SCIENCE CHINA
Information Sciences

• Supplementary File •

Asynchronous amplitude histogram-based multiparameter monitoring for *M*-PAM optical communication channels

Si-Ao LI[1,2], Wenqian ZHAO[1], Yiwen ZHANG[3], Qinru PENG[1], Jian YANG[2], Yuanpeng LIU[2], Runzhou ZHANG[4], Zhongqi PAN[5], and Yang YUE[1,*]

[1]*School of Information and Communications Engineering, Xi'an Jiaotong University, Shaanxi, Xi'an 710049 China*
[2]*Institute of Modern Optics, Nankai University, Tianjin 300350, China*
[3]*Drilling&Production Technology Research Institute, Chuanqing Drilling Engineering Company Limited, Guanghan, Sichuan 618300, China*
[4]*Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA*
[5]*Department of Electrical and Computer Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504 USA*

Received /Revised/Accepted/Published online

Appendix A Introduction

Modern high-speed optical communication has evolved into a multilayered network architecture comprising transmission backbone networks, access networks, and data center interconnections [1-3]. These distinct optical transmission networks have varying high-speed transmission requirements, leading to diverse technical implementations [4-7]. The rapid advancement of cloud computing has substantially increased bandwidth demands for short-reach optical transmission in both data centers (DCs) and access networks, particularly for intra- and inter-data center connections [8-10]. In such scenarios, intensity-modulated direct-detection (IM/DD) systems have been widely adopted due to their low complexity and cost-effectiveness, especially for data center interconnects [11-13].

Figure A1 illustrates the architecture of inter-DC systems based on wavelength-division multiplexed (WDM) IM/DD, incorporating machine learning (ML)-based channel monitoring techniques. Monitoring nodes can be deployed either after the transmitter or before the receiver, enabling the extraction of signal features for processing by various ML algorithms to monitor both signal and channel parameters. Additionally, a software-defined networking (SDN) controller can provide timely feedback to the transmitter, allowing the channel to dynamically adapt to environmental changes and thus enhance overall performance. In IM/DD systems, pulse amplitude modulation (PAM) has emerged as a promising modulation format (MF), with multi-pulse amplitude modulation (*M*-PAM) variants like PAM4 and PAM8 becoming preferred solutions for short-reach high-speed transmission due to their spectral efficiency [14-16]. While PAM4 currently dominates 400 Gb/s data center optical modules, researchers are now exploring higher-order PAM formats with increased baud rates to meet 800 Gb/s and beyond requirements [17].

Optical performance monitoring (OPM) has become a critical technology for the efficient operation and management of dynamic optical networks. By enabling real-time monitoring and analysis of various network parameters, OPM plays a key role in maintaining network stability [18-20], optimizing resource allocation, and facilitating rapid fault localization [21-24]. Traditional OPM approaches often rely on distinct mathematical models for each parameter, making multiparameter monitoring a complex task [23, 25]. Given the increasing complexity of dynamic optical networks, OPM systems must be capable of monitoring not only individual impairments but also multiple interdependent parameters, such as modulation format (MF), baud rate (BR), information rate, and adaptive bandwidth [26-28].

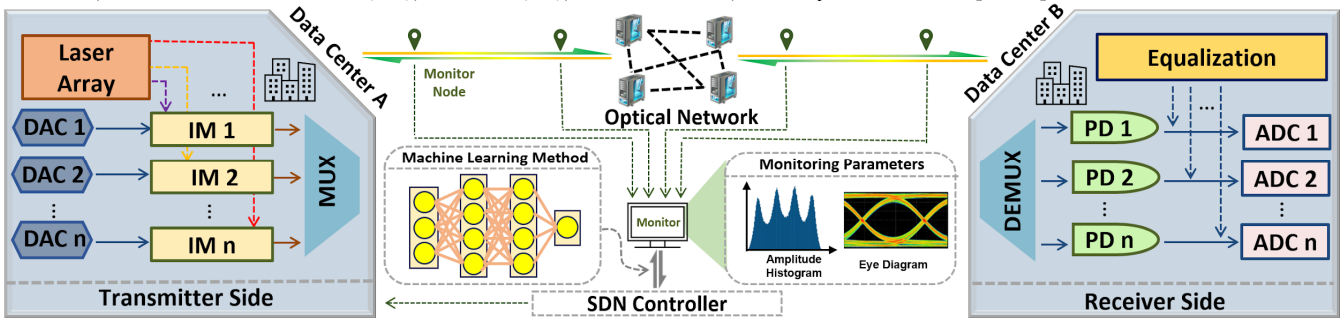

Figure A1 Conceptual diagram of multiparameter performance monitoring of PAM signals in inter-DC systems. DAC: digital-to-analog converter; IM: intensity modulator; PD: photodiode; ADC: analog-to-digital converter; SDN: software-defined networking.

To address these challenges, researchers have turned to machine learning (ML), which offers powerful tools for handling high-dimensional data and uncovering intricate patterns in optical networks. In recent years, ML-based OPM approaches have advanced significantly [29-34]. ML algorithms are particularly effective at feature extraction and learning complex linear and nonlinear parameter relationships [35-38]. Noteworthy developments include the 2019 study by Xia et al., which employed a transfer learning-assisted deep neural network (DNN) method for optical signal-to-noise ratio (OSNR) monitoring, enabling rapid adaptation to changes in various system parameters [39]. Additionally, the 2019 work by Wang et al. proposed an OSNR monitoring scheme with high accuracy and short response time, utilizing a long short-term memory neural network (LSTM-NN) that achieved an error of less than 0.1 dB [40].

Time-domain monitoring methods, which can be classified into synchronous and asynchronous sampling techniques, have also seen significant advancements. While synchronous sampling methods, such as amplitude histograms, require clock recovery, asynchronous approaches—such as asynchronous amplitude histogram (AAH), asynchronous delay-tap sampling (ADTS), and asynchronous single-channel sampling (ASCS)—have gained popularity for multi-parameter monitoring due to their clock-independent operation. Notable studies include the 2015 work by Khan et al., which experimentally demonstrated simultaneous OSNR monitoring and modulation format identification (MFI) in heterogeneous fiber-optic networks. This was achieved using

principal component analysis (PCA) and statistical distance-based pattern recognition applied to scatter plots obtained through ASCS [41]. The 2021 study by Luo et al., which proposed the ADTP/AAH method [42], and the work by Chai et al., which utilized low-bandwidth coherent detection with random forest algorithms [43]. A systematic comparison of these sampling-based monitoring methods is summarized in Table A1.

Table A1 Comparison of sampling-based OPM methods using ML

Aspect	[40] Wang 2019	[41] Khan 2015	[42] Luo 2021	[43] Chai 2021	[49] Cheng 2023	[48] Li 2024	This work
Sampling	Sync. (Coherent)	Async. (ASCS)	Async. (DD)	Async. (Low-BW Coh.)	Async. (DD)	Async. (DD, AAH)	Async. (DD, AAH)
ML/DL Model	LSTM-NN	PCA+Stat. Dist.	Multi-task DNN	Random Forest	MLP/CNN	GBDT	LightGBM
# Params	1	2	4	3	4	4(+CD)	6
Max Classes	N/A (Regr.)	432 (scatter)	Indep. tasks	8/6	Indep. tasks	N/A	1,728
Accuracy	MAE 0.05dB	MFI 98.5%	MFI 100% BRI 99.8%	SR-MFI 100%	MFI 99.4% BRI 98.3%	4-param 97.5%	6-param 93.99%
Clock Recov.	Yes	No	No	No	No	No	No
Sample Rate	High-speed ADC	Low-speed (N/A)	80 GSa/s	1 GSa/s (sim)	DSO (N/A)	N/A	12.5 MSa/s
Signal	PDM-QAM	NRZ-OOK /DPSK	QPSK /16QAM	PDM-QAM	PDM-QAM	PAM 4/6/8	PAM 3/4/6/8

More recently, advanced deep learning architectures have been increasingly explored for OPM. Jiang et al. proposed the first application of Vision Transformer (ViT) combined with fractional Fourier transform (FrFT) for joint estimation of OSNR, SNR, CD, and DGD, achieving MAEs of 0.058 dB, 0.091 dB, 117 ps/nm, and 0.38 ps, respectively, outperforming CNN and DNN baselines [44]. Yang et al. developed an improved multi-task MIMO-ResNet (IMT-MIMO-ResNet) scheme for multi-mode multi-parameter joint OPM in MDM-EON systems, achieving 100% accuracy for baud rate and modulation format identification with computational complexity approximately 1/7 of typical approaches [45]. Lin et al. introduced a physics-informed neural network (PINN) that embeds the nonlinear Schrödinger equation into the loss function for interpretable and nonlinear-tolerant OSNR monitoring, achieving average errors of 0.13–0.26 dB [46]. Kulandaivel and Jeyachitra proposed a dual-stage deep learning model combining CNN-based parameter identification with ANN-based optical modulation index estimation for radio-over-fiber systems, achieving 99.26% multiparameter classification accuracy [47]. These developments highlight the trend toward more sophisticated DL architectures for OPM; however, they generally require high-speed coherent receivers, GPU-accelerated training, or complex preprocessing pipelines. In contrast, the present work focuses on a lightweight, hardware-efficient approach using LightGBM with asynchronous amplitude histograms obtained from a low-cost, low-sampling-rate oscilloscope.

This study proposes a joint monitoring scheme for multi-level PAM signals by integrating AAH analysis with Light Gradient Boosting Machine (LightGBM), enhanced by a cross-correlation and sliding window-based preprocessing strategy. Compared to the initial four-parameter demonstration using GBDT with moving average preprocessing in our preliminary conference work [48], this work introduces a novel AXcorr preprocessing algorithm, expands the monitored parameter space to 6 parameters covering 1,728 classes, and incorporates systematic hyperparameter optimization, progressive evaluation, and accuracy stability analysis with variance quantification. Experimental results using direct detection at a sampling rate of 12.5 MSa/s demonstrate over 93% accuracy across six single-parameter and five multi-parameter classification tasks, including BR, chromatic dispersion (CD), MF, probabilistic shaping (PS), OSNR, and roll-off factor (ROF). The proposed preprocessing method consistently enhances accuracy, achieving 93.99% in a complex six-parameter joint monitoring scenario involving 1,728 classes—representing a 27% improvement over the baseline AAH approach. Additionally, this work offers practical guidance on selecting oscilloscope sampling points and rates, providing a low-cost, transmitter-oriented monitoring solution [48-51].

For practical deployment, as shown in Fig. A1, the proposed monitoring method can be reliably placed at either the transmitter side or the receiver side. When deployed at the transmitter, it enables early-stage signal quality validation, preventing invalid signals from being injected into the network. Moreover, it supports real-time link adaptation and SDN-controlled dynamic adjustment at the source side, such as OSNR optimization and modulation reconfiguration. It is also feasible and effective at receiver side. Since the AAH captures statistical amplitude features that remain stable through transmission, the proposed method can still accurately identify all key parameters at the receiver side, which supports end-to-end channel performance evaluation, including accumulated CD, effective OSNR, and link-induced signal degradation. It is compatible with conventional IM/DD receivers without modifying the front end.

Appendix B Experimental setup and methods

In optical communication signal analysis and performance testing, the AAH serves as a low-cost statistical method for characterizing the frequency and distribution of signal amplitude occurrences [22, 52]. In fiber-optic communication systems, the AAH is highly sensitive to the OSNR and to transmission impairments such as CD and polarization mode dispersion (PMD). By analyzing the occurrence frequency of each amplitude level, AAH can also reveal key signal characteristics, including MF, ROF, and BR. The AAH generation process involves direct detection, real-time sampling via an oscilloscope, amplitude-level statistics, and histogram construction, as illustrated in Fig. B1. Having described the principle of AAH data generation, the following sections present two preprocessing algorithms — Sliding Window (SW) smoothing and Averaged-Cross-Correlation (AXcorr) alignment — that are applied to the AAH data to enhance feature quality before classification.

Physically, chromatic dispersion (CD) causes different frequency components of the optical signal to propagate at different velocities, resulting in temporal broadening of transmitted pulses. For intensity-modulated PAM signals, increasing CD leads to inter-symbol interference (ISI) that blurs the originally well-separated amplitude levels. This manifests in the AAH as broadening and flattening of individual peaks, filling-in of valleys between peaks, and a gradual transition toward a smoother distribution. These CD-dependent shape variations, visible in Fig. C3 (CD panel), constitute the discriminative features that enable the classifier to identify different CD values.

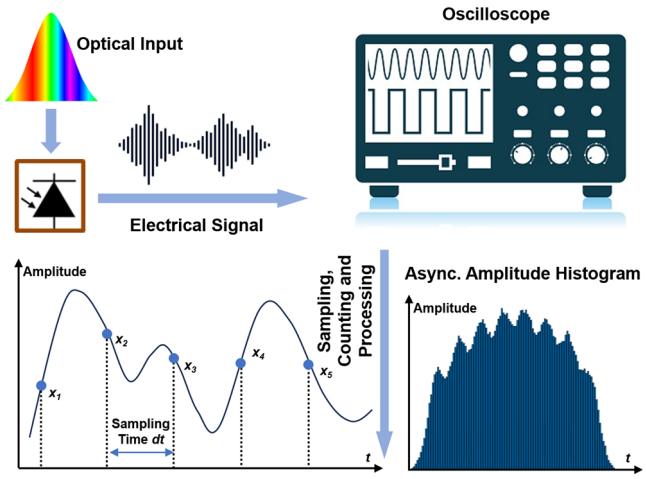

Figure B1 Schematic diagram of the AAH generation process.

The frequency response of the oscilloscope can significantly affect the accuracy of the AAH. To address this, a sliding window (SW) smoothing algorithm is applied directly to the AAH data obtained at the oscilloscope's low sampling rate, thereby reducing processing complexity.

The SW smoothing algorithm is a sequential data processing technique that reduces noise by applying a fixed-size window with a constant step across the data stream [53, 54]. Its key parameters include window length, sliding step size, and the specific smoothing algorithm employed. Fig. 3 presents a comparison of AAH data for a 25 GBd PAM4 signal, sampled using a 50 GSa/s oscilloscope with 500,000 sampling points, after applying various smoothing algorithms. Fig. 3(a) shows the original AAH data measured with an oscilloscope bandwidth of 33 GHz. Fig. B2(b) and Fig. B2(c) depict the AAH results after the sampled signals are processed using a second-order Super-Gaussian (SG) digital filter at bandwidths of 50 GHz and 30 GHz, respectively. Notably, the AAH filtered at 30 GHz exhibits substantial smoothing. Fig. B2(d), Fig B2(e), and Fig. B2(f) display the AAH results smoothed by the simple moving average (SMA), locally weighted scatterplot smoothing (Lowess), and Savitzky–Golay (Sgolay) filters, respectively, using a window length of 5 and a sliding step size of 1. The results indicate that SMA and Lowess provide slightly better smoothing than Sgolay. While Lowess produces the smoothest AAH curves, it may obscure high-frequency features. In contrast, the SMA algorithm offers lower computational complexity and achieves a smoothing effect comparable to that of the 30 GHz SG filter, preserving some high-frequency details while effectively reducing noise. Note: In this work, "SG" refers exclusively to the Super-Gaussian digital filter used for frequency-domain bandwidth limitation, while "Sgolay" denotes the Savitzky–Golay filter used for time-domain smoothing.

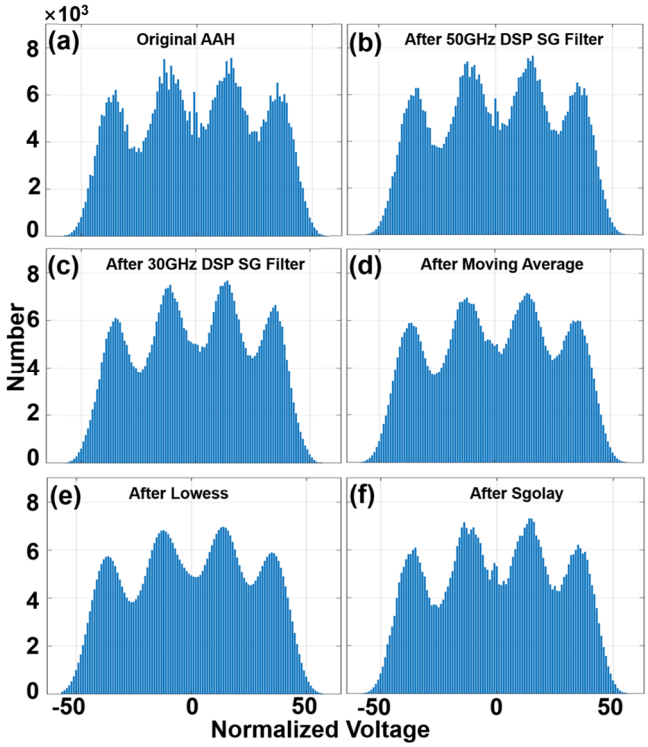

Figure B2 Comparison of the effects of (a) original AAH generated from the sampled signal of 25 GBd PAM4, and AAH generated after (b) 50 GHz DSP filtering, (c) 30 GHz DSP filtering, and raw AAH after (d) SMA smoothing, (e) Lowess smoothing, and (f) Sgolay smoothing (the oscilloscope sampling rate is 50 GSa/s).

When the Savitzky–Golay (SG) filter is used for preprocessing, it is necessary to perform an n-point fast Fourier transform (FFT) on the 500,000 sampling points to convert the data into the frequency domain, followed by an inverse fast Fourier transform (IFFT) to revert to the time domain. In contrast, the SMA method does not require transforming all of the sampled signals. Instead, it directly processes the sliding mean of the 200-length AAH, significantly reducing preprocessing complexity. For this reason, SMA is selected as the smoothing algorithm in the SW method.

Power fluctuations in fiber optic transmission can adversely affect the stability of communication systems. These fluctuations refer to the random variations in the optical signal's power level during transmission, often caused by impairments such as output power jitter at the optical transmitter, multipath transmission interference, and instability in optical amplifier gain [55-58]. Such power variations can lead to jitter in the received peak-to-peak voltage (Vpp), resulting in slight random shifts in the amplitude of the AAH within the same dataset. This phenomenon causes feature misalignment within the same class, particularly when the differences between feature vectors of different classes are small, thereby reducing the accuracy and resolution of classification-based monitoring. To address this issue, an Averaged-Cross-Correlation (Axcorr) data preprocessing method is proposed, which aligns features within the same dataset and improves detection accuracy.

First, all AAH data within the same class are averaged across each bin, as shown in the corresponding equation, to obtain the average AAH feature vector, denoted as $x_{averaged}$.

$$x_{averaged}[n] = \frac{1}{L}\sum_{i=1}^{L} x_i[n] \tag{1}$$

where L denotes the number of data within the dataset. Then all the rest of the AAH data is cross correlated with $x_{averaged}$ as in equation, thus obtaining R.

$$R_{averaged,i}[m] = \sum_{m=1}^{N} x_{averaged}[n] \cdot x_i^*[n+m] \tag{2}$$
$$m \in [-N, N-1]$$

where m denotes the amount of data shift and N denotes the length of the AAH. The position M_i could be found corresponding to the maximum value according to the R. Thus, the corresponding data displacement S_i for the i-th group could be expressed as:

$$M_i = argmax(R_{averaged,i}[m]) \qquad S_i = N - M_i \tag{3}$$

$$x_{shifted}[n] = x[n + S_i] \tag{4}$$

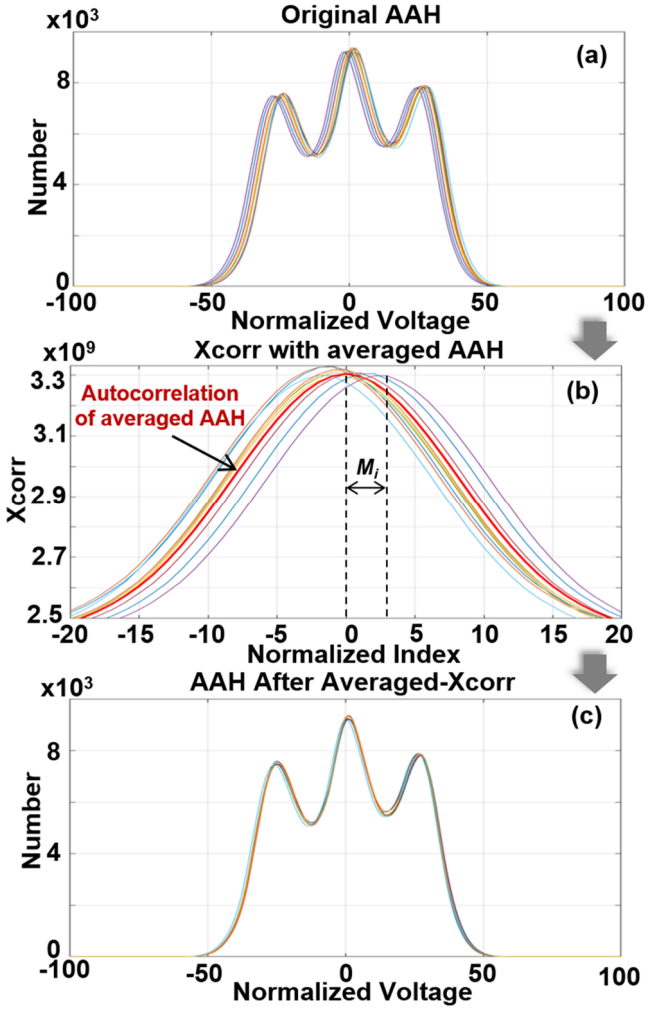

Figure B3 Schematic of the algorithmic flow of AXcorr and the effect of the processed PAM3 AAH.

Figure B3 illustrates the effect of applying the Axcorr preprocessing method to original PAM3 AAH data. As shown in Fig. B3(a), the unprocessed AAH exhibits misalignment and amplitude offsets. Fig. B3(b) presents the cross-correlation results between each AAH sample and the averaged AAH. In this figure, M_i denotes the offset between the peak of each cross-correlation curve and the peak of the autocorrelation curve. After applying the Axcorr method, the aligned AAH data are shown in Fig. B3(c), where the positions of the AAH samples are largely corrected and aligned.

In practical optical transmitter parameter monitoring, it is desirable for the ML model to be both fast and lightweight while

maintaining high accuracy. To meet these requirements, the LightGBM model—an efficient implementation of the gradient boosting decision tree (GBDT) algorithm is adopted for multi-parameter monitoring [59]. LightGBM employs a histogram-based decision tree growth strategy, enabling fast training and low resource consumption. By integrating the two aforementioned AAH preprocessing techniques, an Axcorr + SW-based AAH data preprocessing and LightGBM-based multi-parameter joint monitoring and classification framework is proposed, as illustrated in Fig. B4.

The choice of LightGBM over alternative classifiers (e.g., CNN, LSTM, Random Forest) is motivated by several factors: (1) the AAH input is a compact 1-D tabular vector (200 features), for which gradient boosting decision trees are inherently well-suited, whereas CNN and LSTM are primarily designed for 2-D spatial or sequential data; (2) LightGBM's histogram-based tree growth strategy enables fast training without GPU acceleration, consistent with the low-cost hardware philosophy; (3) the relatively small dataset (100 samples per class) favors ensemble tree methods over deep learning models that typically require larger training sets; and (4) LightGBM's boosting strategy generally yields higher accuracy than bagging-based ensembles such as Random Forest on structured classification tasks.

The sampled PAM signals are first used to statistically generate the original AAHs. These AAH data are initially processed using the Axcorr algorithm to mitigate the effects of power fluctuations and to align feature distributions. Subsequently, the data are smoothed using the SW algorithm to reduce the influence of oscilloscope hardware bandwidth limitations. The resulting preprocessed AAH feature dataset is then divided into training and testing sets and input into a multi-class LightGBM model for classification. The number of output categories corresponds to the total number of parameter combinations targeted for monitoring.

Multiple signal processing steps are performed at the transmitter side to generate *M*-PAM digital signals with varying parameter combinations. These steps include PS, higher-order PAM symbol mapping, two-fold up sampling, root-raised cosine (RRC) roll-off filtering, and frequency-domain equalization (FDE). In the PS process, a specified achievable information rate (AIR) is set, and the shaping coefficient λ is determined according to the Maxwell-Boltzmann probability distribution corresponding to the modulation order. Finally, *M*-PAM signals with a sequence length of 32,768 symbols are generated using the constant composition distribution matching (CCDM) scheme.

$$AIR = N \times \log_2 M \tag{5}$$

$$P_X(x) = \frac{e^{-\lambda x^2}}{\sum_{x' \in X} e^{-\lambda x'^2}} \quad \left(X = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}\right) \tag{6}$$

In the experiments, the symbol *AIR* is changed by varying *N* according to equation *M* is the PAM modulation order, and *N* is referred as the PS coefficient, which is one of the important parameters to be monitored. When *N*=1, it corresponds to the *M*-PAM at uniform distribution. Before Nyquist shaping, the generated higher-order PS-PAM signal needs to be upsampled. Here, a 2× upsampling is used, i.e., 1 zero sample point is inserted between neighboring samples. The upsampled PS-PAM signal will be RRC roll-off filtered to achieve Nyquist shaping, and the ROF is another parameter to be monitored, which is set to 0.1, 0.2, and 0.5. In order to get a better digital-to-analog converter (DAC) output signal at each BR, the bandwidth limitation of the DAC is compensated by performing FDE. Finally, the pre-processed *M*-PAM samples are fed into the DAC with 92 GSa/s sampling rate to obtain the output electrical signal.

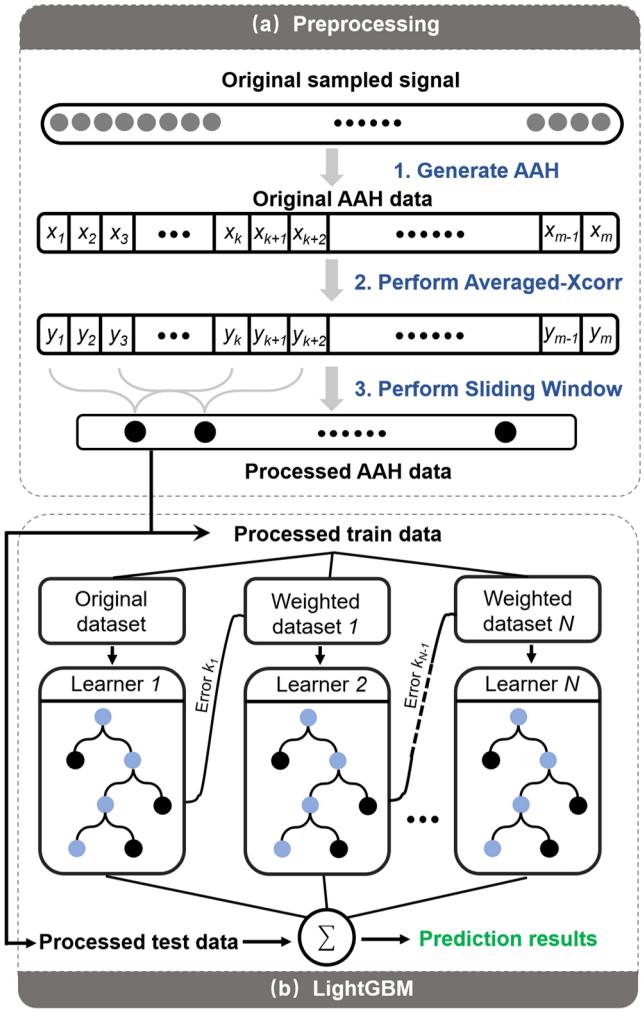

Figure B4 Multi-parameter joint classification scheme based on AAH: (a) AXcorr+SW combined data preprocessing method, (b) LightGBM model for multi-classification tasks Section III. Experimental testing and analysis of the system.

The experimental transmission link, illustrated in Fig. B5, is designed to capture sampled signals of *M*-PAM under varying parameter configurations. A narrow-linewidth laser source with a center wavelength of 1550 nm and an output power of 13 dBm serves as the optical carrier. The bipolar multilevel electrical signals generated by the DAC are modulated onto the optical carrier using an intensity modulator. The modulator's DC bias voltage is automatically adjusted to the quadrature point of the Mach–Zehnder modulator (MZM) transfer function to ensure optimal modulation. The DAC output has a peak-to-peak voltage (Vpp) of 350 mV, which is appropriately matched to the linear operating range of the modulator.

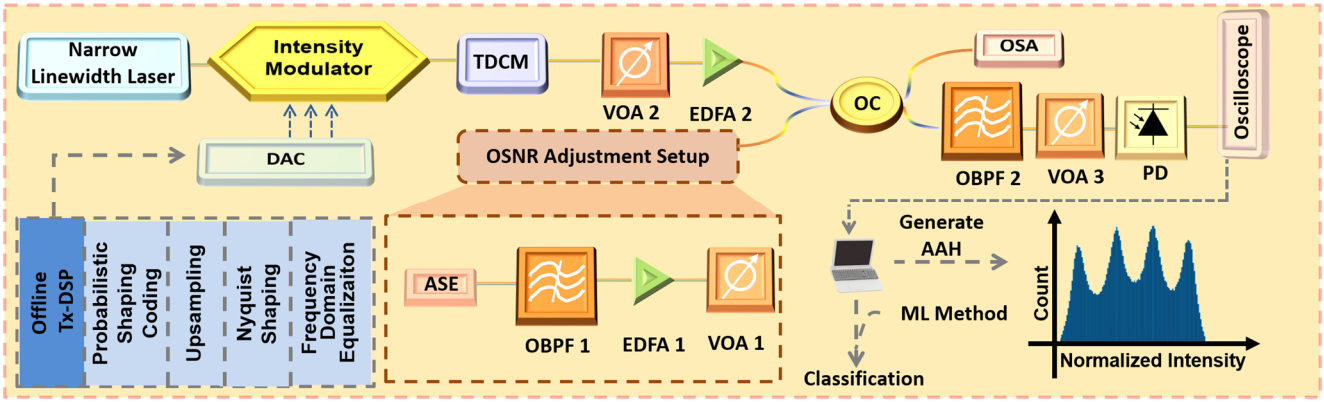

Figure B5 *M*-PAM transmission system experimental link (TBPF: tunable bandpass filter; OC: optical coupler; VOA: variable optical attenuator; TDCM: tunable dispersion module; PD: photodiode; OSA: optical spectrum analyzer).

The modulated optical signal is transmitted through a tunable dispersion compensation module (TDCM), enabling the adjustment of CD to emulate accumulated dispersion effects along the transmission path. Additionally, broadband optical noise, generated by an amplified spontaneous emission (ASE) source, is filtered, amplified, and then attenuated before being combined with the optical signal at a 3-dB coupling ratio to simulate various OSNR conditions.

The combined optical signal is split equally (50:50) into two paths. One path is used for OSNR measurement, while the other is filtered by an optical bandpass filter to suppress the out-of-band noise, then attenuated using a 1-dB variable optical attenuator, yielding an output optical power of approximately 3 dBm. This signal is subsequently detected and sampled by a real-time oscilloscope following photodetection. The settings of the main experimental instruments and the parameters subject to monitoring are summarized in Table B1, where "PS" denotes the PS coefficient *N*, as defined previously.

Table B1 Settings of the *M*-PAM experimental transmission system.

Parameters	Settings
Tx Laser	Power: 13 dBm, Wavelength: 1550 nm
DAC	Vpp: 350 mV
Oscilloscope	Bandwidth: 33 GHz
Bandpass Filter	Bandwidth: 0.27 nm
Attenuator	1 dB
Monitoring Parameters	MF: [PAM3, PAM4, PAM6, PAM8] ROF: [0.1, 0.2, 0.5] PS: [0.8, 0.9, 1] OSNR: [25, 30, 35, 40] (dB) CD: [0, -40, -80, -120] (ps/nm) BR: [25, 26, 27] (GBd)

Appendix C Result

In the LightGBM model, hyperparameters significantly influence classification accuracy and must be carefully tuned. The maximum depth parameter limits the maximum depth of each decision tree and should be set appropriately to avoid overfitting. The number of leaves parameter specifies the maximum number of leaves in a tree and is also critical to controlling model complexity. During the early stages of training, a relatively high learning rate (LR), such as 0.1, can be used to accelerate convergence. Once the other parameters are fixed, the learning rate may be reduced to improve performance. The number of iterations parameter determines the number of boosting rounds. A larger number of iterations generally leads to a more converged loss function, but also increases training time. In practice, achieving higher accuracy typically requires using a smaller LR in combination with a larger number of iterations.

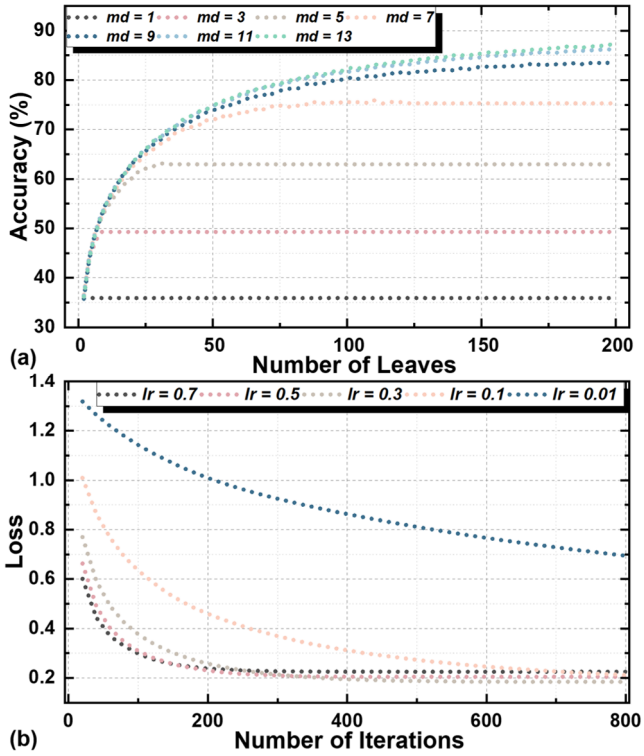

Figure C1 Effect of (a) max depth and number of leaves, (b) learning rate and number of iterations on classification accuracy for CD monitoring in LightGBM.

Initially, max depth and number of leaves are tuned while fixing the learning rate (LR) at 0.1 and the number of iterations at 100. For each max depth in the range [1, 13] and each number of leaves in the range [2, 200], classification accuracy is evaluated under all corresponding parameter combinations. As illustrated in Fig. C1(a), higher values of max depth generally lead to higher peak accuracies for different values of the number of leaves. Moreover, the accuracy curves for max depth = 11 and 13 largely overlap. Additionally, when the number of leaves > 110, accuracy improvements begin to plateau. Considering both training efficiency and classification accuracy, the optimal parameters are selected as max depth = 11 and number of leaves = 110.

Next, the learning rate and number of iterations are optimized with max depth and number of leaves fixed. For each LR value in the set [0.01, 0.1, 0.3, 0.5, 0.7], the number of iterations is varied from 20 to 800 in increments of 5, and the corresponding loss values are recorded, as shown in Fig. C1(b). Results indicate that larger learning rates result in faster convergence. In particular, for LR = 0.3, the loss stabilizes when the number of iterations exceeds 350.

For different classification tasks, achieving an efficient and accurate training model requires sequential tuning of parameters as previously described. In this section, the experiment further optimizes various combinations of LightGBM model parameters according to the specific requirements of each classification task. The optimized parameter settings for each task are summarized in Table C1. It should be noted that the hyperparameter tuning described above is performed offline as a one-time process during the model development phase. Once the optimal parameters are determined for each monitoring task, they are fixed and directly applied during inference. The tuning cost therefore does not affect the operational (inference-time) complexity of the monitoring system.

Table C1 Optimization parameters of LightGBM model corresponding to different monitoring classification tasks.

Monitoring tasks	Max depth	Num of leaves	Learning rate	Num of iterations
CD	11	110	0.3	350
ROF	7	36	0.5	80
OSNR	7	40	0.5	100
MF	7	30	0.5	40
PS	7	30	0.5	40
BR	11	100	0.2	400
ROF&PS&OSNR	7	30	0.1	100
ROF&PS&OSNR&CD	9	40	0.1	90
ROF&PS&OSNR&BR	9	40	0.1	95
ROF&PS&OSNR&BR&MF	7	40	0.1	110

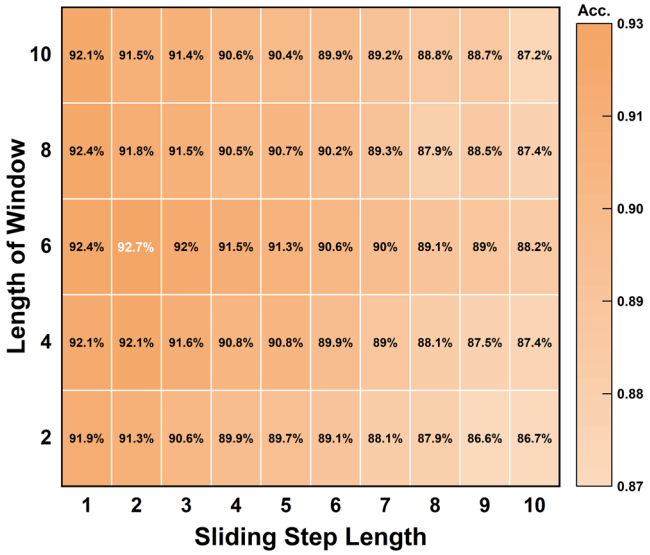

Figure C2 Effect of SW length and sliding step size on classification accuracy of CD monitoring in SMA algorithm.

The window length and sliding step size of the SW algorithm are two critical parameters that must be selected carefully. In the experiments, a grid search is conducted across combinations of window lengths and sliding step sizes to evaluate their impact on monitoring accuracy, as illustrated in Fig. C2. Specifically, the sliding step size is varied from 1 to 10 in increments of 1, while the window length ranged from 2 to 10 in increments of 2. The results indicate that the highest classification accuracy is achieved when the sliding step size is 2 and the window length is 6. Since the SMA algorithm serves as a preprocessing method for AAH feature extraction and does not influence the tuning of LightGBM model parameters, these SW settings can be consistently applied across all classification tasks. Therefore, a sliding step size of 2 and a window length of 6 are adopted for AAH data preprocessing in this work.

The rationale behind these trends is as follows: a small window length (e.g., 2) provides insufficient smoothing and retains excessive noise, while a large window (e.g., 10) risks over-smoothing that blurs discriminative amplitude features between parameter classes. Similarly, a smaller sliding step retains more data points and preserves feature resolution, whereas a large step aggressively reduces dimensionality. The optimal combination (window length = 6, step = 2) achieves the best balance between noise suppression and feature preservation.

In the experiment, each parameter is adjusted according to Table C1, and sampled signals are collected for all class combinations of six parameters: ROF, BR, OSNR, PS, CD, and MF. Amplitude statistics are then applied to generate the original AAH dataset. The oscilloscope sampling rate is set to 12.5 MSa/s, and 500,000 sample points are acquired per signal. For each parameter combination, 100 groups of sampled signals are collected. The number of bins for the AAH is set to 200, resulting in a dataset comprising 1,728 × 100 data samples in total. The LightGBM model is employed for classification tasks, with specific model parameter settings for each task summarized in Table 2. Examples of *M*-PAM AAH under selected parameter combinations obtained from the experiment are illustrated in Fig. C3 and Fig. C4.

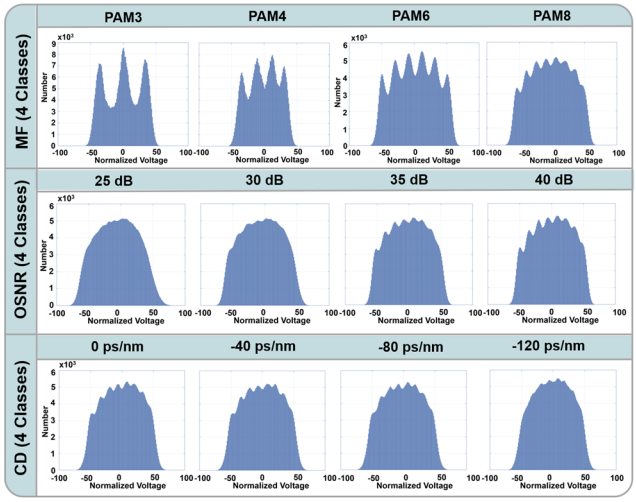

Figure C3 AAHs corresponding to *M*-PAM signals with different MFs, BRs, PS, ROFs, OSNRs, and CD in the experiment.

To evaluate the performance of multi-parameter monitoring, it is essential to first examine single-parameter monitoring tasks. This approach facilitates understanding of how each individual parameter affects monitoring accuracy. In these tasks, the AAH dataset is divided based on the number of classes associated with each parameter. For each of the 1,728 parameter combination classes, 70 out of 100 data samples are randomly selected for training, and the remaining 30 are used for testing, ensuring a stratified split that guarantees balanced class representation in both sets. Notably, during single-parameter monitoring, the other five parameters are randomly varied rather than held constant. This dataset division strategy allows for an assessment of the method's robustness against variations in non-target parameters. To ensure the stability of accuracy evaluations, the monitoring accuracy for each task is averaged over 30 independent trials.

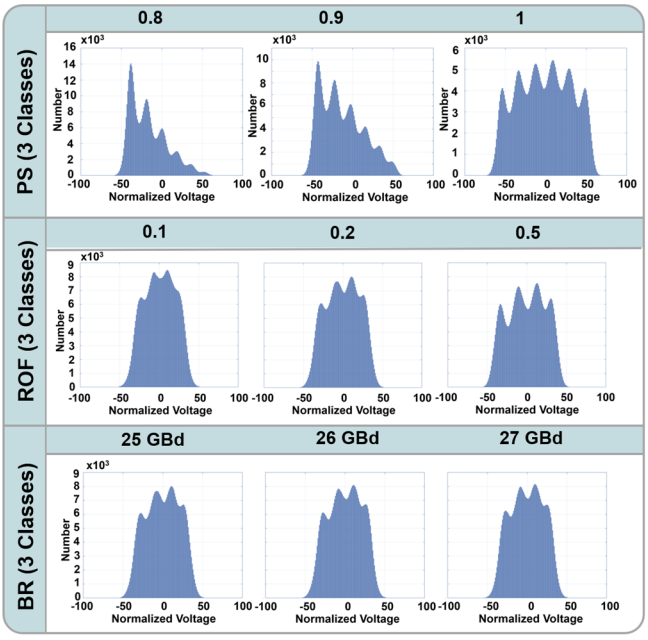

Figure C4 AAHs corresponding to *M*-PAM signals with different, BRs, PS and ROFs in the experiment.

The classification accuracy of LightGBM is compared between using the original AAH data and AAH data preprocessed with different schemes as input features. The results for each single-parameter classification task are presented in Fig. C5(a). Using the original AAH data, high classification accuracy is achieved for MF, PS, ROF, and OSNR; however, the performance for CD and BR classification is suboptimal, with accuracies falling below 90%. After applying either AXcorr or SW preprocessing, accuracy improves across all tasks. In these experiments, the SW method employs the SMA algorithm with a window length of 6 and a sliding step of 2. Notably, AXcorr significantly enhances the accuracy of CD and BR classification. When both AXcorr and SW preprocessing are applied together, the model's performance improves further, achieving near 100% accuracy across all six classification tasks.

The effects of AXcorr and SW are characterized as approximately additive (independent) rather than synergistic, as they address two distinct sources of AAH feature degradation. AXcorr is the dominant contributor, correcting feature misalignment caused by optical power fluctuations — for example, elevating CD accuracy from ~85% to ~98% and BR accuracy from ~84% to ~98% (Fig. C5(a)). SW provides a modest but consistent additional gain (~2–3%) through general-purpose noise suppression. The incremental gain from adding SW on top of AXcorr is comparable in magnitude to the gain from applying SW alone on the original AAH, confirming their independent contributions. This pattern is consistently observed across four-parameter (Figs. C6(a), C7(a)), five-parameter (Fig. C8(b)), and six-parameter (Fig. C9(a)) joint monitoring results.

To further investigate the tasks with lower baseline accuracy (CD and BR), their classification performance is evaluated across different *M*-PAM formats, as shown in Fig. C5(b) and Fig. C5(c). In the BR classification task, the accuracy using original AAH data declines by 3.5% from PAM3 to PAM8; for the CD task, the decrease is 7.12%. After applying AXcorr and SW preprocessing,

the decline in classification accuracy for higher-order PAM formats is significantly mitigated, with differences reduced to 0.27% and 1.26% for BR and CD, respectively. These results demonstrate that the proposed preprocessing scheme not only enhances classification accuracy in general but also effectively reduces performance degradation in higher-order modulation formats.

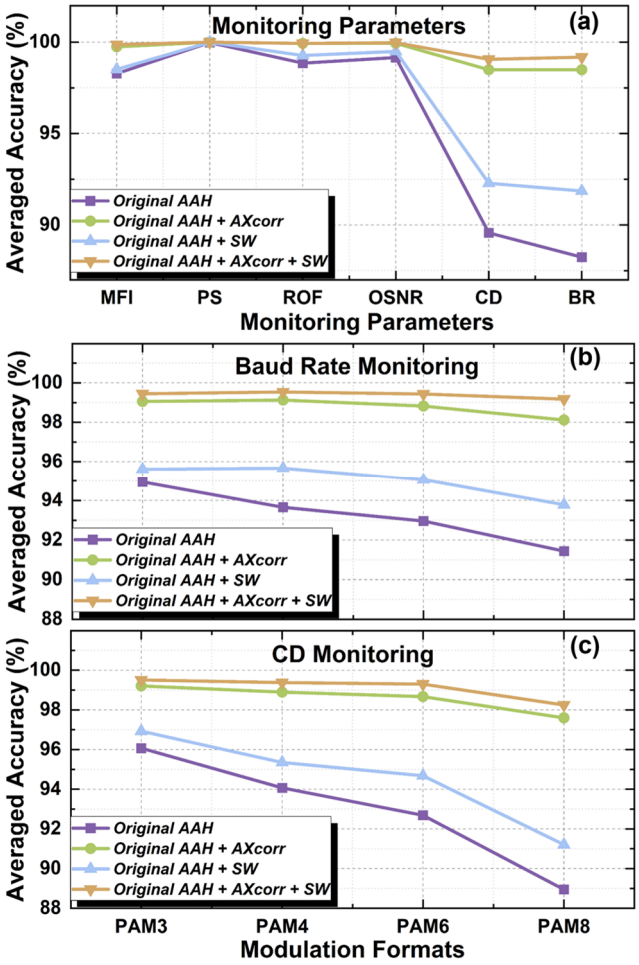

Figure C5 Comparison of classification accuracy using original AAH and preprocessed AAH in monitoring tasks.

Table C2 Accuracy of single-parameter classification of CD and BR for different MFs and OSNRs.

Monitoring tasks	OSNR (dB)	Used AAH data		PAM3	PAM4	PAM6	PAM8
		original	AXcorr + SW preprocessed				
BR Monitoring	25	√		92.08%	91.10%	89.87%	85.74%
	30	√		95.49%	93.43%	93.07%	92.37%
	35	√		96.43%	96.08%	95.13%	94.49%
	40	√		97.06%	96.89%	96.36%	94.99%
	25		√	99.01%	98.93%	98.62%	97.57%
	30		√	99.44%	99.40%	99.37%	99.38%
	35		√	99.44%	99.51%	99.41%	99.42%
	40		√	99.66%	99.63%	99.65%	99.41%
CD Monitoring	25	√		93.41%	90.57%	88.98%	85.54%
	30	√		96.31%	94.02%	91.91%	91.06%
	35	√		97.40%	96.42%	95.84%	92.14%
	40	√		97.96%	96.62%	96.48%	92.71%
	25		√	99.28%	99.07%	98.98%	97.30%
	30		√	99.38%	99.21%	99.20%	98.52%
	35		√	99.55%	99.47%	99.42%	98.65%
	40		√	99.68%	99.63%	99.61%	99.12%

To more clearly investigate the factors affecting classification accuracy in CD and BR monitoring tasks, the experiment evaluates the average classification accuracy of each *M*-PAM format over 30 test runs, under four OSNR conditions: 25 dB, 30 dB, 35 dB, and 40 dB, as shown in Table C2. Results indicate that higher-order *M*-PAM formats are more susceptible to misclassification under lower OSNR conditions, primarily due to their denser amplitude levels. When using original AAH data as

input, the average accuracy gap between 25 dB and 40 dB OSNR for PAM8 reaches 7.17% and 9.25% in CD and BR classification tasks, respectively. Considering all PAM orders together, the maximum accuracy deviation across different OSNR conditions reaches 12.42% for CD and 11.32% for BR.

These findings highlight that both OSNR and modulation format significantly influence classification performance in CD and BR monitoring. However, when AAH data are preprocessed using the proposed AXcorr + SW scheme, classification accuracy improves consistently across all OSNR values. More importantly, the maximum accuracy differences are substantially reduced to 2.38% for CD and 0.28% for BR. This improvement is particularly notable for higher-order *M*-PAM formats under low OSNR, demonstrating the robustness of the proposed preprocessing approach against both OSNR fluctuations and modulation complexity.

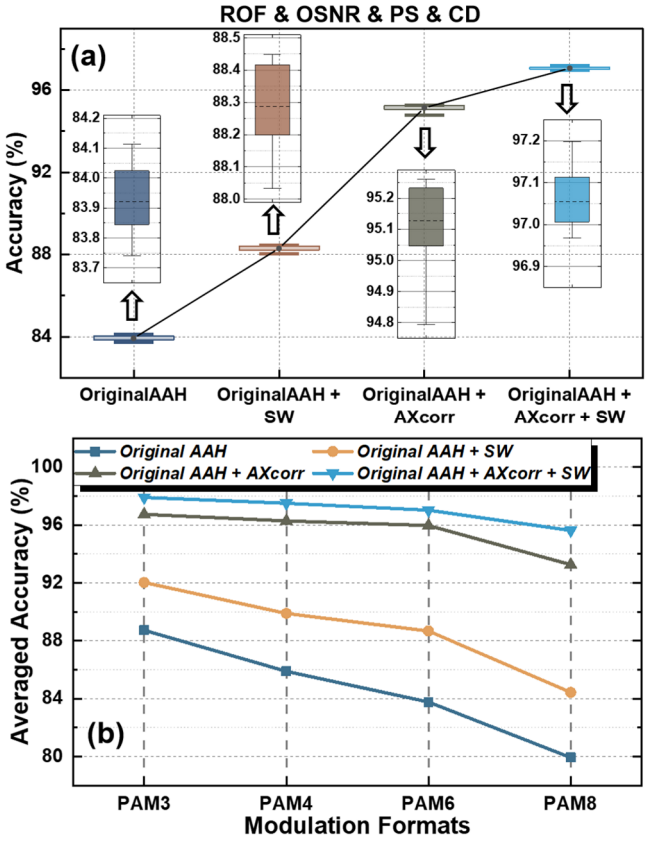

Figure C6 (a) Distribution of the four-parameter joint classification accuracies of ROF, PS, OSNR, and CD and (b) the average classification accuracies in different MFs.

Building upon the results of single-parameter monitoring, this study further investigates joint monitoring tasks involving four parameters. Since CD and BR exhibited relatively low classification accuracy in single-parameter scenarios, two types of four-parameter joint monitoring tasks are designed to evaluate performance under more complex conditions: (1) ROF, PS, OSNR, and CD, encompassing 144 distinct categories; and (2) ROF, PS, OSNR, and BR, encompassing 108 categories. The classification accuracy distributions corresponding to different AAH data preprocessing schemes for these tasks are shown in Fig. C6(a) and Fig. C7(a), respectively. In each box plot, the dashed line represents the average classification accuracy across all test cases.

When the original AAH data is used without preprocessing, the classification accuracy in the four-parameter monitoring tasks significantly declines, particularly due to the inclusion of CD or BR, which are more challenging to classify. Introducing the AXcorr preprocessing alone yields a notable accuracy improvement of approximately 12% for both monitoring tasks, demonstrating that the AXcorr algorithm effectively aligns feature regions within the AAH dataset and enhances classification performance. When AXcorr is further combined with SW smoothing, an additional accuracy gain of around 2% is achieved. As a result, both four-parameter joint monitoring tasks attain high classification accuracies exceeding 97%, confirming the robustness and effectiveness of the AXcorr + SW preprocessing scheme in handling complex multi-parameter classification scenarios.

Furthermore, based on the accuracy distributions shown in Fig. C6(a) and Fig. C7(a), the AXcorr + SW preprocessing not only elevates mean classification accuracy but also enhances stability, with accuracy fluctuations below 0.25% for both tasks. Additionally, Fig. C6(b) and Fig. C7(b) present the mean classification accuracies of the two four-parameter monitoring tasks across various *M*-PAM formats. These results illustrate a marked decline in accuracy with increasing PAM order; however, following AAH data preprocessing, classification accuracies across all formats rise to approximately 97%, effectively mitigating the performance degradation associated with higher-order modulation.

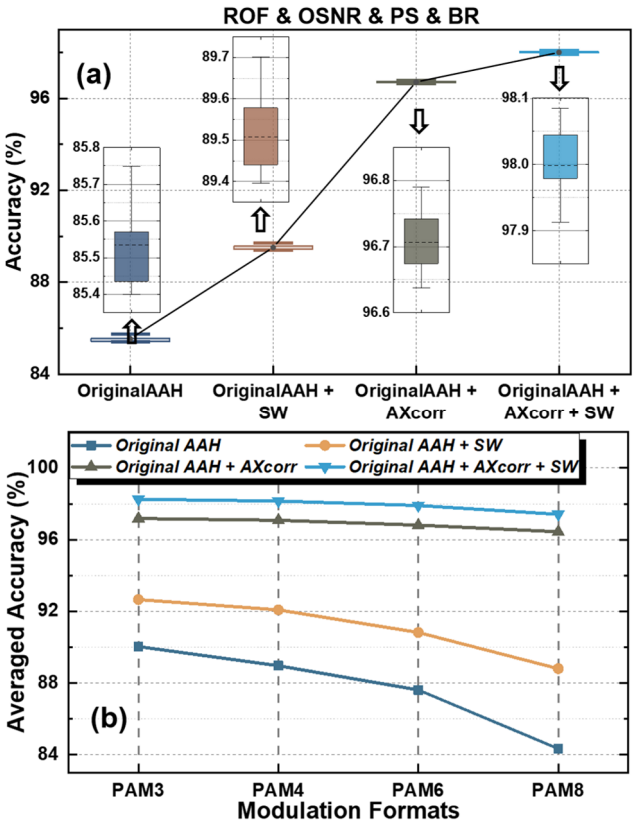

Figure C7 (a) distribution of the four-parameter joint classification accuracies of ROF, PS, OSNR, and BR and (b) the average classification accuracies in different MFs.

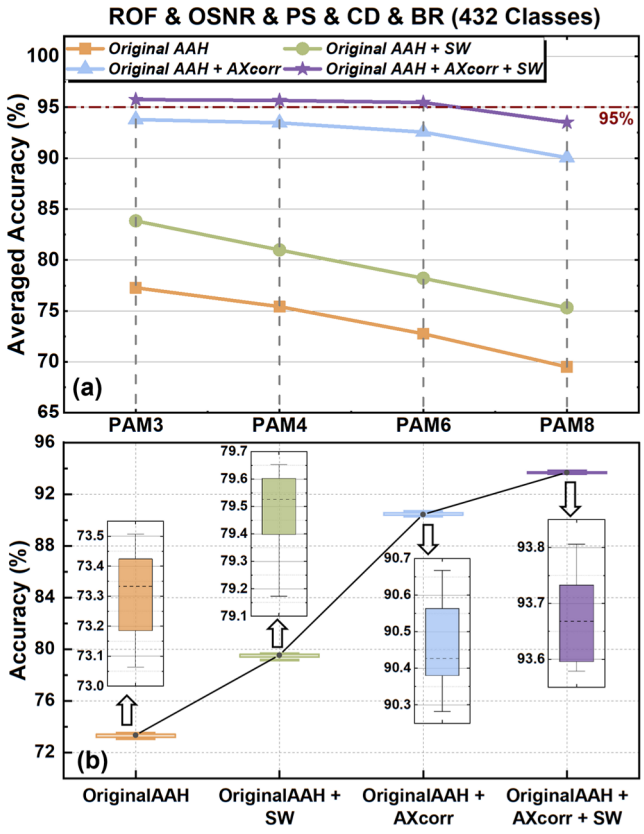

Figure C8 Comparison of classification accuracy of ROF, PS, OSNR, BR, and CD five-parameter joint monitoring with (a) different MF datasets, and (b) all MF datasets.

Subsequently, a more complex five-parameter joint monitoring classification task is explored, incorporating BR and CD alongside ROF, PS, and OSNR, resulting in 432 classification categories. The classification accuracies for this task across four higher-order *M*-PAM formats are evaluated and shown in Fig. C8(a). When the original AAH data is used as input features for the LightGBM model, the classification accuracies for all modulation formats (MFs) remain below 80%, with a pronounced decline as the PAM order increases. Introducing AXcorr preprocessing significantly improves performance, elevating accuracies across all

MFs to above 90%. This demonstrates that power-induced AAH drift is a key factor affecting joint classification performance, and the AXcorr algorithm effectively mitigates this issue. Upon further applying SW preprocessing, classification accuracies rise to approximately 95%, confirming the robustness of the combined AXcorr+SW scheme across various MFs in complex multi-parameter classification tasks. Fig. C8(b) illustrates the distribution of classification accuracies under all MFs, where AXcorr+SW achieves an overall accuracy exceeding 93.5%. Although classification accuracy is primarily constrained by the PAM8 format, the proposed preprocessing scheme still yields an accuracy improvement of approximately 20% compared to using the original AAH data.

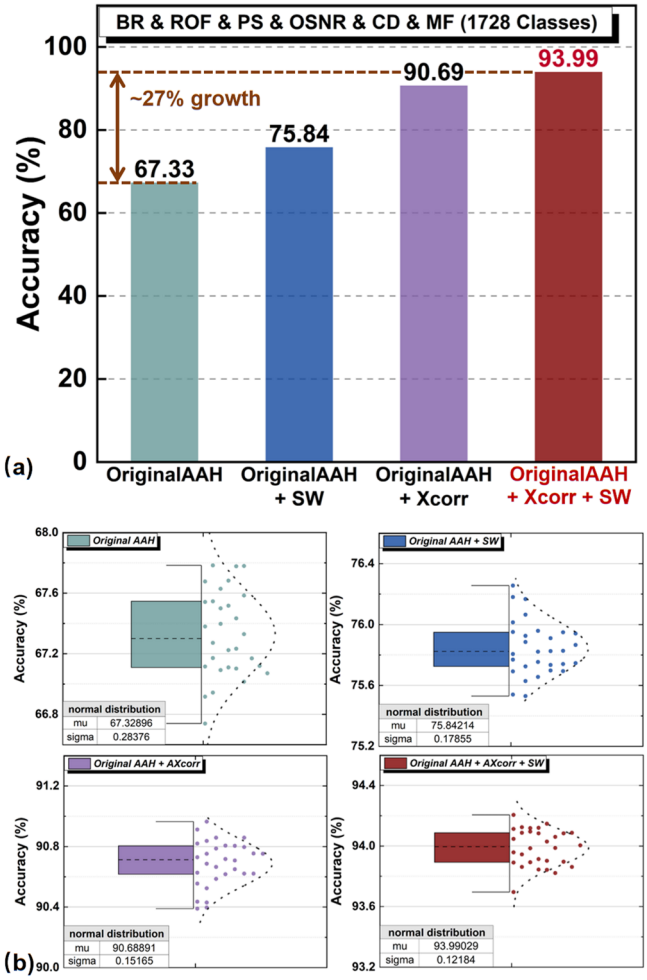

Figure C9 Comparison of the accuracy distribution of the six-parameter joint monitoring classification of ROF, PS, OSNR, BR, CD, and MF.

Finally, the most complex six-parameter joint monitoring task is considered, encompassing all previously discussed parameters: ROF, PS, OSNR, BR, CD, and MF, resulting in a total of 1,728 classification categories. The performance of the LightGBM model is evaluated using AAH data under various preprocessing schemes, as shown in Fig. C9(a). In this scenario, the proposed AXcorr+SW preprocessing scheme continues to demonstrate strong performance, achieving a classification accuracy of approximately 93%. This represents a substantial improvement of about 27% compared to the results obtained using the original, unprocessed AAH data. These findings confirm that the AXcorr+SW preprocessing approach not only significantly enhances joint monitoring accuracy but also exhibits robust generalization capabilities across complex, high-dimensional multi-parameter monitoring tasks.

In addition, the distribution of classification accuracies for the six-parameter joint monitoring task is analyzed, as illustrated in Fig. C9(b). The dotted line on the right side of the boxplot represents the fitted normal distribution derived from the measured accuracy data points, with the corresponding mean and variance values indicated in the upper right corner. It is evident that the use of AAH data preprocessed with the AXcorr + SW scheme not only enhances the monitoring accuracy but also significantly reduces the variance in accuracy. Specifically, the variance of the classification accuracy decreases from 0.28 (using original AAH data) to 0.12 after preprocessing. This result confirms that the proposed preprocessing method not only improves overall accuracy but also enhances the stability and reliability of the multi-parameter monitoring performance.

Appendix D Discussion

The sampling rate and the number of sampling points of the real-time oscilloscope are critical factors that influence the characteristics of the AAH. A higher sampling rate combined with an increased number of sampling points enhances the resolution of the signal's amplitude distribution, thereby improving the monitoring and classification accuracy. However, these improvements also raise the hardware complexity and cost requirements. As illustrated in Fig. D1, the original AAHs of PS-PAM4 under varying sampling conditions show that the number of sampling points has a more pronounced effect on the AAH quality than the sampling rate. In practical applications, it is important to strike a balance between monitoring performance and device cost. Hence, selecting an optimal combination of sampling rate and sampling points is essential for achieving efficient, low-cost monitoring. Despite its importance, existing studies have seldom explored the joint impact of sampling rate and sampling points on the AAH's role in multi-parameter monitoring accuracy. This work addresses this gap by systematically investigating the monitoring performance under various sampling configurations, aiming to identify the optimal oscilloscope parameter settings.

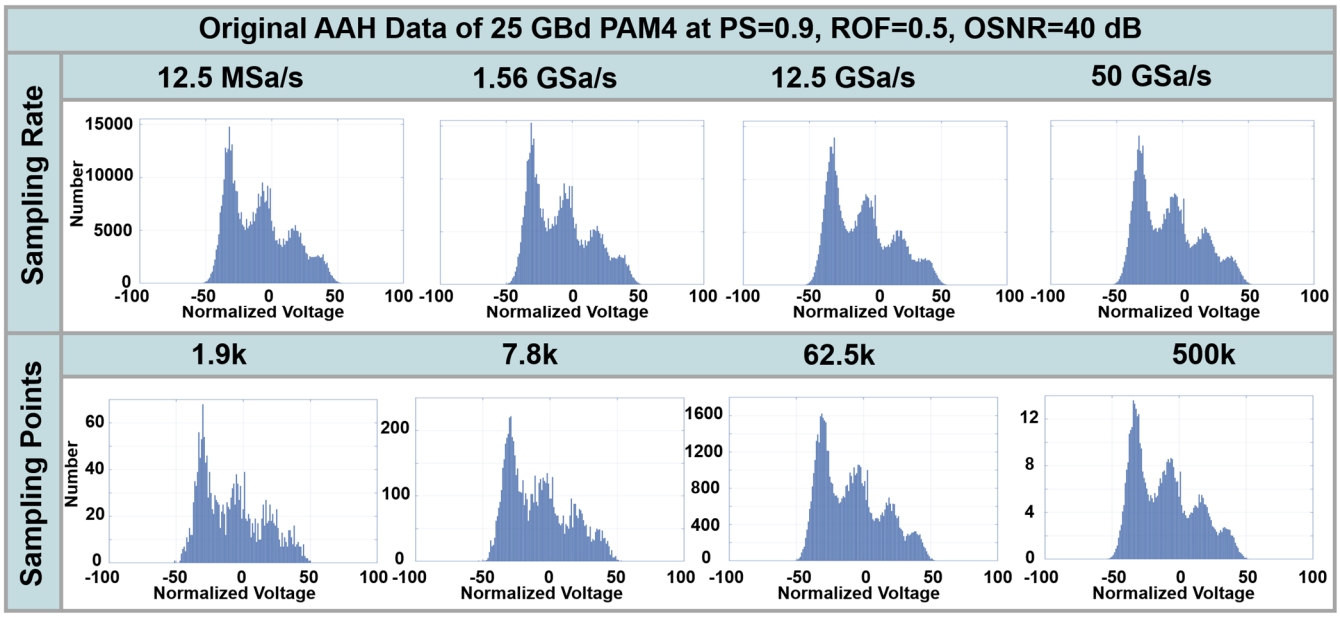

Figure D1 Schematic comparison of original PAM4 AAH at different oscilloscope sample rates and number of sample points (BR: 25 GBd; probability shaping factor: 0.9; ROF: 0.5; channel OSNR: 40 dB.

In the experimental setup, a 25-GBd PAM4 signal is employed at the transmitter side, with the CD in the transmission channel fixed at −40 ps/nm to enable the joint monitoring of three parameters: ROF, PS, and OSNR. This results in a total of 36 distinct monitoring categories. To assess the impact of oscilloscope configuration on monitoring performance, the sampling rate is varied from 12.5 MSa/s to 50 GSa/s. The selected set of sampling rates includes: 12.5 MSa/s, 25 MSa/s, 39 MSa/s, 78 MSa/s, 0.25 GSa/s, 0.31 GSa/s, 0.78 GSa/s, 1.56 GSa/s, 3.13 GSa/s, 12.5 GSa/s, 25 GSa/s, 50 GSa/s, while the number of sampling points is varied across the set: 1.9k, 3.9k, 7.8k, 15.6k, 31.3k, 62.5k, 125k, 250k, 500k. These two sets yield a total of 108 distinct oscilloscope parameter combinations.

For each combination, 100 sets of sampled signals are collected from the oscilloscope. The corresponding original AAH datasets are generated through statistical processing, with the number of histogram bins set to 200. Considering the 36 monitoring categories and 108 oscilloscope configurations, a total of 388,800 AAH samples are generated, forming a large-scale dataset for subsequent classification using the LightGBM model. The scale of this dataset provides a strong foundation for robust model training and evaluation. For clarity in the analysis, the set of oscilloscope parameter combinations is denoted as C1, while the set of monitoring parameter combinations is denoted as C2. For each oscilloscope configuration in the C1 set, 70% of the AAH data from each class in C2 is randomly selected for training, while the remaining 30% is used for testing. The average classification accuracy across multiple runs is then computed and used as the final accuracy for each C1 configuration. Initially, the original AAH data—without any preprocessing—is used as the feature set. The resulting monitoring classification accuracies for all C1 combinations are visualized in Fig. D2(a).

In the heat map presented in Fig. D2(a), the horizontal axis represents the number of sampling points, while the vertical axis denotes the sampling rate. To further aid interpretation, threshold contour lines corresponding to classification accuracies of 0.80, 0.90, 0.95, and 0.99 are overlaid on the heat map. This visualization provides a clear reference for identifying which oscilloscope settings yield acceptable monitoring performance and highlights the trade-offs between sampling resolution and classification accuracy.

Overall, when using the original AAH data, none of the oscilloscope parameter combinations are able to achieve classification accuracies exceeding 99%. The combinations that yield accuracies above 95% are primarily concentrated within the range of 125k to 500k sampling points and sampling rates between 12.5 MSa/s and 3.13 GSa/s. However, approximately half of the C1 configurations result in monitoring accuracies below 80%, highlighting the limitations of using unprocessed AAH data as input features.

The number of sampling points has a more pronounced impact on classification accuracy than the sampling rate. The accuracy increases monotonically with the number of sampling points, demonstrating a consistent positive correlation. For instance, at both the lowest (12.5 MSa/s) and highest (50 GSa/s) sampling rates, the classification accuracy exhibits a variation of approximately 40% when the number of sampling points increases from 1.9k to 500k, as illustrated in Fig. D1. These results underscore the importance of selecting appropriate sampling parameters to ensure high-quality feature extraction for accurate multi-parameter monitoring.

In contrast to the sampling points, the classification accuracy does not exhibit a monotonic relationship with the sampling rate when the number of sampling points is held constant. For example, at 1.9k sampling points, the classification accuracy initially decreases as the sampling rate increases from 12.5 MSa/s. As the sampling rate approaches 0.78 GSa/s, the accuracy stabilizes, showing minimal variation across the intermediate range of 0.78 GSa/s to 12.5 GSa/s. However, when the sampling rate continues to increase beyond 12.5 GSa/s, the accuracy begins to gradually improve again, with a notable increase observed at 50 GSa/s. This non-linear behavior suggests that excessively high sampling rates may begin to recapture finer signal details, particularly under low-resolution conditions (i.e., few sampling points), thereby improving classification performance. Nonetheless, the effect of sampling rate is less consistent and more sensitive to the interplay with the number of sampling points compared to the more dominant impact of sampling resolution alone.

This trend is particularly evident when the number of sampling points is low. This is because, for a fixed number of sampling points N, a lower sampling rate results in a longer observation time window, allowing the captured amplitude distribution to more closely approximate the true amplitude distribution of the signal $x(t)$. As the sampling rate increases—especially approaching 12.5 GSa/s, which is close to the signal baud rate of 25 GBd—the sampled points become more temporally correlated, and finer amplitude variations are captured in the AAH. This enhances classification performance by enriching the feature space. However, higher sampling rates impose stricter hardware requirements, increasing system cost. Consequently, employing low-cost oscilloscopes with reduced sampling rates—when paired with appropriate preprocessing and adequate sampling points—can serve as a practical and cost-effective approach for implementing AAH-based multi-parameter monitoring on the transmitter side.

It is important to note that the bandwidth limitation effect, while affecting received signal quality, does not undermine the

proposed AAH-based monitoring scheme. The fundamental distinction lies between signal recovery and signal monitoring: AAH captures the statistical distribution of signal amplitudes rather than the time-domain waveform, and does not require signal recovery or clock synchronization. Bandwidth-induced distortions (e.g., smoothing of amplitude peaks, merging of histogram modes) are themselves encoded as distinguishable features in the AAH, which can actually aid parameter classification. For instance, a signal that falls within the system bandwidth produces a different AAH profile than one that exceeds it, facilitating baud rate identification.

From the perspective of the Nyquist sampling theorem, the 12.5 MSa/s sampling rate is indeed far below the signal baud rate (up to 27 GBd). However, the Nyquist criterion applies to waveform reconstruction, not to amplitude statistics. Asynchronous sampling at a sub-Nyquist rate can be viewed as an independent and identically distributed (i.i.d.) random sampling process of the signal's amplitude. As long as the number of sampling points is sufficiently large (500,000 in our experiments), the resulting amplitude histogram converges to the true statistical distribution of the signal — regardless of the sampling rate. This principle is confirmed by the experimental results in Figs. D1–D3, which show that monitoring accuracy exceeds 95% at 12.5 MSa/s when paired with sufficient sampling points (>62.5k) and AXcorr+SW preprocessing.

Meanwhile, the SW method, as proposed, is applied to preprocess the original AAH data, as shown in Fig. D2(b). In this case, the window length for the SW algorithm is set to 6, with a sliding step of 2, and the Simple Moving Average (SMA) algorithm is used within the window. The results indicate that the overall trend of monitoring accuracy remains consistent with the original AAH data; however, there is a noticeable improvement in accuracy across all combinations of C1. The accuracy has been enhanced by 2% to 15%, and the proportion of C1 combinations achieving accuracy above 95% has increased to nearly one-third. This demonstrates the effectiveness of the SW algorithm in improving the accuracy of multi-parameter monitoring tasks.

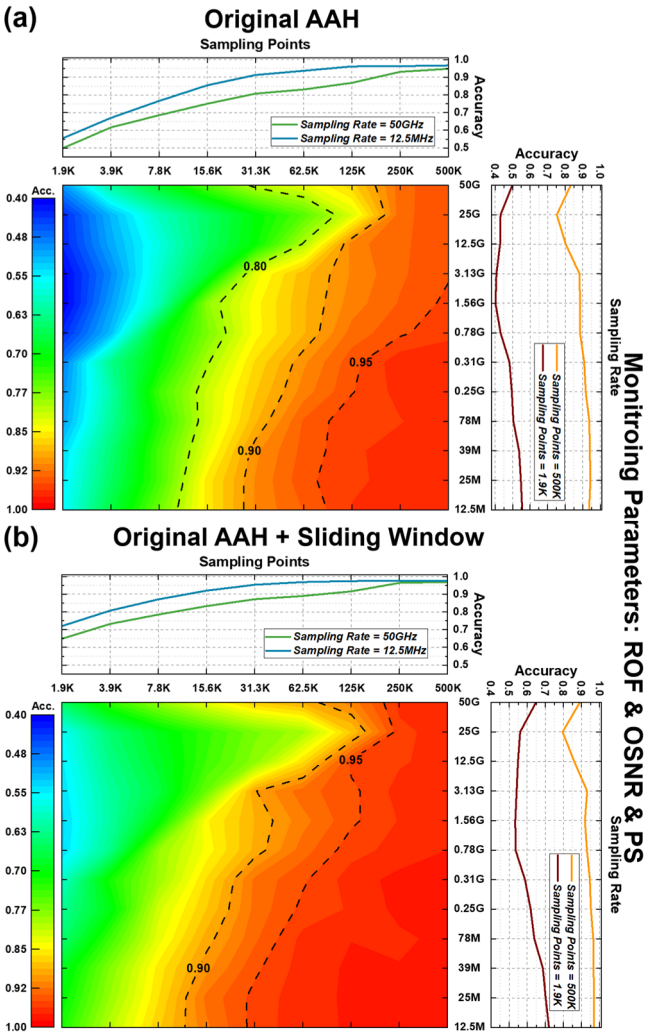

Figure D2 Classification accuracies obtained at different combinations of sampling rates and number of sampling points using (a) raw AAH data and (b) SW-processed AAH data in the ROF, OSNR and PS joint monitoring classification.

Subsequently, the AAH data processed using the AXcorr technique is used as the dataset, as shown in Fig. D3(a). It is evident that there is a substantial improvement in the accuracy across all combinations of sampling rate and sampling points. Approximately two-thirds of the C1 combinations achieve over 90% accuracy, with about one-third surpassing 99% accuracy in the monitoring tasks. This highlights the effectiveness of the AXcorr algorithm in significantly enhancing the accuracy of AAH multi-parameter monitoring.

Finally, when AXcorr and SW are combined to preprocess the original AAH data, as shown in Fig. D3(b), it is clear that the accuracy improves further compared to using AXcorr alone. More than half of the C1 combinations now achieve over 95% accuracy. Compared to the original AAH dataset, the accuracy for each C1 combination has improved by 4% to 30%. These experimental results confirm that the proposed preprocessing method, combining AXcorr and SW, provides significant accuracy improvements and is adaptable to a wide range of oscilloscope parameter combinations.

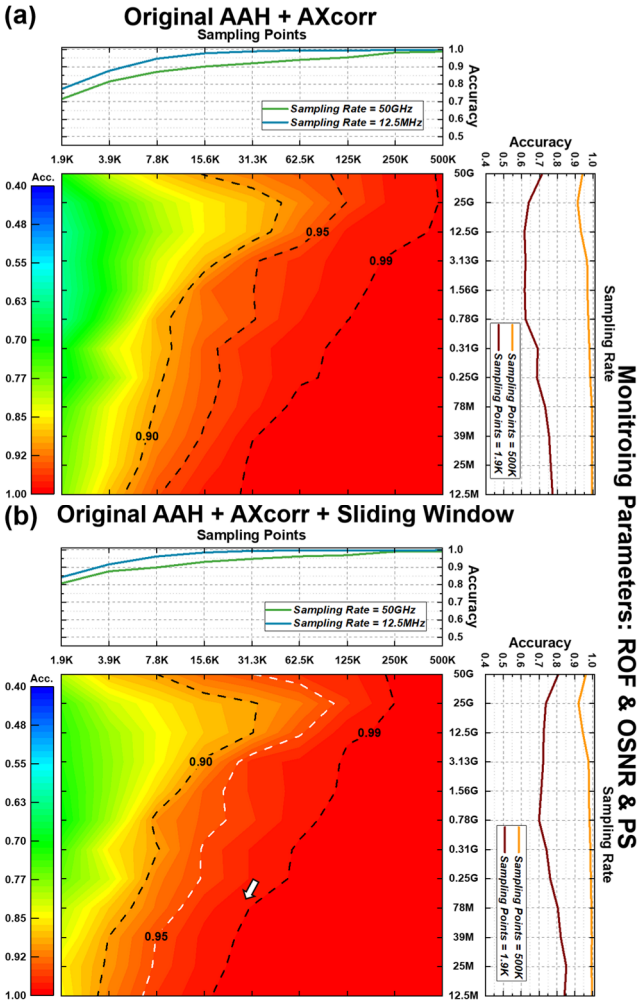

Figure D3 Classification accuracies obtained at different combinations of sampling rate and number of sampling points using (a) AAH processed by AXcorr and (b) AAH pre-processed by AXcorr + SW in the ROF, OSNR, and PS joint monitoring classification.

Based on the experimental results above, the baseline for oscilloscope parameter selection is provided for 25-GBd PAM4, where the joint monitoring accuracy exceeds 95%, as indicated by the white line in Fig. D3(b). In summary, the joint monitoring accuracy increases monotonically with the number of sampling points, so it is recommended to select as many sampling points as possible based on application requirements, with more than 62.5k sampling points being ideal. When the sampling rate is either greater than 1× or less than 0.01× the signal BR, the joint monitoring accuracy can still be achieved. To maintain "low-cost" oscilloscope usage, lower sampling rates are preferred. Finally, while larger sampling points and lower BRs result in longer data acquisition times, oscilloscope parameters can be optimized by referring to the 95% baseline in Fig. D3(b). These results demonstrate that by carefully selecting oscilloscope parameters and combining AXcorr and SW preprocessing methods, high-accuracy multi-parameter monitoring and classification can be achieved even under low-cost conditions.

Appendix E Conclusion

In this paper, a low-cost joint M-PAM signal monitoring scheme based on AAH and LightGBM models is proposed, along with a low-complexity AAH data preprocessing algorithm designed to mitigate the impact of channel power fluctuations on the monitoring accuracy of parameters such as BR, CD, MF, PS, ROF, and OSNR. The proposed method significantly enhances the monitoring accuracy and robustness of multiple parameter combinations, even under high-order MFs of PAM3/4/6/8. By using the combined AXcorr and SW preprocessing algorithm, the method improves the accuracy and robustness of monitoring across various parameter combinations in high-order PAM formats. Experimental results demonstrate that the proposed method, utilizing a low-cost real-time oscilloscope with a sampling rate of just 12.5 MSa/s, achieves over 93% accuracy in five-parameter joint classification tasks (except for MF), covering a total of 432 monitoring classes. Furthermore, the method achieves a classification accuracy of 93.99% in six-parameter joint monitoring tasks, with 1728 classes. Compared to the use of original AAH data, the AXcorr+SW preprocessing scheme leads to a 27% improvement in accuracy, while reducing the variance in accuracy distribution from 0.28 to 0.12. This approach provides a cost-effective and stable solution for multi-parameter monitoring in optical networks. It should be clarified that the low-complexity advantage of this scheme primarily refers to the hardware deployment level — the proposed method requires only a low-sampling-rate oscilloscope and a photodetector, without clock recovery, high-speed ADC, or GPU acceleration.

Since the AAH-based feature extraction is inherently independent of the specific modulation format, the proposed framework can in principle be extended to other modulation formats such as QAM, where different signal configurations produce distinguishable amplitude distributions after direct detection.

Acknowledgment

This work is jointly supported by the Key Laboratory of Opto- electronic Information Technology, Ministry of Education (Tianjin University) (No. 2024KFKT005), the Natural Science Foundation of Shaanxi Province (No. 2023-JC-YB-587) and the National Key Research and Development Program of China (No. 2019YFB1803700).

References

1 Winzer P J, Neilson D T, Chraplyvy A R. Fiber-optic transmission and networking: the previous 20 and the next 20 years [Invited]. Opt Express, 2018, 26: 24190-24239
2 Le S T, Schuh K, Dischler R, et al. Beyond 400 Gb/s direct detection over 80 km for data center interconnect applications. J Lightwave Technol, 2020, 38: 538-545
3 Sludds A, Bandyopadhyay S, Chen Z, et al. Delocalized photonic deep learning on the internet's edge. Science, 2022, 378: 270-276
4 Martella G R, Nespola A, Straullu S, et al. Scaling laws for unamplified coherent transmission in next-generation short-reach and access networks. J Lightwave Technol, 2021, 39: 5805-5814
5 Pedro J, Costa N, Sanders S. Cost-effective strategies to scale the capacity of regional optical transport networks. J Opt Commun Netw, 2022, 14: A154-A165
6 High-speed flexible coherent optical transport network. In: Datacenter Connectivity Technologies: Principles and Practice. River Publishers, 2017. 621-658
7 Zhang J, Li G, Xing S, et al. Flexible and adaptive coherent PON for next-generation optical access network. Opt Fiber Technol, 2023, 75: 103190
8 Pang X, Udalcovs A, Schatz R, et al. High-speed short reach optical communications: technological options and challenges. In: Proceedings of 2020 Asia Communications and Photonics Conference (ACP) and International Conference on Information Photonics and Optical Communications (IPOC), 2020. S3H.5
9 Ozolins O, Joharifar M, Salgals T, et al. Optical amplification-free high baudrate links for intra-data center communications. J Lightwave Technol, 2023, 41: 1200-1206
10 Zhou X, Lam C F, Urata R, et al. State-of-the-art 800G/1.6T datacom interconnects and outlook for 3.2T. In: Proceedings of 2023 Optical Fiber Communications Conference and Exhibition (OFC), 2023. W3D.1
11 Pang X, Udalcovs A, Schatz R, et al. Short reach communication technologies for client-side optics beyond 400 Gbps. IEEE Photonics Technol Lett, 2021, 33: 1046-1049
12 Hu W, Zhu Y. 100G and beyond for PON and short reach optical networks. In: Proceedings of 2023 Optical Fiber Communications Conference and Exhibition (OFC), 2023. Th3G.5
13 Häger C, Pfister H D. Physics-based deep learning for fiber-optic communication systems. IEEE J Sel Areas Commun, 2021, 39: 280-294
14 Pang X, Ozolins O, Lin R, et al. 200 Gbps/lane IM/DD technologies for short reach optical interconnects. J Lightwave Technol, 2020, 38: 492-503
15 Ostrovskis A, Salgals T, Krüger B, et al. 106.25 Gbaud on-off keying and pulse amplitude modulation links supporting next generation ethernet on single lambda. J Lightwave Technol, 2024, 42: 1272-1280
16 Tranter A D, Slatyer H J, Hush M R, et al. Multiparameter optimisation of a magneto-optical trap using deep learning. Nat Commun, 2018, 9: 4360
17 Wei J, Cheng Q, Penty R V, et al. 400 Gigabit ethernet using advanced modulation formats: performance, complexity, and power dissipation. IEEE Commun Mag, 2015, 53: 182-189
18 Meng X, Zhang Y, Zhang X, et al. Machine learning assisted vector atomic magnetometry. Nat Commun, 2023, 14: 6105
19 Nevin J W, Nallaperuma S, Shevchenko N A, et al. Machine learning for optical fiber communication systems: an introduction and overview. APL Photonics, 2021, 6: 121101
20 Pan X, Wang X, Tian B, et al. Machine-learning-aided optical fiber communication system. IEEE Network, 2021, 35: 136-142
21 21 Wan Z, Yu Z, Shu L, et al. Intelligent optical performance monitor using multi-task learning based artificial neural network. Opt Express, 2019, 27: 11281-11291
22 Pan Z, Yu C, Willner A E. Optical performance monitoring for the next generation optical communication networks. Opt Fiber Technol, 2010, 16: 20-45
23 Wang D, Jiang H, Liang G, et al. Optical performance monitoring of multiple parameters in future optical networks. J Lightwave Technol, 2021, 39: 3792-3800
24 Dong Z, Khan FN, Sui Q, et al. Optical performance monitoring: a review of current and future technologies. J Lightwave Technol, 2016, 34: 525-543
25 Lalam N, Bukka S, Bhatta H, et al. Achieving precise multiparameter measurements with distributed optical fiber sensor using wavelength diversity and deep neural networks. Commun Eng, 2024, 3: 121
26 Peng S, Jiang H, Wang H, et al. Modulation classification based on signal constellation diagrams and deep learning. IEEE Trans Neural Networks Learn Syst, 2019, 30: 718-727
27 Wang D, Zhang M, Li J, et al. Intelligent constellation diagram analyzer using convolutional neural network-based deep learning. Opt Express, 2017, 25: 17150-17166
28 Jiang K, Zhang J, Wu H, et al. A novel digital modulation recognition algorithm based on deep convolutional neural network. Appl Sci, 2020, 10: 1166
29 Tizikara D K, Serugunda J, Katumba A. Machine learning-aided optical performance monitoring techniques: a review. IEEE Commun Surv Tutor, 2022, 2: 2839-2882
30 Wang D, Zhang M. Artificial intelligence in optical communications: from machine learning to deep learning. Front Commun Netw, 2021, 2: 656786
31 Du J, Yang T, Chen X, et al. A CNN-based cost-effective modulation format identification scheme by low-bandwidth direct detecting and low rate sampling for elastic optical networks. Opt Commun, 2020, 471: 126007
32 Saif W S, Esmail M A, Ragheb A M, et al. Machine learning techniques for optical performance monitoring and modulation format identification: a survey. IEEE Commun Surv Tutor, 2020, 22: 2839-2882
33 Musumeci F, Rottondi C, Nag A, et al. An overview on application of machine learning techniques in optical networks. IEEE Commun Surv Tutor, 2019, 21: 1383-1408
34 Lu H, Jiang M, Cheng J. Deep learning aided robust joint channel classification, channel estimation, and signal detection for underwater optical communication. IEEE Trans Commun, 2021, 69: 2290-2303
35 Barbastathis G, Ozcan A, Situ G. On the use of deep learning for computational imaging. Optica, 2019, 6: 921-943
36 Chen X, Liu C Y, Proietti R, et al. Automating optical network fault management with machine learning. IEEE Commun Mag, 2022, 60: 88-94
37 Cimini V, Valeri M, Polino E, et al. Deep reinforcement learning for quantum multiparameter estimation. Adv Photonics, 2023, 5: 016005
38 38 Khan FN, Fan Q, Lu C, et al. Chapter 21 - Machine learning methods for optical communication systems and networks. In: Optical Fiber Telecommunications VII. Academic Press, 2020. 921-978
39 Xia L, Zhang J, Hu S, et al. Transfer learning assisted deep neural network for OSNR estimation. Opt Express, 2019, 27: 19398-19406
40 Wang C, Fu S, Xiao Z, et al. Long short-term memory neural network (LSTM-NN) enabled accurate optical signal-to-noise ratio (OSNR) monitoring. J Lightwave Technol, 2019, 37: 4140-4146
41 Khan F N, Yu Y, Tan M C, et al. Experimental demonstration of joint OSNR monitoring and modulation format identification using asynchronous single channel sampling. Opt Express, 2015, 23: 30337-30346
42 Luo H, Huang Z, Wu X, et al. Cost-effective multi-parameter optical performance monitoring using multi-task deep learning with adaptive ADTP and AAH. J Lightwave Technol, 2021, 39: 1733-1741
43 Chai J, Chen X, Zhao Y, et al. Joint symbol rate-modulation format identification and OSNR estimation using random forest based ensemble learning for intermediate nodes. IEEE Photonics J, 2021, 13: 7900106
44 Jiang T, Gao Z, Chen Y, et al. FrFT-based estimation of linear and nonlinear impairments using Vision Transformer. J Opt Commun Netw, 2024, 16: 419-431
45 Yang F, Bai C, Yue Y, et al. ORT image and IMT-MIMO-ResNet-based multi-mode multi-parameter joint optical performance monitoring in MDM-EON systems. Opt Express, 2025, 33: 20530-20548
46 Lin M, Pan X, Wang F, et al. Physics-informed neural network for OSNR monitoring based on the nonlinear Schrödinger equation and fractional Fourier transform. Opt Express, 2026, 34: 8382-8402
47 Kulandaivel S, Jeyachitra R K. Dual-stage deep learning model for multiparameter optical performance monitoring and optical modulation index estimation for radio over fiber systems. Arab J Sci Eng, 2025, 50: 11383-11393

48 Li S-A, Yang J, Liu Y, et al. Multiparameter optical performance monitoring of PAM channels using asynchronous amplitude histogram. In: Proceedings of CLEO 2024, 2024. JTh2A.150
49 Cheng Y, Luo H, Xiang M, et al. Nonlinear activation function-free artificial neural network enabled optical performance monitoring. IEEE Commun Mag, 2023, 61: 156-160
50 Khan F N. Machine learning-enabled intelligent fiber-optic communications: major obstacles and the way forward. IEEE Commun Mag, 2023, 61: 122-128
51 Khan F N. Non-technological barriers: the last frontier towards AI-powered intelligent optical networks. Nat Commun, 2024, 15: 5995
52 Luong T V, Zhang X, Xiang L, et al. Deep learning-aided optical IM/DD OFDM approaches the throughput of RF-OFDM. IEEE J Sel Areas Commun, 2022, 40: 212-226
53 Karanov B, Chagnon M, Aref V, et al. Experimental investigation of deep learning for digital signal processing in short reach optical fiber communications. In: Proceedings of 2020 IEEE Workshop on Signal Processing Systems (SiPS), 2020. 1-6
54 Koç C K. Analysis of sliding window techniques for exponentiation. Comput Math Appl, 1995, 30: 17-24
55 Wen Y J, Cui Y, Bai Y. Mitigation of optical multipath interference impact for directly detected PAMn system. Opt Express, 2020, 28: 38317-38333
56 Shami A, Bai X, Assi C M, et al. Jitter performance in ethernet passive optical networks. J Lightwave Technol, 2005, 23: 1745-1753
57 Wang D, Sui Q, Li Z. Toward universal optical performance monitoring for intelligent optical fiber communication networks. IEEE Commun Mag, 2020, 58: 54-59
58 Zibar D, Wymeersch H, Lyubomirsky I. Machine learning under the spotlight. Nat Photonics, 2017, 11: 749-751
59 Ke G, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst, 2017: 3146-3154