Research Article

Digital twin modeling and controlling of optical power evolution enabling autonomous-driving optical networks: a Bayesian approach

Xiaomin Liu, Yihao Zhang, Yuli Chen, Yichen Liu, Meng Cai, Qizhi Qiu, Mengfan Fu, Lilin Yi, Weisheng Hu, and Qunbi Zhuge ⓘ*

Shanghai Jiao Tong University, State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai, China

Abstract Optical networks are evolving toward ultrawide bandwidth and autonomous operation. In this scenario, it is crucial to accurately model and control optical power evolutions (OPEs) through optical amplifiers (OAs), as they directly affect the signal-to-noise ratio and fiber nonlinearities. However, a fundamental contradiction arises between the complex physical phenomena in optical transmission and the required precision in network control. Traditional theoretical methods underperform due to ideal assumptions, while data-driven approaches entail exorbitant costs associated with acquiring massive amounts of data to achieve the desired level of accuracy. In this work, we propose a Bayesian inference framework (BIF) to construct the digital twin of OAs and control OPE in a data-efficient manner. Only the informative data are collected to balance the exploration and exploitation of the data space, thus enabling efficient autonomous-driving optical networks (ADONs). Simulations and experiments demonstrate that the BIF can reduce the data size for modeling erbium-doped fiber amplifiers by 80% and Raman amplifiers by 60%. Within 30 iterations, the optimal controlling performance can be achieved to realize target signal/gain profiles in links with different types of OAs. The results show that the BIF paves the way to accurately model and control OPE for future ADONs.

1 Introduction

Optical fiber has been widely utilized in the fields of communication[1] and sensing,[2–4] playing an critical role in daily-life communications, the military, scientific research, and other fields. Specifically, in the field of communication, fiber-optic communication carries a majority of global data traffic due to its low attenuation and large capacity.[5] In recent years, due to the information explosion caused by high-definition video streaming, cloud computing, artificial intelligence, and so forth, global communication traffic has grown exponentially.[6,7] To sustain such dramatic traffic growth, optical transmission technologies such as coherent transceivers with multilevel modulation formats[8] and advanced digital signal processing[9]

have been developed, pushing the fiber capacity to approach the Shannon limit[10] for a given spectrum bandwidth.

Today, to further increase the fiber capacity, multiband[11] and spatial-division multiplexing[12] systems have been introduced to exploit more spectrum resources. The multiband solution is more appealing, as it upgrades existing fiber infrastructures in a cost-effective manner. Recently, commercial systems have been extended from C-band to C+L-band, scaling the available bandwidth from ∼4 to ∼12 THz.[13] On the other hand, autonomous-driving optical networks (ADONs)[14,15] are being investigated and developed to improve network performance and lower operational expenditures (OpEx). By leveraging extensive physical layer data, ADON is expected to achieve automated service provisioning, power optimization, and failure management, eliminating the need for human intervention.

In a fiber-optic communication system, the optical power of signals evolves over fiber and also varies at different wavelengths,

*Address all correspondence to Qunbi Zhuge, qunbi.zhuge@sjtu.edu.cn

exhibiting a complex two-dimensional process. Modeling and controlling the optical power evolution (OPE) are crucial for enabling multiband ADONs.[16] First, optical power determines the optical signal-to-noise ratio (OSNR) and fiber nonlinear effects,[17] both of which significantly affect the signal transmission quality and achievable information rate. This is especially crucial for multiband systems, since wider bandwidth usually involves more severe Kerr nonlinearity[18] and inter-channel stimulated Raman scattering (ISRS).[19] Furthermore, in future ADONs, the static fiber channel should be upgraded into a programmable paradigm, which allows for dynamic configuration of the optical powers for signals with different transmission requirements, enabling more flexible and efficient utilization of network resources.

In multiband ADONs, OPE is mainly influenced by fiber propagation and amplification process. Since the fiber propagation process, including attenuation and ISRS, can be accurately calculated,[20–22] the main challenge in modeling and controlling OPE lies in optical amplifiers (OAs). However, modeling and controlling an erbium-doped fiber amplifier (EDFA)[23] in current C-band systems are already difficult tasks[24–27] due to the complex wavelength-dependent gain characteristics of EDFAs in dynamic link conditions.[28,29] In multiband systems, the complexity further increases due to the adoption of multiple homogeneous and/or heterogeneous OAs,[30] which are used to provide broader gain bandwidth and mitigate severe ISRS. In this case, different types of OAs amplify signals through various nonlinear effects, such as the stimulated emission or SRS. These nonlinear effects can be expressed by a set of ordinary differential equations (ODEs) with no closed-form solutions. Consequently, modeling and controlling OPE with diverse OAs pose considerable challenges.

In the past decades, extensive research has been carried out to achieve precise modeling and controlling of OPE through OAs in fiber-optic communication systems. The first direction is to rely on human intelligence.[31,32] However, such approaches are limited in accuracy, as they fail to consider the distinctive characteristics of each OA caused by manufacturing discrepancies and operating conditions.[33,34] For instance, with the center of mass model,[35] the gain profile of an EDFA can be modeled with merely one measurement of a baseline spectrum. However, this approach only achieves a root-mean-squared error (RMSE) of about 0.4 dB.[26] Such a level of accuracy is insufficient for assisting the ADON.

Recently, the concept of the digital twin (DT) has emerged to mirror the real-time status of each optical device based on collected data. This is especially crucial for OA modeling due to the diverse designs, manufactory discrepancies, uncertain parameters, and device aging impacts in OAs. In this context, data-driven techniques, especially neural networks (NNs), have attracted increasing attention in recent years.[24,26,27,36–38] However, a bottleneck in implementing data-driven models is the requirement for large training data sets to achieve high levels of accuracy. For example, with about 12,000 pieces of measured data, the RMSE of the gain model for an EDFA can be reduced to 0.1 dB.[26] In some cases, the training data set includes 40,000 data samples.[24] In a real system, the procedure for collecting such a large amount of data can be costly and time-consuming. Even though some methods utilize transfer learning to reduce the needed data size,[36] the data size of the pretraining is still high, which is hard to be achieved in a real system with various types of OAs from multiple vendors. Therefore, a reliable

method that can assist the accurate modeling and controlling for OA with data as little as possible is desired for future ADONs.

In this paper, we propose a Bayesian inference framework (BIF) to construct the DT of OAs and control OPE in a reliable and data-efficient manner. In this framework, a few initially collected data are used to train a Gaussian process regression (GPR)[39] surrogate model, which can provide the mean and variance of the estimation. By designing the acquisition functions based on the surrogate model, only the most informative data is sequentially collected. This approach can help obtain accurate digital models for modeling OAs and finding the optimal system configurations for controlling OA systems. Compared with the traditional methods for modeling and controlling OAs, the BIF can achieve higher accuracy and significantly reduce the amount of needed data. In this paper, the performance of the BIF is evaluated in both the EDFA system and RA system through simulations or experiments. In terms of modeling, the BIF can reduce more than 80% and 60% data to model the EDFA and RA, respectively. For the online controlling, the target gain/signal power spectra can be realized within 30 iterations. The optimal performance can be achieved with an RMSE of less than 0.5 dB in most cases. In the next section, the architecture of the BIF is introduced. The real-time experiment and simulation investigations for modeling and controlling are demonstrated.

2 Architecture of BIF for OA Modeling and Controlling

When modeling and controlling OAs, the aim of BIF is to optimize the sampling or operating strategies to achieve the best modeling or controlling performance. Based on prior observations, the BIF sequentially selects the next to-be-measured signal spectra or amplifier configurations. As shown in Fig. 1(a), first, a training data set containing fewer than five sets of pre-collected spectra with the corresponding initial amplifier configurations is constructed. Afterwards, a surrogate model is trained based on the GPR to quantify the optimization performance. According to the output of the GPR, the next to-be-sampled data are selected by an acquisition function and then measured automatically. After adding the newly measured data to the training data set, the surrogate GPR model can be updated to decide the next data for sampling. If the BIF is used for modeling, an accurate model for the OA can be obtained iteratively. If the BIF is used for online controlling, the optimal system configuration to achieve the target signal/gain spectra can be realized.

When constructing the GPR surrogate model, the training dataset can be written as $D = [\boldsymbol{x}_i, \boldsymbol{y}_i]$, where $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ represent the i'th input and output of the GPR model, respectively. The mapping between the input and output, denoted as $f(\boldsymbol{x})$, is described by the Gaussian process (GP) of

$$f(\boldsymbol{x}) \sim \mathrm{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')), \tag{1}$$

where $m(\boldsymbol{x})$ is the mean function of the GP and $k(\boldsymbol{x}, \boldsymbol{x}')$ is the covariance function, which is the 'kernel' for evaluating the similarity of each sample. $m(\boldsymbol{x})$ and $k(\boldsymbol{x}, \boldsymbol{x}')$ can be learned from data during training. For a new input, which is written as x_*, the estimation of the output, denoted as $f(x_*)$, follows the joint Gaussian process of
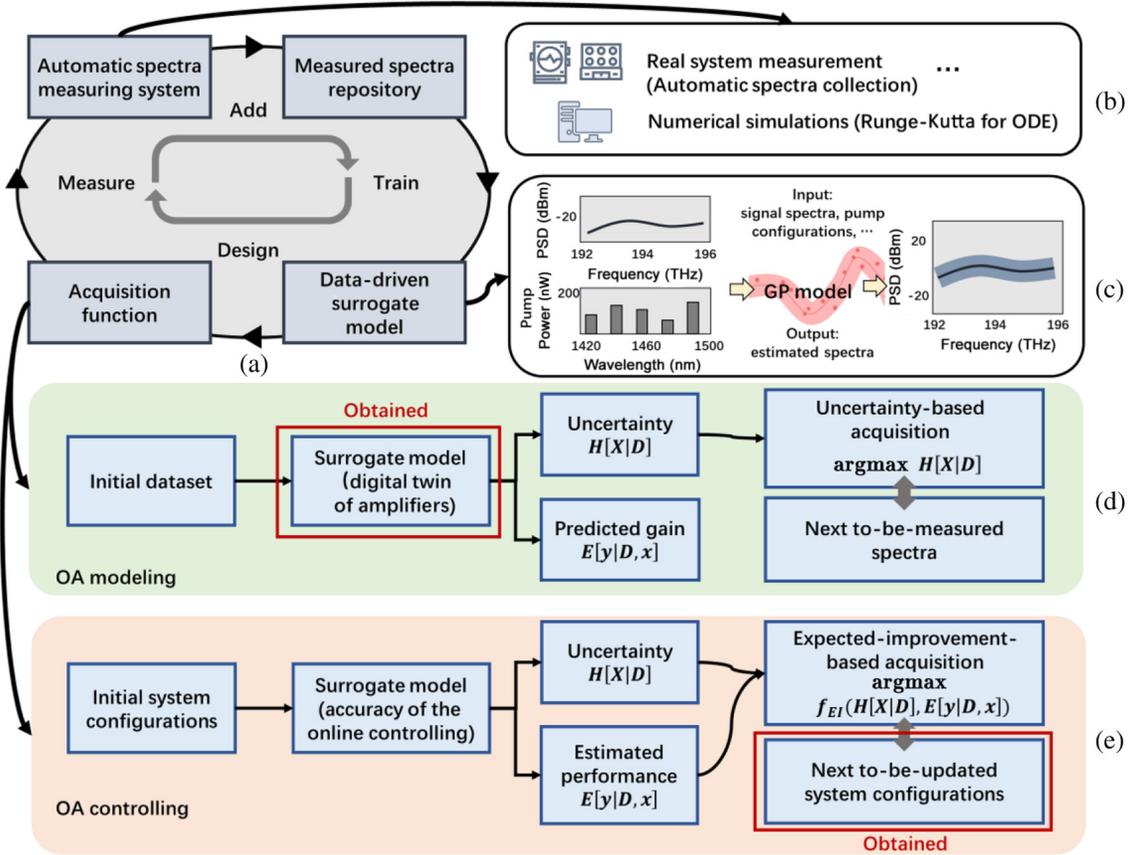
Fig. 1 The architecture of the proposed BIF for modeling and controlling OAs. (a) The general workflow of the BIF. (b) The employed data collecting method in real systems and simulations. (c) The input and output of the surrogate GPR model. (d) The exploration-preferred acquisition workflow of the OA modeling. (e) The exploitation-preferred acquisition workflow of the OA controlling.

$$\begin{bmatrix} \boldsymbol{Y} \\ f(x_*) \end{bmatrix} \sim N\left(\begin{bmatrix} m(\boldsymbol{x}) \\ m(x_*) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K} + \beta I & \boldsymbol{k} \\ \boldsymbol{k}^T & k(x_*, x_*) \end{bmatrix}\right), \qquad (2)$$

where $\boldsymbol{K}$ is the covariance matrix of the training dataset. $\boldsymbol{k}$ and $\boldsymbol{k}^T$ are the covariance between the training set $\boldsymbol{x}$ and x_*. β is the hyperparameter representing the noise of the measurement. Therefore, the estimated mean μ_* and variance σ_*^2 of x_* are

$$\mu_* = m(x_*) + \boldsymbol{k}^T \left(\boldsymbol{K} + \beta \boldsymbol{I}\right)^{-1} (\boldsymbol{Y} - m(\boldsymbol{x})), \qquad (3)$$

$$\sigma_*^2 = k(x_*, x_*) - \boldsymbol{k}^T \left(\boldsymbol{K} + \beta \boldsymbol{I}\right)^{-1} \boldsymbol{k}. \qquad (4)$$

Based on the GPR surrogate model, the next to-be-sampled data are decided by an acquisition function. Since the optimization target is different among tasks, the sampling strategies are different, resulting in the customized design of the acquisition functions.

When modeling OAs, the surrogate GPR model is constructed as a DT of an OA. The input of the GPR-based DT is the input signal spectra and amplifier configuration parameters. The output is the gain spectra of the OA under test. To achieve data-efficient modeling, as shown in Fig. 1(d), the BIF focuses

on the exploration (sampling the places with a high σ_*^2) of the whole feature space. Therefore, the acquisition target is to sample the most informative data, which can be categorized as a problem of Bayesian active learning.[40] In our work, the acquisition function is designed by uncertainty sampling, which means sampling where the uncertainty is high. The sampling strategy can be written as

$$\arg \max_{\boldsymbol{x}} \sigma^2(\boldsymbol{x}), \qquad (5)$$

where $\sigma(\boldsymbol{x})$ represents the estimation variance of the sample $\boldsymbol{x}$. In this way, the most uncertain candidate spectra are selected for the next round of measurement. After several iterations, the surrogate model can converge to a high accuracy.

For controlling OAs, the aim of the BIF is to exploit limited reconfigurations (input spectrum, operating conditions, etc.) to optimize the output signal/gain spectrum, which can be categorized as a problem of Bayesian optimization.[41–44] In this case, the surrogate model is the objective function, which quantifies the controlling performance. The input of the surrogate model is the OA configuration. The output is the estimated controlling performance such as the error between the current system value and the target. The commonly used acquisition functions are

expected improvement (EI),[41,45] probability of improvement (PI),[41] and upper confidence bound.[46] In our study, we find that these acquisition functions achieve a similar performance, and we choose the EI for the following evaluations. The acquisition formulation of the EI can be written as

$$\text{EI}(\boldsymbol{x}) = (\mu(\boldsymbol{x}) - f(\boldsymbol{x}_t^+) - \xi)\Phi(Z) + \sigma(\boldsymbol{x})\phi(Z), \qquad (6)$$

where $Z = \frac{\mu(\boldsymbol{x}) - f(\boldsymbol{x}_t^+) - \xi}{\sigma(\boldsymbol{x})}$. $\mu(\boldsymbol{x})$ and $\sigma(\boldsymbol{x})$ represent the current estimated mean and variance value of the sample $\boldsymbol{x}$, respectively. $f(\boldsymbol{x}_t^+)$ is the current maximum with the optimal sample $\boldsymbol{x}_t^+$. Therefore, $\mu(\boldsymbol{x}) - f(\boldsymbol{x}_t^+)$ represents the expected improvement of $\boldsymbol{x}$ compared with $\boldsymbol{x}_t^+$. $\Phi(Z)$ is the cumulative distribution function, which represents the probability distributions of the improvement. $\phi(Z)$ is the probability distribution function following the standard normal distribution. ξ is a hyperparameter controlling the balance between searching the global optimal and exploring the whole data space. Based on EI, the acquisition strategies can be represented as

$$\underset{\boldsymbol{x}}{\arg\max}\,\text{EI}(\boldsymbol{x}). \qquad (7)$$

After iterative searches, the most suitable system configuration, i.e., $\boldsymbol{x}$, is obtained.

3 Results

3.1 Constructing DT of EDFAs Based on BIF

To further improve the performance of OA modeling, other methods have been proposed previously with some data sets.[47,48] For example, the transfer-learning-based modeling scheme[36] proposes to initially train a basic model, which is then transferred to the specific EDFA. In addition, the hybrid modeling method[49] integrates the estimations of the analytical models to the input for higher precision. These methods have successfully improved the modeling performance by designing the training scheme and the input features. However, the design of the data selection scheme is not thoroughly investigated. For modeling OAs, we aim to build an accurate DT with data as little as possible. In this case, the BIF pays more attention to the exploration, and the uncertainty-based acquisition function is utilized. To evaluate the performance of the BIF, an experimental validation for modeling EDFA is conducted.

First, we show the performance of modeling a commercial EDFA based on the BIF in an experiment. As shown in Fig. 2(a), an automatic EDFA measuring system is built. An ASE noise source is used for simulating the flat full C-band spectrum. After filtered by a Finisar WaveShaper 4000A optical spectrum processor, 80 channels with 50 GHz spacing in the C-band from 192.1 to 196.1 THz are generated. Among them, 40 odd-numbered channels are selected to establish signals while other channels are filtered out. Two optical spectrum analyzers (OSAs) are used to measure the input and output power spectra of the EDFA. The operating mode of the EDFA under test is set as the automatic gain control (AGC) mode with a gain of 16 dB. When generating the repository of the gain spectra, the 40 odd-numbered channels are assumed to be occupied or idle randomly. A random deviation of each signal power from -2 dB to 2 dB with a step size of 1 dB is involved by setting the attenuation of the WaveShaper. In total, 9578 to-be-measured input spectra are generated as the candidate repository. Additionally, to evaluate the forward modeling performance, a testing data set containing 1002 pairs of input and output spectra is generated randomly by the automatic measuring system.

When building the DT for the EDFA, the input is a vector representing the power spectrum of the signals before amplification. The output is the corresponding gain spectrum. In this experiment, we compare the modeling performance with the traditional models based on NNs that are trained with data sampled randomly as the baseline. To investigate the performance, the RMSEs of each model on the testing data set are calculated. Considering the impact of various initial random data sets,
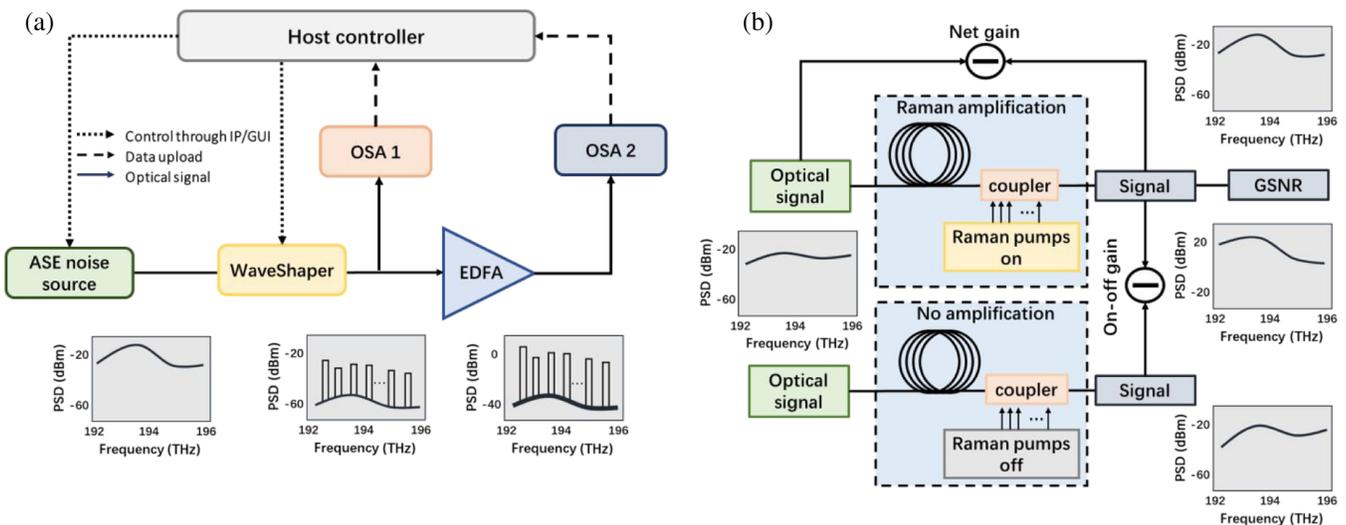
Fig. 2 The diagram of collecting training data for OA modeling. (a) The experimental system for measuring EDFA spectra. (b) The simulation workflow for calculating the GSNR of an RA system.
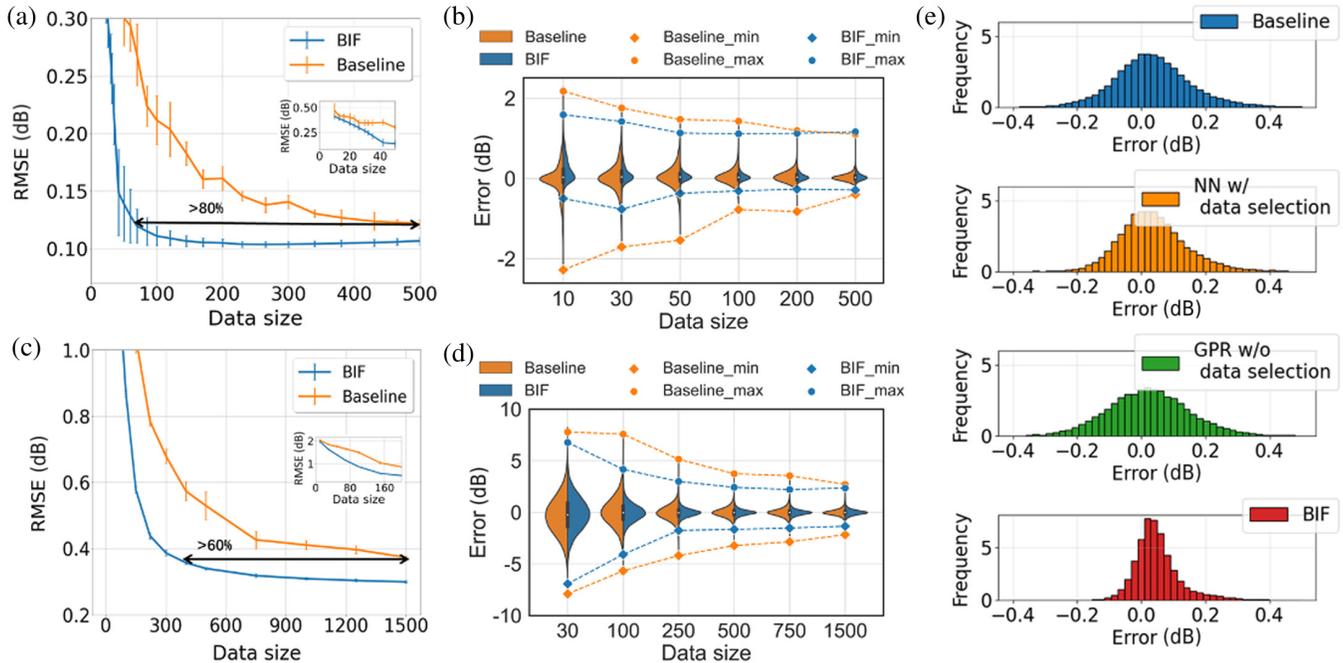
Fig. 3 The modeling performance of the proposed BIF for EDFA and RA. (a) and (c) The RMSEs of the models under different data sizes for EDFA modeling and RA modeling, respectively. (b) and (d) The violin plots of the modeling errors under different data sizes for EDFA modeling and RA modeling, respectively. The maximum and minimum errors are plotted. (e) The error histograms of the EDFA gain models trained with 70 data with or without BIF-based modeling design and data selection.

each training process undergoes five iterations to mitigate performance fluctuations. The mean RMSEs are then plotted in Fig. 3(a), while the error bars represent the performance fluctuations across different training iterations. For the NN-based models with random data selection, the estimation errors are large when the data size is small. The best accuracy achieved with 500 training instances is about 0.12 dB. In contrast, the proposed BIF-based model converges at a high speed and can reduce the RMSE to about 0.1 dB with fewer than 200 instances, demonstrating its significant learning ability. To achieve the same RMSE on the same testing data set, the proposed method can largely reduce the training data size by 80%, making it possible to prepare a customized tiny data set for building the precise gain model for each EDFA. Moreover, the performance of the proposed model is relatively stable because the error bar is smaller when the data size is large. The violin plot of the errors of the models trained by different methods with different amounts of data is plotted in Fig. 3(b). The maximum and minimum errors are plotted. The results show that the model trained by the proposed BIF has lower estimation errors and converges faster, demonstrating its strong ability of selecting data and learning.

As mentioned before, to achieve the data-efficient forward modeling, the BIF employs the data-efficient GPR modeling method and data selection strategy. Here we further investigate the contributions of the GPR algorithm and the data selection strategy individually. In Fig. 3(e), we plot the RMSEs of the GPR-based and NN-based models using the data selected by the BIF or randomly. Four cases are considered: (1) the baseline

NN, (2) the NN with the BIF-based data selection, (3) the GPR model without the BIF-based data selection, and (4) the proposed BIF-based model. The error histograms of models trained with different methods on the testing data set are shown in Fig. 4. Results show that compared with NN, the GPR-based model shows a better learning ability. By using the training data selected by the BIF, both the NN and GPR can have a higher accuracy compared with the model learned from the randomly selected data.

Considering the training time, we train these models with 48 GB of 2400 MHz RAM and an Intel Core i9-9900k 3.6 GHz CPU. For the NN-based and GPR-based EDFA models, the needed training time with 500 training data is 28 and 15 s, respectively. The training time of the NN is primarily based on the configuration of the training procedure, including factors such as the batch size, the number of epochs, and the patience threshold. The training time of the GPR model mainly depends on the size of the covariance matrix, of which the complexity is $O(N^3)$. If the training data size is large, the calculation time of the GPR can be long. Nevertheless, since the BIF significantly reduces the needed data size, the training time of the GPR in our experiment can be effectively managed within a reasonable range.

3.2 Constructing DT of RAs Based on BIF

Besides EDFA modeling, the performance of the BIF for RA modeling is also investigated through simulations. We consider a more complex situation by modeling the generalized signal-to-

noise ratio (GSNR) of arbitrary signals under a certain pump configuration of an RA. The GSNR can be expressed as

$$\text{GSNR} = \frac{P_{\text{S}}^{\text{Rx}}}{P_{\text{ASE}}^{\text{Rx}} + P_{\text{NLI}}^{\text{Rx}}}, \qquad (8)$$

where P_{S}^{Rx}, $P_{\text{ASE}}^{\text{Rx}}$, and $P_{\text{NLI}}^{\text{Rx}}$ denote the power of signal, ASE noise, and NLI noise before the receiver, respectively. In such a situation, both the ASE noise and fiber nonlinearity are modeled, which is more challenging, since they are strongly related to both the signal and pump configurations. Similar to the EDFA modeling, we apply the exploration-preferred BIF using synthetic data for evaluation.

The data set for training is generated by simulations based on the GNPy,[50] which is a commonly-utilized Python tool for calculating the fiber nonlinearity based on Gaussian noise model.[22,51] The simulation setup is shown in Fig. 2(b). Arbitrary C+L-band signal spectra are generated by randomly selecting a flat launch power from −3 to 4 dBm, with a ripple of ±3 dB of each channel. For each signal, the baud rate is 142.8 GBaud and the channel spacing is 150 GHz. The transmitted fiber is the standard single-mode fiber (SSMF). The pump number in the RA is 6, of which the wavelengths are 1513, 1496, 1477, 1458, 1432, and 1420 nm, respectively. As with the EDFA modeling, the pump powers are fixed to set a relatively flat gain spectrum of 10 dB. The pump powers are 40, 30, 20, 120, 300, and 300 mW, respectively. In total, 3997 data are generated. We use 500 data as the testing data set and 3497 are used as the data repository for the BIF-based data selection.

Similar to EDFA modeling, the compared baseline model is based on NN and trained by a data set generated randomly. For both the NN-based model and the BIF-based model, the input features are the 80-dimensional vector representing the signal power of each channel. The outputs are the GSNR of each signal. Considering the influence caused by the randomness of the initial data set, the training process for each method is conducted 5 times to reduce the performance fluctuations. The mean RMSEs of the models obtained under different data sizes are plotted in Fig. 3(c). The differences among these training processes are shown as the error bar. The results show that the proposed method can achieve higher accuracy with different data sizes. To achieve a similar accuracy, the BIF can reduce more than 60% of the training data, demonstrating its efficient learning ability. Additionally, the error bar of the BIF-based model is much smaller, demonstrating its higher stability. The violin plot of the estimation error is shown in Fig. 3(d) with the maximum and minimum errors. The results show that the BIF-based model can converge faster with smaller extreme errors, proving its better learning ability.

3.3 Controlling EDFAs Based on BIF

For ADONs, efficiently shaping the signal/gain power spectrum is desired to assist dynamic network optimizations. To achieve this, BIF-based online controlling is proposed. For EDFA, the flat or tilted signal spectrum after amplification can be realized by adjusting the input signal power spectrum. For RA, the target gain spectrum can be realized by adjusting the pump powers. For both EDFA and RA systems, we perform experiments to demonstrate the effectiveness of the BIF for controlling the OPE.

First, Fig. 4(a) shows the experimental verification with a C-band EDFA. We employ an experimental setup similar to the one used in the previous section for modeling EDFA. The EDFA is configured in the AGC mode with a gain of 17 dB. During online controlling, we adjust the signal spectrum prior to amplification to obtain the target signal spectrum after amplification. Specifically, we adjust the attenuation factors of the WaveShaper as the control parameters, with one attenuation factor for every five consecutive WDM signals. So, in total, there are eight parameters to control. This type of adjustment can be realized in real systems by controlling the wavelength-selective switch (WSS) at the beginning of each optical multiplex section (OMS).

Figure 4(b) shows the amplified signal power spectra with and without the BIF-based online control. The first line shows the measured spectra controlled with traditional simple adjustments. Specifically, the mean value and tilt values of the measured signal spectrum are calculated through linear fitting. Subsequently, the differences in mean and tilt values between the measured and the target spectra are calculated and adjusted by the optical spectrum processor. The second line shows the spectra achieved by BIF. Three scenarios are considered with different target spectra. The corresponding RMSEs are shown in Fig. 4(c). Both the signal spectra and the error histograms can prove that, compared with the traditional spectrum controlling method, the BIF can achieve a better performance. To illustrate the changes of the spectrum during online control, Fig. 4(d) shows the changes of the RMSEs between the measured spectra and the target spectra when the target spectrum is flat with a value of −3 dB. Results show that the BIF can quickly adjust the signal spectrum within 30 iterations, demonstrating the efficiency of the BIF for online controlling.

3.4 Controlling RAs Based on BIF

The performance of the BIF for online controlling is also evaluated in systems with RA. For the OA controlling schemes, previous controlling strategies for controlling RAs can be categorized into two types. The first type relies on online heuristic algorithms, especially evolutionary algorithms.[52–54] For this type of method, one round of generation could include tens of candidates, necessitating several rounds of measurements for one iteration. The second type is based on the offline pretrained NN for optimization.[55–59] However, this type of method requires a pretrained model that needs a substantial number of measured spectra (hundreds to thousands). In our work, the BIF is utilized to adjust the power of each pump to obtain the target gain profile in a data-efficient manner. The experimental setup of the RA system is shown in Fig. 5(a). First, an ASE source is used to emulate the C+L-band signal spectra. After attenuation, the total signal power is set as 15.5 dBm. The transmitted fiber length is 82.8 km. In our experiment, we consider the counter-Raman amplification, which means only the backward pumps are utilized for amplification. The RA has four pumps, of which the wavelengths are 1428, 1454, 1490, and 1509 nm. Since the power of each Raman pump cannot be directly controlled, we control the pump power by adjusting the current of the digital-to-analog converter (DAC). Then, the on-off gain spectra are collected by an OSA. All the control and data processing are through a host computer.

In Fig. 5(d), 460 sets of pump configurations are generated randomly, and the corresponding on–off gain spectra are
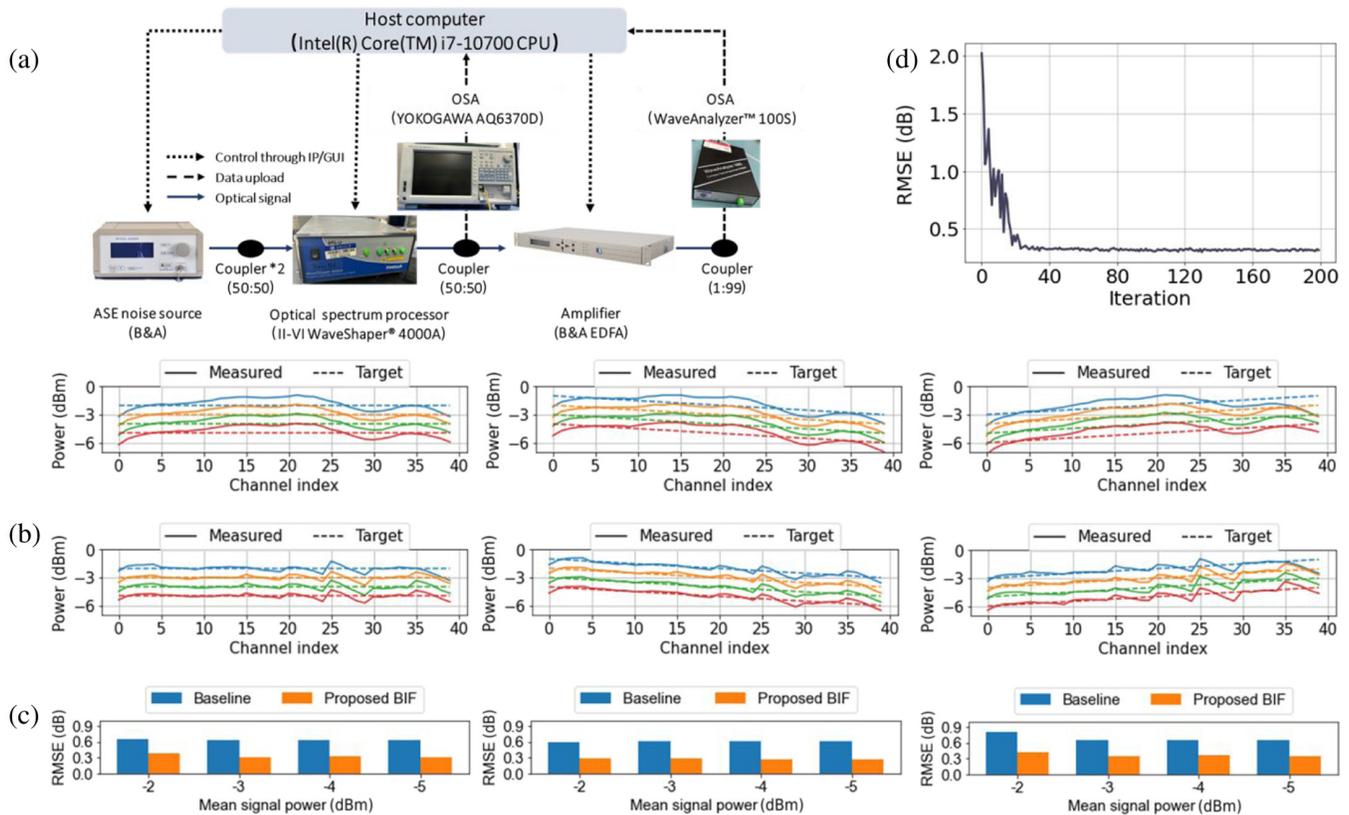
Fig. 4 The experimental setup for online controlling in an EDFA system. (a) The controlling work-flow of the automatic EDFA measuring system. (b) The signal power of each channel after amplification with different controlling targets. The solid lines are the results obtained by BIF and the dashed lines are the target signal spectra. Figures in the first line show the measured spectra when only adjusting the mean power and the tilt of the transmitted signal spectra. The second line shows the measured spectra when utilizing BIF. (c) The RMSEs' comparison between the traditional adjustment and BIF-based online controlling. (d) The changes of RMSEs between the measured spectra and the target spectra during BIF-based online controlling when the target gain is flat and set each channel as –2 dBm.

plotted. As shown in Fig. 5(d), the on–off gain constructed by different pump powers can range from 0 to 10 dB with different shapes. This result proves that, if directly searching for the pump configuration, the best configuration may be hard to achieve. Therefore, an efficient online controlling method is desired to realize various gain spectra without human intervention.

By applying the BIF in this use case, the size of the precollected initial data set is set as five. Afterwards, the exploitation-preferred BIF is employed with the EI optimizer. The RMSE of each iteration when setting a flat 7-dB gain spectrum as the target gain is plotted in Fig. 5(e). The results show that the BIF can quickly find the correct direction to adjust the pump power combinations and then converge to a low RMSE of ~0.3 dB. In Fig. 5(f), the on–off gain spectra of each iteration are plotted. We observe that the gain spectra gradually converge to the target gain, which demonstrates the effectiveness of the proposed method. In most cases, the BIF can approach the target spectrum within 10 iterations and then slightly fine-tune the pump configurations to obtain the optimum performance.

As shown in Figs. 5(g)–5(j), we plot some gain spectra during the online controlling as examples. First, the gain spectra are far away from the target gain spectrum before the fifth iteration. But then it quickly gets closer to the target in the tenth iteration by increasing the powers of all pumps. Afterwards, it starts fine-tuning the pump configuration and gradually achieves the optimal design. In Figs. 5(b) and 5(c), the online controlling performance based on the BIF under different target gain spectra is shown. The BIF can work well in multiple scenarios, and the convergence speed is relatively stable.

To further investigate the generalization of the proposed BIF, we validate the online controlling performance by considering situations with different pump numbers and wavelengths. We conduct simulations considering four, five, and six pumps in an RA system based on the GNPy. In simulations, more scenarios with different types of target gain spectra are investigated. First, the flat on–off gain spectra of 6, 8, 10, 12, and 14 dB are set for evaluation. In addition, the tilted gain spectra for compensating the SRS are set as the target gain. As shown in Fig. 6, the dashed lines are the target spectra and the solid lines are the spectra obtained by the BIF. The results show that the BIF can identify the best pump configurations to generate the desired gain spectra with various pump configurations.
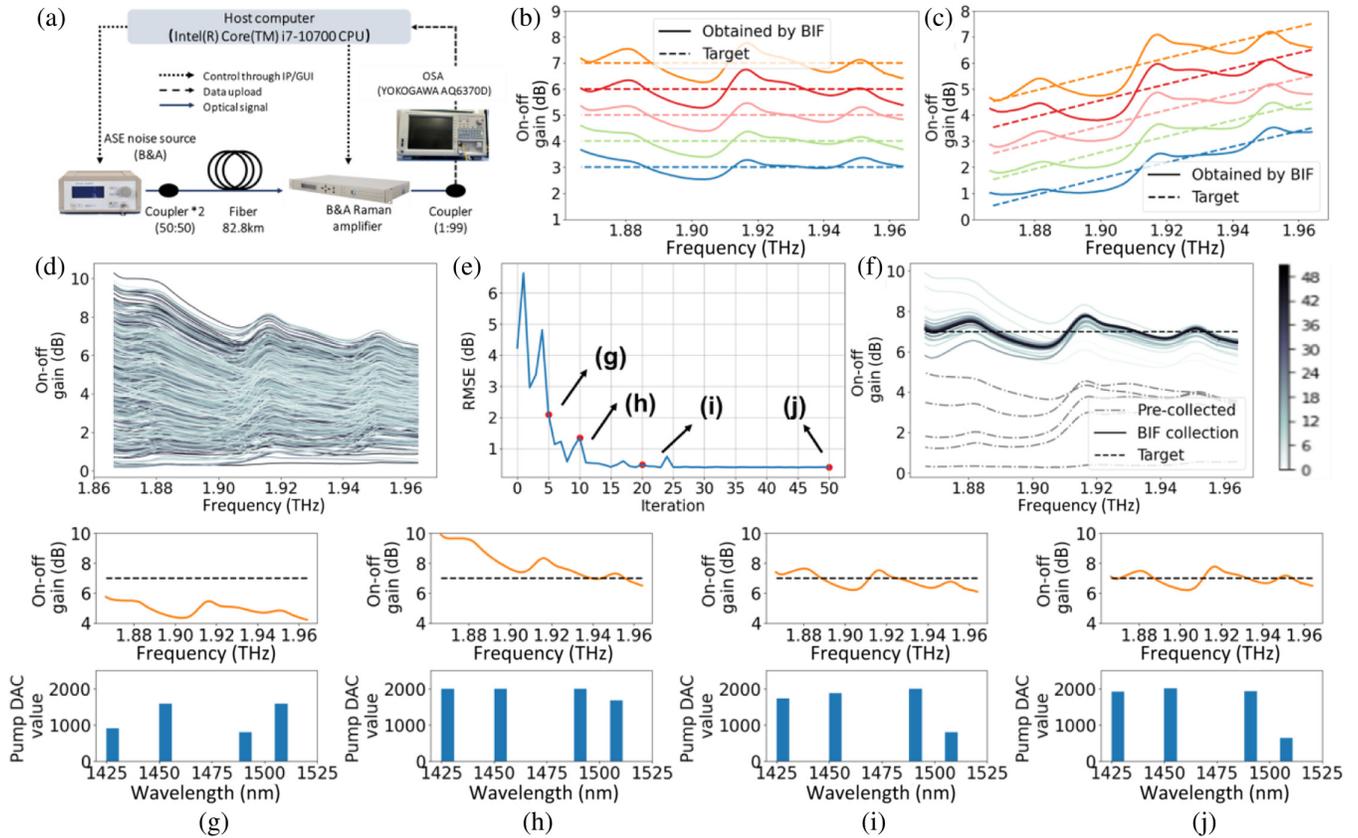
Fig. 5 The experimental setup for controlling the gain spectrum of an RA. (a) The controlling work-flow of the automatic RA controlling and gain measurement system. (b) The obtained flat gain spectra of the RA based on proposed BIF. (c) The obtained tilted gain spectra of the RA based on proposed BIF. (d) The randomly generated gain spectra of the RA under various combinations of pump powers. (e) The RMSEs for each iteration during the online controlling. (f) The measured gain spectra after each iteration during the online controlling. The gray lines are the precollected gain spectra, and the solid lines are the collected gain spectra during the online controlling. Different colors denote different iterations. The dashed black line is the target gain spectrum. Some obtained gain spectra during the online controlling, i.e., the red points in panel (e) are plotted in (g), (h), (i), and (j), respectively. The pump DAC values of each iteration are plotted in the second line accordingly.

4 Discussion

In this work, the Bayesian inference is utilized for modeling and controlling the OPE in optical fiber communication systems. The incorporation of Bayesian probability enriches the model's output by estimating both the mean and variance. Therefore, more comprehensive information is available for data selection. Additionally, the online iterative sampling scheme ensures that each piece of data is utilized to guide the subsequent data collection. Consequently, the efficiency of the data collection is enhanced to reduce the needed data size. Moreover, the BIF allows for flexible design of diverse data collecting objectives for both DT modeling and OA control. It shows the potential in constructing a DT during OA controlling, thereby facilitating future autonomous network operations.

The next critical step involves modeling the OPE during signal transmission across cascaded fiber spans and OAs. Considering a practical long-haul transmission system, the optical power is attenuated by fibers, connectors, and other devices, such as WSSs, and amplified by OAs. Therefore, modeling the OPE over a long-haul link requires the accurate modeling of each optical device and its cascaded effects.

Moreover, the complexity of OPE control arises due to the heterogeneous parameters from various devices in these long-haul links. This complexity is further compounded in scenarios with different wavelength loadings in each OMS. The control sequence and step size among different parameters in different OMSs should be carefully designed.

Additionally, frequent network operations impose higher reliability requirements. The control of OPE should not disrupt existing services, highlighting the need for a precise assessment of reliability in both modeling and controlling processes.

The BIF holds the potential to address the above challenges effectively. Its inherent ability to estimate probabilities makes it well suited for reliable assessment and data selection. Therefore, it can contribute to achieving efficient and highly reliable autonomous OPE modeling and control, aligning with the demands of future ADONs.
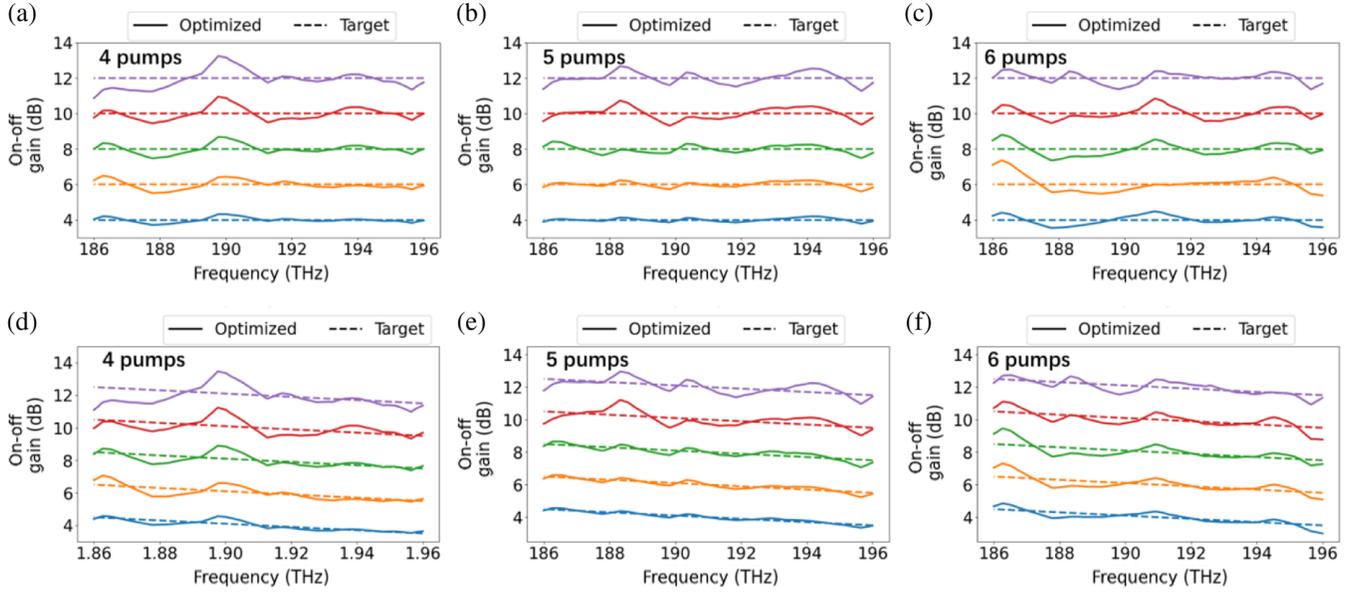
Fig. 6 The simulation performance of the BIF for online controlling with different on–off gain spectra. (a)–(c) The performance of the BIF with flat target gain. The simulated RAs have four, five, and six pumps, respectively. The solid lines are the experimental results and the dashed lines are the target spectra. (d)–(f) The performance of the BIF with tilt target gain. The simulated RAs have four, five, and six pumps, respectively. The solid lines are the results obtained by BIF, and the dashed lines are the target gain spectra.

5 Conclusion

Constructing the DT and controlling OPE is crucial for enabling multiband ADONs. To accomplish these goals, modeling and controlling OAs are the main challenges. In this work, we propose BIF to model and control OAs in a data-efficient manner. The BIF employs a selective data collection strategy that effectively balances exploration and exploitation in the search space.

Simulations and experiments have demonstrated the effectiveness of the BIF in modeling OAs. Compared to traditional NN-based models that use randomly selected data, the proposed BIF significantly reduces the data requirements for accurate modeling. Specifically, it can reduce the required data by 80% for EDFA and 60% for RA. This reduction in data requirements enhances the feasibility of deploying data-driven models in commercial OAs.

In terms of controlling, the BIF assists network controllers in adjusting OA configurations and transmitted signal profiles to achieve a target profile. Within a maximum of 30 iterations, the BIF successfully realizes the desired signal/gain profiles with RMSEs of <0.5 dB in most cases. Importantly, the proposed BIF is not limited to specific link conditions or OA types, making it applicable to a wide range of scenarios in various ADON systems.

6 Appendix: Methods

6.1 Simulation Details for RA Systems

The power propagation along a fiber with RA can be described by ODEs,[60] which can be written as

$$\frac{dP_s}{dz} = -2\alpha_s P_s + C_R(f_s, f_p)(P_p^+ + P_p^-)P_s, \quad (9)$$

$$\pm \frac{dP_p^{\pm}}{dz} = -2\alpha_p P_p^{\pm} - \left(\frac{f_p}{f_s}\right) C_R(f_s, f_p) P_s P_p^{\pm}, \quad (10)$$

where P_p and P_s represent the Raman pump power and signal power, respectively. z is the transmission distance. C_R is the Raman gain efficiency. f_s and f_p are the frequency of the signal and pump, respectively. α_s and α_p are the attenuation of the signal and pump, respectively. The on–off gain can be calculated by

$$G_{\text{ON-OFF}} = \frac{P_{s,\text{pumps ON}}(L)}{P_{s,\text{pumps OFF}}(L)}, \quad (11)$$

where $P_{s,\text{pumps ON}}(L)$ and $P_{s,\text{pumps OFF}}(L)$ represent the signal power at the distance of L with Raman pumps on and off, respectively.

The simulation verifications are conducted using the GNPy.[50] The fiber in simulations is SSMF with an attenuation of 0.2 dB/km, a nonlinearity coefficient of 1.3 W^{-1} km^{-1}, and a chromatic dispersion coefficient of 16.7 ps nm^{-1} km^{-1}.

6.2 Training Details for EDFA DT Models

In the use case of EDFA modeling, two types of models are trained. First, the baseline NN-based model has two hidden fully connected layers with 40 neurons in each layer. The activation functions are sigmoid.[61] The optimizer is Adam.[62] During training, 80% of the available data are used for training, and 20% are used for validation. The total epoch number is set as 10^6, and early stop is employed with a patience of 100. For the proposed GPR-based model, the utilized kernel is the radial basis function (RBF).[61] The noise hyperparameter, i.e., alpha, is set as 10^{-5}.

To conduct a fair comparison, the two models share the same input and output. Specifically, the input feature is a vector representing the power value of each WDM signal before amplification. Since 40 channels are considered, the feature vector has 40 dimensions. The output is the gain value of the signal in each WDM channel. Min-max normalization is conducted for both the input features and labels because we found that this type of preprocessing can achieve the best performance. RMSE is calculated for accuracy evaluation. Moreover, since some of the channels are idle, the estimations of these channels are deleted for RMSE calculation.

6.3 Training Details for RA DT Models

The training processes of RA modeling are similar to those for EDFA modeling but with different input/output data dimensions and hyperparameters. For training data, the input features include the signal power of each WDM channel, represented as an 80-dimensional vector. The outputs are the GSNRs of each WDM signal, represented as an 80-dimensional vector as well. The NN-based baseline models have two hidden layers with 80 neurons in each layer. The activation function is sigmoid. The optimizer is Adam. 80% of the data are used for training and 20% of the data are used for validation. The total epoch number is set as 10^6, and early stop is employed with a patience of 1000. The GPR-based surrogate model employs an RBF kernel. Min-max normalization is utilized to process both the input and output of the data set.

6.4 Parameters for Controlling the EDFA System

The controlling objectives are the attenuation factors of the WaveShaper for every five consecutive WDM signals. For the total 40 channels, there are eight parameters. The online controlling is realized by utilizing the Bayesian-optimization Python tool.[63] During online controlling based on the BIF, the surrogate GPR model has a noise hyperparameter of 10^{-4}. The number of initial sampling data is 2, and an EI optimizer with a ξ of 10^{-9} is employed. Domain reduction[64] with a minimum window length of 0.2 is used to speed up convergence.

6.5 Parameters for Controlling the RA System

To control Raman pump powers, the configured parameters are the DAC values of four pumps. The Bayesian-optimization Python tool[63] is utilized for conducting online controlling, and the initial sampling number is 5. The hyperparameter of the surrogate GPR model is 10^{-5}. The EI optimizer with a ξ of 10^{-5} is employed. Domain reduction with a minimum window length of 0.01 is used.

Code and Data Availability

The source data and code are available from authors upon the reasonable request.

References

1. P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: the previous 20 and the next 20 years [Invited]," *Opt. Express* **26**, 24190–24239 (2018).
2. B. Lee, "Review of the present status of optical fiber sensors," *Opt. Fiber Technol.* **9**, 57–79 (2003).
3. D. M. Chow et al., "Distributed forward Brillouin sensor based on local light phase recovery," *Nat. Commun.* **9**, 2990 (2018).
4. C. Pang et al., "Opto-mechanical time-domain analysis based on coherent forward stimulated Brillouin scattering probing," *Optica* **7**, 176–184 (2020).
5. B. Warf, "International competition between satellite and fiber optic carriers: a geographic perspective," *Prof. Geogr.* **58**, 1–11 (2006).
6. Cisco, "Global_2021_Forecast_Highlights," https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2021_Forecast_Highlights.pdf (2016).
7. Cisco, "Cisco annual internet report - Cisco annual internet report (2018–2023) white paper," https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internetreport/white-paper-c11-741490.html (2020).
8. J. Cho et al., "Shaping lightwaves in time and frequency for optical fiber communication," *Nat. Commun.* **13**, 785 (2022).
9. A. P. T. Lau et al., "Advanced DSP techniques enabling high spectral efficiency and flexible transmissions: toward elastic optical networks," *IEEE Signal Process. Mag.* **31**, 82–92 (2014).
10. C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.* **27**, 379–423 (1948).
11. A. Napoli et al., "Towards multiband optical systems," in *Adv. Photonics 2018 (BGPP, IPR, NP, NOMA, Sens., Networks, SPPCom, SOF)*, p. NeTu3E.1 (2018).
12. R. G. H. Van Uden et al., "Ultra-high-density spatial division multiplexing with a few-mode multicore fibre," *Nat. Photonics* **8**, 865–870 (2014).
13. M. Cantono et al., "Opportunities and challenges of C+L transmission systems," *J. Light. Technol.* **38**(5), 1050–1060 (2020).
14. D. Rafique and L. Velasco, "Machine learning for network automation: overview, architecture, and applications [Invited tutorial]," *J. Opt. Commun. Networking* **10**, D126 (2018).
15. H. Zheng et al., "From automation to autonomous: driving the optical network management to fixed fifth-generation (F5G) advanced," in *IEEE 9th Int. Conf. Network Softwarization (NetSoft)*, pp. 385–389 (2023).
16. M. P. Yankov, U. C. de Moura, and F. D. Ros, "Power evolution modeling and optimization of fiber optic communication systems with EDFA repeaters," *J. Light. Technol.* **39**, 3154–3161 (2021).
17. G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed., Academic Press (2001).
18. E. Temprana et al., "Overcoming Kerr-induced capacity limit in optical fiber transmission," *Science* **348**, 1445–1448 (2015).
19. A. R. Chraplyvy, "Optical power limits in multi-channel wavelength-division-multiplexed systems due to stimulated Raman scattering," *Electron. Lett.* **20**, 58–59 (1984).
20. S. Tariq and J. C. Palais, "A computer model of non-dispersion-limited stimulated Raman scattering in optical fiber multiple-channel communications," *J. Light. Technol.* **11**, 1914–1924 (1993).
21. D. Semrau, R. I. Killey, and P. Bayvel, "The Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Light. Technol.* **36**, 3046–3055 (2018).
22. H. Buglia et al., "An extended version of the ISRS GN model in closed-form accounting for short span lengths and low losses," in *Eur. Conf. Opt. Commun. (ECOC)*, pp. 1–4 (2022).
23. A. K. Srivastava et al., "EDFA transient response to channel loss in WDM transmission system," *IEEE Photonics Technol. Lett.* **9**, 386–388 (1997).
24. Y. You, Z. Jiang, and C. Janz, "Machine learning-based EDFA gain model," in *Eur. Conf. Opt. Commun. (ECOC)*, pp. 1–3 (2018).

25. F. da Ros, U. C. de Moura, and M. P. Yankov, "Machine learning-based EDFA gain model generalizable to multiple physical devices," in *Eur. Conf. Opt. Commun. (ECOC)*, p. Tu1A-4 (2020).

26. J. Yu et al., "Machine-learning-based EDFA gain estimation [Invited]," *Commun. Networking* **13**, B83–B91 (2021).

27. Z. Jiang, J. Lin, and H. Hu, "Machine learning based EDFA channel in-band gain ripple modeling," in *Opt. Fiber Commun. Conf. (OFC)*, p. W4I.2 (2022).

28. B. Pedersen et al., "Experimental and theoretical analysis of efficient erbium-doped fiber power amplifiers," *IEEE Photonics Technol. Lett.* **3**, 1085–1087 (1991).

29. R. Sommer et al., "Multiple filter functions integrated into multi-port GFF components," in *OFC/NFOEC 2007-2007 Conf. Opt. Fiber Commun. and the Natl. Fiber Opt. Eng. Conf.*, pp. 1–3 (2007).

30. L. Rapp and M. Eiselt, "Optical amplifiers for multi–band optical transmission systems," *J. Light. Technol.* **40**, 1579–1589 (2022).

31. A. A. M. Saleh et al., "Modeling of gain in erbium-doped fiber amplifiers," *IEEE Photonics Technol. Lett.* **2**, 714–717 (1990).

32. C. R. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *J. Lightwave Technol.* **9**, 271–283 (1991).

33. M. Hashimoto, M. Yoshida, and H. Tanaka, "The characteristics of WDM systems with hybrid AGC EDFA in the photonics network," in *Opt. Fiber Commun. Conf. and Exhibit*, pp. 517–518 (2002).

34. J. Junio, D. C. Kilper, and V. W. S. Chan, "Channel power excursions from single-step channel provisioning," *J. Opt. Commun. Networking* **4**, A1–A7 (2012).

35. K. Ishii, J. Kurumida, and S. Namiki, "Experimental investigation of gain offset behavior of feed forward-controlled WDM AGC EDFA under various dynamic wavelength allocations," *IEEE Photonics J.* **8**, 7901713 (2016).

36. Z. Wang, D. Kilper, and T. Chen, "Transfer learning-based ROADM EDFA wavelength dependent gain prediction using minimized data collection," in *Opt. Fiber Commun. Conf. and Exhibit. (OFC)*, pp. 1–3 (2023).

37. Y. You, Z. Jiang, and C. Janz, "OSNR prediction using machine learning-based EDFA models," in *45th Eur. Conf. Opt. Commun. (ECOC 2019)*, pp. 1–3 (2019).

38. Y. Liu et al., "Modeling EDFA gain: approaches and challenges," *Photonics* **8**, 417 (2021).

39. K. Débicki, "Gaussian process: overview," in *Wiley StatsRef: Statistics Reference Online*, N. Balakrishnan et al., Eds., 1st ed., Wiley (2014).

40. K. Kandasamy, J. Schneider, and B. Poczos, "Bayesian active learning for posterior estimation," in *24th Int. Joint Conf. Artif. Intell. (IJCAI)*, pp. 3605–3611, AAAI Press, Menlo Park (2015).

41. E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," arXiv:1012.2599 (2010).

42. J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Adv. Neural Inf. Process. Syst.*, Curran Associates, Inc., Vol. **25** (2012).

43. P. I. Frazier, "A tutorial on Bayesian optimization," arXiv:1807.02811 (2018).

44. Z. Zhong et al., "BOW: first real-world demonstration of a Bayesian optimization system for wavelength reconfiguration," in *Opt. Fiber Commun. Conf. (OFC)*, p. F3B.1 (2021).

45. D. Zhan and H. Xing, "Expected improvement for expensive optimization: a review," *J. Global Optim.* **78**(3), 507–544 (2020).

46. A. Garivier and E. Moulines, "On upper-confidence bound policies for switching bandit problems," *Lect. Notes Comput. Sci.* **6925**, 174–188 (2011).

47. M. P. Yankov and F. Da Ros, "Input-output power spectral densities for three C-band EDFAs and four multi-span inline EDFAd fiber optic systems of different lengths," Technical University of Denmark (2020).

48. Z. Wang, D. C. Kilper, and T. Chen, "Open EDFA gain spectrum dataset and its applications in data-driven EDFA gain modeling," *J. Opt. Commun. Networking* **15**, 588 (2023).

49. S. Zhu et al., "Hybrid machine learning EDFA model," in *Opt. Fiber Commun. Conf. (OFC)*, p. T4B.4 (2020).

50. A. Ferrari et al., "GNPy: an open source application for physical layer aware open optical networks," *J. Opt. Commun. Networking* **12**, C31–C40 (2020).

51. P. Poggiolini et al., "The GN-model of fiber non-linear propagation and its applications," *J. Lightwave Technol.* **32**, 694–721 (2014).

52. B. Neto et al., "Efficient use of hybrid genetic algorithms in the gain optimization of distributed Raman amplifiers," *Opt. Express* **15**, 17520 (2007).

53. S. Singh and R. S. Kaler, "Performance optimization of EDFA–Raman hybrid optical amplifier using genetic algorithm," *Opt. Laser Technol.* **68**, 89–95 (2015).

54. J. Chen and H. Jiang, "Optimal design of gain-flattened Raman fiber amplifiers using a hybrid approach combining randomized neural networks and differential evolution algorithm," *IEEE Photonics J.* **10**, 7101915 (2018).

55. J. Zhou et al., "Robust, compact, and flexible neural model for a fiber Raman amplifier," *J. Lightwave Technol.* **24**, 2362–2367 (2006).

56. D. Zibar et al., "Inverse system design using machine learning: the Raman amplifier case," *J. Lightwave Technol.* **38**, 736–753 (2020).

57. X. Ye et al., "Experimental prediction and design of ultra-wide-band Raman amplifiers using neural networks," in *Opt. Fiber Commun. Conf. and Exhibit. (OFC)*, p. W1K.3 (2020).

58. G. Marcon et al., "Model-aware deep learning method for Raman amplification in few-mode fibers," *J. Lightwave Technol.* **39**, 1371–1380 (2021).

59. M. P. Yankov et al., "Flexible Raman amplifier optimization based on machine learning-aided physical stimulated Raman scattering model," *J. Lightwave Technol.* **41**, 508–514 (2023).

60. M. N. Islam, "Raman amplifiers for telecommunications," *IEEE J. Sel. Top. Quantum Electron.* **8**, 548–559 (2002).

61. C. M. Bishop, *Pattern Recognition and Machine Learning, Information Science and Statistics*, Springer (2006).

62. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2017).

63. https://github.com/fmfn/BayesianOptimization.

64. N. Stander and K. J. Craig, "On the robustness of a simple domain reduction scheme for simulation-based optimization," *Eng. Comput.* **19**, 431–450 (2002).

Xiaomin Liu received her BE degree in information engineering from Shanghai Jiao Tong University (SJTU) in 2020. She is currently pursuing a PhD in the Department of Electronic Engineering, SJTU. Her current research interests include modeling, monitoring, and optimization in optical networks.

Yihao Zhang received his BS degree in information engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2021. He is currently pursuing a PhD in electronic and information engineering at SJTU. His research interests include the optical amplifier modeling and optimization, optical networks modeling and optimization, and fiber nonlinearity modeling.

Yuli Chen received his BS in material science and engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2023. He is currently working toward his MS degree in electronic and information engineering at SJTU. His research interests include optical amplifier modeling and optimization, optical networks modeling and optimization, and optical performance monitoring.

Yichen Liu received her BS degree in electronic engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2021. She is currently pursuing an MS degree in electronic and information engineering at SJTU. Her research interests include the optical amplifier modeling and optimization.

Meng Cai received her BS degree in optoelectronic information science and engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2017. She received her master's degree in electronic and information engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2020. She is currently pursuing her PhD in electronic and information engineering from SJTU. Her research interests include optical network modeling, monitoring, and failure management in optical networks.

Qizhi Qiu received his BS degree in electronic engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2022. He is currently pursuing a PhD in information and communication engineering at SJTU. His research interests include optical network monitoring and optimization, and the application of machine learning techniques in optical networks.

Mengfan Fu received her BE degree in information engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China in 2019. She is currently pursuing a PhD in the State Key Laboratory of Advanced Optical Communication Systems and Networks of SJTU. Her current research interests include optical digital coherent technologies in ultra-wideband long-haul systems and low-cost systems for data center interconnect.

Lilin Yi, distinguished professor, vice director of the State Key Laboratory of Advanced Optical Communication Systems and Network of Shanghai Jiao Tong University. He received a PhD from the Ecole Nationale Supérieure des Télécommunications, currently named as Telecom ParisTech, France, and SJTU, China, on March and June 2008, respectively as a joint-educated PhD student. His main research topics include intelligent optical fiber communication systems and intelligent fiber laser systems.

Weisheng Hu: Biography is not available.

Qunbi Zhuge is an associate professor in the Department of Electronic Engineering at Shanghai Jiao Tong University. His current research interests include long-haul optical communication, intelligent optical network, data center interconnects, and optical-wireless convergence. He has published more than 220 journal and conference papers. He is PI and co-PI of several national research grants. He was named as "Innovators Under 35" in China by MIT Technology Review in 2020 and is a recipient of many international awards.