Traffic-grooming- and multipath-routing-enabled impairment-aware elastic optical networks

(article starts on next page)

Traffic-grooming- and multipath-routing-enabled impairment-aware elastic optical networks

Nishan Dharmaweera, Juzi Zhao, Li Yan, Magnus Karlsson, and Erik Agrell

Abstract—**Traffic grooming and multipath routing are two techniques that are widely adopted to increase the performance of traditional wavelength division multiplexed networks. They have been recently applied in elastic optical networks to increase spectral efficiency. In this study, we investigate the potential gains by jointly employing the two techniques in combination with a realistic physical impairment model. To allocate resources and quantify spectral efficiency gains over existing impairment-aware schemes, we present an analytical optimization formulation for small networks and a heuristic for large networks. Through numerical simulations, we demonstrate that traffic grooming and multipath routing, together, increase spectral efficiency and reduce resource consumption over existing schemes. We show that the proposed scheme offers significant performance improvements in networks with low degrees of connectivity, high traffic loads, and long links.**

I. Introduction

The exponential growth of Internet traffic and emerging bandwidth-intensive applications have driven the demand for more advanced, high-capacity optical fiber transmission technologies. Recent advancements in optical technologies have enabled transmission rates in excess of 400 Gb/s. To increase utilization of this available high bandwidth, next-generation optical networks are expected to evolve into more transmission-rate-flexible, reconfigurable, and capacity-efficient elastic optical networks (EONs).

In an EON, a traffic connection is accommodated by assigning a set of consecutive spectrum slots (i.e., a band) with "just enough" bandwidth [1]–[3]. As spectrum slots are spaced at a finer granularity (e.g., 6.25 GHz) and combined to accommodate connections of different bandwidths, EONs significantly reduce bandwidth waste vis-à-vis traditional wavelength division multiplexed (WDM) networks. Due to these differences between WDM networks and EONs, traditional routing and wavelength assignment schemes developed for the former are not applicable to the latter [1]. Instead, EONs require more sophisticated routing, modulation, and spectrum allocation (RSA) schemes.

Over the last decade, a number of RSA schemes were developed by researchers to achieve different objectives. These were often tested in networks with static (i.e., offline) or dynamic

N. Dharmaweera and M. Karlsson are with the Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: nishan@chalmers.se).

J. Zhao, L. Yan, and E. Agrell are with the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden.

(i.e., online) traffic. While increasing spectral efficiency is the primary goal of most RSA schemes, reducing resource usage, power consumption, and bandwidth blocking ratio (BBR) were also considered as alternative objectives [4]–[6]. The performance of every RSA scheme strongly depends on 1) the selected routing scheme, 2) modulation format allocation, and 3) spectrum allocation. Spectrum continuity and spectrum contiguity are two critical constraints of every RSA algorithm.

In an EON, physical-layer impairments (PLI) hinder network performance. Specifically, both linear and non-linear impairments arise when traffic connections are co-propagated with different modulation formats and bit rates.

A. Increasing spectral efficiency

To increase spectral efficiency in an EON, it is vital to reduce spectrum fragmentation. Fragmentation can be minimized either by using complex spectrum reshuffling algorithms [7] or by splitting traffic connections into smaller connections and routing them via multiple paths [8]. Using the latter *multipath routing* (MPR) technique, researchers demonstrated that spectral efficiency can be increased and BBR can be reduced in EONs [8]–[12]. For example, by adopting MPR, the authors in [8] increased spectral efficiency by 50% and reduced BBR by 50%.

In addition to increasing spectral efficiency, MPR helps to minimize network cost. As highlighted in [8], subconnections generally require a narrower spectrum width and can, therefore, utilize spectral fragments more efficiently. As a result, MPR is considered a simpler and more cost-effective solution than existing network-wide spectrum reshuffling schemes [8]. Furthermore, by being able to split traffic connections and route them across multiple paths (i.e., MPR), the network inherently becomes resilient to link failures. Because the network then relies less on dedicated protection schemes that require an overbuild of backup resources, network cost is lower. Moreover, when free spectrum is limited along a single path, instead of assigning a higher modulation format to the connection that has a smaller unregenerated reach, the connection can be split into smaller subconnections with a longer unregenerated reach and assigned lower modulation formats. This method would reduce both network cost and power consumption. However, the drawback of MPR is that splitting increases the number of guardbands (GBs), transponders, and spectrum waste due to large-granularity subconnections [10]– [12]. In order to overcome this limitation, *traffic grooming* (TG) is used in conjunction with MPR [4].

To increase bandwidth efficiency and resource utilization, TG has been extensively used in traditional WDM networks

[6]. By aggregating multiple low-speed traffic connections into a single high-speed stream, TG essentially reduces the number of wavelengths (i.e., bandwidth) and devices needed to accommodate traffic in a WDM network. As shown in [13], traffic grooming increases the utilization of network devices and, thus, reduces network cost and power consumption. Specifically the multi-hop with bypass grooming scheme in [13] reduces the number of router ports and used bandwidth in optical fibers, thus minimizing fiber and router-port cost and power consumption. However, excessive traffic grooming would increase optical-to-electrical-to-optical (O-E-O) conversions and traffic processing at the intermediate core routers. In recent studies, TG has also been successfully applied to EONs to increase spectral efficiency, thereby reducing BBR and power consumption [3], [5], [6]. In an EON, TG increases spectral efficiency by allowing multiple traffic connections to be assigned consecutive spectrum slots without interim GBs. Moreover, TG increases the utilization of bandwidth variable transponders (BVTs) and multi-flow transponders (MFTs) in advanced EONs [3], [5].

For the first time, MPR and TG were jointly employed in [4] to increase throughput and reduce BBR. In that study and other attempts to increase spectral efficiency, the quality of transmission (QoT) degradation encountered due to PLI were resolved by adding a fixed number of GBs (i.e., an integer number of spectrum slots) between co-propagating bands of different connections. Yet, a fixed GB is not related to the actual network state and may, therefore, overestimate or underestimate the PLI [14]. As such, existing impairment-aware, distance-adaptive RSA schemes [15], [16] that utilize a fixed GB are unlikely to provide the optimum transmission reach for different modulation schemes. On the other hand, transmission-reach (TR) based RSA schemes in the literature [1], [11], [16], [17] do not characterize both linear and nonlinear interference. Thus, TR-based RSA schemes too fail to accurately account for PLI. To overcome these limitations, the authors in [14] presented a state-of-the-art, realistic, impairment-aware (RIA) RSA scheme that uses an accurate, yet tractable, PLI model [18]. That scheme offers spectral efficiency improvements over existing TR-based RSA schemes that use a fixed GB [14].

B. Our contribution

In this paper, we propose a TG- and MPR-enabled RSA scheme to increase spectral efficiency over that offered in [14]. Our work is important because it demonstrates that spectral efficiency improvements can be achieved in an EON by collectively adopting TG and MPR techniques in combination with a realistic PLI model. In addition to spectral efficiency improvements, we also assess two important performance metrics, namely, port consumption and transponder consumption. We demonstrate to network operators the applicability and superior performance of our proposed RSA scheme under different network sizes by assessing its performance against varying link lengths and degrees of network connectivity.

The remainder of this paper is organized as follows. In Section II, we describe the architecture of an EON and

introduce the PLI model. TG and MPR are then introduced using an illustrative example. In Section III, we formulate the RSA problem of a TG- and MPR-enabled impairment-aware EON using integer linear programming (ILP) and heuristics. In Section IV, we provide numerical results to demonstrate the superior performance of our proposed scheme in comparison to existing schemes. Concluding remarks are provided in Section VI.

II. SYSTEM DESCRIPTION

In this section, we first describe the architecture of an EON. We then present the PLI model. Lastly, using an illustrative example, we demonstrate how TG and MPR enhance spectral efficiency.

A. Network architecture

In the EON architecture depicted in Fig. 1, traffic connections (i.e., connections/requests/demands) enter the optical transmission network via router ports and MFTs [5], [19]. MFTs aggregate these connections at the optical layer and provide flexible bandwidth management; thus, they are preferred to BVTs [21]. An MFT can be logically split into multiple subtransponders. Each subtransponder is assigned its own tunable laser and modulation/demodulation device. End-to-end connections are realized using lightpaths set up between subtransponders of different nodes. These connections are transported along fibers (i.e., links) that are equipped with amplifiers and switched by bandwidth-variable wavelength-selective switches.

In the EON architecture considered in this paper, TG is performed both electronically and optically. At the electronic layer of a grooming node, multiple low-speed traffic connections are electronically groomed into a single connection and dropped to/added from the optical layer by a single router port. Then, at the optical layer, multiple connections arriving from different router ports are optically groomed into an MFT [5], [19]. Furthermore, a low-speed traffic connection can be groomed with other connections at its source node (i.e., source grooming) and regroomed at an intermediate grooming node (i.e., intermediate grooming). A transiting connection is regenerated and retransmitted by the intermediate grooming nodes [22].

B. Physical-layer impairment model

In this study, we estimate the PLI using the Gaussian noise (GN) model presented in [18]. For every connection i, the signal-to-noise ratio (SNR) is expressed as

$$\text{SNR}_i = \frac{G}{G_{\text{ASE}_i} + G_{\text{NLI}_i}} \qquad (1)$$

where G, G_{ASE_i}, and G_{NLI_i} denote the power spectral density (PSD) of the signal, the PSD of the amplified spontaneous emission (ASE) noise, and the PSD of the noise from nonlinear impairments (NLI), respectively. To be able to successfully transmit connection i along the designated path, the condition $\text{SNR}^k \leq \text{SNR}_i$ has to be satisfied when connection i is assigned modulation format k and SNR^k represents the SNR
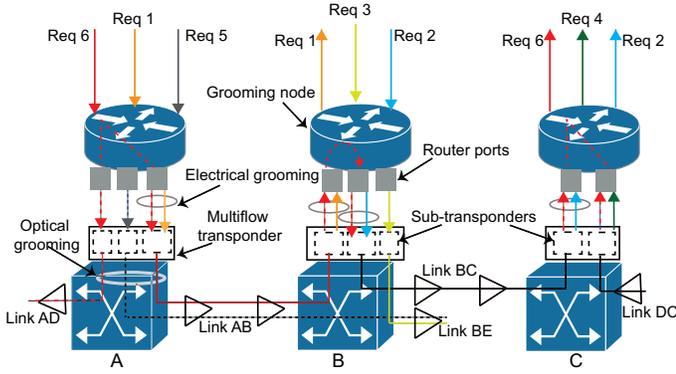
Fig. 1. The fundamental architecture of an EON.

threshold needed to satisfy the bit error rate (BER) performance of the assigned modulation format.

According to [14] and [23, Eq. (1)], G_{ASE_i} and G_{NLI_i} are calculated for a given route r as

$$G_{\text{ASE}_i} = \sum_{l \in r} N^l G_{\text{ASE}}^0 \tag{2}$$

$$G_{\text{NLI}_i} = \sum_{l \in r} N^l G_{\text{NLI}_i}^l \tag{3}$$

where

$$G_{\text{ASE}}^0 = (e^{\alpha L} - 1)n_{sp}hv \tag{4}$$

$$G_{\text{NLI}_i}^l = \mu \left(\text{asinh}(\rho B_i^2) + \sum_{j \in A_l \setminus \{i\}} \ln \left[\frac{(\Delta f_{ij} + B_j/2)}{(\Delta f_{ij} - B_j/2)} \right] \right). \tag{5}$$

Here, G_{ASE}^0 and $G_{\text{NLI}_i}^l$ are ASE and NLI noise encountered by connection i on a single span on link l, respectively; N^l is the number of spans on link l along the designated route r; L is the length of each span; α is the fiber power attenuation; n_{sp} is the spontaneous emission factor; h is Planck's constant; v is the light frequency; $\mu = (3\gamma^2 G^3)/(2\pi\alpha|\beta_2|)$; $\rho = (\pi^2|\beta_2|)/(2\alpha)$; A_l is the set of all connections on link l; B_i is the bandwidth of connection i; Δf_{ij} is the center frequency spacing between connections i and j; γ is the fiber nonlinearity coefficient; and β_2 is the fiber dispersion.

C. Traffic grooming and multipath routing

Given the network architecture and dimensions, the traffic matrix, and quality of service requirements (i.e., BER), the primary objective of RSA schemes is to increase spectral efficiency. In this section, we explain how TG and MPR can further increase spectral efficiency over existing TR-based RSA schemes and the RIA RSA scheme in [14].

Fig. 2 provides an example of a small EON that is required to accommodate 6 offline traffic connections. The span length of each link is indicated beside the respective link in the same figure. It is assumed that every link can serve a maximum of 7 spectrum slots and every connection can be served using modulation format 1 (Mod1) or 2 (Mod2). Mod2 is assumed to have a larger constellation. Therefore, in comparison to Mod1, Mod2 requires fewer spectrum slots to accommodate a given connection. For the TR-based scheme with GB = 1, we

assume that the maximum transmission distances are 5 and 2.5 spans for Mod1 and Mod2, respectively. On the other hand, in the RIA scheme, the spacing between connections could be varied according to the current state of the network. Hence, the reach of Mod2 can be adjusted by increasing GB.

In this example, if the TR-based scheme is used, connection requests (Req) 5 and 6 cannot be served due to the absence of spectrum slots along path 1 or path 2. On the other hand, if the RIA RSA scheme is used, Req 2 and 3 can be served using Mod2, by accounting for actual impairments and adjusting the size of GB accordingly. Consequently, Req 5 can now be accommodated. By incorporating MPR and TG into this scheme, every traffic connection can now be accommodated. Req 6 is now routed along two separate paths and is groomed at node A and node B with Req 1 and Req 2, respectively. A closer look at Fig. 1 reveals how connections are routed from nodes A to C in the considered EON.

III. TG- AND MPR-ENABLED IMPAIRMENT-AWARE RSA

The primary objective of this study is to increase the spectral efficiency of an impairment-aware EON by appropriately incorporating TG and MPR. In this section, the above objective is expressed as an RSA problem that is solved using ILP and heuristics. The complexity of the proposed ILP was reduced by predetermining K shortest loopless routing paths for every connection using Yen's algorithm and performing spectrum and modulation conversion at the source nodes only.

A. Problem statement

The problem of increasing spectral efficiency of a TG- and MPR-enabled impairment-aware RSA contains the following inputs:

- (V, E) : The physical topology of the network consisting of a set of nodes, V, and a set of links, E. Every link $l \in E$ has two fibers that carry traffic in opposite directions.
- C : The spectral width of each subcarrier.
- M : The set of modulation formats.
- c_k : The spectral efficiency of modulation format $k \in M$.
- T : The set of all connections in the network.
- Λ_i : The data rate of connection $i \in T$.
- P_i : The set of K precomputed routing paths for connection $i \in T$.
- SNR^k : The required SNR threshold of the modulation format $k \in M$.

Given the above input parameters, the proposed algorithms perform routing, modulation format allocation, and spectrum allocation to maximize spectral efficiency, while adhering to the following constraints: 1) serve all traffic connections, 2) ensure spectrum continuity and contiguity on transparent links for every connection, 3) perform non-overlapping spectrum allocation, and 4) guarantee SNR performance.

The proposed ILP can be used by network operators to find the routing paths, modulation format, and subcarriers for every connection i in the connection set T to arrive at the desired objective. Every connection i can be split into multiple connections (e.g., i_a and $i_{\tilde{a}}$, where a and $\tilde{a}$ are indices of the subconnections) and routed along several routing
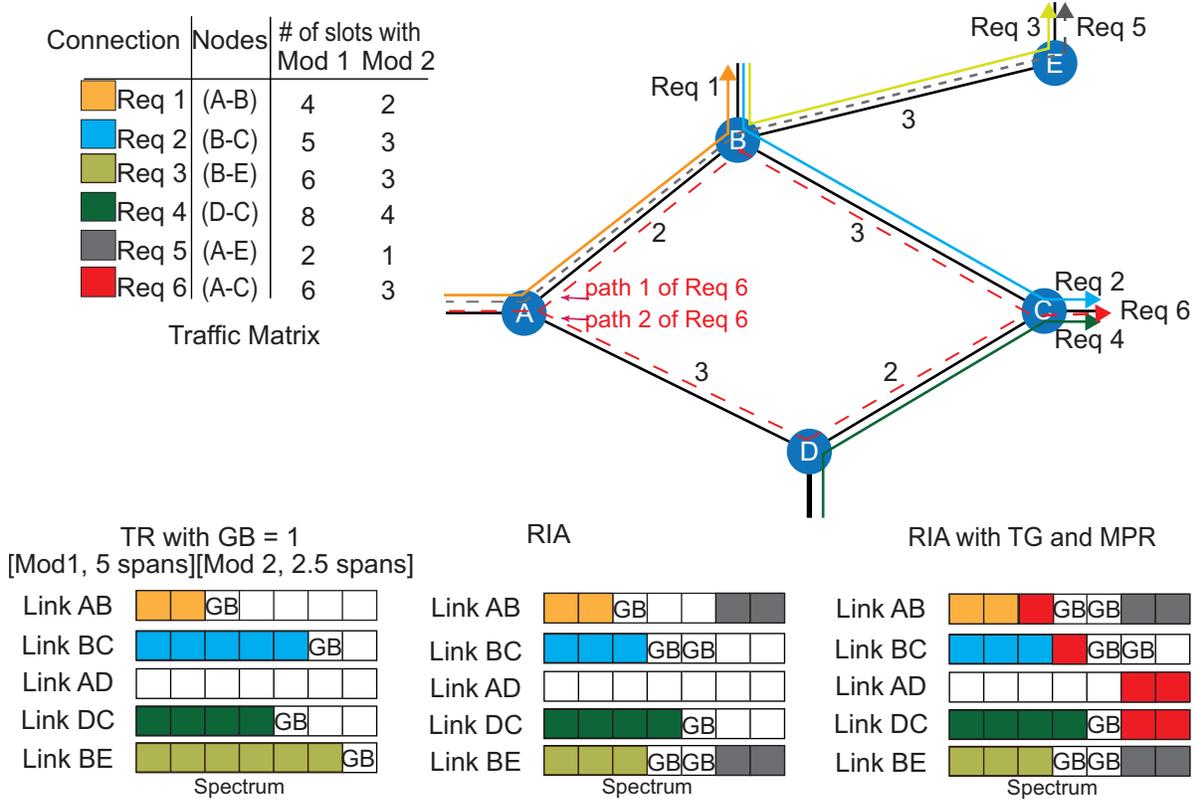
Fig. 2. Spectrum allocation with the TR, RIA, and modified RIA schemes in a small network serving 6 connections. Note that the inclusion of TG and MPR (i.e., modified RIA) allows all connections to be served.

paths selected from P_i. The data rate of traffic carried by a subconnection is specified by B_{i_a}. For example, in Fig. 2, Req 6 (i.e., connection i) is split into two subconnections (e.g., i_3 and i_6) that carry $1/3^{rd}$ and $2/3^{rd}$ of traffic along path 1 and path 2, respectively. The binary variable $p_{i_a}^l$ for $\forall l \in E$ specifies the links that are used by subconnection i_a. The binary variable $M_{i_a}^k$ determines the modulation format $k \in M$ assigned to subconnection i_a. The number of subcarriers along the designated path is determined by variable b_{i_a}. Finally, the lowest subcarrier index for the traffic of subconnection i_a is specified by variable f_{i_a}. Subconnections i_a and j_b that are assigned the same modulation format and travel together via a common set of links are groomed at the converging node.

B. ILP formulation

The proposed ILP takes the following additional input parameters:

- θ : A large inequality coefficient (in this study, we let $\theta = 10^4$).
- SR : The list of predefined splitting ratios (e.g., SR = $[1, 1/2, 1/2, 1/3, 1/3, 1/3, 2/3, 1/4, 1/4, 1/4, 1/4, 3/4, \ldots]$).
- δ : The set of possible indices of subconnections (e.g., $\delta = [0, 1, 2, 3, 4, 5, 6, \ldots]$ where $|\delta| = |\text{SR}|$). Notation i_a denotes the ath subconnection of connection i, where a is an index in δ.
- B_{i_a} : The data rate of the ath subconnection of i, which

is $B_{i_a} = \text{SR}[a] \times \Lambda_i$. [1]
- $T_{i_a}^k$: The number of subcarriers required to serve B_{i_a} when assigned the modulation format $k \in M$.

The variables are:

- ξ : The highest subcarrier index.
- X_{i_a} : is 1 if connection $i \in T$ uses subconnection i_a.
- b_{i_a} : The number of subcarriers required to transport i_a.
- $M_{i_a}^k$: is 1 if i_a is assigned modulation format $k \in M$; else, 0.
- $q_{i_a}^r$: is 1 if i_a is routed along path $r \in P_i$; else, 0.
- $p_{i_a}^l$: is 1 if i_a is routed through link $l \in E$; else, 0.
- f_{i_a} : The lowest subcarrier index of i_a.
- y_{i_a,j_b} : is 1 if i_a and j_b share a link when routed along their respective paths; else, 0.
- u_{i_a,j_b} : is 1 if i_a and i_b share a link and the spectrum of i_a is to the left of the spectrum of j_b, i.e., $y_{i_a,j_b} = 1$ and $f_{i_a} + b_{i_a} \leq f_{j_b}$; else, 0.
- Δ_{i_a,j_b}^{kh} : is 1 if the center frequency spacing between i_a and j_b in a shared link $l \in E$ is $hC/2$ when $M_{i_a}^k = 1$ and $0 \leq h \leq \xi$; else, 0.
- w_{i_a,j_b}^l : The cross-channel interference on i_a by j_b on shared link $l \in E$.
- $t_{i_a}^l$: The total NLI encountered by i_a in link $l \in E$.
- GR_{i_a,j_b}^k : is 1 if i_a and j_b are groomed together at modulation format k.

[1]The total data rate of all the subconnections $a \in \delta$ is greater than the traffic demand Λ_i, but only a subset of δ will be actually used.

The NLI defined in (3) depends on:

- $J_{i_a}^{kh} = \mu \ln \frac{|h+T_{i_a}^k|}{|h-T_{i_a}^k|}$: The partial nonlinear effect (cross-channel interference) introduced by i_a on another subconnection when $M_{i_a}^k = 1$ and the center frequency spacing between i_a and the other subconnection is $hC/2$.
- $H_{i_a}^k = \mu \operatorname{asinh}(\rho(CT_{i_a}^k)^2)$: The partial nonlinear effect (self-channel interference) of i_a when assigned modulation format $k \in M$.

Using the above input parameters and variables, we formulate the following optimization problem. Spectral efficiency is increased by minimizing the highest subcarrier index ξ assigned to any given connection.

Objective:

$$\text{Minimize} \quad \xi \tag{6}$$

Constraints:

$$\Lambda_i \leq \sum_{a \in \delta} B_{i_a} X_{i_a} \quad \forall i \in T \tag{7}$$

$$\sum_{r \in P_i} q_{i_a}^r - X_{i_a} = 0 \quad \forall a \in \delta, i \in T \tag{8}$$

$$\sum_{k \in K} M_{i_a}^k - X_{i_a} = 0 \quad \forall a \in \delta, i \in T \tag{9}$$

$$p_{i_a}^l - q_{i_a}^r \geq 0 \quad \forall a \in \delta, i \in T, r \in P_i, l \in r \tag{10}$$

$$p_{i_a}^l + p_{j_b}^l = 1 + y_{i_a,j_b} \quad \forall a,b \in \delta, i,j \in T, j_b \neq i_a \tag{11}$$

$$b_{i_a} \geq \sum_{k \in M} T_{i_a}^k M_{i_a}^k \quad \forall a \in \delta, i \in T \tag{12}$$

$$f_{i_a} + b_{i_a} - f_{j_b} \leq \theta(1 - u_{i_a,j_b}) \quad \forall a,b \in \delta, i,j \in T, j_b \neq i_a \tag{13}$$

$$f_{j_b} + b_{j_b} - f_{i_a} \leq \theta(1 - y_{i_a,j_b} + u_{i_a,j_b}) \tag{14}$$

$$\xi \geq f_{i_a} + b_{i_a} - 1 \quad \forall a \in \delta, i \in T \tag{15}$$

$$\sum_h \Delta_{i_a,j_b}^{kh} = M_{j_a}^k \quad \forall a,b \in \delta, i,j \in T, j_b \neq i a, k \in M \tag{16}$$

$$\sum_{k,h} h \Delta_{i_a,j_b}^{kh} \leq \quad \theta(1 - u_{i_a,j_b}) + 2(f_{j_b} - 1) + b_{j_b}$$
$$- 2(f_{i_a} - 1) - b_{i_a} \leq 2\theta(2 - \mathrm{GR}_{i_a,j_b}^k - u_{i_a,j_b}) + T_{i_a}^k + T_{j_b}^k \tag{17}$$

$$\sum_{k,h} h \Delta_{i_a,j_b}^{kh} \leq \quad \theta(1 - y_{i_a,j_b} + u_{i_a,j_b}) + 2(f_{i_a} - 1) + b_{i_a}$$
$$- 2(f_{j_b} - 1) - b_{j_b} \leq 2\theta(1 - \mathrm{GR}_{i_a,j_b}^k + u_{i_a,j_b}) + T_{i_a}^k + T_{j_b}^k$$
$$\forall a,b \in \delta, i_a \neq j_b, i,j \in T, k \in M \tag{18}$$

$$\mathrm{GR}_{i_a,j_b}^k \geq M_{i_a}^k + M_{j_b}^k - 2$$
$$\forall a,b \in \delta, j_b \neq i_a, k \in M, i,j \in T \tag{19}$$

$$\mathrm{GR}_{i_a,j_b}^k \leq M_{i_a}^k \quad \forall a,b \in \delta, j_b \neq i_a, k \in M, i,j \in T \tag{20}$$

$$\mathrm{GR}_{i_a,j_b}^k \leq M_{j_b}^k \quad \forall a,b \in \delta, j_b \neq i_a, k \in M, i,j \in T \tag{21}$$

$$w_{i_a,j_b}^l \geq \theta(p_{j_b}^l - 1) + \sum_{k,h}(\Delta_{i_a,j_b}^{kh} - \mathrm{GR}_{i_a,j_b}^k)J_{j_b}^{kh}$$
$$\forall a,b \in \delta, j_b \neq i_a, i,j \in T \tag{22}$$

$$t_{i_a}^l \geq \theta(p_{i_a}^l - 1) + \sum_k M_{i_a}^k H_{i_a}^k + \sum_{\forall j_b \neq i_a} w_{i_a,j_b}^l$$
$$\forall a \in \delta, l \in E, i \in T \tag{23}$$

$$G \sum_k M_{i_a}^k / \mathrm{SNR}^k \geq \theta(q_{i_a}^r - 1) + G_{\mathrm{ASE}}^0 \sum_{l \in r} p_{i_a}^l N^l + \sum_{l \in r} N^l t_{i_a}^l$$
$$\forall a \in \delta, r \in P_i, i \in T \tag{24}$$

where constraint (7) appropriately distributes traffic among multipaths; (8) and (9) constrain a subconnection to a single path and a single modulation format, respectively; (10) specifies the used links along the selected path; (12) ensures subcarrier allocation for the selected subconnection at the assigned modulation format; (11) ensures that y_{i_a,j_b} is 1 if subconnections i_a and j_b travel along a common link; (13) and (14) guarantee non-overlapping spectrum allocation; and (15) guarantees that the objective ξ is no less than the maximum subcarrier index of every subconnection. Constraints (16)–(18) define the center frequency spacing h between i_a and j_b along their respective paths when they are groomed and not groomed. If i_a and j_b are groomed at modulation format k, the spacing between their center frequencies would be half of $T_{i_a}^k + T_{j_b}^k$. Constraints (19)–(21) allow two subconnections that are assigned an identical modulation format to be groomed. Lastly, (22)–(24) guarantee that the SNR requirement is satisfied for every subconnection on its assigned path.

C. Proposed heuristic: TMG

Due to many integer constraints, the complexity of the proposed ILP scales poorly with the number of connections and the size of the networks. Therefore, as optimum solutions can only be found for small networks, we propose the TMG heuristic algorithm that effectively combines **TG**, **MPR**, and the **GN** model.

The proposed TMG heuristic increases spectral efficiency by accommodating traffic connections that cannot be routed due to excessive PLI or due to the absence of spectrum along a single path using TG and MPR. Before we explain the TMG heuristic in detail, we introduce the additional input parameters, variables, and important functions:

Additional input parameters for TMG:
- σ : The maximum number of MPR attempts allowed for every connection, $(\sigma > 0)$.
- ω : The maximum number of TG attempts, $(\omega > 0)$.

Variables:
- mp_i : The number of routing paths travelled by connection $i \in T_u$ $(0 < \mathrm{mp}_i \leq \sigma)$.
- tg : The number of connections that are groomed at an intermediate node $(0 \leq \mathrm{tg} \leq \omega)$.
- ζ : A positive integer number.
- T_u : The set of updated connections (i.e., manipulated T).
- R_i : The set of paths for connection $i \in T_u$.
- Υ_i : The data rate of connection $i \in T_u$.
- T_i^k : The number of contiguous spectrum slots (i.e., band size) required to serve connection $i \in T_u$ at modulation format $k \in M$.
- f_i : The lowest subcarrier index of connection $i \in T_u$.
- ϕ_i : The highest subcarrier index of connection $i \in T_u$ $(f_i < \phi_i < \infty)$.
- r_i^{sel} : The selected path for connection $i \in T_u$ $(r_i^{\mathrm{sel}} \in R_i)$.
- $\mathrm{Cost}_{i,r,s}$: The PLI of connection $i \in T_u$ along path $r \in P_i$ when assigned a band of contiguous spectrum slots s.

- k_i : The modulation format assigned to connection i.
- Γ : The set of information of the groomed connections.
- GrN : is 0 or takes the returned value of function TestG(i, n).

Functions:

- SatM(Υ_i, k) : Returns the number of subcarriers needed to accommodate Υ_i at modulation format $k \in M$.
- Spect(T_i^k, r) : All possible contiguous bands of size T_i^k found on every link (consisting of 0 to ξ subcarriers) in path $r \in R_i$. If not found, then return 0.
- Clink(T_i^k, s, l) : Returns the total impairments encountered by connection i (found using (25)) on link l when assigned modulation format $k \in M$ and band $s \in$ Spect(T_i^k, r). $s[0]$ is the first slot of band s.
- TestG(i, n) : Returns $\{\Upsilon_i, a-n, n-d, n\}$ if connection i (between node pair $a - d$) can be split at intermediate node (not source node) n and groomed with connections of node pairs $a-n$ and $n-d$ without employing additional router ports. Else, return 0.
- Foundpath(i) : is 1 if a path with sufficient spectrum is found for connection i with any modulation format; else, 0.
- Served(i) : is 1 if connection i can be routed successfully; else, 0.
- CondMP(i, GrN) = **True** if (Served$(i) \neq 1$ & FoundPath$(i) \neq 1$ & mp$_i \leq \sigma$) or (Served$(i) \neq 1$ & FoundPath$(i) == 1$ & mp$_i \leq \sigma$ & GrN $== 0$)
- CondTG(i, GrN) = **True** if (FoundPath$(i) == 1$ & tg $\leq \omega$ & Served$(i) \neq 1$ & GrN $\neq 0$)

The proposed heuristic that begins with an empty network involves a number of stages. First, the input parameter σ is assigned a small value to ensure that a connection is not split into a large number of small-sized connections that reduce spectral efficiency [10]–[12]. ω is assigned the value of 20 as multiple trials showed that assigning a larger value (e.g., $\omega > 20$) did not provide any improvements in spectral gains. Then, the temporary variables mp$_i$ and tg are initialized to zero. Variables $T_u, R_i,$ and Υ_i mirror the input parameters $T, P_i,$ and Λ_i, respectively. While the output parameters r_i^{sel} and k_i are set to 0, ϕ_i and f_i are set to ∞. Γ begins with an empty set.

Second, the connections are sorted in descending order of their bit rate. Third, the connections are served one after the other while accounting for PLI. The general flow diagram and the pseudocode of the algorithm are shown in Fig. 3 and Algorithm 1, respectively.

If a connection cannot be routed (with any modulation format) due to excessive PLI, the algorithm performs TG. Alternatively, if the connection cannot be routed due to insufficient subcarriers, the algorithm then performs MPR. If the algorithm fails to accommodate all connections before mp$_i$ or tg exceed their respective thresholds (σ and ω, respectively), ξ is increased. As in the work of [14], an adaptive SNR margin, ζX_i^k, where $X_i^k = \mu \ln(1 + T_{\text{avg}}^1/(T_i^k/2))$, is introduced to every served connection to protect them against interference from future connections, where T_{avg}^1 denotes the average
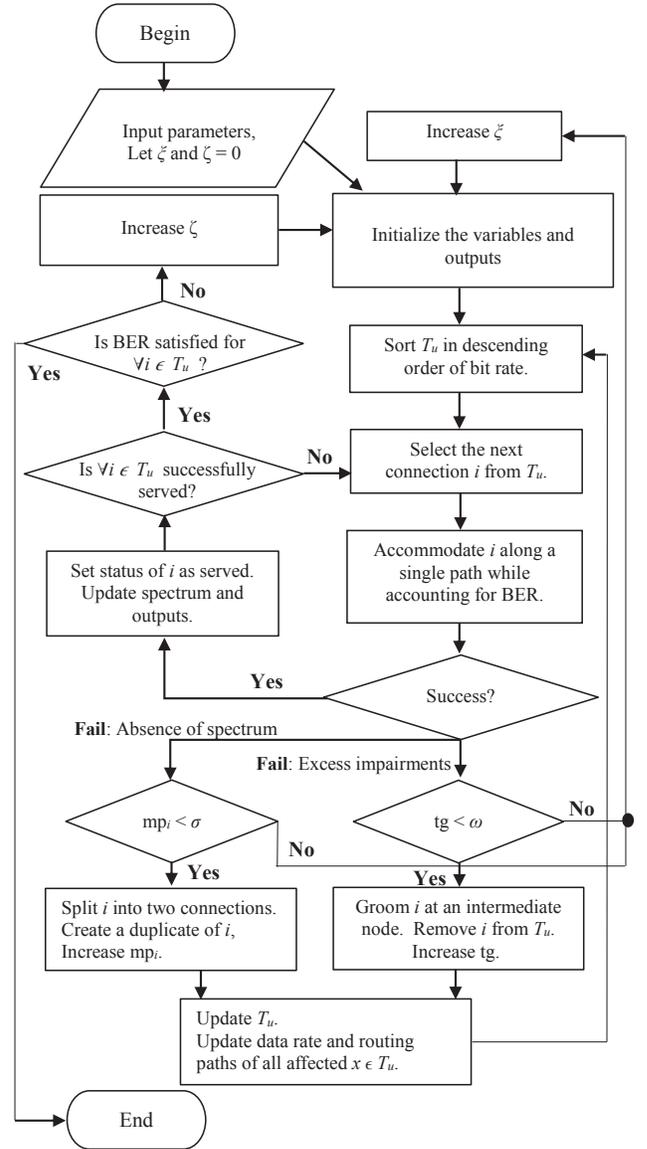
Fig. 3. Flow chart of the proposed algorithm.

number of subcarriers needed to serve a given connection when the assigned modulation format is 1. The total PLI encountered by connection i when assigned modulation format k on link l (PLI$_i^l$) is calculated using

$$\text{PLI}_i^l = N^l \left(G_{\text{ASE}}^0 + G_{\text{NLI}_i}^l + \zeta X_i^k \right). \tag{25}$$

Given the set of input parameters defined in III-A and III-C, the proposed TMG heuristic algorithm specifies the routing path r_i^{sel}, the modulation format k_i, and the assigned subcarriers $f_i - \phi_i$, for every connection i in the updated connection set T_u, which are the desired outputs for network operators. Additionally, information in Γ can be used to groom traffic from one connection to another at an intermediate node.

IV. PERFORMANCE EVALUATION

Numerical simulations were carried out to evaluate the spectral efficiency of the proposed TG- and MPR-enabled impairment-aware RSA scheme. The results of the proposed

---

Algorithm 1 Allocate one connection at a time

Inputs : $(V, E), C, c_k, M, T, \Lambda_i, P_i, \text{SNR}_k, \sigma, \omega$. Let $\{\xi, \zeta\} \leftarrow 0$.
Outputs: $\{r_i^{\text{sel}}, k_i, \phi_i, f_i\} \forall i \in T_u$, Γ, and final value of ξ.

1: **Initialization:**
2: $\quad T_u \leftarrow T$, $\Gamma = [\ ]$, tg $\leftarrow 0$
3: $\quad R_i \leftarrow P_i$, $\Upsilon_i \leftarrow \Lambda_i$ $\qquad\qquad\qquad\qquad\qquad \triangleright \forall i \in T$
4: $\quad \{\text{mp}_i, r_i^{\text{sel}}, k_i\} \leftarrow 0$ and $\{\phi_i, f_i\} \leftarrow \infty$ $\qquad \triangleright \forall i \in T$
5: **Sort** T_u in descending order of bit rate.
6: **for** $i \in T_u$ **do** $\qquad\qquad\qquad\qquad \triangleright$ every connection
7: $\quad$ Served$(i) \leftarrow 0$
8: $\quad$ FoundPath$(i) \leftarrow 0$
9: $\quad$ GrN $\leftarrow 0$
10: $\quad$ **for** $\forall k \in M$ **do** $\qquad\qquad \triangleright$ every modulation format
11: $\qquad T_i^k \leftarrow \text{SatM}(\Upsilon_i, k)$
12: $\qquad$ **for** $\forall r \in R_i$ **do** , $\qquad\qquad\qquad\qquad \triangleright$ every path
13: $\qquad\quad$ if $|\text{Spect}(T_i^k, r)| \neq 0$ then let FoundPath$(i) \leftarrow 1$
14: $\qquad\quad$ **for** $\forall s \in \text{Spect}(T_i^k, r)$ **do** $\qquad \triangleright$ every band
15: $\qquad\qquad$ Cost$_{i,r,s} \leftarrow 0$
16: $\qquad\qquad$ **for** every link l in path r **do**
17: $\qquad\qquad\quad$ Cost$_{i,r,s} = \text{Cost}_{i,r,s} + \text{Clink}(T_i^k, s, l)$
18: $\qquad\qquad\quad$ **if** Cost$_{i,r,s} \geqslant \frac{G}{\text{SNR}^k}$ **then**
19: $\qquad\qquad\qquad n \leftarrow$ source node of link l.
20: $\qquad\qquad\qquad$ **if** TestG$(i, n) == 1$ **then**
21: $\qquad\qquad\qquad\quad$ GrN $\leftarrow$ TestG(i, n)
22: $\qquad\qquad\qquad$ **end if**
23: $\qquad\qquad\qquad$ Discard current s, go to line 14.
24: $\qquad\qquad\quad$ **end if**
25: $\qquad\qquad$ **end for**
26: $\qquad\qquad$ **if** Cost$_{i,r,s} \leqslant \frac{G}{\text{SNR}^k}$ and $s[0] + T_i^k - 1 \leqslant \phi_i$ **then**
27: $\qquad\qquad\quad r_i^{\text{sel}} \leftarrow r$, $k_i \leftarrow k$, $\phi_i \leftarrow s[0] + T_i^k - 1$, $f_i \leftarrow s[0]$
28: $\qquad\qquad\quad$ Served$(i) \leftarrow 1$
29: $\qquad\qquad$ **end if**
30: $\qquad\quad$ **end for**
31: $\qquad$ **end for**
32: $\quad$ **end for**
33: $\quad$ **if** CondMP(i, GrN) true **then**
34: $\qquad$ Halve the traffic of connection i (let $\Upsilon_i \leftarrow \Upsilon_i/2$).
35: $\qquad$ Create a duplicate connection j (let $j \leftarrow i$, $\Upsilon_j \leftarrow \Upsilon_i$,
36: $\qquad$ and $R_j \leftarrow R_i$).
37: $\qquad$ Add the duplicate connection j to T_u.
38: $\qquad$ Let $\text{mp}_i \leftarrow \text{mp}_i + 1$, Go to line 5
39: $\quad$ **else if** CondTG(i, GrN) true **then**
40: $\qquad$ Groom connection i at node GrN$[-1]$ onto other connections.
41: $\qquad$ Increase data rate of these other connections by Υ_i.
42: $\qquad$ Remove connection i from T_u.
43: $\qquad$ Append GrN$[-1]$ to Γ.
44: $\qquad$ Let tg $\leftarrow$ tg $+ 1$, Go to line 5
45: $\quad$ **else if** Served$(i) \leftarrow 1$ **then**
46: $\qquad$ Update spectrum on links along r_i^{sel}.
47: $\qquad$ Store $\{r_i^{\text{sel}}, k_i, f_i, \phi_i\}$
48: $\qquad$ (set the states of spectrum f_i to ϕ_i as unavailable)
49: $\quad$ **else**
50: $\qquad$ Increase subcarrier index ($\xi \leftarrow \xi + 1$);
51: $\qquad$ Go to line 1
52: $\quad$ **end if**
53: **end for**
54: **if** any connection does not meet the BER requirement **then**
55: $\quad$ increase the value of adaptive margin by letting $\zeta \leftarrow \zeta + 1$.
56: $\quad$ Go to line 1
57: **else**
58: $\quad$ Output
59: **end if**

---

optimal and suboptimal schemes were compared against the benchmarks: the RIA and existing TR-based RSA schemes. In addition to spectral efficiency, the performance of the proposed scheme was evaluated in terms of resource consumption and performance against different network dimensions.

A. Benchmarks

The performance of our proposed ILP and TMG heuristic were compared against the benchmarks: the TR-based RSA scheme [1] and the RIA scheme [14]. These benchmark algorithms are summarized here, while readers may refer to [1] and [14] for further elaboration.

In the TR-based RSA scheme presented in [1], the PLI are

dealt with by assigning a fixed TR for different modulation formats and by adjusting the GB size. Specifically, a connection that is assigned modulation format m cannot travel beyond the TR of the assigned modulation format R^m. A fixed-size GB is used to avoid interference between connections, without estimating the actual NLI. The proposed heuristic in [1] works in the following sequence. First, for each pair of nodes, $K = 3$ paths are pre-computed. Next, the connections are sorted by (i) the descending order of their Λ_i, (ii) the descending order of the number of links on their shortest paths, and (iii) using a simulated annealing method. Then, for each of the K routing paths, a connection is assigned the highest-order modulation format m whose R^m is longer than the length of the path. Subsequently, for each connection, the algorithm calculates the spectrum availability of all the candidate K paths, taking into account the spectrum continuity and contiguity constraints. Lastly, the algorithm selects the path that requires the lowest subcarrier index. When comparing our proposed heuristic against this benchmark TR-based RSA scheme, we considered the latter's third sorting criterion, which provided the optimum results in their work. Furthermore, we calculated the fixed TR used in their scheme for different values of PSD G. At modulation format m, the TR was $\lfloor G/\text{SNR}^m G_{\text{ASE}}^0 \rfloor$ spans.

Unlike the TR-based RSA scheme, the RIA heuristic (corresponding to the 'connection list (CL)' heuristic [14]) accounts for impairments by calculating the actual linear and nonlinear impairments using the GN model [14]. This heuristic involved the following stages. First, the connections are sorted in decreasing order of Λ_i. Then, the heuristic sequentially provisions connections and an adaptive SNR margin is introduced to counter the impairments of future connections. Next, for every connection, the algorithm investigates every path, every modulation format, and every possible sub-carrier band. Among the many possibilities, the heuristic then selects the path, modulation format, and spectrum-band that satisfy the SNR requirement and produce the smallest subcarrier index. However, if a feasible allocation cannot be found for any connection, the adaptive margin is increased, and the heuristic is repeated.

B. Parameters and assumptions

The values assigned to the input parameters are listed in Table I. In the considered EONs, each link consists of two fibers in opposite directions. In this study, we assume that the EON supports both source grooming and intermediate grooming. Thus, a connection that travels from node a to node d can be split at an intermediate node y and groomed onto connections that travel between node pairs $a - y$ and $y - d$, respectively. Furthermore, the actual effect of differential delay between traffic routed along multiple paths is not considered [11]. However, to reduce the effect of differential delay, the value of K in Yen's algorithm is constrained to a small number (e.g., 5) [4]. In this study, bandwidth is measured in terms of spectrum slots and, thus, spectral efficiency is measured in terms of the maximum subcarrier index to support a certain traffic matrix. A higher spectral efficiency thus implies a reduced maximum subcarrier index. In our work, SR $= [1, 1/2, 1/2, 1/3, 1/3, 1/3, 2/3, 1/4, 1/4, 1/4, 1/4, 3/4]$.

TABLE I
VALUES ASSIGNED TO THE PARAMETERS [14], [21]

Parameter	Value
α	0.22 dB/km
γ	1.3 Wkm^{-1}
β_2	-21.7ps^2/km
n_{sp}	1.8
v	93 THz
L	100 km
M	{1,2,3,4} b/s/Hz
BER requirement	10^{-3}
σ	3
ω	20
Router port bit rate	400 Gb/s $\times 1$
MFT bit rate	400 Gb/s
C	6.25 GHz

In Matrakidis et al. [28], the authors use multi-granular transponders (10 Gb/s, 40 Gb/s fixed line rate, and 400 Gb/s elastic line rate) that are connected to 400 Gb/s line cards, but their results show that using flexible 400 Gb/s elastic transponders leads to higher cost savings. In our study, we only use 400 Gb/s elastic transponders (i.e., MFTs) connected to 400 Gb/s $\times$ 1 port line cards [20], [21], as we are attempting to optimize RSA and modulation format allocation through the incorporation of TG and MPR, and the ensuing complexity of the problem is very high. Moreover, in our work, 400 Gb/s ports are required to accommodate traffic connections that carry high bit rates (above 343.75 Gb/s), and according to Tanaka et al. [20], [21], a 400 Gb/s MFT can only use a single type (400 Gb/s x 1 port) line card.

In our work, the considered MFT has a maximum data rate of 400 Gb/s and a subcarrier width of 6.25 GHz. As elaborated in [20] and [21], 400 Gb/s MFTs can accommodate traffic flows or connections of different data rates by slicing the capacity of the MFT. We define the capacity of the MFT in terms of the data rate (i.e., 400 Gb/s is sliced according to each connection's bit rate requirement) and not in terms of subcarriers. It is difficult to define the capacity of the MFT in terms of subcarriers as the subcarrier capacity (in Gb/s) would be dissimilar for the multiple modulation schemes that are used in our work, which have different spectral efficiencies.

Simulations were run on a desktop computer with a four-core, i7-4790 CPU and 16 GB RAM. The typical running time for the ILP was 1.5 hours, while the TMG heuristic took, on average, 11 minutes to find a solution.

C. Spectral efficiency increase

First, the spectral efficiency of the proposed scheme is evaluated numerically for a randomly-generated 5-node small network (with average link length 650 km) and the 14-node Deutsche Telekom backbone network [24]. To represent the heterogeneity of network traffic [14], [25], in each traffic matrix, all connections (i.e., connections between every node pair) are assigned a bit rate that is randomly selected from the range of 6.25–190 Gb/s in steps of 6.25 Gb/s with a uniform distribution, which corresponds to between 1 and 30 binary phase shift keying (BPSK) subcarriers. To strengthen the validity of our findings, the test results in each figure
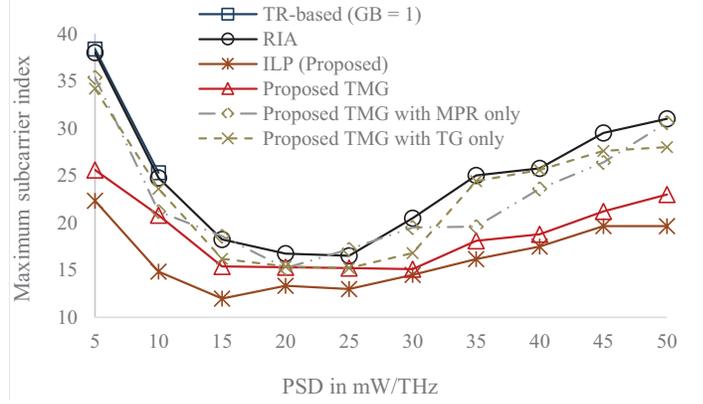
Fig. 4. Maximum subcarrier index vs. PSD for different schemes in the 5-node network.
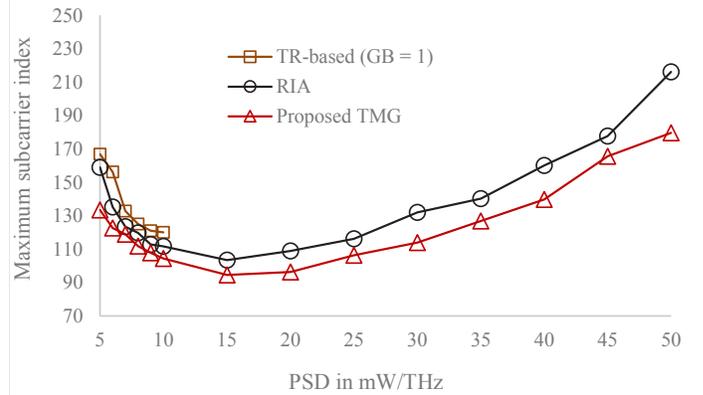
Fig. 5. Maximum subcarrier index vs. PSD for different schemes in the 14-node network.
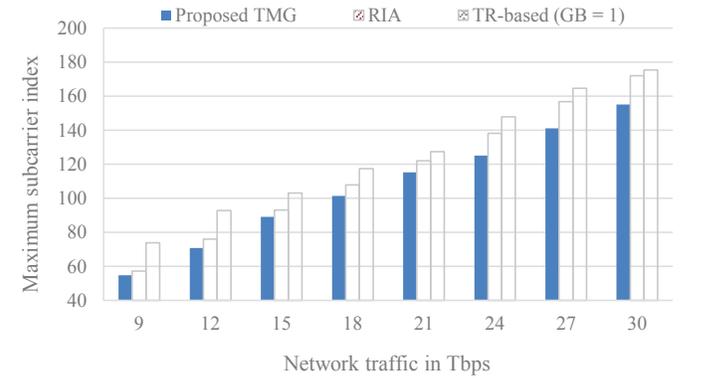
Fig. 6. Maximum subcarrier index vs. increasing network traffic load for different schemes in the 14-node network ($G = 10$ mW/THz).

are presented as an average after repeating a simulation with different randomly-generated traffic matrices.

Figs. 4 and 5 depict the maximum subcarrier index of different schemes against increasing PSD (averaged over ten traffic matrices) for the two considered networks. For the small network (Fig. 4), the proposed ILP and the TMG heuristic offer increased spectral efficiency over the benchmark schemes at all PSD values. While the proposed ILP offers the maximum spectral efficiency (i.e., reduced maximum subcarrier index), the TMG heuristic follows closely with an average difference of 13%. In comparison to the benchmark schemes, the TMG heuristic offers an increase in spectral efficiency of up to 28%. Fig. 4 also shows that the TMG heuristic that incorporates both TG and MPR outperforms the same heuristic with only TG or MPR in terms of spectral savings.

For both networks, the benchmark TR-based scheme (with a single GB) is only feasible for small PSD values. Among the three plotted schemes for the large network (Fig. 5), it is observed that the proposed TMG heuristic can offer spectral savings of up to 17% and 20% over the RIA and TR-based schemes, respectively. It is also observed that the spectral savings of TMG over the benchmark schemes decreases as the PSD G is increased from 5 to 15 mW/THz, but increases gradually as the PSD is raised over 20 mW/THz. This pattern indicates that the proposed scheme is highly preferable, specifically at higher PSD values.

In Fig. 6, the maximum subcarrier index is plotted against increasing network traffic for the 14-node large network (when the PSD is fixed at $G = 10$ mW/THz). The differing traffic volumes indicated in Figs. 6–8 were generated by varying the random bit rate assigned to a connection between 6.25 Gb/s and 343.75 Gb/s (which does not exceed the port capacity of 400 Gb/s). While the maximum generated network traffic was found to be 62 Tb/s, the average over multiple simulations is 30 Tb/s. Regardless of the network traffic volume, the proposed TMG heuristic outperforms the benchmark schemes and provides higher spectral savings. Especially at high traffic volumes, the TMG heuristic consumes approximately 15 and 18 fewer spectrum slots than the RIA and existing TR-based schemes, respectively.

D. Resource consumption

In this study, resource consumption is evaluated in terms of router ports and MFTs, the two network components that consume the most power and cost [21], [26]. Figs. 7 and 8 depict the number of router ports and MFTs consumed, respectively, by each different scheme against increasing traffic load. In traditional WDM networks, TG combines several connections into one, thereby significantly reducing the number of router ports required to serve a given set of connections. However, in the proposed scheme, MPR splits a connection into smaller connections, which increases the demand on router ports. Therefore, a significant reduction in router port consumption cannot be expected with the proposed TMG heuristic. The TMG heuristic will perform intermediate grooming only if additional router ports do not need to be activated. In comparison to the RIA and TR-based benchmark RSA schemes,
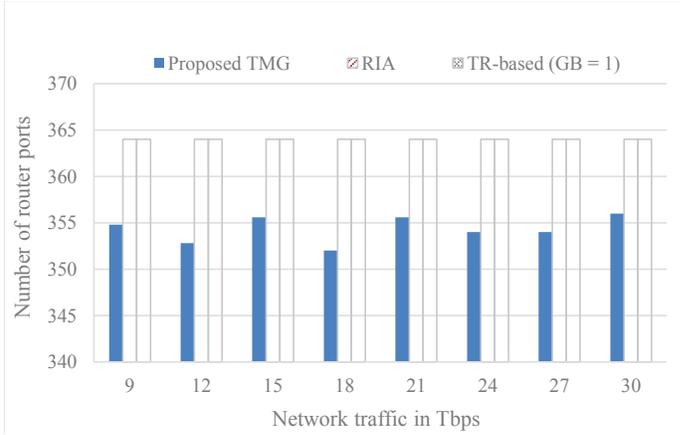
Fig. 7. Router port consumption vs. increasing network traffic (14-node network, $G = 10$ mW/THz).
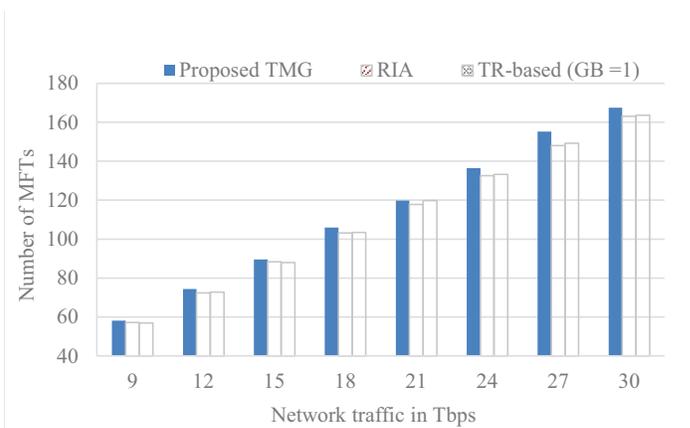
Fig. 8. MFT consumption vs. increasing network traffic (14-node network, $G = 10$ mW/THz).

however, the TMG consumes between 8 and 12 fewer router ports, on average.

An MFT has a fixed bit rate capacity. Then, as the amount of traffic added to or dropped from the optical transmission layer increases, more MFTs have to be deployed. The two benchmark schemes assign end-to-end connections. Each MFT is then only required to process local add/drop traffic with the demand on each MFT therefore being much less. On the other hand, when traffic is aggregated with the TMG heuristic, MFTs not only have to serve local traffic, but also have to serve bypass traffic. Furthermore, when a connection is allowed to split into small-sized connections and travel along multiple paths, and these small-sized connections are groomed at intermediate nodes, the traffic load on the MFTs of intermediate nodes would increase. Thus, the TMG heuristic increases the demand on MFTs. Consequently, as observed in Fig. 8, the TMG scheme consumes between 1 and 7 extra MFTs compared to the benchmark schemes.

Note that if a single modulation scheme is used, the capacity of an MFT could be defined in terms of subcarriers [5], [6]. Then, grooming would increase the utilization of subcarriers, reducing the number of subcarriers needed to accommodate traffic and, in turn, lowering the demand on the MFTs. This
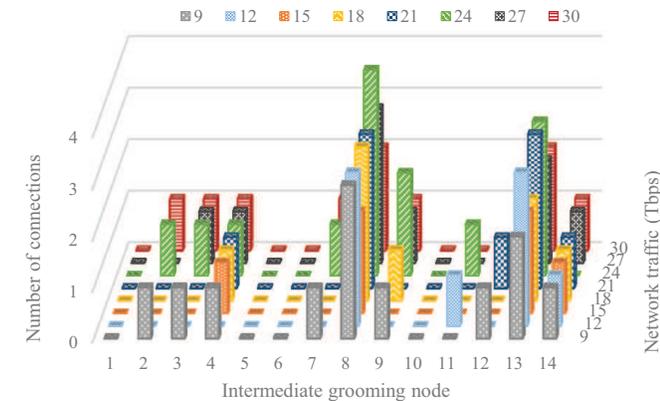
Fig. 9. The number of connections being groomed by each intermediate grooming node at different traffic volumes. Connections groomed at source nodes and transiting connections are not depicted.

Fig. 10. Maximum subcarrier index (bandwidth) vs. stretching factor for the 10-node network that has a low degree of connectivity ($G = 20$ mW/THz).

Fig. 11. Maximum subcarrier index vs. stretching factor for the 10-node network that has a high degree of connectivity ($G = 20$ mW/THz).

effect would be investigated in our future work.

Fig. 9 depicts the number of connections (on average) that are groomed at each intermediate grooming node for different traffic loads. While there could be a large number of connections traversing an intermediate node, there is only a need to groom the few connections that are blocked due to excessive impairments. In Fig. 9, it is observed that a higher number of connections are groomed at intermediate nodes 8 and 13 at all traffic loads. On the other hand, nodes 1, 5, 6, and 10 are not used as intermediate grooming nodes by any connection. As the demand on different nodes are not equal, in our opinion, network cost and power consumption can be reduced by employing high-capacity devices at nodes 8 and 13 and low-capacity devices at nodes 1, 5, 6, and 10, respectively. As the results of Fig. 9 correspond to the average of multiple simulations with different traffic matrices, it is difficult to determine a relationship between the traffic matrices and the utilization of grooming nodes. However, by comparing the results in Fig. 9 with the actual network topology, we observed that nodes with a higher degree of connectivity were often used as intermediate grooming nodes, and edge nodes situated farther away from other nodes were less likely to be used as grooming nodes.

E. Network size

To evaluate how the degree of connectivity and link length would affect the performance of the proposed TMG scheme, simulations were performed on a randomly-generated 10-node network with a connectivity degree of 2.1 (low connectivity) and 3.1 (high connectivity) and various link lengths. The link length was increased from 350 km to 2100 km by varying the stretching factor (from 1 to 6). Connections were assigned a bit rate between 1 and 30 BPSK subcarriers. For each stretching factor, the simulation was repeated with 20 different traffic matrices and the average, highest, and lowest values were recorded accordingly. Figs. 10 and 11 depict the maximum subcarrier index values of the proposed TMG heuristic and the RIA scheme for different stretching factors in low-connectivity and high-connectivity networks, respectively.
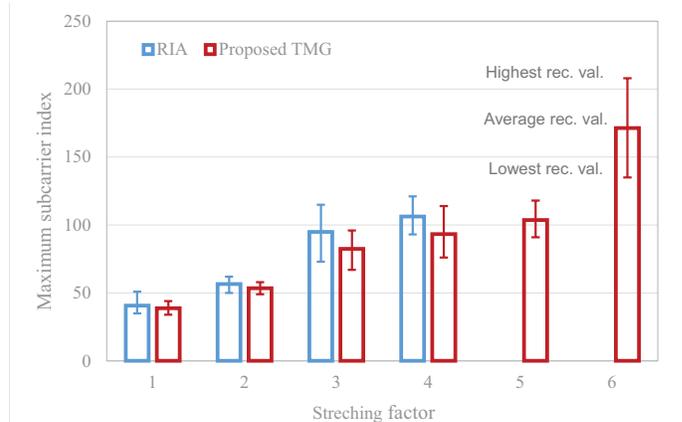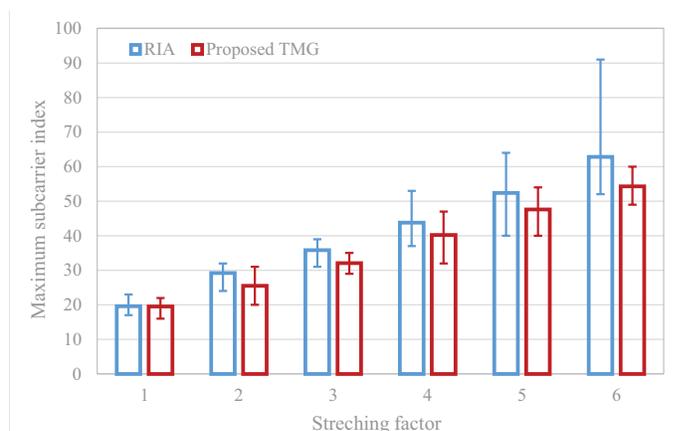
In both Fig. 10 and Fig. 11, it is observed that spectral savings of the proposed TMG scheme increases as the link length is further stretched. Specifically, when the link length exceeds 1050 km, the proposed scheme consumes approximately 10% less spectrum than the RIA in both low- and high-connectivity networks. It is also observed that the TMG heuristic offers higher spectral savings in the low-connectivity network. Importantly, when the average link length exceeds 1750 km, the RIA scheme becomes infeasible in the low-connectivity network as the transmission distance of certain connections exceeds the reach of available modulation schemes. The above limitation can be neutralized by employing regenerators. On the other hand, when the proposed TMG heuristic is used, connections with long transmission distances can be groomed electronically at intermediate nodes, where the transmission reach of the groomed connections can be increased [22].

V. LIMITATIONS AND FUTURE WORK

One limitation of the proposed TMG heuristic is that it splits a connection into equal-sized halves when there is insufficient spectrum along a single path. Instead, if the connection could have been split according to the spare spectrum capacity in alternative paths, MPR would be more effective, which can

TABLE II
MAXIMUM SUBCARRIER INDEX OF THE OPTIMIZED 14-NODE DT
NETWORK WITH DIFFERENT SORTING POLICIES AGAINST DIFFERENT PSD
VALUES

Policy	PSD (in mW/THz)			
	1	2	3	4
(i)	109	93	115	142
(ii)	112	105	118	144
(iii)	106	97	110	140
Used in our work	104	96	114	139

be considered in future work. However, in this study, there are several reasons why splitting a connection asymmetrically would increase the complexity of the problem. First, connections and split connections are assigned multiple modulation formats. Second, when accommodating a new connection in a path, a spectral gap has to be enforced between the new connection and existing connections to avoid nonlinear interferences. Therefore, to successfully split a connection asymmetrically, the heuristic has to select alternative paths among the many possibilities available, decide the modulation format for each alternative path, and predetermine the spectral gap that has to be enforced on each path, which requires comprehensive knowledge of the existing connection. These processes significantly increase the computational complexity of the problem and the running time of the algorithm.

Our proposed scheme primarily reduces bandwidth waste by accurately compensating for the PLI and, secondarily, by carrying out TG and MPR. To reduce the dependency on opaque devices, the TMG heuristic only performs grooming when a connection cannot be transmitted due to excessive impairments. Else, the connections will be served as end-to-end transparent connections. However, if intermediate grooming is performed on all connections to reduce bandwidth waste regardless of impairments, most connections will be assigned the highest modulation format (as multihop grooming would reduce the transmission distance of transparent connections) and groomed at the core routers of many intermediate nodes along their respective paths, mimicking an opaque IP-over-WDM network [27]. While this method would definitely increase spectral efficiency, it would also increase the amount of traffic being processed at the grooming nodes and increase O-E-O conversions, which would elevate network energy consumption [13].

In the proposed TMG heuristic, connections are sorted in decreasing order of their bit rate. Although not presented in the paper, we also considered three additional sorting policies: (i) decreasing order of the number of links on the connections' shortest paths, (ii) decreasing order of their shortest path length, (ii) decreasing order of their bit rate multiplied by the number of links on their shortest path. While the sorting policy considered in this study offered better results compared to policy (i) and (ii), policy (iii) offered marginally higher improvements for certain PSD values. The comparison of results between the considered sorting policy and policies (i)–(iii) are presented in Table II.

At the start of the algorithm, the adaptive SNR margin (ζ) is assigned a small value to ensure that the heuristic does not overestimate interferences from future connections,

and is increased if a single connection does not satisfy the BER requirement at the end of an iteration. When ζ is increased, the interferences among all connections are reduced in the subsequent iteration. As such, the heuristic can find a solution much faster as ζ is increased. However, if we were to isolate the single connection that does not satisfy the BER requirement and increase ζ for that connection only, in the following iteration, the modified connection (with increased ζ) would consume more bandwidth and be shifted further away from existing connections along the links of the designated path. Resultantly, to reduce interferences on future connections and reduce bandwidth waste, the heuristic may select different routing paths, modulation formats, and spectrum slots for future connections. Thus, at the end of the iteration, another connection might fail the BER requirement due to the effect of the modified single connection. If this chain reaction is to continue, we could expect the heuristic to take longer to find a solution.

VI. CONCLUSIONS

Existing RSA schemes incorporating TG and/or MPR have been previously designed and evaluated without considering realistic PLI models. To overcome this limitation, a novel TG- and MPR-enabled impairment-aware RSA scheme was proposed in this study. The primary objective was to increase spectral efficiency, which was formulated as an ILP and a heuristic. Through numerical simulations, we demonstrate that spectral efficiency can be improved beyond that offered by the RIA and existing TR-based RSA schemes by routing connections across multiple routes and aggregating them at intermediate nodes. Repeated tests indicated that the proposed TMG scheme outperforms the RIA scheme and TR-based schemes at all network sizes, assigned PSD values, and traffic volumes. An added advantage of the proposed TMG scheme is that it is able to extend the reach of traffic connections via intermediate grooming without requiring the service of regenerators. Furthermore, the proposed scheme achieves considerable router port savings over the existing schemes, although consuming slightly more transponders.

REFERENCES

[1] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-based optical networks," *Journal of Lightwave Technology*, vol. 29, no. 9, pp. 1354–1366, 2011.

[2] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, and S. Matsuoka, "Spectrum-efficient and scalable elastic optical path network : architecture, benefits, and enabling technologies," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 66–73, 2009.

[3] J. Zhang, Y. Ji, M. Song, X. Yu, J. Zhang, and B. Mukherjee, "Dynamic traffic grooming in sliceable bandwidth-variable transponder-enabled elastic optical networks," *Journal of Lightwave Technology*, vol. 33, no. 1, pp. 183–191, 2015.

[4] Z. Fan, Y. Qiu, and C.-K. Chan, "Dynamic multipath routing with traffic grooming in OFDM-based elastic optical path networks," *Journal of Lightwave Technology*, vol. 33, no. 1, pp. 275–281, 2015.

[5] J. Zhang, Y. Zhao, X. Yu, J. Zhang, M. Song, Y. Ji, and B. Mukherjee, "Energy-efficient traffic grooming in IP-over-elastic optical networks," *Journal of Optical Communications and Networking*, vol. 7, no. 1, pp. A142–A152, 2015.

[6] G. Zhang, M. De Leenheer, and B. Mukherjee, "Optical traffic grooming in OFDM-based elastic optical networks," *Journal of Optical Communications and Networking*, vol. 4, no. 11, pp. B17–B25, 2012.

[7] Y. Wang, J. Zhang, Y. Zhao, J. Zhang, J. Zhao, X. Wang, and W. Gu, "Dynamic spectral defragmentation based on path connectivity in flexible bandwidth networks," in *European Conference and Exhibition on Optical Communication*, Amsterdam, the Netherlands, 2012, p. P5.10.

[8] S. Dahlfort, M. Xia, R. Proietti, and S. J. B. Yoo, "Split spectrum approach to elastic optical networking," in *European Conference and Exhibition on Optical Communication*, 2012, p. Tu.3.D.4.

[9] G. Shen, A. Cai, and L. Peng, "Benefits of sub-band virtual concatenation (VCAT) in CO-OFDM optical networks," in *International Conference on Transparent Optical Networks*, Coventry, England, 2012, pp. 2–5.

[10] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *Journal of Lightwave Technology*, vol. 31, no. 1, pp. 15–22, 2013.

[11] X. Chen, Y. Zhong, and A. Jukan, "Multipath routing in elastic optical networks with distance-adaptive modulation formats," in *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, 2013, pp. 3915–3920.

[12] A. Cai, L. Peng, and G. Shen, "Sub-band virtual concatenation lightpath blocking performance evaluation for CO-OFDM optical networks," in *Asia Communications and Photonics Conference*, Guangzhou, China, 2012, p. AS4D.3.

[13] M. Dharmaweera, R. Parthiban, and Y. Sekercioglu, "Towards a power-efficient backbone network: The state of research," in *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 198–227, 2015.

[14] J. Zhao, H. Wymeersch, and E. Agrell, "Nonlinear impairment aware resource allocation in elastic optical networks," in *Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, USA, 2015, p. M2I.1.

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

[16] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 138–145, 2010.

[17] Z. Ding, Z. Xu, X. Zeng, T. Ma, and F. Yang, "Hybrid routing and spectrum assignment algorithms based on distance-adaptation combined coevolution and heuristics in elastic optical networks," *Optical Engineering*, vol. 53, no. 4, p. 046105, Apr. 2014.

[18] P. Johannisson and E. Agrell, "Modeling of nonlinear signal distortion in fiber-optical networks," *Journal of Lightwave Technology*, vol. 32, no. 23, pp. 4544–4552, 2014.

[19] S. Zhang, C. Martel, and B. Mukherjee, "Dynamic traffic grooming in sliceable bandwidth-variable transponder-enabled elastic optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 1, pp. 4–12, 2013.

[20] T. Tanaka, A. Hirano, and M. Jinno, "Impact of multi-flow transponder on equipment requirements in IP over elastic optical networks," *European Conference and Exhibition on Optical Communication*, 2013, p.We.1.E.3.

[21] T. Tanaka, A. Hirano, and M. Jinno, "Advantages of IP over elastic optical networks using multi-flow transponders from cost and equipment count aspects," *Optics Express*, vol. 22, no. 1, pp. 62–70, Jan. 2014.

[22] A. N. Patel, C. Gao, J. P. Jue, X. Wang, Q. Zhang, P. Palacharla, and T. Naito, "Cost efficient traffic grooming and regenerator placement in impairment-aware optical WDM networks," *Optical Switching and Networking*, vol. 9, no. 3, pp. 225–239, 2012.

[23] L. Yan, E. Agrell, H. Wymeersch, P. Johannisson, R. Taranto, and M. Brandt-Pearce, "Link-level resource allocation for flexible-grid nonlinear fiber-optic communication systems," *IEEE Photonics Technology Letters*, vol. 27, no. 12, pp. 1250–1253, 2015.

[24] M. Angelou, Y. Pointurier, D. Careglio, S. Spadaro, and I. Tomkos, "Optimized monitor placement for accurate QoT assessment in core optical networks," *Journal of Optical Communications and Networking*, vol. 4, no. 1, pp. 15–23, 2012.

[25] X. Wang, M. Brandt-Pearce, and S. Subramaniam, "Impact of wavelength and modulation conversion on transluscent elastic optical networks using mILP," *Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 644–655, 2015.

[26] F. Idzikowski, "Power consumption of network elements in IP over WDM networks," Telecommunication Networks Group (TKN), Berlin, Germany, Tech. Rep., July 2009. [Online]. Available: http://www.tkn.tu-berlin.de/fileadmin/fg112/Papers/powerNumbers_final.pdf

[27] F. Musumeci, M. Tornatore, A. Pattavina, and A. Lord, "A power consumption analysis for IP-over-WDM core network architectures," *Journal of Optical Communications and Networking*, vol. 4, no. 2, pp. 108–117, 2012.

[28] C. Matrakidis, T. Orphanoudakis, A. Stavdas, and A. Lord, "Network optimization exploiting traffic grooming techniques under fixed and elastic spectrum allocation," *European Conference and Exhibition on Optical Communication*, 2014, p. Th.1.2.2.