Stokes Space-Based Modulation Format Recognition for Autonomous Optical Receivers

Pierre Isautier, Jie Pan, Richard DeSalvo, and Stephen E. Ralph

***Abstract*—We propose a new modulation format recognition method for optical links employing OOK, M-PSK, M-PAM, and M-QAM modulation formats. The method combines the Stokes space analysis and higher order statistics as a part of a universal digital coherent receiver. Experimental investigation demonstrates successful format recognition of single-carrier optical signals modulated at 31.5 GBd with OOK, BPSK, QPSK, or 16QAM, and transported through 810 km of large area fiber. These results enable the development of autonomous receivers capable of identifying and demodulating unknown signals from noncooperating transmitters, including legacy optical transmitters.**

***Index Terms*—Coherent communications, dynamic networks, flexible receivers, optical fiber communication.**

I. Introduction

LABORATORY and commercially deployed optical fiber systems with digital coherent receivers have demonstrated record capacity and reach. Due to the continuous improvement in high-speed electronics, advances in fiber optic network optical elements, and the utilization of advanced signaling techniques [1], optical communication systems have reached a state of greater signal waveform flexibility and higher single channel data rates that now exceed 200 Gb/s [2]. As a result, the next generation of fiber optic communication systems will offer a larger degree of dynamic adaptability in reach, data rate, and spectral occupancy [3], thus shifting away from legacy fixed single purpose transceivers.

These advances in hardware and signal processing capability have enabled increased levels of heterogeneity allowing a wide range of services and transmission technologies. A key requirement of these new capabilities is the efficient provisioning and management of all available network resources. Although an omniscient supervisory control layer may ensure coordination between all elements of the network enabling dynamic adaptation to current traffic needs using available resources, there are clear benefits from deploying autonomous network elements that require no intervention from a network operator or a supervisory control layer.

The signal transmission parameters (such as the carrier frequency, the signal bandwidth, the symbol rate, the states of polarization, and the modulation format) are typically communicated to the receiver through the supervisory control layer, however there are situations where network elements are not properly synchronized. For example, these events occur when the monitoring elements have failed, or when the supervisory channel is disrupted or deemed too slow to allow fast and flexible provisioning.

Thus, it becomes necessary in some situations, to deploy optical receivers that operate independently of the supervisory control layer by blindly retrieving the transmission parameters of an incoming signal and optimally demodulate it. Such receivers will have the capability to identify, without prior knowledge and with minor processing delay, signal transmission parameters.

The concept of autonomous optical receivers [4] consists of a universal intelligent receiver that does not require manual configuration or reconfiguration, and which can identify and demodulate nearly any signal, including legacy optical signals not requiring coherent detection, and modern high-spectral efficiency optical signals that require coherent detection technology.

Autonomous receivers enable rapid recovery from network failures without the need for slower supervisory control layer coordination. Furthermore, autonomous receivers can be used for system performance monitoring, network diagnostic, and other surveillance applications.

We have previously presented results demonstrating the successful implementation of autonomous receivers [4]–[10], capable of decoding, without prior knowledge, incoming signals after blind identification of the signals transmission parameters, including the modulation format. A first algorithm based exclusively on higher order statistics (HOS) was introduced to perform modulation format recognition (MFR) and showed robust performance over a restricted number of conventional optical modulation formats [4], [5]. An alternate implementation using Stokes space analysis and HOS was introduced and experimentally verified to robustly recognize optical OOK, BPSK, QPSK, and 16QAM [6], [7]. These two MFR algorithms were implemented in distinct autonomous architectures, whose merits were experimentally compared after transmission of single-carrier signals in an optical communication link composed of 810 km of large area fiber (LAF) [8]. A third algorithm based on the statistical distribution of the incoming signal's radius was proposed to identify optical time-domain hybrid modulation formats and experimentally validated with single-carrier

Manuscript received August 7, 2015; revised October 7, 2015; accepted October 16, 2015. Date of publication October 18, 2015; date of current version November 20, 2015. This work was supported in part by Harris Corporation and the Georgia Institute of Technology Terabit Optical Networking Consortium.

P. Isautier and S. E. Ralph are with the Georgia Institute of Technology, Atlanta GA 30318 USA (e-mail: pierre.isautier@gatech.edu; stephen.ralph@ece.gatech.edu).

J. Pan is with the ADVA Optical Networking, Norcross, GA 30092 USA (e-mail: jpan@gatech.edu).

R. DeSalvo is with the Harris Corporation, Melbourne, FL 32919 USA (e-mail: rdesalvo@harris.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2015.2492858

signals received after transmission in an optical communication link comprising 810 km of LAF [9], [10].

Other works have investigated various optical MFR methods. Gonzalez *et al*. introduced a likelihood-based approach that monitors the constellation diagram with the use of k-means, which utilizes probabilistic and hypothesis testing to identify the incoming signal modulation format [11]. Khan *et al*. presented a feature-based method that uses artificial neural networks based on direct detection amplitude histograms (which requires prior training to obtain correct classification results) to extract prominent features from the received signal that are utilized for the identification of signal modulation format [12]. Borkowski *et al*. proposed a method based exclusively on Stokes space analysis that uses statistical signal processing techniques to extract the modulation format signature [13], [14]. OSNR limitations encountered in this method, inherent to Stokes space mapping which distorts the noise probability density function [15], [16], were mitigated by revising the statistical signal processing methods [16]. Liu *et al*. proposed a feature-based approach that uses the statistical distribution of the received signal's power [17]. Tan *et al*. presented a novel technique for simultaneous identification of the bit-rate and the modulation format using principal component analysis (PCA) [18]. Cui *et al*. introduced a feature-based approach that uses an improved method for asynchronous amplitude histogram [19].

Here, we present the details of the MFR algorithm initially reported in [7] and new experimental results. The proposed algorithm uses a hybrid method combining Stokes space analysis with HOS to robustly identify OOK, M-PSK, M-PAM, and M-QAM signals. We further examine and quantify the relationship between received OSNR and key metrics fundamental to the decision process within the MFR algorithm (i.e., Stokes space constellation dimension coefficients, magnitudes of normalized weighted estimates of fourth-order cumulants, and spatial cross-correlation indices).

This article is organized in four sections. In Section II, we describe the proposed MFR method and discuss its placement in the receiver digital signal processing (DSP) architecture. Section III describes in detail the performance of the Stokes space-based MFR method using experimentally captured signals in an optical communication link. Conclusion is made in Section IV.

II. Proposed Stokes Space-Based MFR Method

The proposed MFR method is based on a hybrid method combining Stokes space analysis with HOS. It is designed to identify OOK, M-PSK, M-PAM, and M-QAM signals.

In the receiver DSP architecture, we use conventional chromatic dispersion (CD) equalization and timing recovery algorithms that are independent of the received signal modulation format. Therefore, the proposed MFR module is placed after timing recovery and can be placed prior to polarization demultiplexing (in the case of polarization division multiplexing (PDM) systems), Fig. 1.

The MFR algorithm operates on the digitized X (signal 1) and Y polarizations (signal 2), resampled after timing recovery to 1 sample/symbol at maximum eye opening, of the incoming PDM signal, Fig. 2.

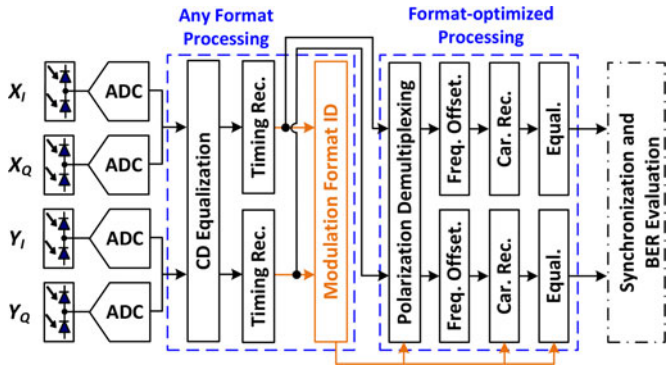

Fig. 1. Receiver DSP architecture with Stokes space-based MFR performed after CD equalization and timing recovery.

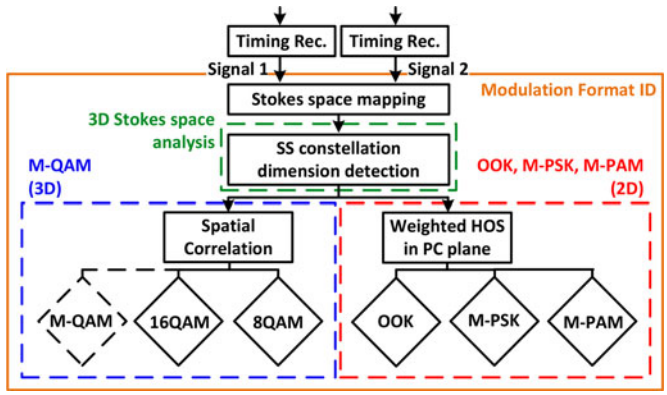

Fig. 2. Stokes space-based MFR method. First, the incoming signal is mapped to Stokes space. Second, the constellation dimension is used to distinguish between {M-QAM} ($M \neq \{2, 4\}$) and {OOK + M-PSK + M-PAM}. Third, {OOK + M-PSK + M-PAM} formats are further distinguished using HOS, whereas {M-QAM} formats are separately distinguished via a spatial cross-correlation method.

In the case of a received single polarization signal, the MFR algorithm operates on an emulated PDM signal. Specifically, the even symbols are used to create a virtual X polarization (signal 1) while the virtual Y polarization contains the odd symbols (signal 2). For optimum performance, the virtual X and Y polarizations signals are truncated over a time period during which the state of polarization (SOP) of the received single polarization signal is fixed.

To determine if the signal is polarization multiplexed, we developed, in previous work [5], the capability to blindly identify the number of multiplexed polarizations. The method is based on a cross correlation between the received X and Y polarizations.

Thus, to identify formats amongst OOK, M-PSK, M-PAM, and M-QAM, the proposed MFR method, Fig. 2, typically uses 10 000 symbols or less, and operates in three consecutive steps: i) Stokes space mapping, ii) estimation of the Stokes space constellation dimension, and iii) format recognition among either {OOK + M-PSK + M-PAM} formats or {M-QAM} formats.

A. Stokes Space Mapping

Stokes parameters fully describe the electric field properties, including total power and SOP. Stokes parameters can be calculated from samples (to include symbols transitions information),

or from symbols of the received signal. Noting $x[k]$ ($y[k]$) the X (Y) polarization received digital signal, the 4-D Stokes vector $\vec{s}$, which contains the Stokes parameters, is determined in the conventional manner:

$$\vec{s}[k] = \begin{pmatrix} s_0[k] \\ s_1[k] \\ s_2[k] \\ s_3[k] \end{pmatrix} = \begin{pmatrix} x[k]^2 + y[k]^2 \\ x[k]^2 - y[k]^2 \\ 2Re\{x\bar{y}\} \\ 2Im\{x\bar{y}\} \end{pmatrix}. \quad (1)$$

The 3-D sub-vector $(s_1\ s_2\ s_3)^t$ is used to visualize the received signal SOP on the Poincaré sphere. Stokes space constellations are the representation of the sub-vector $(s_1\ s_2\ s_3)^t$ in a 3-D space, and are contained in a lens-like surface [20].

The proposed MFR method maps the received signals' symbols to Stokes space, obtaining a Stokes space constellation comprised of clusters corresponding to the nominally discrete set of allowed electric fields.

There are multiple advantages for performing MFR in Stokes space: first, the mapping process is unique to a modulation format whose signature is expanded in a 3-D space, allowing unique identification. Second, the polarization demultiplexing process is bypassed because the Stokes space constellation is independent of the received SOP, since the variation of the received SOP will simply cause a rotation of the Stokes space constellation with time [20]. Finally, the mapping process is independent of the residual carrier-local oscillator (LO) frequency and phase offsets contained in the received signal.

On the other hand, performing MFR in Stokes space presents some challenges. Indeed, Stokes space constellations are sensitive to noise, since Stokes space mapping distorts the noise probability density function [15]. Moreover, Stokes space constellations are perturbed by polarization mode dispersion (PMD), and polarization dependent loss (PDL) [20, 21], which are common within optical fiber systems. Channel nonlinearities also impact Stokes space constellations [8].

Note that in the case of a received single polarization signal, the emulated PDM signal satisfies the first and third properties of Stokes space-based MFR mentioned above. Additionally, the emulation of the PDM signal waives the limitations resulting from PMD or PDL.

B. Stokes Space Constellation Dimension Estimation

In theory, each Stokes space constellation has a plane of symmetry. The study of ideal noiseless Stokes space constellations leads to a natural distinction between 2-D constellations that are wholly contained in a symmetry plane in Stokes space (OOK, M-PSK, M-PAM) and 3-D constellations (M-QAM with $M \neq \{2, 4\}$).

We use this property to distinguish 2-D from 3-D Stokes space constellations by evaluating the absolute distance of Stokes space constellation clusters' mean positions relative to the symmetry plane P.

Hence, the first decision of the proposed MFR algorithm is to detect the dimension of the Stokes space constellation to distinguish if the received signal modulation format is amongst $\{\text{OOK} + \text{M-PSK} + \text{M-PAM}\}$ or $\{\text{M-QAM}\}$.

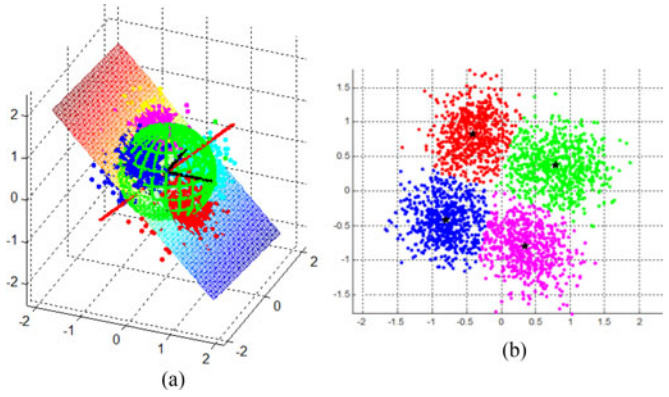

Fig. 3. 32 GBd-PDM-QPSK signal after 810 km transmission in LAF at 0 dBm launch power with a received OSNR of 22 dB: (a) represented in Stokes space with the estimate Q of its symmetry plane; (b) represented after projection on the plane Q.

TABLE I
THEORETICAL VALUES OF THE COEFFICIENT d FOR IDEAL NOISELESS STOKES SPACE CONSTELLATIONS

Format	OOK	M-PSK	M-PAM	8QAM	16QAM	64QAM
d	0	0	0	0.1648	0.1720	0.1966

To first estimate the 2-D linear plane P, we implemented the PCA method. In Stokes space, projected clusters on P have the characteristic of exhibiting a maximized variance. So, using PCA, we search for the principal components that maximize the projected data variance. The main advantage of PCA lies in its robustness to noise variance. Consequently, the plane Q spanned by the first two principal components enables a better estimation of the symmetry plane P than the least-squares plane (we found the existence of degenerate cases using the least-squares method [6]).

After having identified the estimate Q of the symmetry plane P, we define a coefficient$d$that serves as a proxy to estimate the dimension of the Stokes space constellation.

In a similar approach as the one introduced by Borkowski *et al.*, [13], [14], we assume that the 3-D Stokes space constellation cluster distribution follows a Gaussian mixture model (GMM), and identify the 3-D Stokes space clusters parameters (mean position $\vec{\mu}_i$ and Dirichlet concentration parameter λ_i of the ith cluster) by applying a machine learning algorithm (variational Bayesian expectation maximization for GMMs) [22], Fig. 3(a).

The coefficient$d$evaluates the aggregate squared distance of projected clusters mean positions on the straight line $\text{Q}^{\perp}$ orthogonal to the plane Q spanned by the first two principal components, and is given by: $d = \sum_{cluster\,i} \alpha_i | < \vec{\mu}_i, \vec{u} > |^2$, where: $\vec{u}$ is the unitary vector of $\text{Q}^{\perp}$, $\vec{\mu}_i$ and α_i ($= \lambda_i / \sum_i \lambda_i$) the mean position and the weight of the ith cluster, respectively.

Theoretical values of the coefficient d for ideal noiseless Stokes space constellation are summarized in Table I, showing

that, 2-D Stokes space constellations (OOK, M-PSK, M-PAM) exhibit $d = 0$, thus allowing discrimination from 3-D constellations (M-QAM with $M \neq \{2, 4\}$) which exhibit $d \neq 0$.

C. Format Recognition

At this point in the MFR algorithm, we have mapped the received signal to Stokes space, identified the estimate Q of the symmetry plane P of the Stokes space constellation, and determined if the latter is 2-D (OOK, M-PSK, M-PAM) or 3-D (M-QAM with $M \neq \{2, 4\}$).

Next, we need to identify in each situation the final modulation format of the received signal.

1) Amongst {OOK + M-PSK + M-PAM}: To discriminate modulation formats within {OOK + M-PSK + M-PAM} (2-D Stokes space constellations), we developed a method based on HOS and GMMs.

HOS provides a statistical toolset that enables the identification of a signal's modulation format by capturing its unique statistical signature in the signal constellation diagram. Conventionally, the statistical signature of the modulation format is captured from the signal constellation represented in the I/Q plane. But in our method, the modulation format statistical signature is captured from the signal constellation represented in the estimate Q of the symmetry plane P of the Stokes space constellation.

The method is initiated with the orthogonal projection of 3-D Stokes space data points on the plane Q.

Then, assuming the 2-D projected clusters distribution follows a GMM, we identify their parameters in the plane Q (mean position $\vec{\mu}_{Q,i}$ and Dirichlet concentration parameter $\lambda_{Q,i}$ of the ith cluster) by applying the same machine learning algorithm as previously, Fig. 3(b).

Conventional normalized estimates of fourth-order cumulants [23] are reformulated (see Appendix A) in order to take into account of the 2-D projected clusters mean positions $\{\vec{\mu}_{Q,i}\}$ and weights $\{\alpha_{Q,i}\}$ ($\alpha_{Q,i} = \lambda_{Q,i} / \sum_i \lambda_{Q,i}$) and. The modified statistics, introduced as normalized weighted estimates $\tilde{\mathrm{C}}_{4,j}$ of fourth-order cumulants, provide a more reliable interpretation of the statistics of the distribution of the projected constellation in the plane Q.

Finally, to identify the signal modulation format among OOK, M-PSK and M-PAM, normalized weighted estimates $\tilde{\mathrm{C}}_{4,j}$ of fourth-order cumulants are employed in a hierarchical classification scheme, Fig. 7(a), similar to the approach presented by Swami and Sadler [23].

The theoretical magnitudes of the normalized weighted estimates $\tilde{\mathrm{C}}_{4,j}$ of fourth-order cumulants are summarized in Table II for the modulation formats of interest.

2) Amongst {M-QAM} ($M \neq \{2, 4\}$).: To discriminate modulation formats within {M-QAM} with $M \neq \{2, 4\}$ (3-D Stokes space constellations), we use a cross-product statistic Γ_{XY}, defined as spatial cross-correlation index, inspired from spatial autocorrelation models [24]. The spatial cross-correlation index Γ_{XY} evaluates, in the 3-D Stokes space, the correlation between the mapped signal (observation X) and a given reference M-QAM Stokes space constellation (observation Y). Γ_{XY} is based on the deviation between the two sets of observations X and Y, and is given by: $\Gamma_{\mathrm{XY}} = \sum_{i,j} d_{i,j} / \sum_{i,j} \frac{1}{d_{i,j}}$, where $d_{i,j}$ is the Euclidian distance between observations $(x_i, y_j) \in$ (X,Y).

TABLE II
THEORETICAL MAGNITUDES OF THE NORMALIZED WEIGHTED ESTIMATES OF FOURTH-ORDER CUMULANTS FOR IDEAL NOISELESS STOKES SPACE CONSTELLATIONS

Format	OOK	BPSK	QPSK
$\lvert\tilde{C}_{4,0}\rvert$	~0.56	2	1
$\lvert\tilde{C}_{4,1}\rvert$	0	2	0
$\lvert\tilde{C}_{4,2}\rvert$	~0.72	2	−1

Spatial observations sets X and Y are highly correlated as the spatial cross-correlation index Γ_{XY} tends to 0.

Because the variation of the received SOP will cause a rotation of the Stokes space constellation (observation X) with time, we need to ensure observations sets X and Y are spatially synchronized (i.e., have a maximum of overlap in space) prior calculating Γ_{XY}. Overlapping observations sets X and Y in space is performed in two steps: first by aligning the observations sets respective symmetry planes, and then, by rotating one observation set with respect to the other one, around the normal to the aligned symmetry planes, to minimize the sum of squared errors between each pair of observations (x_i, y_k) (with $y_k = \arg \min_{y_j} \{d_{i,\mathrm{j}}\} = \{y_j | \, \forall y_l \colon d_{i,\mathrm{j}} \leq d_{i,1}\}$: the closest observation $y_k \in \mathrm{Y}$ to $x_i \in \mathrm{X}$).

After having calculated the cross-correlation indices Γ_{XY} between the mapped signal and multiple reference M-QAM Stokes space constellations, the identification of the M-QAM format is given by the minimization of the multiple normalized cross-correlation indices.

As a conclusion to this section, we note that it will be useful to study the robustness of the Stokes space methods subject to large impairments, including residual CD, differential group delay (DGD), PDL, and/or channel nonlinearities. Specifically, Stokes space constellations may change shape for DGD comparable to the optical signal symbol period [21]. Additionally, the proposed MFR method can be refined by choosing more complex probability density functions to model clusters distribution.

III. EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate the performance of the proposed MFR method, we use our well calibrated 32 GBd fiber test bed with custom conventional demodulation algorithms. Details about our custom conventional demodulation architecture can be found in previous work [5]. We first assessed the demodulating capabilities (constrained by the link performance) of the fundamental (non-blind) demodulation methods, Fig. 1, by evaluating the BER versus OSNR performance of the experimental link. We then quantified the classification performance of the proposed Stokes space-based MFR method by investigating the percentage of correctly identified waveforms versus OSNR.

The experimental test bed, Fig. 4, is composed of a single-channel optical baseband transmitter, a fiber optic link consisting

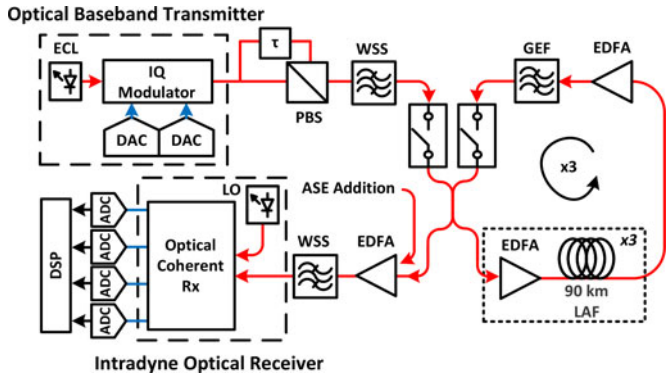

Fig. 4. Optical communication link experimental configuration.

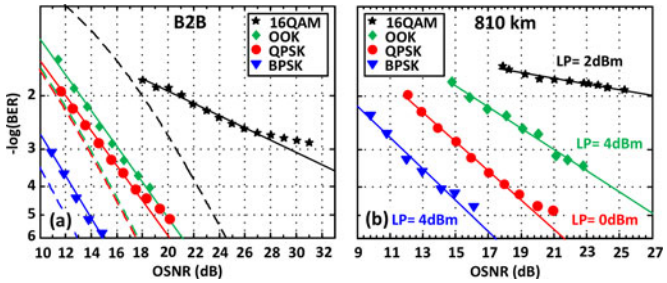

Fig. 5. Demodulation performance of the optical communication link for 31.5 GBd-PDM-signals: (a) in the back-to-back configuration; (b) after 810 km transmission in LAF at different launch powers.

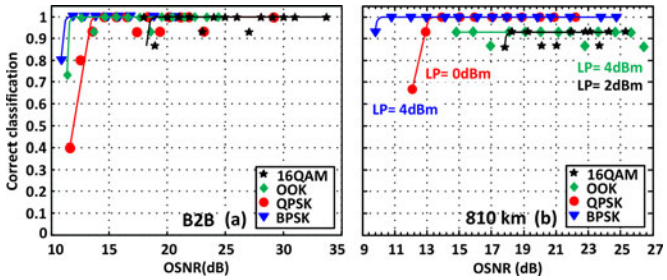

Fig. 6. Classification performance of the proposed Stokes space-based MFR method for 31.5 GBd-PDM-signals: (a) in the back-to-back configuration; (b) after 810 km transmission in LAF at different launch powers.

of an EDFA-based recirculating loop comprised of three 90 km LAF spans, and an intradyne coherent optical receiver with outputs digitized at 80 GSa/s by synchronized 30 GHz bandwidth real-time analog-to-digital converters (ADCs).

The transmitter generates a 31.5 GBd single-carrier signal centered at 1550.92 nm using a LN-IQ modulator with either OOK/ BPSK/ QPSK/16QAM. A pulse pattern generator is used to create the OOK and BPSK signals, whereas a 63 GSa/s high-speed digital-to-analog converter (DAC) is used to generate the QPSK and 16-QAM signals with a RRC-0.01-roll-off pulse shape. For DAC-generated signals, our previously described transmitter-based pre-equalization scheme [25], wherein the tap weights are determined by the receiver-side LMS equalization algorithm, is used to compensate for the band-limitations and other linear imperfections of the overall channel.

All optical signals were channelized for a 50 GHz grid, using a reconfigurable optical filter with 46 GHz 3 dB bandwidth, and polarization multiplexed before transmission. After polarization multiplexing, the data is circulated three times through the loop for a total transmission of 810 km. ASE is added to the optical signal prior detection to control the received OSNR. BER is quantified using PRBS-15.

Signal OSNRs were measured at the input of the coherent optical receiver. For the 810 km LAF transmission experiment, launch powers were chosen close to the optimum launch power for each modulation format (minimum of the BER versus launch power curve).

At a given OSNR, both in the back-to-back configuration and after 810 km transmission in LAF, the demodulation performance (quantified by the BER), Fig. 5, and the classification performance (quantified by the percentage of correctly identified waveforms), Fig. 6, were examined over 15 discontinuous data captures, each containing 1 00 000 samples/polarization (40 000 symbols/polarization).

The demodulation performance of the optical communication link is excellent in the back-to-back configuration, Fig. 5(a), as the BER versus OSNR curves approach the theoretical limit (dashed lines) [26] for 31.5 GBd-PDM-OOK, BPSK, and QPSK signals with a small implementation penalty. However, the implementation penalty for 31.5 GBd-PDM-16QAM signals is greater, and a BER floor is reached for very high OSNRs, which we attribute to hardware limitations. The multi-span 810 km LAF transmission, Fig. 5(b), exhibits good performance with an added link implementation penalty.

The classification performance for all the modulation formats studied (OOK, BPSK, QPSK, 16QAM) is exceedingly robust in the back-to-back configuration, Fig. 6(a), and after 810 km transmission in LAF, Fig. 6(b). The Stokes space-based MFR method yields the correct format with very high confidence for reasonable OSNRs. Indeed, the percentage of correctly identified waveforms is nearly 100% for OSNRs greater or equal to the OSNR required to achieve a 10^{-2} BER. Increasing the number of waveforms examined will increase the confidence level. We note that after 810 km transmission, the percentage of correctly identified waveforms for OOK or 16QAM signals does not approach 100%, Fig. 6(b), most likely due to a combination of residual CD, DGD, PDL, and nonlinearities.

We further examine, Figs. 7 and 8, the decision metrics used in the MFR process (coefficients d, magnitudes of normalized weighted estimates $\tilde{C}_{4,j}$ of fourth-order cumulants, and spatial cross-correlation indices) versus OSNR.

Fig. 7(a) represents the decision tree describing the process; the first decision separating the {M-QAM} formats from {OOK + M-PSK + M-PAM} was described previously. Figs. 7(b), (c), (d), and 8(a) depict the decision metrics for the four modulation formats in the back-to-back configuration, and Figs. 7(e), (f), (g), and 8(b) depicts the same metrics after 810 km transmission in LAF. Specifically, Fig. 7(b) and (e) illustrates the average coefficient$d$ versus OSNR in the back-to-back (after transmission). The 2-D Stokes space constellation coefficient d clearly lies below the displayed threshold, with an increased reliability for higher OSNR regimes. Fig. 7(c) and (f) shows the

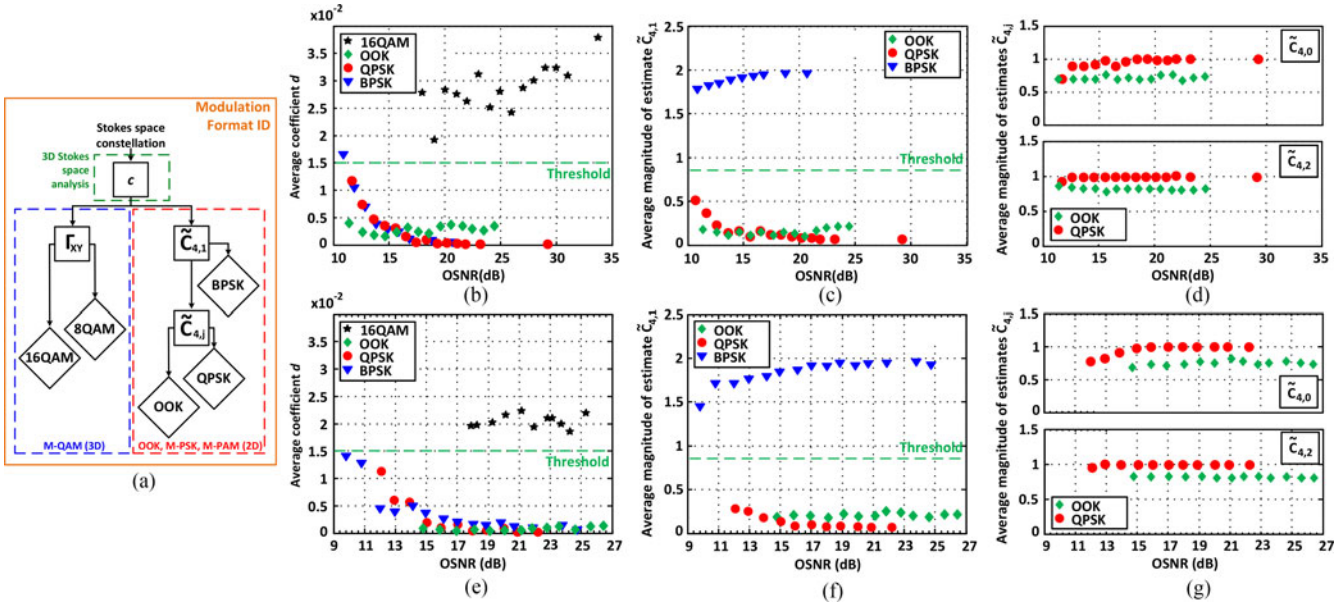

Fig. 7. (a) Decision metrics used in the proposed Stokes space-based MFR method. (b) Average coefficient d versus OSNR; (c) average magnitude of the normalized weighted estimate $\tilde{C}_{4,1}$ versus OSNR; (d) average magnitude of the normalized weighted estimates $\tilde{C}_{4,j}$ versus OSNR ($j = 0$ or 2) for 31.5 GBd-PDM-signals in the back-to-back configuration. (e) Average coefficient d versus OSNR; (f) average magnitude of the normalized weighted estimate $\tilde{C}_{4,1}$ versus OSNR; (g) average magnitude of the normalized weighted estimates $\tilde{C}_{4,j}$ versus OSNR ($j = 0$ or 2) for 31.5 GBd-PDM-signals after 810 km transmission in LAF.

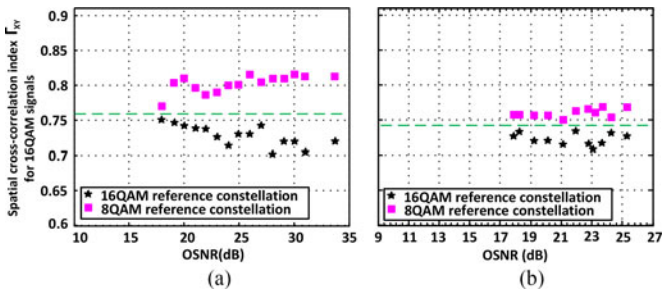

Fig. 8. Average spatial cross-correlation index calculated for 31.5 GBd 16 QAM signals versus OSNR: (a) in the back-to-back configuration; (b) after 810 km transmission in LAF at different launch powers.

magnitudes of normalized weighted estimate $\tilde{C}_{4,1}$ versus OSNR in the back-to-back (after transmission) used to distinguish BPSK from {OOK + QPSK}. Experimental measurements for BPSK lie in a non-ambiguous fashion above the displayed threshold. Fig. 7(d) and (g) exhibits the magnitudes of normalized weighted estimates $\tilde{C}_{4,j}$ versus OSNR ($j = 0$ or 2) in the back-to-back (after transmission) used to distinguish OOK from QPSK. Experimental measurements converge closely to theoretical values calculated for ideal noiseless constellations (see Table II). Finally, Fig. 8(a) and (b) shows the spatial cross-correlation index Γ_{XY} versus OSNR for 16QAM in the back-to-back (after transmission). The minimization process robustly results in the identification of 16QAM formats across the range of investigated OSNR.

IV. Conclusion

We presented a new MFR method that exploits digital coherent receivers to enable identification of signals modulated with M-PSK, M-PAM, M-QAM, and legacy formats. The method is based on a hybrid approach combining Stokes space analysis and HOS. Via experiment we demonstrated robust identification of a variety of signal formats over a wide range of received OSNR including OOK, BPSK, QPSK, and 16 QAM. The advantages of the proposed algorithm include: the capacity to autonomously identify single and dual polarization formats, a resistance to low-OSNR conditions inherent to HOS, and an immunity to the received SOP variations. Furthermore the Stokes space method is inherently insensitive to residual carrier-LO frequency and phase offsets.

The proposed algorithm does not require training, and is performed relatively early in the receiver DSP architecture. It is envisioned that the format recognition is only required at start up or when there is a loss of coordination between transmitter and receiver.

In other investigations, we successfully validated the performance of the proposed MFR method with multi-gigabit wireless signals generated with a high-speed optical transmitter, demonstrating the scalability of the method to photonic assisted wireless systems [27].

Previous work on autonomous receivers [4]–[10] combined with the results presented in this paper enable the deployment of autonomous architectures capable of identifying and demodulating unknown signals from noncooperating transmitters, including legacy optical transmitters.

Appendix

Weighted Estimates of Fourth-Order Cumulants

Let $x[k]$ be a digital signal of length N, with zero-mean, comprised of noisy complex-valued symbols. We assume $x[k]$ is a stationary random process and the symbol distribution follows

a GMM. Let $\vec{\mu}_i$, λ_i, $\alpha_i (= \lambda_i / \sum_i \lambda_i)$ be the mean position, the Dirichlet concentration parameter, and the weight of the ith cluster, respectively, in the complex plane.

Let $\mathrm{C}_{4,i}$ be the ith fourth-order cumulant of $x[k]$. Let $\hat{\mathrm{C}}_{4,i}$ be an estimator of the fourth-order cumulants $\mathrm{C}_{4,i}$. Note $\tilde{\mathrm{C}}_{4,i}$ the normalized version of the estimator $\hat{\mathrm{C}}_{4,i}$. $\tilde{\mathrm{C}}_{4,i}$ is expressed as: $\tilde{\mathrm{C}}_{4,i} = \hat{\mathrm{C}}_{4,i} \,/\, \hat{\mathrm{C}}^2_{2,1}$, where $\hat{\mathrm{C}}_{21}$ is the estimator of the second-order moment $\mathrm{C}_{2,1}$ and is given by: $\hat{\mathrm{C}}_{2,1} = \frac{1}{N}\sum_{k=1}^{N} |x[k]|^2$.

In the general case, we can define estimators $\hat{\mathrm{C}}_{4,i}$ as a function of $x[k]$ and its higher order moments [23]. For example, the estimate $\hat{\mathrm{C}}_{4,2}$ of the $2nd$ fourth-order cumulant $\mathrm{C}_{4,2}$ is express as: $\hat{\mathrm{C}}_{4,2} = \frac{1}{N}\sum_{k=1}^{N} |x[k]|^4 - |\hat{\mathrm{C}}_{2,0}|^2 - 2\hat{\mathrm{C}}^2_{2,1}$, where: $\hat{\mathrm{C}}_{2,0} = \frac{1}{N}\sum_{k=1}^{N} x[k]^2$.

We can reformulate normalized estimates $\tilde{\mathrm{C}}_{4,i}$ of fourth-order cumulants $\mathrm{C}_{4,i}$ in order to take into account of the complex plane clusters mean positions $\{\vec{\mu}_i\}$ and weights $\{\alpha_i\}$. For example, the normalized weighted estimate $\tilde{\mathrm{C}}_{4,2,w}$ of the $2nd$ fourth-order cumulant $\mathrm{C}_{4,2}$ can be express as: $\tilde{\mathrm{C}}_{4,2,w} = \hat{\mathrm{C}}_{4,2,w} / \hat{\mathrm{C}}^2_{2,1,w}$, where:

$$\begin{cases} \hat{\mathrm{C}}_{4,2,w} = \displaystyle\sum_{cluster\, i} \alpha_i |\mu_i|^4 - \left|\hat{\mathrm{C}}_{2,0,w}\right|^2 - 2\hat{\mathrm{C}}_{2,1,w} \\ \hat{\mathrm{C}}_{2,0,w} = \displaystyle\sum_{cluster\, i} \alpha_i u_i^2 ; \hat{\mathrm{C}}_{2,1,w} - \sum_{cluster\, i} \alpha_i \, |\mu_i|^2 \end{cases}$$

REFERENCES

[1] X. Zhou, L. E. Nelson, and P. Magill, "Rate-adaptable optics for next-generation long-haul transport networks," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 41–49, Mar. 2013.

[2] C. Laperle and M. O'Sullivan, "High-speed DACs and ADCs for next generation flexible transceivers," presented at the Signal Processing Photonic Communications Conf., San Diego, CA, USA, Jul. 2014.

[3] K. Roberts and C. Laperle, "Flexible transceivers," presented at the European Conf. Optical Communication, Amsterdam, The Netherlands, Sep. 2012, Paper We.3.A.3.

[4] P. Isautier, A. Stark, K. Mehta, R. de Salvo, and S. E. Ralph, "Autonomous software-defined coherent optical receivers," presented at the Optical Fiber Communication Conf. Exposition, Nat. Fiber Optical Engineering Conf., Anaheim, CA, USA, Mar. 2013, Paper OTh3.B.4.

[5] P. Isautier, A. Stark, K. Mehta, R. de Salvo, and S. E. Ralph, "Robust architecture for autonomous coherent optical receivers," *J. Opt. Commun. Netw.*, vol. 7, no. 9, pp. 864–874, Sep. 2015.

[6] P. Isautier, A. Stark, J. Pan, K. Mehta, and S. E. Ralph, "Autonomous software-defined coherent optical receivers performing modulation format recognition in stokes-space," presented at the European Conf. Exhibition Optical Communication, London, U.K., Sep. 2013, Paper P.3.30.

[7] P. Isautier, J. Pan, and S. E. Ralph, "Robust autonomous software-defined coherent optical receiver," presented at the Optical Fiber Communication Conf., San Francisco, CA, USA, Mar. 2014, Paper W1.G.7.

[8] P. Isautier, J. Pan, J. Langston, and S. E. Ralph, "Performance comparison of autonomous software-defnied coherent optical receivers," presented at the IEEE Photon. Conf., San Diego, CA, USA, Oct. 2014, Paper MG3.2.

[9] P. Isautier, J. Pan, J. Langston, R. DeSalvo, and S. E. Ralph, "Autonomous receivers for complex format identification and demodulation," presented at the Avionics Vehicle Fiber-Optics Photonics Technology Conf., Atlanta, GA, USA, Nov. 2014, Paper TuB2.

[10] P. Isautier, J. Langston, J. Pan, and S. E. Ralph, "Agnostic software-defined coherent optical receiver performing time-domain hybrid modulation format recognition," presented at the Optical Fiber Communication Conf. Exhib., , Los Angeles, CA, USA, Mar. 2015, Paper Th2A.21.

[11] N. G. Gonzalez, D. Zibar, and I. T. Monroy, "Cognitive digital receiver for burst mode phase modulated radio over fiber links," presented at the European Conf. Optical Communication, Torino, Italy, Sep. 2010, Paper P6.11.

[12] F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lao, "Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks," *Opt. Exp.*, vol. 20, no. 11, pp. 12422–12431, May 2012.

[13] R. Borkowski, D. Zibar, A. Caballero, V. Arlunno, and I. T. Monroy, "Optical modulation format recognition in stokes space for digital coherent receivers," presented at the Optical Fiber Communication Conf. Exposition National Fiber Optic Engineers Conf., Anaheim, CA, USA, Mar. 2013, Paper OTh3B.3.

[14] R. Borkowski, D. Zibar, A. Caballero, V. Arlunno, and I.T. Monroy, "Stokes space-based optical modulation format recognition for digital coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 25, no. 21, pp. 2129–2132, Nov. 2013.

[15] G. Bosco, M. Visintin, P. Poggiolini, A. Nespola, M. Huchard, and F. Forghieri, "Experimental demonstration of a novel update algorithm in stokes space for adaptive equalization in coherent receivers," presented at the European Conf. Optical Communication, Cannes, France, Sep. 2014.

[16] R. Boada, R. Borkowski, and I. T. Monroy, "Clustering algorithms for stokes space modulation format recognition," *Opt. Exp.*, vol. 23, no. 12, pp. 15521–15531, Jun. 2015.

[17] J. Liu, Z. Dong, K. Zhong, A. P. T. Lau, C. Lu, and Y. Lu, "Modulation format identification based on received signal power distribution for digital coherent receivers," presented at the Optical Fiber Communication Conf. Exposition National Fiber Optic Engineers Conf., San Francisco, CA, USA, Mar. 2014, Paper Th4D.3.

[18] M. C. Tan, F. N. Khan, W. H. Al-Arashi, Y. Zhou, and A. P. T. Lau, "Simultaneous optical performance monitoring and modulation format/bit-rate identification using principal component analysis," *J. Opt. Commun. Netw.*, vol. 6, no. 5, pp. 441–448, May 2014.

[19] S. Cui, S. He, J. Shang, C. Ke, S. Fu, D. Liu, "Method to improve the performance of the optical modulation format identification system based on asynchronous amplitude histogram," *Opt. Fiber Technol.*, vol. 23, pp. 13–17, Jun. 2015.

[20] B. Szafraniec, B. Nebendahl, and T. S. Marshall, "Polarization demultiplexing in stokes space," *Opt. Exp.*, vol. 18, no. 17, pp.17928–17939, Aug. 2010.

[21] B. Szafraniec, T. S. Marshall, and B. Nebendhal, "Performance monitoring and measurement technique for coherent optical systems," *J. Lightw. Technol.*, vol. 31, no. 4, pp. 648–663, Feb. 2013.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning*, 3rd ed. New York, NY, USA: Springer-Verlag, 2006.

[23] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[24] A. Getis, "Spatial interaction and spatial autocorrelation: A cross product approach," *Environ. Planning A*, vol. 23, no. 9, pp. 1269–1277, Sep. 1991.

[25] J. Pan, P. Isautier, and S. E. Ralph, "Digital pre-shaping for narrowband filtering impairment compensation in superchannel applications," presented at the Advanced Photon. Optical. Sens.ors Conf., Rio Grande, PR, USA, Jul. 2013, Paper JT3A.1.

[26] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston, MA, USA: McGraw-Hill, 2008.

[27] P. Isautier and S. E. Ralph, "Autonomous identification and detection for multi-gigabit photonic assisted wireless links," presented at the IEEE Int. Topical Meeting Microw. Photon., Paphos, Cyprus, Oct. 2015.

Authors' biographies not available at the time of publication.