

Received December 2, 2019, accepted December 26, 2019, date of publication January 1, 2020, date of current version January 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963504

# Identifying Probabilistically Shaped Modulation Formats Through 2D Stokes Planes With Two-Stage Deep Neural Networks

WENBO ZHANG<sup>1</sup>, DINGCHENG ZHU<sup>1</sup>, NANNAN ZHANG<sup>2</sup>, HENGYING XU<sup>3,4</sup>,  
XIAOGUANG ZHANG<sup>2</sup>, (Senior Member, IEEE), HU ZHANG<sup>2</sup>, AND YONG LI<sup>5</sup>

<sup>1</sup>School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>2</sup>State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>3</sup>Physics Science and Information Engineering College, Liaocheng University, Liaocheng 252000, China

<sup>4</sup>Shandong Provincial Key Laboratory of Optical Communication Science and Technology, Liaocheng 252059, China

<sup>5</sup>School of Electronic Engineering, Beijing Key Laboratory of Work Safety and Intelligent Monitoring, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Wenbo Zhang (zhangwb@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571057, Grant 61575082, and Grant 61527820.

**ABSTRACT** A lightweight two-stage convolutional (deep) neural network (CNN) based modulation format identification (MFI) scheme is proposed and demonstrated for the polarization domain multiplexing (PDM) fiber communication system with probabilistically shaped (PS) modulation formats. The scheme is tested on a PDM system at a symbol rate of 28 GBaud. Six probabilistically shaped (PS) modulation formats (of 3 bit/symbol PS-16QAM, PS-32QAM, and PS-64QAM, of 4 bit/symbol PS-32QAM and PS-64QAM, and of 5 bit/symbol PS-64QAM) along with six standard modulation formats (BPSK, QPSK, 8PSK and three uniformly shaped (US) QAM: US-16QAM, US-32QAM and US-64QAM) are identified by the trained CNN. By taking advantage of computer vision, the results show that the proposed scheme can provide very high accuracy and significantly improve the identification performance over the existing techniques. The influences of the learning rate of the CNN are also discussed.

**INDEX TERMS** Optical fiber communication, modulation format identification, convolutional neural networks.

## I. INTRODUCTION

To meet the demand of various data services, an intelligent optical network [1], which is called elastic optical networks (EON), has drawn considerable interest in these years. This network has the capability of adjusting transceiver configurations, such as changing modulation format and data rates intelligently [2] according to different data demands. The key to implementing EON is to design a hitless flexible transceiver that can identify the modulation format at the receiver end (Rx) to ensure proper demodulation.

A number of modulation format identification (MFI) methods were proposed, which can be used to design the hitless flexible Rx, including some blind algorithms [3], [4] and pilot aided [5] schemes. Other MFI methods like the number of clusters or the higher-order statistics [6]–[10] of

the power distributions in the Stokes space or constellation planes [11]–[14] of received signals. However, these methods are both designed to handle the problem of MFI of standard modulation formats such as PSK signals and uniformly shaped (US) QAM signals. To the best of our knowledge, there is no existing method that can be used to identify the modulation format when the probabilistically shaped QAM signals are involved.

Probabilistic shaping (PS) has been receiving a significant amount of attention since it can realize higher capacity and higher spectral efficiency on optical communications [15]. The challenge of the MFI scheme is that the accuracies of the above MFI methods will tend to some lower values since the PS-mQAM will tend to some other kinds of standard modulation formats or other PS-mQAM for some special cases. For instance, as in some specific conditions, in the Jones space, PS-16QAM may have very small differences between US-4QAM or QPSK signals, and PS-64QAM may

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

also appear similar to PS-16QAM. Although the kind of PS-QAM can be determined after we obtain the channel quality (OSNR or something else) in the receiver side, and receivers seem to know what PS-QAM will be sent, the quality of the channel is not always constant in a real system, especially in an EON system, and it is hard for a receiver to have any preknowledge of the modulation format used by various transmitters with different OSNRs.

Machine learning methods have attracted much attention, and some ANNs methods are used as identification techniques [16]. However, ANN is not suitable for complicated situations, such as having too many modulation formats or high-order QAM signals. Recently, convolutional neural networks (CNNs) have been used in the area of artificial vision, and their performances are very good [17], [18]. An image-based method using a 6-layer CNN proposed in [19] improved the identification performance. However, the new method that uses Jones constellations has poor tolerance to frequency offset and phase noise, and the OSNR values for identifying high-order QAM signals precisely remain relatively high. Additionally, the identification rate will be even worse if PS-QAM signals are added. In [20], we proposed a high-performance MFI scheme by successfully using lightweight convolutional neural networks (CNNs). By changing the received signals into Stokes space and projection constellations in Stokes space onto three coordinate planes, the CNN-based MFI scheme can achieve very high accuracy, even if the OSNR is very low. It is also reasonable to design the MFI scheme following a similar step introduced in [20].

In this paper, we propose our MFI scheme using two-stage CNNs, which can also handle the situations in which some commonly used PS modulation formats are involved. By mapping received signals into the Stokes space and projecting the constellations onto three coordinate planes, we can obtain a series of images, which are used as inputs of the CNNs. Twelve modulation formats are tested. They are BPSK, QPSK, 8PSK, US-16QAM, US-32QAM, US-64QAM, PS-16QAM (3 bit/symbol), PS-32QAM (3 and 4 bit/symbol), PS-64QAM (3, 4, 5 bit/symbol). The results show that our CNN-based MFI scheme performs very well. All modulation formats can be identified when OSNR is higher than 15 dB with accuracy higher than 95%. The tolerances of the proposed MFI scheme with respect to chromatic dispersion (CD), polarization-dependent loss (PDL), and polarization mode dispersion (PMD) are also discussed.

In section II, the two-stage CNN-based MFI scheme is introduced. The setup of the experimental system is shown in section III. In section IV, the performances of the scheme are discussed, and we explain our conclusions in section V.

## II. MODULATION FORMATS IDENTIFICATION SCHEME

### A. GENERATION OF PS-QAM SIGNALS

The probabilistic amplitude shaping (PAS) architecture [21] is proposed for making PS practical. In general, the standard

PAS process mainly contains three steps. In the first step, a distribution matcher (DM) is used. The outputs of the DM can be used to form PS  $\sqrt{m}$  pulse amplitude modulation ( $\sqrt{m}$ -PAM) symbol sequence. The DM generates the shaped positive amplitude  $A = \{1, 2, \dots, \sqrt{m}/2\}$  of the constellation symbols  $X$  following a probability distribution  $P_A$  with corresponding entropy  $H(P_A)$  [22], where  $m$  is the order of QAM symbols. From  $H(P_A)$ , we can obtain the entropy  $H(P_X)$  (in bit/symbol) with all the amplitude portions of all constellation symbols  $X$ . The entropy  $H(P_X)$  decreases when the distributions of PS-QAM signals become more shaped (see Fig. 1). In this article, the constant composition distribution matching (CCDM) [23] is used. In the second step, a forward error correction (FEC) encoder might be used to ensure error-free transmission. Finally, in the third step, a symbol mapping maps shaped stream of bits into a stream of symbols. A brief architecture of PAS is shown in Fig. 2. Different from other PS-mQAM, the probabilistic fold shaping (PFS) method is used to generate PS-32QAM [24].

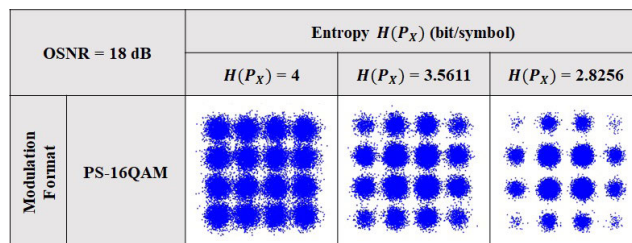


FIGURE 1. Constellations of 18 dB PS-16QAM signals at different  $H(P_X)$ . From left to right, the value of  $H(P_X)$  decreases, and the distributions of PS-QAM signals become more shaped.

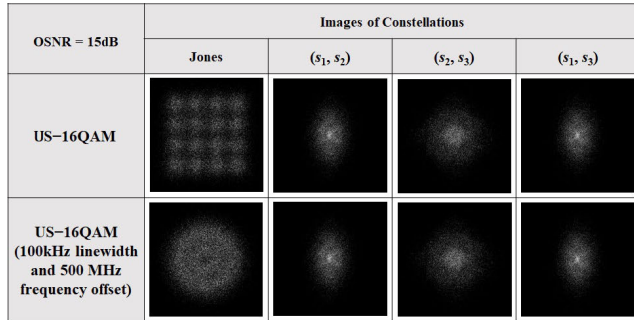
### B. STOKES MAPPING AND IMAGE GENERATION

In a PDM system, after a coherent receiver, the corresponding constellation can be obtained. In [19], the MFI method used images generated in a Jones constellation diagram. However, constellation points randomly distribute around the ideal location or shift in real optical communication systems with signal degradation, such as phase noise and frequency offset. In this case, especially when OSNR is relatively low, MFI schemes may not obtain a reliable performance. In [20], we proposed a CNN-based method by projecting constellations in the Stokes space onto the Stokes plane, which can overcome phase noise and frequency offset. In detail, the received PDM signals can be mapped into Stokes space using the formula

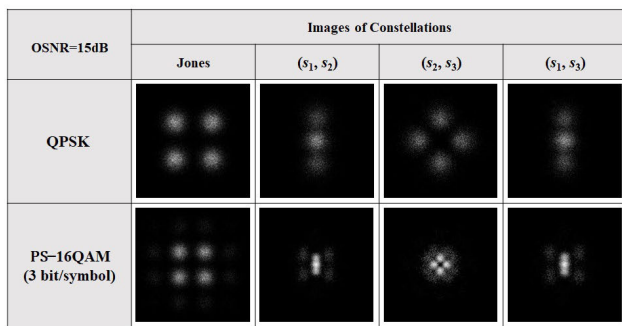
$$S = \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} e_x e_x^* + e_y e_y^* \\ e_x e_x^* - e_y e_y^* \\ e_y e_x^* - e_x e_y^* \\ -j e_y e_x^* + j e_x e_y^* \end{pmatrix} = \begin{pmatrix} a_x^2 + a_y^2 \\ a_x^2 - a_y^2 \\ 2a_x a_y \cos \delta \\ 2a_x a_y \sin \delta \end{pmatrix} \quad (1)$$



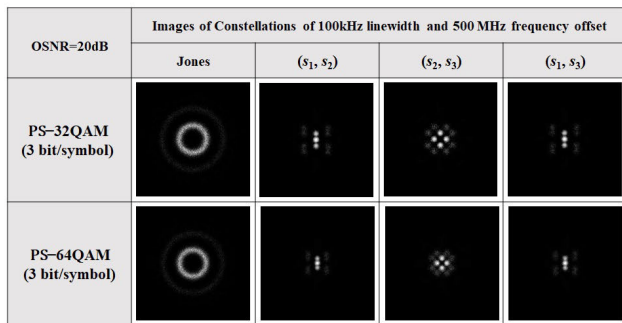
FIGURE 2. Brief architecture of PAS.



(a)



(b)



(c)

FIGURE 3. Images of constellations in Jones space and three different 2D Stoke planes generated by using 20,000 signals. (a) US-16QAM with and without a 100 kHz laser linewidth and 500 MHz frequency offset at 15 dB OSNR. (b) QPSK and 3 bit/symbol PS-16QAM at 15 dB OSNR. (c) 3 bit/symbol 32QAM and 64QAM at OSNR 20 dB with 100 kHz laser linewidth and 500 MHz frequency offset.

where  $(e_x, e_y)$  are received PDM complex signals after algorithms and  $(ax, ay)$  are the amplitudes of the complex signals.  $\delta$  is the phase difference between  $e_x$  and  $e_y$ . A 3D Stokes space can be obtained by the last three components  $(s_1, s_2, s_3)$  of the Stokes vector  $S$  in formula (1). By noticing that the amplitudes and relative phase of the signals are retained, phase noise and frequency offset vanish, and the constellations of signals in Stokes space are independent from the phase noise and frequency offset (see Fig. 3 (a)). As in [20],

each constellation point in the Stokes space,  $(s_1, s_2, s_3)$ , can be projected onto three Stokes planes, plane  $(s_1, s_2)$ ,  $(s_2, s_3)$  and  $(s_1, s_3)$ , then three images for the MFI scheme can be generated.

In Fig. 3, QPSK and PS-16QAM are considered. In Fig. 3(b), PS-16QAM (3 bit/symbol) might not be easily distinguished from the QPSK in the Jones constellation plane, but they can still be separated by using constellations in Stokes space. In Fig. 3(c), as phase noise and frequency offset are encountered, PS-16QAM (3 bit/symbol) and PS-32QAM (3 bit/symbol) are too close, but they can also be separated in Stokes space. In Fig. 4 and Fig. 5, images generated from signals in the Jones constellation plane and Stokes space of all modulation formats considered in this paper are listed.

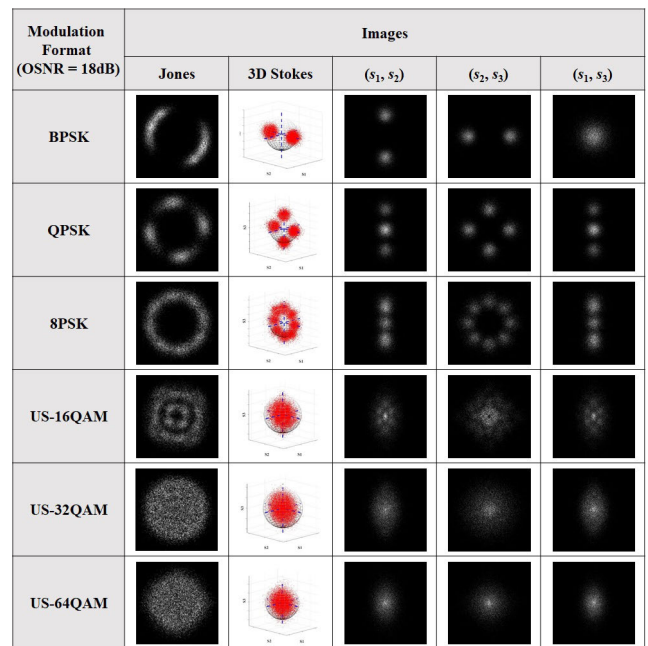
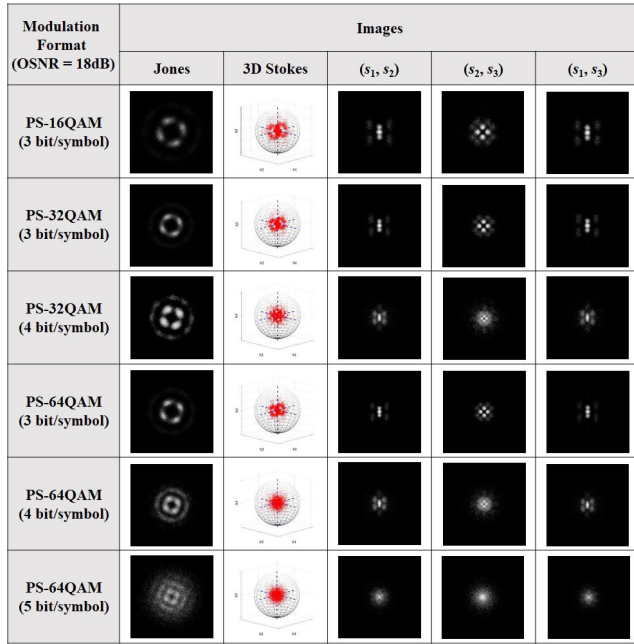


FIGURE 4. Images of 20,000 symbols of BPSK, QPSK, 8PSK, US-16QAM, US-32QAM, US-64QAM signals in 2D Stokes planes with their corresponding 3D Stokes space constellations. Images in the first column are Jones constellations of the corresponding signals. The laser linewidth is 100 kHz.

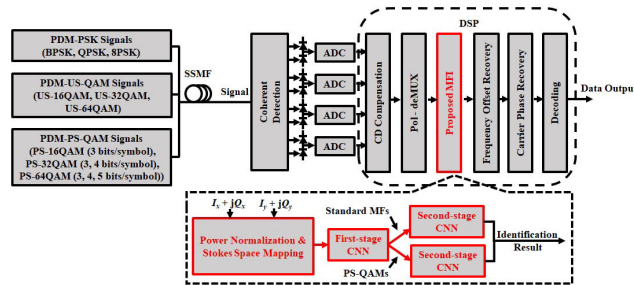
### C. THE PROPOSED SCHEME

The proposed CNN-MFI scheme is designed for a PDM coherent system, which is shown in Fig. 6. Behind the modulation format-independent chromatic dispersion (CD) compensation and Pol-deMUX, the proposed CNN-MFI scheme is processed. After the modulation format is determined successfully, other modulation format-dependent equalization algorithms can be carried out.

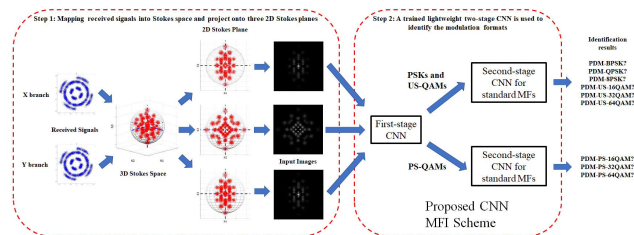
To implement the proposed CNN-MFI scheme, there are two main steps (see Fig. 7). In the first step, received signals are mapped into Stokes space, and images are generated by projecting 3D Stokes onto three 2D Stokes planes, and then in the second step, modulation formats are identified by a two-stage trained neural network in real-time.



**FIGURE 5.** Images of 20,000 symbols of PS-16QAM (3 bit/symbol), PS-32QAM (3, 4 bit/symbol), PS-64QAM (3, 4, 5 bit/symbol) signals in 2D Stokes planes with their corresponding 3D Stokes space constellations. Images in the first column are Jones constellations of the corresponding signals. The laser linewidth is 100 kHz.



**FIGURE 6.** The architecture of the proposed scheme. After modulation format-independent processes, the proposed MFI scheme is processed, and then all modulation format-dependent processes can be implemented after the MFI process.



**FIGURE 7.** The steps of the proposed MFI scheme (from left to right): step one, all received signals are mapped into Stokes space and then projected onto three 2D Stokes planes; step two, a trained lightweight two-stage deep neural network MobileNet V2 is used to identify the modulation formats.

The first-stage CNN is used to distinguish whether the signal belongs to a standard modulation format or is a PS-QAM signal. Then, the second-stage CNN obtains the specific

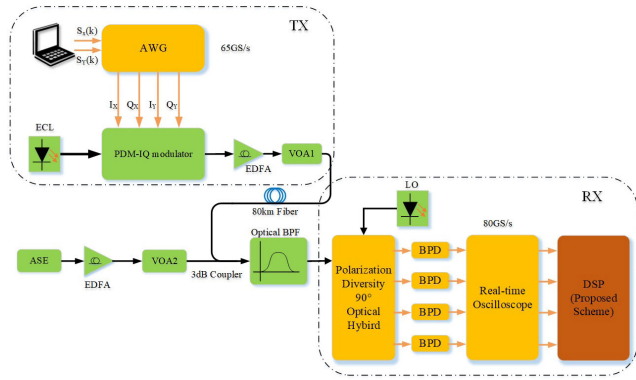
modulation format. The lightweight deep neural network MobileNet V2 [25], [26] is used as our CNN scheme, which helps to improve the efficiency of our deep neural network. The basic unit of a MobileNet is called a depthwise separable convolution, which mainly separates a standard convolution into a depthwise convolution and a pointwise convolution. The new convolution layer is called a bottleneck, which can reduce the computational cost 8 or 9 times less than a standard layer. Each stage is an independent MobileNet V2 with 21 layers: 17 residual bottleneck layers, 3 convolutional layers, and 1 average pooling layer. The two-stage CNNs are used here because if we use one-stage CNNs, it will be hard for the CNN to learn all the MFs well in the case of so many formats. However, if we first distinguish the standard MFs and PS-QAMs, it will be easier for the second-stage CNN to determine specific modulation formats as they just learn standard MFs or PS-QAMs.

**III. EXPERIMENTAL SETUP**

To verify the feasibility of the proposed scheme, a 28 Gbaud coherent experiment platform is shown in Fig. 8. The 28 Gbaud signals of 12 modulation formats were generated by a 65 GS/s arbitrary wave generator (AWG) with an offline MATLAB program [15]. At the transmitter, the external cavity lasers (ECLs) with a linewidth of 100 kHz were utilized to produce light at the central wavelength of 1550 nm and were modulated by an integrated polarization-multiplexing I/Q modulator. Then, the transmitted signals were amplified by an EDFA and sent into a span of an 80 km standard single-mode fiber (SSMF). The loss of fiber was 0.2 dB/km, the chromatic dispersion parameter was 16.9 ps/(nm·km) and the nonlinear coefficient was 1.27 km<sup>-1</sup>·W<sup>-1</sup>. Additionally, an ASE noise source and a variable optical attenuator (VOA) were coupled to change the OSNR condition of the fiber link. At the receiver side, a bandwidth optical bandpass filter (OBPF) of approximately 33 GHz was used to filter out the out-band noise, and the filtered signals were fed into a 42 GHz electrical bandwidth polarization diversity coherent receiver. The local oscillator (LO) was a 100 kHz bandwidth. After balanced photoelectric detection (BPD) and 80 GS/s analog-to-digital sampling of a real-time oscilloscope, electrical signals were processed by an offline DSP module, where our proposed CNN-MFI was embedded. Inside the DSP module, the CD compensation and Pol-deMUX [27] were performed first, then signals were mapped into Stokes space, and images were generated for three 2D Stokes planes. These three images are called an image combination, which is used as input of our three-channel CNN.

The modulation format was identified by feeding the generated images to the proposed MFI scheme, and then modulation format-dependent equalizations could be carried out.

We did not adopt data augmentation methods, including image flip, and translation since the generated images are not images of natural color scenes. To train the first-stage network, we used 3,240 image combinations in the training process (each modulation format has 270 image combinations),



**FIGURE 8.** Experimental setup of a 28 Gbaud coherent platform. The signals of twelve modulation formats are generated by an AWG. The OSNR of the fiber link is simulated by an ASE noise source and VOA. After balanced photoelectric detection, samples are fed into the DSP where the proposed MFI scheme is performed.

and for the second-stage network, 64,800 image combinations were used (each modulation format had 5,400 image combinations). OSNR ranged from 9 dB to 35 dB for these two-stage training images. The size of the input images was  $224 \times 224$  pixels and had 256-level grayscale, so the largest memory size of an image was 49 kb. Additionally, only one generated image needs to be stored in the memory once a time; it is small enough to implement the algorithm in a DSP. The training process of the CNN can be performed offline. To generate images, 20,000 symbols were used, and 100 epochs were conducted to converge the neural network. We selected  $10^{-5}$  as the initial learning rate, and the Adam algorithm was used to adaptively adjust the learning rate during the training process.

In the test stage, 32,400 image combinations for 12 modulation formats (OSNR from 9 dB to 35 dB) were used to test the accuracy of the trained model.

#### IV. RESULTS AND DISCUSSION

Fig. 9 shows the result of our MFI system. Fig. 9 (a) shows that the first-stage CNN can easily distinguish the standard MFs and PS-QAMs at all OSNRs. After the second-stage CNN, the proposed system can recognize all 12 modulation formats at relatively low OSNRs. In our opinion, the criteria of the evaluation for whether the performance of our system was good enough that the identification accuracy for each modulation format was higher than 95%. Fig. 11 shows that the minimum value of OSNR when the criteria can be reached was approximately 15 dB for standard MFs and PS-QAMs, so we give the identification result for OSNR range from 15 dB to 35 dB in Fig. 9 (b) and Fig. 9 (c). Theoretical OSNR values under the FEC threshold of the standard 28 GBaud 16QAM, 32QAM, and 64QAM are approximately 18 dB, 21 dB, and 24 dB, respectively. The proposed scheme reduced the required OSNR value to 16 dB for all of the QAM modulation formats. For PSK signals, all PSK modulation formats were identified precisely when the OSNR value was

		Identified Modulation Format		Accuracy
		Six standard MFs	Six PS-QAMs	
Actual Modulation Format	Six standard MFs	16200 50%	0 00.00%	16200 100%
	Six PS-QAMs	0 00.00%	16200 50%	1620 100%
Accuracy		16200 100%	1620 100%	32400 100%

(a)

		Identified Modulation Format						Accuracy
		BPSK	QPSK	8PSK	US-16QAM	US-32QAM	US-64QAM	
Actual Modulation Format	BPSK	2100 16.67%	0 00.00%	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 100%
	QPSK	0 00.00%	2100 16.67%	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 100%
	8PSK	0 00.00%	0 00.00%	2100 16.67%	0 00.00%	0 00.00%	0 00.00%	2100 100%
	US-16QAM	0 00.00%	0 00.00%	0 00.00%	2094 16.62%	0 00.00%	6 00.05%	2100 99.71%
	US-32QAM	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 16.67%	0 00.00%	2100 100%
	US-64QAM	0 00.00%	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 16.67%	2100 100%
Accuracy		2100 100%	2100 100%	2100 100%	2094 100%	2100 100%	2106 99.72%	12600 99.95%

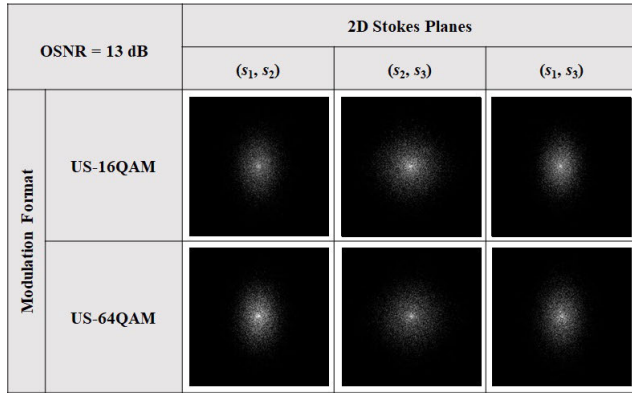
(b)

		Identified Modulation Format						Accuracy
		PS-16QAM (3 bit/s)	PS-32QAM (3 bit/s)	PS-64QAM (3 bit/s)	PS-32QAM (4 bit/s)	PS-64QAM (4 bit/s)	PS-64QAM (5 bit/s)	
Actual Modulation Format	PS-16QAM (3 bit/s)	2096 16.63%	0 00.00%	4 00.04%	0 00.00%	0 00.00%	0 00.00%	2100 99.81%
	PS-32QAM (3 bit/s)	0 00.00%	2099 16.65%	2 00.02%	0 00.00%	0 00.00%	0 00.00%	2100 99.90%
	PS-64QAM (3 bit/s)	0 00.00%	1 00.01%	2099 16.66%	0 00.00%	0 00.00%	0 00.00%	2100 99.95%
	PS-32QAM (4 bit/s)	0 00.00%	0 00.00%	0 00.00%	2099 16.66%	1 00.01%	0 00.00%	2100 99.95%
	PS-64QAM (4 bit/s)	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 16.67%	0 00.00%	2100 100%
	PS-64QAM (5 bit/s)	0 00.00%	0 00.00%	0 00.00%	0 00.00%	0 00.00%	2100 16.67%	2100 100%
Accuracy		2096 100%	2099 99.95%	2105 99.71%	2099 100%	2101 99.95%	2100 100%	12600 99.94%

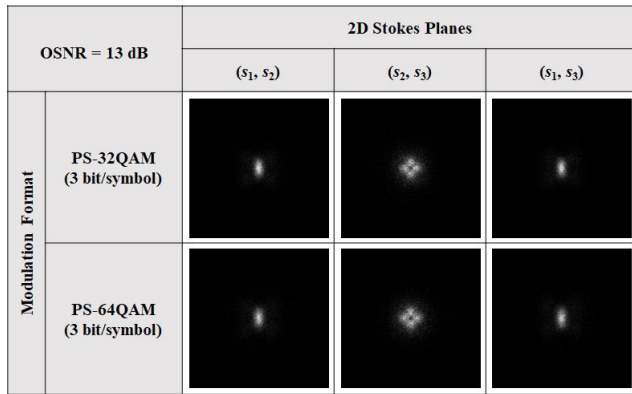
(c)

**FIGURE 9.** The confusion matrix of the proposed MFI scheme. (a) Number of test image combinations and accuracy matrix for first-stage CNN. All images can be identified precisely. (b) Number of test image combinations and accuracy matrix for second-stage CNN of standard MFs. Each modulation format contains 2,100 test image combinations (OSNR from 15 dB to 35 dB) of six standard MFs, including three PSK signals and three US-QAM signals. Six image combinations of US-16QAM are misclassified as US-64QAM. (c) Number of test image combinations and accuracy matrix for second-stage CNN of PS-QAMs. Each modulation format contains 2,100 test image combinations (OSNR from 15 dB to 35 dB) of six PS-QAM signals. Four image combinations of 3 bit/symbol PS-16QAM are misclassified as 3 bit/symbol PS-64QAM, and two image combinations of 3 bit/symbol PS-32QAM are misclassified as 3 bit/symbol PS-64QAM. One image combination of 3 bit/symbol PS-64QAM is also misclassified as 3 bit/symbol PS-32QAM.

higher than 9 dB, which was lower than the FEC threshold of 28 GBaud BPSK (9.73 dB). According to [28]–[30], the required OSNRs under the FEC threshold for PS-16QAM, PS-32QAM, PS-64QAM are approximately 17 dB, 19.7 dB and 22.5 dB, respectively, which are also lower than the



(a)



(b)

FIGURE 10. Images of 16QAM and 64QAM in three different 2D Stokes planes at OSNR 13 dB. (a) US-16QAM and US-64QAM. (b) PS-32QAM (3 bit/symbol) and PS-64QAM (3 bit/symbol).

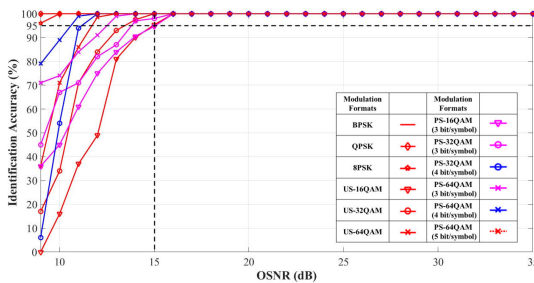


FIGURE 11. Identification accuracy at different OSNRs (9 dB to 35 dB) with image size 224 x 224 pixels.

required OSNR value 16 dB to recognize all PS-QAM signals precisely.

It can be seen in Fig. 9 (b) that six image combinations of US-16QAM were misclassified as US-64QAM. The reason is that the visual appearance of the two formats were very similar to each other when the OSNR was too low, e.g. 13 dB, and it is very hard to distinguish them from each other (see Fig. 10 (a)). For similar reasons, for PS-QAM signals, four image combinations of 3 bit/symbol PS-16QAM were misclassified as 3 bit/symbol PS-64QAM and two image combinations of 3 bit/symbol PS-32QAM

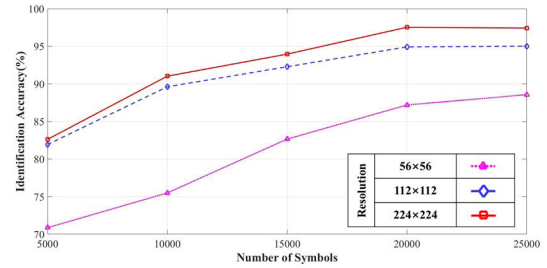


FIGURE 12. Identification accuracy for input images of different numbers of symbols from 5,000 to 20,000 at 5,000 with different resolutions.

were misclassified as 3 bit/symbol PS-64QAM and one image combination of 3 bit/symbol PS-64QAM was also misclassified as 3 bit/symbol PS-32QAM (see Fig. 10 (b) for example), which are shown in Fig. 9 (c). However, as OSNR was higher than 16 dB, the CNN-MFI scheme identified all twelve modulation formats precisely (Fig. 11).

As in Fig. 11, the proposed MFI scheme reached a precise identification accuracy at a very low OSNR requirement compared to conventional Stokes space cluster algorithms described in [12] and [31]. US- and PS-mQAM signals were successfully identified when OSNR was higher than 14 dB; for PSK signals only, all PSK modulation formats were recognized when OSNR was as low as 9 dB.

Meanwhile, the proposed MFI scheme also runs quickly and requires very small storage memory. In a 28 Gbaud system, the time to achieve 20,000 samples was approximately 0.174  $\mu$ s, and it took approximately 5 ms to determine the modulation format. The hardware platform was an Intel Xeon E5 v4, frequency 1.7 GHz, 8 cores, and a GTX TITAN Xp GPU card, and 12 GB memory.

In Fig. 12, the effects of sizes of images and numbers of symbols in each combination of images is considered. Three sizes of images are shown: 56 x 56 pixels, 112 x 112 pixels, and 224 x 224 pixels. The range of numbers of symbols in each combination of images are from 5,000 to 25,000 at 5,000. For the same number of symbols in each combination of images, the identification accuracy increases as the size of the images increases. However, a larger size of images implies that more computational resources are needed. Here, we take images of 224 x 224 pixels as a reasonable size, which is a tradeoff of efficiency and complexity. Additionally, using a specific size of images in a combination of images increases the identification accuracy as the number of symbols increases. Other than typical MFI schemes, such as schemes proposed in [12] and [31], the complexity and requirement of memory of the proposed MFI scheme does not change considerably as the number of symbols increases because they are used to generate images only. From Fig. 12, the identification accuracy does not improve considerably as the number of symbols becomes larger than 20,000. Therefore, we use 20,000 symbols in the proposed MFI scheme.

In Fig. 13, we investigated the change in identification accuracy with the increasing number of epochs as we trained

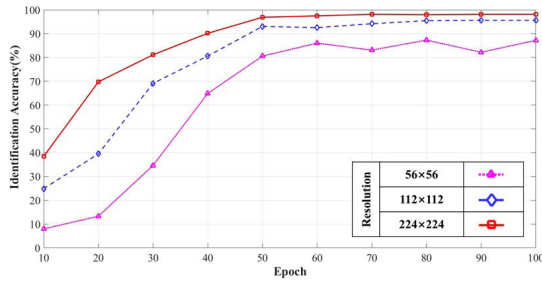


FIGURE 13. The total accuracy at different epochs for input images with different resolutions of 56 × 56 pixels, 112 × 112 pixels, 224 × 224 pixels.

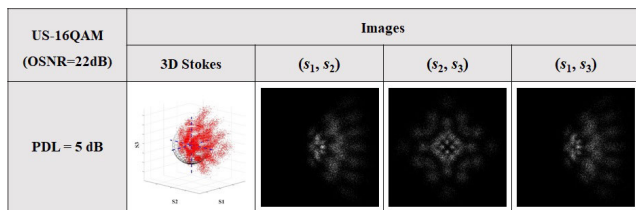


FIGURE 14. Images of constellations in 3D Stokes space and three 2D Stokes planes of US-16QAM signal with 5 dB PDL effect.

the CNN. The conclusion is that the accuracy achieves a stable value after 50 or 60 epochs for three sizes of images with 20,000 symbols. Therefore, it is enough to train the CNN with 60 epochs, and the accuracy can achieve a reliable value.

Although CD can be assumed to be compensated precisely [32] and polarization demultiplexing can be performed before the MFI scheme, in practice, the influence of impairments such as polarization-dependent loss (PDL), residual CD, and polarization mode dispersion (PMD) in optical transmission systems still exist and should be considered. In the following, the tolerances of the proposed MFI scheme with respect to these impairments are considered. Two uniformly shape modulation formats (US-16QAM, US-64QAM) and two probabilistically shaped modulation formats (4 bit/symbol PS-32QAM, 4 bit/symbol PS-64QAM) at OSNR 22 dB are discussed as examples. All impairments are simulated using the offline MATLAB program.

To the effect of PDL, all the Stokes constellation points in 3D Stokes space will linearly shift in the same direction but in the 2D Stokes plane (s<sub>2</sub>, s<sub>3</sub>), it will not change (see Fig. 14). Therefore, benefiting from our three channels input, the scheme can accurately identify signals with PDL effect. PDL is set in the range from 0 dB to 10 dB. In Fig. 15 (a), the proposed MFI scheme can tolerate PDL until 6.5 dB for US-16QAM, 6.5 dB for US-64QAM, 4.5 dB for PS-32QAM (4 bit/symbol) and 5 dB for PS-64QAM (4 bit/symbol), respectively.

For the residual CD (-300 – 300 ps/nm), the results are shown in Fig. 15 (b). The tolerances of the proposed MFI scheme of the four modulation formats are -170 – 160 ps/nm, -200 – 210 ps/nm, -220 – 230 ps/nm and -240 – 250 ps/nm, respectively.

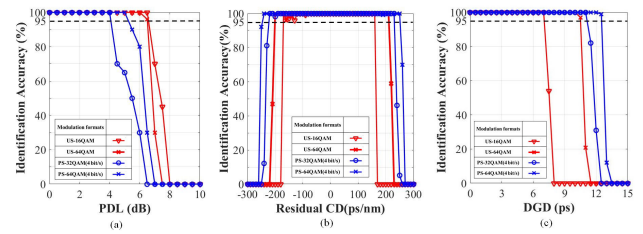


FIGURE 15. Tolerances of the proposed MFI scheme of four modulation formats (US-16QAM, US-64QAM, 4 bit/symbol PS-32QAM and PS-64QAM) under three following effects: (a) PDL values from 0 dB to 10 dB. (b) Residual CD values from -300 to 300 ps/nm. (c) DGD values from 0 ps to 15 ps.

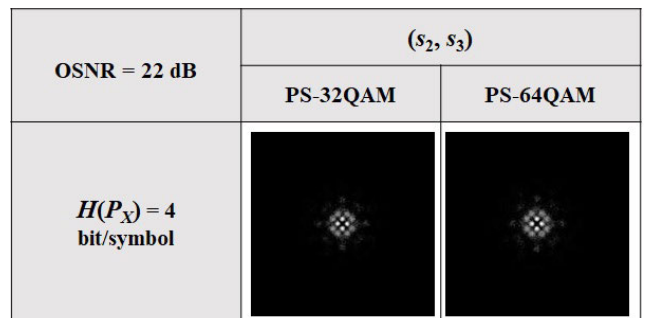


FIGURE 16. Images of constellations in 2D Stokes plane (s<sub>2</sub>, s<sub>3</sub>) of 4 bit/symbol PS-32QAM and PS-64QAM signal.

To the effect of PMD, the DGD is set in a range from 0 to 15 ps. In addition, the tolerances of the CNN-MFI system for the four modulation formats are 7 ps, 10.5 ps, 11 ps and 12.5 ps, respectively (see Fig. 15 (c)).

From Fig. 15, the proposed MFI scheme can tolerate residual CD, PDL and PMD effects in a wide range. It is interesting to see that for residual CD and DGD, PS-QAM signals perform better than US-QAM signals. The reason is as follows: the inner part of constellations of PS-QAM signals is more likely to appear than the outer part, which makes signals more powerful against the residual CD and DGD effects. However, for the PDL effect, PS-QAM signals perform worse than US-QAM signals. This is because the images of PS-QAM in the 2D Stokes plane (s<sub>2</sub>, s<sub>3</sub>) are too similar to each other with the same low H(P<sub>X</sub>) (see Fig. 16). Additionally, we notice that higher-order modulation formats often have better tolerance than the lower order modulation formats since lower order modulation formats tend to be misclassified as higher-order modulation formats. However, this does not influence the identification accuracy of US-64QAM or 4 bit/symbol PS-64QAM, as they are the highest order modulation format for US- and 4 bit/symbol PS-mQAM signals.

V. CONCLUSION

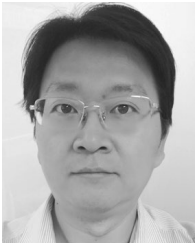
For twelve modulation formats, BPSK, QPSK, 8PSK, US-16QAM, US-32QAM, US-64QAM, PS-16QAM (3 bit/symbol), PS-32QAM (3 and 4 bit/symbol), PS-64QAM (3, 4, 5 bit/symbol), a deep learning-based modulation

identification scheme is proposed. By mapping all received signals into three 2D Stokes planes, images are generated that are fed into the two-stage CNN-MFI scheme. The results show that the proposed CNN-MFI scheme can achieve a relatively high identification accuracy (>95%) with relatively low OSNR burden (>15 dB). Additionally, the tolerances of the proposed scheme with respect to the effects of residual CD, PDL and PMD are also discussed. The results show that the scheme has a wide range of tolerances to these effects.

## REFERENCES

- [1] Q. Zhuge, M. Morsy-Osman, X. Xu, M. Chagnon, M. Qiu, and D. V. Plant, "Spectral efficiency-adaptive optical transmission using time domain hybrid QAM for agile optical networks," *J. Lightw. Technol.*, vol. 31, no. 15, pp. 2621–2628, Aug. 1, 2013. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-31-15-2621>
- [2] A. Nag, M. Tornatore, and B. Mukherjee, "Optical network design with mixed line rates and multiple modulation formats," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 466–475, Feb. 15, 2010, doi: [10.1109/jlt.2009.2034396](https://doi.org/10.1109/jlt.2009.2034396).
- [3] S. M. Bilal, G. Bosco, Z. Dong, A. P. T. Lau, and C. Lu, "Blind modulation format identification for digital coherent receivers," *Opt. Express*, vol. 23, no. 20, pp. 26769–26778, Oct. 2015, doi: [10.1364/oe.23.026769](https://doi.org/10.1364/oe.23.026769).
- [4] L. Jiang, L. Yan, A. Yi, Y. Pan, T. Bo, M. Hao, W. Pan, and B. Luo, "Robust and blind modulation format identification for elastic optical networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2018, doi: [10.1109/ecoc.2018.8535367](https://doi.org/10.1109/ecoc.2018.8535367).
- [5] M. Xiang, Q. Zhuge, M. Qiu, X. Zhou, M. Tang, D. Liu, S. Fu, and D. V. Plant, "RF-pilot aided modulation format identification for hitless coherent transceiver," *Opt. Express*, vol. 25, no. 1, p. 463, Jan. 2017, doi: [10.1364/oe.25.000463](https://doi.org/10.1364/oe.25.000463).
- [6] A. Nandi and E. Azzouz, "Automatic analogue modulation recognition," *Signal Process.*, vol. 46, no. 2, pp. 211–222, Oct. 1995, doi: [10.1016/0165-1684\(95\)00083-p](https://doi.org/10.1016/0165-1684(95)00083-p).
- [7] R. Boada, R. Borkowski, and I. T. Monroy, "Clustering algorithms for Stokes space modulation format recognition," *Opt. Express*, vol. 23, no. 12, pp. 15521–15531, Jun. 2015, doi: [10.1364/oe.23.015521](https://doi.org/10.1364/oe.23.015521).
- [8] L. Cheng, L. Xi, D. Zhao, X. Tang, W. Zhang, and X. Zhang, "Improved modulation format identification based on Stokes parameters using combination of fuzzy C-means and hierarchical clustering in coherent optical communication system," *Chin. Opt. Lett.*, vol. 13, no. 10, Oct. 2015, Art. no. 100604. [Online]. Available: <http://col.osa.org/abstract.cfm?URI=col-13-10-100604>
- [9] P. Isautier, A. Stark, K. Mehta, R. de Salvo, and S. E. Ralph, "Autonomous software-defined coherent optical receivers," in *Proc. Opt. Fiber Commun. Conf./Nat. Fiber Optic Eng. Conf.*, Anaheim, CA, USA, Mar. 2013, Paper OTh3B.4. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2013-OTh3B.4>, doi: [10.1364/OFC.2013.OTh3B.4](https://doi.org/10.1364/OFC.2013.OTh3B.4).
- [10] P. Isautier, J. Pan, R. Desalvo, and S. E. Ralph, "Stokes space-based modulation format recognition for autonomous optical receivers," *J. Lightw. Technol.*, vol. 33, no. 24, pp. 5157–5163, Dec. 15, 2015. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-33-24-5157>
- [11] J. Liu, Z. Dong, K. Zhong, A. P. T. Lau, C. Lu, and Y. Lu, "Modulation format identification based on received signal power distributions for digital coherent receivers," in *Proc. Opt. Fiber Commun. Conf., San Francisco, CA, USA, Mar. 2014, Paper Th4D.3*. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2014-Th4D.3>, doi: [10.1364/OFC.2014.Th4D.3](https://doi.org/10.1364/OFC.2014.Th4D.3).
- [12] L. Jiang, L. Yan, A. Yi, Y. Pan, T. Bo, M. Hao, W. Pan, and B. Luo, "Blind density-peak-based modulation format identification for elastic optical networks," *J. Lightw. Technol.*, vol. 36, no. 14, pp. 2850–2858, Jul. 15, 2018. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-36-14-2850>
- [13] L. Jiang, L. Yan, A. Yi, Y. Pan, M. Hao, W. Pan, and B. Luo, "An effective modulation format identification based on intensity profile features for digital coherent receivers," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5067–5075, Oct. 1, 2019, doi: [10.1109/jlt.2019.2927748](https://doi.org/10.1109/jlt.2019.2927748).
- [14] X. Lin, Y. A. Eldemerdash, O. A. Dobre, S. Zhang, and C. Li, "Modulation classification using received signal's amplitude distribution for coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 29, no. 21, pp. 1872–1875, Nov. 1, 2017, doi: [10.1109/LPT.2017.2754501](https://doi.org/10.1109/LPT.2017.2754501).
- [15] J. Shi, J. Zhang, X. Li, N. Chi, Y. Zhang, Q. Zhang, and J. Yu, "Improved performance of high-order QAM OFDM based on probabilistically shaping in the datacom," in *Proc. Opt. Fiber Commun. Conf. Exposit. (OFC)*, Mar. 2018, pp. 1–3.
- [16] A. S. Kashi, A. I. A. El-Rahman, J. C. Cartledge, and S. A. Etemad, "Extending a nonlinear SNR estimator to include shaping distribution identification for probabilistically shaped 64-QAM signals," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3252–3260, Jul. 1, 2019. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-37-13-3252>
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [18] F. N. Khan, K. Zhong, W. H. Al-Arashi, C. Yu, C. Lu, and A. P. T. Lau, "Modulation format identification in coherent receivers using deep machine learning," *IEEE Photon. Technol. Lett.*, vol. 28, no. 17, pp. 1886–1889, Sep. 2016, doi: [10.1109/lpt.2016.2574800](https://doi.org/10.1109/lpt.2016.2574800).
- [19] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express*, vol. 25, no. 15, Jul. 2017, Art. no. 17150, doi: [10.1364/oe.25.017150](https://doi.org/10.1364/oe.25.017150).
- [20] W. Zhang, D. Zhu, Z. He, N. Zhang, X. Zhang, H. Zhang, and Y. Li, "Identifying modulation formats through 2D Stokes planes with deep neural networks," *Opt. Express*, vol. 26, no. 18, Sep. 2018, Art. no. 23507, doi: [10.1364/oe.26.023507](https://doi.org/10.1364/oe.26.023507).
- [21] G. Bocherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015, doi: [10.1109/tcomm.2015.2494016](https://doi.org/10.1109/tcomm.2015.2494016).
- [22] J. Cho, X. Chen, S. Chandrasekhar, and P. Winzer, "On line rates, information rates, and spectral efficiencies in probabilistically shaped QAM systems," *Opt. Express*, vol. 26, no. 8, p. 9784, Apr. 2018, doi: [10.1364/oe.26.009784](https://doi.org/10.1364/oe.26.009784).
- [23] P. Schulte and G. Bocherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016, doi: [10.1109/tit.2015.2499181](https://doi.org/10.1109/tit.2015.2499181).
- [24] Z. Qu, S. Zhang, and I. B. Djordjevic, "Universal hybrid probabilistic-geometric shaping based on two-dimensional distribution matchers," in *Proc. Opt. Fiber Commun. Conf., San Diego, CA, USA, Mar. 2018, Paper M4E.4*. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2018-M4E.4>, doi: [10.1364/OFC.2018.M4E.4](https://doi.org/10.1364/OFC.2018.M4E.4).
- [25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation," 2018, *arXiv:1801.04381*. [Online]. Available: <https://arxiv.org/abs/1801.04381>
- [26] M. Z. Andrew G. Howard, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: <https://arxiv.org/abs/1704.04861>
- [27] N. J. Muga and A. N. Pinto, "Adaptive 3-D Stokes space-based polarization demultiplexing algorithm," *J. Lightw. Technol.*, vol. 32, no. 19, pp. 3290–3298, Oct. 2014. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-32-19-3290>
- [28] J. Cho, L. Schmalen, and P. J. Winzer, "Normalized generalized mutual information as a forward error correction threshold for probabilistically shaped QAM," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3, doi: [10.1109/ECOC.2017.8345872](https://doi.org/10.1109/ECOC.2017.8345872).
- [29] S. Okamoto, F. Hamaoka, M. Nakamura, and Y. Kisaka, "Laser phase noise tolerance of probabilistically-shaped constellations," in *Proc. Opt. Fiber Commun. Conf., San Diego, CA, USA, Mar. 2018, Paper W2A.51*. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2018-W2A.51>, doi: [10.1364/OFC.2018.W2A.51](https://doi.org/10.1364/OFC.2018.W2A.51).
- [30] D. Pileri, L. Bertignono, A. Nespola, F. Forghieri, and G. Bosco, "Comparison of probabilistically shaped 64qam with lower cardinality uniform constellations in long-haul optical systems," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 501–509, Jan. 15, 2018. [Online]. Available: <http://jlt.osa.org/abstract.cfm?URI=jlt-36-2-501>
- [31] X. Mai, J. Liu, X. Wu, Q. Zhang, C. Guo, Y. Yang, and Z. Li, "Stokes space modulation format classification based on non-iterative clustering algorithm for coherent optical receivers," *Opt. Express*, vol. 25, no. 3, p. 2038, Feb. 2017, doi: [10.1364/OE.25.002038](https://doi.org/10.1364/OE.25.002038).
- [32] B. Chomyczc, *Planning Fiber Optics Networks*. New York, NY, USA: McGraw-Hill, 2009, ch. 7.





**WENBO ZHANG** was born in Beijing, China, in September 1974. He received the B.S. and M.S. degrees in mathematics from Hebei University, Baoding, Hebei, China, in 1997 and 2000, respectively, and the Ph.D. degree in applied mathematics from Nankai University, Tianjin, China, in 2003.

Since 2006, he has been an Associate Professor with the Department of Mathematics, School of Science, Beijing University of Posts and Telecommunication, Beijing, China. He is the author or coauthor of more than 80 international journals (including IEEE and OSA journals) and peer-reviewed conference papers, and holds three China patents.

Dr. Zhang was a recipient of the Science and Technology Award from Ministry of Education of the people's Republic of China, in 2015, for his work Monitoring, Equalization and Adaptive Compensation for Polarization Mode Dispersion in High Speed Optical Communication Systems.



**DINGCHENG ZHU** was born in Suzhou, Jiangsu, China, in 1995. He received the B.S. degree from the Nanjing University of Posts and Telecommunications, in 2017. He is currently pursuing the M.S. degree with the School of Science, Beijing University of Posts and Telecommunications. His research interests include the parameter identifications like modulation format identification of elastic optical networking with artificial intelligence technology and probabilistic constellation shaping of fiber transmission.



**NANNAN ZHANG** was born in Shijiazhuang, Hebei, China, in 1993. She received the B.S. degree in communications engineering from Hebei University, in 2016. Since 2006, she has been studying with the State Key Laboratory of Information Photonics and Optical Communication, Beijing University of Posts and Telecommunications. Her research interests include linear impairments in coherent optical communication systems, including polarization impairments, carrier impairments and Kalman filters, and constellation probabilistic shaping.



**HENGYING XU** was born in Liaocheng, Shandong, China, in 1981. He received the M.S. degree from the Beijing University of Posts and Telecommunications and the Ph.D. degree from the Harbin University of Science and Technology, in 2018 and 2006, respectively. He is currently an Associate Professor with the Shandong Provincial Key Laboratory of Optical Communication Science and Technology, Liaocheng University. He has published more than 20 related articles. His research

interests include the optical performance monitoring, parameter identifications of elastic optical networking, and coherent receiving of fiber transmission. He is a member of Optical Society of America and a Senior Member of China Communications Society.



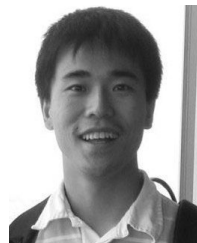
**XIAOGUANG ZHANG** (Senior Member, IEEE) was born in Beijing, China, in June 1961. He received the B.S. and M.S. degrees in physics from Peking University, Beijing, in 1985 and 1988, respectively, and the Ph.D. degree in physical electronics from the Beijing University of Posts and Telecommunication, Beijing, in 2004.

From 1988 to 2008, he was a Professor in physics with the Department of Physics, School of Science, Beijing University of Posts and Telecommunications, where he is currently a Professor with the Institute of Information Photonics and Optical Communications. He is the author or coauthor of more than 200 international journals (including IEEE and OSA journals) and peer-reviewed conference papers (including OFC, ECOC, CLEO, and APOC), and holds more than 20 China patents. His research interests include polarization effects in fibers, multilevel phase modulation formats, recovery of optical signal distortions, and generation of optical frequency comb.

Dr. Zhang is a member of OSA and a Senior Member of Chinese Optics Society, the China Institute of Communications, and the China Institute of Electronics. He is also a Syndic of Beijing Optics Society (BOS). He was a recipient of the Science and Technology Progress Award from the China Ministry of Information Industry, in 1999, for his work on the project The Experimental and Theoretical Research on Optical Soliton Transmissions, the Science and Technology Award from the China Institute of Communications, in 2005, for his work Adaptive Compensation for Polarization Mode Dispersion in High Speed Optical Communication Systems, and the Science and Technology Award from Ministry of Education of the people's Republic of China, in 2015, for his work Monitoring, Equalization and Adaptive Compensation for Polarization Mode Dispersion in High Speed Optical Communication Systems.



**HU ZHANG** received the master's degree in optics from Henan Normal University, Xinxiang, China, in 2002, and the Ph.D. degree in electromagnetic field and microwave technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2009. He is currently a Teacher with the School of Optoelectronic Information, Beijing University of Posts and Telecommunications. His research interests include special optical fiber, fiber devices, and orbital angular momentum.



**YONG LI** received the M.S. degree in applied mathematics, M.S.E.E., and Ph.D. degree in electrical engineering from the University of Notre Dame. He is currently an Associate Professor with the School of Electronic Engineering, Beijing University of Posts and Telecommunications. His research interests include computer vision, multispectral images, deep learning, and differential geometry.

...