

Digital Twin Modelling of Cascaded Amplifiers in the COSMOS Testbed

Rishu Raj*, Shuang Xie*, Zehao Wang[#], Tingjun Chen[#], and Daniel Kilper*

*CONNECT Centre, Trinity College Dublin, Ireland; [#]Duke University, Durham, NC, USA.

Abstract— Digital twins provide a cost-effective means of evaluating performance, predicting network changes, and enhancing network administration and decision-making processes. However, acquiring detailed data for digital twin development remains a challenge due to commercial system constraints. City-scale testbeds, like COSMOS, offer practical solutions, aiding in data collection for modelling of digital twins. In this paper, we utilize data from experiments on the COSMOS testbed to design a digital twin model for the accumulation of gain ripple in cascaded Erbium-doped fiber amplifiers (EDFAs). We quantify the gain ripple in terms of its peak-to-valley ratios and identify optimal EDFA combinations. Moreover, we explore the impact of system parameters, and validate the proposed digital model through comparison with experimental results. We show that the differences between the digital twin predictions and experimental results are ≤ 0.1 dB.

Keywords—digital twins, EDFA, gain ripple, COSMOS testbed.

I. INTRODUCTION

The ever-increasing demands for high bandwidth and low latency in emerging wireless technologies (like 6G) and in wide-area networks (such as data center interconnects) rely heavily on optical infrastructure deployed in metro, regional, and long-haul networks [1]. Presently, optical transmission systems find application in metro networks, and are also a promising candidate for high-capacity radio fronthaul and backhaul edge networks. These high-capacity systems span from the network edge to the core, each presenting diverse performance and control requirements. However, an experimental emulation platform to address this complexity has not yet been established. At the same time, the introduction of open and disaggregated optical transmission systems creates additional experimental needs. The majority of research challenges associated with such systems revolve around their control and management. Control complexities arise due to factors such as non-linear impairments, component anomalies, and optical power dynamics resulting from wavelength and polarization-dependent effects in amplifiers and fiber spans [2].

Data-driven approaches and machine learning (ML) methods have garnered attention as potential solutions for managing the increased control complexity in fully disaggregated systems [3],[4]. Enhanced data collection and data-driven methods, designed to utilize such data, are crucial for making advancements in these disaggregated systems. As such, addressing the challenges of fully disaggregated systems necessitates experimental platforms for examining the interactions between new control and management systems and the physical transmission effects.

Traditional optical networking experiments involved a small number (typically three to six) of nodes over short distances (e.g., a single transmission span). However, modern

optical networks can scale to include hundreds or thousands of nodes, supporting signal transmission across dozens of nodes [5]. With the increase in scale and cost of early experiments, laboratory experiments quickly adopted experimental emulation methods, such as the widely used recirculating loop method [6]. While recirculating loops have proven effective for studying impairments in large-scale systems, they are generally incompatible with optical networking experiments. Hence, new methods are imperative for experimentation to study the physical effects on a large scale and how they interact with innovative software controls, ML algorithms, or control hardware innovations. This shift towards data-driven controls for enhanced management and increased automation also necessitates new experimental emulation methods such as digital twins.

A digital twin is a virtual model of a communication network that faithfully simulates the devices, communication links, operational conditions, and applications operating on the actual operational network [7]. Through the creation of various settings within a controlled environment and the execution of multiple scenarios, digital twins offer an economical method to assess performance, anticipate the consequences of network alterations, enhance network administration, and facilitate informed decision-making [8]. Therefore, digital twin methods require detailed datasets of the system's characteristics and performance across a wide range of operating conditions [9], [10].

Obtaining data from commercial systems is challenging due to business practices and operating requirements, such as customer privacy regulations. Alternatively, testbeds can play an indispensable role in collecting datasets for experimentation, especially if they include a field-deployed fiber and programmable optical hardware. To this end, city-scale testbeds are being deployed with a focus on experimenting with various networking technologies. For instance, the Japan-wide Orchestrated Smart Sensor Environment (JOSE) testbed [11] in Japan focuses on Internet-of-Things (IoT) services, while Bristol is Open [12] in the United Kingdom was designed with a focus on 5G wireless and optical networking experimentation. Another example is the COSMOS (acronym for Cloud-enhanced Open Software-defined MOBILE wireless testbed for city-Scale deployment) [13] which is one of the four testbeds under the platforms for advanced wireless research (PAWR) program [14] funded by the National Science Foundation (NSF) and an industrial consortium of 30 companies and associations. These testbeds allow investigation into the interoperability of different applications and devices in practical, deployed network scenarios. Additionally, they can be utilized to collect datasets to aid in the development of ML algorithms and data-driven methods.

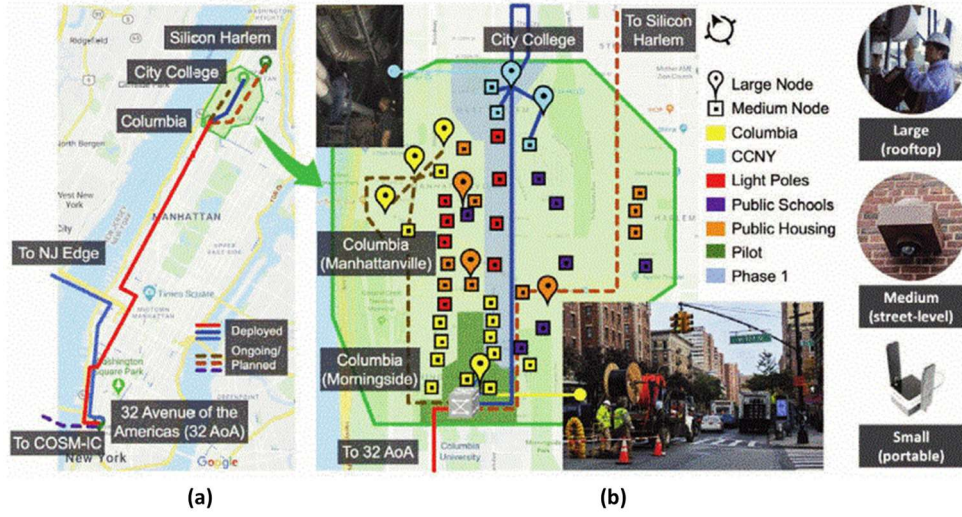


Fig. 1. (a) The envisioned deployment area of COSMOS in West Harlem, NYC, and (b) the deployment sites involving radio nodes of various types and dark fiber connections. The core of COSMOS, depicted here as a grey cube, is in the Computing Research Facilities (CRF) of Columbia University [15].

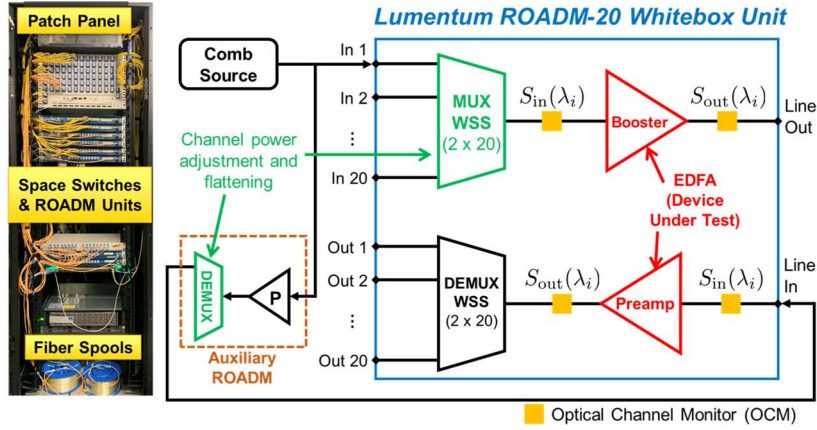


Fig. 2. (Left) The programmable optical network in the COSMOS testbed. (Right) Block diagram of the Lumentum ROADM-20 whitebox unit with built-in optical power monitors, which are used for the EDFA gain spectrum measurements.

In this paper, we use the experimental dataset for the spectral performance of individual Erbium-doped fiber amplifiers (EDFAs) (in the COSMOS testbed) to design a digital twin model of the gain ripple accumulation in cascaded EDFAs. We quantify the gain ripple in terms of the peak-to-valley ratio and identify the combination with the highest and lowest gain ripple accumulation. We also study the effect of system parameters on the gain ripple. Furthermore, we validate the proposed digital model by comparison with experimental results.

The remainder of this paper is organized as follows. In Section II, we explain the methodology adopted in our work. We discuss the results of our study in Section III and present a brief conclusion in Section IV.

II. METHODOLOGY

In this section, we briefly discuss the features of the COSMOS testbed and the set up therein which has been used for experimentation and data collection. We then describe the modelling of cascaded EDFAs in the digital twin.

A. COSMOS Testbed

COSMOS is a city-scale open-access programmable

testbed. It is an ambitious project designed for research and experimentation with cutting-edge wireless and optical technologies in real-world urban environments [13]. The project's scope encompasses a vast area as illustrated in Fig. 1 which provides an overview of its extensive planned deployment within the bustling Harlem neighborhood of Manhattan, spanning approximately 1 square mile. This initiative is a collaborative effort involving multiple stakeholders, including the NYC CTO, City College of New York, Silicon Harlem, and the local community [15].

The envisioned deployment strategy is both extensive and intricate, comprising three distinct categories of nodes, each playing a critical role within the testbed ecosystem [15]. Firstly, there are the 9 large nodes, serving as macrocellular base stations, strategically perched upon rooftops. These pivotal installations provide broad coverage and support to the network's overarching functionality. Secondly, complementing these large nodes are approximately 40 medium-sized nodes thoughtfully positioned at street level. These nodes, functioning as access points or small cells, ingeniously harness building facades and light poles to optimize their placement, enhancing network accessibility

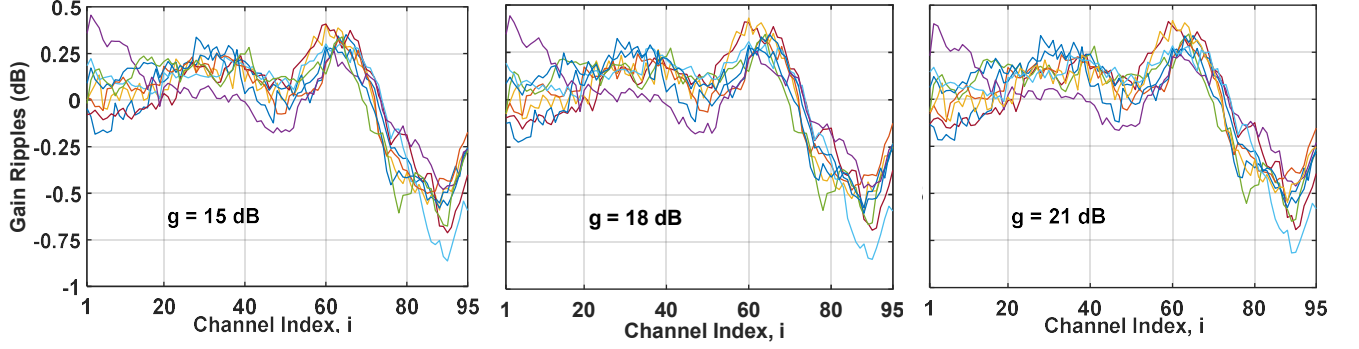


Fig. 3. Gain ripple waveforms for 8 booster EDFAs at three different target gain (g) values. Here, channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

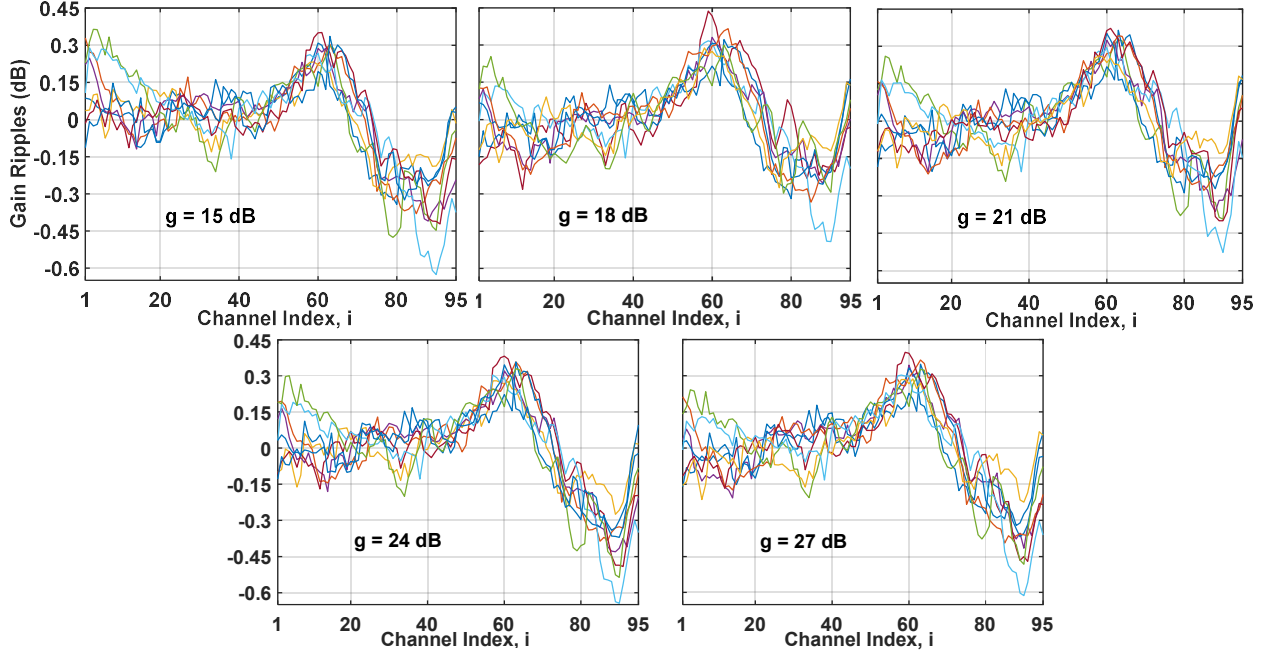


Fig. 4. Gain ripple waveforms for 8 preamp EDFAs at five different target gain (g) values. Here, channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

and performance. Thirdly, there are approximately 200 small nodes, prized for their near-portable nature. These nodes exhibit remarkable versatility, capable of acting as either fixed or mobile devices, dispersed strategically throughout the expansive testbed [13]. The primary components of COSMOS' programmable x-haul optical network at Columbia University comprise a Calient S320 320×320 space switch, a DiCon 16×16 space switch, eight Lumentum ROADM-20 whitebox units, a spooled fiber plant, and dark fiber connections leading to the colocation site at 32 Avenue of the Americas (32 AoA) [15].

B. Experimental Set-up

Our focus lies in characterizing the gain spectrum of EDFAs integrated into the commercial-grade Lumentum ROADM-20 whitebox units deployed within the COSMOS testbed. Each ROADM unit is equipped with two optical amplifiers: a receive preamp EDFA at the line in and a transmit booster EDFA at the line out, along with various optical channel monitors (OCMs) as depicted in Fig. 2.

To measure the gain spectrum for each device under test (DUT) EDFA, a comb source with 95×50 GHz channels is

utilized to generate the wavelength division multiplexing (WDM) spectrum in the C-band. In the case of a DUT booster EDFA, the multiplexer (MUX) wavelength-selective switch (WSS) of the same DUT ROADM is employed for channel loading configuration and spectrum flattening. Alternatively, for a DUT preamp EDFA, channel loading configuration and spectrum flattening are carried out using the demultiplexer (DEMUX) WSS of an auxiliary ROADM, with one drop port connected to the line in port of the DUT ROADM.

To create the digital twin, we utilize an openly accessible dataset [16] comprising gain spectra measurements obtained from 16 EDFAs within 8 commercial-grade Lumentum ROADM-20 units deployed in the PAWR COSMOS testbed. The gain spectrum of each booster or pre-amplifier EDFA is recorded at specific target gain settings of 15/18/21 dB and 15/18/21/24/27 dB, respectively. These measurements are conducted in the high gain mode with a 0 dB gain tilt under fully loaded channel configuration across 95 channels, spanning from 191.325 THz (1566.72 nm) to 196.075 THz (1529.16 nm) in the C-band, with a fixed channel spacing of 50 GHz. The input-output relationship of the EDFAs is studied in [16].

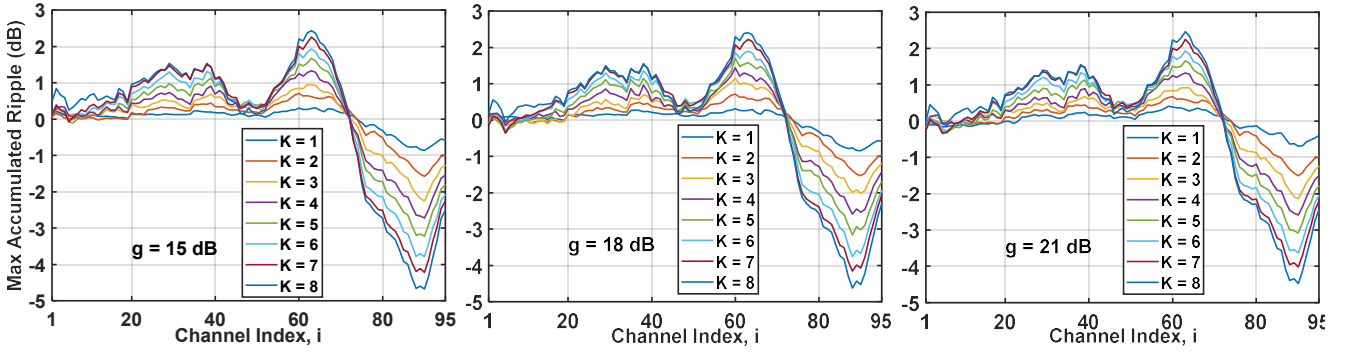


Fig. 5. Maximum accumulated gain ripple waveforms for booster EDFAs at three different target gain (g) values. Here, $N = 8$ and the channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

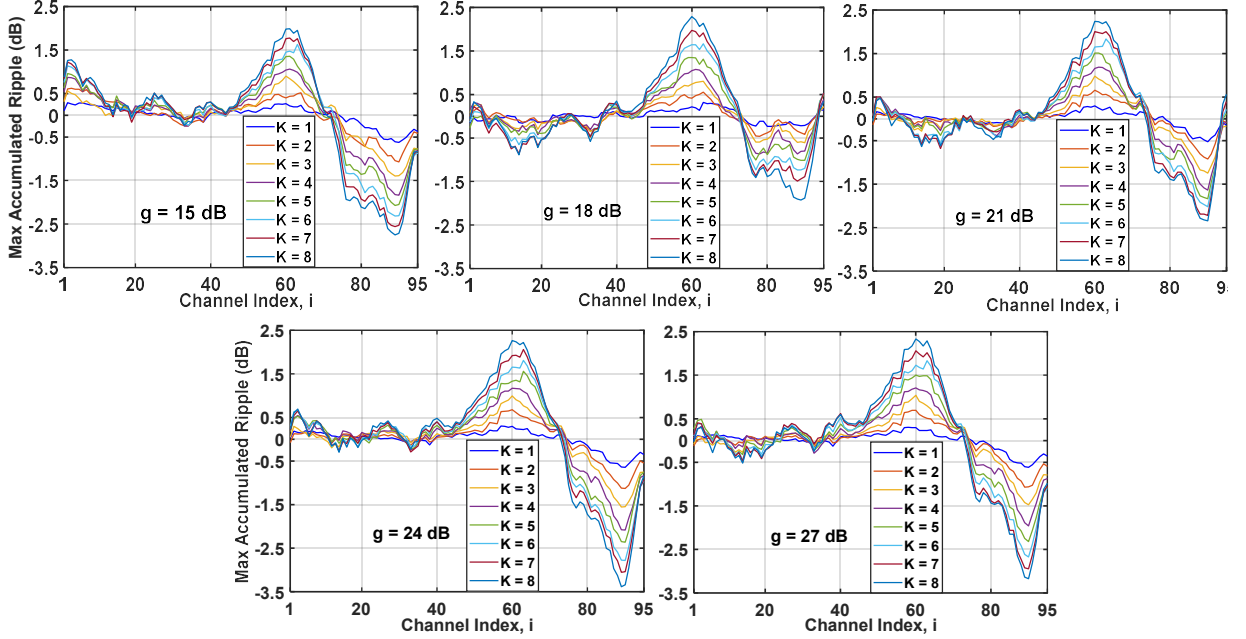


Fig. 6. Maximum accumulated gain ripple waveforms for preamp EDFAs at five different target gain (g) values. Here, $N = 8$ and the channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

The signal gain of each EDFA is calculated from the input and output spectra as

$$G_s(\lambda_i) = S_{\text{out}}(\lambda_i) - S_{\text{in}}(\lambda_i) \quad (1)$$

where $S_{\text{in}}(\lambda_i)$ and $S_{\text{out}}(\lambda_i)$ are the input and output spectra, respectively, measured in dBm at the i^{th} wavelength λ_i . The zero mean gain ripple waveform of is then obtained as

$$\mathcal{G}(\lambda_i) = G_w(\lambda_i) - \overline{G_w} \quad (2)$$

where $\overline{G_w}$ is the mean of $G_w(\lambda_i)$ which is the wavelength-dependent ripple gain spectrum calculated as $G_w(\lambda_i) = G_s(\lambda_i) - g$ where g is the target gain of the EDFA. In this work, for brevity, we refer to the zero mean gain ripple as simply the gain ripple. Note that all calculations are in the decibel scale.

C. Cascading of EDFAs

When EDFAs are connected in cascade, the individual gain ripple accumulates (or adds up in the decibel scale). We may choose to connect K out of N available amplifiers, such that $K \in \{1, 2, \dots, N\}$ and there are ${}^N C_K$ possible cascade combinations, where ${}^N C_K$ is the binomial coefficient defined as ${}^N C_K = N! / (K! (N - K)!)$. The accumulated gain ripple in the j^{th} combination of K amplifiers is

$$\mathcal{R}_K^{(j)}(\lambda_i) = \sum_{k=1}^K \mathcal{G}_{k,j}(\lambda_i) \quad (3)$$

where $\mathcal{G}_{k,j}(\lambda_i)$ is the gain ripple of the k^{th} EDFA in the j^{th} cascade combination $\forall k \in \{1, 2, \dots, K\}, j \in J$, and J is the set of indices of all possible cascade combinations, so $J \triangleq \{1, 2, \dots, {}^N C_K\}$. In general, each combination of K EDFAs has a unique accumulation of gain ripple. So, for each value of K , there are ${}^N C_K$ different types of accumulated gain ripple. We are interested in identifying the combinations with the highest and lowest gain ripple corresponding to the worst-case and best-case scenarios, respectively. To quantify the gain ripple, we define the peak-to-valley ratio (PVR) as the difference between the maxima and minima values of the gain ripple waveform. This is mathematically expressed as

$$\mu_K^{(j)} \triangleq \max(\mathcal{R}_K^{(j)}) - \min(\mathcal{R}_K^{(j)}) \quad (4)$$

Further, with K EDFAs connected in cascade, the highest accumulated gain ripple is

$$\mathcal{H}_K = \mathcal{R}_K^{(m)} | m \in J \text{ and } \forall n \in J, \mu_K^{(m)} \geq \mu_K^{(n)} \quad (5)$$

Hence, the m^{th} combination has the largest PVR $\mu_K^{(m)}$ and is adjudged the worst possible combination of K out of N EDFAs. Similarly, the lowest accumulated gain ripple is

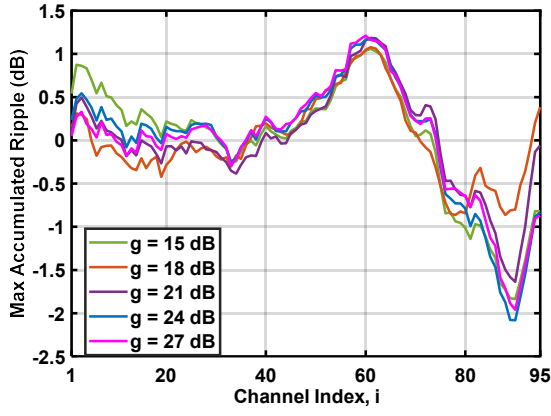


Fig. 7. Maximum accumulated gain ripple waveforms for preamp EDFAs at five different target gain (g) values. Here, $K = 4$, $N = 8$, and the channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

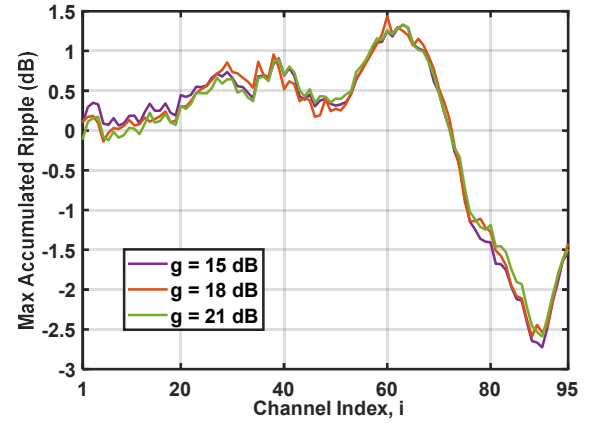


Fig. 9. Maximum accumulated gain ripple waveforms for booster EDFAs at three different target gain (g) values. Here, $K = 4$, $N = 8$, and the channel indices denote increasing wavelengths such that $\lambda_1 = 1529.16$ nm (196.075 THz) and $\lambda_{95} = 1566.72$ nm (191.325 THz).

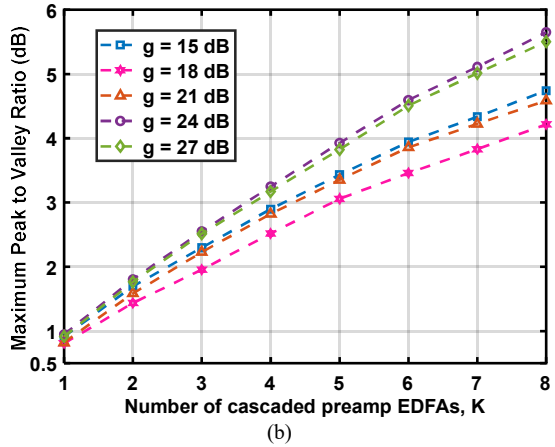
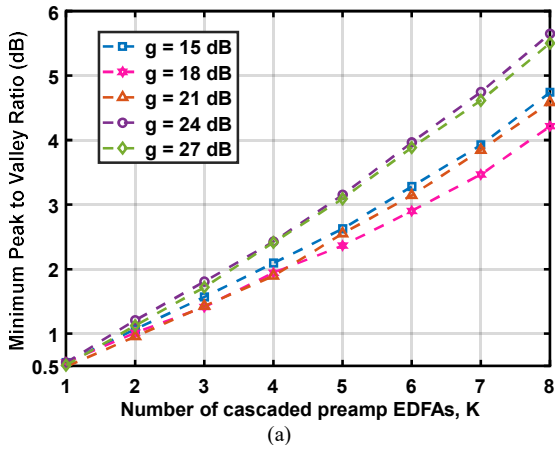


Fig. 8. Variation in (a) minimum, and (b) maximum peak to valley ratios with increase in number of preamp EDFAs at different target gains (g).

$$\mathcal{L}_K = \mathcal{R}_K^{(l)} | l \in J \text{ and } \forall n \in J, \mu_K^{(l)} < \mu_K^{(n)} \quad (6)$$

Ergo, the l^{th} combination has the smallest PVR $\mu_K^{(l)}$ and is considered the best possible combination of K out of N EDFAs.

III. RESULTS AND DISCUSSION

In this section, we present the results obtained in our work using the experimental methodology described in Section II.

A. Gain Ripple of Individual EDFAs

We obtain the gain ripple waveforms of the EDFAs using

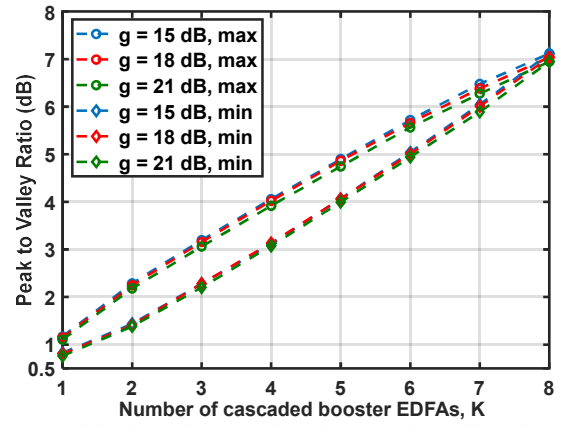


Fig. 10. Variation in minimum, and maximum peak to valley ratios with increase in number of booster EDFAs at different target gain (g) values.

(2) and the spectral data available in [16]. In Fig. 3, we show the amplifier gain ripple in the booster EDFAs at three different target gain values, and in Fig. 4, the amplifier gain ripple in preamp EDFAs is shown for five different target gain values. In both cases, we observe that the EDFAs add considerable gain ripple to the signal passing through them.

B. Accumulation of Gain Ripple in Cascaded EDFAs

To emulate the cascaded combination of K EDFAs, we determine the accumulated gain ripple of each possible combination indexed in J using (3). We then deduce the combination with the highest gain ripple (or maximum PVR) using (5) to obtain the digital twin of the worst-case scenario. The highest accumulated gain ripple waveforms at each value of K are depicted in Fig. 5 for cascaded booster EDFAs, and in Fig. 6 for cascaded preamp EDFAs with different target gain values. As anticipated, the gain ripple accumulation is higher when more number of EDFAs are connected. A similar analysis can be done to obtain the digital twin of the best-case scenario. This has not been shown for the sake of brevity.

C. Effect of System Parameters

We now analyse the effect of number of EDFAs and target gain on the gain ripple accumulation of the cascade combination. In Fig. 7, we fix $K = 4$ and plot the highest accumulated gain ripple waveforms at for preamp EDFAs at

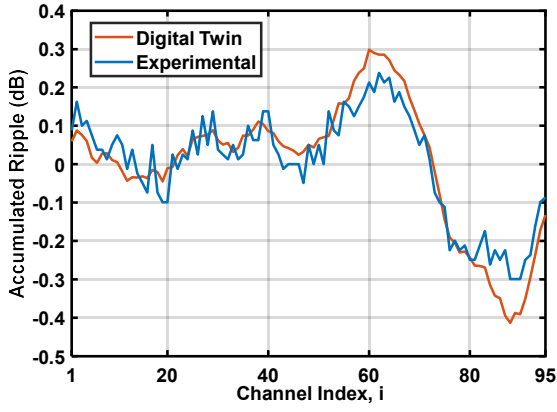


Fig. 11. Comparison of results from digital twin model and experimental study for 4 cascaded ROADMs (4 preamps and 4 boosters).

different target gain values. We observe that the gain ripple does not vary much when the target gain (g) value is changed, and the maximum variation (due to g) in the gain ripple waveforms is ~ 1 dB. To study the effect of number of cascaded EDFAs (K), we quantify the gain ripple using PVRs and in Fig. 8, we plot the variation in maximum and minimum PVRs with increase in K at different target gain values. As K is increased from 1 to 8, the minimum PVR increases from ~ 0.5 dB to ~ 5.5 dB, whereas the maximum PVR increases from ~ 1 dB to ~ 5.5 dB. The same analysis is done for booster EDFAs in Figs. 9 and 10 and they exhibit a similar trend as preamp EDFAs. Although, in both cases, a variation in the target gain values does not affect the gain ripple much, the effect is more pronounced in case of preamps than boosters. Increasing K from 1 to 8, results in an enhancement of PVR from ~ 1 dB to ~ 7 dB for boosters.

D. Validation of the Digital Twin Model

To validate the digital twin model, we verify our results against the accumulated gain ripple obtained using the experimental set-up in [17]. We deduce the accumulated gain ripple in 4 cascaded ROADMs (each ROADM comprises of a preamp and a booster EDFA) using the digital twin proposed in our work and compare it with the experimental results for the combination of the same ROADMs in a multi-span scenario [17]. We observe in Fig. 11, that the deductions from digital twin model are in good agreement with the experimental results. Note that the OCMs in the ROADMs are limited by a resolution of 0.1 dB, which is one of the sources of the error between the two curves plotted in Fig. 11. Moreover, the launch power into the fiber span is around -8 dBm per channel. At such low values of power, the impact of Raman tilt is minor and so, as mentioned earlier, we adopt a 0 dB gain tilt approach for all EDFAs.

IV. CONCLUSION

We propose a digital twin model for cascaded amplifiers in the COSMOS testbed. We analyze the gain ripple waveforms of individual EDFAs and then obtain the accumulated gain ripple in cascaded EDFAs. We also quantify the gain ripple using PVRs and infer the worst and best combination depending on the highest and lowest PVRs, respectively. Moreover, we study the effect of system parameters on the accumulated gain ripple. We observe that

the PVR of gain ripple increases substantially by ~ 6 dB for boosters and ~ 5 dB for preamps when the number of EDFAs is increased from 1 to 8. However, the gain ripple are not affected much the target gain values of the EDFAs. Furthermore, we show that the error between the gain ripple estimated by the digital twin model and actual experimental set-up is negligible (≤ 0.1 dB). This verifies the proposed digital twin model for the cascaded EDFAs. As a future extension of this work, it would be interesting to study the gain ripple waveforms with non-zero gain tilt in the presence of Raman scattering effects.

ACKNOWLEDGEMENTS

This work was supported in part by NSF grants CNS-1827923, OAC-2029295, CNS-2112562, and CNS-2330333, and Science Foundation Ireland under Grant #13/RC/2077 P2.

REFERENCES

- [1] M. Ruffini, "Multidimensional convergence in future 5G networks," *Journal of Lightwave Technology*, vol. 35, no. 3, pp. 535-549, 2017.
- [2] E. Akinrintoyo *et al.*, "Reconfigurable topology testbeds: a new approach to optical system experiments," *Optical Fiber Technology*, vol. 76, pp. 10324, 2023.
- [3] F. Musumeci *et al.*, "An overview on application of machine learning techniques in optical networks," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1383-1408, 2019.
- [4] E. Seve, J. Pestic and Y. Pointurier, "Accurate QoT estimation by means of a reduction of EDFA characteristics uncertainties with machine learning," in *Proc. International Conference on Optical Network Design and Modeling (ONDM)*, Barcelona, Spain, 2020, pp. 1-3.
- [5] A. Lord, "The future of optical transport: architectures and technologies from an operator perspective," in *Proc. Optical Fiber Communication Conference*, San Diego, CA, USA, 2022, pp. 1-18.
- [6] N. S. Bergano and C. R. Davidson, "Circulating loop transmission experiments for the study of long-haul transmission systems using erbium-doped fiber amplifiers," *Journal of Lightwave Technology*, vol. 13, no. 5, pp. 879-888, 1995.
- [7] M. S. Rodrigo *et al.*, "Digital twins for 5G networks: A modeling and deployment methodology," *IEEE Access*, vol. 11, pp. 38112-38126, 2023.
- [8] Y. Hui *et al.*, "Digital twins for intelligent space-air-ground integrated vehicular network: Challenges and solutions," *IEEE Internet of Things Magazine*, vol. 6, no. 3, pp. 70-76, 2023.
- [9] DTC Innovation Forum, Digital twin computing (white paper), [online] Available: <https://iowngf.org/white-papers/> (accessed Oct. 5, 2023).
- [10] S. Mihai *et al.*, "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2255-2291, 2022.
- [11] Y. Teranishi, Y. Saito, S. Muroto and N. Nishinaga, "JOSE: An open testbed for field trials of large-scale IoT services," *Journal of the National Institute of Information and Communications Technology*, vol. 62, no. 2, pp. 151-159, 2016.
- [12] D. Simeonidou, "Bristol is open," in *Proc. 5G Radio Technology Seminar*, London, UK, 2015, pp. 1-32.
- [13] COSMOS Testbed Main Site [online] Available: <https://www.cosmos-lab.org/> (accessed Oct. 5, 2023).
- [14] Platforms for Advanced Wireless Research – PAWR. [online] Available: <https://advancedwireless.org/> (accessed Oct. 5, 2023).
- [15] T. Chen *et al.*, "A software-defined programmable testbed for beyond 5G optical-wireless experimentation at city-scale," *IEEE Network*, vol. 36, no. 2, pp. 90-99, 2022.
- [16] Z. Wang, D. C. Kilper and T. Chen, "Open EDFA gain spectrum dataset and its applications in data-driven EDFA gain modeling," *Journal of Optical Communications and Networking*, vol. 15, no. 9, pp. 588-599, 2023.
- [17] Z. Wang *et al.*, "Optical signal spectrum prediction using machine learning and in-line channel monitors in a multi-span ROADM system," in *Proc. European Conference on Optical Communication*, Basel, Switzerland, 2022, pp. 1-4.