dB. As a consequence, depending on whether the power excursion increases or decreases, the power predistortion module receives a reward. By taking actions and receiving rewards, designed agent learns which actions maximize the rewards, and therefore it learns to minimize the power excursion (Fig. 5).

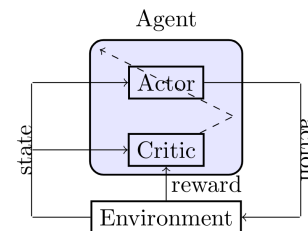


Fig. 5: Block diagram of reinforcement learning.

### A. Environment: Model based on simulations/ Agent: Power predistortion module

Simulations are replaced by a neural network model already successfully used to predict power excursion [3], [10]. As indicated before, this model includes one neural network per channel sharing the same parameters optimized after several trials: one hidden layer with rectified linear unit activation (Relu) and an output layer with linear activation [9].

To complete the environment, a reward is required. After testing continuous and discontinuous reward functions, the best results were obtained using an exponential of the distance to the desired zero power excursion [11].

The power predistortion module objective is to adjust the input channel power at the transmitter side in order to minimize the power excursion at the output of the line system. The RL model permits to solve this complex optimization difficulty by directly learning on the data collected from the system. The small 0.1-dB step tuning of the input powers allows smooth power transitions, reducing undesired impact on the established lightpaths.

As RL algorithm, an “agent” or “actor-critic” method is used [12]. The actor decides which action has to be taken depending on the current state, whereas the critic evaluates the actions of the actor. During the training, the critic learns a reward function which defines how accurate it is for the agent to choose an action based on the current state.

From above, “state and action” important concepts are to be considered. Here, the state corresponds to the current power excursion whereas the action corresponds to the fine-tuning of the optical input power done by the power predistortion module.

The “actor-critic” method based on neural networks was also developed in TensorFlow [9]. For the actor, the neural network has as input the state. After one hidden layer, the output layer, using hyperbolic tangent (tanh) activation, returns the action. For the critic, the neural network has as inputs the state and the action, followed by one hidden layer. After, both are merged in a second one, connected to the output layer with linear activation. Both neural networks use Relu activation in the hidden layers and Adam method for optimization [9].

### B. Reinforcement learning implementation

The RL method was trained for three different channel loads: 6, 12 and 24. Following the training process of 2 hours in a processor Intel core i7-8654U at 1.90 GHz, Fig. 6 shows the evolution of the input channel power and output power excursion for the 12-channel scenario. Starting from a flat input power per channel, the input powers are modified in steps of 0.1 dB until the maximum absolute power excursion is below 0.5 dB. As an example (Fig. 6), in the first channel, power excursion is reduced from 2.03 dB before starting the adjustment to approximately -0.34 dB after applying the power predistortion. Further training is under progress to consider a higher number of channels.

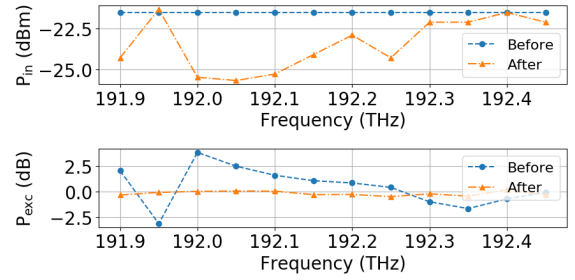


Fig. 6: Input power and power excursion for 12 channels.

## IV. CONCLUSIONS

In this work, we have proposed a RL-based power predistortion method suitable to work under wavelength continuity constraint. The validation of the double-stage amplifier VPI model used to generate the dataset for training the RL model was performed with amplifier measurements performed on a testbed. The results obtained with a 12-channel case show that the power excursion can be reduced from 2.03 dB to -0.34 dB. Future work includes testing the RL model at higher channel loads and its capability of adaptation to variable channel loads.

## REFERENCES

- [1] E. Desurvire, “Erbium-doped fiber amplifiers : principles and applications,” 1994, “A Wiley-Interscience publication.”.
- [2] W. Mo, S. Zhu, Y. Li, and D. C. Kilper, “EDFA wavelength dependent gain spectrum measurement using weak optical probe sampling,” *IEEE Photonics Technology Letters*, vol. 30, no. 2, pp. 177–180, 2017.
- [3] W. Mo, C. L. Gutterman, Y. Li, S. Zhu, G. Zussman, and D. C. Kilper, “Deep neural network based wavelength selection and switching in ROADM systems,” *Journal of optical communications and networking*, vol. 10, no. 10, pp. D1–D11, 2018.
- [4] E. d. A. Barboza, C. J. Bastos-Filho, J. F. Martins Filho, M. J. da Silva, L. D. Coelho *et al.*, “Local and global approaches for the adaptive control of a cascade of amplifiers,” *Photonic Network Communications*, vol. 33, no. 2, pp. 194–207, 2017.
- [5] A. S. Ahsan, C. Browning, M. S. Wang, K. Bergman, D. C. Kilper, and L. P. Barry, “Excursion-free dynamic wavelength switching in amplified optical networks,” *Journal of Optical Communications and Networking*, vol. 7, no. 9, pp. 898–905, 2015.
- [6] Y. Huang, P. B. Cho, P. Samadi, and K. Bergman, “Power excursion mitigation for flexgrid defragmentation with machine learning,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 1, pp. A69–A76, Jan 2018.
- [7] VPI Photonics, “VPI Photonics Design suite,” <http://www.vpiphotonics.com>.
- [8] J. Burgmeier, A. Cords, R. Rz, C. Ffer, and B. Stummer, “A Black Box Model of EDFA’s Operating in WDM Systems,” *Journal of Lightwave Technology*, vol. 16, no. 7, pp. 1271–1275, Jul 1998.
- [9] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis *et al.*, “Tensorflow: A system for large-scale machine learning,” in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265–283.
- [10] M. Freire, S. Mansfeld, D. Amar, F. Gillet, A. Lavignotte, and C. Lepers, “Predicting optical power excursions in erbium doped fiber amplifiers using neural networks,” in *2018 Asia Communications and Photonics Conference (ACP)*, Oct 2018, pp. 1–3.
- [11] L. Matignon, G. J. Laurent, and N. Le Fort-Piat, “Reward function and initial values: Better choices for accelerated goal-directed reinforcement learning,” in *International Conference on Artificial Neural Networks*. Springer, 2006, pp. 840–849.
- [12] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.