Recurrent Neural Network Soft-Demapping for Nonlinear ISI in 800Gbit/s DWDM Coherent Optical Transmissions

Maximilian Schaedler[(1,2)], Fabio Pittalà[(1)], Georg Böcherer[(1)], Christian Bluemm[(1)], Maxim Kuschnerov[(1)], and Stephan Pachnicke[(2)]

[(1)] Huawei Munich Research Center, Riesstr. 25, 80992 Munich, Germany
[(2)] Kiel University (CAU), Chair of Communications, Kaiserstr. 2, 24143 Kiel, Germany
maximilian.schaedler@huawei.com

Abstract *Optical transmission systems suffer from ISI induced by nonlinear components. As a countermeasure, a recurrent neural network soft-demapper is proposed and benchmarked against a combination of Volterra equalizer, noise whitening filter and soft-output Viterbi-algorithm. A substantial gain up to 1.52dB in 800G-DWDM-600km measurements is achieved.*

Introduction

Modern communication networks build upon a backbone of optical systems use fiber as transmission medium. Due to its physical properties, including the Kerr effect, fiber transmissions are always affected by a certain amount of nonlinearities[1]. However, this is not the only major source of nonlinear distortion in optical communication systems. Regardless of the individual technology choices, the optical transceiver architecture will always comprise optical/electrical (O/E) components with nonlinear transfer characteristics, often with memory effects, which results in intersymbol interference (ISI). Their impact on the achievable capacity is aggravated towards higher data rates. Advanced digital signal processing (DSP) can reduce the impact of these imperfections. Some demonstrations have shown that Volterra nonlinear equalizers (VNLE) are able to mitigate both, fiber nonlinearities[2] and component nonlinearities[3],[4], but, only up to a certain extent. In recent years, machine learning (ML) methods, especially neural networks (NN), have demonstrated excellent performance gains in various applications. In the optical communication community, NNs[5],[6] and convolution neural networks (CNNs)[7] have been used to tackle nonlinear distortions.

This contribution suggests a bidirectional recurrent neural network[8] soft-demapper (BRNN-SD) capable of nonlinear ISI compensation. In contrast to previous publications[9],[10], which used symbol-wise optimization, our approach is based on bit-wise equalization. The performance of the BRNN-SD is benchmarked against a stacked combination of a VNLE and a Forney's detector[11] structure, which is capable to handle colored noise and ISI effects. The Forney's detector comprises a symbol-spaced whitening filter (WF) and a channel-model-based soft-output Viterbi algorithm (SOVA)[12] and we refer to its by WF + SOVA in the following. The evaluation is done with offline processing on basis of captured samples in the case of optical back-to-back (BtB) and 600km 32-channel dense wavelength division multiplexing (DWDM) fiber transmission of coherent 92GBd 800Gb/s dual polarization (DP)-32QAM.

Bidirectional Recurrent Neural Network Soft-Demapper

Estimating the embedded soft bits from received signals with ISI effects can be performed using deep NN architectures through supervised learning[9],[13]. Common feed-forward NNs with inputs comprising current and historical captures, often implemented as tapped delay lines, are limited in functionality since they rely on a predetermined memory architecture. A more powerful architecture can be implemented via BRNNs[8],[14], which enable self learning of memory utilization during the training phase. The particular memory units of BRNNs are capable of keeping the relevant information. Fig. 1 illustrates the structure of the pro-

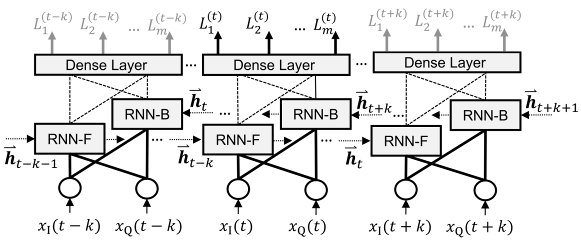

Fig. 1: BRNN Soft-Demapper: The final estimated soft information is provided by the centered dense output.

posed BRNN-SD for a sequence of length $2k+1$, where k denotes the time steps[14]. The received symbol sequences are fed in the forward direction into one RNN-F unit and in backward direction

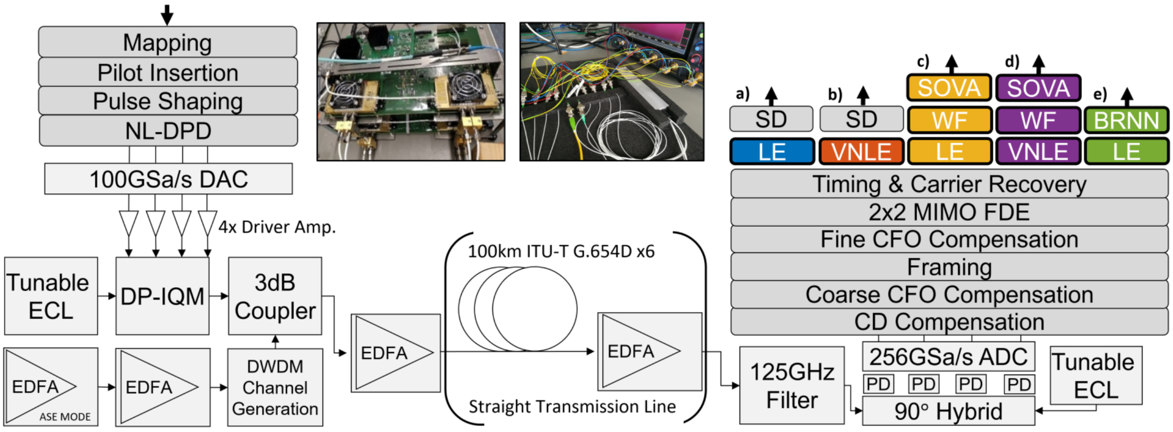

Fig. 2: Schematic of the DWDM experimental setup including the Tx and RX DSP architectures with a) Linear Equalizer (LE) & soft-demapper (SD) b) Volterra Equalizer (VNLE) & SD c) LE, whitening filter (WF) & accompanied by a soft-output Viterbi-algorithm (SOVA) d) VNLE, WF & SOVA and e) LE & BRNN soft-demapper.

into a RNN-B unit[15]. The exchanging memory state vectors are denoted by $\overrightarrow{\boldsymbol{h}}$ and $\overleftarrow{\boldsymbol{h}}$, respectively. The outputs are then concatenated and fed into multiple fully connected dense layers terminated by a final linear layer of length equal to m, the number of bits per symbol. This architecture accounts for past and future symbols when estimating the centered log-likelihood ratios (LLRs), $\hat{\boldsymbol{L}}(t) = \{L_1^{(t)}, ..., L_m^{(t)}\} \in \mathcal{R}^m$, corresponding to the t^{th} received symbol.

During training, known transmission sequences of bits $\boldsymbol{b}(t) \in \mathcal{R}^m$, are used to obtain the target LLRs[13]. Therefore, the cross-entropy loss function[16] is directly applied on the LLRs, namely

$$\mathcal{L}_{\text{LLR}}(\boldsymbol{b}, \hat{\boldsymbol{L}}) = \log(1 + \exp(-(1 - 2\boldsymbol{b})\hat{\boldsymbol{L}}). \quad (1)$$

Experimental Setup and DSP Configurations

The experimental setup including the offline DSP stack is shown in Fig. 2. The channel under test (CUT) carries a 96GBd DP-32QAM with gross data rate of 960Gb/s. Assuming 15% overhead for FEC, 3.47% for pilot symbols, framing and other training sequences, the net bit rate is 800Gb/s.

At the transmitter, the DSP inserts a CAZAC training sequence, which is used for framing, carrier frequency offset estimation, 2×2 MIMO equalization and residual chromatic dispersion compensation[17]. To compensate for transmitter impairments, a static nonlinear digital predistortion (NL-DPD) is applied after pulse-shaping. The four 100GSa/s Micram digital-analog converters (DACs) with 40GHz bandwidth generate a repeated pattern of 76800 samples. Subsequently, SHF S804A amplifiers (60-GHz bandwidth) are used. In the optical domain, a tunable 100kHz external cavity laser (ECL) source generates a continuous wave signal which is modulated by a Fujitsu DP-I/Q modulator (32 GHz bandwidth). The WDM system is emulated by generating 31 channels loaded with noise and shaped with a 96GHz root-raised cosine (RRC) optical filter with 0.2 roll-off factor having central frequencies ranging from 192.095THz (1529.774nm) to 195.970THz (1560.633nm) on a 125GHz grid. The channels are multiplexed together with the CUT by using a 3dB coupler and sent to an Erbium-doped fiber amplifier (EDFA) acting as a booster. The DWDM signal is launched into a transmission line consisting of six spans, each of 100km length. In use are pure-silica core fibers, with ultra low attenuation, compliant to ITU-T G.654D[18]. The seven C-band EDFAs operate with input and output powers of 4dBm and 20dBm, respectively. Fig. 3-left depicts the spectrum of the 96GBd DP-32QAM CUT, which wavelength is varied over the whole band. Fig. 3-right illustrates the DWDM spectra at the input and output of the 600km transmission line. The spectrum at the output of the booster amplifier is fairly flat with a ripple below 0.5dB. After the transmission line, at the pre-amplifier, a ripple of 5dB is observed. The receiver consists of an optical 90°-hybrid and four 70GHz balanced

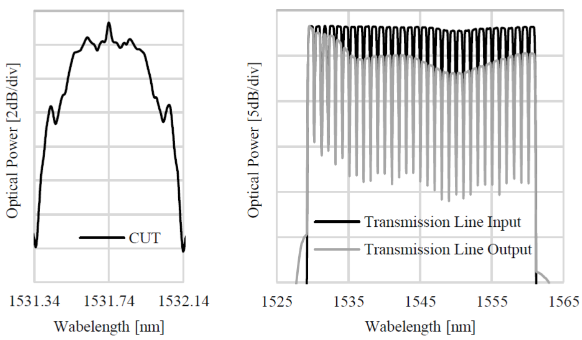

Fig. 3: DWDM experiment: spectrum of the 96GBd DP-32QAM CUT (left) and DWDM spectrum (right).

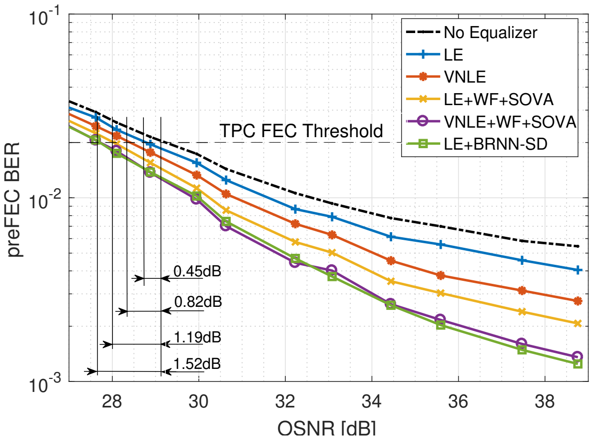

Fig. 4: 800G-96GBd-DP-32QAM optical BtB performance.

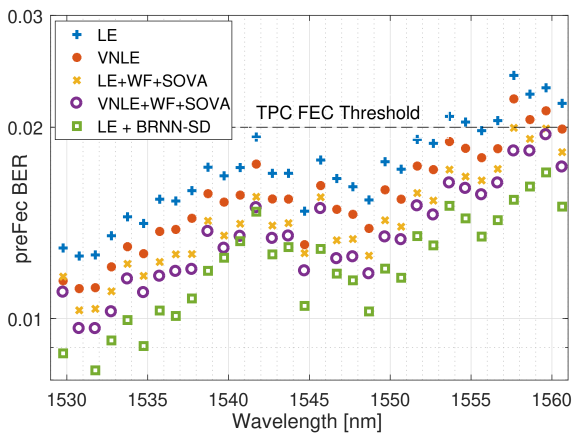

Fig. 5: 800G-96GBd-DP-32QAM Pre-FEC BER for the 32-channel DWDM system over a 600km G.654D fiber link.

photodiodes. The electrical signals are digitized using four 10-bits analog-digital converters operating at 256GSa/s with 110GHz bandwidth. In order to compensate nonlinearities, ISI, memory, as well as bandwidth limitation, the DSP stack includes next to the classical coherent signal recovery blocks several further options for enhanced signal processing, i.e., VNLE[19], WF + SOVA[11] and the proposed BRNN-SD. After the timing and carrier recovery block, five architectures are selectable: **a)** Linear equalizer (LE) **b)** VNLE **c)** LE accompanied by WF + SOVA. In particular, the WF has one tap with impulse response $1 + \alpha D$, where D denotes the tap delay and α the weight, with α optimized w.r.t. BER. The SOVA assumes Gaussian additive noise with mean and variance adapted individually for each of the 32 QAM symbol, i.e., for QAM symbol x, the noise density is $\mathcal{N}(\mu(x), \sigma^2(x))$. The means and variances are trained data aided using one frame with received and transmitted 32QAM symbols. **d)** VNLE accompanied by WF + SOVA and finally **e)** LE plus a BRNN-SD. In a lab setup, a long LE is required to handle the pronounced cable reflections. However, in an integrated setup, it can be omitted and included directly in the BRNN-SD.

The LE, VNLE, SOVA channel model and the BRNN-SD are trained on the first received frame. Once trained, all parameters remain static for the following frames. This ensures strict separation of training and testing data.

Results and Discussion

The performance of the individual DSP architectures is analysed on the optical BtB measurements and shown in Fig. 4. All architectures are optimized regarding BER performance (e.g. the equalizers are optimized with respect to the number of taps). The black dashed line indicates the baseline where no additional equalizer at the receiver is applied. By switching on the LE (161 taps) a gain of 0.45dB is observed at the assumed FEC[18] threshold of $2 \cdot 10^{-2}$. A further gain of 0.4dB is obtained when a VNLE with nonlinear taps (11/7 taps of order 2/3) is used. The combination of LE, WF + SOVA exceed the performance of the VNLE by another 0.4dB. Altogether the required OSNR is reduced to 27.8dB resulting in a gain of 1.52dB with respect to the baseline configuration. The green curve indicates the performance of the BRNN-SD. It achieves the same performance as the full DSP stack, in particular no nonlinear equalization or noise whitening is required. This indicates that the BRNN-SD is capable to handle mixed signal memory and colored noise effects in time domain.

The 800G DWDM transmission results with reach of 600km are shown in Fig. 5. It can be observed that in the case of fiber transmission the BRNN-SD outperforms the conventional DSP combination of VNLE, and WF + SOVA. While in the BtB scenario the classical algorithms can handle most of the impairments, in the case of transmission over fiber, they cannot model all effects optimally. In comparison, the BRNN-SD is capable to learn the channel and to adjust its model.

Conclusions

This paper proposes a bidirectional recurrent neural network bitwise soft-demapper for nonlinear channels. In optical back-to-back the proposed soft-demapper matches the performance of the reference DSP, consisitng of a Volterra nonlinear equalizer accompanied by a symbol-spaced whitening filter and a soft-output Viterbi algorithm. However, in 800Gb/s 32-channel DWDM transmissions over a 600km G.654D fiber link the proposed approach substantially outperforms the reference DSP.

References

[1] K. Kikuchi, “Fundamentals of coherent optical fiber communications”, *JLT*, vol. 34, no. 1, pp. 157–179, 2015.

[2] A. Amari, O. A. Dobre, R. Venkatesan, O. S. Kumar, P. Ciblat, and Y. Jaouën, “A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems”, *Commun. Surv. Tutor*, vol. 19, no. 4, pp. 3097–3113, 2017.

[3] A. Rezania, J. Cartledge, A. Bakhshali, and W.-Y. Chan, “Compensation schemes for transmitter-and receiver-based pattern-dependent distortion”, *PTL*, vol. 28, no. 22, pp. 2641–2644, 2016.

[4] J. C. Cartledge, “Volterra equalization for nonlinearities in optical fiber communications”, in *SPPCom*, JOSA, 2017, SpTu2F–2.

[5] O. Sidelnikov, A. Redyuk, and S. Sygletos, “Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems”, *OE*, vol. 26, no. 25, pp. 32 765–32 776, 2018.

[6] C. Bluemm, M. Schaedler, S. Calabrò, G. Charlet, C. Xie, F. Pittalà, and M. Kuschnerov, “Equalizing nonlinearities with memory effects: Volterra series vs. deep neural networks”, in *45th ECOC*, IET, 2019, pp. 1–4.

[7] P. Li, L. Yi, L. Xue, and W. Hu, “56 Gb/s IM/DD PON based on 10G-class optical devices with 29 dB loss budget enabled by machine learning”, in *OFC*, IEEE, 2018, pp. 1–3.

[8] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures”, *Neural networks*, vol. 18, no. 5-6, pp. 602–610, 2005.

[9] N. Farsad and A. Goldsmith, “Neural network detection of data sequences in communication systems”, *Trans. Signal Process*, vol. 66, no. 21, pp. 5663–5678, 2018.

[10] X. Dai, X. Li, M. Luo, Q. You, and S. Yu, “LSTM networks enabled nonlinear equalization in 50-gb/s pam-4 transmission links”, *Applied optics*, vol. 58, no. 22, pp. 6079–6084, 2019.

[11] G. Forney, “Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference”, *Trans. Inf. Theory*, vol. 18, no. 3, pp. 363–378, 1972.

[12] J. Hagenauer, “A viterbi algorithm with soft-detection outputs and its applications”, *Trans. Inf. Theory*, 1989.

[13] M. Schaedler, S. Calabrò, F. Pittalà, C. Bluemm, M. Kuschnerov, and S. Pachnicke, “Neural network-based soft-demapping for nonlinear channels”, in *OFC*, IEEE, 2020, pp. 1–3.

[14] S. Hochreiter and J. Schmidhuber, “Long short-term memory”, *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[15] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks”, *Trans. Signal Process*, vol. 45, no. 11, pp. 2673–2681, 1997.

[16] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[17] F. Pittala, I. Slim, A. Mezghani, and J. A. Nossek, “Training-aided frequency-domain channel estimation and equalization for single-carrier coherent optical transmission systems”, *JLT*, vol. 32, no. 24, pp. 4849–4863, 2014.

[18] F. Pittala, M. Schaedler, C. Bluemm, G. Goeger, S. Calabrò, M. Kuschnerov, and C. Xie, “800ZR+ DWDM demonstration over 600km G. 654D fiber enabled by adaptive nonlinear triplex equalization”, in *OFC*, JOSA, 2020, M4K–5.

[19] C. Bluemm, M. Schaedler, M. Kuschnerov, F. Pittalà, and C. Xie, “Single carrier vs. OFDM for coherent 600Gb/s data centre interconnects with nonlinear equalization”, in *2019 OFC*, IEEE, 2019, pp. 1–3.