Feedforward Neural Network-Based EVM Estimation: Impairment Tolerance in Coherent Optical Systems

Citation for the original published paper (version of record):
Fan, Y., Pang, X., Udalcovs, A. et al (2022). Feedforward Neural Network-Based EVM Estimation: Impairment Tolerance in Coherent Optical
Systems. IEEE Journal of Selected Topics in Quantum Electronics, 28(4).
http://dx.doi.org/10.1109/JSTQE.2022.3177004

(article starts on next page)

Feedforward Neural Network-Based EVM Estimation: Impairment Tolerance in Coherent Optical Systems

Yuchuan Fan, *Graduate Student Member, IEEE*, Xiaodan Pang, *Senior Member, IEEE*, Aleksejs Udalcovs, *Senior Member, IEEE*, Carlos Natalino, *Member, IEEE*, Lu Zhang, *Member, IEEE*, Vjaceslavs Bobrovs, *Member, IEEE*, Richard Schatz, Xianbin Yu, *Senior Member, IEEE*, Marija Furdek, *Senior Member, IEEE*, Sergei Popov, and Oskars Ozolins, *Member, IEEE*

***Abstract*—Error vector magnitude (EVM) is commonly used for evaluating the quality of m-ary quadrature amplitude modulation (mQAM) signals. Recently proposed deep learning techniques for EVM estimation extend the functionality of conventional optical performance monitoring (OPM). In this article, we evaluate the tolerance of our developed EVM estimation scheme against various impairments in coherent optical systems. In particular, we analyze the signal quality monitoring capabilities in the presence of residual in-phase/quadrature (IQ) imbalance, fiber nonlinearity, and laser phase noise. We use feedforward neural networks (FFNNs) to extract the EVM information from amplitude histograms of 100 symbols per IQ cluster signal sequence captured before carrier phase recovery. We perform simulations of the considered impairments, along with an experimental investigation of the impact of laser phase noise. To investigate the tolerance of the EVM estimation scheme to each impairment type, we compare the accuracy for three training methods: 1) training without impairment, 2) training one model for all impairments, and 3) training an independent model for each impairment. Results indicate a good generalization of the proposed EVM estimation scheme, thus providing a valuable reference for developing next-generation intelligent OPM systems.**

***Index Terms*—Optical communication, optical fiber communication, feedforward neural networks, signal processing, monitoring.**

I. Introduction

OPTICAL communication networks are evolving towards simultaneous support of various types of data traffic driven by fifth-generation (5G) mobile networks, cloud services, and data centers. The massive increase in data traffic requires advanced optical networks to meet the ever-growing capacity demand [1]. These complex network architectures make better use of available resources by integrating key enabling technologies such as reconfigurable optical add-drop multiplexers (ROADMs), flexible grids, and advanced modulation formats [2]. Optical performance monitoring (OPM) is crucial for adaptive, reliable, and efficient management and operation of optical networks. Estimating transmission performance is required for both the intermediate network nodes and the receiver [3], by monitoring various network performance parameters to ensure the best utilization of available resources. Monitored parameters can include optical power, optical signal to noise ratio (OSNR), error vector magnitude (EVM), etc. EVM is a measure of signal error statistics for m-ary quadrature amplitude modulation (mQAM) formats, which can be related to OSNR and bit error rate (BER) [4], [5] for each specific modulation format within a certain range of practical impairments. Monitoring EVM can extend the functionality of OPM modules by providing intuitive error statistics of the system to network analytics functions. Traditionally, EVM calculation requires receiving millions of symbols [4]. Such a cumulative process is time-consuming and inconvenient for OPM implementation. In addition, the prohibitive cost of OPM equipment limits its placement to only a subset of network nodes. Therefore, the development of lower-complexity OPM strategies can enable its ubiquitous

Manuscript received October 1, 2021; revised February 28, 2022 and April 28, 2022; accepted May 19, 2022. Date of publication May 23, 2022; date of current version June 3, 2022. This work was supported in part by the China Scholarship Council under Grant 201807930003, in part by the Swedish Research Council (VR) projects under Grants 2019-05197 and 2016-04510, in part by the RISE SK funded project Optical Neural Networks under Grant P109599, in part by the ERDF-funded CARAT Project under Grant 1.1.1.2/VIAA/4/20/660, in part by the National Key Research and Development Program of China under Grant 2018YFB2201700, and in part by the VINNOVA funded CELTIC-NEXT Project AI-NET PROTECT under Grant 2020-03506. *(Corresponding author: Oskars Ozolins.)*

Yuchuan Fan and Xiaodan Pang are with the Applied Physics Department, KTH Royal Institute of Technology, 106 91 Stockholm, Sweden, and also with the RISE Research Institutes of Sweden, 164 40 Kista, Sweden (e-mail: yuchuanf@kth.se; xiaodan@kth.se).

Aleksejs Udalcovs is with the RISE Research Institutes of Sweden, 164 40 Kista, Sweden (e-mail: aleksejs.udalcovs@gmail.com).

Carlos Natalino and Marija Furdek are with the Electrical Engineering Department, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: carlos.natalino@chalmers.se; furdek@chalmers.se).

Lu Zhang and Xianbin Yu are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and also with Zhejiang Lab, Hangzhou 310000, China (e-mail: zhanglu1993@zju.edu.cn; xyu@zju.edu.cn).

Vjaceslavs Bobrovs is with the Institute of Telecommunications, Riga Technical University, 1048 Riga, Latvia (e-mail: vjaceslavs.bobrovs@rtu.lv).

Richard Schatz and Sergei Popov are with the Applied Physics Department, KTH Royal Institute of Technology, 106 91 Stockholm, Sweden (e-mail: rschatz@kth.se; sergeip@kth.se).

Oskars Ozolins is with the Applied Physics Department, KTH Royal Institute of Technology, 106 91 Stockholm, Sweden, with the RISE Research Institutes of Sweden, 164 40 Kista, Sweden, and also with the Institute of Telecommunications, Riga Technical University, 1048 Riga, Latvia (e-mail: ozolins@kth.se).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSTQE.2022.3177004.

Digital Object Identifier 10.1109/JSTQE.2022.3177004

department at optical network nodes while enhancing the monitoring capabilities and reduce the associated costs [3], which motivates the need for a versatile and simple EVM estimation scheme.

Deep learning (DL) techniques have been shown to achieve state-of-art accuracy and efficiency [6]–[8] and are, therefore, widely studied as enablers of intelligent OPM schemes. Recently, DL-based EVM estimation schemes have been proposed to monitor coherent system signal quality [9]–[12]. The first attempt at intelligent EVM estimation used convolutional neural networks (CNNs) in conjunction with constellation diagram of received signal [9]. The study in [10] provided insight into the possibility of skipping carrier phase recovery (CPR) to simplify and speed up EVM estimation. In [11], a scheme based on a feedforward neural network (FFNN) was proposed to infer the EVM from an amplitude histogram (AH) of a short signal sequence captured before the CPR module. This approach improves agility and energy efficiency for implementation thanks to the simplified signal processing and light neural network structure. An initial experiment of laser linewidth tolerance of FFNN-based scheme study was conducted in [12]. However, it is still unclear what is the contribution of individual degradation sources. To address this gap, in this paper, we focus on assessing the impact of physical layer impairments on the performance of an EVM estimation scheme by using datasets obtained via simulations and experiments. We use FFNN conjunction with vectorized AHs as the EVM estimation scheme, since it particularly relaxes the training and inference time, and energy consumption, compared with CNN-based scheme [11].

As the digital signal processing (DSP) techniques mature, coherent optical systems can recover linear and nonlinear transmission impaired signals [13]. However, practical imperfections common in deployed transceivers and systems might still impact the recovered signal, which can induce impairments such as in-phase/quadrature (IQ) imbalance, fiber nonlinearity, laser phase noise, etc. IQ imbalance (IQI) refers to gain/phase mismatch of the I and Q channels, which can be introduced at the transmitter (Tx) or receiver (Rx) side [13], [14]. An imperfectly compensated system will result in the signal containing a small amount of residual IQ imbalance. Besides, the long-haul transmission systems require higher launch power (LP), which exacerbates fiber nonlinearity. Moreover, semiconductor lasers with Lorentzian linewidths (LW) in the 1-10 MHz range can be used in coherent optical transceivers for metro and access range applications [15]. These cost-efficient type lasers can induce high phase noise to the systems. All these impairments trigger transmission penalty even after DSP and may result in inaccurate signal quality monitoring. Therefore, it is necessary to analyze the performance of the signal quality monitoring scheme in such systems.

In this paper, we investigate the robustness of the previously proposed FFNN-powered EVM estimation scheme [11] for coherent transceivers in the presence of practical imperfections. The main contribution of this work are as follows: (i) different causes of degradation, i.e., residual IQ imbalance, fiber nonlinearity, and laser phase noise, are studied separately via simulation; (ii) the capability of signal AH to represent laser phase noise features is examined; and (iii) a detailed insight into residual imperfection evaluation of experimental systems is provided. The considered modulation formats are square 64QAM (Sq-64QAM) and circular 64QAM (C-64QAM) [16], with 64 constellation clusters. The included C-64QAM is an example of geometric shaping to test the generalization of the EVM estimation model. We simulate the considered impairments individually with ASE noise and collect corresponding AH datasets for further signal quality monitoring investigation. Moreover, we conduct an experimental study of the effect of laser phase noise on EVM estimation. The results show that the proposed EVM estimator performs well when the system is impaired by residual IQ imbalance or fiber nonlinearity. For laser phase noise impaired systems, opposite estimation results are observed when the FFNN models are trained with the simulation and the experimental datasets. The trained model does not perform well for the simulation dataset. This interesting finding indicates that AH-based performance monitoring works better when the signal phase distortions are coupled with other practical impairments like the IQ imbalance, revealing extractable phase information for estimation. This paper is organized as follows. Section II describes setups for different impairment scenarios, dataset collection, and the implemented EVM estimation scheme. In Section III, the EVM estimation results are presented and analyzed. Finally, Section IV concludes the paper.

II. Operating Principles

This section begins by presenting signal collection setups for IQ imbalance, fiber nonlinearity, and laser phase noise. The collected signals are further processed into AH datasets as the input of the EVM estimator. We use VPItransmissionMaker^{TM} [17] to simulate all considered impairments, while an experiment is conducted for laser phase noise in addition. At the end of the section, we describe the operation principle of the EVM estimator and the FFNN structure.

A. IQ Imbalance

IQ imbalance is a typical impairment in coherent transceivers induced by the hardware components introducing the signal's amplitude mismatch and phase deviation [13], [14], [18]–[20]. The amplitude mismatch sources include an electrical amplifier, electro-optical response between I and Q channels, photodetector, and transimpedance amplifiers. Phase deviation in the phase modulator and optical hybrid can introduce IQ imbalance for Tx and Rx, respectively. The IQ imbalance of Tx is normally pre-calibrated and compensated in commercial systems. In this work, we mainly focus on residual IQ imbalance resulting from the coherent receiver. The IQ imbalance scenarios include phase imbalance $\varphi = [-5{:}\ +5]$ deg and 1% amplitude mismatch in the optical hybrid. Fig. 1 shows the 32 Gbaud single-polarization coherent back-to-back simulation setup for residual IQ imbalance signals collection of both Sq-64QAM and C-64QAM. Each IQ imbalance case contains 10 different OSNR values ranging from 26 to 44 dB. The Tx and local oscillator (LO) laser linewidth are 100 kHz in the simulation setup for IQ imbalance and fiber nonlinearity studies.

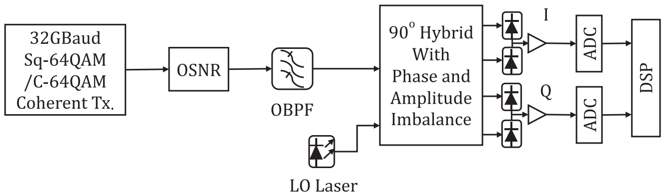

Fig. 1. IQ imbalance simulation setup. OBPF: optical bandpass filter, LO: local oscillator, ADC: Analog-to-digital converter.

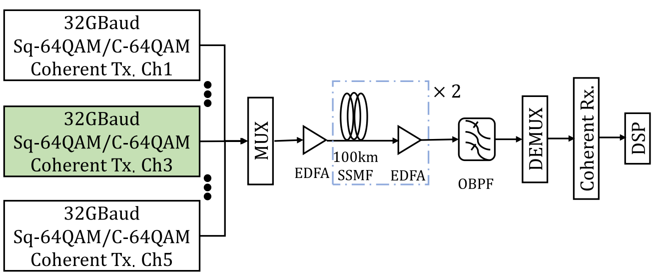

Fig. 2. Launch power simulation setup. MUX: multiplexer, EDFA: erbium-doped fiber amplifier, SSMF: standard single-mode fiber, DEMUX: demultiplexer.

B. Fiber Nonlinearity

The optical fiber nonlinearities are non-negligible for wavelength division multiplexing systems with high launch power and long transmission distances, where they degrade the signal quality [21]. In this regard, we investigate a 32 Gbaud 5- channel optical fiber transmission system using the 50 GHz ITU grid, illustrated in Fig. 2. The central channel is under the test. The modulated optical signal is amplified by an erbium-doped fiber amplifier (EDFA) and transmitted over a 200 km standard single-mode fiber (SSMF) link. The chromatic dispersion coefficient, attenuation coefficient, and nonlinear refractive index of SSMF are set to 16e-6 s/m^2, 0.2 dB/km, 2.6e-20 m^2/W, respectively. The optimized launch power per channel for the system is 6 dBm. In total, we consider 30 simulation scenarios with 1 dBm increments of launch power values ranging between -4 dBm and $+10$ dBm per channel for both Sq-64QAM and C-64QAM.

C. Laser Phase Noise

We collect two types of datasets, namely, simulation and experiment, for the laser phase noise study. The simulation setup is similar to the one in Fig. 1, except for omitting the IQ imbalance. 28 Gbaud symbol rate is employed in the simulation to match the experimental configuration. We sweep the transmitter laser linewidth from 100 kHz to 4.1 MHz and keep the LO laser linewidth at 200 kHz. The total linewidth in the system ranges from 300 kHz to 4.3 MHz. For each simulated linewidth case, we change the OSNR from 24 to 44 dB with a 4 dB step size.

The corresponding experimental setup is shown in Fig. 3. The transmitter consists of 50 GSa/s arbitrary waveform generators (AWG), IQ modulator (IQM), and a less than 100 kHz linewidth external cavity laser (ECL). We map repeated pseudo-random bit sequence of 2^{15}-1 word length (PRBS-15) to Sq-/C-64QAM

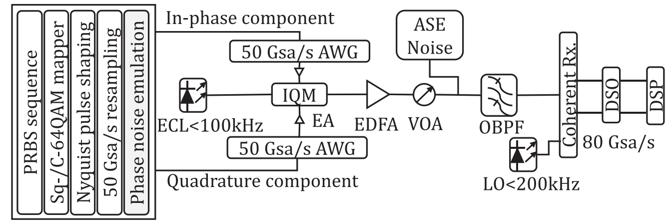

Fig. 3. Experimental setup for laser phase noise dataset collection. PRBS, pseudorandom bit sequence; ECL, external cavity laser; EA, electrical amplifier; IQM, in-phase and quadrature modulator; VOA, variable optical attenuator; ASE, amplified spontaneous emission; OBPF, optical bandpass filter; DSO, digital sampling oscilloscope.

symbols in a complex plane. Then, the symbol sequence is filtered by Nyquist pulse shaper with a 0.15 roll-off factor and resampled to match the AWG rate. We emulate the transmitted signal $x_{laser}(t)$ with semiconductor laser linewidths from 300 kHz to 4.3 MHz

$$x_{laser}\ (t) = A \cdot e^{j(\omega_{Tx}(t)+\varphi_{pn}(t))} \tag{1}$$

where A, ω_{Tx} and φ_{pn} denote the amplitude, central emitting frequency and phase noise of the laser, respectively [22], [23]. The digitally generated phase noise sequences use Wiener's phase noise model [22]

$$\varphi_{pn}\ (k) = \varphi_{pn}\ (k-1) + \Delta\theta. \tag{2}$$

Where $\Delta\theta$ is an independent and identically distributed random Gaussian variable with zero mean and variance [22]

$$\sigma_\theta^2 = \ 2\pi\left(\Delta f \cdot T_s\right). \tag{3}$$

The Δf is corresponding laser linewidth of the signal, and T_s is symbol period. The modulated optical signal is then amplified by an EDFA. The signal OSNR is configured with a variable optical attenuator (VOA) and an amplified spontaneous emission (ASE) noise source. We collect four linewidth scenarios, and each of them contains 6 OSNR values between 25 dB and 44 dB and two modulation formats.

D. Neural Network Structure and Training

In the simulation, we collect 2^{20} number of symbols for each transmission scenario. After collecting the signals with different impairment types, we generate datasets with 100 AHs for each EVM true label. The EVM true labels are calculated from the received symbols after CPR. Each AH in the dataset is expressed as 64 bin counts, which is generated by 6400 symbols before CPR. Fig. 4 captures the constellation diagrams of collected signals after/before CPR, and before CPR AH examples. Besides, this figure also provides information about the examples of EVM versus OSNR curves and BER versus OSNR curves for the collected signals. The EVM values are calculated against received symbols centroids, which are obtained from K-means clustering method. The BER counting adopts a hard-decision method based on the Voronoi boundaries between received symbol centroids, and differential decoding between quadrants. Besides, we use Gray coding in each quadrant for symbol-to-bit demapping in

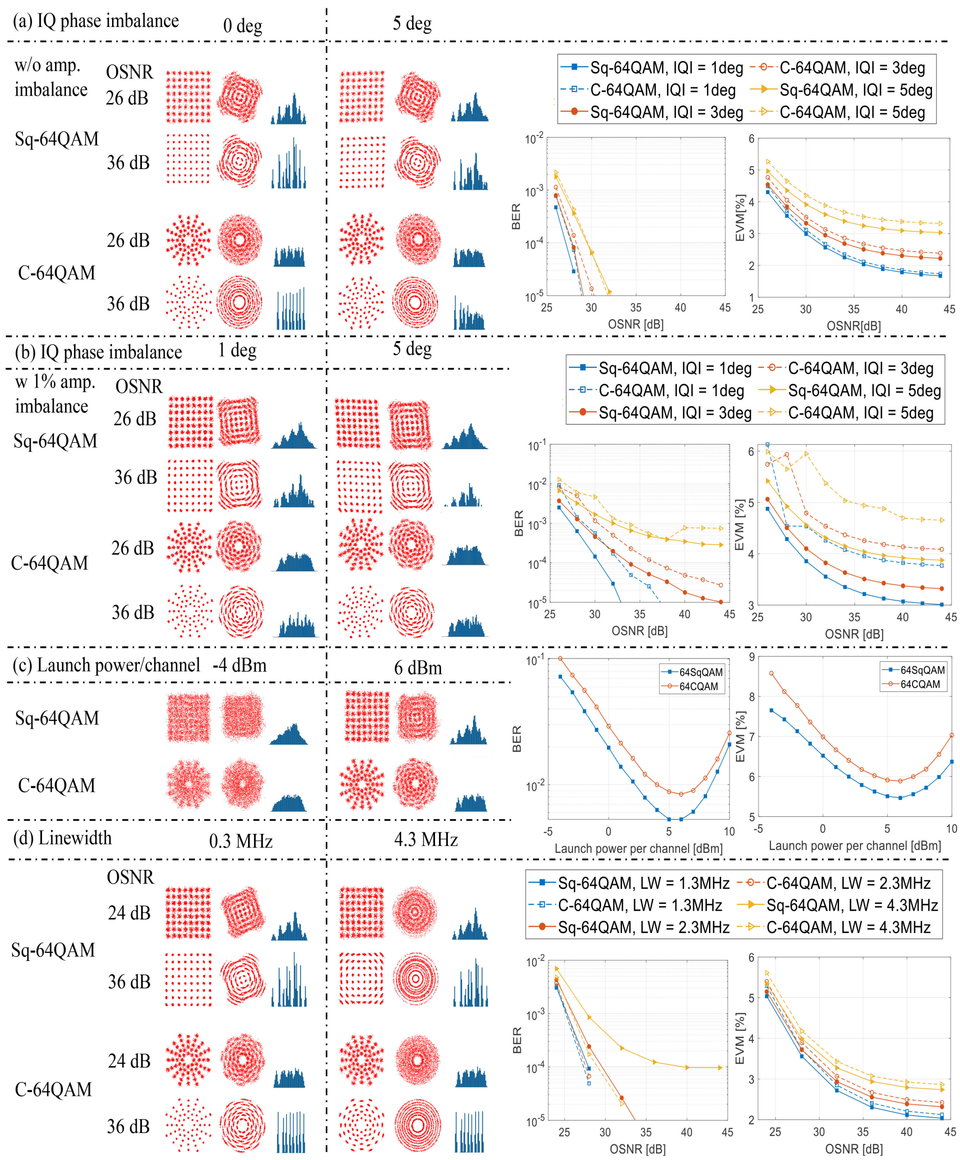

Fig. 4. The examples of before/after CPR constellations, before CPR AHs, BER versus OSNR curves, and EVM versus OSNR curves for different degradation source: (a) IQ imbalance without amplitude imbalance, (b) IQ imbalance with 1% residual amplitude imbalance, (c) fiber nonlinearity, (d) laser phase noise.

the case of square QAM. In the case of circular QAM, we use the encoding configuration from [16] as Gray encoding is not feasible to be implemented. With the Voronoi boundary-based hard-decision method, the impairments (including IQ imbalance and phase noise inducing displacements of received symbol centroids) are taken into consideration to a certain extent. One should notice that across the modulation formats, one EVM value may link to a different BER level at different impairments. For example, in Fig. 4(b), the EVM-chart of 'Sq-64QAM, IQI = 3deg' performances better than that of 'C-64QAM, IQI = 1deg', whereas the BER-charts are the opposite. To obtain an accurate EVM to BER mapping, the information about the modulation format needs to be extracted, which can be done with the same model [24], and then the information about the received signal modulation format can be combined with the predicted EVM values. In such a way, a lookup table can be made in order to obtain a BER for monitoring purposes. However, it requires diagnosis studies since these impairments are correlated to each other. Although the correlations between EVM and BER across different degradation source is not directly linked, the EVM can still indicate signal quality under the same system conditions. In general, a higher EVM is qualitatively correlated to a higher BER. Statistical studies for accurately mapping EVM to BER in practical impairments are worth for future work.

Fig. 5 depicts a schematic diagram of the FFNN-based EVM estimation scheme. The deployed FFNN consists of an input layer, four hidden layers, and an output layer [11]. The number of neurons for each layer is shown in Fig. 5. An FFNN is a simple inference structure mapping the input space to an output space by nonlinear output of the weighted sum operations in the neurons [25], [26]. The operation of the neurons in layer m can

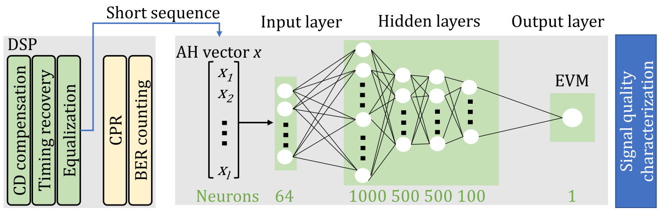

Fig. 5. Block diagram of FFNN-based EVM estimation scheme.

be denoted as

$$y^m = f_m \left(W^m \cdot x^{m-1} + b^m\right) \tag{4}$$

where W^m and b^m are weight and bias matrices from layer *m-1* to layer *m*. The $f_m(\cdot)$ is the activation function of layer *m*, and we use the rectified linear unit (ReLU) in our case [27]. The weight and bias matrix are updated by minimizing the loss between the estimated and the true EVM values in each epoch. To grasp small estimation errors, we utilize the mean squared logarithmic error (MSLE) as the loss function.

An effective and adequate dataset for training contributes to a better generalization of the EVM estimator. Therefore, we implement three training schemes to investigate the tolerance to each impairment type: (i) the case with insignificant impairment (e.g., IQI = 0 deg, LP = 6 dBm, LW = 300 kHz) to use as a benchmark, (ii) all cases: training one model to incorporate all impairments, and (iii) separate cases: training an independent model for each impairment case. The number of training epochs is 200. We use 50%, 25%, and 25% of the dataset for training, validation, and test purposes, respectively. The neural network model is constructed using the Keras framework [28] and TensorFlow library [29]. All experiments reported in this paper are performed on a 2.4 GHz Intel Xeon E5-2630-v3 with 64 GB of RAM and a GTX TITAN Black graphic card.

III. Results and Discussion

This section analyzes the estimation performance obtained for the testing datasets as mean estimation error and EVM estimation error for different impairment scenarios. The mean estimation error is calculated by averaging the mean absolute error of each OSNR case in the impairment scenarios.

A. Effect of IQ Imbalance

Fig. 6 shows the mean EVM estimation error for different training methods, and both Sq-64QAM and C-64QAM modulation formats are equally mixed in the datasets. When the model is trained only on the dataset without IQ imbalance, the test error increases along with the increase of IQ imbalance degree as shown in Fig. 6(a). Besides, the residual amplitude imbalance can degrade the model accuracy (see Fig. 6(b)). Fig. 6(c) shows an example of Sq-64QAM estimated mean absolute error versus OSNR levels. It can be observed that the mean absolute error is slightly higher in the cases with lower OSNR, and the amplitude imbalance can magnify this difference. The averaged mean absolute error can show the trend of model performance for each impairment. We collect further insights by analyzing the EVM estimation error, displayed as violin plots in Fig. 7. The blue vertical line represents the range of estimation errors in the violin plot, and the bottom, median, and top dashes correspond to the minimum, median, and maximum error, respectively. The violin shape around the median dash denotes the estimation error distribution of test samples. The shorter and broader violin shape means that the estimation errors are more concentrated around the median error. Without amplitude imbalance, the trained model achieves below 0.25% EVM deviation for up to 2 degrees IQ imbalance, as depicted in Fig. 7(a). Moreover, the performance of the model trained on all IQ imbalance scenarios is as good as the model trained separately for each impairment case. It can be observed that the mean estimation error and the estimation error deviation are 0.05% and 0.25%, respectively. When the systems include 1% residual amplitude imbalance, although the estimation error deviation is below 0.25% for most test samples, there are outliers for each phase imbalance case.

B. Effect of Fiber Nonlinearity

Figs. 8 and 9 show the EVM estimation capability when fiber nonlinearity is present in the system. For the model trained only on the scenario with optimum launch power per channel, the estimation performance degrades when the launch power increases or decreases with respect to the optimum, denoted with the red lines in Fig. 8. Training the model for all launch power cases jointly or separately results in a mean estimation error of less than 0.5% for all test launch power regions, denoted with blue and yellow curves, respectively. It is observed that some violin shapes of launch power scenarios are broken into two parts in Fig. 9(a) and (c). We attribute such results to the use of two EVM true labels (one for each modulation format) in a single model, enhanced by the higher additive noise at the lower launch powers. This indicates that it is hard to extract useful features from AHs with a limited-size dataset, resulting in poor generalization for some launch power scenarios.

C. Effect of Laser Phase Noise

Figs. 10 and 11 illustrate the EVM estimation capability of different models trained with simulation and experimental datasets. From the results of the benchmarking model denoted as LW = 300 kHz, we can observe that the performance for both the simulation and the experimental data deteriorates when we gradually increase the linewidth value. When a single model is trained on all LW scenarios, we achieve a mean estimation error below 0.5% and 0.2% for the simulation and the experimental datasets, respectively. A bending shape for training a single model on all linewidth scenarios is observed when the simulation dataset is used, as shown in Fig. 10(a). The reason is that the phase noise information in ideal IQ constellations is removed by the AHs, which makes FFNN lose the ability

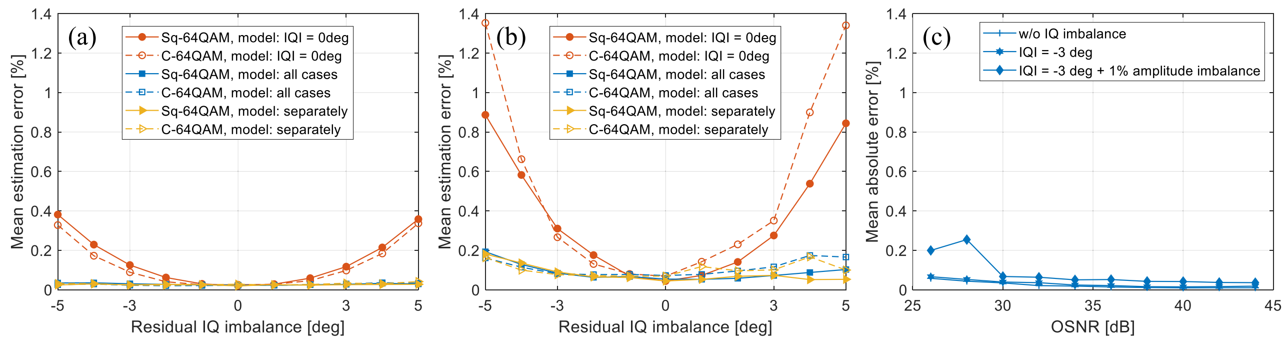

Fig. 6. The mean estimation error versus IQ imbalance scenarios for different datasets: (a) without amplitude imbalance, (b) with 1% amplitude imbalance. (c) The mean absolute error versus OSNR curves for Sq-64QAM when training for all impairment cases together.

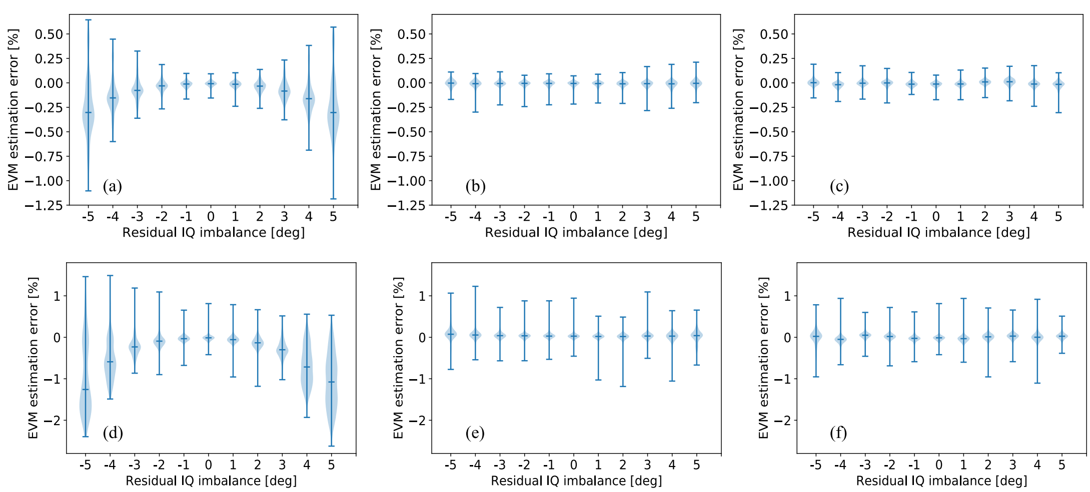

Fig. 7. Distribution of EVM estimation errors under different IQ imbalance scenarios: (a)–(c) without amplitude imbalance, (d)–(f) with 1% amplitude imbalance. The model is trained on (a), (d) IQI = 0deg; (b), (e) all cases; (c), (f) separate cases.

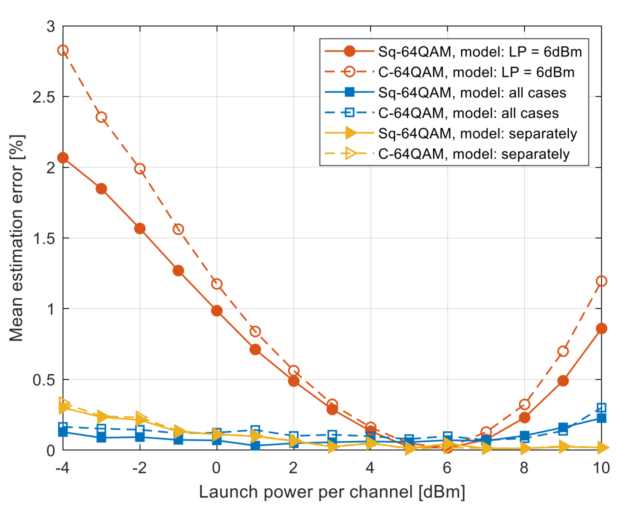

Fig. 8. The mean estimation versus launch power per channel for different training schemes.

to distinguish among different linewidth scenarios. In contrast, the results associated with the corresponding experimental dataset show a good generalization. Different from the simulation environment where impairments are ideally separated, the signals contained in the experimental dataset are impaired simultaneously by several implementation imperfections, such as laser phase noise, transmitter IQ imbalance, etc. We attribute such experimental results to this hybrid impairment, which uncovers the features that contain the phase information. This helps the FFNN recognize different linewidth scenarios, which have a similar error floor with individual training (see Fig. 11(e) and (f)). Finally, as shown in Fig. 11(c), training the model separately for different linewidth cases gives better estimation performance on the simulation dataset, which is considered reasonable but less practical. The added laser phase noise has less impact on FFNN inferring EVM from ASE noise.

In this experiment, we use Lithium Niobate-based IQ modulator, where the bias drift is significant. This results in a variation of transmitter IQ imbalance, which can be observed in constellation diagrams as shown in Fig. 12(a). Most of practical systems use the Indium Phosphide-based IQ modulator and are pre-compensated by a feedback loop, thus the huge variations in IQ imbalance are not expected. Finally, after studying the impairments separately via simulations, we include IQ imbalance in the 0.3 MHz (1 deg and 0.0231% amplitude imbalance) and the 4.3 MHz (1 deg and 1.16% amplitude imbalance) simulation

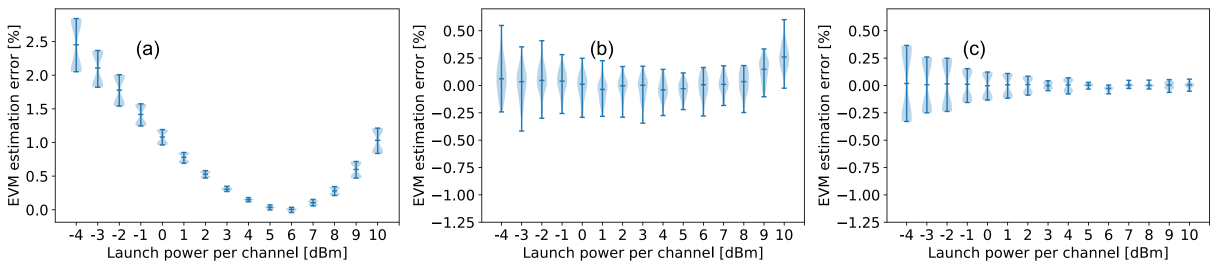

Fig. 9. Distribution of EVM estimation errors under different launch power. The model is trained on (a) LP = 6 dBm, (b) all cases, (c) separate cases.

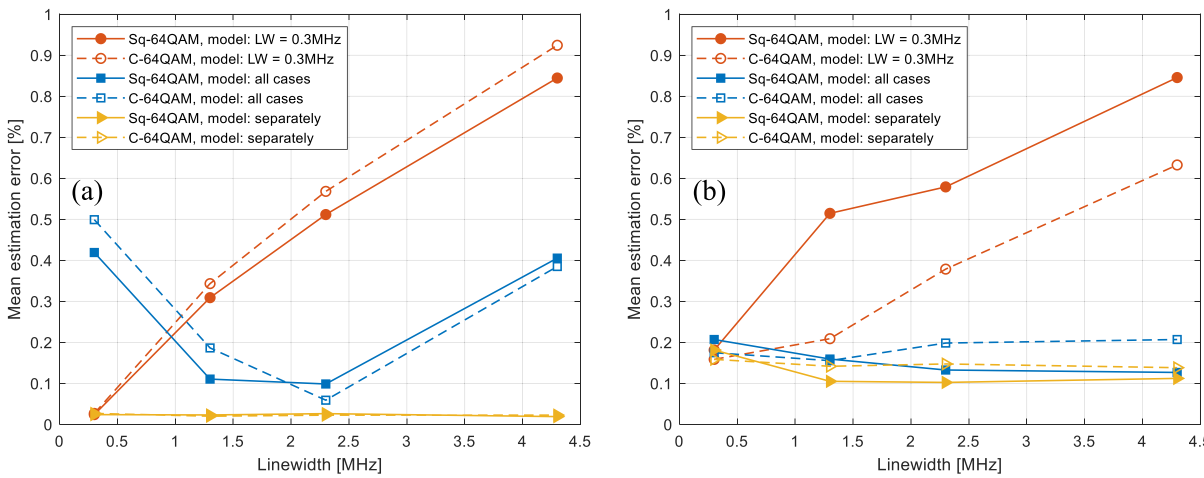

Fig. 10. The mean estimation error versus linewidth for different training schemes: (a) simulation and (b) experimental results.

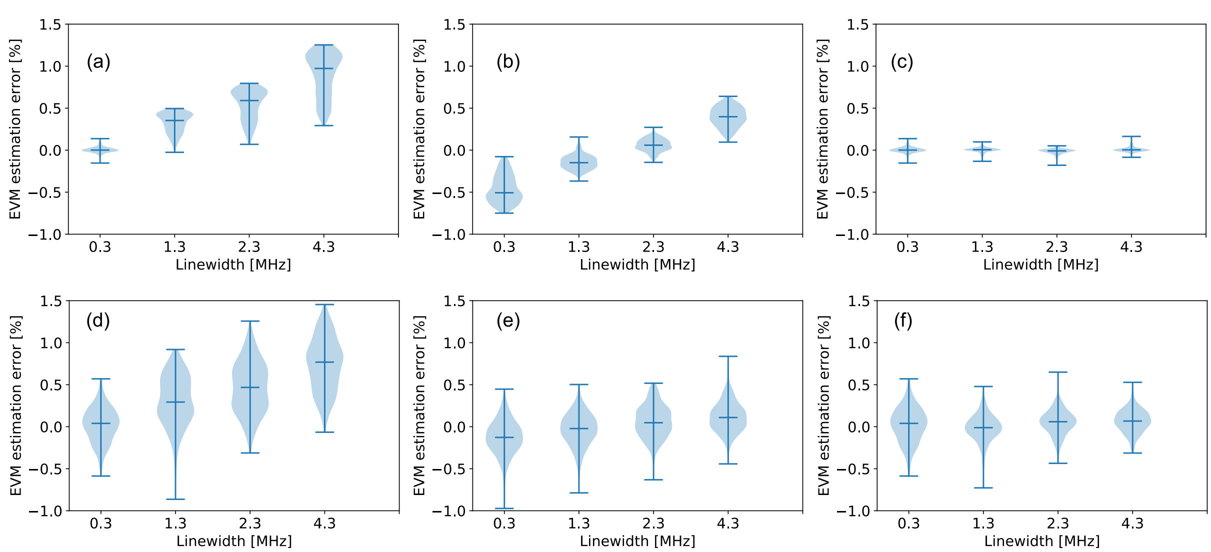

Fig. 11. Laser phase noise simulation (a)–(c) and experimental (d)–(f) results. The model is trained on: (a), (d) LW = 300 kHz; (b), (e) all cases; (c), (f) separate cases.

datasets. Fig. 12(b) illustrates the comparison of estimation results on different datasets when we train one model with mixed modulation formats and linewidth cases. It can be observed that if more impairments are included in simulations, the estimation tendency of simulation scenarios can match the experimental setting much better. Our next phase study will concentrate on the interactions between impairments and how those affect the performance of the model.

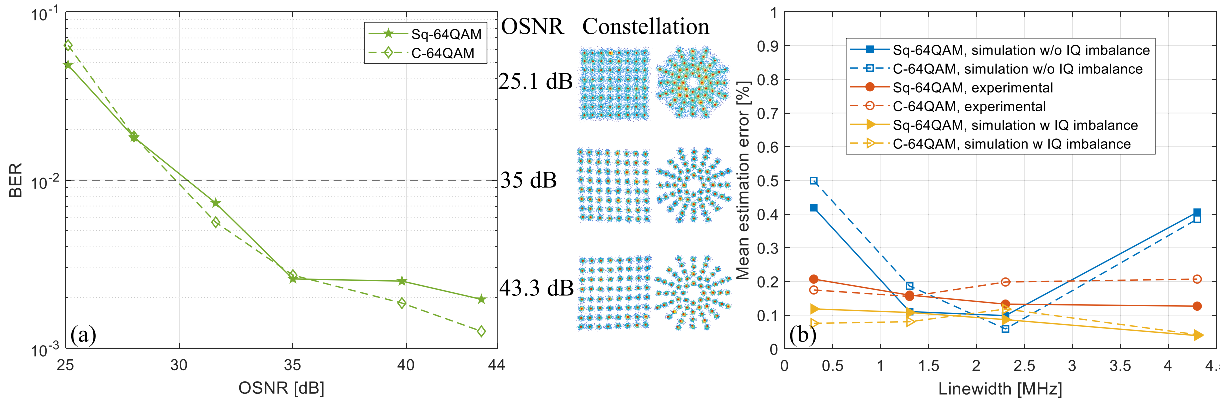

Fig. 12. (a) The BER versus OSNR in the 0.3 MHz experimental dataset and constellation diagram examples in the experiment. (b) Comparison of simulation and experimental dataset results when training one model for all linewidth scenarios.

IV. Conclusion

We have studied the performance of the FFNN-enabled EVM estimation scheme for the coherent system in the presence of practical impairments common in deployed transceivers and systems. The impact of residual IQ imbalance, fiber nonlinearity, and laser phase noise on the EVM estimation performance is studied in detail. Amplitude histograms represented by 100 symbols per cluster signal sequence captured before the CPR module are used to train the FFNN regression model. The simulation results show that the proposed EVM estimator is robust against IQ imbalance and fiber nonlinearity. The mean estimation error is below 0.05% and 0.2% for IQ imbalance without and with residual amplitude imbalance, respectively. The performance of fiber nonlinear is closer to the residual amplitude imbalance, which is below 0.3%. However, when the transmission system is mainly restricted by laser phase noise, the simulation results demonstrate the limitations of extracting the phase information from AHs. In this case, the model needs to be trained separately for different linewidth scenarios to achieve desirable performance. The proposed EVM estimator exhibits good generalization potential when applied to experimental datasets of laser phase noise. In this case, besides the laser phase noise accounted for in the simulations, the experimental setup also has other implementation penalties which reveal phase information on AHs. Thus, the proposed EVM estimation scheme can be used for a practical system in this situation. This study is one step forward before looking into all the interactions among them. In future work, we will conduct a full analysis of IQ amplitude imbalance and phase noise interacting with other impairments. Besides, a more effective signal representation for laser phase noise is an interesting topic for future studies.

References

[1] Z. Dong *et al.*, "Optical performance monitoring: A review of current and future technologies," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 2016.

[2] A. E. Willner, Z. Pan, and C. Yu, "Optical performance monitoring," in *Optical Fiber Telecommunications V B*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. New York, NY, USA: Academic, 2008, ch. 7.

[3] F. N. Khan, Z. Dong, C. Lu, and A. P. T. Lau, "Optical performance monitoring for fiber-optic communication networks," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*. Hoboken, NJ, USA: Wiley, 2016, ch. 14.

[4] R. Schmogrow *et al.*, "Error vector magnitude as a performance measure for advanced modulation," *IEEE Photon. Technol. Lett.*, vol. 24, no. 1, pp. 61–63, Jan. 2012.

[5] I. Fatadin, "Estimation of BER from error vector magnitude for optical coherent systems," *Photonics*, vol. 3, 2016, Art. no. 21.

[6] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 493–516, Jan. 2019.

[7] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 13, pp. B60–B71, 2021.

[8] T. Tanaka, T. Inui, S. Kawai, S. Kuwabara, and H. Nishizawa, "Monitoring and diagnostic technologies using deep neural networks for predictive optical network maintenance [Invited]," *J. Opt. Commun. Netw.*, vol. 13, pp. E13–E22, 2021.

[9] Y. Fan *et al.*, "Fast signal quality monitoring for coherent communications enabled by CNN-based EVM estimation," *J. Opt. Commun. Netw.*, vol. 13, pp. B12–B20, 2021.

[10] Y. Fan *et al.*, "Deep learning assisted pre-carrier phase recovery EVM estimation for coherent transmission systems," in *Proc. Conf. Lasers Electro-Opt.*, 2021, Paper STh1F.2.

[11] Y. Fan *et al.*, "Experimental validation of CNNs versus FFNNs for time- and energy-efficient EVM estimation in coherent optical systems," *J. Opt. Commun. Netw.*, vol. 13, pp. E63–E71, 2021.

[12] Y. Fan *et al.*, "Laser linewidth tolerant EVM estimation approach for intelligent signal quality monitoring relying on feedforward neural networks," in *Proc. Eur. Conf. Opt. Commun.*, 2021, Paper Th1E.2.

[13] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1125–1141, Mar. 2017.

[14] Q. Zhang *et al.*, "Modulation-format-transparent IQ imbalance estimation of dual-polarization optical transmitter based on maximum likelihood independent component analysis," *Opt. Exp.*, vol. 27, pp. 18055–18068, 2019.

[15] A. Kakkar *et al.*, "A path to use large linewidth LO in 28 Gbd 16-QAM metro links," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.

[16] J. R. Navarro *et al.*, "Carrier phase recovery algorithms for coherent optical circular mQAM systems," *J. Lightw. Technol.*, vol. 34, no. 11, pp. 2717–2723, Jun. 2016.

[17] VPIphotonics GmbH, "VPItransmissionMaker11," 2022. [Online]. Available: https://www.vpiphotonics.com/

[18] T. H. Nguyen *et al.*, "Blind adaptive transmitter IQ imbalance compensation in m-QAM optical coherent systems," *J. Opt. Commun. Netw.*, vol. 9, pp. D42–D50, 2017.

[19] Q. Wang, Y. Yue, and J. Anderson, "Detection and compensation of power imbalance, modulation strength, and bias drift in coherent IQ transmitter through digital filter," *Opt. Exp.*, vol. 26, pp. 23069–23083, 2018.

[20] Q. Zhang *et al.*, "Algorithms for blind separation and estimation of transmitter and receiver IQ imbalances," *J. Lightw. Technol.*, vol. 37, no. 10, pp. 2201–2208, May 2019.

[21] D. Semrau, E. Sillekens, P. Bayvel, and R. I. Killey, "Modeling and mitigation of fiber nonlinearity in wideband optical signal transmission [Invited]," *J. Opt. Commun. Netw.*, vol. 12, pp. C68–C76, 2020.

[22] F. Munier, E. Alpman, T. Eriksson, A. Svensson, and H. Zirath, "Estimation of phase noise for QPSK modulation over AWGN channels," in *Proc. GigaHertz 2003 Symp.*, Linköping, Sweden, 2003, pp. 4–5.

[23] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-Efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[24] F. N. Khan *et al.*, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Exp.*, vol. 25, pp. 17767–17776, 2017.

[25] T. Hastie, R. Tibshriani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009, Ch. 11.

[26] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feedforward neural networks," *Chemometrics Intell. Lab. Syst.*, vol. 39, pp. 43–62, 1997.

[27] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural network," *J. Mach. Learn. Res.*, vol. 15, pp. 315–323, 2011.

[28] F. Chollet, "Keras," 2022. [Online]. Available: https://keras.io

[29] M. Abadi, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," in *Proc. Conf. Lang. Resour. Eval.*, 2016, pp. 3243–3249.

Yuchuan Fan (Graduate Student Member, IEEE) received the M.Sc. degree in electrical engineering from the Tampere University of Technology, Tampere, Finland, in 2017. He is currently working toward the Ph.D. degree in applied physics from the School of Engineering Sciences, KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include optical communication systems and machine learning applications for the physical layer aspect in fiber-optic and networking.

Xiaodan Pang (Senior Member, IEEE) received the M.Sc. degree from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2010, and the Ph.D. degree from the DTU Fotonik, Technical University of Denmark, Kongens Lyngby, Denmark, in 2013. He was a Postdoc with the RISE Research Institutes of Sweden (formerly ACREO Swedish ICT) from October 2013 to March 2017 and then was a Researcher with KTH Optical Networks Lab (ONLab) from March 2017 to February 2018. From March 2018 to February 2020, he was a Staff Opto Engineer and a Marie Curie Research Fellow with Infinera Corporation. Since March 2020, he has been a Senior Researcher with the Department of Applied Physics, KTH Royal Institute of Technology. He authored or coauthored more than 190 publications in journals and conferences. His research interests include ultrafast communications with millimeter-wave/terahertz, free-space optics, and fiber-optics. He has been the PI of a Swedish Research Council Starting Grant, EU H2020 Marie Curie Individual Fellowship Project NEWMAN, and a Swedish SRA ICT-TNG Postdoc project. He has been a TPC Member of in total of 19 conferences, including OFC'20-22, ACP'18-20, and GLOBECOM'20-21. He is a Senior Member of OSA, and a Board Member of IEEE Photonics Society Sweden Chapter.

Aleksejs Udalcovs (Senior Member, IEEE) received the M.Sc. degree in telecommunications and the Dr.Sc.Ing. degree in electronic communications from Riga Technical University, Riga, Latvia, in 2011 and 2015, respectively. From 2012 to 2016, he was with the KTH Royal Institute of Technology, Stockholm, Sweden, as a Ph.D. Researcher within Swedish Institute's Visby Program and as a Postdoctoral Researcher within the EU Project GRIFFON, Stockholm, Sweden. During this period, he became a member of Kista High Speed Transmission Laboratory, a research unit jointly owned and operated by KTH and by the RISE Research Institutes of Sweden. After receiving the research grant named SCENE–Spectrum, cost, and energy trade-offs in optical networks from Swedish ICT TNG consortium in 2016, he moved to RISE where he is a Senior Scientist providing his expertise in communication technologies. He is the author or coauthor of more than 100 papers in peer-reviewed international journals and conferences. His main research interests include the PHY-layer aspects of optical and photonic-wireless networks. He has participated in numerous experimental activities in a number of research groups, including KTH, Sweden, RISE, Sweden, VPIphotonics GmbH, Germany, Ghent University, Belgium, III-V Lab, France, DTU Fotonik, Denmark on fiber-optic transmission experiments, (sub-)system modeling, and optical network planning.

Carlos Natalino (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the Federal University of Pará, Pará, Brazil, in 2011 and 2016, respectively. From June 2016 to March 2019, he was a Postdoctoral Fellow with the KTH Royal Institute of Technology, Stockholm, Sweden, where he was a Visiting Researcher from 2013 to 2014. He is currently a Postdoctoral Researcher with the Chalmers University of Technology, Gothenburg, Sweden. He has authored or coauthored more than 40 papers published in international conferences and journals. His research focuses on the application of machine learning techniques for the optimization and operation of optical networks. He was a TPC Member in several international conferences and workshops, and a reviewer in several journals.

Lu Zhang (Member, IEEE) received the bachelor's degree from Southeast University, Nanjing, China, in 2014, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2019. He is currently a Research Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. From 2016 to 2017, he was a Visiting Ph.D. Student with the KTH Royal Institute of Technology, Stockholm, Sweden, sponsored by China Scholarship Council. Since 2018, he has been a Visiting Research Engineer with the KTH Royal Institute of Technology and Kista High-Speed Transmission Lab of RISE AB. His research interests include ultrafast THz communications, fiber-optic communications, digital signal processing algorithms for optical, and THz transmission systems.

Vjaceslavs Bobrovs (Member, IEEE) received the M.Sc. degree in telecommunications and the Dr.Sc.Ing. degree in electronics and telecommunications from Riga Technical University, Riga, Latvia, in 2005 and 2010, respectively. He is currently a Professor and the Director of Institute of Telecommunications, Head of Transmission Systems Department, and Member of RTU Science council. He is also engaged with the Latvian Council of Science, as an Expert in the area of physics and engineering and computer science, sub-area electronics and telecommunications. He is the coauthor of more than 170 articles in international journals and conferences. His research interests include nonlinear optics, metaphotonics, laser physics, fiber optical metro and access networks, radio-over-fiber technologies, digital signal processing, and quantum communication.

Richard Schatz was born in 1963. He received the Ph.D. degree in photonics from the Laboratory of Photonics and Microwave Engineering, Royal Institute of Technology, Stockholm, Sweden, in 1995. Since 1987, he has been conducting research and has been a Senior Researcher and Lecturer with the Laboratory of Photonics and Microwave Engineering, Royal Institute of Technology. He is also a part of Acreo-KTH Kista HST-Laboratory, Stockholm, Sweden. During 1992–1993, he was a Visiting Scientist with AT&T Bell Laboratories, Murray Hill, NJ, USA. He has developed laser simulation software for the photonics industry and authored or coauthored more than 150 journal papers and conference contributions. His research interests include modeling, design, and characterization of fiber-optical transmitters (edge emitter lasers, VCSELs, and modulators) and links, both for ON-OFF keying and more advanced modulation formats.

Xianbin Yu (Senior Member, IEEE) received the Ph.D. degree from Zhejiang University, Hangzhou, China, in 2005. From 2005 to 2007, he was a Postdoctoral Researcher with Tsinghua University, Beijing, China. Since November 2007, he has been with the DTU Fotonik, Technical University of Denmark, Kongens Lyngby, Denmark, where he became an Assistant Professor in 2009, and was promoted to a Senior Researcher in 2013. He is currently a Research Professor with Zhejiang University. He has coauthored more than 180 peer-reviewed international journal and conference papers within the fields of microwave photonics and optical fiber communications. He has given more than 40 invited conference presentations and was the session Chair/TPC Member of a number of international conferences. His research interests include mm-wave/THz photonics and its applications, THz communications, ultrafast photonic RF signal processing, and high-speed photonic wireless access technologies.

Marija Furdek (Senior Member, IEEE) received the Ph.D. and Dipl. Ing. degrees in electrical engineering from the University of Zagreb, Zagreb, Croatia, in 2012 and 2008, respectively. Since 2019, she has been an Assistant Professor with the Chalmers University of Technology, Gothenburg, Sweden. From 2013 to 2019, she was a Postdoc and then a Senior Researcher with the KTH Royal Institute of Technology, Stockholm, Sweden. She was a Visiting Researcher with Telecom Italia, Italy, Massachusetts Institute of Technology, Cambridge, MA, USA, and Auckland University of Technology, Auckland, New Zealand. Her research interests include optical network design and automation, with a focus on resiliency and physical-layer security. Dr. Furdek is the PI of the Project Safeguarding optical communication networks from cyber-security attacks, funded by the Swedish Research Council. As the PI or co-PI, Workgroup Leader, and Researcher, she has participated in several European, Swed, and Croatian research projects with a wide network of collaborators from industry and academia. She has coauthored more than 100 scientific publications in international journals and conferences, five of which was the recipient of best paper awards. She is currently the General Chair of the Optical Network Design and Modeling conference and was the General Chair of the Photonic Networks and Devices conference, part of the OSA Advanced Photonics Congress 2016–2019. She is an Associate Editor for the IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING and *Photonic Network Communications*, and the Guest Editor of IEEE/OSA JOURNAL OF LIGHTWAVE TECHNOLOGY and IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS. She is a Senior Member of OSA.

Sergei Popov received the M.Sc. degrees in applied physics and computer science from Russia, in 1987 and 1989, respectively, and the Ph.D. degree in applied physics from Finland, in 1999. He is currently a Professor with the Applied Physics Department, KTH Royal Institute of Technology, Stockholm, Sweden. He has authored or coauthored 300 papers and conference contributions. His research interests include optical communication, laser physics, plasmonics, and optical materials. Before joining KTH, he stayed with Ericsson Telecom AB and Acreo AB, Sweden. Dr. Popov is a OSA Fellow and the Editor-in-Chief of the *Journal of the European Optical Society: Rapid Publications* journal.

Oskars Ozolins (Member, IEEE) received the M.Sc. degree in telecommunications and the Doctor of engineering science (Dr.Sc.Ing.) degree in electronics and telecommunications from Riga Technical University, Riga, Latvia, in 2009 and 2013, respectively. He is a Senior Scientist and Technical Lead with the Kista High-speed Transmission Lab (Kista HST-Lab), RISE Research Institutes of Sweden, Sweden. He is also a Senior Researcher with the Department of Applied Physics, KTH Royal Institute of Technology. He is an Academician (foreign member) with the Latvian Academy of Science. In his professional career, he has been a Guest Researcher with III-V Lab, Keysight Technologies GmbH, Technical University of Denmark, Kongens Lyngby, Denmark, Ghent University, Ghent, Belgium, and University of Rennes 1, Rennes, France. He is the author of about 210 international journal publications, conference contributions, invited talks/tutorials/keynotes/ lectures, nine patents, and book chapters. His research interests include digital and photonic-assisted signal processing techniques, high-speed short-reach communications and devices, optical and photonic-wireless interconnects, and single-photon quantum communication. He was a regular Designated Reviewer of OSA Journals: *Journal of Optical Communications and Networking*, IEEE/OSA JOURNAL OF LIGHTWAVE TECHNOLOGY, *Photonics Technology Letters*, *Chinese Optics Letters*, *Applied Optics*, *Optics Express*, *Optics Letters*, and *Photonics Research*. He was the recipient of the Docent title from the KTH Royal Institute of Technology, Stockholm, Sweden in 2021.