

DNN-assisted phase distance tuned PSK modulation for PAM4-to-QPSK format conversion gateway node

Kodama, Takahiro; Koike-Akino, Toshiaki; Millar, David S.; Kojima, Keisuke; Parsons, Kieran

TR2022-091 August 06, 2022

Abstract

This paper proposes an optical gateway that converts pulse amplitude modulation (PAM) format to phase shift keying (PSK) modulation format, enabling flexible intensity-to-phase mapping without relying on optical-electrical-optical data conversion at heterogeneous network connections. A proof-of-principle experiment shows that optically converted PSK signals from regular PAM signals will induce non-uniform irregular phase noise distortion. The proposed optical gateway is designed to provide an optimized phase distance for PSK signals such that an achievable information rate is maximized by a deep learning-based decision on the receiver side. The phase distance-tuned PSK signals offer performance improvement of about 4 dB gain at a target generalized mutual information by making use of a digital neural network recovery.

Optics Express 2022

DNN-assisted phase distance tuned PSK modulation for PAM4-to-QPSK format conversion gateway node

TAKAHIRO KODAMA,^{1,*} TOSHIAKI KOIKE-AKINO,² KEISUKE KOJIMA,² AND KIERAN PARSONS,²

¹*Faculty of Engineering and Design, Kagawa University, 2217-20 Hayashi-cho, Takamatsu, Kagawa 761-0396 Japan*

²*Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139 USA*

**kodama.takahiro@kagawa-u.ac.jp*

Abstract: This paper proposes an optical gateway that converts pulse amplitude modulation (PAM) format to phase shift keying (PSK) modulation format, enabling flexible intensity-to-phase mapping without relying on optical-electrical-optical data conversion at heterogeneous network connections. A proof-of-principle experiment shows that optically converted PSK signals from regular PAM signals will induce non-uniform irregular phase noise distortion. The proposed optical gateway is designed to provide an optimized phase distance for PSK signals such that an achievable information rate is maximized by a deep learning-based decision on the receiver side. The phase distance-tuned PSK signals offer performance improvement of about 4 dB gain at a target generalized mutual information by making use of a digital neural network recovery.

© 2022 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

In current optical communication networks (NWs), different types of modulation formats and detection schemes are used, specifically intensity modulation/direct detection (IM/DD) and phase modulation/coherent detection, respectively, for short-reach and long-reach fiber NWs. When bridging a short-distance NW (such as optical access and intra-datacenter NW) to a long-distance NW (such as metro and inter-datacenter NW), IM/coherent detection may be used at the transmitter (Tx)/receiver (Rx) pair, whereas IM modulation such as pulse-amplitude modulation (PAM) is not energy/spectrally efficient for coherent detection. In general, an optical-electrical-optical (O/E/O) conversion is employed at heterogeneous NW gateway nodes (HNGNs) that connect short- and long-reach optical fiber communication NWs [1-4]. Nevertheless, such indirect modulation format conversions may cause non-negligible traffic delays and power consumption for electrical signal processing.

An emerging technology to avoid the O/E/O conversion is a direct optical conversion which can flexibly convert a PAM signal to a phase-shift keying (PSK) signal at a gateway node. The optical-domain conversion from a PAM signal to an in-phase (I) and quadrature (Q) modulated signal has been realized by using nonlinear optical effects such as cross-phase modulation (XPM) and four-wave mixing (FWM) for optical-based HNGNs [5-9]. Investigations on PSK-to-PAM signal conversion [10-11] and PSK to quadrature-amplitude modulation (QAM) conversion [12-15] have also been reported using nonlinear optical effects. To use nonlinear optical effects, an active device such as a semiconductor optical amplifier is superior in terms of miniaturization. However, active devices usually suffer from a limitation in speed of operations, and hence passive devices such as a highly nonlinear fiber (HNLF) are often more advantageous. A definite advantage of optical-domain conversion is that it seamlessly transmits signals from short-reach optical NWs to long-reach optical NWs

while maintaining the packet format, modulation depth, and symbol rate without using driver amplifier or linear optical IQ modulator [16]. In addition, the optical conversion may achieve a low latency processing since the use of erbium doped fiber amplifier (EDFA) and HNLF may just have a lightwave propagation delay of about 0.15 μs /unit and 4.9 $\mu\text{s}/\text{km}$. If the fiber length of HNLF is within 20 km, the processing time of HNLF-based optical signal processing will be shorter than that of the O/E/O conversion which requires a delay of at least 100 μs [17] even without forward error correction. As the realization of low latency systems is highly demanded in the next-generation optical networks (e.g., for mobile broadband, automatic guided vehicle, and industrial automation), all-optical processing approach is attractive for nodes connected to the optical access network where traffic delay is likely to occur. In particular for industrial automation applications and rigorous cloud computing, the latency requirement is much stringent, within one millisecond [18].

Although a recent work of optical-domain convertor [19] shows a decent performance when converting four-level PAM (PAM4) to quadrature PSK (QPSK) format by keeping a Gray code mapping, a relatively high signal-to-noise ratio (SNR) is required at the IM signal to prevent undesired distortion. When the IM signal has a low SNR due to poor characteristics of the light source and the optical amplifiers, the optical conversion from a noisy PAM signal to a PSK signal will face two phenomena: (1) intensity noise is converted to phase noise; and (2) larger phase noise occurs at higher amplitude due to the nonlinear optics. Accordingly, the converged PSK signals suffer from a non-uniform irregular distortion. This brings up two questions: (1) how to optimally adjust the irregular distortion; and (2) how to recover the distorted QPSK signals. We introduce a deep learning-based joint optimization of the optical gateway and the receiver detection. Specifically, we use a deep neural network (DNN) framework [20-22] to quantify the signal quality so that the optical gateway can tune the phase noise properly, and to recover the irregular constellation so that the end-to-end achievable information rate is maximized.

This paper is an extended work to solve the problems caused by the PAM4-QPSK conversion proposed in our preliminary study [19]. In the present paper, we propose a flexible generation of phase distance tuned (PDT) QPSK signals by adjusting the optical amplifier level of a regular PAM4 signal, and optimum decision of the PDT-QPSK signal using coherent detection with a DNN in the digital signal processing (DSP). To confirm the problem after noise additional PAM4-QPSK conversion, we experimentally verify the generation of a non-uniform phase noise distributed QPSK signal from a noisy PAM4 signal using XPM generated in a HNLF. We show that the DNN-based decision can significantly outperform the conventional methods and that the penalty due to the non-optimal amplifier level can be effectively compensated by the DNN for a variable PDT-QPSK signal.

2. Operating principle

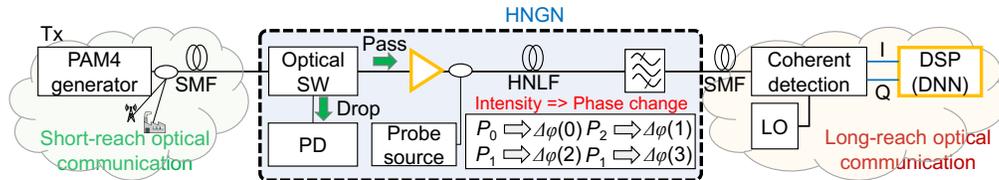


Fig. 1. PDT-QPSK generation and DNN decision.

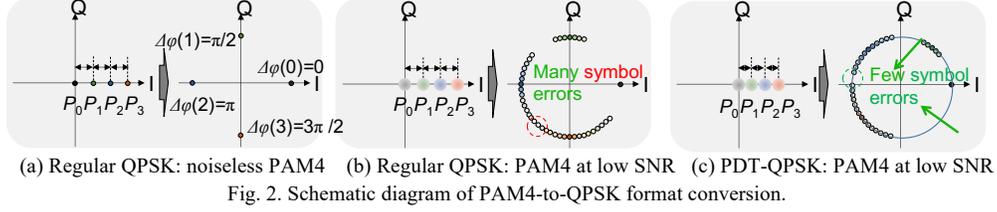


Fig. 2. Schematic diagram of PAM4-to-QPSK format conversion.

Figure 1 shows a configuration of the proposed PAM4-to-QPSK modulation format conversion, using PDT-QPSK and DNN-based decision. On the Tx side, a non-return-to-zero (NRZ) PAM4 signal is generated using a laser diode (LD) or Lithium niobate intensity modulator (LN-IM) at wavelength λ_1 in the short-reach optical communication NW.

At the HNGN, an optical switch selects whether to pass the PAM4 signal through the long-reach optical communication NW or to drop it. When the signal is dropped, a photodetector (PD) receives the PAM4 signal in the same way as for the standard DD detection. When passing through, the PAM4 signal and continuous wave (CW) as a probe light at wavelength λ_2 are launched into the HNLF. The probe light is phase modulated due to XPM induced by the PAM4 signal whose power is tuned by an optical amplifier at the EDFA.

The amount of phase change of the probe light due to XPM can be expressed as

$$\Delta\varphi(k) \cong k \left(2\gamma_1 L_{eff} P_k \right), \quad (1)$$

where γ_1 and L_{eff} are the nonlinear coefficient and the effective interaction length of the HNLF, with P_k being the power of the k -th level of the PAM4 signal ($k=0, \dots, 3$). We assume that the PAM4 signal is equally spaced as a standard modulation. When PAM4 is noiseless (sufficiently high SNR), the XPM with a proper power amplification can generate a clean regular QPSK signal with an equal phase distance as $\Delta\varphi(k) = \frac{\pi}{2}k$, as shown in Fig. 2(a). In

our previous study [19], the possibility of low power consumption operation was experimentally demonstrated at high SNR regimes by tuning the HNLF input power.

In the presence of noise in the PAM4 signal, non-uniform phase noise occurs at the optically converted QPSK signal, where the quality of constellation points having more phase shifts is drastically degraded, as shown in Fig. 2(b). Figure 2(c) shows the PDT-QPSK signal that can reduce the phase noise by tuning the optical amplifier level at the cost of smaller Euclidean distance between the irregular QPSK points. In this paper, we optimize the tradeoff between the phase noise distortion and the Euclidean distance of the PDT-QPSK signal to maximize the achievable information rate. To this end, we propose the use of a DNN model at the end of NW to quantify the quality of the PDT-QPSK signal as an estimator of generalized mutual information (GMI), which is known as a relevant metric for soft-decision forward error correction.

If λ_1 and λ_2 are excessively close, there is a low crosstalk tolerance for the PAM4 signal. If we increase the wavelength detuning, the XPM effect becomes small due to the walk-off between the two wavelengths. The parameters of the HNLF such as the nonlinear coefficient, length, and chromatic dispersion should be selected carefully to prevent undesired FWM and unstable phase modulation in the PAM4-QPSK conversion. An optical bandpass filter is placed after the HNLF to divide the generated QPSK signal from the PAM4 signal. At the end node of a long-reach optical communication NW, the QPSK signal is detected by coherent detection with high receiver sensitivity, and processed by the DNN-based decision.

3. Experimental verification of non-uniform distortion in optical gateway

Figure 3 shows the experimental configuration for the proof of concept of PDT-QPSK generation from a PAM4 signal. The PDT-QPSK is realized by adjusting a non-linear optical effect to make the phase difference between adjacent signal points less than $\pi/2$. Here, we focus only on the nonlinear optical effects generated by the HNGN, ignoring other nonlinear effects in a transmission over short-reach and long-reach optical NWs.

On the Tx side of an edge node in the short-reach optical NW, a narrow 100 kHz line-width tunable LD at a central wavelength of 1550 nm generates unmodulated light. The output of the LD is connected to an LN-IM. We use a narrow line-width LD at 1550 nm for both a probe light source in the HNGN and a local oscillator in the receiver. In the electrical domain, on the Tx side ONU in the short-reach optical NW, a 10 Gbaud NRZ-PAM4 signal with a 10-Gsample/s pseudo-random bit sequence data pattern is generated using an arbitrary waveform generator (AWG) (7122C, Tektronix, Tokyo, Japan) and driver amplifier. The LN-IM modulates the CW light at 1550 nm to obtain the optical NRZ-PAM4 signal. A low SNR PAM4 signal is emulated by reducing the gain of the driver amplifier. The clock signal output from the AWG is used to synchronize the AWG and digital storage oscilloscope (DSO) (6154C, Tektronix, Tokyo, Japan). We assume the utilization of the cost-effective quad small form-factor pluggable 28 (QSFP28) for 100G Ethernet for the optical short-reach transmission system, so the Nyquist shape does not limit the band. Although it is ideal for performing all-optical modulation format conversion with the Nyquist spectrum-shaped PAM4 signal, it is difficult to give the amount of phase change to the probe light at the desired signal point due to the influence of walk-off. Aliasing could be reduced because the frequency band of DSO is limited to 15 GHz.

At the HNGN, the optical PAM4 signal and the 2.4 dBm average power probe light at 1545 nm are launched into the HNLFs. The HNLF parameters are summarized in Table 1. Two types of HNLF are connected in series, but the two types of nonlinear fibers are fused, and there is no apparent effect of dispersion. The overall loss of the HNLFs including propagation loss and connection loss is 3.2 dB. The amounts of noise and phase change due to XPM in the HNLF are controlled using the gain from an EDFA before sending the signals through the HNLFs. The polarization controllers are arranged and adjusted so that the polarization states of the two input lights match in the HNLFs. After passing through the HNLFs, the optical filter with a transmission band of 1 nm passes only the QPSK signal generated at a center wavelength of 1545 nm.

On the Rx side of an edge node in the long-reach optical NW, the average power of the received PDT-QPSK signal is adjusted using a variable optical attenuator (VOA), mixed with the local oscillator light, and coherently detected. An analog-to-digital converter in the DSO samples and quantizes the received signal. The 20 Gsample/s output signal from the DSO is processed offline in terms of down sampling, frequency offset compensation (FOC), and decision directed-based carrier phase recovery (CPR). Figures 3(b) and 3(c) show the constellation and phase distribution of the received PDT-QPSK signal with different phase noises for the four levels. Based on the results in Fig. 3(c), we confirm that the amount of phase noise increases as the amount of phase change increases. The signal distribution of the signal points without phase change due to XPM becomes shot noise generated in the receiver.

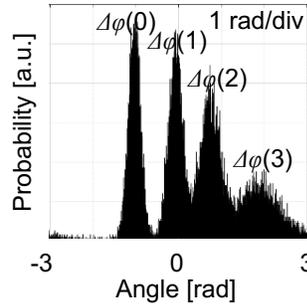
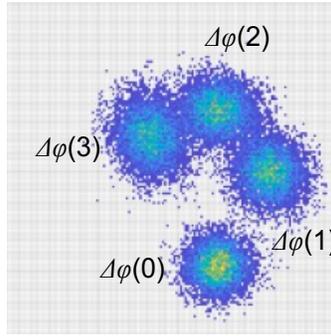
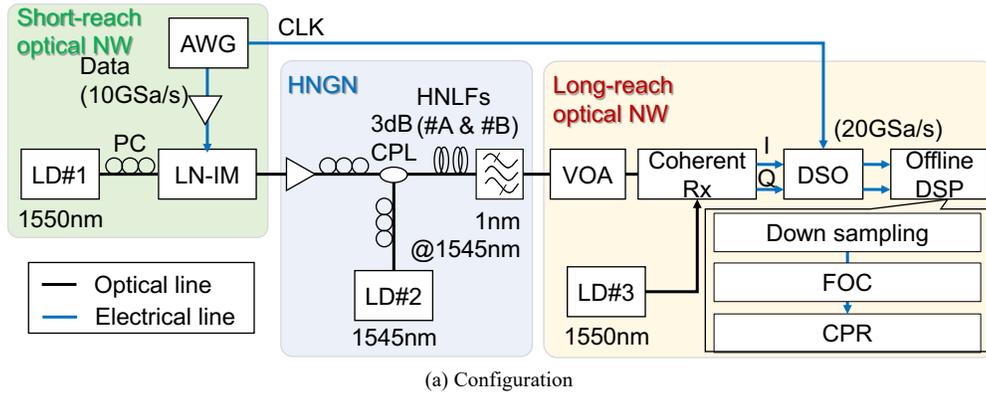


Fig. 3. Experimental results.

Table 1. HNLF Parameters

HNLF parameters at 1550 nm	#A	#B
Length [km]	2.5	1
Nonlinearity [1/W/km]	10	10
Dispersion [ps/nm/km]	0.57	0.52
Slope [ps/nm/km ²]	0.018	0.016
Loss [dB/km]	1.07	0.76

4. Simulation configuration and results

To verify the performance of the 10 Gbaud XPM-based NRZ-PAM4-to-QPSK optical conversion with PDT and DNN-based decision, we conduct a numerical simulation. Figure 4 shows the configuration for the simulation, which is basically same as the experimental setup. Some differences from the experimental setup include the central wavelength of probe light, the optical amplifier adjustment to vary the amount of additive white Gaussian noise (AWGN), the gain in front of the non-linear fiber, and DNN in the DSP. The bit stream generation is based on the Mersenne Twister having a repetition period of $2^{19937}-1$. The NRZ-PAM4 signal with 960 samples/symbol considers the fiber's non-linear effects over a wide spectral range. We evaluate the noise tolerance of the PAM4 signal by changing the constellation shape. The standard PAM4 signal is distorted before HNLFs by AWGN caused by the amplified spontaneous emission noise from the output of the EDFA. The HNLF parameters are the same as in the previous experiment and a loss-less connection between HNLFs is assumed. The central wavelengths for the probe light and signal light are set to

1540 nm and 1550 nm, respectively. The walk-off effect is ignored because the dispersion slope value is low. An additional noise that may occur after modulation format conversion is not considered for simplicity in the analysis.

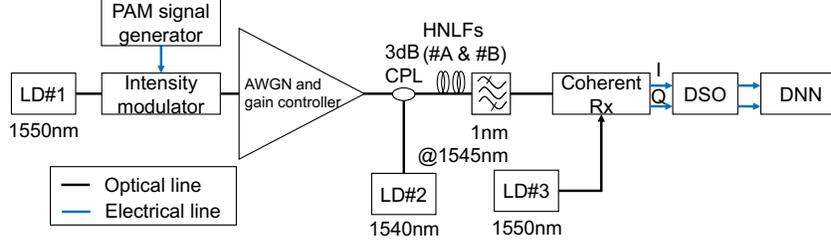


Fig. 4. Simulation setup.

Figure 5 shows constellation maps of the PDT-QPSK signal at an SNR of the PAM4 signal from 20 dB to 30 dB, and an HNLF input power from 38.5 mW to 60.5 mW. When the HNLF input power is 55 mW, the standard QPSK signal with $\Delta\varphi(k) = \frac{\pi}{2}k$ is generated.

However, noisy PAM4 signal causes considerable distortion in the optically-converted QPSK signal especially at $\Delta\varphi(3) = \frac{3\pi}{2}$. We aim at designing the phase shift of the QPSK

constellations to take advantage of the non-uniform phase distortion characteristics caused by the optical gateway.

To compensate for such irregular distortion of PDT-QPSK constellations, we employ a DNN whose architecture is based on residual network (ResNet) depicted in Fig. 6, where multiple fully-connected linear layers, batch normalization layers, rectified linear unit (ReLU) activation layers, and skip connections are configured with 10% dropouts. The DNN feeds distorted QPSK symbols to generate two-bit log-likelihood ratios (LLRs), whose sigmoid cross-entropy loss is minimized by a stochastic gradient descent based on adaptive momentum (Adam) with a learning rate of 0.001 over 300 epochs. We use a mini-batch size of 50. The data patterns for the DNN learning and testing are 120,600 bits and 13,400 bits long, respectively. We use the trained DNN to measure the signal quality of irregular constellations as shown in Fig. 5, and an optimal HNLF input power of around 38.5 mW is found by sweeping the power from 15 mW to 80 mW so that a maximum achievable rate is realized via joint optimization of gateway input power and DNN receiver model in an end-to-end fashion.

Figure 7(a) shows the bit error ratio (BER) performance of the PDT-QPSK signal as a function of the SNR for the PAM4 signal. In the case of conventional linear equalization, the best BER is obtained when the HNLF input power is 55 mW (i.e., regular QPSK). When the HNLF input power is reduced to 38.5 mW, BER performance is degraded in particular for higher SNR regimes. Note that the decision boundary is adjusted by the Gaussian Naïve Bayes (GNB) which analyzes the mean and variance for the irregular constellations. Even though the GNB can handle non-uniform noise statistics for soft-decision calculation, the performance is considerably degraded especially for PDT-QPSK signals because the GNB still assumes element-wise independent AWGN models. Whereas, in the case of the DNN-based decision, the BER degradation can be fully compensated and the BER performance at the HNLF input power of 38.5 mW is slightly improved compared to the standard 55 mW case. This is a benefit of the DNN decision and optical constellation tuning, which controls phase shifts according to non-uniform phase distortion.

Since the DNN-based signal recovery produces soft-decision LLR values, the shaping gain is more remarkable in terms of the GMI metric against the traditional BER metric. Figure 7(b) shows the GMI performance of the PDT-QPSK signal as a function of the PAM4 signal SNR.

When the optical HNLF input power is 38.5 mW, the achievable gain at a GMI of 0.8 will be nearly 1 dB compared to that for standard QPSK phase shifts with 55 mW input power. In addition, the DNN signal recovery exhibits outstanding performance improvement of greater than 3 dB over conventional linear equalization that employs the soft-decision GNB. As a consequence, the proposed PDT and DNN-based decision can conserve the required HNLF input power. For reference, we also present the performance with a nonlinear equalization based on 3rd order Volterra kernel before GNB soft decision. Although the Volterra nonlinear equalization can improve the performance over the linear counterpart, the DNN still exhibits the remarkable superiority. Even though DNN equalization is computationally complex in general, there exist quite a few methods enabling practical deployment with a low hardware complexity, e.g., sparse multiplier-less DNN architecture [22].

We demonstrated that the DNN-based signal recovery improves BER and GMI performance over the conventional recovery based on GNB. We next compare the DNN method with the other classical machine learning methods such as support vector machine (SVM). Figure 8(a) shows the GMI performance of the PDT-QPSK signal at an optical HNLF input power of 38.5 mW. For comparison, we present the performance of SVM, logistic regression (LR) and quadratic discriminant analysis (QDA), using scikit-learn framework. It is confirmed that the ResNet can significantly outperform them. It is partly because the ResNet is directly optimized to minimize the binary cross-entropy loss which determines the lower bounds of the GMI metric, whereas the other classical machine learning methods are not optimized for GMI maximization and do not provide accurate soft-decision probabilities in general.

Although ResNet achieves the best performance, training an over-parameterized DNN usually requires a massive amount of training data. Figure 8(b) shows the impact of the training dataset size on GMI performance for ResNet in comparison to LR, SVM, and QDA. As expected, the performance of such machine learning techniques degrades when the available amount of training data is decreased. Interestingly, we can observe that training ResNet is relatively robust against the training data deficiency; specifically, ResNet having more than 20,000 trainable parameters could be trained to outperform SVM even when the number of training bits is about 100. Among the classical techniques, LR is the best when sufficient training data is available, QDA is the best for a moderate training size, and SVM is the best for a limited training size. Whereas, ResNet achieves the best GMI performance across the whole range of training data size in consideration. Note that the original ResNet paper [23] also showed the similar capability of training very deep networks by using 400-fold smaller size of training data than the total number of trainable parameters (19.4M).

Finally, we evaluate the impact of network depth in Figure 8(c), where we compare DNN with and without skip connections. The DNN without skip connections is denoted as multi-layer perceptron (MLP). The advantage of skip connections lies in its robustness against data deficiency and excessive depth. When the available training data is deficient, increasing the neural network depth does not always improve the performance, whereas ResNet can prevent a significant degradation due to the excessive depth. Hence, the use of skip connections is beneficial especially when we have no prior knowledge of best hyperparameters (the layer depth, hidden node width, nonlinear activations, etc.). It may be further improved by additional tricks such as stochastic depth, zero-starting batch normalization, and automated machine learning (AutoML). We leave those research topics as a potential future work since

our main focus is a proof-of-concept study of the optical gateway whose irregular phase tuning is designed jointly with a DNN equalization at the receiver in an end-to-end fashion.

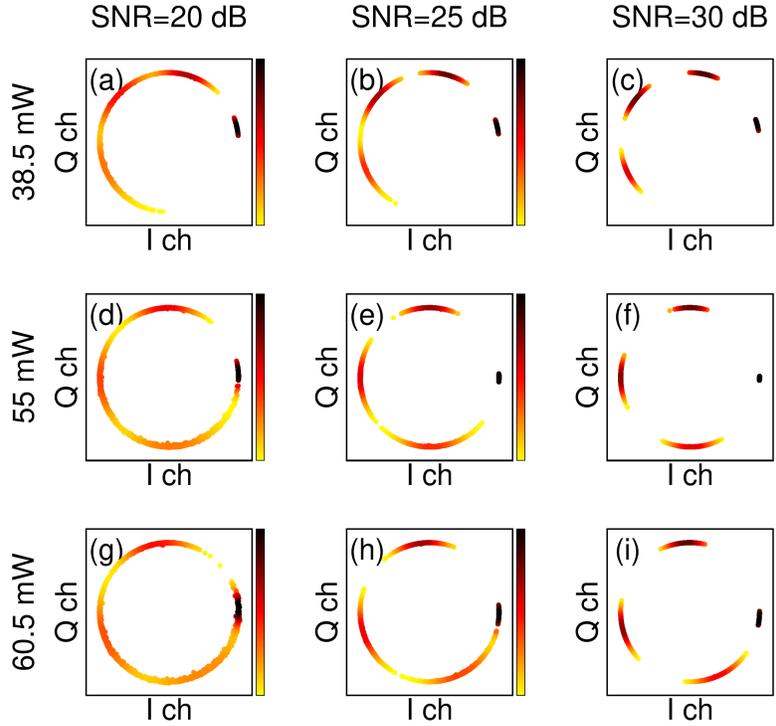


Fig. 5. Constellation maps of PDT-QPSK signal.

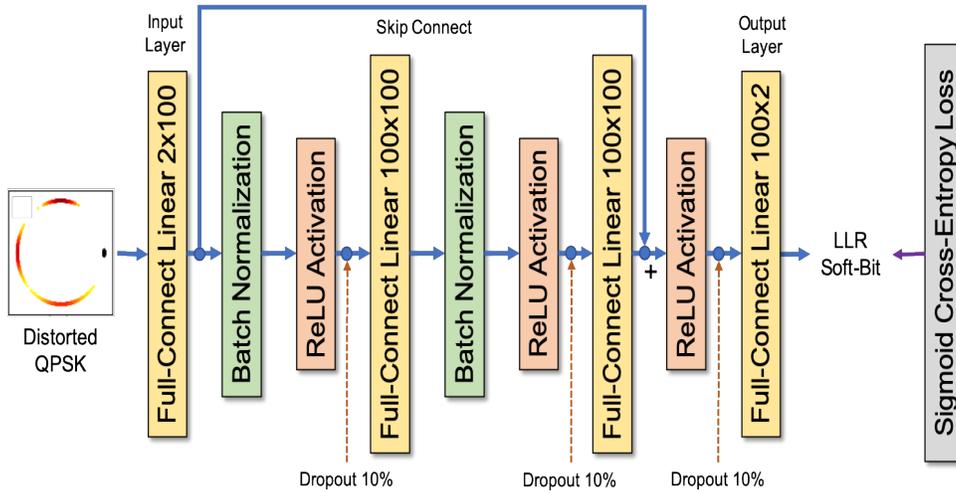
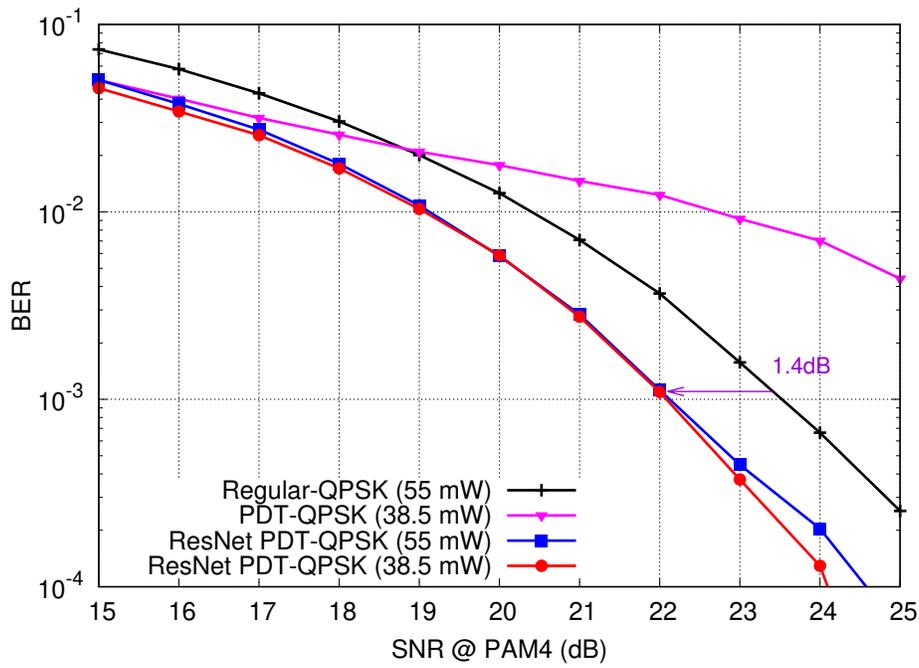
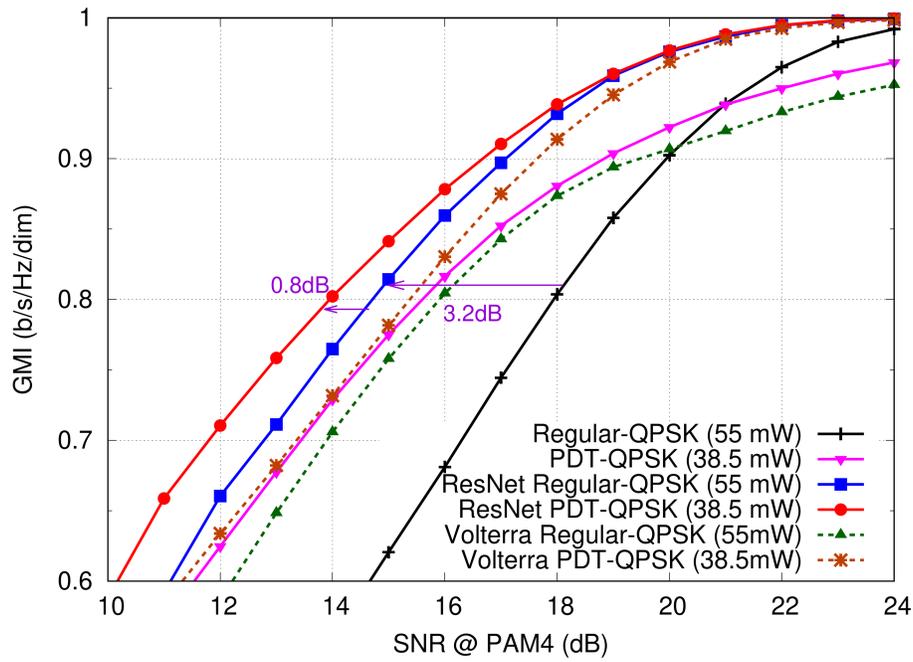


Fig. 6. ResNet-based signal recovery.

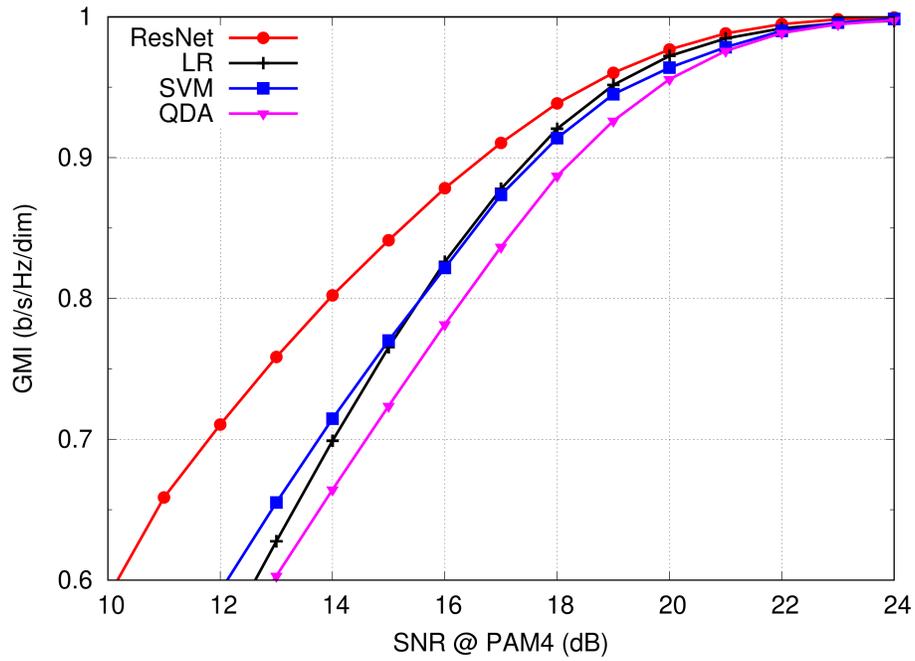


(a) BER

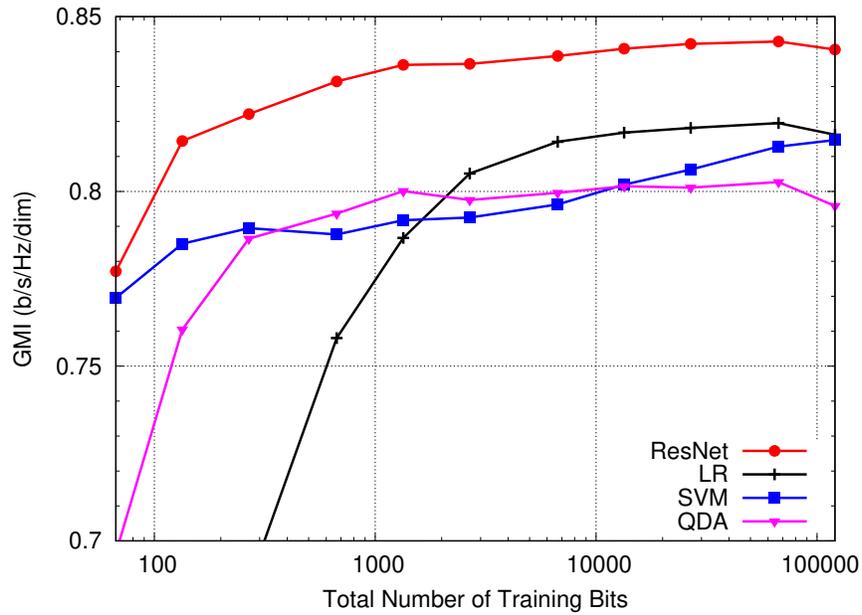


(b) GMI

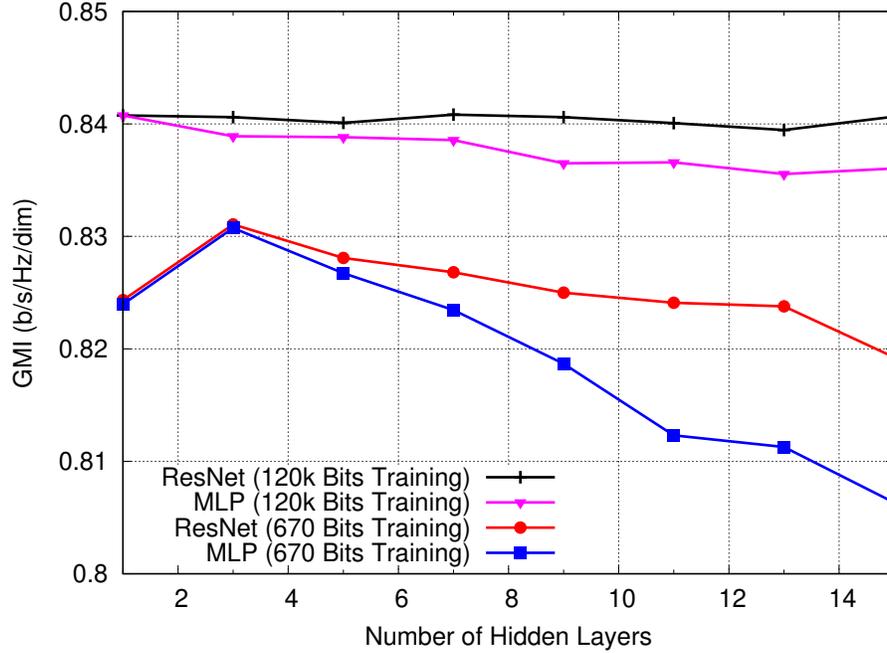
Fig. 7. GMI performance as a function of SNR for PAM4 in HNGN modulation converter.



(a) Comparison of Classical Machine Learning Methods



(b) Training Data Size Impact (15dB SNR @ PAM4)



(c) DNN Depth Impact with/without Skip Connection (15dB SNR @ PAM4)

Fig. 8. GMI performance of PDT-QPSK at 38.5mW.

5. Conclusions

We proposed an irregular PDT-QPSK generation for optical PAM4-to-QPSK conversion at an HNGN. A 10 Gbaud PDT-QPSK signal with DNN-based decision reduces the HNLF input power of the PAM4 signal compared to that for the original QPSK signal with conventional linear equalization or Volterra nonlinear equalization. For soft-decision data recovery of the irregular PDT-QPSK constellations, it was confirmed that ResNet is the best algorithm over various machine learning algorithms. Specifically, the optimal DNN-based signal recovery achieves about 3 dB gain, and additional gain of nearly 1 dB can be achieved by tuning the input power of the optical gateway. To the best of our knowledge, there is no other report that applies deep learning techniques to mitigate nonlinear distortion occurring in the HNGN.

Acknowledgments. We thank T. Miyazaki and M. Hanawa of the University of Yamanashi for their support in the experiments.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. A. Tzanakaki, M. Anastasopoulos, and D. Simeonidou, "Converged access/metro infrastructures for 5G services," *Proc. Optical Fiber Communication Conference (OFC)*, San Diego, USA, Mar. 2018, M2A.3.
2. J. Li, N. Hua, and Y. Yu, "A flexible low-latency metro-access converged network approach based on time-synchronized TWDM-PON," *Proc. Optical Fiber Communication Conference (OFC)*, San Diego, USA, Mar. 2018, Th2A.50.
3. W. Jin, C. F. Zhang, X. Duan, M. R. Kadhum, Y. X. Dong, R. P. Giddings, N. Jiang, K. Qiu, and J. M. Tang, "Improved performance robustness of DSP-enabled flexible ROADMs free from optical filters and O-E-O conversions", *IEEE Journal of Optical Communications and Networking* **8**(8), 521–529 (2016).

4. M. Ruffini, M. Achouche, A. Arbelaez, R. Bonk, A. D. Giglio, N. J. Doran, M. Furdek, R. Jensen, J. Montalvo, N. Parsons, T. Pfeiffer, L. Quesada, C. Raack, H. Rohde, M. Schiano, G. Talli, P. Townsend, R. Wessaly, L. Wosinska, X. Yin, and D. B. Payne, "Access and metro network convergence for flexible end-to-end network design," *IEEE Journal of Optical Communications and Networking* **9**(6), 524–535 (2017).
5. Y-H. Wen and K-M. Feng, "A simple NRZ-OOK to PDM RZ-QPSK optical modulation format conversion by bidirectional XPM," *IEEE Photonics Technology Letters* **27**(9), 935–938 (2015).
6. K. Mishina, S. Kitagawa, and A. Maruta, "All-optical modulation format conversion from on-off-keying to multiple-level phase-shift-keying based on nonlinearity in optical fiber," *Optics Express* **15**(13), 8444–8453 (2007).
7. G. W. Lu, E. Tipsuwannakul, T. Miyazaki, C. Lundstrom, M. Karlsson, and P. A. Andrekson, "Format conversion of optical multilevel signals using FWM-based optical phase erasure," *IEEE Journal of Lightwave Technology*, **29**(16), 2460–2466 (2011).
8. W. Astar and G. M. Carter, "Mitigation of XPM penalty in 10-Gb/s OOK/DBPSK mixed modulation format 50-GHz-spaced WDM transmission by conversion of OOK to BPSK using an SOA," *IEEE Photonics Technology Letters* **20**(20), 1715–1717 (2008).
9. T. Kodama and T. Miyazaki, "CW-probe-less OOK and BPSK to QPSK optical modulation format conversion and SSMF transmission," *Proc. Conference on Lasers and Electro-Optics (CLEO)*, San Jose, USA, May 2019, JTh2A.79.
10. A. E. Wilner, A. Fallahpour, and F. Alishahi, "All-optical signal processing techniques for flexible networks," *IEEE Journal of Lightwave Technology* **37**(1), 21–35 (2019).
11. Z. Yang, H. Wang, and Y. Ji, "Tunable all-optical format conversion for BPSK to OOK based on highly nonlinear optical loop mirror," *Optics Communications* **473**(15), 125907 (2020).
12. B. Zhang, H. Zhang, and C. Yu, "An all-optical modulation format conversion for 8QAM based on FWM in HNLf," *IEEE Photonics Technology Letters*, **25**(4), 327–330 (2012).
13. F. D. Ros, K. Dalgaard, L. Lei, J. Xu, and C. Peucheret, "QPSK-to-2×BPSK wavelength and modulation format conversion through phase sensitive four-wave mixing in a highly nonlinear optical fiber," *Opt. Express* **21**(23), 28743–28750 (2013).
14. H. Kishikawa, N. Goto, and L. R. Chen, "All-optical wavelength preserved modulation format conversion from PDM-QPSK to PDM-BPSK using FWM and interference," *IEEE Journal of Lightwave Technology*, **34**(23), 5505–5515 (2016).
15. J. Cui, H. Wang, and Y. Ji, "Optical modulation format conversion from one QPSK to one BPSK with information-integrity-employing phase-sensitive amplifier," *Applied Optics*, **56**(18), 5307–5312 (2017).
16. H. Yamazaki, H. Takahashi, T. Goh, Y. Hashizume, S. Mino, and Y. Miyamoto, "Linear optical IQ modulator for high-order multilevel coherent transmission," *Proc. Optical Fiber Communication Conference (OFC)*, Anaheim, USA, Mar. 2013, OM3C.1.
17. V. Bobrovs, S. Spolitis, and G. Ivanovs, "Latency causes and reduction in optical metro networks," *Proc. SPIE 9008, Optical Metro Networks and Short-Haul Systems VI*, Feb. 2014, 90080C.
18. G. Brown, "New transport network architecture for 5G RAN," Fujitsu, Kanagawa, Japan, White Paper, 2018. [Online]. Available: https://networking.report/Resources/Whitepapers/34e518ed-c067-4579-adf1-38441f13bc66_New-Transport-Network-Architectures-for-5G-RAN.pdf 38441f13bc66_New-Transport-Network-Architectures-for-5G-RAN.pdf
19. T. Kodama, T. Miyazaki, and M. Hanawa, "Seamless PAM-4 to QPSK modulation format conversion at gateway for short-reach and long-haul integrated networks," *European Conference on Optical Communication (ECOC)*, Rome, Italy, Sept. 2018, We3H.1.
20. M. Schaedler, S. Calabro, F. Pittala, G. Bocherer, M. Kuschnerov, C. Bluemmm, and S. Pachnicke, "Neural network assisted Geometric shaping for 800Gbit/s and 1Tbit/s optical transmission," *Proc. Optical Fiber Communication Conference (OFC)*, San Diego, USA, Mar. 2020, M1G.1.
21. V. Talreja, T. Koike-Akino, Y. Wang, D. S. Millar, K. Kojima, and K. Parsons, "End-to-end deep learning for phase noise-robust multi-dimensional geometric shaping," *European Conference on Optical Communication (ECOC)*, Brussels, Belgium, Sept. 2020, Th1D.2.
22. T. Koike-Akino, Y. Wang, K. Kojima, K. Parsons, and T. Yoshida, "Zero-multiplier sparse DNN equalization for fiber-optic QAM systems with probabilistic amplitude shaping," *European Conference on Optical Communication (ECOC)*, Sept. 2021.
23. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016 (pp. 770–778).