



Achievable information rate optimization in C-band optical fiber communication system

Zheng Liu¹ · Tianhua Xu^{1,2,3} · Ji Qi¹ · Joshua Uduagbomen² · Jian Zhao¹ · Tiegen Liu¹

Received: 14 February 2023 / Accepted: 19 April 2023
© The Author(s) 2023

Abstract

Optical fiber communication networks play an important role in the global telecommunication network. However, nonlinear effects in the optical fiber and transceiver noise greatly limit the performance of fiber communication systems. In this paper, the product of mutual information (MI) and communication bandwidth is used as the metric of the achievable information rate (AIR). The MI loss caused by the transceiver is also considered in this work, and the bit-wise MI, generalized mutual information (GMI), is used to calculate the AIR. This loss is more significant in the use of higher-order modulation formats. The AIR analysis is carried out in the QPSK, 16QAM, 64QAM and 256QAM modulation formats for the communication systems with different communication bandwidths and transmission distances based on the enhanced Gaussian noise (EGN) model. The paper provides suggestions for the selection of the optimal modulation format in different transmission scenarios.

Keywords Optical fiber communication · Achievable information rate · Mutual information · Generalized mutual information

1 Introduction

Over 95% of digital data traffic is carried over optical fiber networks [1]. The information transmission rate of optical fiber communication systems limits the communication rate of the global telecommunication networks. With the development of fiber communication technology, larger communication bandwidth and higher symbol rate are realized to transmit more bits within one second. However, severe nonlinear effects also occur and lead to fewer valid bits transmitted per second. Meanwhile, the equalization enhanced phase noise (EPPN) further decreases the signal quality [2]. In other words, the effective communication rate is limited by nonlinear effects and transmission noise. This

phenomenon is more obvious when higher-order modulation formats are applied. Generally, a higher modulation format means a higher symbol error rate (SER) [3, 4]. However, using high-order modulation format can transmit more bits by each symbol. Therefore, it is not sufficient to use the signal-to-noise ratio (SNR) to evaluate the performance of the communication system. To reasonably measure the communication capability, the transmission bit rate that the system can effectively support should be used as the metric. Generalized mutual information (GMI) can be used to measure the effective transmission bit rate of the system. For wavelength division multiplexed (WDM) systems, more channels can be used to transmit signals at the same time to achieve higher data rates. Although the larger bandwidth will further reduce the SNR due to the inter-channel interactions, the performance penalty is much less than the information rate gain arising from the use of more channels [5]. Therefore, this paper employs the number of bits effectively transmitted in one second as the metric of achievable information rate (AIR). The enhanced Gaussian noise (EGN) model is applied to analyze the performance of the optical fiber system under different conditions. Finally, the optimal modulation format is obtained by comprehensive analyses of different transmission scenarios. Discussions are conducted

✉ Tianhua Xu
tianhua.xu@ieee.org

✉ Jian Zhao
enzhaojian@tju.edu.cn

¹ School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

² School of Engineering, University of Warwick, Coventry CV4 7AL, UK

³ Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK

provide an optimization direction for future high-capacity optical fiber communication systems.

This paper evaluates different communication scenarios in terms of effective bit rates that can be efficiently transmitted. Such metric provides a fair comparison of systems, and the results have fundamental implications and provide insightful suggestions for follow-up research. The conclusions in this paper are based on systems without the application of forward error correction (FEC) techniques [6, 7]. Different types of FEC codes have different error-correction capabilities, and research of AIR in this case only needs to carry out a further-step based on our results. Moreover, the impact of introducing error correction codes on the transmission bit rate is linear, so the conclusions in this paper are insightful and applicable for systems with FECs.

This paper is arranged as follows. The GMI and MI are introduced in Sect. 2. Section 3 discusses the EGN model. The results and discussion can be found in Sect. 4 and some proposals for the future are presented in Sect. 5.

2 Generalized mutual information

Mutual information (MI) is a measure of the amount of information that two random variables share. It quantifies the degree to which knowledge of one variable reduces uncertainty about the other variable. For communication signals, the higher the MI between the transmitter and the receiver is, the better the communication quality is. This means that more information is correctly transmitted. The Shannon limit is used to measure the channel capacity, by calculating the MI between the signals before entering the channel and the signals when leaving the channel. However, the receiver will still cause loss in MI. Therefore, the signals used in the calculation are expanded into bit sequences, as shown in Fig. 1, and the information rate is calculated based on GMI.

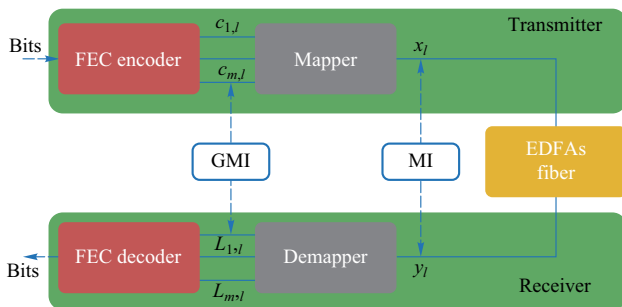


Fig. 1 Schematic of MI and GMI

Suppose that the modulated bit signal at time l is $\{c_{1,l}, c_{2,l}, \dots, c_{m,l}\}$. Every m bits will be mapped into one symbol x_l and $x_i \in X, card(X) = M$, where M is order of the modulation format and X is modulated symbol set. The symbol obtained after passing through the channel is y_i , where $y_i \in Y$. Then the symbols will be demapped to bit sequences $\{L_{1,l}, L_{2,l}, \dots, L_{m,l}\}$. MI and GMI can be estimated using following formulas [8]

$$MI = I(X : Y) = \frac{1}{M} \sum_{i=1}^M \int_{C^N} f_{Y|X}(y|x_i) \log_2 \frac{f_{Y|X}(y|x_i)}{\frac{1}{M} \sum_{j=1}^M f_{Y|X}(y|x_j)} dy, \quad (1)$$

$$GMI = \sum_{k=1}^m \mathbb{E}_{B_k, Y} \left[\log_2 \frac{f_{Y|B_k}(Y|B_k)}{\frac{1}{2} \sum_{b \in \{0,1\}} f_{Y|B_k}(Y|b)} \right] = \frac{1}{M} \sum_{k=1}^m \sum_{b \in \{0,1\}} \sum_{i \in \mathcal{I}_m^b} \int_{C^N} f_{Y|X}(y|x_i) \log_2 \frac{\sum_{j \in \mathcal{I}_k^b} f_{Y|X}(y|x_j)}{\frac{1}{2} \sum_{p=1}^M f_{Y|X}(y|x_p)} dy, \quad (2)$$

where $\mathcal{I}_m^b \subset \{1, 2, \dots, M\}$ with $card(\mathcal{I}_m^b) = M/2$ is the set of indices of constellations points, of which the binary label is b at the bit position m . $f_{Y|X}(y|x)$ is the channel transfer function and C^N is the N -dimension constellation map. B_k is the bit sequence. \mathbb{E} is the expectation. In practice, GMI and MI can be calculated using the Monte-Carlo integration method and the results are shown in Fig. 2 [9].

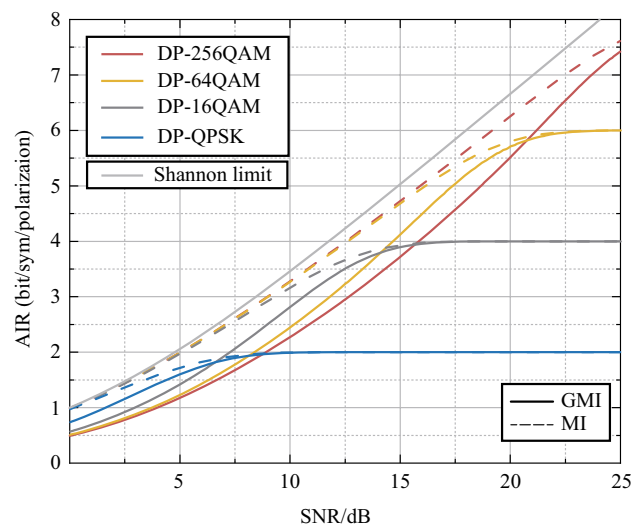


Fig. 2 GMI and MI of DP-QPSK, DP-16QAM, DP-64QAM and DP-256QAM, DP: dual polarization

3 Enhanced Gaussian noise model

Due to the existence of nonlinear effects, signal propagation in the fiber is very complicated. It is impossible to provide explicit expressions for signal transitions. However, the nonlinear effects of the channel are not very strong near the optimal power, where the behavior of signal propagation is close to linear signal propagation. This is the basic assumption of the perturbation-based Gaussian noise model. Poggiolini et al. proposed the EGN model for quickly estimating the SNR of optical fiber communication systems [10, 11]. In this paper, the EGN model is used to quickly calculate the channel SNR, and then the EGN-based evaluation of the corresponding nonlinear interference is added to estimate the system GMI. The EGN model in the C-band can be approximately expressed as [12, 13]

$$SNR = \frac{P}{\sigma^2 + \sigma_{s-s}^2 + \sigma_{s-n}^2}, \tag{3}$$

$$\sigma^2 = \sigma_{TRx}^2 + \sigma_{ASE}^2, \tag{4}$$

$$\sigma_{s-s}^2 = N_s^{\epsilon+1} \eta P^3, \tag{5}$$

$$\sigma_{s-n}^2 \approx 3 \left(\frac{N_s^{\epsilon+1}}{2} + \frac{N_s^{\epsilon+2}}{\epsilon + 2} \right) \eta \sigma_{ASE}^2 P^2 + 3N_s^{\epsilon+1} \eta \kappa P^3, \tag{6}$$

where P , σ_{ASE}^2 , σ_{TRx}^2 are the signal power, the amplifier spontaneous emission (ASE) noise power and the transceiver noise power, respectively. N_s is the number of fiber spans.

$$\epsilon = \frac{3}{10} \log \left[1 + \frac{6}{L_s} \frac{L_{eff}}{\sinh^{-1} \left(\frac{\pi}{2} |\beta_2| R_s^2 N_{ch}^2 L_{eff} \right)} \right], \tag{7}$$

$$\eta \approx \frac{8}{27} \frac{\alpha \gamma^2 L_{eff}}{\pi |\beta_2| R_s^2} \sinh^{-1} \left(\frac{\pi^2}{2} |\beta_2| L_{eff} N_{ch}^2 R_s^2 \right) - \frac{80}{81} \frac{\kappa \gamma^2 L_{eff}^2}{\pi |\beta_2| L_s R_s^2} \left[\Phi \left(\frac{N_{ch} + 1}{2} \right) + C + 1 \right], \tag{8}$$

where $L_{eff} = (1 - e^{-\alpha L_s})/\alpha$, α is the fiber attenuation coefficient, β_2 is the second-order dispersion coefficient, N_{ch} is the number of WDM channels, R_s is the symbol rate, $C \approx 0.557$ is the Euler-Mascheroni constant and γ is fiber nonlinear coefficient. L_s is the span length. $\Phi(x)$ is the digamma function and κ is a constant equals to 1, 17/25, 13/21, and 121/200 corresponding to QPSK, 16QAM, 64QAM, and 256QAM, respectively [11]. The accuracy of EGN model in the C-band has already been verified by other scholars as well in our previous works [14–17].

4 Results and discussion

For a Nyquist-spaced optical communication system, according to the Nyquist sampling theorem, the number of symbols transmitted per second can be measured via the bandwidth of the system. The value of GMI represents the effective number of bits in a symbol. Multiplying the bandwidth by the GMI gives the effective number of bits per second, transmitted over each polarization mode. This paper studies the communication scenario of a 80 km per span 32 GBaud fiber communication system with different modulation formats, transmission distances and bandwidths. Results of AIR versus transmission distances and bandwidths are shown in Fig. 3.

The MI degradation at the receiver is particularly severe for higher-order modulation formats, as shown in Fig. 2. When the SNR is low, the GMI of the high-order modulation format drops sharply, and it can be even lower than that of the low-order format at low SNR region. Moreover, higher-order modulation formats are more significantly affected by the noise, resulting in a more severe GMI degradation. It is demonstrated that higher-order modulation formats show their advantages in the case of shorter transmission distances or smaller communication bandwidths. For systems with long transmission distances and large bandwidths, some low-order modulation formats can be more robust and appropriate. Figure 4 shows the optimal modulation format for different transmission situations.

For terrestrial communication systems, the common fiber span length is 80 km, and the transmission distance is less than 10000 km. When the symbol rate is 32 GBaud and the transmission distance exceeds 2000 km, the modulation format of 16QAM can always obtain the highest AIR. When the transmission distance is reduced to between 240 and 2000 km, the modulation scheme of 64QAM becomes the most suitable format. The 256QAM signal can surpass the other three modulation formats only when the transmission distance is lower than 240 km.

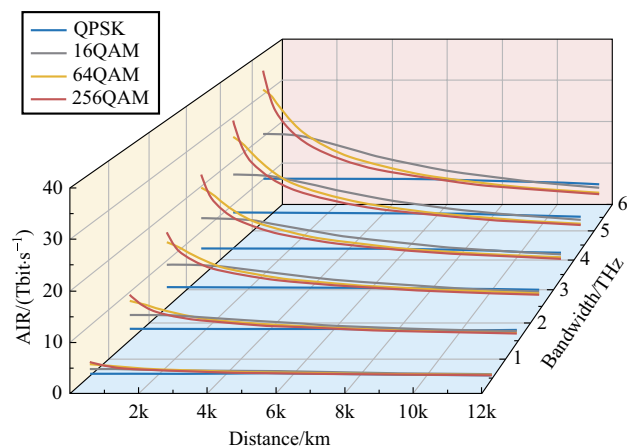


Fig. 3 AIR versus transmission distance and communication bandwidth. Symbol rate is 32 GBaud and each fiber span is 80 km

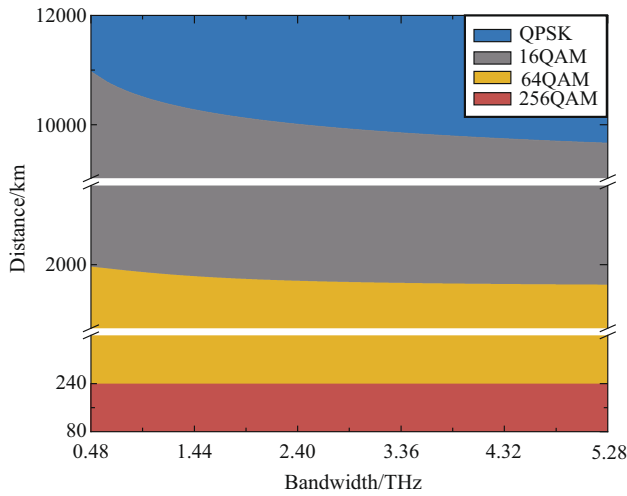


Fig. 4 The optimal modulation formats under different transmission distances and communication bandwidths. Symbol rate is 32 GBaud and each fiber span is 80 km

To study higher symbol rate systems, we fixed the transmission distance as 8000 km. Figure 5 shows the GMI with different symbol rates and different communication bandwidths at transmission distance of 8000 km and fiber span of 80 km.

Every curve in Fig. 5 is almost as straight line, and this means that the GMI is weakly correlated with the symbol rate. However, increase of the communication rate can save the number of channels for WDM transmission and hence saves the related component sets cost. Therefore, higher-speed transmitters have more efficient AIR per transmitter. Meanwhile, the GMI behaves almost independently of the symbol rate, and so 16QAM can still obtain the best performance at 8000 km as shown in Fig. 4.

A submarine communication system with a span length of 50 km is also studied. Compared with the system with 80 km

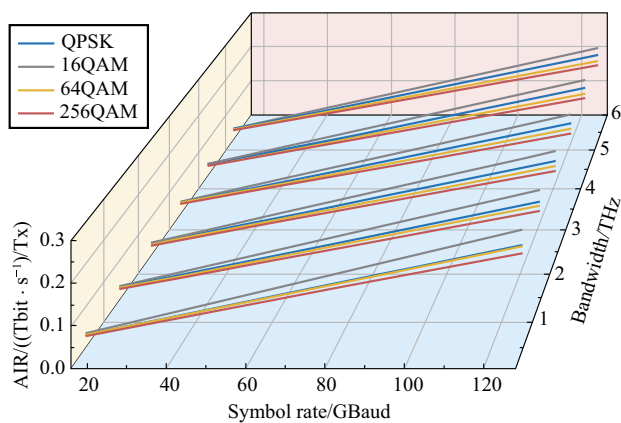


Fig. 5 AIR per transmitter versus symbol rate and communication bandwidth. Transmission distance is 8000 km and each fiber span is 80 km

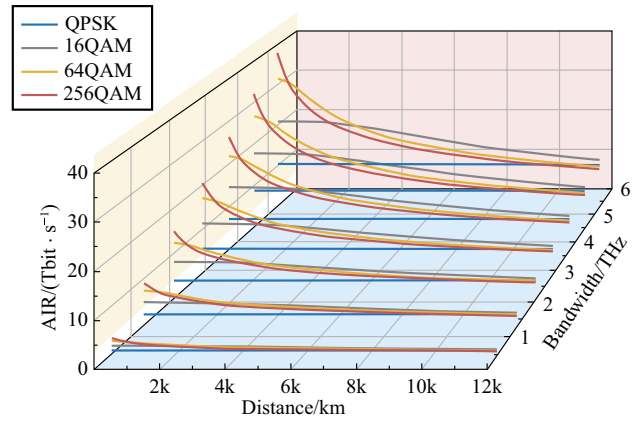


Fig. 6 AIRs versus transmission distances and communication bandwidths. Symbol rate is 32 GBaud and each fiber span is 50 km

span, shortening the span to 50 km can significantly improve the system SNR [14], so the higher-order modulation formats could benefit from this. The result is shown in Fig. 6.

The intersection between curves with different colors in the same group moves towards a longer communication distance when higher-order modulation is used. This proves that the higher-order modulation format obtains more improvement than the lower-order format with the increase of system SNR. Since current transmission scenario refers to a submarine system, we focus on the scenario where the communication distance exceeds 8000 km. When the span length is 50 km, it can be found that the QPSK modulation format can almost reach the maximum GMI (2 bit/sym/polarization). This is also the reason why QPSK format is widely used in current submarine communications. However, the 16QAM modulation format also obtains a great improvement, and the use of the 16QAM format within 12000 km can significantly improve the system AIR, especially for larger bandwidth.

In summary, symbol rate has little effect on the system GMI, but the application of a higher symbol rate can effectively reduce the number of required transceivers and link components. For long distance (2000–10000 km) terrestrial communication systems with 80 km per span, the 16QAM format can get the highest AIR. For the submarine communication systems with each fiber span of 50 km [18], the 16QAM shows more significant performance improvement compared to the QPSK format. In a terrestrial communication system or a submarine communication system, it can be seen that the communication bandwidth has marginal effects on the SNR, as shown in Fig. 5. Therefore, trade-off between high-speed transmitters and the number of channels is important when designing new fiber optic systems. For the convenience of use, we list the results (optimal selection of the modulation format) for bandwidth exceeding 2.4 THz as in the following Tables 1 and 2.

Table 1 Optimal modulation format for bandwidth of over 2.4 THz and span distance of 80 km

Transmission distance	Modulation format
< 240 km	256 QAM
240–2000 km	64 QAM
2000–10000 km	16 QAM
> 10000 km	QPSK

Table 2 Optimal modulation format for bandwidth of over 2.4 THz and span distance of 50 km

Transmission distance	Modulation format
< 500 km	256 QAM
500–3000 km	64 QAM
3000–12000 km	16 QAM
> 12000 km	QPSK

5 Proposals for the future

The MI of the high-order modulation format is always higher than that of the low-order format. However, the GMI of the higher-order modulation format could be lower than that of the lower-order format due to the loss of information caused by the transceiver. Therefore, the use of more advanced transceivers can be an effective solution. In fact, the SNR difference between each modulation format is very small, especially when the modulation order is higher than 4 (equal to or above 16QAM) [19]. Various methods that can reduce the information loss at the receiver side or shift the intersection between solid lines with different colors (modulation formats) to the left (low SNR region) in Fig. 2 will be an interesting research direction for next-generation fiber communication systems. On the other hand, another hot research direction use various approaches, such as constellation shaping and waveform shaping [20], to improve the GMI of the optical fiber system, thereby shifting the dotted line in Fig. 2 closer to the Shannon limit (the gray line). Optical fiber communication systems, despite still having a long way to go, will eventually become the cornerstone of future telecommunication networks.

Acknowledgements This work was supported by the National Key Research and Development Program of China (No. 2022YFE0202100), EU Horizon 2020 MSCA Grant 101008280 and UK Royal Society Grant (IES\R3\223068).

Author contributions All authors contributed to the writing and revision of this paper.

Availability of data and materials The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Declarations

Competing interests The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Cisco Visual Networking Index (VNI) Complete Forecast Update, 2017–2022. Available at website of [cisco.com/c/dam/m/en_us/network-intelligence/service-provider/digital-transformation/knowledge-network-webinars/pdfs/1213-business-services-ckn.pdf](http://www.cisco.com/c/dam/m/en_us/network-intelligence/service-provider/digital-transformation/knowledge-network-webinars/pdfs/1213-business-services-ckn.pdf) (2018)
2. Jin, C., Shevchenko, N.A., Li, Z., Popov, S., Chen, Y., Xu, T.: Nonlinear coherent optical systems in the presence of equalization enhanced phase noise. *J. Lightwave Technol.* **39**(14), 4646–4653 (2021)
3. Ives, D.J., Bayvel, P., Savory, S.J.: Adapting transmitter power and modulation format to improve optical network performance utilizing the gaussian noise model of nonlinear impairments. *J. Lightwave Technol.* **32**(21), 3485–3494 (2014)
4. Saif, W.S., Ragheb, A.M., Nebendahl, B., Alshawi, T., Marey, M., Alshebeili, S.A.: Performance investigation of modulation format identification in super-channel optical networks. *IEEE Photon. J.* **14**(2), 1–10 (2022)
5. Semrau, D., Sillekens, E., Killey, R.I., Bayvel, P.: The benefits of using the s-band in optical fiber communications and how to get there. In: Proceedings of 2020 IEEE Photonics Conference (IPC). IEEE, pp. 1–2 (2020)
6. Tzimpragos, G., Kachris, C., Djordjevic, I.B., Cvijetic, M., Soudris, D., Tomkos, I.: A survey on FEC codes for 100 G and beyond optical networks. *IEEE Commun. Surv. Tutor.* **18**(1), 209–221 (2014)
7. Zokaei, A., Truhachev, D., El-Sankary, K.: Memory optimized hardware implementation of open FEC encoder. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **30**(10), 1548–1552 (2022)
8. Alvarado, A., Fehenberger, T., Chen, B., Willems, F.M.J.: Achievable information rates for fiber optics: applications and computations. *J. Lightwave Technol.* **36**(2), 424–439 (2018)
9. James, F.: Monte Carlo theory and practice. *Rep. Prog. Phys.* **43**(9), 1145 (1980)
10. Carena, A., Bosco, G., Curri, V., Jiang, Y., Poggiolini, P., Forghieri, F.: EGN model of non-linear fiber propagation. *Opt. Express* **22**(13), 16335–16362 (2014)
11. Poggiolini, P.: The GN model of non-linear propagation in uncompensated coherent optical systems. *J. Lightwave Technol.* **30**(24), 3857–3879 (2012)
12. Poggiolini, P., Jiang, Y., Carena, A., Forghieri, F.: A simple and accurate closed-form EGN model formula. *arXiv e-prints*, 1503–04132. [arXiv:1503.04132](https://arxiv.org/abs/1503.04132) (2015)
13. Shevchenko, N.A., Xu, T., Lavery, D., Liga, G., Ives, D.J., Killey, R.I., Bayvel, P.: Modeling of nonlinearity-compensated optical

- communication systems considering second-order signal-noise interactions. *Opt. Lett.* **42**(17), 3351–3354 (2017)
14. Liu, Z., Xu, T., Jin, C., Xu, T., Tan, M., Zhao, J., Liu, T.: Analytical optimization of wideband nonlinear optical fiber communication systems. *Opt. Express* **30**(7), 11345–11359 (2022)
 15. Nespola, A., Huchard, M., Bosco, G., Carena, A., Jiang, Y., Poggiolini, P., Forghieri, F.: Experimental validation of the EGN-model in uncompensated optical links. In: *Proceedings of 2015 Optical Fiber Communications Conference and Exhibition (OFC)*. OSA, pp. 1–3 (2015)
 16. Seve, E., Pesic, J., Pointurier, Y.: Accurate QoT estimation by means of a reduction of edfa characteristics uncertainties with machine learning. In: *Proceedings of 2020 International Conference on Optical Network Design and Modeling (ONDM)*. IEEE, pp. 1–3 (2020)
 17. Liga, G., Chen, B., Barreiro, A., Alvarado, A.: Modeling of nonlinear interference power for dual-polarization 4D formats. In: *Proceedings of 2021 Optical Fiber Communications Conference and Exhibition (OFC)*. OSA, pp. 1–3 (2021)
 18. Zhong, K., Zhou, X., Wang, Y., Gui, T., Yang, Y., Yuan, J., Wang, L., Chen, W., Zhang, H., Man, J., Zeng, L., Yu, C., Lau, A.P.T., Lu, C.: Recent advances in short reach systems. In: *Proceedings of Optical Fiber Communication Conference*. OSA, pp. 2–7 (2017)
 19. Xu, T., Shevchenko, N.A., Lavery, D., Semrau, D., Liga, G., Alvarado, A., Killey, R.I., Bayvel, P.: Modulation format dependence of digital nonlinearity compensation performance in optical fibre communication systems. *Opt. Express* **25**(4), 3311–3326 (2017)
 20. Qu, Z., Djordjevic, I.B.: On the probabilistic shaping and geometric shaping in optical communication systems. *IEEE Access* **7**, 21454–21464 (2019)



Zheng Liu is a Ph.D. student in the Department of Optical Engineering at School of Precision Instrument and Opto-electronics Engineering, Tianjin University, China. His research interests include optical communications, optical networking, artificial intelligence, and digital signal processing.



Tianhua Xu (Member, IEEE) has authored or coauthored more than 120 international journal and conference papers. His research interests include optical communications, optical sensing, optical networking, and machine learn and digital signal processing. He has been the Technical Program Committee (TPC) Co-Chair/Member of more than 20 IEEE conferences, e.g., GLOBECOM and ICC, and the Chair with Optics and Digital Systems Technical Group, Optical Society of America.



Ji Qi is currently studying as a research M.Sc. student in the Department of Optical Engineering at School of Precision Instrument and Opto-electronics Engineering, Tianjin University, China. His research topics include optical communications, signal processing and machine learning.



Joshua Uduagbomen is currently pursuing a Ph.D. degree in Electrical and Electronics Engineering from University of Warwick, UK. He received an M.Sc. degree in Microelectronics and Communications Engineering from Northumbria University, UK in 2020. His research interests include artificial intelligence for high capacity optical communication networks, nonlinear fiber optics, physics informed neural networks and scientific computing.



Jian Zhao received his Ph.D. degree from Hong Kong Polytechnic University, China in 2010. He has led or participated in a few scientific research projects, including the 973 project, the National Major Instrument Project, the National Natural Science Foundation of China Instrument Project, etc. He has published more than 50 high level papers. His research interests include fiber optic communication, space division multiplexing, coherent optical transmission, optical signal processing, nonlinear optics, fiber optic sensing, DSP, fiber optic design, sign, deep learning and AI.



Tiegeng Liu is the director and the professor of the Key Laboratory of Optoelectronic Information Technology of the Ministry of Education at Tianjin University, China, the Deputy Director of the Optoelectronic Technology Professional Committee of the Chinese Optical Society, the Deputy Chair of the Optical Instrument Branch of the China Instrument Association. His research interests include fiber optic sensing technology, optoelectronic detection technology, and embedded image detection technology.