Experimental Demonstration of Learned Pulse Shaping Filter for Superchannels

Zonglong He[(1)], Jinxiang Song[(2)], Christian Häger[(2)], Alexandre Graell i Amat[(2)], Henk Wymeersch[(2)], Peter A. Andrekson[(1)], Magnus Karlsson[(1)], and Jochen Schröder[(1)]

[(1)] *Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg, Sweden*
[(2)] *Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden*
zonglong@chalmers.se

Abstract: We demonstrate a pulse shaping filter enabled by machine learning for spectral superchannels. In contrast to a 1% roll-off root-raised cosine filter, our learned filter reduces the adaptive equalizer length by 47% for the same spectral efficiency.

1. Introduction

Spectral superchannels, which combine wavelength division multiplexing (WDM) with narrow guard-bands and high-order modulation formats [1], have attracted considerable attention due to their potential to achieve high spectral efficiency (SE) at bandwidths larger than the electronic transceiver bandwidth. One key factor to improve SE is minimizing the spectral gap between adjacent channels and eliminating inter-channel interference (ICI). Narrow spacing relies on low roll-off factors, which is difficult to achieve in practice due to finite filter memory and hardware constraints [2]. Various experiments have been proposed to optimize the roll-off factors of the root-raised cosine (RRC) filter to reduce ICI in a WDM system [3–5].

Recently, machine learning (ML) techniques using neural networks (NNs) have been proposed to mitigate transmitter [6, 7] and transmission impairments [8] in fiber communication systems. End-to-end learning based autoencoder (AE) has been proposed to optimize transceivers jointly [9]. However, the NNs proposed in [6, 7] were learned in a single wavelength system by ignoring the ICI, and may not suit narrowly-spaced superchannel systems. Our method in [10] used an AE to perform a joint optimization of the constellation, pulse shaping (PS) filter and modulator nonlinearity mitigation in a simulated superchannel system. In contrast to RRC-filter-shaped superchannels which can be sub-optimum, systems shaped with a ML filter can optimize the signal spectra and then further minimize the ICI.

In this work, we experimentally verify the concepts from [10], taking advantage of the learned PS filter in a superchannel system. We apply a 64QAM signal shaped with the RRC and learned filter to a comb-based superchannel system with three carriers spaced at 25 GHz. For the same DSP complexity, the learned filter improves the spectral efficiency by 2.3% and 0.8% compared with a 10% and 1% roll-off RRC filter, respectively. Surprisingly, in contrast to a 1% roll-off RRC filter, the proposed method can reduce the adaptive equalizer length by 47% at SE 5.684 bits/s/Hz, enabling lower DSP complexity and power consumption.

2. Pulse shaping filter training

The training is finished in simulated model as shown in Fig. 1 (a). The system consists of three 64QAM wavelength channels, where the PS filter is implemented by an NN as in [10]. In each channel, the baseband symbols $\boldsymbol{x} \in \mathbb{C}^N$ are first upsampled with rate R and then convolved with the trainable PS filter, after which a digital pre-distortion block (i.e., arcsine function and clipping) is employed to compensate for the hardware imperfections. Then, the pre-distorted signals $\boldsymbol{s} \in \mathbb{C}^{NR}$ are fed to a digital-to-analog converter with 6 effective-number-of-bits and amplified by a linear power amplifier to drive the IQ-modulator. The modulated signals of the 3 sub-channels are then combined to form the superchannel. The channel model under consideration is the optical-back-to-back channel, and therefore only AWGN noise is simulated as the booster amplifier noise. At the receiver, the channel observations are first passed through a low-pass filter and then convolved with a matched filter (MF). Instead of employing a trainable NN as MF, we apply a RRC filter to enable a faster convergence for the PS filter NN. The resulting signals $\boldsymbol{u} \in \mathbb{C}^{NR}$ are then downsampled with rate R to get the recovered baseband symbols $\boldsymbol{y} \in \mathbb{C}^N$. Training is performed by minimizing the mean-squared-error (MSE) loss, i.e., $\mathcal{L}_{\mathrm{MSE}}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N} |y_k - x_k|^2$, of the center channel by employing the commonly used Adam optimizer [11]. The length of the NN-based PS filter is 201, while 5000 iterations are used for each training.

3. Experimental setup

The comb-based superchannel system is presented in Fig. 1 (b). The electro-optic frequency comb (EO-comb) consists of two cascaded phase modulators and an intensity modulator driven by a 25 GHz RF clock. An optical

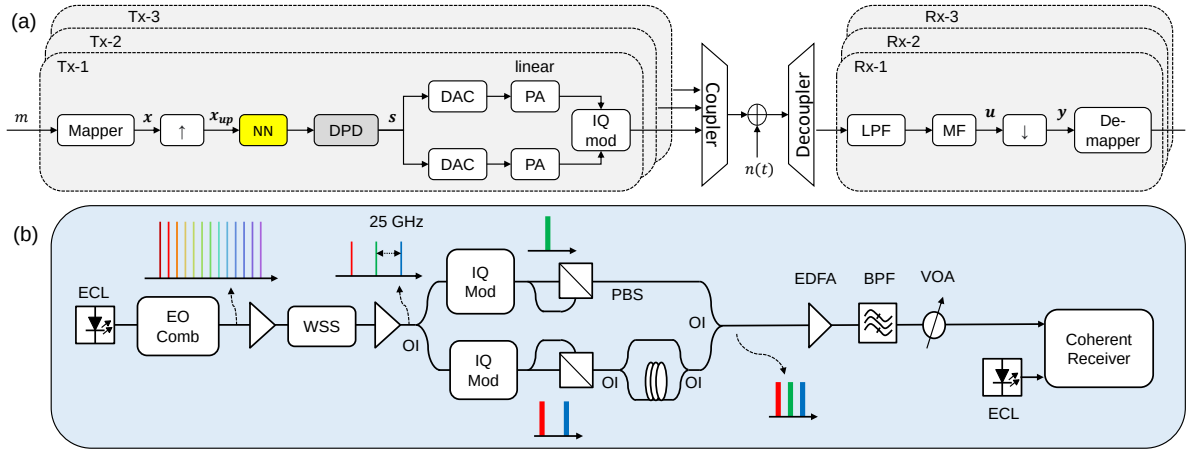

Fig. 1: (a) Block diagram of the training model; (b) Experimental setup of the superchannel system. NN: neural network; DPD: digital pre-distortion; DAC: digital-to-analog converter; PA: power amplifier; LPF: low-pass filter; MF: matched filter; ECL: external cavity laser; EO-comb: electro-optic frequency comb; WSS: wavelength-selective switch; OI: optical interleaver; PBS: polarization beam splitter; VOA: variable optical attenuator; BPF: band-pass filter.

carrier at 1550.12 nm is fed to the EO-comb to generate 25 GHz equally-spaced comb lines. We amplify the comb output using two erbium-doped fiber amplifiers (EDFAs). The wavelength selective switch filters out the three central carriers (1549.92 nm, 1550.12 nm, and 1550.32 nm) and flattens the comb. A 25 GHz optical interleaver (OI) is then used to separate the carriers into the center and side channels. By using an arbitrary waveform generator, we modulate each carrier with a 64QAM signal pulse-shaped with the RRC (10% or 1% roll-off) or the learned filter. Independent data is modulated on each channel. We vary the symbol rate from 23 GBaud to 25 Gbaud to adjust the effective guard-bands of the superchannel. Polarization multiplexing is employed by using the split-delay-combine method [12] for each channel. Signal decorrelation between the side channels is induced by another pair of OIs. The modulated signals are then combined to generate the superchannel before being amplified by a booster followed by a 0.25 nm band-pass filter and a variable optical attenuator to control the received power.

The filtered signal and local oscillator line are combined in a coherent receiver and then sampled by a 80 GS/s real-time oscilloscope before being processed by offline DSP, which performs low-pass filtering, resampling, front-end imbalance compensation, carrier recovery and dynamic equalization [13].

4. Results

During the training, two sets of ML filters (ML-0.1 and ML-0.01) are learned, where 0.1 and 0.01 indicates the roll-off of the RRC matched filter at the receiver. The ML filter operated at each symbol rate is trained under the targeted channel spacing ratio, defined by the ratio between carrier spacing (25 GHz) and symbol rate.

The bit error rate (BER) and SE for a superchannel shaped with a 10% roll-off RRC filter is shown in Fig. 2 (a) and (b), respectively. The performance of the ML-0.1 and ML-0.01 filters are evaluated. Compared to the ML-0.01 filters, the ML-0.1 filters perform better from 23 Gbaud to 25 Gbaud. Therefore we only show the performance of the ML-0.1 filters, indicated by the red curve, in Fig. 2 (a)-(b) and Fig. 3. The frequency response of a learned filter, targeted at the channel spacing ratio of 1.04 and the 10% roll-off RRC matched filter, is shown in Fig. 2 (c). The learned filter has a narrower spectrum compared with the RRC filter and therefore reduces the ICI. We use the single channel system shaped with RRC filter as a reference. The gap between the single and three channel system is probably induced by the hardware imperfections and inter-channel crosstalk. A 45-tap equalizer trained

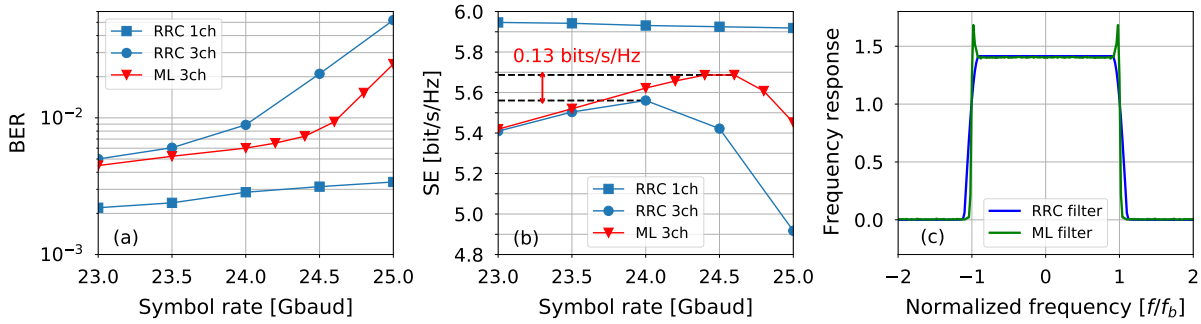

Fig. 2: Comparison between ML-based filter and RRC filter with 10% roll-off. (a) BER versus symbol rate; (b) Spectral efficiency versus symbol rate; (c) Frequency response of learned filter when the channel spacing ratio is 1.04 and the roll-off factor of the matched filter is 10%. The frequency response of the 10% roll-off RRC filter is presented as a reference.

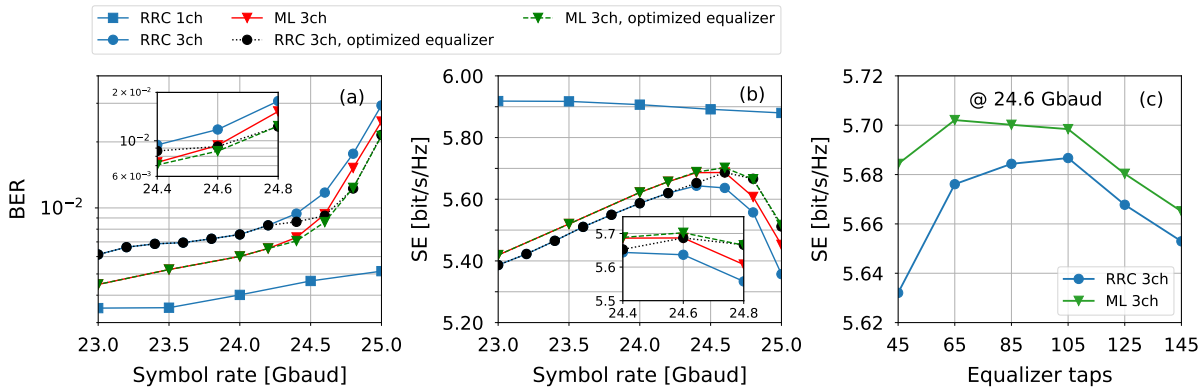

Fig. 3: Comparison between ML-based filter and RRC filter with 1% roll-off. (a) BER versus symbol rate; (b) Spectral efficiency versus symbol rate; (c) Spectral efficiency versus the equalizer length at the symbol rate of 24.6 Gbaud.

with adaptive step-size is applied to perform channel equalization. By employing the learned filter, we achieve a 0.13 bits/s/Hz increase, corresponding to 2.3% improvement, in spectral efficiency. The optimum symbol rate increases from 24 Gbaud to 24.6 Gbaud, further minimizing the spectral gap between adjacent channels.

The system performance for pulse shaping with a 1% roll-off RRC filter is illustrated in Fig. 3. We start from using the same equalizer length (i.e., 45) to evaluate the system performance. As shown in Fig. 3 (b), the spectral efficiency is improved from 5.64 bits/s/Hz (24.4 Gbaud, blue curve) to 5.68 bits/s/Hz (24.6 Gbaud, red curve), resulting in a 0.8% performance increase. However, the SE of the RRC-filter-shaped superchannels is significantly enhanced by optimizing the number of equalizer taps. As presented in Fig. 3 (c), we sweep the equalizer length at 24.6 Gbaud for SE optimization. The SE gap between the RRC and the learned filter decreases to 0.015 bits/s/Hz, while the optimum number of taps is 105 and 65, respectively. Moreover, while the SE gain is minor, it is possible to significantly reduce the equalizer length without penalty. For example, the required number of taps at SE 5.684 bits/s/Hz is reduced from 85 to 45, corresponding to a 47% reduction. For large number of taps (i.e., 145), the performance of the equalizer is degraded due to suboptimal equalizer convergence. Nevertheless, the performance of the equalizer is almost saturated at the equalizer length of 105. The black curve in Fig. 3 (a) and (b) indicates the performance achieved by optimizing the equalizer length for each symbol rate. Currently the learned filter has sub-optimum performance due to the difference between training model and experimental setup.

5. Conclusion

We experimentally demonstrated a pulse shaping filter enabled by machine learning in a comb-based superchannel system. For the same DSP complexity, the proposed method improves the spectral efficiency by 2.3% and 0.8% compared with a 10% and 1% roll-off RRC filter, respectively. More surprisingly, the ML-based pulse shaping filter induces a 47% reduction in the required number of equalizer taps at SE 5.684 bits/s/Hz.

Acknowledgements: This work was supported by the Knut and Alice Wallenberg Foundation (KAW 2018.0090).

References

1. P. J. Winzer *et al.*, "From Scaling Disparities to Integrated Parallelism: A Decathlon for a Decade," *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1099-1115, Mar. 2017.
2. G. Bosco *et al.*, "On the Performance of Nyquist-WDM Terabit Superchannels Based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM Subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53-61, Jan. 2011.
3. Y. Yue *et al.*, "Transmitter Skew Tolerance and Spectral Efficiency Tradeoff in High Baud-rate QAM Optical Communication Systems," *Opt. Express.*, vol. 26, no. 12, pp. 15045-15058, Mar. 2018.
4. J. Fickers *et al.*, "Design rules for pulse shaping in PDM-QPSK and PDM-16QAM Nyquist-WDM coherent optical transmission systems," *ECOC*, 2012.
5. R. Maher *et al.*, "Digital Pulse Shaping to Mitigate Linear Crosstalk in Nyquist-spaced 16QAM WDM Transmission Systems," *ECOC*, 2014.
6. G. Paryanti *et al.*, "Recurrent Neural Network for Pre-distortion of Combined Nonlinear Optical Transmitter Impairments with Memory," in *Proc. SPPCom*, 2018.
7. V. Bajaj *et al.*, "Single-channel 1.61 Tb/s Optical Coherent Transmission Enabled by Neural Network-based Digital Pre-distortion," *ECOC*, 2020.
8. C. Häger *et al.*, "Model-Based Machine Learning for Joint Digital Backpropagation and PMD Compensation,", *OFC*, 2020.
9. T. O'Shea et al., "An Introduction to Deep Learning for the Physical Layer," *Trans. on Cogn. Commun. and Netw.*, vol. 3, no. 4, pp. 563–575, 2017.
10. J. Song *et al.*, "End-to-end Autoencoder for Superchannel Transceivers with Hardware Impairment," *OFC*, 2021.
11. D. P. Kingma *et al.*, "Adam: A method for stochastic optimization,", *Proc. ICLR*, 2014.
12. M. Mazur *et al.*, "Overhead-optimization of Pilot-based Digital Signal Processing for Flexible High Spectral Efficiency Transmission," *Opt. Express.*, vol. 27, no. 17, pp 24654-24669, Aug. 2019.
13. J. Schröder *et al.*, "QAMPy a DSP chain for optical communications, DOI: 10.5281/zenodo.1195720".