

# Online Simultaneous Modeling and Gain Profile Optimization for Multi-pump Raman Amplifiers in C+L-band Optical Systems

Xiaomin Liu<sup>(1)</sup>, Yihao Zhang<sup>(1)</sup>, Yichen Liu<sup>(1)</sup>, Meng Cai<sup>(1)</sup>, Lilin Yi<sup>(1)</sup>,  
Weisheng Hu<sup>(1) (2)</sup>, Qunbi Zhuge<sup>\*(1) (2)</sup>

<sup>(1)</sup> State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, [qunbi.zhuge@sjtu.edu.cn](mailto:qunbi.zhuge@sjtu.edu.cn)

<sup>(2)</sup> Peng Cheng Laboratory, Shenzhen, China

**Abstract** We propose an online paradigm to simultaneously model and optimize Raman gain profiles. Real-time experiments demonstrate the proposed scheme can approach the optimal modelling and optimization performance within 30 data, reducing ~90% of the data compared with the scheme based on an offline-built model. ©2023 The Author(s)

## Introduction

With the rapid development of the Internet applications, the global data traffic is increasing dramatically. To avoid capacity crunch and further exploit the existing fiber infrastructures, future optical transmission bandwidth will be scaled from C-band to C+L, or even C+L+S-band [1,2]. In such a scenario, Raman amplifiers (RAs) can be an appealing choice for fiber attenuation compensation [3]. Besides low noise figures, RAs have a potential in providing a broad gain profile with arbitrary shapes by adjusting the power of each Raman pump.

To realize the gain profile optimization (GPO) of RAs efficiently, heuristic algorithms are firstly utilized, which are time-consuming since massive gain profile calculations or measurements are needed [4,5]. Recently, neural networks (NNs) are adopted to build a digital twin (DT) for fast gain spectrum estimations. Then, the GPO can be conducted quickly based the differentiable DT [6-9]. However, current NN-based methods follow a paradigm of firstly offline modelling and then online optimization, requiring an offline training procedure to construct an accurate DT before optimization. Therefore, a large dataset with hundreds to thousands of randomly-collected data is required. Actually, the extensive measurement for an RA is quite expensive and

time-consuming, hindering the practical utilization in a commercial optical network with a large number of RAs.

To address this issue, in this paper, a novel paradigm is proposed to conduct modelling and GPO simultaneously in an online system. By freezing and unfreezing the inner weights of a NN-based DT, the DT and pump configurations are updated based on gradient descent (GD) iteratively. The proposed scheme does not require constructing an accurate NN-based DT before GPO, so that the burden of the offline data collection can be significantly alleviated. An online experimental validation in a C+L-band amplification system is conducted. The proposed scheme can simultaneously build the DT and conduct the GPO, and both of their optimal performances can be achieved using ~30 data, reducing ~90% data compared to the offline-trained model with randomly-collected data.

## Principle

The proposed scheme is illustrated in Fig. 1. When the GPO starts, in step ①, an NN-based DT for the RA is initialized by randomization or training with a small dataset, which can be written as  $G = f_{NN}(P)$ .  $P$  and  $G$  are the vectors representing the pump configurations and gain profiles, respectively. Then, in the online

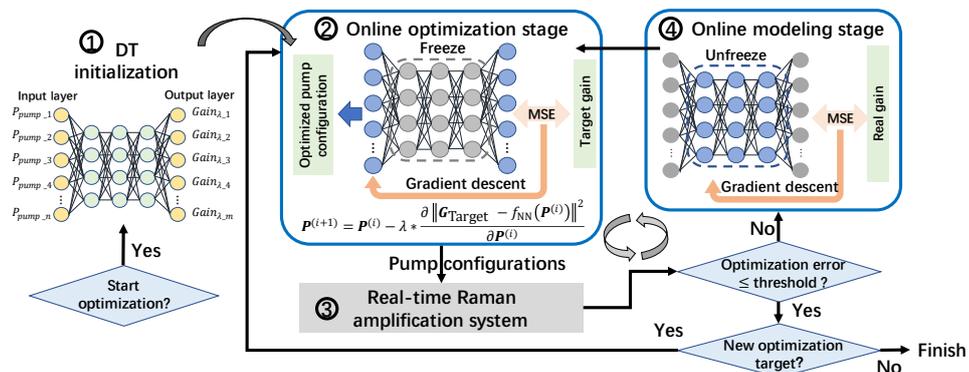
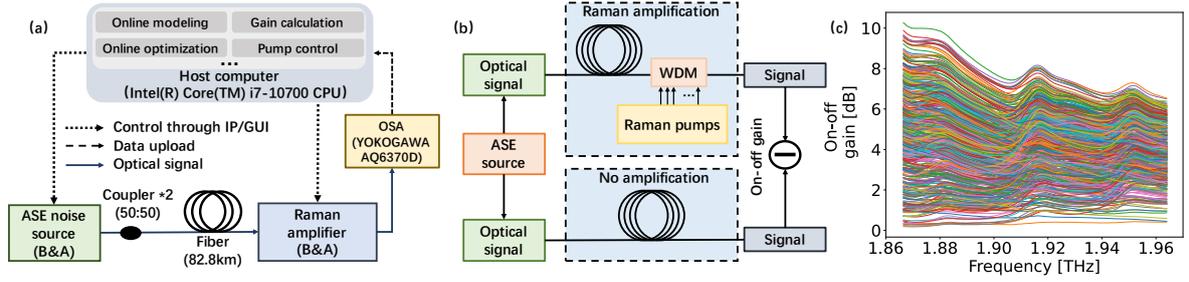


Fig. 1: The workflow of the proposed scheme for simultaneous online modeling and optimization for an RA.



**Fig. 2:** (a) The automatic experimental setup for online modeling and optimization. (b) The measurement method for the on-off gain. (c) The experimentally measured on-off gain spectra under various pump configurations.

optimization stage (step ②), the inner weights of the NN-based DT are frozen. With the GD, the input  $\mathbf{P}$  are updated by minimizing the mean-squared-error (MSE) between the NN's output and target gain, denoting as  $\mathbf{G}_{\text{Pred}}$  and  $\mathbf{G}_{\text{Target}}$  respectively. Afterwards, the optimized pump configuration is obtained and fed to the real system. Then, the corresponding gain spectrum, denoted as  $\mathbf{G}_{\text{Real}}$ , is measured (step ③).

If the error between the  $\mathbf{G}_{\text{Real}}$  and  $\mathbf{G}_{\text{Target}}$  is larger than the threshold  $\delta_{\text{threshold}}$ , the measured data is added to the training dataset to update the DT in the online modelling stage (step ④). In this stage, same as the traditional NN training, the inner weights of the NN are unfrozen and updated based on GD. Afterwards, the updated DT is utilized to reconduct the online optimization in step ② and derive a new optimized pump configuration. In this way, step ② to ④ are conducted iteratively until the error between  $\mathbf{G}_{\text{Real}}$  and  $\mathbf{G}_{\text{Target}}$  is smaller than  $\delta_{\text{threshold}}$ .

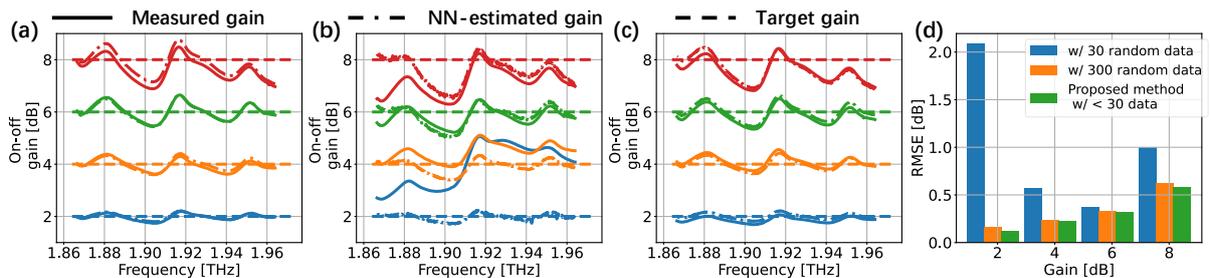
When finishing the GPO for one target gain, a DT trained along with the optimization is obtained. If the network status changes and a new gain profile is required, the DT obtained during the previous optimization procedure can be directly employed without initialization. In this way, the knowledge learned from the previous optimization procedures can be used to speed up the next round of optimization.

### Experimental setup

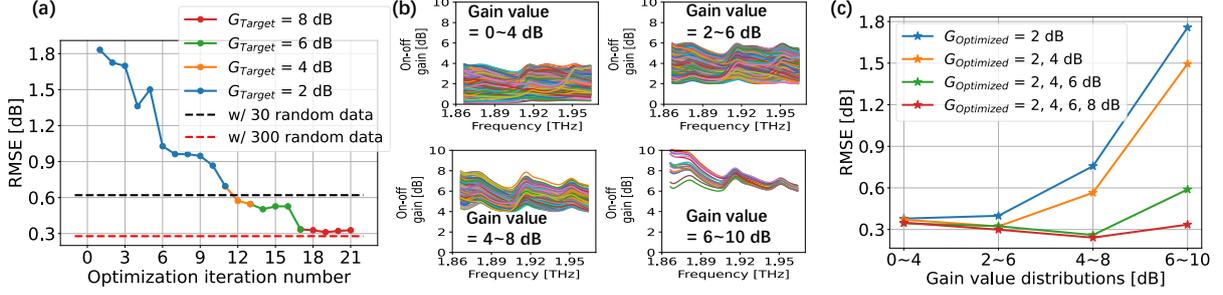
A real-time experiment is conducted to demonstrate the effectiveness of the proposed scheme. The experimental setup is shown in Fig.

2(a). First, an amplified spontaneous emission (ASE) noise source is utilized to simulate a flat C+L-band signal spectrum from 186.6 THz to 196.4 THz. Two 50:50 couplers are used to provide suitable attenuations and set the total signal power to 15.5 dBm. Then, the signal is transmitted in an 82.8-km standard single mode fiber. A distributed RA with four counter-propagating pumps is employed and the wavelengths of the pumps are 1428 nm, 1454 nm, 1490 nm, and 1509 nm. The power of each pump can be adjusted individually by controlling the pump current. The signal spectra with and without amplification are measured by an optical spectrum analyser (OSA) to calculate the on-off gain. The measurement method is shown in Fig. 2(b). The OSA and RA are controlled by a host computer through IP networking. Automatically, procedures such as the device controlling, online modeling, and GPO are processed with Python on the host computer.

The NN-based DT has four input features to represent the controlling values of the pump currents. The output is a 400-dimensional vector depicting the gain spectra. The NN has two hidden layers with 17 and 60 nodes. The activation functions are Sigmoid. To initialize the DT in step ①, we collect 5 randomly sampled data for pre-training. During training, early-stop with a patience of 1000 is utilized. 70% and 30% of the data are used for training and validation, respectively. We collect 460 gain spectra under different pump configurations randomly as the testing dataset for following accuracy evaluations, which are shown in Fig. 2(c).



**Fig. 3:** (a) The optimized gain profiles of the proposed scheme. (b) and (c) are the optimized gain profiles of the NN-based baseline models trained with 30 data and 300 data, respectively. In each figure, the measured gain, target gain and the gain estimated by the NN-based DT are shown. (d) The optimization RMSEs of each method with different target gains.



**Fig. 4:** (a) The DT's accuracy on the testing dataset during the GPO with the proposed scheme. The lines with different colors show the changes of the DT's accuracy during the GPO with different target gains. (b) The datasets with different gain value distributions. (c) The DT's accuracy on the datasets with different gain value distributions after finishing each optimization target.

When the real-time optimization starts, the first  $G_{\text{Target}}$  is assumed to be 2 dB. Then, we further consider a situation that the target gain changes from 2 dB to 4 dB, 6 dB, and 8 dB sequentially. The changes of the gain profiles are assumed to simulate network reconfigurations due to the changes of the link loss or power optimizations.

### Optimization Performance Evaluations

We first analyze the GPO performance. In Fig. 3(a), the gain profile obtained by the proposed scheme is shown. In the experiment, the online GPO finishes the first optimization target of 2 dB in 11 iterations and finishes the four assumed optimization targets in 21 iterations. So, only 26 measured data are collected (5 for initialization), which is data-efficient. The experiments are conducted for three times with different DT initializations for stability analysis. The GPO performance is stable with a root-mean-square error (RMSE) fluctuation less than 0.15 dB. Additionally, the proposed scheme can realize all the assumed optimization targets within 25 iterations, *i.e.*, within 30 collected data.

The baseline GPO method for comparison is to perform GD on an offline-trained DT without online modelling. As shown in Fig. 3(b), if relying on the offline-trained DT with 30 randomly-collected data, the DT's accuracy is low, resulting in unsatisfactory performance of GPO. By increasing the training data size to 300, the offline DT and GPO can both converge to a relatively stable and optimal precision, which are plotted in Fig. 3(c). The optimization RMSE of each method under different target gains are shown in Fig. 3(d). The results show that, compared with the offline-trained DT with randomly-collected data, the proposed scheme can largely reduce the data needed to achieve the optimal GPO performance.

### Modeling Performance Evaluations

We further investigate the online modelling performance. In Fig. 4(a), we illustrate the changes of DT's accuracy on the testing dataset during the whole optimization processes with the four target gains. The results show that, the DT's

accuracy is gradually improved along with the GPO. With 21 iterations, the DT can reach an optimal accuracy. Additionally, when  $G_{\text{Target}}$  changes, the GPO is reconducted and the DT's accuracy can be further improved under the new optimization target. At the same time, the DT with higher accuracy can speed up the convergence of GPO with less iterations.

It should be noted that, when  $G_{\text{Target}}$  is 2 dB, the accuracy of the DT obtained by the proposed scheme is 0.65 dB on the testing dataset in Fig. 2(c). Meanwhile, the RMSE of the GPO performance is 0.12 dB, which is quite different from the modeling accuracy. To explain this, we test the DT on the datasets with different gain value distributions. As shown in Fig. 4(b), we generate four datasets containing gain profiles around the target gains with a ripple of 2 dB. The DTs after finishing each optimization target are tested on these datasets. The RMSEs are plotted in Fig. 4(c). We observe that after GPO, the obtained DT reaches a high accuracy on the dataset with gain values around  $G_{\text{Target}}$ . When the gain value deviates from  $G_{\text{Target}}$ , the DT's accuracy reduces because the DT has not seen these data. By carefully learning about the region around  $G_{\text{Target}}$ , the DT built by the proposed scheme can reach a high accuracy in a certain subspace to effectively assist the GPO.

### Conclusions

We propose a new paradigm which conducts online modelling and GPO simultaneously for an RA. A real-time experiment demonstrates the effectiveness of the proposed scheme in a C+L-band amplification system. Compared with the optimization methods based on an offline training model with a large number of data, the proposed scheme can significantly reduce the needed data size for modelling to assist the efficient GPO.

### Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFB1800902), the NSFC (62175145), and Shanghai Pilot Program for Basic Research - Shanghai Jiao Tong University (21TQ1400213).

## References

- [1] M. Cantono, R. Schmogrow, M. Newland, V. Vusirikala and T. Hofmeister, "Opportunities and challenges of C+L transmission systems," *Journal of Lightwave Technology*, vol. 38, no. 5, pp. 1050-1060, 2020, DOI: [10.1109/JLT.2019.2959272](https://doi.org/10.1109/JLT.2019.2959272)
- [2] V. Lopez, B. Zhu, D. Moniz, N. Costa, J. Pedro, X. Xu, A. Kumpera, L. Dardis, J. Rahn and S. Sanders, "Optimized design and challenges for C&L band optical line systems," *Journal of Lightwave Technology*, vol. 38, no. 5, pp. 1080-1091, 2020, DOI: [10.1109/JLT.2020.2968225](https://doi.org/10.1109/JLT.2020.2968225)
- [3] L. Rapp and M. Eiselt, "Optical amplifiers for multi-band optical transmission systems," *Journal of Lightwave Technology*, vol. 40, no. 6, pp. 1579-1589, 2022, DOI: [10.1109/JLT.2021.3120944](https://doi.org/10.1109/JLT.2021.3120944)
- [4] B. Neto, A. L. J Teixeira, N. Wada, and P. S. André, "Efficient use of hybrid Genetic Algorithms in the gain optimization of distributed Raman amplifiers," *Optics Express*, vol. 15, pp. 17520-17528, 2007, DOI: [10.1364/OE.15.017520](https://doi.org/10.1364/OE.15.017520)
- [5] 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, DOI: [10.1109/JPHOT.2011.2140366](https://doi.org/10.1109/JPHOT.2011.2140366)
- [6] 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, DOI: [10.1109/JPHOT.2018.2817843](https://doi.org/10.1109/JPHOT.2018.2817843)
- [7] 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, DOI: [10.1109/JLT.2019.2952179](https://doi.org/10.1109/JLT.2019.2952179)
- [8] X. Ye, A. Arnould, A. Ghazisaeidi, D. Le Gac, and J. Renaudier, "Experimental prediction and design of ultra-wideband Raman amplifiers using neural networks," *Optical Fiber Communication Conference (OFC)*, 2020, W1K.3, DOI: [10.1364/OFC.2020.W1K.3](https://doi.org/10.1364/OFC.2020.W1K.3)
- [9] U. C. de Moura, M. A. Iqbal, M. Kamalian, L. Krzczanowicz, F. Da Ros, A. M. Rosa Brusin, A. Carena, W. Forsyia, S. Turitsyn, and D. Zibar, "Multi-band programmable gain Raman amplifier for high-capacity optical networks," *2021 IEEE Photonics Conference*, 2021, pp. 1-2, DOI: [10.1109/IPC48725.2021.9592938](https://doi.org/10.1109/IPC48725.2021.9592938)