Joint Symbol Rate-Modulation Format Identification and OSNR Estimation Using Random Forest Based Ensemble Learning for Intermediate Nodes

Jia Chai, Xue Chen, Yan Zhao, Tao Yang, Danshi Wang, and Sheping Shi

***Abstract*—In this paper, a novel joint symbol rate-modulation format identification (SR-MFI) and optical signal-to-noise ratio (OSNR) estimation scheme using the low-bandwidth coherent detecting and random forest (RF)-based ensemble learning is proposed for intermediate nodes in the flexible dense wavelength division multiplexing (F-DWDM) networks. By leveraging low-bandwidth coherent detecting with small bulk wavelength scanning, no chromatic dispersion compensation and low-complexity RF, the proposed scheme could serve as a reduced-complexity and cost-effective option to realize joint SR-MFI and OSNR estimation at intermediate nodes in F-DWDM networks. To verify the feasibility of the proposed scheme, the comprehensive simulations of 8/16 GBaud polarization division multiplexing (PDM)-4/16/32/64 quadrature amplitude modulation (QAM) systems are conducted. The simulation results show that the identification accuracy of SR-MFI reaches 100% and the mean absolute error of OSNR estimation is within 1 dB. Moreover, the proposed monitoring scheme is verified by 8/16 GBaud PDM-4/16/32QAM coherent transmission experiments.**

***Index Terms*—Symbol rate and modulation format identification, OSNR estimation, random forest, intermediate nodes, low bandwidth.**

I. Introduction

THE capacity of fiber-optic communication systems is increasing ceaselessly and the network architecture is becoming more complex, transparent, and dynamic in nature. The intermediate nodes will be capable of adjusting multiple channel parameters dynamically according to time-varying traffic demands and channel conditions in the flexible dense wavelength division multiplexing (F-DWDM) networks [1]. Consequently, it is imperative to ensure the consistency of the multiple channel parameters such as modulation formats (MFs) or symbol rates (SRs) and that arranged by the control plane in real time. Besides, at intermediate nodes, it is also crucial to estimate the optical signal-to-noise ratio (OSNR) which is directly related to the quality of the transmitted service [2]. The real-time channel knowledge enables the current channel conditions to be suitable for the transmission, which is conducive for networks operating efficiently and reliably.

To implement reliable monitoring on the whole network, two different optical performance monitoring (OPM) modes are equipped, one of which is the end-to-end channel monitoring at destination node. It shares coherent receiver and analog-to-digital conversion (ADC) with service signals' recovery to generate sampling points for monitoring. The other is the intermediate node monitoring at intermediate node which does not require service signals' recovery, but incorporates large inline chromatic dispersion (CD) and can only handle the limited implementation complexity [3]. A lot of schemes have been proposed to implement end-to-end channels monitoring, such as the neural-network based algorithms for coherent receivers [4]–[8] and the classical algorithms that employ statistical features of the signals [9]–[13], but only this kind of monitoring is far from enough to perceive the status of the whole network in order to carry out operation, administration and maintenance (OAM) functions, such as the accurate location of fault points. Therefore, some works have been done by full-bandwidth direct-detecting receiving to achieve joint multiple parameters monitoring of all wavelength signals in each fiber input at intermediate nodes [14]–[18]. But they are sensitive to CD, which is not suitable for the intermediate node monitoring with large inline CD. To cope with these issues, an OSNR monitoring scheme with large inline CD using an artificial neural network (ANN) has been proposed based on low-bandwidth direct-detecting receiver at intermediate notes [19]. It has been proved that CD affects the high frequency components of the service signals more than the low frequency components which can potentially be used for CD insensitive OSNR monitoring. However, the full connections among the neurons of different layers requiring numerous multipliers, adders and comparators, resulting in the complexity is much larger than the random forest (RF) which can achieve good performance with small computational overhead [8], [20]–[22]. Recently, an OSNR estimation scheme based on low-bandwidth coherent receiver was proposed [23]. However, OSNR monitoring based on a discrete classifier with 1 dB interval leads to high monitoring error.

Manuscript received October 1, 2021; accepted October 2, 2021. Date of publication October 6, 2021; date of current version October 21, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61871415. *(Corresponding author: Xue Chen.)*

Jia Chai, Xue Chen, Tao Yang, and Danshi Wang are with the State Key Laboratory of Information Photonics and Optical Communications, University of Posts and Telecommunications, Beijing 100876, China (e-mail: cjchaijia@163.com; xuechen@bupt.edu.cn; yangtao@bupt.edu.cn; danshi_wang@bupt.edu.cn).

Yan Zhao and Sheping Shi are with the ZTE Corporation, Beijing 100191, China (e-mail: buptyanzhao@bupt.edu.cn; 13501025782@163.com).

Digital Object Identifier 10.1109/JPHOT.2021.3117984

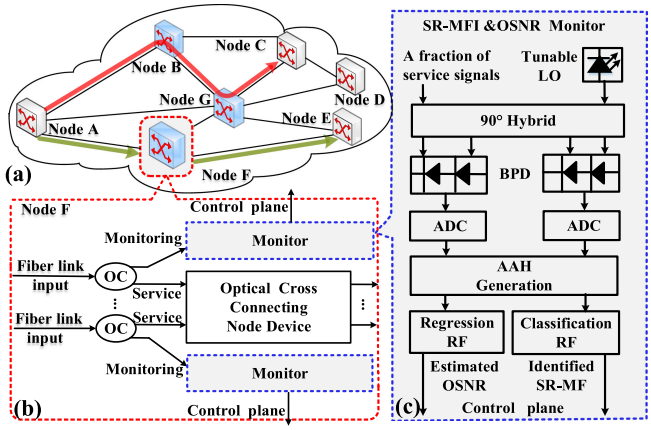

Fig. 1. The proposed joints SR-MFI and OSNR estimation scheme at intermediate nodes.

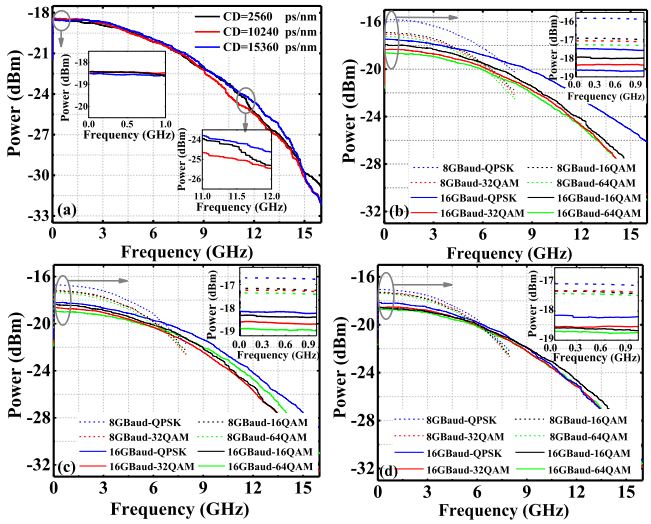

Fig. 2. (a) Power spectrums of different CDs under PDM-16QAM. Power spectrums of different SR-MFIs under various OSNRs @ 2560 ps/nm CD (b) @ SD-FEC threshold, (c) 8 dB, (d) 15 dB greater.

In this paper, a joint SR-MF identification (SR-MFI) and OSNR estimation scheme using low-bandwidth and tunable coherent detecting is proposed. By tuning the frequency of local oscillator (LO), the proposed scheme can conveniently scanning all wavelength channels in one fiber port. Moreover, combining with the RF, the monitoring schemes at intermediate nodes have the advantages of cost-effectiveness, small bulk, and high flexibility.

II. Scheme Principle

A. Intermediate Nodes Monitoring

A large number of end-to-end wavelength channels pass through numerous fiber links and intermediate nodes in the F-DWDM networks where the frequency slot width granularity (12.5 GHz) [24] and limited optical transport unit bit rates [25] specified by ITU-T usually lead to that the number of nominal SRs of wavelength links is limited. As regards the F-DWDM networks' monitoring, the SRs and MFs of a wavelength link are identified to verdict the consistency with its configuration. Here two SRs and four MFs are taken as examples without loss of generality and the joint SR-MFI and OSNR estimation scheme aimed to be used at intermediate nodes is proposed and shown in Fig. 1, where the node B and G are intermediate nodes under the end-to-end channel connection from source node A to destination node C expressed in the red line shown in Fig. 1(a) [26].

As shown in Fig. 1(b), at the intermediate node F, each fiber input signals are divided into two branches through a 1:99 optical coupler (OC). Taking advantage of the high sensitivity of coherent detecting, 1% of the service signals' power is taken out for the proposed SR-MFI and OSNR estimation, thus minimizing the impact on the service signals. The rest 99% of the signals are processed as usual. As exhibited in Fig. 1(c), by tuning LO instead of an optical bandwidth variable tunable filter, all wavelength channels in one fiber can be scanned, leading to small bulk. Besides, a fraction of service signals and tunable LO are mixed by just a 90° hybrid and two pairs of balanced photodiode (BPD) with low bandwidth to obtain the monitoring signals and then digitized by two low-rate sampling ADCs, enabling implementation simplicity and cost effectiveness substantially. The polarization-independent processing is done at the front end of the receiver since the two polarization states generally share the same MFs, SRs and so on. Each asynchronous amplitude histograms (AAHs) extracted from the low bandwidth-filtered and low rate-sampled symbols, which is a distribution histogram of symbol modulus, is minimally affected by the polarization mode dispersion and polarization crosstalk. Subsequently, because the complexity of RF is much lower compared to machine learning algorithms based on fully connected structures and distance metrics, such as deep neural network, and support vector machine under the same monitoring performance [8], the generated AAH data set is simultaneously fed into two different types of RFs which are the classification RF used for discrete SR-MFI and the regression RF used for continuous OSNR estimation. Finally, the information about the identified SR-MF and estimated OSNR is sent to the control plane to make it perceive the underlying parameters and aware of the quality of the current link.

B. Power Spectrums of Different CDs and SR-MFIs

The low-frequency components of power spectrums which have a certain tolerance for CD, as shown in Fig. 2(a), CD affects the high frequency components of power spectrums more than the low frequency components. Besides, to verify the feasibility of SR-MFI under the proposed scheme, the power spectrums of the eight SR-MFs which include two SRs @ 8/16 GBaud and four MFs @ polarization division multiplexing (PDM)-4/16/32/64 quadrature amplitude modulation (QAM) are evaluated. As shown in Fig. 2(b)–(d), under the OSNR of each SR-MFs at the respective soft decision forward error correction (SD-FEC) threshold or 8, 15 dB greater, the power spectrum exhibit unique and distinctive characteristics which are

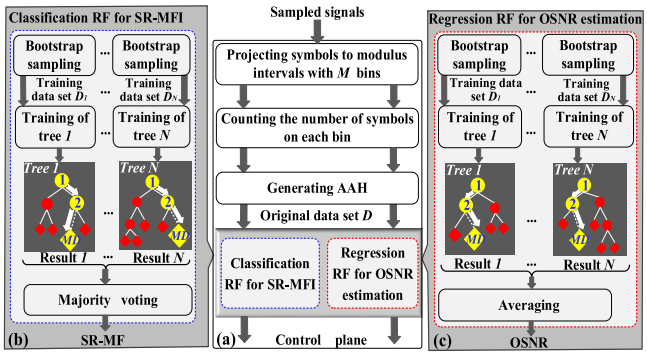

Fig. 3. The RF structure of the proposed joint SR-MFI and OSNR estimation scheme.

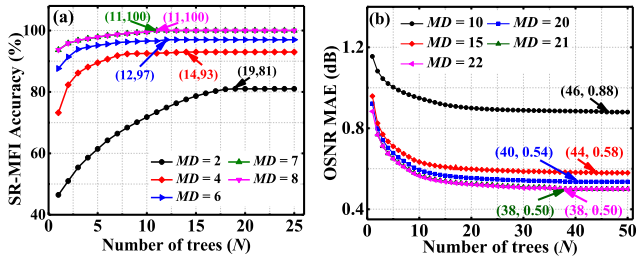

Fig. 4. (a) SR-MFI accuracy, (b) OSNR estimation MAE versus N with different MDs.

sensitive to the SR-MF along the whole band. However, considering Fig. 2(a)–(d) comprehensively, the use of low-frequency components of the monitoring signal can achieve CD-tolerant SR-MFI.

C. RF-Based Ensemble Learning Structure

The RF-based ensemble learning structure of the proposed joint SR-MFI and OSNR estimation scheme is illustrated. As shown in Fig. 3(a), the SR-MFI classification RF and the OSNR estimated regression RF are trained with the original data set D composed of various AAHs extracted from the monitoring signals. To generate AAHs, after calculating the symbol modulus of a polarization state, it is arranged in ascending order and then averagely divided into M bins, where M is taken as 80 according to [5]. The number of samples falling in each bin is counted and represented as a $1\times M$ vector to generate an AAH.

By traversing the channel transmission conditions, the AAH set can be obtained and is input into two independent RF algorithms in the joint monitor. The training of the SR-MFI classification RF is shown in Fig. 3(b). Firstly, the training data set denoted by D_n is obtained by the bootstrap sampling from D and the random feature subset is obtained by randomly selecting k features from D_n according to the rule of thumb, $k = \log_2 M + 1$, where M is the total number of features of AAH [22]. The introduction of two randomness which are bootstrap sampling and random selection method makes the RF less prone to overfitting problems [27]–[29]. Secondly, the Gini index of each feature in the random feature subset is calculated. The smaller the Gini index, the higher the proportion of AAH belonging to the same category, hence the feature with the minimum Gini index is selected as the optimal division criterion to split the data set at the current node into left and right children nodes. With the recursive execution of this step, the depth of decision tree continues to increase until it reaches a preset value denoted by MD which is defined as the number of nodes included in the longest decision path of decision tree. Thirdly, the number of decision trees continues to increase with the recursive execution of all the above steps until it reaches a preset value. Finally, majority voting working as the aggregating strategy is implemented for SR-MFI and the one with the maximum votes is the final identified SR-MF. The training of the OSNR estimation regression RF is shown in Fig. 3(c). The minimum mean absolute error (MAE) is positively correlated with the difference between the true and estimated values, hence the node with the minimum MAE is selected as the optimal division criterion. Besides, the aggregating strategy namely averaging is implemented and the average value is the final estimated OSNR. The monitored information about SR-MF and OSNR is sent to the control plane to achieve the real-time perception and acquisition of the underlying link status.

The number of decision trees represented by N and the maximum depth of RF represented by MD are two important parameters that affect the SR-MFI accuracy, the OSNR estimation MAE, and the RF algorithm complexity. As MD increases initially, each decision tree of RF can learn more features of the samples, but there are no newer features to be learned when MD increases to a certain level. Therefore, the identification accuracy of each decision tree in the SR-MFI classification RF will continue to increase to a certain value and the MAE of each decision tree in the OSNR estimated regression RF will continue to decrease to a certain value well. As shown in Fig. 4, when N is large enough to ensure the convergence, increasing MD can increase the SR-MFI accuracy and decrease the OSNR estimation MAE, however, they will remain unchanged when MD increases to a certain extent. On the other hand, there is a certain similarity among decision trees since the bootstrap sampling, which means that as N continues to increase, the correlation between the newly added decision tree and the existing ones will also increase. Therefore, the degree of performance improvement is continuously reduced to a constant value, which conforms to the convergence theorem of RF. As shown in Fig. 4, when MD is certain, the initial increase of N makes the SR-MFI accuracy continuously increase and the OSNR estimation MAE continuously decrease. However, the degree of performance improvement is decreasing continuously and when it reaches a certain level, the SR-MFI accuracies and OSNR estimation MAEs will remain unchanged. However, the MD and N are positively correlated with the RF complexity determined mostly by the number of comparators which are $N \times (MD - 1) + C - 1$ and $N \times (MD - 1)$ for the SR-MFI classification RF and the OSNR estimated regression RF respectively, where C is the number of SR-MFs to be classified [8].

Comprehensively, the choice of N and MD should be a trade-off between the SR-MFI accuracy, the OSNR estimation MAE, and the complexity of RF. As shown in Fig. 4(a), for the SR-MFI

TABLE I
RESULTS OF JOINT SR-MFI AND OSNR ESTIMATION

	SR(GBaud)	QPSK(dB)	16QAM(dB)	32QAM(dB)	64QAM(dB)
OSNR	8	(3.01-18.09)	(8.86-23.01)	(12.56-27.86)	(16.03-31.52)
range	16	(5.23-20.85)	(11.48-26.27)	(14.15-29.34)	(18.14-32.49)
Lowest	8	3.02	9.25	12.63	16.25
OSNR	16	6.25	12.55	15.88	19.5
Mean MAE	16	0.31	0.39	0.45	0.56

classification RF, the identification accuracy under $N = 11$ and $MD = 7$ reaches 100% and improved by 3% at approximately the same complexity compared with $N = 12$ and $MD = 6$. Fixing $N = 11$ and increasing MD to 8, the identification accuracy maintains 100% but the complexity increases by 15%. Therefore, $N = 11$ and $MD = 7$ are more appropriate. Likewise, for the OSNR estimated regression RF, taking $N = 38$ and $MD = 21$ is appropriate as shown in Fig. 4(b). In summary, considering both complexity and performance of RF, the SR-MFI classification RF with $N = 11$, $MD = 7$ and the OSNR estimated regression RF with $N = 38$, $MD = 21$ are suitable.

III. SIMULATION DEMONSTRATION AND PERFORMANCE EVALUATION

Here we define the valid bandwidth range (VBR) of joint SR-MFI and OSNR estimation monitor as the bandwidth range where the performance of the scheme reaches 100% identification accuracy and at same time OSNR estimation MAE within 1 dB. For OSNR estimation, under 2560/10240/15360 ps/nm CDs and 1 GSa/s sampling rate for 8/16 GBaud PDM-4/16/32/64QAM, the VBR is [60, 500] MHz for the ratio of VBR to SR is 0.0038-0.0625 [26]. To obtain the VBR of SR-MFI, the identification accuracies are obtained by training the RF with four original data sets which are obtained under bandwidths @ 60, 125, 250, and 500 MHz respectively. Each original data set can be obtained by traversing the channel conditions which contain two SRs @ 8/16 GBaud, three CDs @ 2560/10240/15360 ps/nm, and four MFs @ PDM-4/16/32/64QAM. Besides, the 25 OSNRs within the range shown in Table I are contained and the steps are randomly selected within (0.5-1) dB. Moreover, 100 AAHs are collected for each OSNR. Hence the original data set under each bandwidth aggregately contains 60000 AAHs in total, 70% of which are used for training and the rest 30% for testing.

Fig. 5(a) shows that the VBR of joint SR-MFI and OSNR estimation is [60, 500] MHz. In particular, it can be seen that the identification accuracies of SR-MFI are 100% and the OSNR estimation error is the smallest under 125 MHz bandwidth which mainly depends on the range of monitoring SRs. Accordingly, MAE is obtained from around the SD-FEC threshold to 15 dB above for 16 GBaud PDM-4/16/32/64QAM under 125 MHz bandwidth and 2560/10240/15360 ps/nm CDs. As shown in Fig. 5(b), SR-MFI based on the classification RF can achieve 100% identification accuracy under the lowest OSNRs shown in Table I, which is less than the required OSNRs for SD-FEC

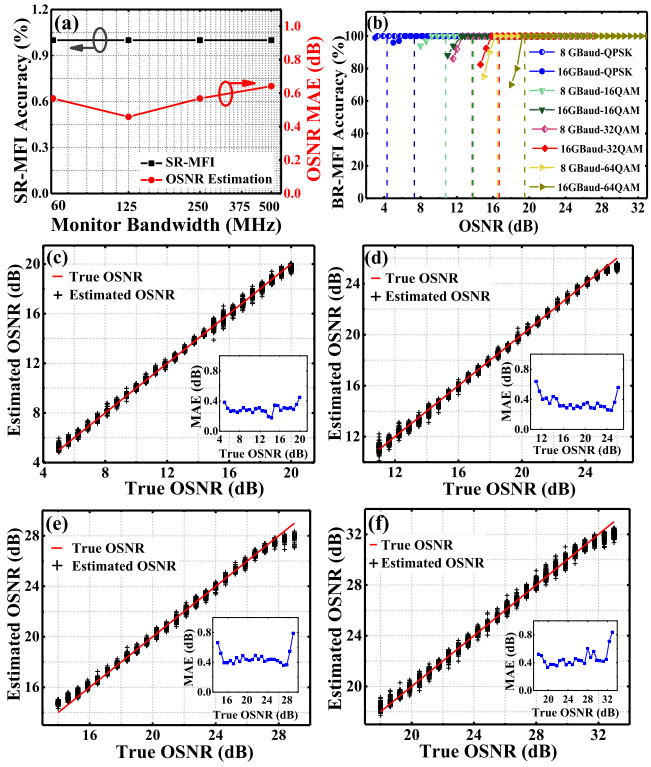

Fig. 5. (a) SR-MFI accuracy (left) and OSNR MAE (right) versus monitor bandwidth. (b) SR-MFI accuracy versus OSNR. Estimated OSNR versus actual OSNR @ 16 GBaud for (c) PDM-QPSK, (d) PDM-16QAM, (e) PDM-32QAM, (f) PDM-64QAM under 2560/10240/15360 ps/nm CDs.

threshold @ 2E-2 bit error rate (BER) for 8/16 GBaud PDM-4/16/32/64QAM. In addition, as shown in Fig. 5(c-f), the MAE of each OSNR is calculated by the difference between estimated OSNR and true OSNR. For 16GBaud PDM-4/16/32/64QAM, the mean MAEs CDs are shown in Table I. Among them, PDM-64QAM shown in Fig. 5(f) has the highest MAE because the higher-order MF is more sensitive to damage. Besides, the OSNR estimation errors are relatively large under the conditions of low OSNR (1–2 dB around the SD-FEC threshold) and high OSNR (13–15 dB higher than the SD-FEC threshold). This is because in the case of OSNRs close to the threshold, the monitoring signals are severely deteriorated and the features become less distinguishable, causing a large estimation error. Compared with the OSNR estimation range based on the RF and full-bandwidth coherent receiver [8], the proposed joint monitor performance is not good at low OSNR which is below the tolerance limit of 2–3 dB. It is since the use of low-bandwidth coherent detecting, some of the signal characteristics including in the high-frequency are discarded, which will deteriorate the monitoring range. When the OSNR is large enough (13–15 dB higher than the SD-FEC threshold), the feature discrimination of different OSNR samples is relatively low, leading to larger estimation errors.

IV. EXPERIMENT SETUP AND RESULTS

In order to investigate the feasibility and performance of the proposed scheme, transmission experiments are constructed to

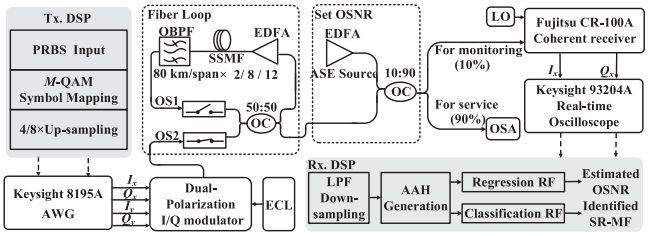

Fig. 6. Experimental setup for 8/16 GBaud PDM-4/16/32QAM coherent detecting systems.

generate kinds of widely-researched optical signals @ PDM-4/16/32QAM at 8/16 GBaud. As shown in Fig. 6, in the transmitter (Tx) DSP block, the length of pseudo random bit sequence (PRBS) utilized to generate 2/4/5-level 4/16/32QAM is $2^{17} - 1$. Then 8/4 time up-sampling are performed without Nyquist shaping and the oversampled signals are sent into arbitrary waveform generator (AWG, Keysight 8195A) working at 64 GSa/s to generate 8/16 GBaud signals respectively. The in-phase and quadrature (IQ) data generated by AWG drive dual-polarization IQ modulators with a tunable external cavity laser (ECL) to generate PDM signals which are transmitted in an optical fiber loop. The service signals with different MFs and SRs are transmitted by 2, 8, and 12 spans in the fiber loop and the transmission distances are 160, 640, and 960 km. Following the fiber loop, the amplified spontaneous emission (ASE) noise and transmitted service signals are input into a 90:10 OC which is easily available in our lab to guarantee the sensitivity of coherent receiver (Fujitsu CR-100A). The 90% power arm which should be used for service signal processing in practical networks is used for actual OSNRs measurement by optical spectrum analyzer (OSA) to evaluate monitoring errors in the experiment. The last 10% power in the other arm should be coherently received polarization-independently by the balanced coherent receiver with low bandwidth, but this kind of receiver is not available in our lab. Therefore, half of the coherent receiver and a digital low-pass-filter (LPF) after sampling are used to be equivalent to it. The sampling is realized by a real time sampling oscilloscope (Keysight 9320A) and sampling rate is dropped down in the following DSP.

Fig. 7(a) shows that the VBR of joint SR-MFI and OSNR estimation is [60, 500] MHz. In particular, the identification accuracy of SR-MFIs is 100% and the OSNR estimation error is the smallest under 125 MHz bandwidth. Accordingly, as shown in Fig. 7(b), SR-MFI based on the classification RF achieves 100% identification accuracy for 8 GBaud PDM-4/16/32QAM when OSNRs are as low as 4.02, 9.72, 12.92 dB and for 16 GBaud PDM-4/16/32QAM when OSNRs are as low as 7.02, 13.87, and 16.67 dB respectively which are less than the required OSNR for SD-FEC threshold @ 2E-2 BER for 8/16 GBaud PDM-4/16/32QAM. In addition, MAEs are obtained from around the SD-FEC threshold to 11–13 dB above for 16 GBaud PDM-4/16/32QAM. As shown in Fig. 7(c-e), the mean MAEs are 0.37, 0.48, and 0.59 dB respectively for PDM-4/16/32QAM under 2560/10240/15360 ps/nm CDs. Among them, PDM-32QAM shown in Fig. 7(e) has the highest MAE because the higher-order

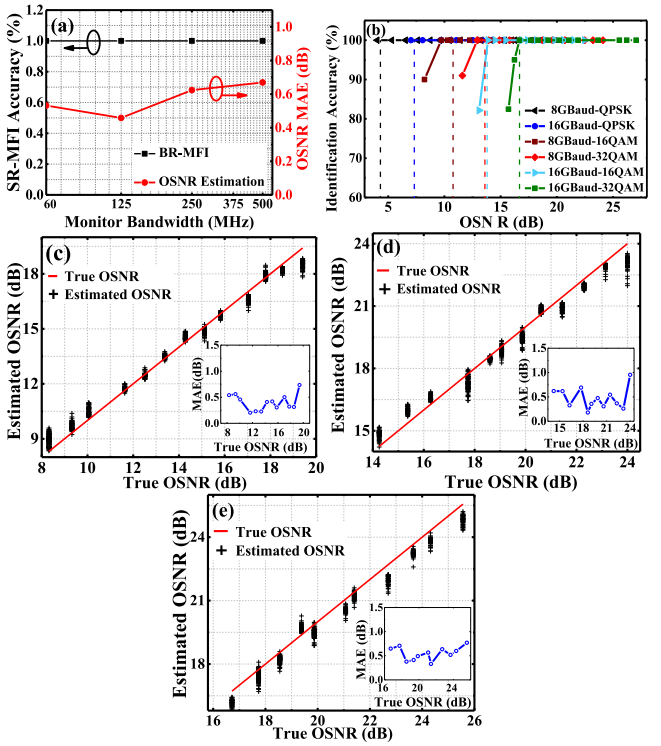

Fig. 7. (a) SR-MFI accuracy (left) and MAE (right) versus monitor bandwidth. (b) SR-MFI accuracy versus OSNR. Estimated OSNR versus actual OSNR @ 16 GBaud for (c) PDM-QPSK, (d) PDM-16QAM, (e) PDM-32QAM under 2560/10240/15360 ps/nm CDs.

MF is more sensitive to damage. Besides, the OSNR estimation error is relatively large under the conditions of low OSNR (1–2 dB around the SD-FEC threshold) and high OSNR (11–13 dB higher than the SD-FEC threshold).

V. Conclusion

In this paper, a low-cost and simplified joint SR-MFI and OSNR estimation scheme based on low-bandwidth coherent detecting with both small bulk, low-rate sampling, and low complexity RF-based ensemble learning algorithm trained with AAHs has been proposed and demonstrated. Simulations are conducted and the results show that for 8/16 GBaud PDM-4/16/32/64QAM, the VBR of joint SR-MFI and OSNR estimation is [60, 500] MHz. The identification accuracies of the eight SR-MFs are 100% and the mean MAEs of 16 GBaud PDM-4/16/32/64 QAM are 0.31, 0.39, 0.45, and 0.56 dB respectively under 2560/10240/15360 ps/nm inline CDs. In addition, the SR-MFI accuracy and OSNR estimation MAE have also been experimentally validated.

References

[1] R. M. Morais and J. Pedro, "Machine learning models for estimating quality of transmission in DWDM networks," *J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D84–D99, 2018.

[2] A. Yi *et al.*, "Modulation format identification and OSNR monitoring using density distributions in stokes axes for digital coherent receivers," *Opt. Exp.*, vol. 27, no. 4, pp. 4471–4479, Feb. 2019.

[3] Z. Dong, Q. Sui, F. N. Khan, K. Zhong, C. Lu, and A. P. Lau, "Optical performance monitoring: A review of current and future technologies," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 2016.

[4] D. Wang *et al.*, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Exp.*, vol. 25, no. 15, pp. 17150–17166, Jul. 2017.

[5] F. N. Khan *et al.*, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Exp.*, vol. 25, no. 15, pp. 17767–17776, Jul. 2017.

[6] C. Wang *et al.*, "Joint OSNR and CD monitoring in digital coherent receiver using long short-term memory neural network," *Opt. Exp.*, vol. 27, no. 5, pp. 6936–6945, Mar. 2019.

[7] 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," *Opt. Exp.*, vol. 26, no. 16, pp. 21346–21357, Aug. 2018.

[8] Y. Zhao *et al.*, "Low-complexity and joint modulation format identification and OSNR estimation using random forest for flexible coherent receivers," *Opt. Commun.*, vol. 457, Feb. 2020, Art. no. 124698.

[9] C. Zhu *et al.*, "Statistical moments-based OSNR monitoring for coherent optical systems," *Opt. Exp.*, vol. 20, no. 16, pp. 17711–17721, 2012.

[10] M. S. Faruk, Y. Mori, and K. Kikuchi, "In-band estimation of optical signal-to-noise ratio from equalized signals in digital coherent receivers," *IEEE Photon. J.*, vol. 6, no. 1, Feb. 2014, Art. no. 7800109.

[11] 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. 2017.

[12] L. Jiang *et al.*, "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. 2019.

[13] X. Lin, O. A. Dobre, T. M. N. Ngatched, and C. Li, "A non-data-aided OSNR estimation algorithm for coherent optical fiber communication systems employing multilevel constellations," *J. Lightw. Technol.*, vol. 37, no. 15, pp. 3815–3825, Aug. 2019.

[14] F. N. Khan, Y. Zhou, Q. Sui, and A. P. T. Lau, "Non-data-aided joint bit-rate and modulation format identification for next-generation heterogeneous optical networks," *Opt. Fiber Technol.*, vol. 20, no. 2, pp. 68–74, Jan. 2014.

[15] X. Fan *et al.*, "Joint optical performance monitoring and modulation format/bit-rate identification by CNN-based multi-task learning, " *IEEE Photon. J.*, vol. 10, no. 5, Sep. 2018, Art. no. 7906712.

[16] X. Fan *et al.*, "Feature fusion-based multi-task convnet for simultaneous optical performance monitoring and bit-rate/modulation format identification," *IEEE Access*, vol. 7, pp. 126709–126719, Sep. 2019.

[17] Y. Cheng, S. Fu, M. Tang, and D. Liu, "Multi-task deep neural network (MT-DNN) enabled optical performance monitoring from directly detected PDM-QAM signals," *Opt. Exp.*, vol. 27, no. 13, pp. 19062–19074, Jun. 2019.

[18] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar. "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, Feb. 2016.

[19] T. S. R. Shen, Q. Sui, and A. P. T. Lau, "OSNR monitoring for PM-QPSK systems with large inline chromatic dispersion using artificial neural network technique," *IEEE Photon. Technol. Lett.*, vol. 24, no. 17, pp. 1564–1567, Sep. 2012.

[20] W. Saif, M. Esmail, A. Ragheb, T. Alshawi, and S. Alshebeili, "Machine learning techniques for optical performance monitoring and modulation format identification: A survey," *IEEE Commun. Surv. Tut.*, vol. 22, no. 4, pp. 2839–2882, Sep. 2020.

[21] S. Marsland, *Machine Learning: An Algorithmic Perspective*. Boca Raton, FL, USA: Chapman and Hall/CRC, 2014.

[22] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.

[23] H. Ye *et al.*, "OSNR monitoring based on a low-bandwidth coherent receiver and LSTM classifier". *Opt. Exp.*, vol. 29, no. 2, pp. 1566–1577, Jan. 2021.

[24] ITU-T "Spectral grids for WDM applications: DWDM frequency grid," *Edition 3.0. ITU-T Recommendation G.694.1*, 2020. [Online]. Available: https://www.itu.int/rec/T-REC-G.694.1/en/

[25] ITU-T "Interfaces for the optical transport network," *Edition 6.0. ITU-T Recommendation G.709/Y.1331*, 2020. [Online]. Available: https://www.itu.int/rec/T-REC-G.709/en/

[26] J. Chai *et al.*, "Cost-effective OSNR monitoring with large chromatic dispersion tolerance using random forest for intermediate nodes," *Opt. Commun.*, vol. 479. no. 13, Jan. 2021, Art. no. 126469.

[27] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.

[28] Y. Liu, Y. Wang, and Z. Jian, "New machine learning algorithm: Random forest," in *Proc. Int. Conf. Inf. Comput. Appl.*, Sep. 2012, pp. 246–252.

[29] G. Biau, "Analysis of a random forests model," *J. Mac. Learn. Res.*, vol. 13, no. 2, pp. 1063–1095, May 2010.