



## Point-to-multipoint coherent PON system modeling based on the diffusion model

XIAOKAI GUAN,<sup>†</sup>  YONGXIN XU,<sup>†</sup>  WENQING JIANG,  XUDONG WANG,  CHENYANG HAN,   
WEISHENG HU,  AND LILIN YI\* 

State Key Laboratory of Photonics and Communications, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>†</sup> Authors contributed equally to this work and should be considered co-first authors.

\*lilinyi@sjtu.edu.cn

Received 26 December 2024; revised 11 March 2025; accepted 15 April 2025; posted 15 April 2025; published 6 May 2025

**Modeling point-to-multipoint (P2MP) coherent passive optical network (PON) systems based on digital subcarrier multiplexing (DSCM) plays a crucial role in global system optimization and overall performance monitoring. The subcarriers could support any rate and modulation format. Besides, subcarriers reach different optical network units (ONUs) with different fiber link distances and received optical power. Therefore, it is necessary to construct a generalization model of the communication system to support arbitrary configuration of link parameters. In this paper, for what we believe to be the first time, we use the diffusion model in machine learning to establish a generalized model of the P2MP system. The SNR error of receiving signals between the modeling channel and the original channel is less than 0.1 dB under a multi-ONU channel with different configurations. It provides a new idea, to our knowledge, for P2MP coherent PON modeling.** © 2025 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.

<https://doi.org/10.1364/OL.553826>

In recent years, with the rapid expansion of user demand, new information services are developing rapidly, which puts forward higher requirements for the communication capacity bandwidth and transmission rate [1–4]. Therefore, coherent detection becomes a technique for passive optical network (PON) in the future due to the better performance and high transmission rate [5]. In a coherent PON system, the point-to-multipoint (P2MP) transmission mode has become an important future evolution direction because of its flexibility, scalability, and cost-efficiency. Digital subcarrier multiplexing (DSCM) technology is a feasible solution to realize flexible P2MP coherent PON.

However, in a traditional communication system, each module needs to be designed separately. It is a complicated process and does not guarantee that each discrete module is optimal for the whole system. In addition, in the P2MP structure, the digital subcarrier signals generated at the optical line terminal (OLT) are given to multiple optical network units (ONUs). There is still room for optimization of the modulation format, power, and constellation points of each subcarrier. Therefore, joint optimization is required in order to achieve system global optimization [6,7].

End-to-end (E2E) learning is a solution to achieve optimal system performance by exploiting the ability of deep learning to approximate any nonlinear function and jointly optimizing the transceiver and the fiber optic channel through artificial neural network (ANN) [8–11]. E2E learning usually uses an auto-encoder (AE) structure. Generally, it uses two neural networks as transmitter and receiver instead of separate signal processing modules [12–14]. The neural networks are trained on a specifically designed differentiable channel model that simulates a real physical channel. The effectiveness of end-to-end learning depends heavily on the accuracy of the differentiable channel model [15,16]. Besides, channel modeling also plays a significant role in system performance monitoring. To build a digital twin system of optical network, it requires the accurate digital channel to simulate the physical system.

Lots of researches on channel modeling focus on machine learning methods. In order to achieve better performance, several different machine learning models used for channel modeling are designed, such as generative adversarial network (GAN) [17], transformer model [18], bidirectional long short-term memory (Bi-LSTM) [19], and other researches [20–23]. While in the short-range channel modeling field, the scenarios are mostly about intensity modulation direct detection (IMDD) or single-carrier coherence. Researches on configuration generalization in P2MP PON remain undone. Besides, the machine learning model used for modeling such as GAN leads to the problem of difficult training. Therefore, it requires a suitable model applied for P2MP modeling.

In order to model P2MP structure based on DSCM, the denoising diffusion probabilistic model (DDPM) [24] is considered in our research. DDPM has attracted much attention in the field of deep learning in recent years. The stable diffusion model [25], which based on DDPM with better performance, has demonstrated strong learning and generalization capabilities in the generative domain. There have been works and researches in the field of communications that have applied diffusion models for channel modeling in the wireless domain with success [26–28]. In this paper, the diffusion model is applied to optical communication channel modeling for the first time. Then the ability of channel modeling and generalization is verified through comparison of SNR and constellation.

The diffusion model constructs a Markov chain of discrete steps for the input raw data, including a forward diffusion process and an inverse denoising process. The forward diffusion process, which is also the training step, involves adding random Gaussian noise until it becomes unrecognizable as pure noise (obeying Gaussian distribution), and the NN model is trained to predict noise during this process. The backward denoising process, as known as the sampling step, predicts the Gaussian noise at each time step. The noise is gradually removed from the noisy data to restore the original data distribution. According to mathematical derivation [24], predicting Gaussian noise is equal to predicting the mean value of it, which can be expressed as follows:

$$\mu_{\theta}(x_t, t) = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(x_t, t) \right), \quad (1)$$

where  $t$  represents time step,  $\alpha_t$  and  $\bar{\alpha}_t$  are hyper-parameters,  $x_t$  represents noisy data at time step  $t$ , and  $\epsilon_{\theta}(x_t, t)$  represents the neural network model where  $\theta$  represents the parameters of the neural network. The loss function  $L_t$  is defined at time step  $t$  as follows:

$$L_t = \mathbb{E}_{\epsilon_t \sim \text{Unif}[7]} \left[ \left\| \epsilon_t - \epsilon_{\theta} \left( \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon_t, t \right) \right\|^2 \right], \quad (2)$$

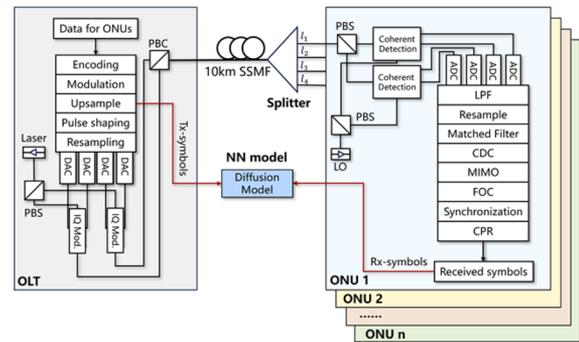
with the aim of minimizing the difference between  $\mu_{\theta}(x_t, t)$  and  $\tilde{\mu}(x_t, x_0, t)$ , where  $\mu_{\theta}(x_t, t)$  means the mean value of Gaussian noise predicted by the trained model and  $\tilde{\mu}(x_t, x_0, t)$  means the label mean value of Gaussian noise added to the signal during the diffusion process and  $x_0$  represents original input data. At each moment  $t$ , the DDPM model returns the predicted noise component  $\epsilon_{\theta}(x_t, t)$  with  $x_t$  as input, where  $\epsilon_t$  denotes the true diffuse noise term at moment  $t$ .

For channel modeling, taking the channel output as  $x_0$  and the channel input  $y_i$  as the condition, the ultimate goal of the diffusion model is to learn the conditional distribution  $p(x_0|y_i)$ . The input  $y_i$  is added to the denoising process, and the rest of the derivation and conclusions remain unchanged; thus (1) becomes the form of (3) as follows:

$$\mu(x_t, y) = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(x_t, y, t) \right). \quad (3)$$

When generalizing the model, such as modeling different ONU channels in P2MP, it is necessary to additionally introduce  $C$  as a channel parameter (e.g., ONU id, fiber link length, launch power, etc.). Although there is a small difference in the structure of the model,  $C$  and  $y_i$  are mathematically equivalent. Therefore, the channel inputs can be written as the condition  $y_i$  along with the channel parameter.

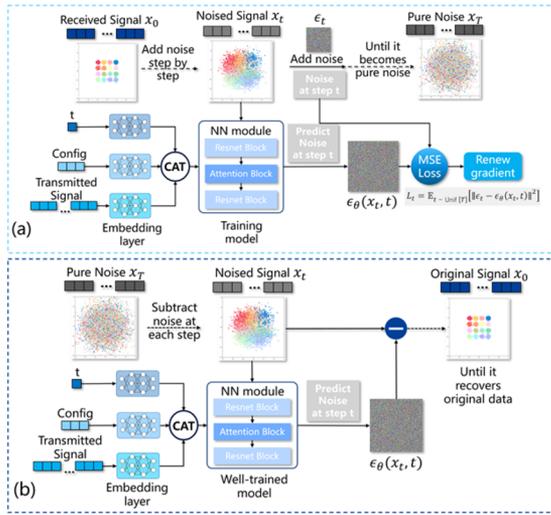
According to the principle of the diffusion model, a method using the diffusion model for channel modeling is proposed. The goal of channel modeling is to learn the channel transfer function  $x = H(y)$  for the specific channel parameter  $C$ , which can be expressed as the conditional distribution  $q(x|y, C)$ . For channel modeling of P2MP coherent PON, the goal is to generalize different ONUs with the same model network. Different ONU links have their own characteristics and correspond to different places on the spectrum. The model network is able to generalize based on the input channel parameters and achieve accurate modeling for each ONU channel. In order to learn the characteristics and impairments of the complete channel, the diffusion model is applied to implement the symbol-to-symbol modeling approach in this work.



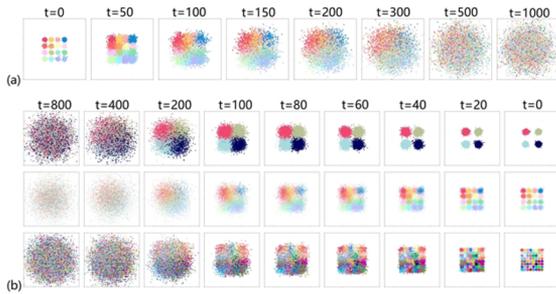
**Fig. 1.** Schematic of the DSCM-based P2MP downstream architecture.

Training the deep learning model requires generating a dataset. The structure of the P2MP based on DSCM with a dual-polarization simulation system is shown in Fig. 1. The transceiver is set to be nonideal to simulate real situations, including the limit bandwidth, quantification noise, and phase noise. The signal is affected by attenuation, chromatic dispersion (CD), polarization mode dispersion (PMD), nonlinear impairment, and other impairments in the fiber. The fiber nonlinearity is simulated using the Manakov channel [29]. At the OLT, a large-bandwidth single-carrier signal is divided into a number of small-bandwidth independent subcarrier signals for transmission using frequency division multiplexing. Optical signals loaded with multiple ONU signals are sent out from the OLT and arrive at the optical distribution network (ODN) through the fiber of the main link. Then the ODN, which refers to splitter in Fig. 1, distributes the signals to different ONU sub-links. At each ONU, the local oscillator (LO) is adjusted to restore the corresponding ONU signals to the baseband. The optical signals are converted to electrical signals by the coherent receiver. The electrical signals of the ONUs are extracted by the low pass filter (LPF), and then the subsequent digital signal process (DSP) operations are performed. The simulation system corresponds with experiment system based on BER, SNR, and constellation. The modeling dataset consists of the final received symbols after DSP and the transmitted symbols after bit coding. In a real system, after the optical link is built and settled, the training data will be sent from OLT to ONU, and the received signal will be collected to train the model.

The specific methods of training and sampling are shown in Fig. 2. Rx-signal, used as  $x_0$  in the equation, is the received symbols after DSP. It gradually adds Gaussian white noise, and the noise added at each time step  $t$  is recorded as the label of the training structure. Then  $x_t$  is directly acquired according to equation [24], which indicates that the data becomes pure noise after the  $T$ th step of adding noise. The transmitted signal Tx-signal used as a control condition corresponds to the receiving signal, and the actual length can be changed as needed. The three control conditions, the transmitted signal Tx-signal, the channel parameter config, and the corresponding time step  $t$ , are spliced together after their respective embedding layers and fed into the NN layer for training. The output of the NN layer is the predicted noise at step  $t$ , and then the predicted noise and the true noise are used to calculate mean square error (MSE) loss in (2), the gradient of which is updated iteratively. The NN layer learns the noise that should be added at step  $t$  during the iteration process. In the sampling step, a pure Gaussian noise is first obtained by randomly sampling from a Gaussian distribution,



**Fig. 2.** Model structure for channel modeling using a diffusion model. (a) Training process. (b) Sampling process.

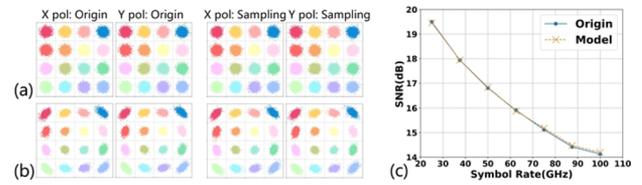


**Fig. 3.** (a) Diffusion process for 16QAM signal. (b) Denoise process for QPSK, 16QAM, and 64QAM signal.

which represents the data after the addition of noise at step  $T$ . The Gaussian noise is input into the diffusion model together with the control conditions, and the model outputs the predicted noise added at the current time step. Subtracting the predicted noise from noisy data yields the data from the previous step. Clean data can be eventually recovered after gradually repeating these operations.

The noise added to the data at each time step satisfies the standard Gaussian distribution  $N(0, 1)$ , which is represented as a signal of purely chaotic points on the constellation diagram. Figure 3(a) illustrates the process of noise addition for the diffusion model with 16 quadrature amplitude modulation (QAM). When the time step is 1000, the constellation diagram has become pure Gaussian white noise. Figure 3(b) shows the sampling process of the diffusion model. The model is able to learn the approximate distribution of original signal data and recover it from Gaussian white noise. Different modulation formats are also verified, such as quadrature phase shift keying (QPSK) and 64QAM, indicating that the model is able to recover signals no matter what the modulation format is.

Firstly, a single ONU is being researched. Figure 4 shows the modeling results. In Fig. 4(a), the original signal data contains only Gaussian white noise. The shape, distribution, and noise size of the model-generated signals and the original signal constellation diagrams are quite close to each other, and the human eye can no longer distinguish the differences. Next, the Kerr nonlinear effect and the phase noise of the transceiver laser are

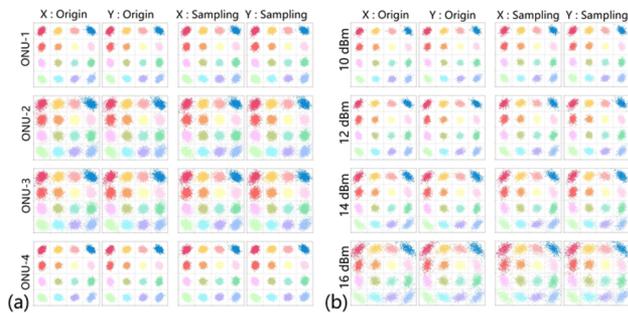


**Fig. 4.** Channel modeling results for diffusion model (a) without nonlinearity, (b) with nonlinearity, (c) curve of SNR, and symbol rate.

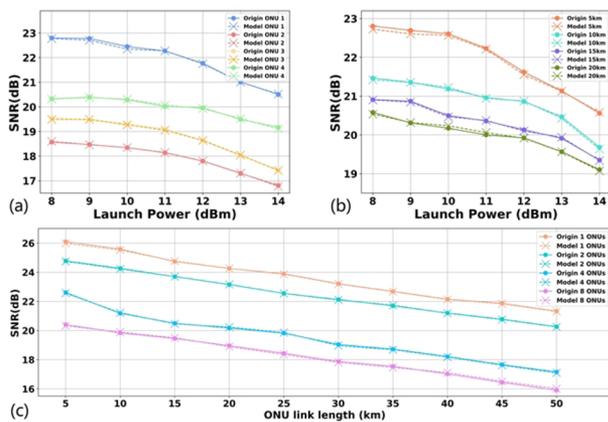
added to check whether the model can learn further nonlinear effects. In Fig. 4(b), the constellation shows the nonlinear distortion, and the model also learns the nonlinear effects. Due to the impact of the device bandwidth limitations and other effects, the received signal noise becomes larger as the transmitted signal rate gets higher, leading to a smaller SNR of the received signal. The curve of SNR and symbol rate between the original received signal and the model-generated received signal is shown in Fig. 4(c). It can be seen that the model models the channel quite accurately and errors are all within 0.1 dB, which is within the acceptable range.

In order to model the P2MP channel and examine the generalization ability of the diffusion model, training data is generated under different configurations. The target parameters include the ONU number, fiber link length, and launch power. The launch power is from 8 dBm to 14 dBm. There are groups of data for four ONUs, fiber length (km) of which is set randomly. It is worth noting that in the short-reach optical transmission, the nonlinear effect occurs when the launch power is greater than 10 dBm. When it is greater than 12 dBm, the distortion of constellation is obvious, and the SNR of the signal drops around 2 dB [30]. Figure 5(a) shows the constellation diagram of four ONUs under the same parameter configuration, with launch power of 12 dBm and fiber lengths (km) of {5,15,10,20} for four ONUs. From the constellation diagram, it can be observed that the noise of two ONUs in the middle is significantly larger than that of the two next to it. The reason is that the signal of an ONU located in the middle of the spectrum is subjected to more band interference from neighboring ONUs than those at the edge of the spectrum. The diffusion model learns the channel characteristics of different ONUs well, including noise distribution and corresponding nonlinear effects. Figure 5(b) shows the modeling results for the same fiber length at the same ONU location for different launch power. As the launch power increases, the nonlinear effect increases, and the distortion becomes more obvious. However, the diffusion model is still able to learn the nonlinearity and the noise.

By changing the launch power while maintaining the fiber length of the ONU link, the curves of model under different ONUs are obtained and compared with that of the original channel, as shown in Fig. 6(a). It can be seen that the modeling results fit the original channel well. In Figs. 6(b) and 6(c), the fiber length (km) of each ONU is the same, and only the result of the first ONU (ONU-1) is shown. Figure 6(b) explores channel modeling ability for four ONUs under different launch power, and Fig. 6(c) investigates the influence of the number of ONUs under different link length with launch power of 10 dBm. Changing the different values of fiber length, launch power, and the number of ONUs, the diffusion model still owns the ability of generalization, and modeling results fit the original one well with average error less than 0.1 dB.



**Fig. 5.** Channel modeling results. (a) Four ONUs under the same configuration. (b) ONU-1 under different launch power.



**Fig. 6.** Channel modeling results. (a) Four different ONUs under different launch power. (b) ONU-1 with different fiber length under different launch power. (c) ONU-1 with different numbers of ONUs under different link length.

In this paper, the diffusion model is successfully applied to optical communication channel modeling, and the effectiveness and practicability of the model are verified. This work explores the accuracy and generalization ability of the model used for the channel modeling of the P2MP coherent PON system. The generalization parameter includes ONU ID (stands for different spectrum positions), fiber link length, and optical launch power. The modeling error of the SNR is less than 0.1 dB at different launch power. Though simulation system matches the experiment well, there are still discrepancies such as error between theory formula and reality, unpredictable random noise in fiber and devices, or ignored influence factors. In the next step, we consider collecting data on a high-speed experimental platform to train the model. The diffusion model can also be compared with different high-performance models. In the future, we aim to apply the model to build a digital twin system in order to realize global system optimization and overall performance monitoring.

**Funding.** National Key Research and Development Program of China (2023YFB2905400); National Natural Science Foundation of China (62025503); Shanghai Jiao Tong University 2030 Initiative.

**Disclosures.** The authors declare no conflicts of interest.

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

## REFERENCES

1. M. Chagnon, *J. Lightwave Technol.* **37**, 1779 (2019).
2. J. S. Wey and J. Zhang, *J. Lightwave Technol.* **37**, 2830 (2019).
3. V. Houtsma, A. Mahadevan, N. Kaneda, *et al.*, *IEEE/OSA J. Opt. Commun. Networking* **13**, A44 (2021).
4. D. van Veen and V. Houtsma, *IEEE/OSA J. Opt. Commun. Networking* **12**, A95 (2020).
5. M. S. Faruk, X. Li, D. Nasset, *et al.*, *IEEE Commun. Mag.* **59**, 112 (2021).
6. D. Welch, A. Napoli, J. Back, *et al.*, *J. Lightwave Technol.* **39**, 5232 (2021).
7. D. Welch, A. Napoli, J. Bäck, *et al.*, in *Optical Fiber Communications Conference (OFC)* (2020), paper Tu3H.1.
8. B. Karanov, M. Chagnon, F. Thouin, *et al.*, *J. Lightwave Technol.* **36**, 4843 (2018).
9. B. Karanov, V. Oliari, M. Chagnon, *et al.*, in *European Conference on Optical Communication (ECOC)* (2020), pp. 1–4.
10. T. O'Shea and J. Hoydis, *IEEE Trans. Cognit. Commun. Networking* **3**, 563 (2017).
11. B. Karanov, M. Chagnon, V. Aref, *et al.*, in *Optical Fiber Communications Conference (OFC)* (2020), paper Th2A.48.
12. S. Dorner, S. Cammerer, J. Hoydis, *et al.*, *IEEE J. Sel. Top. Signal Process.* **12**, 132 (2018).
13. V. Neskorniuik, A. Carnio, V. Bajaj, *et al.*, in *European Conference on Optical Communication (ECOC)* (2021), pp. 1–4.
14. V. Aref and M. Chagnon, in *Optical Fiber Communications Conference (OFC)* (2022), paper W41.3.
15. J. Song, C. Häger, J. Schröder, *et al.*, *IEEE J. Sel. Top. Quantum Electron.* **28**, 1 (2021).
16. Z. Niu, H. Yang, H. Zhao, *et al.*, *J. Lightwave Technol.* **40**, 2807 (2022).
17. H. Yang, Z. Niu, S. Xiao, *et al.*, *J. Lightwave Technol.* **39**, 1322 (2021).
18. N. Zhang, H. Yang, Z. Niu, *et al.*, *J. Lightwave Technol.* **40**, 7779 (2022).
19. D. Wang, Y. Song, J. Li, *et al.*, *J. Lightwave Technol.* **38**, 4730 (2020).
20. M. Li, D. Wang, Q. Cui, *et al.*, in *Opto-Electronics and Communications Conference (OECC)* (2020), pp. 1–3.
21. Q. Zhang, Z. Wang, S. Duan, *et al.*, *Micromachines* **13**, 31 (2022).
22. E. Ip and J. M. Kahn, *J. Lightwave Technol.* **26**, 3416 (2008).
23. M. Li and S. Wang, *IEEE Commun. Lett.* **26**, 1829 (2022).
24. J. Ho, A. Jain, and P. Abbeel, in *Advances in Neural Information Processing Systems (NeurIPS)* (2020), pp. 6840.
25. R. Rombach, A. Blattmann, D. Lorenz, *et al.*, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021), p. 10674.
26. M. Kim, R. Fritschek, and R. F. Schaefer, in *26th International ITG Workshop on Smart Antennas and 13th Conference on Systems, Communications, and Coding (WSA & SCC)* (2023), pp. 1–6.
27. T. Wu, Z. Chen, D. He, *et al.*, in *IEEE Global Communications Conference (GLOBECOM)* (2023), p. 7429.
28. M. Arvinte and J. I. Tamir, in *IEEE Wireless Communications and Networking Conference (WCNC)* (2021), p. 453.
29. G. P. Agrawal, *Nonlinear Fiber Optics* (Springer, 1995).
30. P. Torres-Ferrera, M. S. Faruk, I. B. Kovacs, *et al.*, in *Optical Fiber Communications Conference (OFC)* (2023), paper Th2A.21.