



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## Wideband Time-Domain Digital Backpropagation via Subband Processing and Deep learning

Downloaded from: <https://research.chalmers.se>, 2019-08-13 09:40 UTC

Citation for the original published paper (version of record):

Häger, C., Pfister, H. (2018)

Wideband Time-Domain Digital Backpropagation via Subband Processing and Deep learning  
2018 European Conference on Optical Communication (ECOC), 2018-September

<http://dx.doi.org/10.1109/ECOC.2018.8535251>

N.B. When citing this work, cite the original published paper.

# Wideband Time-Domain Digital Backpropagation via Subband Processing and Deep Learning

Christian Häger<sup>(1,2)</sup> and Henry D. Pfister<sup>(2)</sup>

<sup>(1)</sup> Department of Electrical Engineering, Chalmers University, Sweden

<sup>(2)</sup> Department of Electrical and Computer Engineering, Duke University, USA, ✉ [ch303@duke.edu](mailto:ch303@duke.edu)

**Abstract** We propose a low-complexity sub-banded DSP architecture for digital backpropagation where the walk-off effect is compensated using simple delay elements. For a simulated 96-Gbaud signal and 2500 km optical link, our method achieves a 2.8 dB SNR improvement over linear equalization.

## Introduction

Real-time digital backpropagation (DBP) based on the split-step Fourier method (SSFM) is widely considered to be impractical due to the complexity of the chromatic dispersion (CD) steps. To address this problem, finite impulse response (FIR) filters may be used instead of fast Fourier transforms (FFTs) to perform time-domain CD filtering<sup>1–7</sup>. Indeed, the FIR filters can be as short as 3 taps per SSFM step, provided that the step size is sufficiently small (i.e., many steps are used) and the filters in all steps are jointly optimized<sup>6</sup>.

The complexity of time-domain DBP (TD-DBP) is dominated by the total number of CD filter taps in all steps. Recent work has focused on relatively narrowband signals (e.g., 10 Gbaud in<sup>6</sup> and 20 Gbaud in<sup>3–5</sup>) for which the overall CD memory is low. Since the memory increases quadratically with bandwidth, it is not clear if TD-DBP can be scaled gracefully also to more wideband signals.

In this paper, we consider a 96-Gbaud signal where the delay spread per 100 km amounts to 125 symbol periods. It is shown that TD-DBP can still offer a good performance–complexity trade-off by leveraging digital subband processing. In particular, the group delay difference in different subbands can be compensated almost entirely using delay elements. A fractional delay filter is only needed after the last SSFM step.

## Subband Processing and Related Work

Subband processing has been previously studied for both linear<sup>8–11</sup> and nonlinear<sup>12–14</sup> impairment compensation. The idea is to split the received signal into  $N$  parallel signals using a filter bank. Assuming a bandwidth reduction by  $N$ , the delay spread per subband signal is reduced by  $N^2$ . This can allow for significant complexity savings.

We consider a uniformly modulated filter bank as shown in Fig. 1. The subband signals are obtained by filtering a downconverted version of  $u[k]$  with a prototype filter  $A(z)$ , where  $w_i \triangleq \frac{2\pi i}{NT}$  is the

frequency shift. The signals are then downsampled by  $K < N$  and jointly processed. Finally, a synthesis filter bank reassembles the signal  $\tilde{u}[k]$ . Certain subbands may be inactive if they do not contain useful signal components. Active subbands are indexed symmetrically around the central subband according to  $i \in \{-S, \dots, S\} \triangleq \mathcal{S}$ .

*Example 1:* Consider a 96-Gbaud signal sampled at  $1/T = 192$  GHz. For  $N = 12$  subbands, most of the spectrum falls within the central 7 subbands, see Fig. 1. Thus, one may set  $S = 3$ .  $\triangle$

## Proposed DSP Architecture

A theoretical foundation for DBP based on subband processing can be obtained by inserting the split-signal assumption  $u = \sum_i u_i$  into the NLSE. This leads to a set of coupled equations which can then be solved numerically. Our approach is based on the SSFM proposed in<sup>15</sup>. The method is essentially equivalent to the standard SSFM for each subband, except that all sampled intensity signals are jointly processed with a multiple-input multiple-output (MIMO) filter prior to the nonlinear phase rotation step. This accounts for cross-phase modulation (XPM) between subbands but not four-wave mixing (FWM) because no phase information is exchanged. The method can also be used for DBP of noncoherent subband signals,

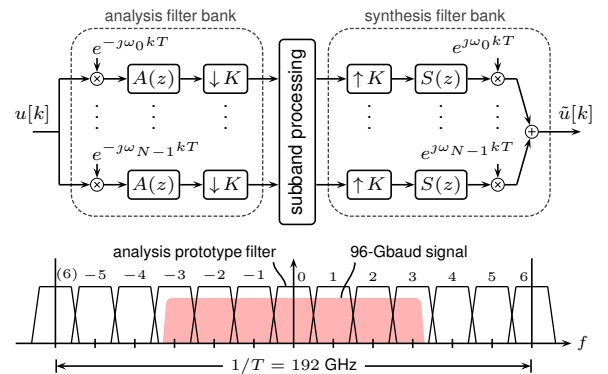


Fig. 1: Filter bank (top) and signal spectrum (bottom)

e.g., in wavelength division multiplexing scenarios with different local oscillators. This was done in <sup>12</sup>.

Fig. 2 shows the proposed architecture, which consists of the following three main components:

- (A) Short filters compensate for pulse broadening in each subband and step  $\ell = 1, \dots, M$ .
- (B) Delay elements are used to compensate the group delay difference in different subbands.
- (C) The MIMO filtering is performed in the time domain using sparse tensor decompositions.

Compared to <sup>12,15</sup>, no FFT/IFFT pairs are used. In the following, the individual components (A)–(C) in Fig. 2 are described in more detail.

### (A) Pulse-Broadening FIR Filters

The frequency response of an ideal CD compensation filter is  $H(\omega) = e^{j\kappa\omega^2}$  where  $\kappa \triangleq \frac{\beta_2}{2}\xi$  and  $\xi$  is the propagation distance. Since the subband signals are downconverted relative to the carrier frequency, the filter responses have to be shifted as well. The ideal response for subband  $i$  is

$$H(\omega + \omega_i) = e^{j\kappa\omega^2} e^{j\kappa 2\omega\omega_i} e^{j\kappa\omega_i^2} = H(\omega) D_i(\omega) e^{j\phi_i},$$

where  $D_i(\omega) \triangleq e^{-j t_i \omega}$  with  $t_i \triangleq -\beta_2 \xi \omega_i$  compensates the walk-off relative to the central subband, and  $\phi_i \triangleq \beta_2 \xi \omega_i^2 / 2$ . The filters  $H^{(\ell)}(z)$  correspond to  $H(\omega)$  and compensate for pulse broadening which is independent of  $i$ . Thus, the same filter can be used in all subbands. Moreover, the filters are symmetric since  $H(\omega)$  is symmetric, allowing for a folded DSP implementation with  $4(L+1)$  real multiplications (RMs) assuming  $2L+1$  complex filter taps. Different filters are used in different steps (even if the step size is the same) to avoid accumulating truncation errors that arise from approximating  $H(\omega)$  with finite-length filters <sup>6</sup>.

### (B) Walk-off Delays and SSFM Step Size

The group delay  $t_i$  depends linearly on the propagation distance  $\xi$ . The step size can thus be chosen such that  $t_i$  for all  $i \in \mathcal{S}$  is an integer multiple of the subband sampling interval  $T_{\text{sub}} \triangleq KT$ :

$$T_{\text{sub}} = |t_1| \implies \xi = \frac{NKT^2}{2\pi|\beta_2|} \triangleq \delta, \quad (1)$$

where  $t_1$  is the group delay difference in two neighboring subbands. Thus, as long as the step size is an integer multiple of  $\delta$ , the walk-off can be compensated exactly using delay elements.

*Example 2:* For the parameters in Ex. 1,  $K = 8$ , and  $\beta_2 = -21.7 \text{ ps}^2/\text{km}$ , we have  $\delta \approx 19.1 \text{ km}$ .  $\triangle$

It is clear that the transmission distance is not necessarily an integer multiple of  $\delta$ . Therefore, a fractional delay filter  $F_i(z)$  is inserted after the

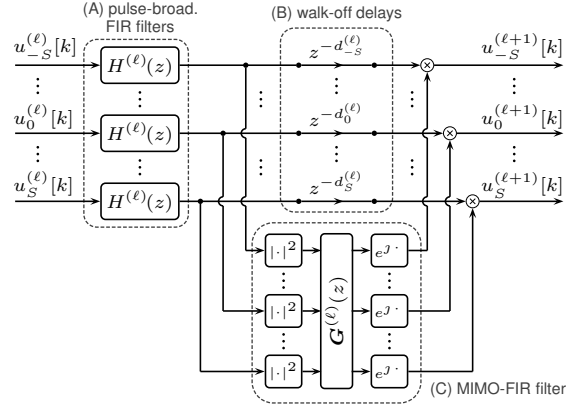


Fig. 2: Proposed DSP architecture for one SSFM step

last step to account for any remaining non-integer delay prior to the synthesis filter bank.

### (C) MIMO-FIR Filter

Let  $a_i^{(\ell)}(z)$  be the  $z$ -transform of the intensity signal after the filter  $H^{(\ell)}(z)$  in subband  $i$  and step  $\ell$  and define  $\mathbf{a}^{(\ell)}(z) \triangleq (a_{-S}^{(\ell)}(z), \dots, a_S^{(\ell)}(z))^T$ . The nonlinear phase shift is computed based on filtered intensity signals whose  $z$ -transform is

$$\mathbf{b}^{(\ell)}(z) = \mathbf{G}^{(\ell)}(z) \mathbf{a}^{(\ell)}(z), \quad (2)$$

where  $\mathbf{G}^{(\ell)}(z)$  is an  $|\mathcal{S}| \times |\mathcal{S}|$  polynomial matrix. The order  $O_\ell$  of the (real and nonsymmetric) MIMO filter, i.e., the largest polynomial degree in  $\mathbf{G}^{(\ell)}(z)$ , is assumed to be equal to the maximum number of walk-off delay elements in step  $\ell$ .

*Example 3:* For the parameters as before and a step size  $2\delta$ , the group delay difference of the outermost subbands corresponds to  $2(|\mathcal{S}| - 1)T_{\text{sub}} = 12T_{\text{sub}}$ . Thus, the order of  $\mathbf{G}^{(\ell)}(z)$  is  $O_\ell = 12$ .  $\triangle$

### Joint Filter Optimization via Deep Learning

“Unrolling” all SSFM steps in Fig. 2 leads to a multi-layer computation graph similar to a deep neural network<sup>5,6</sup>. Thus, joint optimization of all filters can be achieved using tools from machine learning, in particular deep learning via stochastic gradient descent. The tunable parameters are

- the prototype filters  $A(z)$  and  $S(z)$ ,
- the filters  $H^{(\ell)}(z)$  for  $\ell \in \{1, \dots, M\}$ ,
- the MIMO filters  $\mathbf{G}^{(\ell)}(z)$  for  $\ell \in \{1, \dots, M\}$ ,
- the fractional delay filters  $F_i(z)$  for  $i \in \mathcal{S}$ .

The optimization is performed in TensorFlow using the Adam optimizer. The mean squared error between the transmitted and received data symbols is used as a loss function, assuming a matched filter and phase offset rotation after the subband processing. Initially, the prototype filters are raised-cosine filters, the filters  $H^{(\ell)}(z)$  are pre-optimized using least-squares methods<sup>6</sup>, and

the filters  $F_i(z)$  are 8-tap Lagrange interpolation filters. The MIMO filters are randomly initialized.

A potential issue in terms of complexity is the large number of MIMO filter coefficients, e.g., the filter in Ex. 3 is a  $7 \times 7 \times 13$  tensor with 637 real coefficients. We assume that these tensors can be decomposed into a cascade of  $F$  sparse tensors  $\mathbf{G}^{(F,\ell)}(z) \dots \mathbf{G}^{(1,\ell)}(z) = \mathbf{G}^{(\ell)}(z)$ , where all  $\mathbf{G}^{(j,\ell)}(z)$  have dimension  $|S| \times |S|$  and order  $O_\ell/F$ . To encourage sparsity during training, we employ  $L_1$ -norm regularization for all MIMO filters.

## Results and Discussion

We consider a 96-Gbaud signal (root-raised cosine, 0.1 roll-off, Gaussian symbols),  $25 \times 100$  km of fiber ( $\alpha = 0.2$  dB/km,  $\beta_2 = -21.7$  ps<sup>2</sup>/km,  $\gamma = 1.3$  1/W/km), and amplifiers with 4.5 dB noise figure. Forward propagation is simulated using the standard SSFM with 6 samples/symbols and 1000 logarithmic steps per span (StPS). Subband TD-DBP is performed with  $N = 12$ ,  $S = 3$ ,  $K = 8$ , and a uniform step size of  $2\delta = 38.2$  km for the first 65 steps (see Ex. 2). The last step size is 17 km for a total of  $M = 66$  steps (2.6 StPS on average).

The results after training are shown in Fig. 3. Our method achieves a 2.8 dB SNR improvement over linear equalization. The loss with respect to full DBP (2 samples/symbol, 1000 StPS) is mostly due to the incoherent subband processing. To illustrate this, we also show results assuming essentially unrestricted complexity (dashed line), where  $S = 5$ ,  $K = 1$ , and frequency-domain filtering according to<sup>15</sup> with 1000 StPS is used.

To quantify the complexity, we use RMs focusing on the pulse-broadening and MIMO filters which dominate the requirements. The results were obtained with 7-tap pulse-broadening filters ( $L = 3$ ) which can be implemented using  $4(L + 1) = 16$  RMs. For the MIMO filters,  $F = 3$  is used. The learned coefficients were thresholded, after which only 3812 out of  $F|S|^2(O_\ell/F + 1)M = 48510$  total coefficients were nonzero. This gives  $3812/(|S|M) \approx 8$  RMs per subband and step on average, i.e., 24 RMs in total. A similar analysis for frequency-domain overlap-and-add filtering is presented in<sup>12</sup>. Following the same arguments, the number of RMs for our scenario is  $4(2n \log_2 n + 8n)/(n - 13) \approx 98$  per subband and step with an optimized FFT size of  $n = 2^7$ . This is significantly more than required using TD-DBP.

## Conclusions

We have proposed a novel DSP architecture for DBP based on subband processing. Our method uses short FIR filters for the CD compensation to

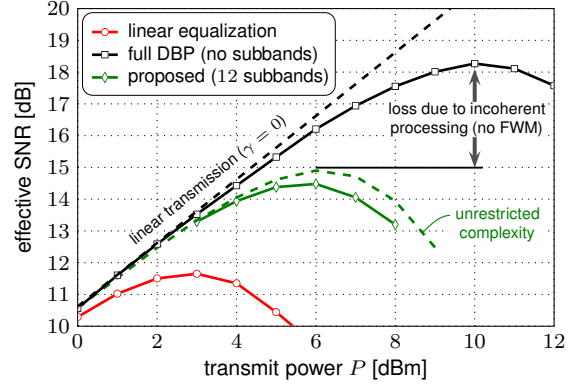


Fig. 3: Simulation results

achieve computational efficiency. It was shown that a proper step size choice can significantly simplify the walk-off compensation by using delay elements. Lastly, the complexity of the XPM MIMO filters proposed in<sup>15</sup> can be reduced by applying sparse tensor decomposition.

## Acknowledgements

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 749798. The work was also supported in part by the National Science Foundation (NSF) under Grant No. 1609327. Any opinions, findings, recommendations, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

## References

- [1] L. Zhu et al., "Complementary FIR filter pair for distributed impairment compensation of WDM fiber transmission," *IEEE Photon. Technol. Lett.* **21**, 292–294 (2009).
- [2] G. Goldfarb and G. Li, "Efficient backward-propagation using wavelet-based filtering for fiber backward-propagation," *Opt. Express* **17**, 814–816 (2009).
- [3] C. Fougstedt et al., "Time-domain digital back propagation: Algorithm and finite-precision implementation aspects," in *Proc. OFC*, (Los Angeles, CA, 2017).
- [4] C. Fougstedt et al., "Finite-precision optimization of time-domain digital back propagation by inter-symbol interference minimization," in *Proc. ECOC*, (Gothenburg, Sweden, 2017).
- [5] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Proc. OFC*, (San Diego, CA, 2018).
- [6] —, "Deep learning of the nonlinear Schrödinger equation in fiber-optic communications," in *Proc. ISIT*, (Rome, Italy, 2018).
- [7] C. S. Martins et al., "Efficient time-domain DBP using random step-size and multi-band quantization," in *Proc. OFC*, (San Diego, CA, 2018).
- [8] M. G. Taylor, "Compact digital dispersion compensation algorithms," in *Proc. OFC*, (San Diego, CA, 2008).
- [9] K.-P. Ho, "Subband equaliser for chromatic dispersion of optical fibre," *Electronics Lett.* **45**, 1224–1226 (2009).
- [10] I. Slim et al., "Delayed single-tap frequency-domain chromatic-dispersion compensation," *IEEE Photon. Technol. Lett.* **25**, 167–170 (2013).
- [11] M. Nazarathy and A. Tolmachev, "Subbanded DSP architectures based on underdecimated filter banks for coherent OFDM receivers: Overview and recent advances," *IEEE Signal Processing Mag.* **31**, 70–81 (2014).
- [12] E. F. Mateo et al., "Efficient compensation of inter-channel nonlinear effects via digital backward propagation in WDM optical transmission," *Opt. Express* **18**, 15,144 (2010).
- [13] E. Ip et al., "Complexity versus performance tradeoff for fiber nonlinearity compensation using frequency-shaped, multi-subband backpropagation," in *Proc. OFC*, (Los Angeles, CA, 2011).
- [14] T. Oyama et al., "Complexity reduction of perturbation-based nonlinear compensator by sub-band processing," in *Proc. OFC*, (Los Angeles, CA, 2015).
- [15] J. Leibrich and W. Rosenkranz, "Efficient numerical simulation of multichannel WDM transmission systems limited by XPM," *IEEE Photon. Technol. Lett.* **15**, 395–397 (2003).