



## **Fast signal quality monitoring for coherent communications enabled by CNN-based EVM estimation**

Downloaded from: <https://research.chalmers.se>, 2025-06-18 04:13 UTC

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

Fan, Y., Udalcovs, A., Pang, X. et al (2021). Fast signal quality monitoring for coherent communications enabled by CNN-based EVM estimation. *Journal of Optical Communications and Networking*, 13(4): B12-B20. <http://dx.doi.org/10.1364/JOCN.409704>

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

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

# Fast signal quality monitoring for coherent communications enabled by CNN-based EVM estimation

YUCHUAN FAN,<sup>1,2</sup> ALEKSEJS UDALCOVS,<sup>2</sup> XIAODAN PANG,<sup>1,2</sup> CARLOS NATALINO,<sup>3</sup> MARIJA FURDEK,<sup>3</sup> SERGEI POPOV,<sup>1</sup> OSKARS OZOLINS<sup>1,2\*</sup>

<sup>1</sup> School of Engineering Sciences, KTH Royal Institute of Technology, Isafjordsgatan 22, 164 40 Kista, Sweden

<sup>2</sup> RISE Research Institutes of Sweden, Isafjordsgatan 22, 164 40 Kista, Sweden

<sup>3</sup> Department of Electrical Engineering, Chalmers University of Technology, Chalmersplatsen 4, 412 96 Gothenburg, Sweden

\*Corresponding author: oskars.ozolins@ri.se

Received 8 September 2020; revised 27 November 2020; accepted 27 November 2020; published 18 January 2021 (Doc. ID 409704)

**We propose a fast and accurate signal quality monitoring scheme that uses convolutional neural networks (CNN) for error vector magnitude (EVM) estimation in coherent optical communications. We build a regression model to extract EVM information from complex signal constellation diagrams using a small number of received symbols. For the additive white Gaussian noise (AWGN) impaired channel, the proposed EVM estimation scheme shows a normalized mean absolute estimation error of 3.7% for quadrature phase shift keying (QPSK), 2.2% for 16-ary quadrature amplitude modulation (16QAM), and 1.1% for 64QAM signals, requiring only 100 symbols per constellation cluster in each observation period. Therefore, it can be used as a low-complexity alternative to conventional bit-error-rate (BER) estimation, enabling solutions for intelligent optical performance monitoring.**

<https://doi.org/10.1364/JOCN.409704>

## 1. INTRODUCTION

Coherent optical communications are widely used in metro and long-haul networks, where physical-layer impairments result in dynamically varying optical signal degradation. Advanced modulation formats, such as quadrature phase-shift keying (QPSK) and quadrature amplitude modulation (QAM) are now commonly used in coherent transceivers operating at 100 Gbps and beyond [1]. As optical networks become more heterogeneous and dynamic, accurate optical performance monitoring (OPM) characterized by, e.g. signal-to-noise ratio (SNR) or bit error rate (BER), is important to ensure reliable data transmission [2]. Error vector magnitude (EVM), as one of the commonly used performance metrics, contains signal quality information of high order modulated signals. For an additive white Gaussian noise (AWGN) channel, the EVM can be mapped to a BER and SNR [3]. Normally, millions of received symbols are used in EVM calculation [1],[3]. However, such a cumulative process is time-consuming and unsuitable to track fast network dynamics. Instead, a fast and accurate EVM monitoring scheme that requires only a small number of received symbols is needed.

Deep learning is a promising technique for OPM thanks to its ability to extract knowledge from high-dimensional data [4]–[6]. Various neural network types have been exploited for OPM, such as deep neural

networks (DNN), recurrent neural networks (RNN), and convolutional neural networks (CNN) [7]–[11]. For example, in [7], a DNN is used to estimate the optical signal-to-noise-ratio (OSNR) and perform modulation format recognition (MFR) based on signal amplitude histograms. In [8], an RNN architecture called long short-term memory (LSTM) is used to estimate the OSNR and nonlinear noise power in the frequency domain, achieving 1 dB test error both in OSNR and nonlinear noise power estimation. The LSTM capabilities in the monitoring of chromatic dispersion (CD) have also been demonstrated [9]. However, CNN attracts even more attention thanks to its powerful capabilities of extracting knowledge from graphical information [12], [13]. AlexNet [14], ResNet [15], and VGGNet [16] are well-known CNN structures that show high accuracy on pattern recognition tasks. CNNs have been used for MFR and OSNR estimation based on graphical information in [10],[11]. This method has been shown to extract signal features from complex signal constellation diagrams, with a 100 % MFR accuracy and less than 0.7 dB OSNR estimation error. However, to the best of our knowledge, CNN-based EVM estimation using constellation diagram images has not been studied.

In this paper, we propose an EVM estimation scheme that uses a CNN together with complex signal constellation diagrams for fast and accurate OPM. In coherent optical communications, constellation diagrams are commonly used to qualitatively evaluate the received signal as they provide a human-friendly visualization format whose

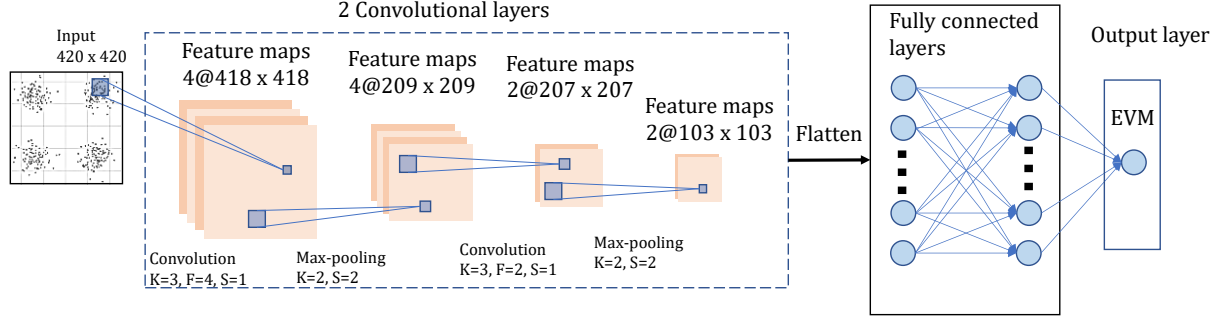


Fig. 1. Schematic diagram of the CNN structure (2 convolutional layers) for EVM estimation from constellation diagrams. K denotes the kernel size, F is the number of filters, and S is the stride step.

observation provides insight into, e.g. the modulation format, OSNR, EVM, phase noise, etc. However, distilling exact values of, e.g. BER from the diagram requires complex calculations that demand improvements in terms of automation and scalability. To this end, we apply supervised learning to analyze the visual representation of the channel state. The CNN structure is designed in such a way that each convolutional layer is followed by a max-pooling operation for complexity reduction. The resulting CNN can be applied directly to the constellation images for EVM estimation. Additional conversions, e.g. amplitude histogram [7], transfer frequency domain [8], or two-tap delay sampling [17], are avoided, enabling simplified data acquisition directly from transceivers. Moreover, unlike the previous demonstrations of CNN-based OPM schemes (e.g., [10]), the proposed approach allows for utilizing only a small number of received symbols to estimate the EVM. The accuracy of the proposed EVM estimator is verified using 32 Gbaud QPSK, 16QAM and 64QAM signals across the OSNR range of interest for practical operations. While using only 100 symbols per cluster in the complex signal constellation diagram, we achieve the accurate EVM estimation with a normalized mean absolute error (MAE) of 3.7% for QPSK ( $M = 4$  clusters), 2.2% for 16QAM ( $M = 16$  clusters), and 1.1% for 64QAM ( $M = 64$  clusters).

The rest of the paper is organized as follows. Section 2 gives an overview of the proposed technique and the methodology developed to obtain the constellation diagram dataset. Section 3 analyzes the CNN performance with respect to its architecture, i.e., the number and size of the convolutional layers, and the number of symbols/cluster. Section 4 provides concluding remarks.

## 2. OPERATION PRINCIPLE AND DATASET PREPARATION

In this section, we introduce the main characteristics of CNNs and the key hyperparameters for model optimization. Moreover, we describe the simulation setup used to collect the dataset comprising images of complex signal constellation diagrams for QPSK, 16QAM and 64QAM at different OSNR levels.

### 2.1 The employed CNN structures

The proposed CNN structure used for EVM estimation is shown in Fig. 1. In general, a CNN consists of an input layer that receives an  $n$ -dimensional array and a number of convolutional, pooling, and fully connected layers [18]. For instance, ordinary digital images (composed of red, blue and green components) are represented as a 3-dimensional array. Two of the dimensions represent the image in the vertical and horizontal directions, where each position in the array represents one pixel in the image. The third dimension contains the colors, where

usually three positions represent the three color components (red, blue and green) of each pixel. In single-color images (e.g., grayscale), only two dimensions are needed (for the horizontal and vertical directions).

CNNs have been successfully applied to image recognition problems in many research areas [16],[18]-[21] thanks to their excellent capabilities of capturing spatial correlations. Each convolutional layer contains multiple kernels, which are used to scan the entire image (or feature maps) from left to right and from top to bottom to obtain the output feature maps. Convolutional layers generate rich feature maps by convolving the input image (or feature maps from previous layers) and filters. Filter kernels are updated during the training process. The  $i^{\text{th}}$  feature map in convolutional layer  $l$  calculation can be expressed as:

$$x_i^l = f \left( \sum_{j=1}^{m^{l-1}} x_j^{l-1} * k_{i,j}^l + b_i^l \right), \quad (1)$$

where  $x_j^{l-1}$  and  $x_i^l$  are feature maps of previous layer  $l-1$  and current layer  $l$ ,  $k_{i,j}^l$  denotes the filter kernel connecting the  $j^{\text{th}}$  feature map in the previous layer with the  $i^{\text{th}}$  feature map in the current layer,  $b_i^l$  represents a bias matrix, and  $f(\cdot)$  is the activation function. The number of trainable parameters grows sharply as the number of convolutional layers increases.

We add max-pooling layers to grasp the main features of a certain region, reducing the dimensions of the feature map passed to the following layer. This allows us to effectively reduce the number of trainable parameters in the network. After the series of convolutional layers, the resulting  $n$ -dimensional array is transformed into a one-dimensional array (flattened) and passed to fully connected layers. Finally, the output layer outputs the estimated EVM value. In our case, we set one neuron in the output layer indicating an EVM value with a linear activation function. A rectified linear unit (ReLU) is selected as an activation function for the convolutional and fully connected layers. The ReLU operation for the input  $x$  of a neuron is given as [21]:

$$f(x) = \max(0, x). \quad (2)$$

It is well known that the estimation error of a defined estimator is related to the model complexity. A low complexity estimator may result in large prediction bias and poor generalization. Conversely, a high complexity model may adapt too closely to the training data, while estimation results on the test set have a high variance [22]. Therefore, it is important to adjust the model structure (the number of layers and filters) for the proposed signal quality monitoring scheme so that the model obtains a good balance between training and testing estimation error.

One particularity of the EVM estimation is that even small estimation errors are relevant. Therefore, during training, small errors should have a relevant contribution to the CNN updates. For this purpose, we use the

mean squared logarithmic error (MSLE) as the error function. The MSLE between the true EVM ( $EVM_t$ ) and the estimated EVM ( $EVM_e$ ) can be expressed as:

$$MSLE = \frac{1}{k} \sum_{i=1}^k (\log(EVM_{t_i} + 1) - \log(EVM_{e_i} + 1))^2 \quad (3)$$

where  $k$  represents the total number of samples.

## 2.2 Simulation setup for data collection

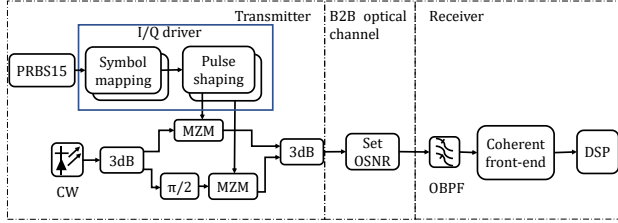


Fig. 2. Simulation setup for data collection where PRBS—pseudo-random binary sequence, CW—continuous wave, I/Q—in-phase/quadrature, MZM—Mach-Zehnder modulator, B2B—back-to-back, OBPF—optical bandpass filter, DSP—digital signal processing.

To collect the data, we set up a 32 Gbaud coherent optical system in VPItransmissionMaker [23] as shown in Fig. 2. The transmitter includes a continuous wave (CW) laser and a dual-parallel Mach-Zehnder modulator (MZM) driven by an In-phase/Quadrature (I/Q) driver for symbol mapping and pulse shaping. The pulse shaper is a root-raised cosine filter with a 0.15 roll-off factor. We use QPSK, 16QAM and 64QAM modulation formats. For each format, we simulate  $2^{19}$  symbols choosing 10 OSNR values to cover the EVM range of practical interest: OSNR = [12 : 30] dB for QPSK, OSNR = [20 : 38] dB for 16QAM, and OSNR = [26 : 44] dB for 64QAM, as shown in Fig. 3. These selected values ensure a BER below the hard-decision forward error correction (HD-FEC) threshold of  $3.8 \times 10^{-3}$  for QPSK and below the soft-decision FEC (SD-FEC) threshold of  $1 \times 10^{-2}$  for 16QAM and 64QAM. Figure 3 shows the modulated signal EVM conditions with respect to their OSNR values and the corresponding constellation diagrams. The true EVM values are computed from received symbols using the k-means clustering to obtain the constellation cluster centroids, which allows achieving high accuracy while avoiding the use of pilot symbols [24].

Figure 4 illustrates the dataset preparation stages. First, we define a number of symbols per constellation cluster ( $N$ ) to plot on a diagram.

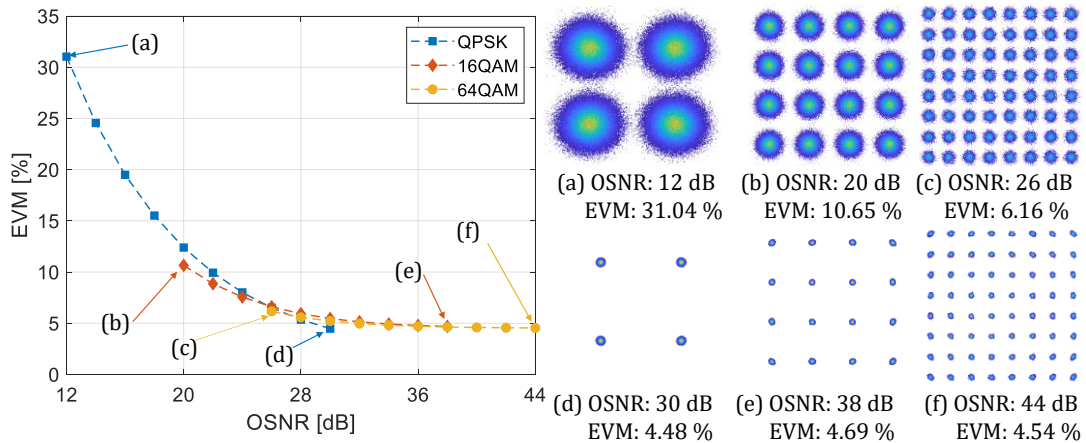


Fig. 3. The true EVM values of simulated QPSK, 16QAM, 64QAM signals with respect to the OSNR. (a)-(f) are the corresponding constellation diagrams for the end-points of the considered OSNR ranges.

Next, we plot constellation diagrams without fixed elements (axis boxes, labels and ticks) as shown in Fig. 3 (a)-(f). Then, we save  $L = 100$  such images during the training period  $T$ . Later, these input images are fed to a CNN scheme. A smaller  $N$  corresponds to a shorter training period  $T$  for signal quality monitoring. We train the CNN model over the constellation diagrams collected during the training period. The trained model is then applied to estimate the EVM by using data collected during the observation period. The monitoring interval shown in the figure is set by the network management system.

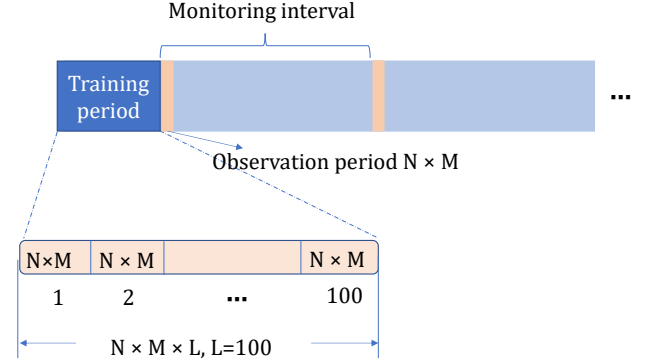


Fig. 4. Schematic diagram of the dataset collection.  $N$  is the number of symbols per cluster,  $M$  is the number of clusters in a complex signal constellation diagram,  $L$  is the number of constellation diagrams used for training.

To explore how many symbols are enough for accurate EVM estimation, we generate dataset options with  $N = 10, 100, 300, 500$  symbols per cluster in the complex signal constellation diagram. We refer to this as  $N$ -symbol/cluster dataset. Each  $N$ -symbol/cluster dataset consists of 30 simulation scenarios (three modulation formats and 10 different OSNR values), and each of them contains 100 images of constellation diagrams accumulated during the training period. Therefore, each  $N$ -symbol/cluster dataset contains 3000 images and 30 EVM labels. The training period and the observation window are set such that 50% and 25% of the  $N$ -symbol/cluster datasets are divided for training and testing, respectively. The remaining 25% of the dataset is used for validation during training.

Table 1. Configuration of convolutional layers for the considered CNN structures.

Structure	Structure info.	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
1	2+2 (3,3)	2 filters size (3,3)	2 filters size (3,3)			
2	4+2 (3,3)	4 filters size (3,3)	2 filters size (3,3)			
3	8+16+8 (3,3)	8 filters size (3,3)	16 filters size (3,3)	8 filters size (3,3)		
4	8+16+16+8 (3,3)	8 filters size (3,3)	16 filters size (3,3)	16 filters size (3,3)	8 filters size (3,3)	
5	16+16+16+8 (3,3)	16 filters size (3,3)	16 filters size (3,3)	16 filters size (3,3)	8 filters size (3,3)	
6	8+16+16+8+8 (5,5)	8 filters Size (5,5)	16 filters Size (5,5)	16 filters Size (5,5)	8 filters Size (5,5)	8 filters Size (5,5)
7	8+16+16+8+8 (3,3)	8 filters Size (3,3)	16 filters Size (3,3)	16 filters Size (3,3)	8 filters Size (3,3)	8 filters Size (3,3)
8	16+16+16+8+8 (3,3)	16 filters Size (3,3)	16 filters Size (3,3)	16 filters Size (3,3)	8 filters Size (3,3)	8 filters Size (3,3)

### 3. RESULTS AND DISCUSSION

We first investigate the impact of the CNN structure on the EVM estimation accuracy. For this purpose, the proposed monitoring scheme is evaluated with the dataset containing constellation diagrams of 300 symbols per constellation cluster. We use the Adam algorithm with a learning rate  $1e-4$  as the optimizer [25]. The CNN is built using Keras framework and TensorFlow library [26], [27]. The python code together with the entire dataset used to obtain the results presented in this paper is available for download [28].

Table 1 summarizes the configurations of the tested CNN structures. We vary the number of convolutional layers, the number of filters and their kernel size. The tested CNNs consist of up to 5 convolutional layers. The kernel size is 3 by 3 (3,3) for all structures except Structure 6, where it is set to 5 by 5 (5,5). As an example, Fig. 1 shows a schematic diagram of one of the CNN structures that we use for EVM estimation. It corresponds to Structure 2. The convolutional layers are followed by two fully connected layers with 500 and 100 nodes, respectively.

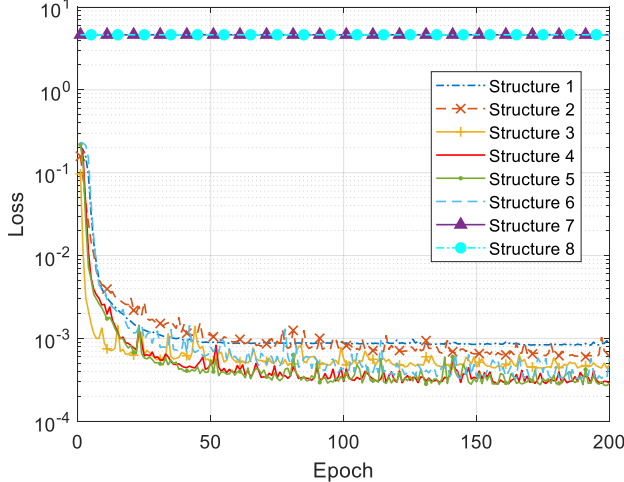


Fig. 5. The validation loss across training epochs for different CNN structures.

Figure 5 shows the validation loss for the tested configurations. From the figure, one can see that Structures 1 to 6 converge, whereas Structures 7 and 8 fail to learn EVM information. Further investigation indicates that training these structures with mean square error (MSE)

as the loss function would enable the network to decrease the loss over the epochs. The MSE is defined as:

$$MSE = \frac{1}{K} \sum_{i=1}^K (EVM_{t_i} - EVM_{e_i})^2, \quad (4)$$

Yet, after convergence, these structures still present the worst performance among the structures tested, i.e., structures 7 and 8 still have poorer performance than structures 1 to 6. Once converged, the loss performance does not improve over subsequent epochs. By comparing Structures 1 to 6, it is observed that the loss performance improves with the structure complexity up to a certain extent, while a further increase of complexity degrades the performance. Structures 4 and 5 ensure similarly low validation losses. However, considering the computational complexity, Structure 4 is more favorable as it balances the trade-off between the model complexity and its EVM estimation accuracy.

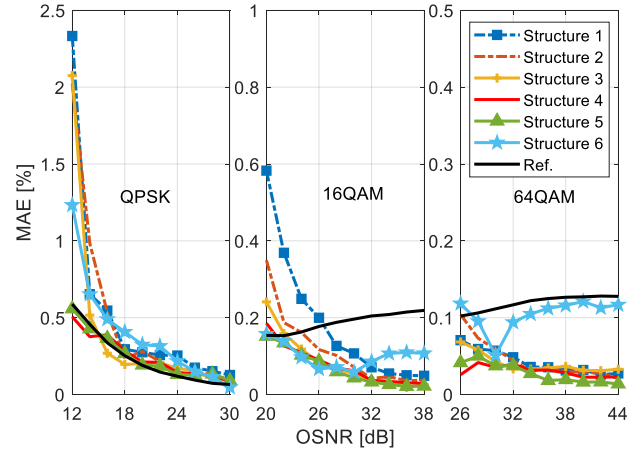


Fig. 6. Mean absolute error (MAE) of the estimated EVM values for the 300-symbol/cluster dataset with different layer configurations. The black solid line is the reference (Ref.) of conventional method.

Figure 6 illustrates the estimation mean absolute error (MAE) for the 300-symbol/cluster dataset. The MAE is defined as:

$$MAE [\%] = \frac{1}{n} \sum_{i=1}^n |EVM_{t_i} [\%] - EVM_{e_i} [\%]|, \quad (5)$$

where  $n$  is the number of images in the test dataset. For Structures 1 to 5, the estimation MAE steadily decreases with the increase of OSNR for



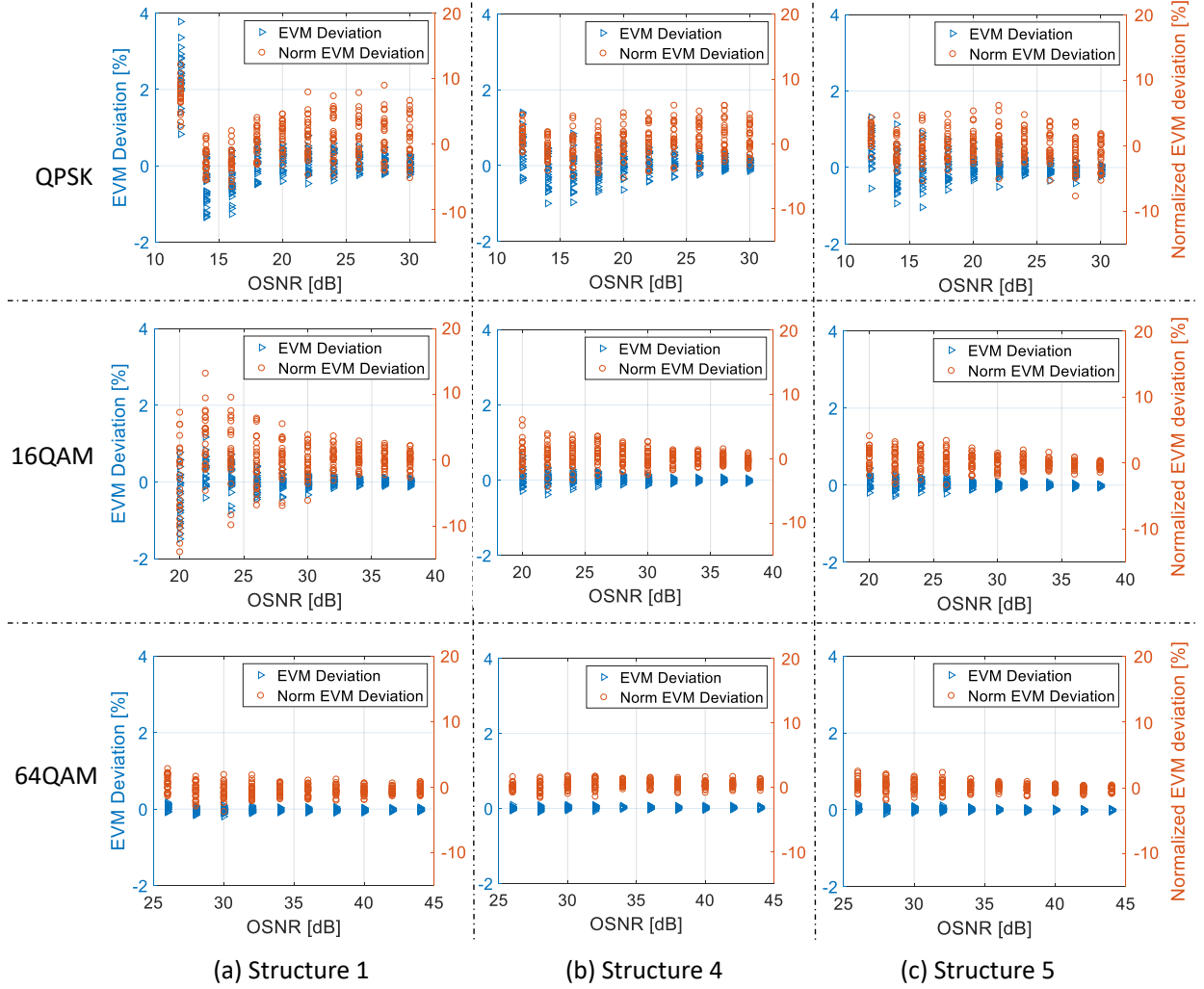


Fig. 7. Estimation performance for the test dataset of 300-symbol/cluster.

all three modulation formats. Figure 6 also shows at which conditions Structure 6 performance worse than Structure 4 and 5. For OSNRs above 30 dB, EVM of 16QAM and 64QAM signals cannot be estimated as accurately with Structure 6 as using Structures 4 and 5. Therefore, these 4-layer structures balance out the trade-off between the model complexity and its EVM estimation accuracy. Besides, these results show that the conventional EVM estimation method, shown as “Ref.” curves in Fig. 6, provides comparable performance for QPSK signals. However, its accuracy is worse for 16QAM and 64QAM signals, which can be explained by the higher number of clusters. As the modulation order increases, the CNN-based estimation has more information available for feature extraction and thus more accurate EVM estimation. In contrast, this additional information might be averaged out when the conventional centroid based EVM estimation is used.

We further evaluate the deviation of the EVM estimation for the 300-symbol/cluster dataset. In Fig. 7, we show 2-layer (Structure 1) and 4-layer (Structure 4 and 5) CNN performance for the test dataset in terms of EVM deviation and its normalized value. The normalized EVM deviation is calculated as follows:

$$\text{Normalized EVM deviation} [\%] = \frac{EVM_e - EVM_t}{EVM_t} \times 100. \quad (6)$$

The results in Fig. 7 show that all three structures ensure below 4% EVM deviation, which corresponds to the normalized EVM deviation of 15%. Since in practice BER is a more commonly used measure for signal quality monitoring, we select Structure 4 and the 300-symbol/cluster dataset to quantify the estimation accuracy under the assumption of an AWGN channel. The distribution of BER calculated from the estimated EVM is shown in Fig. 8. The solid lines are BER calculated from the true EVM ( $EVM_t$ ). For each considered combination of OSNR and modulation format, 25 estimations are obtained. One can see that BER fluctuations, observed due to the accuracy of the EVM estimation, are not significant enough to trigger a false alarm, especially when operating close to the BER threshold of a certain FEC code. Therefore, the proposed scheme represents an accurate tool for signal quality monitoring relying on EVM estimation.

To investigate how long the observation period should be for accurate EVM estimation, we numerically evaluate the proposed scheme using datasets with 10- to 1000-symbol/cluster. We use Structure 4 as the best-performing structure and show its estimation accuracy in Fig. 9 expressed with MAE and normalized MAE. The normalized MAE is defined as:

$$\text{Normalized MAE} [\%] = \frac{MAE}{EVM_t} \times 100. \quad (7)$$

The normalized MAE allows us to compare the performance of the proposed scheme across different modulation formats and their true EVM values. From Fig. 9, one can observe that the estimation is more accurate for higher OSNRs as the symbols are located closer to a reference point. Yet, we see that only 100-symbol/cluster is sufficient to ensure a normalized MAE of 3.7% for QPSK, 2.2% for 16QAM, and 1.1% for 64QAM signals. Considering the order of the estimated values, such small errors result in a negligible EVM fluctuation that does not impact the system status even when operating close to the FEC limit.

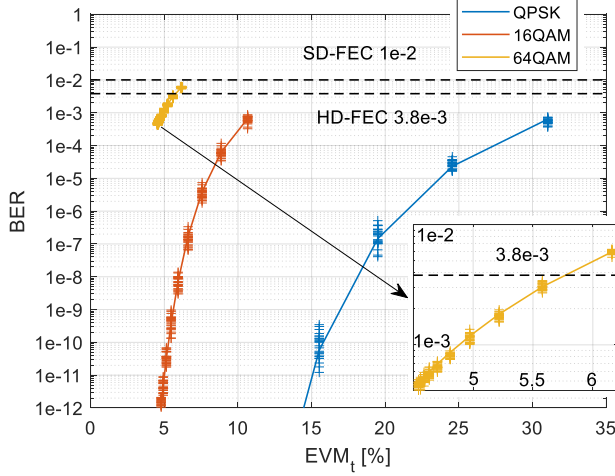


Fig. 8. BER calculated from the true (solid line) and estimated (+) EVMs for a 300-symbol/cluster dataset using the proposed CNN scheme with Structure 4.

To test the performance of the proposed EVM estimation scheme for signal quality monitoring after a long-haul transmission, we build an amplified fiber-optic link using 100 km long spans of the standard single-mode fiber (SSMF, chromatic dispersion coefficient  $D = 16\text{e-}6$  s/m<sup>2</sup>, attenuation coefficient  $\alpha = 0.2$  dB/km, and nonlinear refractive index  $n = 2.6\text{e-}20$  m<sup>2</sup>/W). As previously, we set the OSNR at the transmitter to 45 dB and we control it after every two to four 100 km spans using an optical spectrum analyzer. In such a way, we know the exact OSNR value at a specific point of the link where we also collect constellation diagrams for EVM estimation and accuracy analysis. Figure 10 shows how the true EVM, estimated EVM, and MAE evolve with the transmission distance. The proposed scheme provides a normalized MAE below 6.2% (QPSK), 2.6% (16QAM), and 2.8% (64QAM) using Structure 4 and only 100 symbols per constellation cluster. These results indicate that the proposed scheme achieves good generalization capability. However, lower errors might be achieved by a structure that is specialized for long-haul transmission. This will be addressed in further research as it deserves special attention.

The computation time is also evaluated for Structure 4 and conventional approach with a computer powered by an Intel Xeon E5-2630-v3 processor running at 2.4 GHz, 64 GB of RAM, and GTX TITAN Black graphics card. For the 100-symbol/cluster dataset, the inference time of the conventional estimation method is 11.2 ms, 14.5 ms, and 35.7 ms for QPSK, 16QAM and 64QAM, respectively. For the proposed EVM estimation scheme, the inference of a single constellation diagram takes approximately 2.7 ms regardless of the modulation type. Such a fast inference time makes the proposed scheme a viable candidate for real-time OPM. Note that the training time for 1500 training samples (constellations) is 1400 seconds or 7 seconds per epoch but it's done offline.

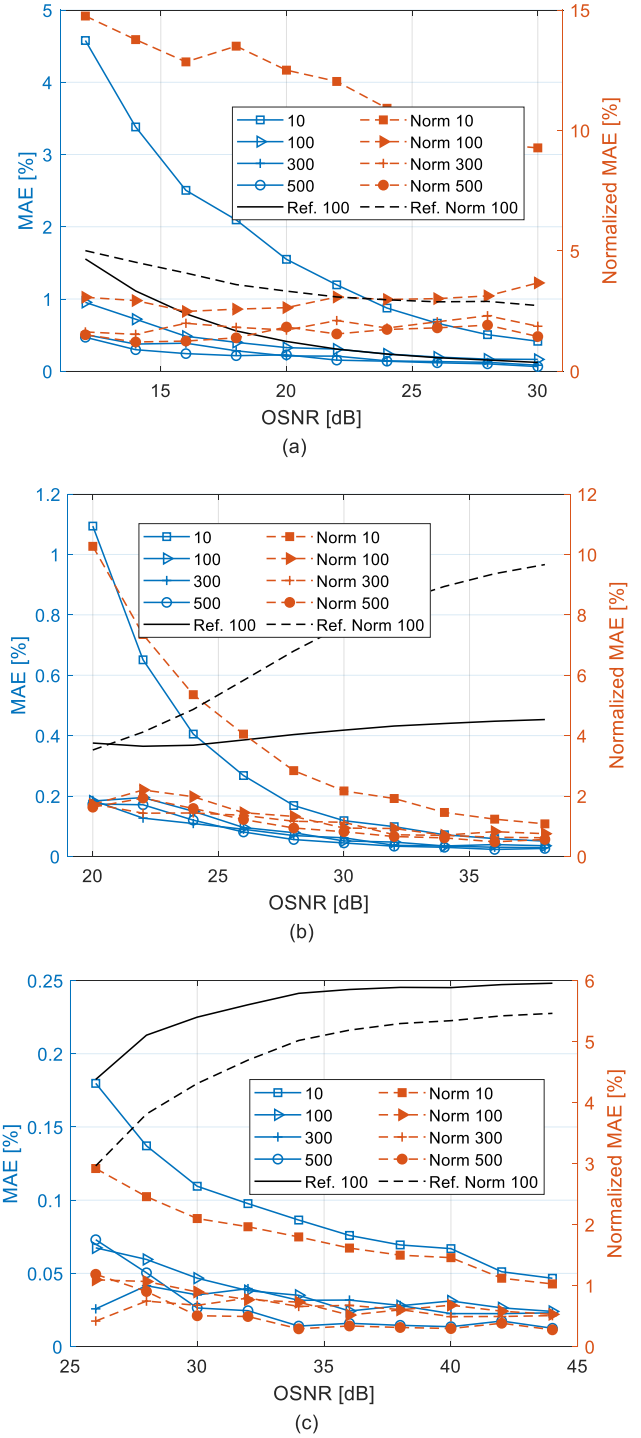


Fig. 9. Test performance of the proposed EVM monitoring scheme relying on the CNN Structure 4 and datasets containing constellation diagrams with 10- to 500-symbol/cluster. The Ref. curves are baselines obtained using the conventional method applied for the 100-symbol/cluster dataset. (a): QPSK, (b): 16QAM, (c): 64QAM.

#### 4. CONCLUSIONS

A CNN-based EVM estimation scheme is proposed for signal quality monitoring in coherent optical communication systems. It relies on

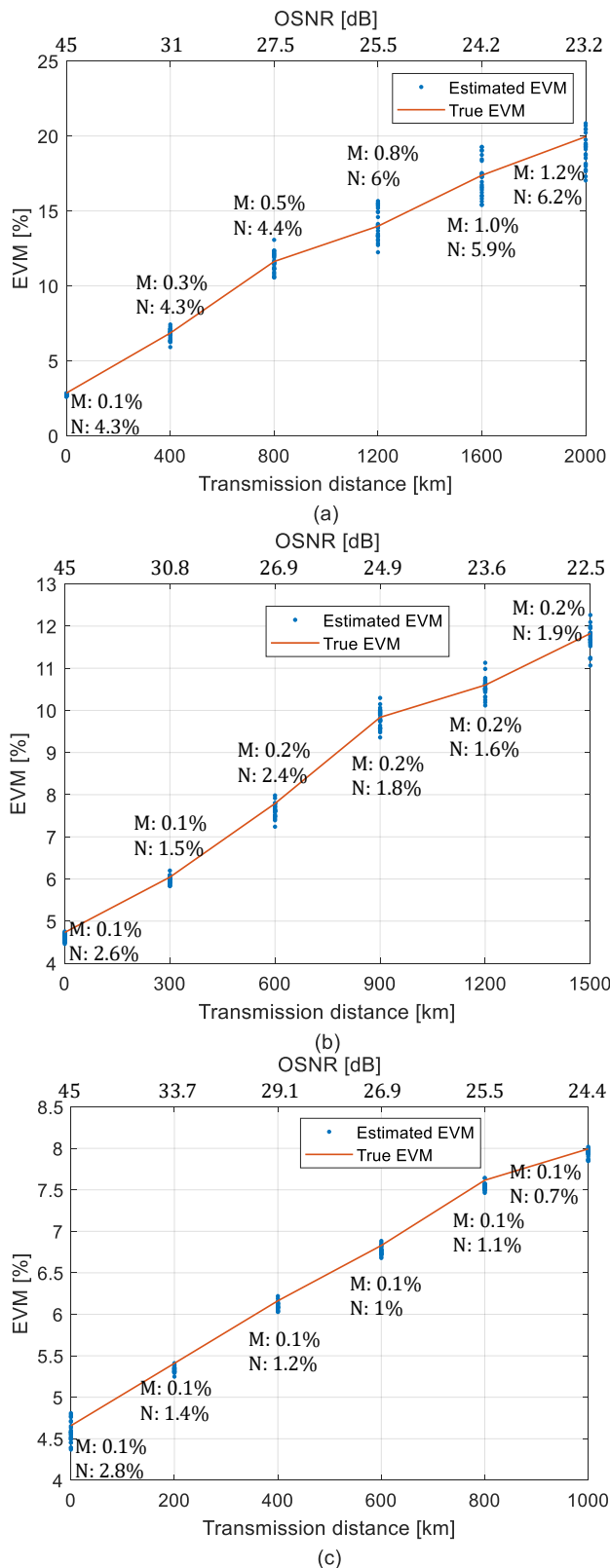


Fig.10. True versus estimated EVM with respect to transmission distance. (a) QPSK, (b) 16QAM, (c) 64QAM. M denotes MAE, N—normalized MAE.

images of complex signal constellation diagrams fed into the low-complexity regression model that consists of interleaved convolutional

layers and max-pooling layers. The performance of the proposed scheme is validated with 32 Gbaud QPSK, 16QAM, and 64QAM signals at different OSNR values of practical interest. Besides, two different transmission configurations are tested: an optical-back-to-back, representing an AWGN-impaired optical channel, and a long-haul (>1000 km) fiber transmission, including the AWGN and fiber nonlinearity induced noise. The estimation accuracy is investigated considering CNN architecture and the number of symbols in the constellation diagrams. The results show that CNN structures consisting of 2 to 5 convolutions layers ensure the best performance in terms of the computational complexity and EVM estimation accuracy. Further increase in complexity might be inadequate for the specific problem and thus lead to the degradation of estimation accuracy. The 4-layer CNN architecture with Structure 4 (8+16+16+8 filters per layer and (3,3) kernel) provides the most accurate EVM estimation regardless to the OSNR values in the system. For the AWGN-impaired channel, the normalized mean absolute error of 3.7% for QPSK, 2.2% for 16QAM, and 1.1% for 64QAM is achieved with only 100 symbols per cluster in the complex signal constellation diagram. The corresponding values for a long-haul fiber transmission are 6.2% for QPSK after 2000 km, 2.6% for 16QAM after 1500 km, and 2.8% for 64QAM after 1000 km. Such accuracy together with a 2.7 ms short observation period allows considering the proposed scheme as an enabler for intelligent optical performance monitoring.

**Funding Information.** This work was supported by the ERDF-funded project CARAT (1.1.1.2/VIAA/4/20/660), the Swedish Research Council (Vetenskapsrådet) within the project PHASE (2016-04510) and the project 2019-05197, and RISE project “AI in optical transmission”.

## References

1. R. Schmogrow et al., “Error vector magnitude as a performance measure for advanced modulation,” *IEEE Photonics Technology Letters*, vol. 24, no. 1, pp. 61-63, Jan. 2012.
2. D.C. Kilper et al., “Optical performance monitoring,” *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 294-304, Jan. 2004.
3. R. Shafik, Rahman, and Ahm Razibul Islam, “On the extended relationships among EVM, BER and SNR as performance metrics,” in *Proc. International Conference on Electrical and Computer Engineering*, Dhaka, Bangladesh, Dec. 2006, pp. 408-411.
4. F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, “An overview on application of machine learning techniques in optical networks,” *IEEE Commun. Surv. Tut.*, vol. 21, no. 2, 2019.
5. F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, “An optical communication’s perspective on machine learning and its applications,” *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 2, Jan. 2019.
6. F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, “Machine learning methods for optical communication systems and networks,” in *Optical Fiber Telecommunications*, 7th ed., Chapter 21, Academic Press, 2019.
7. F. N. Khan et al., “Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks,” *Optics express*, vol. 25, no. 15, pp. 17767-17776, Jul. 2017.
8. Z. Wang, A. Yang, P. Guo, and P. He, “OSNR and nonlinear noise power estimation for optical fiber communication systems using LSTM based deep learning technique,” *Optics express*, vol. 26, no. 16, pp. 21346-21357, Aug. 2018.
9. C. Wang et al., “Joint OSNR and CD monitoring in digital coherent receiver using long short-term memory neural network,” *Optics Express*, vol. 27, no. 5, pp. 6936-6945, Mar. 2019.
10. D. Wang et al., “Intelligent constellation diagram analyzer using convolutional neural network-based deep learning,” *Optics express*, vol. 25, no. 15, pp. 17150-17166, Jul. 2017.



11. C. Natalino, A. Udalcovs, L. Wosinska, O. Ozolins, M. Furdek, "One-Shot Learning for Modulation Format Identification in Evolving Optical Networks," in OSA APC 2019 (IPR, Networks, NOMA, SPPCom, PVLED), OSA Technical Digest (Optical Society of America, 2019), paper JW4A.2.
12. J. Deng et al., "ImageNet: a large-scale hierarchical image database," in IEEE Conf. Comput. Vision and Pattern Recognit., pp. 248–255, 2009.
13. O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vision*, 115 (3), pp. 211–252, 2015.
14. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
15. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770–778.
16. K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556.
17. S. D. Dods and T. B. Anderson, "Optical performance monitoring technique using delay tap asynchronous waveform sampling," in Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference, Anaheim, CA, 2006, pp. 175–192.
18. Y. LeCun, and Y. Bengio. "Convolutional networks for images, speech, and time series." *The handbook of brain theory and neural networks*, 3361.10 (1995): 1995.
19. Y. Lecun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in Proc. IEEE International Symposium on Circuits and Systems (ISCAS), 2010, pp. 253–256, doi: 10.1109/ISCAS.2010.5537907.
20. K. Simonyan, and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in Proc. NIPS, 2014, CoRR, abs/1406.2199.
21. X. Glorot, A. Bordes, Y. Bengio, "Deep Sparse Rectifier Neural Networks," *Journal of Machine Learning Research*, vol. 15, pp. 315–323, Jun. 2011.
22. T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: data mining, inference, and prediction," Springer Science & Business Media, 2nd edition, 2009.
23. VPIphotonics GmbH, "VPItransmissionMaker10," <https://www.vpi Photonics.com/>, accessed May 2020.
24. Zhang, Qun et al., "Accurate BER Estimation Scheme Based on K-Means Clustering Assisted Gaussian Approach for Arbitrary Modulation Format," *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2152–2157, 2020.
25. D. P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," in International Conference for Learning Representations (ICLR), San Diego, 2015.
26. F Chollet et al., "Keras," <https://keras.io>, accessed May 2020.
27. M. Abadi, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," in Proc. Conf. Lang. Resour. Eval., Mar. 2016, pp. 3243–3249.
28. The code and dataset on GitHub: [https://github.com/JhoneFan/2020\\_JOEN\\_EVM\\_Estimation\\_using\\_CN\\_N.git](https://github.com/JhoneFan/2020_JOEN_EVM_Estimation_using_CN_N.git)

**Yuchuan Fan** received the M.Sc. degree in Electrical Engineering from Tampere University of Technology, Finland, in 2017. He is currently pursuing his Ph.D. degree in Applied Physics from School of Engineering Sciences at KTH Royal Institute of Technology in Stockholm, Sweden. His research interests include optical communication systems and machine learning application for the physical layer (PHY) aspect in fiber-optic and networking.

**Aleksejs Udalcovs** (M'16) received his M.Sc. degree in telecommunications and the Ph.D. degree (Dr.sc.ing.) in electronic communications from Riga Technical University in 2011 and 2015, respectively.

In 2012–2016, he worked at KTH Royal Institute of Technology as a Ph.D. Researcher within the Swedish Institute's Visby program and then as a Postdoctoral Researcher within the EU project GRIFFON in Stockholm, Sweden. During this period, he became a member of Kista High Speed Transmission Lab jointly owned and operated by KTH and by RISE Research Institutes of Sweden. In 2016, after receiving the research grant named "SCENE -- Spectrum, cost and energy trade-offs in optical networks" from the Swedish ICT TNG consortium, he moved to RISE. Since 2019 he is a Senior Scientist at RISE, who is working in cooperation with industry and academia providing his expertise in communication technologies. His main research interests are within the PHY-layer aspects in optical transport and photonic-wireless networks.

Dr. Udalcovs is a (co-)author of more than 100 papers in peer-reviewed international journals and conferences. He has participated in numerous experimental activities in a number of research groups (in Sweden—KTH, RISE; in Germany—VPIphotonics GmbH; in Belgium—Ghent University; in France—III-V Lab; in Denmark—DTU Fotonik) on fiber-optic transmission experiments, modeling of fiber-optic links and (sub-)systems, and optical network planning. Dr. Udalcovs is a Member of IEEE.

**Xiaodan Pang** (SM'19) received the M.Sc. degree from KTH Royal Institute of Technology, Sweden in 2010, and the PhD degree from DTU Fotonik, Technical University of Denmark in 2013.

He was a Post Doc at RISE Research Institutes of Sweden (former ACREO Swedish ICT) from October 2013 to March 2017 and then worked as a researcher with the KTH Optical Networks Lab (ONLab) from March 2017 to February 2018. From March 2018 to February 2020 he worked as a Staff Opto Engineer and a Marie Curie Research Fellow at Infinera Corporation. Since March 2020 he has been working as a Senior Researcher at Department of Applied Physics, KTH Royal Institute of Technology, Sweden. He is/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 Post Doc project.

Dr. Pang's research focuses on ultrafast communications with MMW/THz, free-space optics and fiber-optics. He authored and coauthored over 170 publications in journals and conferences. He has been a TPC member of in total 17 conferences including OFC'20–21, ACP'18–20 and GLOBECOM'20. He is an IEEE Senior Member, an OSA Senior Member and a Board Member of IEEE Photonics Society Sweden Chapter.

**Carlos Natalino** (M'17) received the M.Sc. and Ph.D. degrees in electrical engineering from the Federal University of Pará, Brazil, in 2011 and 2016, respectively.

He was a Post-Doctoral Fellow with the KTH Royal Institute of Technology from June 2016 to March 2019, where he was a Visiting Researcher from 2013 to 2014. He is a Post-Doctoral Researcher with the Chalmers University of Technology.

Dr. Natalino authored/coauthored over 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 has served as a TPC member in several international conferences and workshops, and a reviewer in several journals

**Marija Furdek** (M'09-SM'17) obtained her PhD and Dipl. Ing. degrees in electrical engineering from the University of Zagreb, Croatia, in 2012 and 2008, respectively.

She has been an Assistant Professor at Chalmers University of Technology in Gothenburg, Sweden, since 2019. From 2013 to 2019 she was a Postdoc and then Senior Researcher at KTH Royal Institute of Technology in Stockholm, Sweden. She was a visiting researcher at Telecom Italia, Italy, Massachusetts Institute of Technology, USA, and Auckland University of Technology, New Zealand. Her research interests encompass 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 (co)PI, work group leader and researcher, Marija has participated in several European, Swedish and Croatian research projects with a wide network of collaborators from the industry and academia. She has co-authored more than 100 scientific publications in international journals and conferences, 5 of which received best paper awards. She is currently serving as a General Chair of the Optical Network Design and Modeling (ONDM) conference and was a General Chair of the Photonic Networks and Devices conference, part of the OSA Advanced Photonics Congress 2016-2019. She is an Associate Editor of IEEE/OSA Journal of Optical Communications and Networking and Photonic Network Communications, and a Guest Editor of IEEE/OSA Journal of Lightwave Technology and IEEE Journal of Selected Topics in Quantum Electronics. She is a Senior Member of IEEE and OSA.

**Sergei Popov** is Professor at Applied Physics Department at KTH Royal Institute of Technology, Stockholm, Sweden. He holds M.Sc. degrees in Applied Physics (1987), Computer Science (1989), Russia; and Ph.D. in Applied Physics (1999), Finland.

His expertise covers optical communication, laser physics, plasmonics, and optical materials. He stayed with Ericsson Telecom AB and Acreo AB (both in Sweden) before joining KTH.

Prof. Popov is OSA fellow, Editor-in-Chief of JEOS:RP journal (EOS), and he published over 300 papers and conference contributions.

**Oskars Ozolins** (M'09) received his M.Sc. degree in telecommunications and his degree of Doctor of Engineering Science (Dr.sc.ing.) in electronics and telecommunications from Riga Technical University in Riga, Latvia, in 2009 and 2013, respectively.

He is a Senior Scientist and a Technical Lead of KTH/RISE Kista High-speed Transmission Lab at RISE Research Institutes of Sweden where he is working under Swedish Research Council starting grant project "Photonic-assisted signal processing techniques (PHASE)". He is also appointed as an Affiliated Faculty and Senior Researcher on optical communication at the Department of Applied Physics at KTH Royal Institute of Technology. His research interests are in the areas of digital and photonic assisted signal processing techniques, high-speed short-reach communications and devices, optical and photonic-wireless interconnects, and single photon quantum communication. In his professional career Dr. Ozolins has been guest researcher at III-V Lab (Nokia Bell Labs and Thales, France), Keysight Technologies (Boblingen Germany), DTU Fotonik (Technical University of Denmark, Denmark), IDLab (Ghent University--imec, Belgium), OFO (KTH Royal Institute of Technology, Sweden), and FOTON laboratory (University of Rennes 1, France).

Dr. Ozolins is (co-)author of more than 195 international journal publications, conference contributions, invited talks/tutorials/keynote/lecture, patents, and book chapters.