# Symbol-Based Supervised Learning Predistortion for Compensating Transmitter Nonlinearity

Zonglong He(1, *), Jinxiang Song(2,*), Christian Häger(2), Kovendhan Vijayan(1), Peter Andrekson(1), Magnus Karlsson(1), Alexandre Graell i Amat(2), Henk Wymeersch(2), and Jochen Schröder(1)

(1) Photonics Laboratory, Dept. of Microtechnology and Nanoscience, Chalmers Univ. of Technology, Gothenburg, Sweden
(2) Dept. of Electrical Engineering, Chalmers Univ. of Technology, Gothenburg, Sweden
(*) Authors with equal contributions, zonglong@chalmers.se

***Abstract*** *We experimentally demonstrate a symbol-based nonlinear digital predistortion (DPD) technique utilizing supervised learning, which is robust against a change of modulation format. Back-to-back transmission of 30 Gbaud 32, 64 and 256QAM confirms that our scheme significantly outperforms the baseline of arcsine-based predistortion.*

## Introduction

Both linear[1]–[3] and nonlinear[4]–[9] digital predistortion (DPD) have been widely applied to compensate for transmitter impairments and are particularly important when transmitting high-order modulation formats. The most common nonlinear DPD is the cancellation of the modulator response using its inverse transfer function often in conjunction with clipping to reduce the peak-to-average power ratio. More complex nonlinear DPD can further alleviate pattern-dependent distortions caused by the combined effect of bandwidth limitation and RF amplifier nonlinearity[10].

Several sample-based DPD techniques which modify the oversampled signal have been proposed to pre-compensate for the pattern-dependent impairments[4]–[6]. One of the most widely studied DPD is based on the Volterra series[4] or its low-complexity variants such as memory polynomials[6]. Another type of sample-based DPD is based on neural networks (NN)[11]–[15]. In general, such sample-based DPD is challenging to implement in practical systems due to e.g. required hardware modeling (e.g., modulator and RF amplifier)[4] or sample-level digital-signal-processing (DSP)[6],[16] which makes it challenging to combine with the regular DSP toolchain. As an alternative, symbol-based techniques which compensate pattern errors using a look-up-table (LUT) are attractive due to their low-complexity[7]–[9]. The main idea behind LUT-based DPD is to add a predistortion term, based on the surrounding symbol pattern, to the transmitted symbols to mitigate pattern dependent distortions. However, LUT-based DPD depends on modulation format and must be trained separately for each format. Moreover, the size of the LUT increases exponentially with modulation order. To lessen memory requirements, reduced-size LUT schemes have been proposed, either by only considering high-error patterns[8] or taking advantage of the periodicity found in the LUT[9].

In this paper, we propose a novel symbol-based supervised learning DPD for compensating transmitter nonlinearity. We experimentally demonstrate the proposed technique in back-to-back transmission of 30 Gbaud 64QAM and show that, for memory 3, it can achieve nearly identical performance to the LUT-based DPD. Further experiments show that the NN trained for a 64QAM signal can be applied to other modulation formats (32QAM/256QAM) without retraining and no performance penalty compared to a NN trained for the tested format.

## Neural network based DPD

The purpose of the NN-based DPD is to emulate the LUT-based approach, while avoiding its memory requirements.

*Pattern error characterisation:* Before training the NN-DPD it is necessary to first characterise the pattern-dependent error of the transmitter. This closely follows the LUT implementation[9] and is done separately for in-phase and quadrature components. To construct the in-phase pattern error we identify all unique length-$n$ input real valued symbol patterns $x_k^I = [x^I_{k-\frac{n-1}{2}}, ..., x^I_k, ..., x^I_{k+\frac{n-1}{2}}]$, and estimate the in-

978-1-6654-3868-1/21/$31.00 ©2021 IEEE

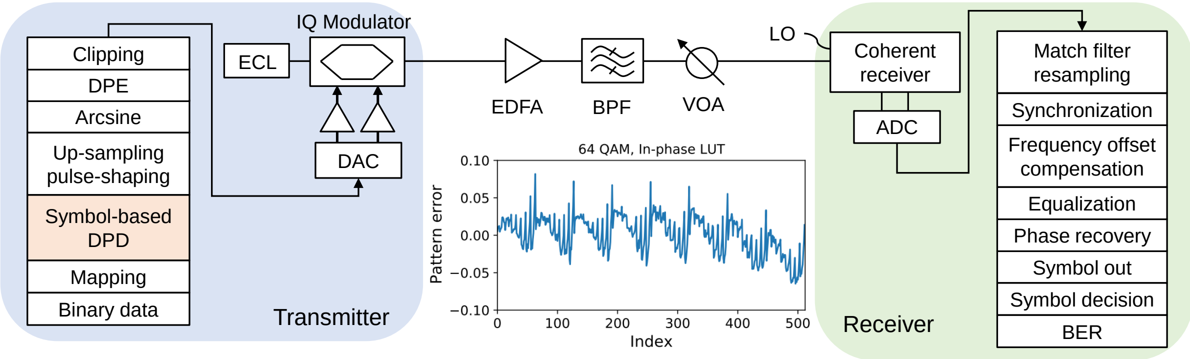


**Fig. 1:** Experimental setup for optical back-to-back transmission. DPE: digital pre-emphasis; ECL: external cavity laser; DAC: digital-to-analog converter; EDFA: erbium-doped fiber amplifier ; BPF: band pass filter; VOA: variable optical attenuator; LO: local oscillator; ADC: analog-to-digital converter.

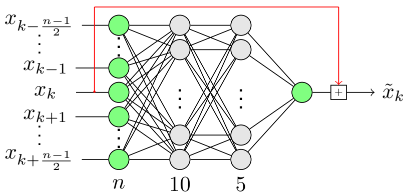

**Fig. 2:** Architecture of the proposed NN-based DPD.

phase pattern error $\delta_k^I = \delta^I(\boldsymbol{x}_k^I) = x_k^I - y_k^I$ by averaging over all occurrences of $\boldsymbol{x}_k^I$ inside the transmitted signal. Here $y_k^I$ is the post-processed in-phase component of the received symbol before symbol decision corresponding to $x_k^I$. The same process is then repeated to calculate the pattern error of the quadrature component $\delta_k^Q$ from the unique $\boldsymbol{x}_k^Q$ inputs and $y_k^Q$ output. The pattern error for a given symbol is then used to adjust the transmitted symbol $x_k^{tx} = x_k^I + \delta_k^I + i\left(x_k^Q + \delta_k^Q\right)$ in the LUT-DPD and for training in the NN-based DPD.

*NN-based predistortion:* The architecture of the proposed NN-based DPD is depicted in Fig. 2. Similar to LUT-based DPD, two NNs with length–$n$ input, denoted by $f_{\theta_I}$ and $f_{\theta_Q}$, where $\theta_I$ and $\theta_Q$ are the set of trainable parameters, need to be separately trained for the in-phase and quadrature part of the signal. In our experiment, both NNs consist of an input layer, 2 fully-connected hidden layers, an output layer, and a shortcut connection that directly adds the middle symbol of the input sequence to the output. The detailed number of neurons in each layer is shown in Fig. 2.

Training of the NN-based DPD follows the standard supervised learning approach using stochastic gradient descent. Given the length-$n$ patterns $\boldsymbol{x}_k^I$ in the real (or $\boldsymbol{x}_k^Q$ in the imaginary) part of the transmitted signal as inputs and the corresponding in-phase (or quadrature) component of the LUT-based DPD outputs $\Re(x_k^{tx})$ (or

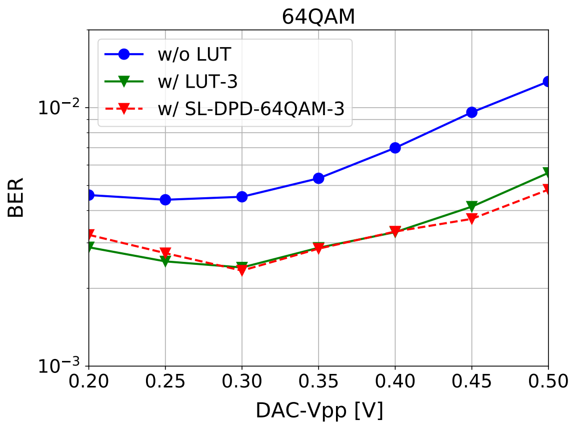


**Fig. 3:** BER against DAC-Vpp for 64QAM. SL: supervised learning.

$\Im(x_k^{tx})$) as the training labels, each of the two NNs is trained separately by minimizing the corresponding mean square error (MSE) loss defined by $\ell_I(\theta_I) = \frac{1}{N}\sum_{k=1}^{N}|\tilde{x}_k^I - \Re(x_k^{tx})|^2$ for $f_{\theta_I}$ and $\ell_Q(\theta_Q) = \frac{1}{N}\sum_{k=1}^{N}|\tilde{x}_k^Q - \Im(x_k^{tx})|^2$ for $f_{\theta_Q}$, using the Adam optimizer[17], where $N = 2^{16}$ is the number of symbols in each training iteration.

The length of a memory-3 LUT is 216, 512, and 4096 for 32QAM, 64QAM, and 256QAM, respectively, while in comparison the memory-3 NN only has 85 weights to be stored.

## Experimental setup

The experimental setup is presented in Fig. 1. Binary data is first mapped to QAM symbols before performing symbol-based supervised learning or LUT-based DPD. The corrected symbols are then upsampled and pulse-shaped with a root-raised cosine (RRC) filter with 10% roll-off. An arcsine function and a digital pre-emphasis (DPE) filter are applied to compensate for modulator non-linearity and the DAC frequency response, respectively, while optimized clipping reduces the peak-to-average power ratio (PAPR) before feeding the signal to a 60 GS/s DAC. The analog sig-

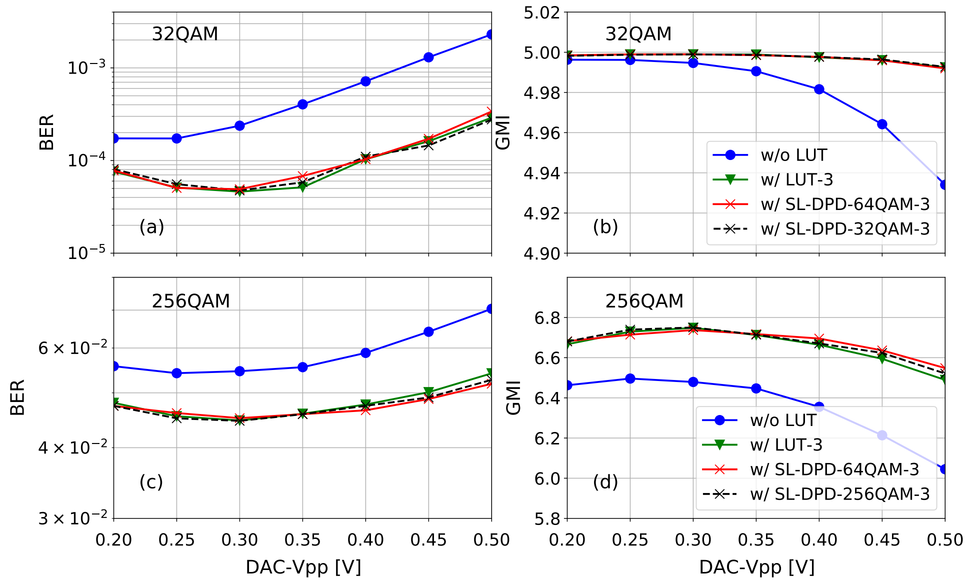


**Fig. 4:** Performance comparison of LUT and NN based DPD for 32 and 256QAM signals. SL-DPD-$X$QAM-3: supervised learning based DPD trained for $X$QAM signal and with memory 3; GMI: generalized mutual information.

nals are amplified and applied to the IQ modulator. The optical carrier is generated by an external cavity laser (ECL) operating at 1550nm. An erbium-doped fiber amplifier (EDFA) is added as a booster. A bandpass filter (BPF) and variable-optical-attenuator (VOA) filter the out-of-band noise and control the input power to the receiver, respectively. The signal is detected by a coherent receiver and sampled by an analog-to-digital converter (ADC) operating at 80 GS/s. Pilot-based DSP[18] is used to recover the constellation. All BER measurements are performed with independent random test data.

## Results

To demonstrate the efficiency of the proposed approach, we investigate the back-to-back performance of the system for a 30 Gbaud 64QAM signal by measuring the BER against DAC output peak-to-peak voltage (DAC-Vpp), which determines the strength of system nonlinearity. We first employ linear predistortion and arcsine-based modulator compensation as the baseline scheme. A LUT-3 and NN-based DPD with memory 3 are implemented for performance comparison.

As can be seen from Fig. 3, both the NN and LUT methods significantly outperform the baseline scheme. the BER is reduced by a factor of 2 at DAC-Vpp=300 mV. Importantly, the fully trained NN with memory 3 performs nearly identical to the LUT-3 for all measured DAC-Vpp, showing the possibility of replacing the LUT with a supervised learning NN.

While the LUT-based DPD needs to be retrained for each constellation, the NN approach can be easily applied to other modulation formats once trained for a given format without retraining. We test the robustness of this method for a variety of QAM signals. The memory-3 NN is first learned with a 64QAM signal (DPD-64QAM-3) and then applied to 32QAM and 256QAM. We also directly train the NNs for 32QAM and 256QAM signal (DPD-32/256QAM-3) as a reference. Fig. 4 (a)-(d) shows the measured BER and generalized mutual information (GMI). For the tested QAM signals, DPD-64QAM-3 achieves equivalent BER performance as both LUT-3 and the supervised learning based DPD trained from LUT-3 with the corresponding modulation format. This confirms that the supervised learning based DPD is robust against a change of modulation format which is reflected in the GMI measurements as well [Fig. 4 (b), (d)].

## Conclusions

We demonstrated a novel symbol-based supervised learning DPD scheme for compensating transmitter nonlinearity and evaluate its performance for various 30Gbaud QAM signals in optical back-to-back transmission. The results show that, for memory 3, a single NN can replace three LUTs estimated for different modulation formats, while achieving equivalent nonlinear gain with lower memory requirements.

## Acknowledgements

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

## References


[1] D. Rafique, H. Griesser, and J. Elbers, “Enabling 64Gbaud Coherent Optical Transceivers”, in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, 2017, pp. 1–3.

[2] A. Napoli, S. Calabrö, D. Rafique, R. Palmer, B. Spinnler, and M. Bohn, “Adaptive Digital Pre-emphasis for High Speed Digital Analogue Converters”, in *2016 Optical Fiber Communications Conference and Exhibition (OFC)*, 2016, pp. 1–3.

[3] A. Napoli, M. M. Mezghanni, D. Rafique, V. A. J. M. Sleiffer, T. Rahman, B. Spinnler, S. Calabrò, and M. Bohn, “Novel DAC Digital Pre-emphasis Algorithm for Next-generation Flexible Optical Transponders”, in *2015 Optical Fiber Communications Conference and Exhibition (OFC)*, 2015, pp. 1–3.

[4] P. W. Berenguer, M. Nölle, L. Molle, T. Raman, A. Napoli, C. Schubert, and J. K. Fischer, “Nonlinear Digital Pre-distortion of Transmitter Components”, *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1739–1745, 2016.

[5] L. Guan and A. Zhu, “Low-Cost FPGA Implementation of Volterra Series-Based Digital Predistorter for RF Power Amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 4, pp. 866–872, 2010.

[6] G. Khanna, B. Spinnler, S. Calabro, E. De Man, U. Feiste, T. Drenski, and N. Hanik, “A Memory Polynomial based Digital Pre-distorter for High Power Transmitter Components”, in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, 2017, pp. 1–3.

[7] J. Zhang, P. Gou, M. Kong, K. Fang, J. Xiao, Q. Zhang, X. Xin, and J. Yu, “PAM-8 IM/DD Transmission Based on Modified Lookup Table Nonlinear Predistortion”, *IEEE Photonics Journal*, vol. 10, no. 3, pp. 1–9, 2018.

[8] S. Zhalehpour, J. Lin, W. Shi, and L. A. Rusch, “Reduced-size Lookup Tables Enabling Higher-order QAM with All-silicon IQ Modulators”, *Opt. Express*, vol. 27, no. 17, pp. 24 243–24 259, 2019.

[9] Z. He, K. Vijayan, M. Mazur, M. Karlsson, and J. Schröder, “Look-up Table based Pre-distortion for Transmitters Employing High-Spectral-Efficiency Modulation Formats”, in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.

[10] P. J. Winzer, “High-spectral-efficiency optical modulation formats”, *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3824–3835, 2012.

[11] G. Paryanti, L. Rokach, H. Faig, and D. Sadot, “Recurrent Neural Network for Pre-distortion of Combined Nonlinear Optical Transmitter Impairments with Memory”, in *Signal Processing in Photonic Communications*, Optical Society of America, 2018, SpW4G–5.

[12] M. Schaedler, M. Kuschnerov, S. Calabrò, F. Pittalà, C. Bluemm, and S. Pachnicke, “AI-based Digital Predistortion for IQ Mach-zehnder Modulators”, in *2019 Asia Communications and Photonics Conference (ACP)*, IEEE, 2019, pp. 1–3.

[13] M. Abu-Romoh, S. Sygletos, I. D. Phillips, and W. Forysiak, “Neural-network-based Pre-distortion Method to Compensate for Low Resolution DAC Nonlinearity”, in *45th European Conference on Optical Communication (ECOC 2019)*, IET, 2019, pp. 1–4.

[14] G. Paryanti, H. Faig, L. Rokach, and D. Sadot, “A Direct Learning Approach for Neural Network based Pre-distortion for Coherent Nonlinear Optical Transmitter”, *Journal of Lightwave Technology*, vol. 38, no. 15, pp. 3883–3896, 2020.

[15] V. Bajaj, F. Buchali, M. Chagnon, S. Wahls, and V. Aref, “Single-channel 1.61 Tb/s Optical Coherent Transmission Enabled by Neural Network-Based Digital Pre-Distortion”, in *2020 European Conference on Optical Communications (ECOC)*, IEEE, 2020, pp. 1–4.

[16] G. Khanna, B. Spinnler, S. Calabro, E. De Man, and N. Hanik, “A Robust Adaptive Pre-Distortion Method for Optical Communication Transmitters”, *IEEE Photonics Technology Letters*, vol. 28, no. 7, pp. 752–755, 2016.

[17] K. Da, “A Method for Stochastic Optimization”, *arXiv preprint arXiv:1412.6980*, 2014.

[18] M. Mazur, J. Schröder, A. Lorences-Riesgo, T. Yoshida, M. Karlsson, and P. A. Andrekson, “Overhead-optimization of Pilot-based Digital Signal Processing for Flexible High Spectral Efficiency Transmission”, *Opt. Express*, vol. 27, no. 17, pp. 24 654–24 669, 2019.