Optimizing Gain Shaping Filters with Neural Networks for Maximum Cable Capacity under Electrical Power Constraints

Junho Cho(1)*, Greg Raybon(1), Ellsworth Burrows(1), Jean-Christophe Antona(2), Nicolas Fontaine(1), Roland Ryf(1), Haoshuo Chen(1), Sethumadhavan Chandrasekhar(1), Erixhen Sula(3), Samuel Olsson(4), Steve Grubb(5), and Peter Winzer(1)

(1) Nokia Bell Labs, Holmdel, NJ, USA, *junho.cho@nokia-bell-labs.com,
(2) Alcatel Submarine Networks, Nozay, France
(3) EPFL, Lausanne, Switzerland
(4) Nokia Corporation, Murray Hill, NJ, USA
(5) Facebook, Menlo Park, CA, USA

Abstract *We experimentally demonstrate capacity gains of up to 23% under electrical supply power constraints in a long-haul optical fiber cable by optimizing the gain shaping filters using neural networks.*

Introduction

In submarine optical fiber cables that are commonly built under a strict *electrical supply power (ESP)* constraint, massive spatial parallelism is a key technology to maximize cable capacity[1],[2]. It is because the *aggregate cable capacity* grows linearly with the number of spatial paths, and over-compensates for the logarithmically reduced spectral efficiency of each spatial path. A critical figure of merit in such systems is the *power efficiency*, defined as the capacity C per Watt of ESP $\mathcal{P}_E$, both per spatial path, i.e., $\mathcal{m} = C/\mathcal{P}_E$ [1]. In our previous work[3], we experimentally demonstrated a gain of 19% in $\mathcal{m}$ on a 744-km optical fiber link by removing the *gain flattening filters (GFFs)* from the link, and by optimally allocating the launch power across the C-band, whose optimal shape was found by *neural networks (NNs)*[4]. However, this approach is not scalable with distance, since complete removal of the GFFs causes the erbium-doped fiber amplifiers (EDFAs) to provide useful gains only in a very narrow spectral band at long distances. In this paper, therefore, we use gain *shaping* filters (GSFs) instead of GFFs after every EDFA, in order to maximize $\mathcal{m}$ at transoceanic distances. We build a simulation tool using NNs, which shows much better accuracy than state-of-the-art physical modeling software[3], to predict the output power spectral density (PSD) of a submarine cable that has arbitrarily shaped GSFs. Compared to the NNs of [3] that predict the PSD at the output of a specific optical link, the NNs in this paper are on a per-EDFA basis and can thus predict the evolution of the PSD over arbitrary links, enabling the optimization of in-line GSFs. We demonstrate by experiment that the GSFs optimized using the gradient-descent (GD) method with the NN-based simulation tool provide up to 23% aggregate capacity gain compared to traditional GFFs.

Methodology and Experimental Setups

Given the optical signal-to-noise ratios (OSNRs) across the spectrum of a wavelength-division multiplexed (WDM) system, the capacity is estimated as $C = 2R_s \sum_{k=1}^{K} \log_2(1 + \eta\, SNR_k)$, where SNR_k is the OSNR of the k-th of K WDM channels (normalized to one polarization and a reference bandwidth equal to the symbol rate R_s), and $\eta \leq 1$ accounts for transponder implementation penalties and fiber nonlinearities. We can neglect fiber nonlinearity in a massive spatial diversity context; e.g., the optimized system in this paper uses 10-dB lower optical power per spatial path than typical submarine systems. Although the capacity enhancement obtained by optimized GSFs becomes greater with $\eta < 1$, we use $\eta = 1$ without loss of generality. To experimentally determine SNR_k, we use the methodology developed in [3]; as shown in Fig. 1(a), a WDM transmitter (TX) emulator generates a 4-THz amplified spontaneous emission (ASE) comb using cascaded EDFAs and wavelength selective switches (WSSs). The ASE comb consists of 41 slots of 50-GHz low-power ASE (emulating noise slots) and 40 slots of 50-GHz high-power ASE in between (emulating signal slots), cf. Fig. 1(b). A subsequent EDFA and a variable optical attenuator (VOA) adjust the optical power launched into a typical submarine EDFA under test that is designed for a nominal span loss of 11 dB (excluding the GFF

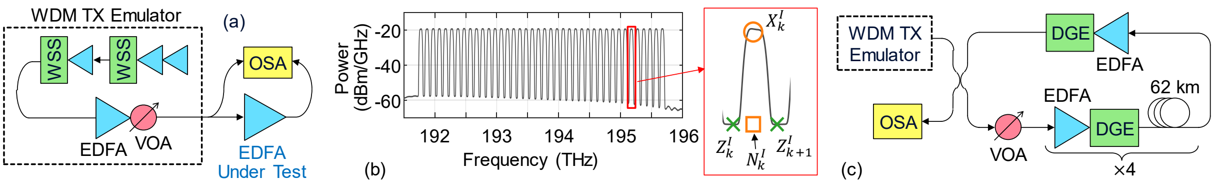

Fig. 1: (a) Experimental setup for training NNs, (b) measured optical spectrum at the TX OSA, and (c) experimental recirculating loop.

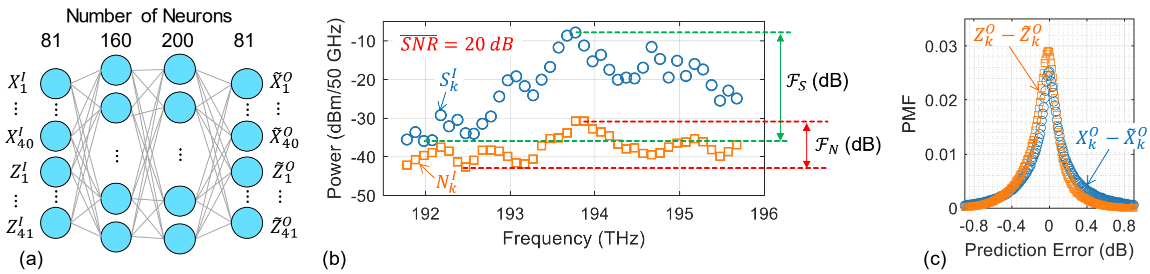

Fig. 2: (a) Structure of the NN, (b) an example of $(\boldsymbol{S}_{1:40}^I, \boldsymbol{N}_{1:41}^I)$, and (c) PMF of the prediction error.

loss) at a nominal pump current $\geq$ 450 mA for a minimum gain tilt; see [3] for the characteristics of the EDFA under test. The input and output PSDs of the EDFA under test are measured by an optical spectrum analyzer (OSA). From the measured powers X_k^I and Z_k^I, cf. inset to Fig. 1(b), with the subscript being the slot index and the superscript being the EDFA's input (I) or output (O), we estimate the signal slot's noise power by $N_k^I = (Z_k^I + Z_{k+1}^I)/2$ and the signal power by $S_k^I = X_k^I - N_k^I$, eventually obtaining $SNR_k^I = S_k^I / N_k^I$.

Training NNs with Measurement Data

In this work, NNs are used to simulate a single EDFA, i.e., to produce the prediction $(\tilde{\boldsymbol{X}}_{1:40}^O, \tilde{\boldsymbol{Z}}_{1:41}^O)$ of measured output powers $(\boldsymbol{X}_{1:40}^O, \boldsymbol{Z}_{1:41}^O)$ when arbitrary powers $(\boldsymbol{X}_{1:40}^I, \boldsymbol{Z}_{1:41}^I)$ are input to the EDFA, where $\boldsymbol{X}_{1:K}^{I/O}$ denotes $[X_1^{I/O}, \dots, X_K^{I/O}]$. The structure of the NNs selected after topology optimization is shown in Fig. 2(a), consisting of 81 input neurons $(\boldsymbol{X}_{1:40}^I, \boldsymbol{Z}_{1:41}^I)$, two hidden layers of 160 and 200 neurons each (both with sigmoid activations), and 81 output neurons $(\tilde{\boldsymbol{X}}_{1:40}^O, \tilde{\boldsymbol{Z}}_{1:41}^O)$ (with softplus activation).

We collect the training data using the constant-current mode of the EDFA at two pump currents I_{Pump} = 150 mA, 450 mA (corresponding to pump powers of 205 mW and 675 mW, respectively). For each I_{Pump}, we launch 21,200 randomly shaped ASE combs, and measure the PSDs at the input and output of the EDFA. It takes 26 days to acquire all the training data with our fully automated system. Depicted in Fig. 2(b) is an example of measured $\boldsymbol{S}_{1:40}^I$ (blue circles) and $\boldsymbol{N}_{1:40}^I$ (orange squares) with *overall SNR* in the C-band (denoted as $\overline{SNR}$) of 20 dB, and with maximum *signal and noise power excursions* (denoted as $\mathcal{F}_S$ and $\mathcal{F}_N$, cf. Fig. 2(b)) of 28 dB and 12 dB, respectively. We pay particular attention to create random ASE profiles such that: *(i)* $\overline{SNR}$ varies in $[-15, 37]$ dB, *(ii)* $\mathcal{F}_S$ varies in $[0, 38]$ dB, *(iii)* $\mathcal{F}_N$ varies in $[0, 27]$ dB, and *(iv)* the input optical power in the C-band varies in $[-3, 3]$ dB. Given $\overline{SNR}$, randomizations for (ii)-(iv) are performed independent of each other. This ensures that the NN learns any EDFA that receives a different OSNR across a submarine cable, whose input spectrum is arbitrarily shaped by the preceding EDFAs, GSFs, and fibers.

We train a set of 20 NNs for each I_{Pump} using the logcosh loss function, and take an average of 20 outputs as a prediction. Monte-Carlo cross-validation is performed with random subsampling[4]. Among the 21,200 data sets, 19,000 are used for training, and the rest for validation. Very close loss values are observed between the training and validation data, indicating no overfitting. Fig. 2(c) shows the probability mass functions (PMFs) of the prediction errors, defined as $X_k^O - \tilde{X}_k^O$ and $Z_k^O - \tilde{Z}_k^O$, whose standard deviations are 0.33 dB for X_k^O (blue circles) and 0.24 dB for Z_k^O (orange squares).

Optimizing the GSF for Maximum Capacity

We measure a net span loss of around 10 dB from 62-km Corning® Vascade® EX3000 fiber (shown in the inset to Fig. 3(a) as $\boldsymbol{\mathcal{L}}_{1:81}^{Fiber}$ for 81 signal and noise slots) including the Raman tilt at the two optical powers obtained with I_{Pump} = 150 mA and 450 mA in the loop experiment de-scribed below. Our NNs are trained with a varying optical input power, allowing 10 dB to 16 dB of fiber attenuation (i.e., 62 km to 99 km of span length); for simulation of a span length $\geq$ 62 km, we add a constant attenuation to the measured fiber loss. As shown in Fig. 3(a), our link simulator consists of a series of fiber, an EDFA, and a GSF in each span of an

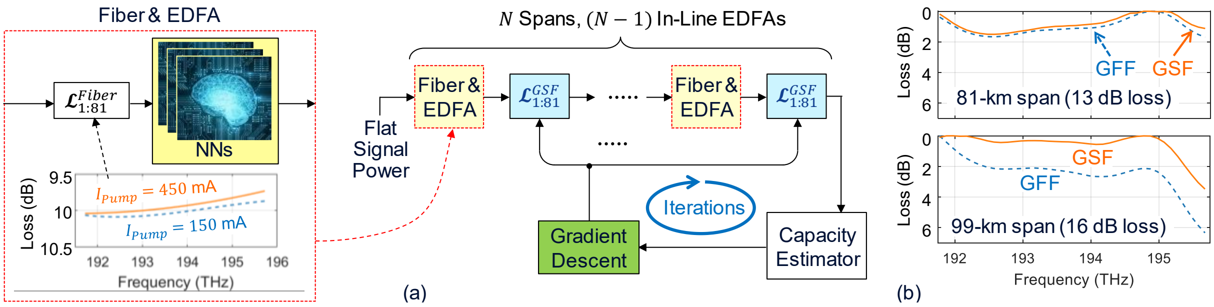

Fig. 3: (a) GD-based optimization (inset: simulator block for fiber and EDFA), and (b) shapes of the GFF and the optimized GSF.

N-span link, where N can be flexibly varied. To simplify the problem, we use identical spectral shapes for all GSFs. Without loss of generality, we launch flat signal powers into the system, then the simulator produces $(\widetilde{\boldsymbol{X}}^{O}_{1:40}, \widetilde{\boldsymbol{Z}}^{O}_{1:41})$ at the receiving end of the link, for a given GSF shape (denoted as $\boldsymbol{\mathcal{L}}^{GSF}_{1:81}$ in Fig. 3(a)). The received SNR (hence the capacity) can then be estimated as aforementioned. We apply the *approximate stochastic GD* method[3] and iteratively update $\boldsymbol{\mathcal{L}}^{GSF}_{1:81}$ towards an ascending trajectory of capacity.

Fig. 3(b) shows the shape of the GFFs (blue dashed) and the optimized GSFs (orange solid) for a 25-span link with two span losses of 13 dB and 16 dB (excluding the losses due to GFFs or GSFs) at I_{Pump} = 150 mA. The optimized GSF for a 13-dB span loss is similar to the corresponding GFF. Remarkably, however, when a gain tilt emerges for a span loss of 16 dB, the GSF no longer flattens the low-frequency regime where the EDFA does not produce as large a gain as in the rest of the C-band; instead, it attenuates the optical power much less than the GFF across the whole frequency band, cf. Fig. 3(b).

Experimental Validation

We validate the NN-based capacity maximization approach using an equivalent 5-span recirculation loop experiments, as shown in Fig. 1(c); four spans are fiber spans and the 5th span of identical loss is emulated by the losses of a loop switch and a VOA. This approach is permissible due to the low launch powers and the resulting absence of noticeable nonlinearities. Dynamic gain equalizers (DGEs) with 4-dB insertion loss are used to realize both GFFs and arbitrarily shaped GSFs. Since the GFFs and GSFs can have a negligible insertion loss when built with the typical manufacturing process for submarine cables, we translate the DGE's 4-dB insertion loss and additional 2-dB loss from 8 connectors per span into an added span length, yielding a total equivalent span length of 99 km. Fig. 4(a) shows the experimentally measured $\boldsymbol{S}^{O}_{1:40}$ (circles), $\boldsymbol{N}^{O}_{1:40}$ (squares), and $\boldsymbol{SNR}^{O}_{1:40}$ (triangles) after 5 loops (at equivalent 2,475 km), when the DGEs realize GFFs (top) and optimized GSFs (bottom), which agree very well with the NN predictions (pluses, crosses, stars), proving the high accuracy of our NN-based simulator. The GSFs enhance the SNR in almost all of the C-band at the expense of a reduced spectral bandwidth.

Fig. 4(b) shows the capacity of systems with GFFs (dashed lines) and GSFs (solid lines) at I_{Pump} = 150 mA, predicted by simulation for span lengths of 62 km, 81 km, and 99 km. Also shown are the experimental results obtained with GFFs (squares) and GSFs (circles) for 99-km equivalent spans, where significant gains of up to 23% (at 4,950 km) are observed. As expected, shorter spans offer a greater capacity, since the OSNR drop is less after every EDFA, but they also consume more ESP per distance. The capacity per ESP of the systems with optimized GSFs is shown in Fig. 4(c), obtained by simulation (lines) for 62-km, 81-km, and 99-km spans, and by experiment (circles) for 99-km equivalent spans, using I_{Pump} =150 mA and 450 mA. It can be seen that a greater capacity per ESP is achieved by: *(i)* a smaller pump current and *(ii)* a longer span length, despite the EDFAs being more deviated from their design target (12.3-dB gain, $I_{Pump} \geq$ 450 mA), cf. latest submarine cables use a small EDFA gain with short spans (e.g., 56.5 km[6]). This experimental result is in line with the finding of [2], and shows that increasing spatial parallelism is better than maximizing per-path capacity under an ESP constraint. Importantly, optimized GSFs produce the highest capacity gain for the longest span length and for the smaller pump current, i.e., when the system has the greatest power efficiency (cf. Fig. 4(c)) with massive spatial parallelism.

Conclusion

We developed a versatile NN-based simulator on a per-EDFA basis, which predicts the evolution of the optical power spectrum across transoceanic distances with high accuracy. We introduced gain shaping filters instead of gain flattening filters and experimentally showed capacity gains of up to 23% at transoceanic distances.

We acknowledge Amonics Corp for supplying the EDFAs with removable GFFs, and Corning for the loan of the EX3000 fiber used in this experiment.

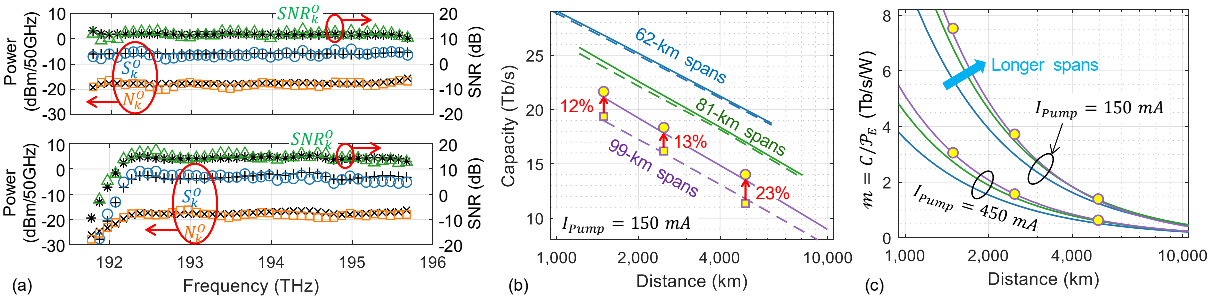

Fig. 4: (a) Measurement after 25 spans (equivalent to 2,475 km) with I_{Pump} = 150 mA using GFFs (top) and optimized GSFs (bottom), (b) capacity of systems with the GFFs (dashed, squares) and optimized GSFs (solid, circles), and (c) power efficiency of the systems with optimized GSFs.

References

[1] O. V. Sinkin et al., “SDM for power-efficient undersea transmission,” *J. Lightw. Technol.*, vol. 36, no. 2, pp. 361–371, 2017.

[2] R. Dar et al., “Cost-optimized submarine cables using massive spatial parallelism,” *J. Lightw. Technol.*, vol. 36, no. 18, pp. 3855–3865, 2018.

[3] J. Cho et al., “Supply-power-constrained cable capacity maximization using multi-layer neural networks,” *J. Lightw. Technol.*, Accepted for publication.

[4] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[5] M. Kuhn and K. Johnson, Applied predictive modeling, New York: Springer, 2013.

[6] S. Grubb et. al, “Real-time 16QAM transatlantic record spectral efficiency of 6.21 b/s/Hz enabling 26.2 Tbps capacity,” in *Proc. OFC*, San Diego, CA, USA, Mar. 2019, Paper M2E.6.