# PtrNet-RSA: A Pointer Network-Based QoT-Aware Routing and Spectrum Assignment Scheme in Elastic Optical Networks

Yuansen Cheng, Shifeng Ding, Yingjie Shao, and Calvin Chun-Kit Chan, *Senior Member, IEEE, Fellow, Optica*

***Abstract*—To enable flexible service provisioning in elastic optical networks (EONs), quality of transmission (QoT) estimation and dynamic routing and spectrum assignment (RSA) are critical network management tasks. Recently, machine learning approaches have been extensively investigated to address these tasks individually. We propose to take an integrative approach to design a Pointer Network (PtrNet)-based QoT-aware RSA (PtrNet-RSA) scheme to optimize the blocking probability and the generalized signal-to-noise ratio (GSNR) for the lightpath allocation in EONs. Given the profiles of the physical network and the service requests, the proposed scheme can generate high-GSNR lightpaths without pre-calculated candidate paths. Reinforcement learning is employed to train the model while the parameters of the PtrNets are optimized by utilizing the actor-critic algorithm. Extensive simulations have been conducted in EON environments based on the Gaussian Noise (GN) model with dynamic traffic. The results show that the proposed PtrNet-RSA scheme consistently outperforms the benchmarks with various network topologies, significantly reducing the blocking probability while guaranteeing a good lightpath QoT.**

***Index Terms*—Elastic optical networks, pointer network, quality of transmission, reinforcement learning, routing and spectrum assignment.**

## I. Introduction

THE continuous growth of network traffic, driven by the emergence of 5G/F5G services, such as cloud services, live videos, and autonomous vehicles, has created tremendous pressure on the current network infrastructure. With the high transmission capacity and flexibility, elastic optical networks (EONs) have been among the promising ways to relieve the pressure for next-generation backbone networks [1]. Equipped with bandwidth-variable transponders, EONs enable dynamic service provisioning with flexible bandwidth allocation. Compared to traditional WDM networks, EONs realize finer bandwidth granularities, which can lead to lower network margins to improve the network efficiency further [2]. However, this would lead to critical requirements on the quality of transmission (QoT) in the lightpath provisioning [3]. Therefore, an efficient resource allocation scheme with a sufficient QoT guarantee is highly desirable for network management in EONs.

In EONs, the QoT of the lightpath is mainly determined by physical-layer impairments, including amplifier spontaneous noise emission (ASE), self-channel interference (SCI), and cross-channel interference (XCI). Among them, XCI contributes the highest computational complexity in the estimation. Generally, it is calculated with exact numerical solvers, such as the split-step Fourier method [4] or the Gaussian noise (GN) model [5]. These methods can provide precise estimation results on QoT at the expense of high computation complexity. Recently, machine learning (ML) based QoT estimation approaches [6] have shown to have reduced computation complexity yet achieve a considerable QoT-estimation accuracy by utilizing lightpath-based [7], [8], [9] or network-wide information [10], [11]. With computation-efficient QoT estimation, the next step is to get effective solutions for the routing and spectrum assignment (RSA). The RSA problem is challenging in EON due to the spectrum continuity and contiguity constraints. Recently, ML has also been employed to tackle this problem. Leveraging either supervised learning or reinforcement learning (RL), one can turn the RSA problem into a classification problem, which can significantly reduce the complexity and obtain a considerably low blocking probability [12], [13].

Although the QoT estimation and the RSA problem have been extensively investigated, applying ML to the integrated problem is challenging [8]. In the static traffic model, the QoT estimation results can be easily incorporated into the RSA problem as the reach constraints of the lightpaths since all the demands are known in advance [14]. However, service requests arrive and expire dynamically in the dynamic traffic model. Evaluating all lightpath candidates for each service request [15] would lead to high computation complexity [11]. Alternatively, QoT estimation can be performed after the RSA model [16], [17], to determine whether the QoT of the lightpath under test (LUT) is above a predefined threshold. This can reduce the computational complexity but may lead to a poor allocation

Manuscript received 4 January 2024; revised 12 April 2024; accepted 22 May 2024. Date of publication 27 May 2024; date of current version 2 September 2024. *(Corresponding author: Calvin Chun-Kit Chan.)*

Yuansen Cheng, Shifeng Ding, and Calvin Chun-Kit Chan are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, SAR, China (e-mail: cy019@ie.cuhk.edu.hk; ds020@ie.cuhk.edu.hk; ckchan@ie.cuhk.edu.hk).

Yingjie Shao is with the Centre for Applied Photonics, Fraunhofer U.K. Research Ltd., G1 1RD Glasgow, U.K. (e-mail: yingjie.shao@fraunhofer.co.uk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2024.3405587.

Digital Object Identifier 10.1109/JLT.2024.3405587

solution. For instance, due to the XCI, future service provisioning may cause signal-to-noise ratio (SNR) degradation in the deployed lightpaths, potentially causing the QoT of the LUT to fall below the threshold.

In this paper, we propose a pointer network (PtrNet)-based QoT-aware RSA scheme (PtrNet-RSA) to integrate the QoT estimation and RL-based RSA for dynamic service provisioning in EONs. In addition to the blocking probability, a QoT metric, namely the generalized signal-to-noise ratio (GSNR), is also considered the optimization objective. Given the network and the service request information, the proposed PtrNet-RSA scheme can directly generate high-GSNR lightpaths without using pre-calculated candidate paths. A preliminary version of PtrNet-RSA was introduced in our previous work [18]. In this paper, we extend the work by exploring the scalability under different network topologies and investigating the tradeoff between the blocking probability and the QoT of the lightpaths. The key contributions can be summarized as follows:

- Considering physical-layer impairments, we propose a PtrNet-RSA scheme that integrates the QoT estimation inside the dynamic RSA process in EONs. PtrNet-RSA utilizes the link-level network information and allows a direct generation of high-QoT lightpaths with the GSNR estimation for the service requests, significantly reducing computational complexity.
- Leveraging the PtrNet, the proposed scheme introduces flexibility in inputted network topologies and enables a node-by-node approach to find the desired lightpaths. It expands the output action space, facilitating the selection of lightpaths from all possible paths between the source and the destination nodes instead of selecting from the pre-calculated candidates, which substantially reduces the blocking probability.
- A novel allocation policy is proposed to handle the trade-off between the blocking probability and the QoT of the lightpaths. By introducing a priority factor, a quantified priority is assigned to lightpaths with lower frequency slot (FS) indices, mitigating the impact of spectrum fragmentations.
- Extensive simulations have been conducted to verify the efficiency of the proposed scheme under different network topologies with extensive network metrics. Results show that the proposed PtrNet-RSA scheme can significantly improve the QoT on lightpaths and reduce the blocking probability.

The rest of the paper is organized as follows. Section II reviews the related work on QoT estimation and RL-based RSA. Section III presents the RSA problem formulation and the analytical QoT model in EONs. Section IV discusses the design and implementation of PtrNet-RSA, including the modeling, operation mechanisms, and training. In Section VI, we evaluate the performance of the proposed PtrNet-RSA scheme and give some discussions. Finally, Section VII concludes the paper.

## II. Related Work

In this section, we review the reported studies on QoT estimation, ML-assisted RSA, and QoT-aware RSA.

### A. QoT Estimation

The split-step Fourier methods were usually employed to solve the nonlinear Schrödinger equation in optical systems, demonstrating high efficiency and robustness [19]. Then, the Gaussian-noise (GN) models were proposed to approximate the non-linear fiber propagation models [5], showing simplicity and reliability in predicting performance for uncompensated coherent systems. Nevertheless, these approaches have extremely high computation complexity.

To alleviate the high computational complexity, ML techniques have been widely studied to transform the QoT estimation into a classification problem. In [20], an ML-based classification approach was proposed to predict the probability that a candidate lightpath's bit error rate (BER) would not exceed a given threshold based on the lightpath characteristics. In [21], a cognitive QoT estimator was proposed to significantly reduce the computational time and achieve over 98% successful classifications by optimizing the dataset through the learning and forgetting techniques in optical networks. To further improve the classification accuracy, a deep convolutional neural network formulation [10] was proposed for network-wide QoT classification tasks in optical networks with significant topological differences, achieving more than 99% accuracy with extensive and diverse test datasets. Apart from classification, regression-based approaches have also been studied. For instance, ML was used to refine existing models, like the GN model, to improve fiber nonlinear impairments (NLI) modeling and monitoring in dynamic optical networks [9]. In [22], a data-driven QoT model was investigated with unicast and multicast connections in metropolitan optical networks, which achieved 95% estimation accuracy. In [23], signal power measurements from the optical amplifiers and the channel monitors in a disaggregated and partially loaded live production network were combined to reduce the standard deviation of the GSNR estimation error significantly.

### B. ML-Assisted RSA

ML approaches have also been employed to tackle the RSA problem. In [24], an ML-based routing and wavelength assignment (RWA) scheme was proposed to convert the original RWA into a supervised classification problem, which achieves near-optimal solutions while significantly reducing the computational time in WDM networks. In [12], a deep reinforcement learning framework for routing, modulation, and spectrum assignment (DeepRMSA) was proposed to significantly reduce the blocking probability compared with the K-shortest path routing and the first-fit (KSP-FF) benchmark. As the routing and spectrum assignment (SA) were completed separately in DeepRMSA, MaskRSA was proposed [13] to integrate the deep reinforcement learning (DRL) based routing with a DRL-based spectrum assignment algorithm. It extended the action space of the path-FS pairs and the mask unassignable choices to achieve better performance. Moreover, to further utilize the link-level features, the graph convolutional neural network (GCN) and the recurrent neural network (RNN) were employed in [25] for the feature extraction of the network topology, demonstrating superior performance compared to DeepRMSA.

### C. QoT-Aware RSA

The QoT estimation and RSA schemes have been shown to attain high estimation accuracy and low blocking probabilities with the aid of ML, respectively. Nevertheless, the performance may not be practical when considering them jointly due to their interactions and mutual constraints. To address this, in [17], the output of a neural network-based Q-factor estimator was exploited by a heuristic algorithm for dynamic RSA in a multicast scenario. This could guarantee the QoT of the LUT. However, the deployed lightpaths might suffer SNR degradation after the new traffic provisioning. An SNR re-verification scheme [26] was further implemented to ensure the SNR values of all deployed requests in the network. In [27], a genetic algorithm-based heuristic was proposed to improve the signal quality tolerance of existing lightpaths.

To further improve the efficiency of the integrated QoT estimation and RSA, the authors in [14] proposed a dual-stage iterative RSA optimization framework that incorporated the QoT estimations provided by an ML regressor into a mixed integer linear programming (MILP) formulation for static traffic allocation. In [28], the authors proposed an RL-based NLI-aware routing, modulation, and spectrum assignment (NLI-RMSA) scheme that could significantly improve the channel state estimation accuracy in EONs. However, the outcome of the RL model was only an action, such as deploying or dropping a channel, while no QoT estimation results were attained for the execution of the RSA algorithm. Moreover, the NLI-RMSA scheme was only applied for sliding scheduled traffic conditions, and the influence of XCI on deployed lightpaths was not considered.

In this paper, we propose a PtrNet-RSA scheme that integrates QoT estimation into the RSA process to generate high-QoT lightpaths along with the estimated GSNR values for EONs in dynamic traffic scenarios. Moreover, unlike the previous RL-based RSA methods relying on fixed paths calculated by the KSP algorithm, PtrNet-RSA finds lightpath node by node, significantly improving allocation flexibility.

## III. Problem Statement and Network Model

In this section, we introduce the proposed network and traffic model for the RSA problem in EONs and the QoT mathematical model for constructing the RL environment.

### A. Problem Formulation

The EON topology is modeled as a graph $G(V, E, F)$, where $V$ and $E$ represent the sets of nodes and fiber links, respectively, and $F$ contains the state of each frequency slot (FS) on each fiber link. For instance, $F = \{F_{e,f} \mid e, f\}$, where $e$ and $f$ are the indices of fiber links and FSs. The total number of FSs is $N$. Service requests are modeled by a tuple $R(s, d, M, \tau)$, where $(s, d)$ is the source-destination pair, $M$ is the number of required frequency slots, and $\tau$ is the service duration. In this paper, we focus on the dynamic RSA problem in EON, which consists of assigning every dynamic traffic request a lightpath connecting its source node to its destination node over a commensurate spectral bandwidth. The objective of the proposed PtrNet-RSA scheme is to jointly minimize the blocking probability and maximize the QoT of the lightpaths, considering the physical-layer impairments.

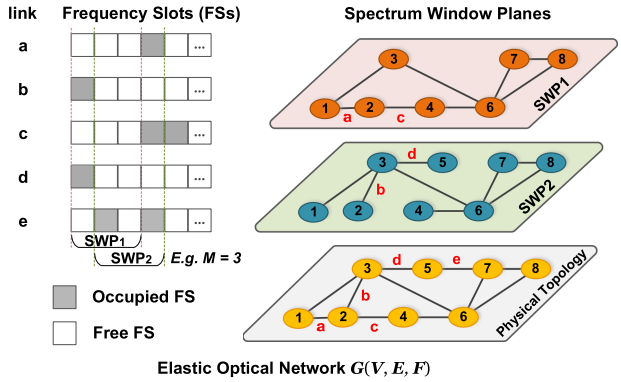



Fig. 1. Spectrum window planes for elastic optical network.

We assume that EONs do not have the spectrum conversion capability. Thus, two vital constraints should be considered in RSA, namely the *spectrum contiguity*, in which the assigned FSs to each service request should be contiguous; and the *spectrum continuity*, in which the indices of the assigned FSs should be the same on all links along the selected path. With these two constraints, we need to assign $M$ consecutive FSs for a service request $R(s, d, M, \tau)$. The concept of spectrum window planes (SWPs) is employed in this paper [29]. As shown in Fig. 1, the EON is split into multiple SWPs, where each of them consists of $M$ consecutive FSs, and in each plane, a virtual link between a pair of nodes is connected if the spectrum in the corresponding fiber link is available [30], [31]. By dividing the whole spectrum into $(N - M + 1)$ SWPs, the RSA problem can be completed in two steps, in which one SWP is first chosen before a path is determined in the selected SWP. In a dynamic scenario, each SWP may have different available links due to the set-up and the tear-down of the lightpaths, leading to varying SWP topologies. Fig. 1 presents an example when the number of required frequency slots $M$ equals 3, in which links $b$ and $d$ are unavailable in SWP1, links $a$ and $c$ are unavailable in SWP2, and link $e$ is unavailable in either SWP1 or SWP2. These dynamics pose a significant challenge for the previous RL-based RSA models, since their action space consists of fixed path candidates. Our proposed PtrNet-RSA scheme aims to find the lightpath node by node, getting rid of the limitation of fixed pre-calculated path candidates.

### B. QoT Mathematic Model

In this paper, the GSNR is adopted as the QoT metric, which is defined as [32], [33], [34].

$$\text{GSNR} = \frac{P_{ch}}{P_{ASE} + P_{NLI}} \cdot FP \tag{1}$$

where $P_{ch}$ is the launch power of the LUT, $P_{ASE}$ and $P_{NLI}$ are the accumulated power of the ASE and NLI noise, respectively, and $FP$ is the penalty introduced by the filtering effects [35].

The accumulated ASE noise is calculated as [36], [37].

$$P_{ASE} = \sum_{l \in L_d} P_{ASE}^{l} \tag{2}$$

$$P_{ASE}^{l} = \sum_{s}^{N_s^l} h \cdot f \cdot NF \cdot \left(e^{-2\alpha z_s^l} - 1\right) \cdot B_N \tag{3}$$

where $L_d$ is the set of link indices that belong to the LUT, $P_{ASE}^{l}$ is the ASE noise power on link $l$, $N_s^l$ is the number of spans in link $l$, $h$ is Planck's constant, $f$ is the central frequency of the channel to which the LUT is assigned, $NF$ is the noise figure of EDFAs, $B_N$ is the noise bandwidth; $\alpha$ is the fiber attenuation, and $z_s^l$ is the span length of the $s$-th span of link $l$.

The accumulated NLI power is a summation of NLI noise power for all links along the LUT, calculated as

$$P_{NLI} = \sum_{l \in L_d} \left(P_{SCI}^{l} + P_{XCI}^{l}\right) \tag{4}$$

where $P_{SCI}^{l}$ is the SCI noise power and $P_{XCI}^{l}$ is the XCI noise power on link $l$.

SCI and XCI are caused by the Kerr effect and the stimulated Raman scattering, which can be approximated by the Gaussian Noise (GN) model as [38]

$$P_{SCI}^{l} = B_N \cdot \frac{16}{27}\gamma^2 z_{eff}^2 G_{ch,i} G_{ch,i} G_{ch,i} \varphi_{i,i} \tag{5}$$

$$P_{XCI}^{l} = B_N \cdot \frac{16}{27}\gamma^2 z_{eff}^2 \sum_{n=1}^{N_{ch}} G_{ch,n} G_{ch,n} G_{ch,i} \varphi_{n,i} \tag{6}$$

where $\gamma$ is the fiber nonlinearity coefficient, $N_{ch}$ is the number of interfering channels, $G_{ch,n}$ and $G_{ch,i}$ are the power spectral density (PSD) of the interfering channel $n$ and the channel under test (CUT) $i$, respectively. Moreover, the effective length $z_{eff}$ is formulated as

$$z_{eff} = \frac{\left(1 - e^{-2\alpha z_s^l}\right)}{2\alpha} \tag{7}$$

The SCI and XCI coefficients are $\varphi_{i,i}$ and $\varphi_{n,i}$ ($n \neq i$), respectively, given by

$$\varphi_{i,i} = N_s^{l\,1+\in} \frac{\alpha}{\pi \left|\beta_2\right|} \text{asinh}\left(\frac{\frac{\pi^2}{2}\left|\beta_2\right| \cdot B_{ch,i}^2}{2\alpha}\right) \tag{8}$$

$$\varphi_{n,i} = N_s^l \frac{\alpha}{2\pi \left|\beta_2\right|} \left[\begin{array}{l} \text{asinh}\left(\frac{\pi^2 |\beta_2| \left[f_{ch,n} - f_{ch,i} + \frac{1}{2} B_{ch,n}\right] B_{ch,i}}{2\alpha}\right) \\ -\text{asinh}\left(\frac{\pi^2 |\beta_2| \left[f_{ch,n} - f_{ch,i} - \frac{1}{2} B_{ch,n}\right] B_{ch,i}}{2\alpha}\right) \end{array}\right] \tag{9}$$

where $\beta_2$ is the second order dispersion parameter, $B_{ch,i}$ and $B_{ch,n}$ are the channel bandwidths of the CUT $i$ and the interfering channel $n$, respectively, $f_{ch,i}$ and $f_{ch,n}$ are the center frequencies of the CUT $i$ and the interfering channel $n$, respectively. The exponent $\in$ for the coherent accumulation is defined as

$$\in = \frac{3}{10} \ln\left(1 + \frac{3}{\alpha z_s^l \text{asinh}\left(\frac{\frac{\pi^2}{2}|\beta_2| \cdot B_{ch,i}^2}{2\alpha}\right)}\right) \tag{10}$$

TABLE I
PARAMETERS OF THE QoT ESTIMATION MODEL

| Symbols | Description |
|---|---|
| $V$ | The set of nodes |
| $E$ | The set of fiber links |
| $F$ | The set of frequency slots |
| $L_d$ | The set of link indices that belong to the LUT |
| $N$ | The total number of frequency slots |
| $M$ | The number of demanded frequency slots |
| $\tau$ | Service duration |
| $P_{ch}$ | Launch power |
| $P_{ASE}$ | Accumulated ASE noise power |
| $P_{NLI}$ | Accumulated NLI noise power |
| $FP$ | Equivalent penalty introduced by the filtering effects |
| $h$ | Planck's constant |
| $f$ | The central frequency of the channel |
| $NF$ | The noise figure of the EDFAs |
| $B_N$ | Noise bandwidth |
| $\alpha$ | Fiber attenuation coefficient |
| $z_s^l$ | Span length of the $s$-th span of link $l$ |
| $P_{SCI}^l$ | SCI noise power on link $l$ |
| $P_{XCI}^l$ | XCI noise power on link $l$ |
| $\gamma$ | The fiber nonlinearity coefficient |
| $z_{eff}$ | The effective length |
| $N_{ch}$ | Number of interfering channels |
| $G_{ch,n}$ | PSD of the interfering channel $n$ |
| $G_{ch,i}$ | PSD of the channel CUT $i$ |
| $\varphi_{i,i}$ | Coefficient accounts for SCI |
| $\varphi_{n,i}$ | Coefficient accounts for XCI |
| $\beta_2$ | The second order dispersion parameter |
| $N_s^l$ | The number of spans in link $l$ |
| $B_{ch}$ | Channel bandwidth |
| $\epsilon$ | The exponential coefficient |
| $f_{ch,n}$ | The center frequency of the interfering channel $n$ |
| $f_{ch,i}$ | The center frequency of the CUT $i$ |

All parameters of the QoT estimation are summarized in Table I.

## IV. PTRNET-RSA SCHEME

In this section, we introduce the overall architecture and the operation principle of the proposed PtrNet-RSA scheme in the RL framework, shown in Fig. 2. When a new traffic request arrives, the network state monitor observes the EON status to extract network state features. Specifically, a network feature matrix that consists of network topology, configuration, status, and service request information is obtained for each SWP. Then, the extracted feature matrices for all SWPs are sent to the RL agent as the input states and processed in parallel. The RL agent is constructed based on the actor-critic algorithm, consisting of two PtrNets, the actor and the critic. After receiving the feature matrices, the actor PtrNet is responsible for generating the desired lightpath $\pi$ for each SWP as the action, while the critic PtrNet model provides an unbiased estimation of the GSNR value for the lightpath generated by the actor PtrNet. Based on

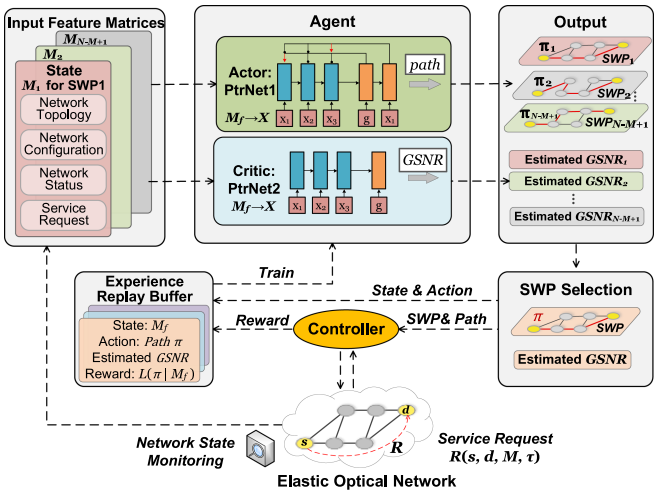



Fig. 2. Overall architecture and operation principle of the PtrNet-RSA model.

the generated GSNR estimations, the SWP selection module will give the policy to select the SWP. Once the SWP is selected, the path is also determined, and the decision is forwarded to the controller. After the controller sets up the lightpath, the reward module receives the outcome of the completed RSA operation as feedback and returns a reward for taking the RSA solution. Finally, the reward is stored together with the input feature matrix (state) and lightpath $\pi$ (action) in an experience replay buffer for the PtrNet training. The following subsection will give details of the PtrNet-RSA, including the input feature abstraction, the mechanism of PtrNet, the SWP selection policy, and the training process.

### A. Feature Matrix Design

Behaved as the input state in the RL framework, the feature matrix $M_f$ should contain enough information for QoT estimation and RSA in EON. Besides, to utilize the link-level network information, the feature matrix $M_f$ is formulated in node-base representation, which can be expressed as

$$M_f = [M_T,\ M_{SL},\ M_{SN},\ M_{ASE},\ M_{SCI}, M_{NC},\ M_{sd}] \tag{11}$$

where $M_T$ is the network topology matrix, $M_{SL}$ is the span length matrix, $M_{SN}$ is the span number matrix, $M_{ASE}$ is the ASE noise matrix, $M_{SCI}$ is the SCI noise matrix, $M_{NC}$ is the neighborhood channel states matrix, and $M_{sd}$ is the source-destination indicator matrix. For instance, the network topology matrix $M_T$ is given by

$$M_T = \begin{bmatrix} M^T_{1,1} & M^T_{1,2} & \dots & M^T_{1,|V|} \\ M^T_{2,1} & M^T_{2,2} & \dots & M^T_{2,|V|} \\ \vdots & \vdots & \ddots & \vdots \\ M^T_{|V|,1} & M^T_{|V|,2} & \dots & M^T_{|V|,|V|} \end{bmatrix}_{|V|\times|V|} \tag{12}$$

where each element $M^T_{i,j}$ is a binary parameter that is equal to 1 if nodes $i$ and $j$ are connected by an available fiber link within the current SWP; 0, otherwise. In $M_T$, row $i$ represents

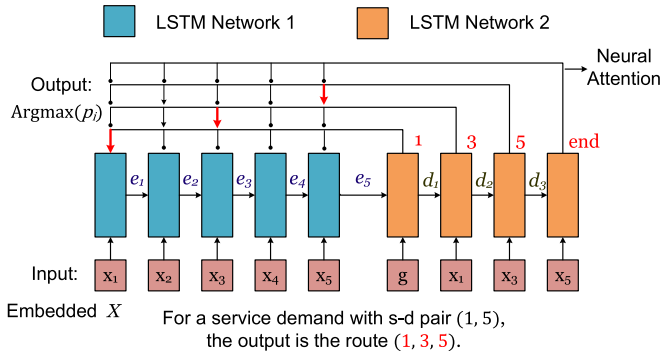



Fig. 3. The pointer network architecture and mechanism.

the connectivity of node $i$ and $|V|$ is the number of nodes in EON.

Similar to $M_T$, other matrices replace the element "1" in $M_T$ with the corresponding value, for instance, the span lengths for each link in $M_{SL}$, except $M_{sd}$, which is formulated with the one-hot encoding representation of the source and destination node indices. In this case, donating $V_i$ as row $i$ of feature matrix $M_f$, it carries the information pertaining to node $i$ in EON.

As Fig. 3 shows, given the service request, the feature matrix $M_f$ is extracted for each SWP in EON and serves as the input state for the RL agent. Besides, to speed up the training process, a synthetic feature matrix $M_f$can be manually generated for a single SWP rather than extracted from the EON environments. Specifically, in the synthetic feature matrix $M_f$, matrix $M_T$ can be generated by randomly blocking some links with a certain FS utilization ratio, matrices $M_{SL}$ and $M_{SN}$ are fixed for a certain network configuration, matrices $M_{ASE}$ and $M_{SCI}$ can be obtained based on the (3) and (5), matrix $M_{NC}$ is randomly generated based neighbourhood channels' FS utilization ratio. Synthetic feature matrices can be easily generated under different traffic loads by adjusting the FS utilization ratios.

### B. Operating Mechanism of PtrNet

Fig. 3 shows the architecture and mechanism of the PtrNet. The PtrNet consists of a neural attention module and two Long Short-Term Memory Network (LSTM) modules, where LSTM Network 1 is the encoder and LSTM Network 2 is the decoder [39]. The neural attention module enables a pointer that selects an input vector index as the output. Given an input feature matrix $M_f$ in node-based representation, the PtrNet sequentially reads the information of each node and gives the sequence of indices of the input nodes as the output.

The detailed procedures are introduced as follows. The feature matrix $M_f$ is initially transformed into an embedded abstract feature matrix $X = [x_1,\ x_2, \dots, x_N]$, where each vector $x_i$ is the result of a linear transformation of $V_i$. The encoder then proceeds to sequentially read and encode the embedded information $x_i$ into a sequence of latent memory states $e_i$. The output $e_i$ of each encoding step will also be used as the input of the next encoding step. The number of encoding steps is equivalent to the number of nodes in the corresponding SWP. Upon the completion of the encoding process, a vector $g$ is fed into the

decoder, acting as a trigger for starting the decoding steps. In each decoding step $i$, a latent memory state $d_i$ is attained by the decoder, and the neural attention selects an input node index $j$. The decoder sequentially reads latent memory states $d_i$ and embedded information $x_j$ of the last selected node $j$. For each decoding step $i$, the attention mechanism produces a probability distribution $u_i$ for the next node $j$ to select:

$$u_j^i = v^T \cdot \tanh\left(W_1 \cdot e_j + W_2 \cdot d_i\right)\ j \in (1, 2, \ldots N) \quad (13)$$

where $v$, $W_1$, and $W_2$ are the learnable parameters of the PtrNet.

To ensure the availability of the generated path, a dual-mask scheme is implemented as

$$p_i = \text{softmax}\left(u_i \cdot Mask1 \cdot Mask2\right) \quad (14)$$

where *Mask1* masks the unavailable nodes and *Mask2* disables the selection of past nodes. Then, the *SoftMax* function normalizes the probability distribution $p_i$ among the available nodes.

Based on the probability distribution $p_i$, PtrNet selects the following node index using *sampling* for training and the *argmax* function for testing. In the RL agent, the actor PtrNet follows the above mechanism and gives the sequence of indices of the input nodes as the output path. In this case, the action space includes all possible paths for a service request rather than being limited to fixed path candidates. The critic PtrNet outputs an estimated GSNR value after its decoder receives the trigger signal $g$.

An example of actor PtrNet is illustrated in Fig. 3, where the EON has five nodes and a service request with the $(s, d)$ pair (1, 5) arrives. First, the feature matrix is embedded into abstract feature matrix $X = [x_1,\ x_2, \ldots, x_5]$. The encoder (LSTM network1) reads from $x_1$ to $x_5$ sequentially, encoding it into latent memory states $e_1$ to $e_5$. The output of the last encoding step is also the input for the next encoding step; for instance, $e_1$ is also an input for encoding $x_2$. After the encoding of $x_5$, the output $e_5$ is fed to the decoder (LSTM network2) with the trigger vector $g$ to start the decoding. In the first decoding step, the decoder output $d_1$, and a probability distribution $p_i$ is produced by the neural attention and mask scheme for the five input nodes. Since node 1 is the source, other nodes are masked when selecting the first node. Thus, only node 1 is available and selected. In the second decoding step, both $d_1$ (the output from the first decoding step) and $x_1$ (The embedded vector for the last selected node 1) are fed into the decoder. Again, the decoder output $d_2$ and a probability distribution $p_2$ are produced. Based on $p_2$, node 3 is selected as the second node in the path. In the third decoding step, being fed with $d_2$ and $x_3$, decoder output $d_2$, and node 5 are selected. Since node 5 is the destination node, the decoding process terminates. Finally, the path (1, 3, 5) is attained through the cooperation of the encoder, decoder, and the attention mechanism in actor PtrNet. At the same time, critic PtrNet gives a GSNR estimation for this lightpath.

### C. SWP Selection

For the EON with $N$ FSs and a service request with $M$ FSs demands, a total number of $(N - M + 1)$ input feature matrices are fed into the agent, where the actor PtrNet generates the desired paths $\Pi = \{\pi_1, \pi_2, \ldots, \pi_{N-M+1}\}$, and the critic PtrNet provides the corresponding GSNR estimation $\text{GSNR} = \{\text{GSNR}_1, \text{GSNR}_2, \ldots, \text{GSNR}_{N-M+1}\}$. Given the lightpath and GSNR estimation within each SWP, a suitable SWP selection policy is required to attain the final allocation solution. Selecting an SWP with a smaller index may introduce more interferences from the neighborhood channels, causing GSNR degradation. However, selecting an SWP with a larger GSNR value may cause more spectrum fragmentation, increasing the blocking probability. In our previous work [18], we chose the SWP with the highest GSNR estimation value. In this paper, we investigate the impact of different SWP selection strategies in the trade-off between the network blocking probability and the QoT of the lightpath. For this, an adjusted GSNR value is used, which assigns a quantified priority to SWPs with smaller indices, expressed as

$$\text{GSNR}_k^{\text{adjusted}} = \text{GSNR}_k + (N - k)\,\varepsilon \quad (15)$$

where $\text{GSNR}_k$ is the estimated GSNR value for the path in $\text{SWP}_k$ from the critic PtrNet and $\varepsilon$ is the priority factor that indicates the priority assigned to the lower-indexed SWPs. For instance, when $\varepsilon$ is set to 0, PtrNet-RSA always selects the lightpath with the largest GSNR value. Conversely, an extremely large $\varepsilon$ implies a preference for selecting the lightpath in the SWP with the smallest index.

The SWP with the largest adjusted GSNR value is selected as the final solution and forwarded to the controller together with the corresponding path $\pi$. The controller then allocates the SWP and path $\pi$ for the traffic request. After that, a reward $L(\pi \mid M_f)$ for the completed RSA operation is returned, given by

$$L\left(\pi \mid M_f\right) = \begin{cases} \text{GSNR}_{real} & \text{if } \pi \text{ assignable} \\ 0 & \text{otherwise;} \end{cases} \quad (16)$$

where the $\text{GSNR}_{real}$ is the measured GSNR value of the established lightpath. In practice, the $\text{GSNR}_{real}$ should be obtained from an OSNR monitor [41], while in this paper, the $\text{GSNR}_{real}$ is calculated through the QoT mathematical model introduced in Section III-B with the input feature matrix $M_f$.

The reward is stored in an experience replay buffer with the state (input feature matrix $M_f$), the action (path $\pi$), and the GSNR estimation value of the selected SWP, donated as $b(M_f)$, for the PtrNet training.

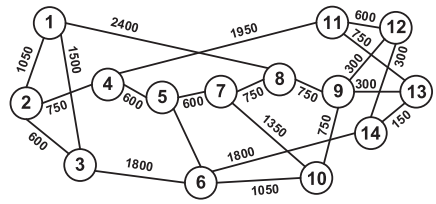
### D. Training

The parameters of the PtrNets are optimized using the actor-critic algorithm [40]. During the training process, the actor aims to minimize the difference between the measured GSNR value of the lightpath generated by the actor and the GSNR estimation provided by the critic. The critic tries to reduce the absolute value of the difference between the estimated GSNR and the actual GSNR. The loss functions of the actor PtrNet and critic PtrNet are defined as

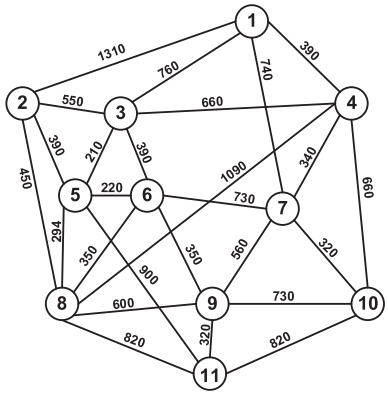
$$f_{\text{loss-actor}} = \left(-L\left(\pi \mid M_f\right) + b\left(M_f\right)\right) \times p_\pi \quad (17)$$

$$f_{\text{loss-critic}} = \left\| b\left(M_f\right) - L\left(\pi \mid M_f\right) \right\| \quad (18)$$

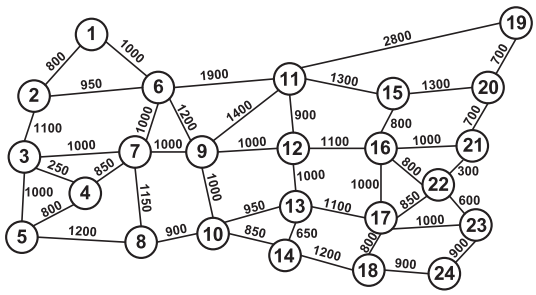
where $L(\pi \mid M_f)$ is the reward of allocating the lightpath $\pi$ for the input state $M_f$, $b(M_f)$ is the GSNR estimation from the




(a) NSF network




(b) COST-239 network




(c) USNET network

Fig. 4. Simulation networks topologies with fiber distance in kilometers. (a) NSF network (b) COST-239 network (c) USNET network.

critic model, and $p_\pi$ is the factorized probability of the selected path $\pi$.

Both actor PtrNet and critic PtrNet are trained with the samples in the experience replay buffer.

### E. Computational Complexity

For the PtrNet agent, the computational complexity of encoding steps is $O(|V|)$, where $|V|$ is the number of nodes in the optical network. The computational complexity of decoding steps with the attention module is $O(|V|^2)$. Thus, the overall computational complexity of PtrNet is $O(|V|^2)$, while the traditional RL-based RSA agents, such as DeepRMSA and MaskRSA agents, have the computational complexity of $O(KN)$, where $K$ is the number of path candidates, and $N$ is the number of FSs.

## V. Simulation Environment and Numerical Analysis

To verify the efficiency of the proposed scheme, we conducted simulations with (a) the 14-node, 22-link NSF network, (b) the 11-node, 26-link COST239 network, and (c) the 21-node, 43-link USNET network, as shown in Fig. 4. We utilized the mathematical model in Section III-B to establish the network RL simulation environment. We compared the performance of PtrNet-RSA with three benchmark schemes, i.e., Shortest Path First Fit (SP-FF), *K*-Shortest Path First Fit (KSP-FF) with K = 5, and MaskRSA proposed in [13]. We evaluated the efficiency of the proposed scheme in both single SWP and dynamic traffic scenarios. Other simulation setup, training process, and validation results are presented in the following subsections.

### A. Simulation Setup

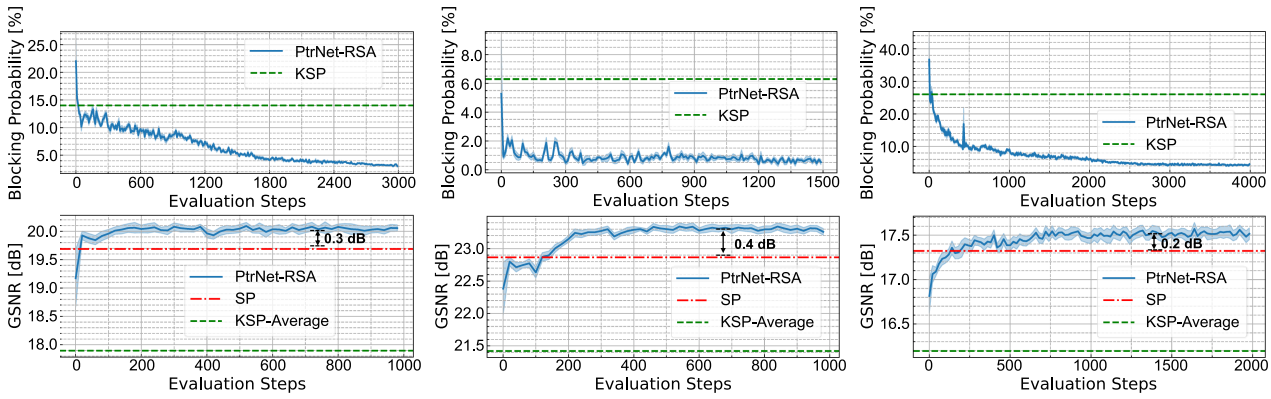
For the EON configuration, the frequency slot width is set to 12.5 GHz with a fixed channel spacing of 17.5 GHz on each link. All transmitted channels have a fixed launch power $P$ of $-2$ dBm and a symbol rate of 12.5 GBaud. For the XCI, we considered four adjacent channels, and $N_{ch}$ was set to be 4. The span length was randomly generated from the set of [40, 60, 80] km, and the number of spans in a link could be calculated with the span length and the link length. The equivalent penalty $FP$, the fiber attenuation, and the second-order dispersion were set to be 1, 0.2 dB/km, and $-21.6$ ps$^2$/km, respectively. The noise figure $NF$ was assumed to be 5 dB for all EDFAs and the fiber nonlinearity coefficient was set to 1.11/W/km.

We set the embedding size and hidden size of the pointer network to 1024. The initial learning rate was set to $10^{-4}$ and decayed by a factor of 0.8 every 500 steps. The batch size was set to 1024, and *Adam* was selected as the optimizer.

For the validation of the single SWP case, we tested the PtrNet-RSA model with synthetic feature matrices. In the dynamic traffic scenario, we verified the PtrNet-RSA model using the input feature matrix observed from the EON environment. Two dynamic traffic cases were considered: (1) forty C-band channels and service requests with only one FS demand; (2) eighty C-band channels and service requests with one to four FS demands. The C-band channels are anchored to 194 THz. The modulation format polarization-multiplexing QPSK (PM-QPSK) is adopted for case (1), and PM-BPSK is adopted for case (2). The dynamic service requests were generated according to a Poisson process following a uniform traffic distribution with the average arrival rate $\lambda$. The service duration time followed a negative exponential distribution with the average service duration time $1/\mu$. Therefore, the network load was expressed as $\lambda/\mu$. All parameters in the simulation are summarized in Table II.

### B. Training Performance

To speed up the training process, the synthetic feature matrices are employed for training the PtrNets. Fig. 5 illustrates the evaluated blocking probability and GSNR after a certain number of training steps, where the model was trained with a batch size of 1024 samples per step and was evaluated with 10000 testing samples. The evaluation of blocking probability and GSNR was performed on a single SWP with an average traffic load of a 20% FS utilization ratio in the channel under test and neighborhood channels, respectively. Fig. 5(a) shows the average blocking probability and GSNR performance after certain training steps in the NSF network. We can see that PtrNet-RSA could outperform SP in the average GSNR of the




(a) NSF network (b) COST-239 network (c) USNET

Fig. 5. Performance comparison of evaluated GSNR [dB] versus training steps: (a) NSF network, (b) COST-239 network, (c) USNET.

TABLE II
SIMULATION PARAMETERS

| Parameters | Value |
|---|---|
| $N$ | [40, 60, 80] |
| $M$ | [1, 2, 3, 4] slots |
| $\lambda$ | [1, 2, 3, …, 45] |
| $\mu$ | 0.1 |
| $P_{ch}$ | -2 dBm |
| $FP$ | 1 |
| $h$ | $6.6 \times 10^{-34}$ J· s |
| $NF$ | 5 dB |
| $\alpha$ | 0.2 dB/km |
| $z_s^l$ | [40, 60, 80] km |
| $\gamma$ | 1.1 1/W/km |
| $N_{ch}$ | 4 |
| $\beta_2$ | -21.6 $ps^2$/km |
| $B_N$ | [12.5, 25, 37.5, 50] GHz |
| $B_{ch}$ | [12.5, 25, 37.5, 50] GHz |

generated path with a gain of up to 0.3 dB after about 150 training steps. For the blocking probability, PtrNet-RSA outperformed KSP within 100 training steps and converged after 2800 training steps. In contrast, Fig. 5(b) depicts the training results for the COST239 network, PtrNet-RSA yielded a larger improvement in GSNR value (0.4 dB) compared to NSF. This difference can be attributed to the fact that the COST239 topology has a higher average node degree, providing more alternative paths for each source-destination pair and enabling more flexible routing. For the blocking probability, it required around 1200 steps to converge to a stable state. In Fig. 5(c), a similar trend is observed for training in the USNET network, where the proposed PtrNet-RSA model demonstrated a maximum improvement of up to 0.2 dB compared to the benchmark. In the USNET network, PtrNet required approximately 3000 training steps to converge, primarily due to the larger number of links and nodes of the USNET topology.

### C. Case With Single SWP

The well-trained PtrNet-RSA model was first applied to the single SWP scenario to test the performance. For the validation of the blocking performance, we changed the FS utilization ratio for the SWP under test and the neighboring frequency channels. For fairness, when evaluating the QoT performance, we set the SWP under test to be empty and only changed the traffic load in the neighboring channels.

Fig. 6 shows the results of the blocking probability under different traffic loads in the three network topologies, where the blocking was computed only considering spectrum unavailability. For the NSF network, PtrNet-RSA achieved a significant reduction of up to 14.3% in the blocking probability compared to KSP. This is because PtrNet-RSA generated paths node by node without being limited to pre-calculated candidates. However, as the traffic further increased, the reduction in blocking probability owing to PtrNet-RSA diminished, due to the gradual reduction of the available alternative paths. In the COST239 network and the USNET network, PtrNet-RSA achieved larger reductions in the blocking probability of 34% and 27% with traffic loads of 0.5 and 0.3, respectively, compared to KSP. This is because these two networks have higher average node degrees, and more alternative paths could be used between each source-destination pair in PtrNet-RSA.

In Fig. 7, the GSNR values for the established lightpaths under different traffic loads are depicted. The proposed PtrNet-RSA scheme achieved up to 0.3-dB, 0.5-dB, and 0.25-dB improvements in GSNR compared to the SP approach for the NSF, the COST239, and the USNET networks, respectively. It effectively mitigated the impact of the crosstalk interferences across various traffic loads and achieved a significant improvement in the QoT of the lightpaths. We could see that the GSNR improvements gradually decreased with the increase in the traffic load. This is mainly because there were more crosstalk interferences in the network with higher loads. Furthermore, the improvement in the USNET topology was relatively smaller than that in the other two topologies due to its longer average link length.

### D. Case With Dynamic Traffic

We also conducted simulations involving the dynamic traffic scenario with different traffic loads, where the average service duration time was fixed at 10 and different arrival rates

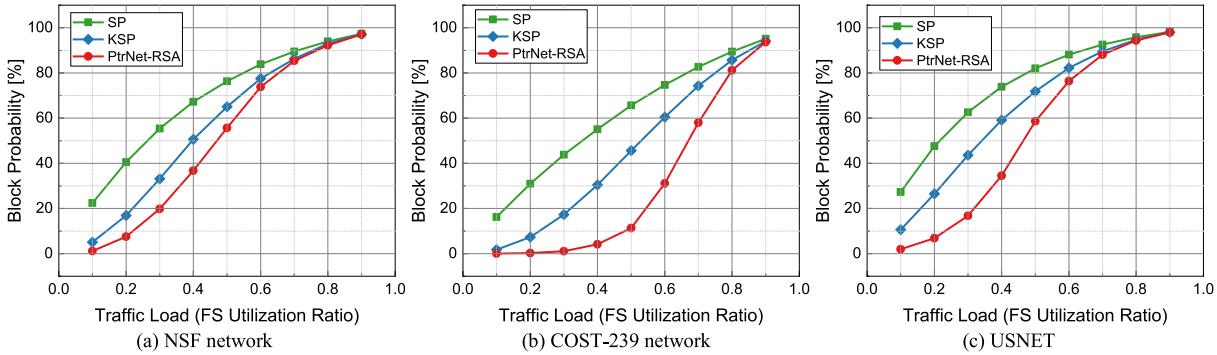



Fig. 6. Blocking probability versus single SWP traffic load for (a) NSF, (b) COST-239, (c) USNET.

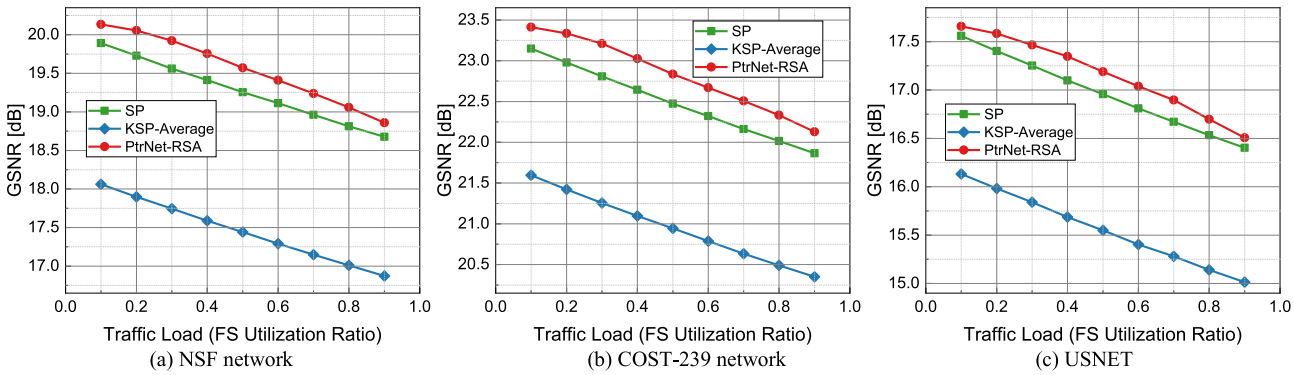



Fig. 7. GSNR [dB] versus single SWP traffic load for (a) NSF, (b) COST239, (c) USNET.

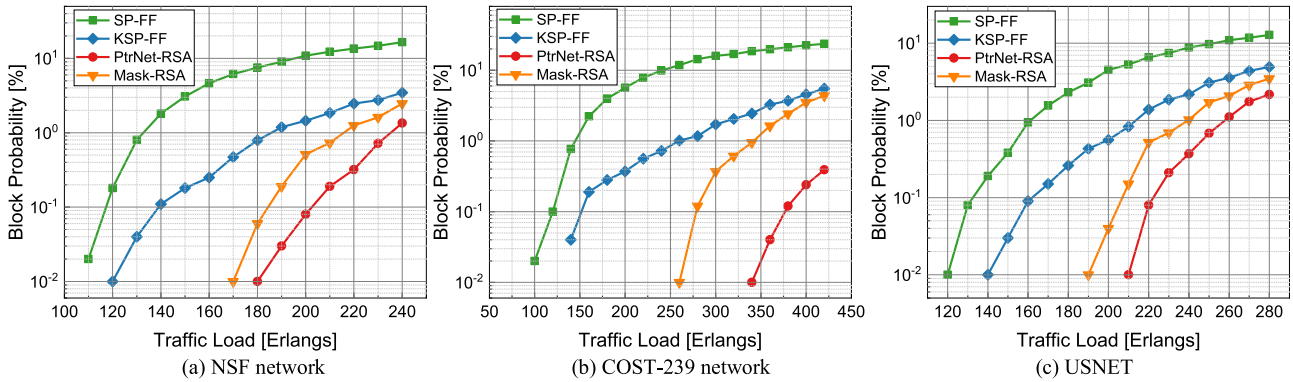



Fig. 8. Blocking probability versus dynamic traffic load (40-channel case) for (a) NSF, (b) COST-239, (c) USNET.

were tested. A total of 10000 requests were evaluated for each traffic load. To ensure a fair comparison, the average blocking probabilities and GSNR values were first computed for each source-destination pair, and then the results of all source-destination pairs were averaged as the final results. Fig. 8 shows the blocking probability for the 40-channel case, where the priority factor $\varepsilon$ was set to 0 in PtrNet-RSA. As depicted in Fig. 8, the proposed PtrNet-RSA significantly reduced the blocking probability compared to benchmarks, including SP-FF, KSP-FF, and MaskRSA. It achieved blocking probability reductions of up to 15.2%, 23.47%, and 11.43% compared to KSP-FF in the NSF, the COST239, and the USNET network topologies, respectively. Although MaskRSA also performed well, PtrNet still outperformed it with up to two orders of magnitude reduction in the blocking probability among the three network topologies. PtrNet-RSA enabled a node-by-node approach to find the desired path and facilitated the selection of lightpaths from all possible paths, resulting in a more significant reduction in the blocking probability compared with MaskRSA. The results also demonstrated that under the same network load, PtrNet-RSA achieved a more substantial improvement in the COST239 network, consistent with the previous training and single SWP validation results.

Fig. 9 shows the GSNR performance for the 40-channel case. The proposed PtrNet-RSA scheme achieved significant improvements compared to the benchmark methods, by up to

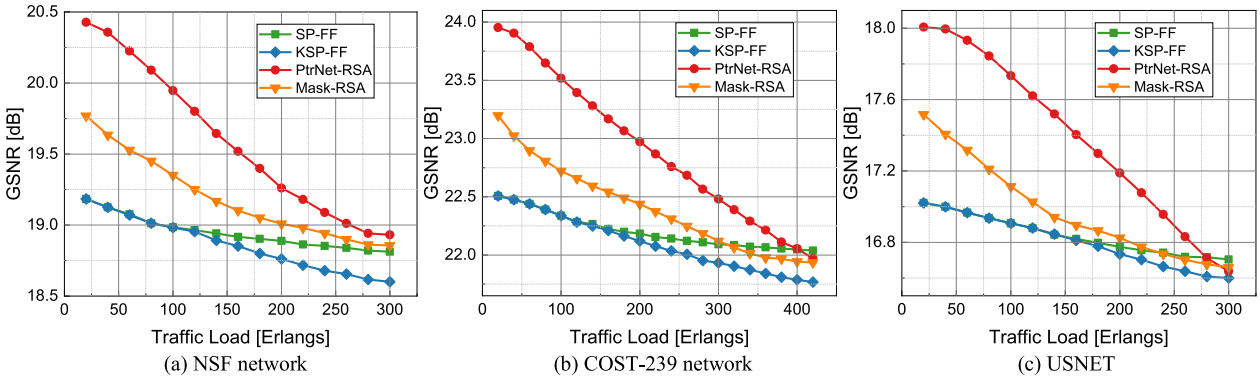



Fig. 9. GSNR [dB] versus dynamic traffic load (40-channel case) for (a) NSF, (b) COST-239, (c) USNET.

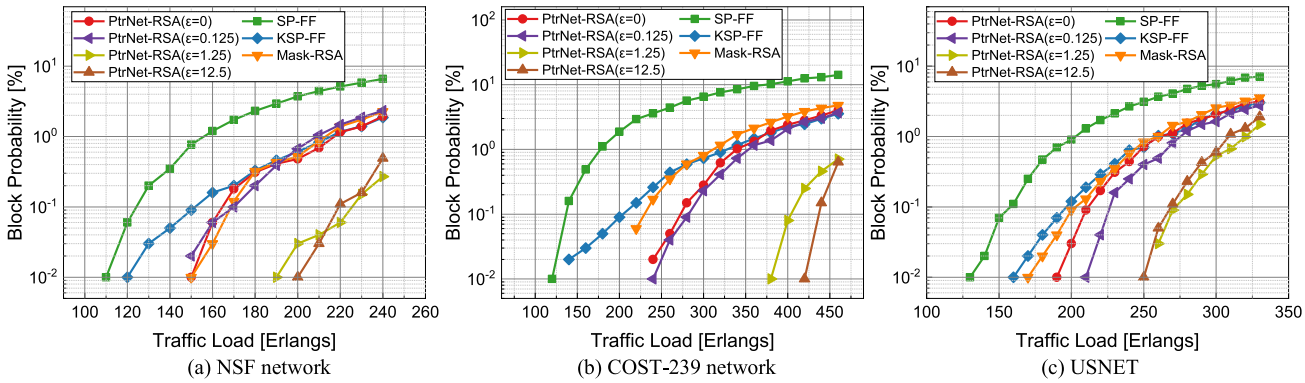



Fig. 10. Blocking probability versus dynamic traffic load (80-channel case) for (a) NSF, (b) COST-239, (c) USNET.

1.2 dB, 1.4 dB, and 1.0 dB compared with SP-FF and 0.7 dB, 0.9 dB, and 0.6 dB compared with Mask-RSA, in the NSF, the COST239, and the USNET network topologies, respectively. This is because PtrNet-RSA integrated the QoT estimation inside the RSA process and considered the GSNR of the lightpath as the optimization objective, reducing the XCI of the selected lightpaths. Specifically, for low traffic load cases, PtrNet-RSA attained a high GSNR value similar to the results observed at 0.1 traffic load in the single SWP verification. This revealed that PtrNet-RSA avoided selecting adjacent SWPs when the traffic load was small. In contrast, SP-FF and KSP-FF tended to select SWPs with lower indices first, causing most of the lower-indexed adjacent channels of the newly established lightpath to be already occupied. Consequently, SP-FF and KSP-FF attained low lightpath GSNR values even when the traffic load was small. Mask-RSA, although performing better than SP-FF and KSP-FF, did not optimize the GSNR performance as effectively as PtrNet-RSA. As the traffic load increased, both the GSNR results of PtrNet-RSA and Mask-RSA decreased due to the increasing XCI from neighborhood channels. At high traffic loads, due to its low blocking probability performance, PtrNet-RSA could allocate lightpaths for the requests that the benchmarks failed to serve. However, these lightpaths generally had longer lengths with lower GSNR values since many links were unavailable, leading to degradation in the GSNR performance of PtrNet-RSA. This is why PtrNet-RSA's GSNR performance was even worse than the SP-FF algorithm in the COST239 network at a traffic load of 420, depicted in Fig. 9(b).

Fig. 10 presents the blocking probability for the 80-channel case, where the traffic requests arrived with FS demand ranging from 1 to 4. For PtrNet-RSA, different SWP selection strategies were investigated with priority factor $\varepsilon$ values of 0, 0.125, 1.25, and 12.5. As shown in Fig. 10, PtrNet-RSA with a zero value in $\varepsilon$ selects SWPs with the largest GSNR value, although it still outperformed the benchmarks with small traffic loads. The performance degraded as the traffic load increased. This is because selecting SWPs with the largest GSNR value caused more spectrum fragmentation, and larger FS demands made lightpath provisioning more sensitive to spectrum fragmentation. When the value of $\varepsilon$ was 0.125, there was no significant improvement in the blocking performance compared with that when the value of $\varepsilon$ was 0. However, as $\varepsilon$ increased to 1.25 and 12.5, a substantial improvement in blocking probability was achieved for all three network topologies. This improvement was attributed to the PtrNet-RSA which selected SWPs with relatively smaller GSNR values but lower indices, thus reducing possible spectrum fragmentation. However, PtrNet-RSA with a larger priority factor might select adjacent SWPs, which could cause a degradation in GSNR performance due to the increase in XCI. Fig. 11 shows the GSNR performance comparison for the 80-channel case for PtrNet-RSA with different SWP selection policies and benchmark algorithms. As shown in Fig. 11, PtrNet-RSA with

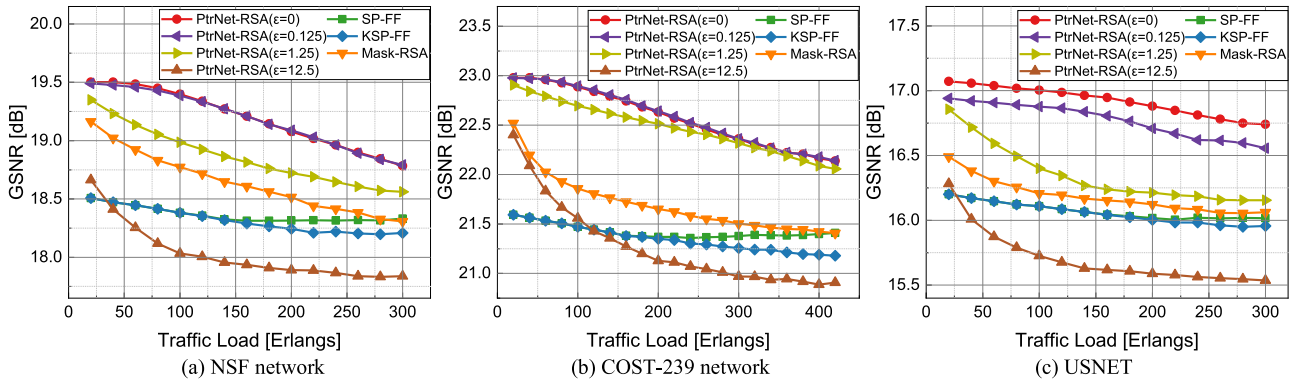



Fig. 11. GSNR [dB] versus dynamic traffic load (80-channel case) for (a) NSF, (b) COST-239, (c) USNET.

zero-valued $\varepsilon$ attained the largest improvement compared with the benchmarks. Similar to the blocking performance, there was no obvious degradation in GSNR for PtrNet-RSA when the value of $\varepsilon$ was 0.125. As $\varepsilon$ increased to 1.25, PtrNet-RSA encountered a slight performance degradation while still outperformed the benchmarks. However, when the value of $\varepsilon$ became 12.5, the GSNR performance of PtrNet-RSA suffered a serious decline and was even worse than all benchmarks. This was attributed to the fact that the larger the priority factor, the higher the spectrum utilization in the adjacent SWPs of the selected SWP, thus leading to a decrease in the GSNR performance.

From Figs. 10 and 11, we can notice there was a trade-off between the network blocking probability and the QoT of the lightpath. A small priority factor significantly improved the QoT of the lightpath but led to a higher blocking rate (e.g., $\varepsilon = 0$ or 0.125). Conversely, a large priority factor could reduce the blocking rate but lead to degradation in QoT (e.g., $\varepsilon = 12.5$). Setting $\varepsilon$ to 1.25 might be a better choice, achieving a low blocking probability and a high QoT performance, simultaneously, for different network topologies.

## VI. Summary

In this paper, we have addressed the QoT-aware RSA problem in EONs by introducing the PtrNet-RSA scheme, which leverages the pointer network and the reinforcement learning. PtrNet-RSA offers an efficient solution to generate the desired lightpath with good QoT and low blocking probability, eliminating the need for pre-calculated paths. Simulation results have demonstrated the effectiveness of PtrNet-RSA with up to two orders of magnitude reduction in the blocking probability and notable improvements in GSNR (up to 1.6 dB, 1.8 dB, and 1.0 dB) in the NSF, the COST239, and the USNET topologies when compared to the baseline approaches. We have further investigated the SWP selection strategies and analyzed the trade-off between the network blocking probability and the QoT of the lightpath, revealing the existence of a suitable priority factor for reducing the blocking probability while maintaining a relatively high GSNR value. Potential future work may include optimization of the modulation formats and the application of transfer learning across different network topologies, aiming to reduce the volume of the required training data.

## References

[1] O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo, "Elastic optical networking: A new dawn for the optical layer?," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s12–s20, Feb. 2012.

[2] Y. Pointurier, "Design of low-margin optical networks," *J. Opt. Commun. Netw.*, vol. 9, pp. A9–A17, 2017.

[3] J. Zhao, H. Wymeersch, and E. Agrell, "Nonlinear impairment-aware static resource allocation in elastic optical networks," *J. Lightw. Technol.*, vol. 33, no. 22, pp. 4554–4564, Nov. 2015.

[4] J. Shao, X. Liang, and S. Kumar, "Comparison of split-step Fourier schemes for simulating fiber optic communication systems," *IEEE Photon. J.*, vol. 6, no. 4, Aug. 2014, Art. no. 7200515.

[5] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[6] M. Ibrahimi et al., "Machine learning regression for QoT estimation of unestablished lightpaths," *J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B92–B101, Apr. 2021.

[7] A. Mahajan, K. Christodoulopoulos, R. Martinez, S. Spadaro, and R. Munoz, "Machine learning assisted EDFA gain ripple modelling for accurate QoT estimation," in *Proc. IEEE 45th Eur. Conf. Opt. Commun.*, 2019, pp. 1–4.

[8] S. Aladin, A. V. S. Tran, S. Allogba, and C. Tremblay, "Quality of transmission estimation and short-term performance forecast of lightpaths," *J. Lightw. Technol.*, vol. 38, no. 10, pp. 2807–2814, May 2020.

[9] Q. Zhuge et al., "Application of machine learning in fiber nonlinearity modeling and monitoring for elastic optical networks," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3055–3063, Jul. 2019.

[10] P. Safari, B. Shariati, G. Bergk, and J. K. Fischer, "Deep convolutional neural network for network-wide QoT estimation," in *Proc. IEEE Opt. Fiber Commun. Conf. Exhib.*, 2021, pp. 1–3.

[11] J. Müller, S. K. Patri, T. Fehenberger, C. Mas-Machuca, H. Griesser, and J. P. Elbers, "A QoT estimation method using EGN-assisted machine learning for network planning applications," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2021, pp. 1–4.

[12] X. Chen, B. Li, R. Proietti, H. Lu, Z. Zhu, and S. J. B. Yoo, "DeepRMSA: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks," *J. Lightw. Technol.*, vol. 37, no. 16, pp. 4155–4163, Aug. 2019.

[13] M. Shimoda and T. Tanaka, "Mask RSA: End-to-end reinforcement learning-based routing and spectrum assignment in elastic optical networks," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2021, pp. 1–4.

[14] M. Salani, C. Rottondi, L. Ceré, and M. Tornatore, "Dual-Stage planning for elastic optical networks integrating machine-learning-assisted QoT estimation," *IEEE/ACM Trans. Netw.*, vol. 31, no. 3, pp. 1293–1307, Jun. 2023.

[15] H. Beyranvand and J. A. Salehi, "A quality-of-transmission aware dynamic routing and spectrum assignment scheme for future elastic optical networks," *J. Lightw. Technol.*, vol. 31, no. 18, pp. 3043–3054, Sep. 2013.

[16] T. Panayiotou, G. Savva, I. Tomkos, and G. Ellinas, "Decentralizing machine-learning-based QoT estimation for sliceable optical networks," *J. Opt. Commun. Netw.*, vol. 12, no. 7, pp. 146–162, Jul. 2020.

[17] T. Panayiotou, S. P. Chatzis, and G. Ellinas, "Performance analysis of a data-driven quality-of-transmission decision approach on a dynamic multicast-capable metro optical network," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 98–108, Jan. 2017.
[18] Y. Cheng, S. Ding, and C.-K. Chan, "Leveraging pointer network for QoT-aware routing and spectrum assignment in elastic optical networks," in *Proc. IEEE Eur. Conf. Opt. Commun.*, 2022, pp. 1–4.
[19] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61–68, Jan. 2003.
[20] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," *J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A286–A297, Feb. 2018.
[21] T. Jimenez et al., "A cognitive quality of transmission estimator for core optical networks," *J. Lightw. Technol.*, vol. 31, no. 6, pp. 942–951, Mar. 2013.
[22] T. Panayiotou, S. P. Chatzis, and G. Ellinas, "Performance analysis of a data-driven quality-of-transmission decision approach on a dynamic multicast-capable metro optical network," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. 98–108, Jan. 2017.
[23] Y. He et al., "Improved QoT estimations through refined signal power measurements and data-driven parameter optimizations in a disaggregated and partially loaded live production network," *J. Opt. Commun. Netw.*, vol. 15, no. 9, pp. 638–648, Sep. 2023.
[24] I. Martín et al., "Machine learning-based routing and wavelength assignment in software-defined optical networks," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 3, pp. 871–883, Sep. 2019.
[25] L. Xu, Y. C. Huang, Y. Xue, and X. Hu, "Deep reinforcement learning-based routing and spectrum assignment of eons by exploiting GCN and RNN for feature extraction," *J. Lightw. Technol.*, vol. 40, no. 15, pp. 4945–4955, Aug. 2022.
[26] Q. Yao et al., "SNR re-verification-based routing, band, modulation, and spectrum assignment in hybrid C-C+L optical networks," *J. Lightw. Technol.*, vol. 40, no. 11, pp. 3456–3469, Jun. 2022.
[27] Z. Luo et al., "Survivable routing, spectrum, core and band assignment in multi-band space division multiplexing elastic optical networks," *J. Lightw. Technol.*, vol. 40, no. 11, pp. 3442–3455, Jun. 2022.
[28] Y. Zou et al., "Nonlinear impairment-aware RMSA under the sliding scheduled traffic model for eons based on deep reinforcement learning," *J. Lightw. Technol.*, vol. 41, no. 22, pp. 6854–6864, Nov. 2023.
[29] A. Cai, G. Shen, L. Peng, and M. Zukerman, "Novel node-arc model and multiiteration heuristics for static routing and spectrum assignment in elastic optical networks," *J. Lightw. Technol.*, vol. 31, no. 21, pp. 3402–3413, Nov. 2013.
[30] Y. Li, H. Dai, G. Shen, and S. K. Bose, "Adaptive FEC selection for lightpaths in elastic optical networks," in *Proc. IEEE Opt. Fiber Commun. Conf. Exhib.*, 2014, pp. 1–3.
[31] C. Wang, G. Shen, and S. K. Bose, "Distance adaptive dynamic routing and spectrum allocation in elastic optical networks with shared backup path protection," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 2955–2964, Jul. 2015.
[32] V. Curri, "GNPy model of the physical layer for open and disaggregated optical networking," *J. Opt. Commun. Netw.*, vol. 14, no. 6, pp. C92–C104, Jun. 2022.
[33] M. Filer, M. Cantono, A. Ferrari, G. Grammel, G. Galimberti, and V. Curri, "Multi-vendor experimental validation of an open source QoT estimator for optical networks," *J. Lightw. Technol.*, vol. 36, no. 15, pp. 3073–3082, Aug. 2018.
[34] A. Ferrari et al., "GNPy: An open source application for physical layer aware open optical networks," *J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. C31–C40, Jun. 2020.
[35] T. Rahman et al., "On the mitigation of optical filtering penalties originating from ROADM cascade," *IEEE Photon. Technol. Lett.*, vol. 26, no. 2, pp. 154–157, Jan. 2014.
[36] G. Bosco, P. Poggiolini, A. Carena, V. Curri, and F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection," *Opt. Exp.*, vol. 19, pp. B440–B451, 2011.
[37] D. Semrau, R. I. Killey, and P. Bayvel, "The Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightw. Technol.*, vol. 36, no. 14, pp. 3046–3055, Jul. 2018.
[38] P. Poggiolini, "The GN Model of Non-Linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.
[39] O. Vinyals, M. Fortunato, and N. Jaitly, "Pointer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1–9.
[40] I. Bello, H. Pham, Q. V. Le, M. Norouzi, and S. Bengio, "Neural combinatorial optimization with reinforcement learning," 2016, *arXiv preprint*.
[41] R. Proietti et al., "Experimental demonstration of machine-learning-aided qot estimation in multi-domain elastic optical networks with alien wavelengths," *J. Opt. Commun. Netw.*, vol. 11, no. 1, pp. A1–A10, Jan. 2019.

**Yuansen Cheng** received the B.Eng. degree in electronic information engineering from The Chinese University of Hong Kong, Shenzhen, China, in 2019. He is currently working toward the Ph.D. degree in information engineering with The Chinese University of Hong Kong, Hong Kong SAR, China. His research interests include machine learning and optical networking.

**Shifeng Ding** received the M.Eng. degree in information and communication engineering from Soochow University, Suzhou, China, in 2019. He is currently working toward the Ph.D. degree in information engineering with The Chinese University of Hong Kong, Hong Kong SAR, China. His research interests include quantum key distribution and optical networking.

**Yingjie Shao** received the B.Eng. degree from Shandong University, Jinan, China, in 2017, and the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, in 2021. She is currently a Researcher with the Centre for Applied Photonics, Fraunhofer U.K., and also a Visiting Researcher with the University of Strathclyde, Glasgow, U.K. Her research interests include digital signal processing for optical communication and free space optical communications.

**Calvin Chun-Kit Chan** (Senior Member, IEEE) received the B.Eng., M.Phil., and Ph.D. degrees in information engineering from The Chinese University of Hong Kong, Hong Kong SAR, China. After graduation, he joined the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, as a Research Assistant Professor. Later, he joined the Bell Laboratories, Photonic Networks Department, Lucent Technologies, Crawford Hill Holmdel, NJ, USA, as a Member of Technical Staff. He was a Senior Optical System Engineer with Jedai Broadband Networks, Inc., Red Bank, NJ, before he joined the Department of Information Engineering, The Chinese University of Hong Kong. He is currently a Professor. He has authored or coauthored more than 320 technical papers in refereed international journals and conferences. He holds two issued U.S. patents, one edited book, and two book chapters. His main research interests include WDM optical metro or access networks, high-speed optical and digital signal processing, optical performance monitoring, and optical network planning and optimization. He was the Co-Chair and TPC Member of several international conferences, including OFC, OECC, CLEO/Pacific Rim, and ACP. He was an Associate Editor for the *OSA Journal of Optical Networking*, IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING, and *Frontiers in Optoelectronics*. He is a Fellow of Optica.