Performance Comparison of Optical Networks Exploiting Multiple and Extended Bands and Leveraging Reinforcement Learning

Performance Comparison of Optical Networks Exploiting Multiple and Extended Bands and Leveraging Reinforcement Learning

Rasoul Sadeghi*[1], Bruno Correia[1], Elliot London[1], Antonio Napoli[2], Nelson Costa[3], João Pedro[3,4], and Vittorio Curri[1]

[1]*Department of Electronics and Telecommunications (DET) Politecnico di Torino, Corso Duca degli Abruzzi, Torino, Italy*
[2]*Infinera, Munich, Germany*; [3]*Infinera Unipessoal Lda, Rua da Garagem 1, 2790-078 Carnaxide, Portugal*;
[4]*Instituto de Telecomunicações, Instituto Superior Técnico, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*

*rasoul.sadeghi@polito.it

***Abstract*—Exploiting additional low loss bands of optical fibres is a promising solution to expand the capacity of optical transport networks. Recently, extended bandwidth bands (super bands) have been proposed, having a total bandwidth of 6 THz, instead of the regular 4.8 THz. We compare network performance for bands with regular and extended bandwidths when employing transparent and translucent network designs with and without reinforcement learning on the US-NET reference network topology. A total of four MBT scenarios are considered, namely super C, C+L, super C+L, and C+L+S1-band, where S1 denotes half of the S-band bandwidth. We show that the use of super bands and reinforcement learning significantly improves network capacity compared to the use of regular bands and traditional network design methods.**

***Index Terms*—Optical Networks, Multi-band Transmission, Reinforcement Learning, Routing and Wavelength Assignment**

I. Introduction

The ever-increasing volume of IP data traffic requires network service providers to increase capacity in their wavelength-division multiplexing (WDM) networks. Multi-band transmission (MBT) optical networks are being investigated as a cost-effective solution to achieve this objective [1]. With MBT, the full low-loss spectrum of the widely-deployed single-mode ITU-T G.652.D optical fibers is explored, increasing the bandwidth of WDM systems from $\approx$ 4.8 THz (C-band only) up to $\approx$ 50 THz (when considering L- to O-band transmission [2]–[4]). Extending the bandwidth of existing bands (i.e., using super-bands) is another option to increase the available capacity, as an alternative or to complement MBT. Currently, this consists of adding 1.2 THz of bandwidth to each of the traditional C- and L-bands, resulting in a total transmission bandwidth of $\approx$ 6 THz [5]. Furthermore, the available network capacity can also be augmented by exploiting state-of-the-art coherent transceivers (TRXs) to improve spectral efficiency [4], [6] and/or by performing optical signal regeneration at intermediate node(s), i.e., using a translucent network design [7]–[10], where long lightpaths (LPs) are divided into several shorter and more spectral efficient LPs. Although signal regeneration increases network throughput, it does so at the expense of increasing the cost and energy consumption of a network, since additional TRX pairs are required to realize the regenerator function [4]. Several works have investigated how to manage regenerator placement to limit network costs while increasing capacity. For instance, the authors have investigated the planning and deployment of re-amplification, re-shaping, and re-timing (3R) regenerators based on the network and traffic data in [11]. In [12], an optimized regenerator assignment strategy based on the quality of transmission (QoT) of the LPs was proposed. Considering optical networks exploiting MBT, in [4] three QoT-based regenerator placement algorithms were proposed – $General$, $Power\ Optimized$ ($Pow.\ Opt.$), and $Hybrid$ – for a translucent network design. This work showed that performing signal regeneration in a band with a poor QoT such as the S-band can increase network capacity while curbing the increase in cost and energy consumption [3]. On the topic of power consumption, Multi-Source-Agreement (MSA) Open ZR+ TRX have been proposed as a power-efficient and cost-effective TRX [13]. A MSA Open ZR+ TRX supports different modulation formats with a symbol rate of $\approx$ 60 GBaud. Routing and wavelength assignment (RWA) methods play a key role in efficiently exploiting the available network bandwidth. Recently, several machine learning (ML)/reinforcement learning (RL) methods have been proposed [14]–[16] to maximize the allocation of requested services (i.e., reduce blocking), aiming to replace traditional algorithms such as the k-Shortest Path/First-Fit (KSP-FF). Early works suggest that implementing RL techniques can result in better resource management in a network and consequently leads to an increase in the optical network throughput.

In this work, we explore these approaches in MBT systems, considering "super" and "regular" band transmission in transparent and translucent network designs. In Sec. II, the generalized signal to noise ratio (GSNR) as a QoT metric is evaluated in a single span for the investigated MBT scenarios. Moreover, network assessment process with KSP-FF and RL methods are described in this section. Sec. III presents the network simulation results in terms of network throughput as well as link congestion with different percentages for the investigated MBT network scenarios and using KSP-FF and

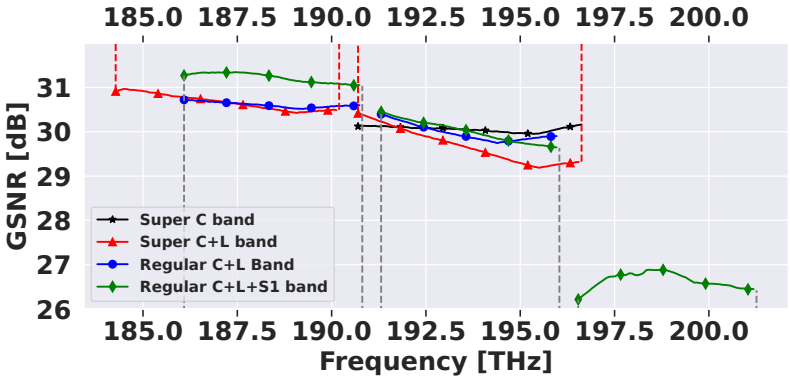


Fig. 1: GSNR value for a single span of 75km in an SSMF.

RL methods.

Network simulations show that not only do super-bands increase capacity, as expected, but that when combined with RL for RWA, a more efficient use of the network infrastructure can be achieved. These findings also highlight that the benefits of employing RL techniques for RWA are likely more significant in MBT networks due to the fact that this optimization problem becomes significantly more complex.

II. QoT Abstraction, Network Assessment and Reinforcement Learning

For all scenarios considered we assume a WDM grid with a spacing of 75 GHz and channels transmitting 64 GBaud signals. Four different scenarios are modelled: (1) super C with 80 channels, (2) regular C+L with 128 channels, (3) super C+L with 160 channels, and (4) regular C+L+S1 with 192 channels. The notation S1 emphasizes that only half of the S-band is used. Noteworthy, the bandwidth of super bands is higher than that of regular bands. More precisely, the bandwidth of super and regular bands is assumed to be 6 and 4.8 THz, respectively. QoT can be estimated using the GSNR, which considers both amplified spontaneous emission (ASE) and nonlinear interference (NLI) generation as noise sources [17], [18]. The NLI is computed using the generalized Gaussian noise (GGN) model, which also accounts for the interaction between the NLI and the stimulated Raman scattering (SRS) effect [19]. In a disaggregated optical network, the total GSNR of a LP, consisting of multiple spans (s), can be computed based on the GSNR of each span in each frequency, i, as expressed in (1).

$$\text{GSNR}_{i,LP} = \frac{1}{\sum_{s \in LP} (\text{GSNR}_{i,s})^{-1}} \,. \tag{1}$$

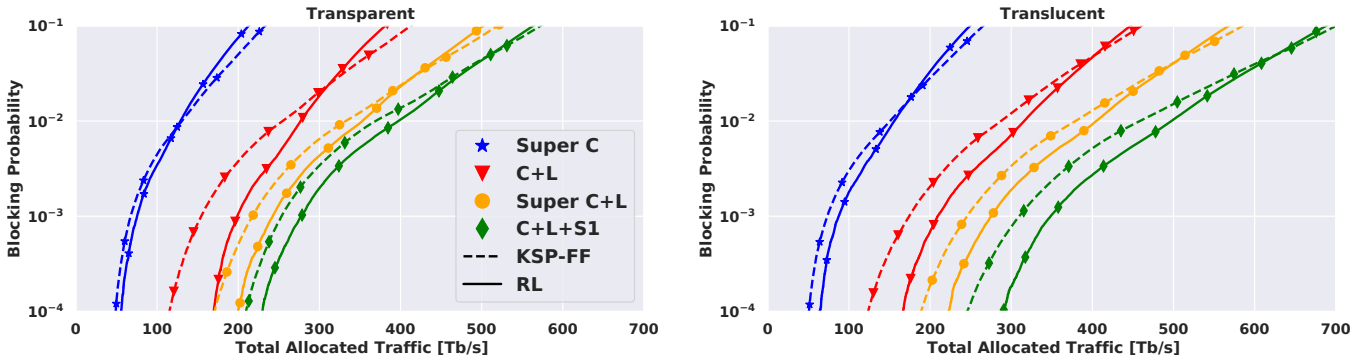
The GSNR value for the investigated MBT scenarios in a single span of 75 km of standard single mode fiber (SSMF) is shown in Fig. 1. To illustrate the width of each band, the frequency limits of regular bands (C, L, and S1-band) is shown via dashed vertical gray lines connecting the corresponding GSNR curves and the bottom horizontal axis, whereas the frequency limits of super bands (super C and L-band) are highlighted by vertical red lines connecting the GSNR curves to the upper horizontal axis.

TABLE I: Average GSNR value for each band in investigated MBT scenarios.

Scenario	L-band	C-band	S1-band
Super C	–	30.0	–
C+L	30.6	29.9	–
Super C+L	30.6	29.6	–
C+L+S1	31.2	30.0	26.6

The average GSNR in each band for all MBT scenarios is summarized in Tab. I. This figure is obtained by averaging the GSNR value over all channels of a given band. According to this table the average GSNR value for the super C-band is 30.0 dB. For the C+L-band scenario the average GSNR values are 29.9 and 30.6 dB in the C- and L-band, respectively, which changes to 29.6 (C-band) and 30.6 dB (L-band) in the super C+L-band scenario. In the C+L+S1-band configuration the average GSNR values are 30.0, 31.2 and 26.6 dB in the C-, L- and S1-band, respectively. Note that the noise figure (NF) assumed for the optical amplifiers is the same as in [4], with the NF profile interpolated for the appropriate frequency ranges for the super band scenarios. As expected, adding more bands/channels to the transmission system may lead to performance degradation, which can be observed in Fig. 1 for the case of an individual fiber span.

Network assessment is carried out using the Statistical Network Assessment Process (SNAP) framework [2], [3], considering the US-NET network topology, which consists of 24 ROADM nodes and 43 links, with an average link length of 971 km and an average nodal degree of 3.58 [4]. The analysis considers progressively loading the network with 100 Gb/s traffic requests. For the transparent network design, the most spectral efficient modulation format among the ones that allow to bridge the source and destination nodes is selected. We assume usage of MSA OpenZR+ TRX [13], which supports 16QAM, 8QAM, and QPSK dual-polarization modulation formats. Each of these modulation formats has



(a) (b)

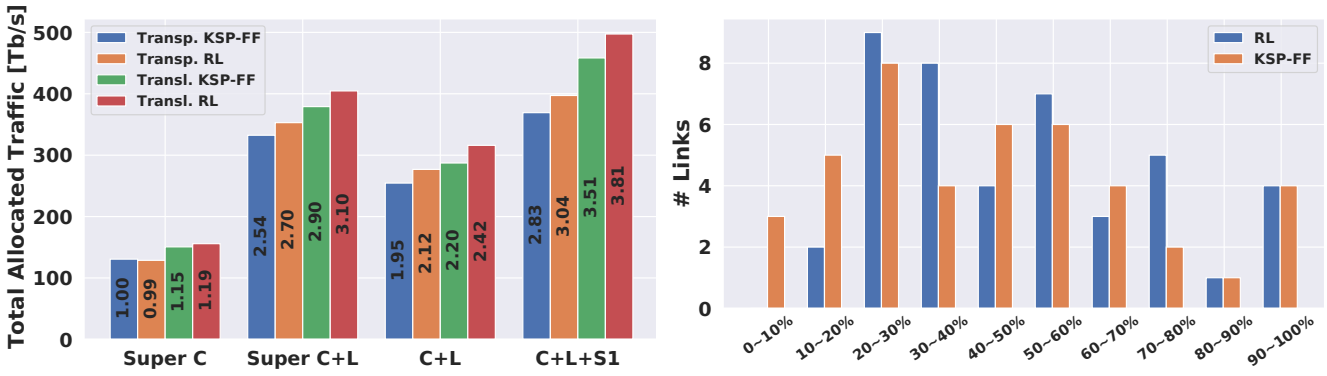
Fig. 2: Total allocated traffic as a function of blocking probability for the US-NET topology, using the KSP-FF or RL methods, and considering (a) transparent and (b) translucent network designs.

a minimum required GSNR (RGSNR) [20], which must be observed to allow their utilization over a given LP. For the translucent network design, a regenerator assignment is used, corresponding to "$Algorithm$ 1" in [4]. This algorithm divides the candidate LP into the minimum number of sub-paths that allow using the most spectral efficient modulation format, without considering wavelength conversion after each signal regeneration. In this work, two different RWA strategies are also employed and their performance compared. Firstly, we considered a k-shortest path (assuming $k_{max} = 5$) routing algorithm and a first-fit wavelength assignment policy (KSP-FF). The KSP-FF strategy allocates the request to the first available channel of the shortest available path, following a best-effort approach. Secondly, we considered a RL approach, with the key components: (i) environment, (ii) action space, (iii) observation space, and (iv) reward. For the environment, the SNAP framework is used and the KSP-FF method is replaced with an RL agent. In the action space (A), the RL agent selects a path among the k candidate paths (for fairness of comparison, the same total of $k_{max} = 5$ possibilities is considered). The observation space is an array defined as $f = \{s, d, \{\text{QoT}_1, ch_1, n_1, G_1\}, ..., \{\text{QoT}_k, ch_k, n_k, G_k\}\}$ for all $k \in K$, where s and d are the source and destination nodes, respectively, in the one-hot format. QoT_k is the QoT of the candidate LP, ch_k and n_k are the number of common free channels and number of intermediate nodes in the candidate LP, respectively, whereas G_k represents the available free capacity in existing LPs with the same source and destination nodes. When there is free capacity from already allocated LPs between the same source and destination G_k equals 1, otherwise it is set to 0. To be more clear regarding G_k, if an LP establishes for a first time, G_k=0, and if its QoT supports 16QAM (400 Gb/s) – the size of used traffic in this work is 100 Gb/s –, the vacant capacity of this LP, 300 Gb/s, can be used for the next requests. In that case for the same source and destination, G_k value in the observation space considered equal to 1 which improves its probability to establish among other paths. All values of the observation space, f, which have different magnitudes are normalized between -1 and 1 before being fed to the deep neural network (DNN). In RL, an agent takes an action every time step (following a policy) in an environment in order to maximize the expected cumulative reward.

For RL, the shaping of a reward function plays a crucial role in an agent's learning, and so RL gives reshaped rewards for an action taken [21]. The reward returned by the software-defined network (SDN) unit is 1 if the request is allocated to a new LP. As we wish to better utilize existing LPs, a reward of 10 is returned if the request can be allocated to free capacity in existing LPs. A reward of -1 is instead returned if the request is blocked. The RL hyperparameters were set to 0.95, 10^{-5}, 50, 2 and 512 for discount factor, learning rate, batch size, number of hidden layers and neurons set, respectively. Moreover, for the training state the "$Adam$" algorithm [22] is used. When using RL, first the state array, f_t, is collected for an arriving connection request, t. The DNN unit [14] reads this data, generating a policy, $\pi_t(A|f_t, \theta)$, which is a set containing the probability of each action, $a_t \in A$, and where θ is the DNN value between the source and destination. Then, according to the probabilities of π_t, the SDN unit takes an action, a_t, to attempt to establish the LP with the corresponding path. After taking an action, the SDN unit returns a reward, and stores the state, the reward, and the action in a buffer, as an array, for the training section. Since the objective of RL methods is to maximize long-term cumulative rewards, RL attempts to improve this factor. The SNAP framework is executed using the KSP-FF and RL methods for each iteration/epoch, until the specific threshold blocking probability (BP) to stop loading the network is achieved (set to 20%).

III. Network Simulation Results and Discussion

This section presents and discusses the network performance results for the KSP-FF and RL methods using transparent and translucent designs for all transmission scenarios considered. Fig. 2 shows the total allocated traffic in the BP range between 10^{-4} and 10^{-1} for the MBT scenarios



(a) (b)

Fig. 3: (a) Multiplicative factor of the total allocated traffic for all investigated scenarios at BP=1%, (b) link congestion for KSP-FF and RL approaches in the translucent C+L+S1 network design, for a delivered traffic of 400 Tb/s.

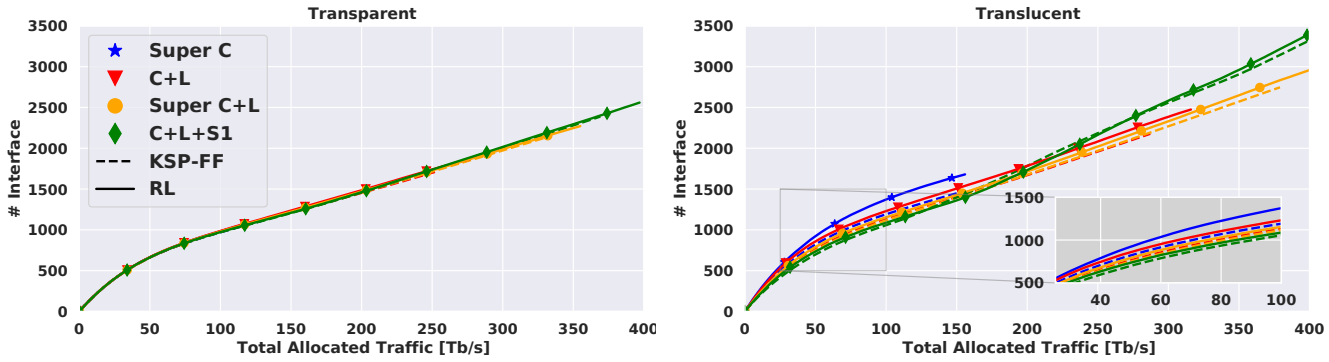
considered, using the KSP-FF (dashed) and RL (solid) RWA approaches. Fig. 2a reports the results of the transparent network design, showing that in the case of super C-band there is only minor capacity improvement from using RL instead of the KSP-FF algorithm, for small BPs (BP$<$1%). As expected, the capacity in the C+L-band scenario is higher than that for the super C-band. In the former MBT scenario, the delivered network capacity when using the RL algorithm is 277 Tb/s (versus 254 Tb/s with KSP-FF) for a target BP of 1%. Moreover, for the super C+L-band, network throughput increases by ≈80 Tb/s in comparison to the regular C+L-band scenario. For instance, the delivered capacity by exploiting super C+L-band is 353 and 332 Tb/s with the RL and KSP-FF methods, respectively. Finally, the total allocated traffic in the C+L+S1-band scenario is 397 and 369 Tb/s when using the RL and KSP-FF algorithms, respectively. A key finding is that RL achieves almost the same capacity in the super C+L MBT scenario, compared to the scenario utilizing the KSP-FF algorithm and the C+L+S1-band, for small target BP. This suggests that the benefit of using RL tends to increase proportionally to the number of channels available per fiber. This is due to the ability of RL to better exploit a more complex solution space, when compared to the simpler KSP-FF approach.

The results plotted in Fig. 2b, which are obtained when considering a translucent network design, are similar to those in Fig. 2a. However, two key differences can be observed when comparing both figures. Firstly, when considering the same combination of MBT scenario and RWA algorithm, the translucent network design results in lower blocking, i.e., higher network capacity for a given target BP. This is a consequence of enforcing more spectrally efficient modulation formats by using shorter LPs. Secondly, the network capacity difference between the RL and KSP-FF methods increases for all transmission scenarios. This can be explained by the ability of the RL algorithm to better exploit the available capacity in the larger number of deployed LPs.

From Fig. 2b, the traffic allocated in the super C-band, when using the KSP-FF and RL methods, is 150 and 156 Tb/s for a target BP of 1%, respectively. Deploying a C+L-band system leads to a total allocated traffic of 287 and 316 Tb/s when using KSP-FF and RL, respectively. Moreover, deploying the extended bandwidth configuration (super C+L-band) increases the network throughput to 379 and 404 Tb/s. Finally, these figures are further raised to 458 and 497 Tb/s in the C+L+S1-band scenario, again when using KSP-FF and RL, respectively.

A more concise representation of the differences in network capacity is shown in Fig. 3a, which presents the network capacity multiplicative factor (MF) for all scenarios at the BP of 1%, using the transparent super C-band with KSP-FF algorithm as the reference case (i.e., with MF = 1). It can be seen that the super C-band increases the network capacity $\times$1.19 in the best case (translucent network design with the RL method). Although these improvements are not large, they become more relevant when considering additional bands. For instance, considering transparent network design and KSP-FF, it can be observed that the network capacity increases $\times$2.54 when instead using the super C+L-band, albeit the amount of spectrum made available increased only $\times$2, and the impact of SRS is higher. Adopting a translucent network design and RL for RWA, the increase in capacity is more than threefold. It is also interesting to note that the network capacity obtained in this configuration (super C+L-band, translucent design, RL) is higher than that reported with C+L+S1-band using a transparent design and also RL. This occurs even though the latter MBT scenario exploits a wider spectrum (approximately 15.6 versus 12 THz). In this case, there is a trade-off in terms of number of optical amplifiers and TRXs: fewer amplifiers are required with super C+L than with C+L+S1 (one third less), but more TRXs are used to enforce the translucent design.

In order to gain further insight on the how RL outperforms KSP-FF, link congestion is shown in Fig. 3b for the translucent C+L+S1-band network design for a delivered traffic of 400 Tb/s. This figure shows that RL makes use of the network links more evenly compared to KSP-FF. For example, considering less congested links, i.e., with loading that ranges



(a) (b)

Fig. 4: The number of demanded interfaces versus total allocated traffic for both KSP-FF and RL methods for (a) transparent and (b) translucent network designs in the BP$\leq$1% range.

from 0 to 20%. There are 8 links with this load range with KSP-FF, whereas with RL there are only 2 links. Considering links with medium congestion, defined here as varying from 20% to 80% of link occupation, the number of links present are higher when using RL than KSP-FF. In the higher congestion range, the number of links are the same, indicating a particular characteristic of the investigated topology. Overall, it can be seen that RL is capable to more effectively use the link capacity, which translates into lower blocking and higher traffic load for a given BP.

Fig. 4 shows the number of demanded interfaces versus total allocated traffic when using KSP-FF and RL for each network design and scenario for the BP range of $\leq$1%. Considering the transparent network design, given in Fig. 4a, it is visible that the number of demanded interfaces for all scenarios is almost equal for both algorithms, and exploiting more bands permits a higher traffic allocation, and correspondingly a higher number of demanded interfaces. We however remark that adding extra bands requires extra amplifiers for the newly exploited bands. For instance, the number of interfaces used to deliver 100 Tb/s capacity is about 1000, however to double this capacity the number of demanded interfaces is approximately 1500. Additionally, the maximum demanded interface count is for the C+L+S1-band scenario, which is approximately 2500 interfaces at the delivered capacity of 400 Tb/s. Furthermore, the delivered capacity by deploying a super C+L-band with the RL method delivers almost the same capacity as a C+L+S1-band scenario using KSP-FF (refer to Fig. 2a); however, the number of deployed amplifiers in the C+L+S1-band is $\times$1/3 higher than super C+L-band scenario.

Considering the translucent network design, given in Fig. 4b, it is visible that, in general, the number of required interfaces for the same amount of traffic is higher than for the transparent network design. For example, in the transparent network design, the required number of interfaces to achieve 100 Tb/s network capacity for the super C-band is approximately 1000 for both the KSP-FF and RL methods. In contrast, in the translucent network design this value increases to 1200 and 1373 by using KSP-FF and RL methods, respectively. Focusing on the C+L+S1-band scenario, when the network is first loaded we observe that the number of demanded interfaces is less than the other MBT scenarios, as the first loaded channels have a higher QoT. However, the difference between the number of used interfaces between the KSP-FF and RL implementations is smaller, as a larger number of channels/bands are utilized. Additionally, for the C+L+S1-band scenario, after the initial loading period is finished, at approximately 150 Tb/s, it is visible that the number of required interfaces progressively increases. This is due to the last allocated channels lying within the S1-band, which have a much lower QoT than those within other bands, consequently requiring a larger number of regenerators.

Focusing on the difference between the two algorithms for any total allocated traffic value, it is visible that this difference decreases proportionally to the number of channels. This because the RL algorithm makes better use of a larger number of available channels and any vacant capacity of already-established LPs – this is also visible in the C+L+S1-band scenario in Fig. 2b, where the RL algorithm provides a larger capacity than the KSP-FF implementation.

IV. Conclusion

In this work, we have compared network capacity and link congestion between bands with regular and extended bandwidths in transparent and translucent network designs. We have shown that super bands provide a significant capacity improvement compared to their regular counterparts, despite the increased impact of SRS. Importantly, a super C+L-band translucent network design can provide a higher level of delivered traffic, compared to the regular C+L+S1-band with a transparent network design, all the while requiring 33% less amplifiers than the latter. We have found that the benefits of using RL instead of KSP-FF are more pronounced in cases where more channels are available.

ACKNOWLEDGMENT

This work was partially funded by the EU H2020 within the ETN WON, grant agreement 814276 and by the Telecom Infra Project. A. Napoli, N. Costa, and J. Pedro thank the H2020 B5G-OPEN (G.A. 101016663) for partially funding their activities.

REFERENCES

[1] F. Hamaoka, M. Nakamura, S. Okamoto, K. Minoguchi, T. Sasai, A. Matsushita, E. Yamazaki, and Y. Kisaka, "Ultra-wideband wdm transmission in s-, c-, and l-bands using signal power optimization scheme," Journal of Lightwave Technology, vol. 37, no. 8, pp. 1764–1771, 2019.

[2] A. Ferrari, E. Virgillito, and V. Curri, "Band-division vs. space-division multiplexing: a network performance statistical assessment," Journal of Lightwave Technology, pp. 1–1, 2020.

[3] R. Sadeghi, B. Correia, E. Virgilito, E. London, N. Costa, J. Pedro, A. Napoli, and V. Curri, "Optimized translucent s-band transmission in multi-band optical networks," in 2021 European Conference on Optical Communication (ECOC), pp. 1–4, 2021.

[4] R. Sadeghi, B. Correia, A. Souza, N. Costa, J. Pedro, A. Napoli, and V. Curri, "Transparent vs translucent multi-band optical networking: Capacity and energy analyses," J. Lightwave Technol., vol. 40, pp. 3486–3498, Jun 2022.

[5] J. Pedro, N. Costa, and S. Sanders, "Cost-effective strategies to scale the capacity of regional optical transport networks," Journal of Optical Communications and Networking, vol. 14, no. 2, pp. A154–A165, 2022.

[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] G. Shen and R. S. Tucker, "Translucent optical networks: the way forward [topics in optical communications]," IEEE Communications Magazine, vol. 45, no. 2, pp. 48–54, 2007.

[8] E. Karasan and M. Arisoylu, "Design of translucent optical networks: Partitioning and restoration," Photonic Network Communications, vol. 8, no. 2, pp. 209–221, 2004.

[9] G. Shen, W. Grover, T. Cheng, and S. Bose, "Sparse placement of electronic switching nodes for low blocking in translucent optical networks," Journal of Optical Networking, vol. 1, no. 12, pp. 424–441, 2002.

[10] G. Shen and W. Grover, "Segment-based approaches to survivable translucent network design under various ultra-long-haul system reach capabilities," Journal of Optical Networking, vol. 3, no. 1, pp. 1–24, 2004.

[11] J. Pedro, "Predeployment of regenerators for fast service provisioning in DWDM transport networks [invited]," J. Opt. Commun. Netw., vol. 7, pp. A190–A199, Feb 2015.

[12] M. Kanj, E. L. Rouzic, J. Meuric, and B. Cousin, "Optical power control in translucent flexible optical networks with GMPLS control plane," J. Opt. Commun. Netw., vol. 10, pp. 760–772, Sep 2018.

[13] "Open ZR+ MSA Technical Specification." https://openzrplus.org/site/assets/files/1075/openzrplus$_1p0.pdf$.

[14] X. Chen, B. Li, R. Proietti, H. Lu, Z. Zhu, and S. B. Yoo, "Deeprmsa: a deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks," Journal of Lightwave Technology, vol. 37, no. 16, pp. 4155–4163, 2019.

[15] J. W. Nevin, S. Nallaperuma, N. A. Shevchenko, Z. Shabka, G. Zervas, and S. J. Savory, "Techniques for applying reinforcement learning to routing and wavelength assignment problems in optical fiber communication networks," Journal of Optical Communications and Networking, vol. 14, no. 9, pp. 733–748, 2022.

[16] N. Di Cicco, E. F. Mercan, O. Karandin, O. Ayoub, S. Troia, F. Musumeci, and M. Tornatore, "On deep reinforcement learning for static routing and wavelength assignment," IEEE Journal of Selected Topics in Quantum Electronics, vol. 28, no. 4: Mach. Learn. in Photon. Commun. and Meas. Syst., pp. 1–12, 2022.

[17] 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," Journal of Lightwave Technology, vol. 36, pp. 3073–3082, Aug 2018.

[18] V. Curri, "Gnpy model of the physical layer for open and disaggregated optical networking," J. Opt. Commun. Netw., vol. 14, pp. C92–C104, Jun 2022.

[19] M. Cantono, D. Pilori, A. Ferrari, C. Catanese, J. Thouras, J.-L. Augé, and V. Curri, "On the interplay of nonlinear interference generation with stimulated Raman scattering for QoT estimation," Journal of Lightwave Technology, vol. 36, no. 15, pp. 3131–3141, 2018.

[20] J. Pedro and S. Pato, "Capacity increase and hardware savings in DWDM networks exploiting next-generation optical line interfaces," in 2018 20th International Conference on Transparent Optical Networks (ICTON), pp. 1–6, 2018.

[21] A. Dayal, L. R. Cenkeramaddi, and A. Jha, "Reward criteria impact on the performance of reinforcement learning agent for autonomous navigation," Applied Soft Computing, vol. 126, p. 109241, 2022.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.