Optical Networking Gym: an open-source toolkit for resource assignment problems in optical networks

Citation for the original published paper (version of record):
Natalino Da Silva, C., Magalhaes, T., Arpanaei, F. et al (2024). Optical Networking Gym: an open-source toolkit for resource assignment problems in optical networks. Journal of Optical Communications and Networking, 16(12): G40-G51.
http://dx.doi.org/10.1364/JOCN.532850

Optical Networking Gym: An Open-Source Toolkit for Resource Assignment Problems in Optical Networks

CARLOS NATALINO [1,*], TALLES MAGALHÃES [2], FARHAD ARPANAEI [3,1], FABRICIO R. L. LOBATO [2], JOÃO C. W. A. COSTA [2], JOSÉ ALBERTO HERNÁNDEZ [3], AND PAOLO MONTI [1]

[1] Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden.
[2] Federal University of Pará (UFPA), Belém, Pará, Brazil.
[3] Department of Telematic Engineering, Universidad Carlos III de Madrid (UC3M), Leganés, Madrid, 28911 Spain.
[*] carlos.natalino@chalmers.se

Compiled December 9, 2024

The dynamic provisioning of optical network services requires algorithms to find a suitable solution given the specific service requirements and the current network state. These algorithms are usually evaluated using a software simulator developed ad hoc, which may require different levels of detail depending on the problem addressed and how realistic the evaluation needs to be. Moreover, to demonstrate they are a significant contribution to the field, these new algorithms must be benchmarked against the best-performing previously-proposed solutions. Due to the large set of parameters and their wide range of possible values, benchmarking algorithms from the literature is not straightforward and can quickly become challenging and time-consuming. This work introduces the *Optical Networking Gym*, an open-source toolkit that simplifies implementing optical resource assignment simulations and benchmarking new solutions against previously-published algorithms. The toolkit provides environments modeling relevant optical networking scenarios, common algorithms for solving problems related to these scenarios, and a set of scripts to prepare and execute simulations for various use cases. Currently, four environments are available, with the possibility of increasing this number through contributions from the co-authors and the community. This paper describes the architecture, interface, environments, and scripts included with the toolkit. We adopt the Quality of Transmission (QoT)-aware dynamic resource allocation of optical services as the network scenario under exam. Three use cases highlight the toolkit's modularity, flexibility, and performance. The toolkit allows researchers to streamline the process of developing simulation scenarios and algorithms, enhancing their ability to benchmark their algorithms.

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

Resource assignment problems in optical networks have been studied for several decades. In particular, the dynamic provisioning of optical services has been studied since the dawn of wavelength division multiplexing (WDM) optical networks [1]. Despite this, as new optical technologies and architectures are developed, the research on resource assignment in optical networks shows no signs of slowing down or diminishing relevance.

Currently, research on optical network resource allocation problems is increasingly relevant, but also challenging and time consuming due to a number of factors. Firstly, optical networks are widely adopted in various scenarios, e.g., core, metro, edge, datacenter, and access networks. Secondly, with the growing relevance of physical layer impairments, a plethora of different analytical models for QoT estimation are available depending on the specific assumptions of the physical characteristics of the network. Thirdly, building a new simulator requires a number of critical decisions, such as the programming language, architecture, and libraries to use. Finally, this problem is aggravated by the numerous solutions in the literature that are potentially suitable for being benchmarked against the newly-proposed solution. As a result, new solutions are usually evaluated against a few arbitrary baseline works.

Several tools that address part of these complexities have been published over the past years. For instance, network design tools have been made available [2, 3], but they focus on the static planning, not suitable for the dynamic simulation of optical networks. Simulators focusing on resource assignment problems have been published [4, 5], but these are not open source, limiting its applicability. A similar observation can be made about digital twins [6], where the great majority is not open source. Open-source models focusing on physical layer impairment computation have been published [7, 8], but the simulation of a dynamic optical network is outside of their scope due to the substantially different business logic and concerns involved in discrete event simulation. Optical network emulators such as [9] focus on the software-defined networking (SDN) aspects of operating optical networks, but incur a substantial signaling overhead that does not add value to certain use cases. Two of the co-authors of this paper proposed the Optical RL-Gym [10], a toolkit inspired by the OpenAI Gym for facilitating the application of reinforcement learning (RL) and deep reinforcement learning (DRL) to solve resource assignment problems. The tool has influenced research related to routing and wavelength assignment (RWA) algorithms [11], and other related tools [12]. The tool has been extended for multi-band [13, 14] and defragmentation [15] problems. Other approaches to take advantage of graphic processing units (GPUs) for executing such environments have also been proposed [16]. The popularity reached by RL-focused tools indicates that the approach of following well-known interfaces from the machine learning community is promising. However, RL-focused tools do not cover the use cases where traditional heuristics are involved.

Another important research topic over the past few years has been the application of artificial intelligence & machine learning (AI/ML) models for the estimation of QoT of unestablished lightpaths [17, 18]. This research requires a substantial and accurately collected dataset, which can be very challenging and expensive to collect. To address this issue, open datasets have been published recently [19, 20]. However, these datasets have a set number of lightpath information included, limiting the possibilities to innovate by adopting other pieces of information. Ideally, researchers should be able to generate their own datasets, selecting from a wide range of information that can be included, and controlling the network scenario such as topology, load, and resource assignment algorithm.

In this paper, we introduce the Optical Networking Gym (ONG), an open-source toolkit that takes advantage of well-known RL interfaces while extending them for use in a wide set of use cases in optical networks. This results in a flexible and extensible toolkit that lowers the barrier of entry for new researchers in the field, and makes it more feasible to benchmark the performance of new algorithms against a wide range of solutions from the literature. The toolkit includes a set of interfaces, environments, and scripts that streamline the development and execution of new optical network resource assignment scenarios and use cases. More specifically, the ONG takes inspiration from Optical RL-Gym [10], and extends it in the following ways:

- The architecture is updated with the latest developments from the RL research community, as described in Sec. 2;
- The updated interface makes it more modular, simplifying the development of new environments and algorithms, as described in Sec. 2;
- The standard Python implementation is combined with Cython, allowing for critical parts of the source-code to be compiled and speeding up the execution of simulations, as described in Sec. 2;
- The use of ONG with standard (non-RL) algorithms is simplified, as described in Sec. 3;
- A new environment for the QoT-aware dynamic routing, modulation format, and spectrum assignment (RMSA) problem is provided, implementing an state-of-the-art physical layer impairment model, as described in Sec. 3.

In addition to describing the properties and contributions of the ONG, this paper analyzes its use in three relevant use cases. The first use case analyzes the use of ONG for global launch power optimization, where we aim to find the launch power that results in the highest number of optical services deployed in the network. This use case is essential for QoT-aware RMSA studies, due to the strong influence of the launch power on the QoT, and its dependency on physical layer parameters such as span length. The second use case is related to the assessment and benchmarking of QoT-aware RMSA algorithms. This use case represents one of the most typical benchmarking needed when proposing a new resource allocation algorithm. The third use case illustrates how ONG can be used to create custom datasets for AI/ML QoT estimation research. This use case allows researchers to expand the set of information analyzed, and easily include the generated datasets as complementary material to their papers. The results show that the proposed ONG can properly tackle the use cases, enabling the detailed analysis of network performance and optical parameters.

The rest of the paper is organized as follows. Sec. 2 introduces the ONG toolkit, discussing its architecture and common interface. Sec. 3 discusses the scenario and environment selected to showcase the functionalities and use cases of ONG. The use cases are described in detail, and a few performance metrics are illustrated in Sec. 4. A few open challenges and for future work are discussed in Sec. 5 Sec. 6 concludes the paper.

2. THE OPTICAL NETWORKING GYM (ONG)

The ONG is a toolkit whose main contribution is the definition of a common set of software interfaces that enable the modular development of scenarios and use cases. Due to its modular interfaces, the toolkit enables researchers to focus on the specific challenge while taking advantage of standard implementations and algorithms from the literature. The toolkit builds upon the Optical RL-Gym [10], updating and enhancing its functionalities. The toolkit is implemented using modern Python, making it suitable for integrating popular frameworks and other languages such as C/C++/MatLab. One of the improvements compared to the Optical RL-Gym is the reduction of inheritance in favor of composition, which generally improves the clarity and locality of the code, simplifying its understanding and extension.

Another critical change compared to the Optical RL-Gym is related to the adoption of the Cython language, which works as an extension of Python. Cython has a syntax that resembles Python, and can have files that mix Python and Cython. The code is compiled to obtain more performance, but runs well in conjunction with Python as a compiled module. In this way, parts of the code where performance is critical can be seamlessly integrated with other parts of the code where the simplicity of Python is preferred. Moreover, researchers that use ONG can select in which language they prefer to write their extensions. The source code of the ONG can be found in [21].

A. Environments

In the context of RL, an environment represents a real-world system from which an intelligent agent needs to learn from. ONG defines a set of environments that correspond to network scenarios where common resource assignment problems arise. The defined environments can be used to train RL agents. However, this paper highlights the fact that these environments can also be useful for benchmarking strategies and algorithms that not necessarily are built using RL.

An environment is commonly comprised of a topology, a set of supported modulation formats, a traffic generator, a list of current and past services served by the network, and a random number generator (RNG). The topology represents the inventory of nodes, edges, spans, paths, and spectrum resources. The current occupation state of each resource is also tracked by the environment.

In its current version, the ONG provides the following four environments:

- Routing and wavelength assignment (RWA): Environment that models a WDM optical network where each service has one wavelength of fixed bandwidth which needs to meet the spectrum continuity constraint.
- Routing, modulation format, and spectrum assignment (RMSA): Environment that models an elastic optical network (EON) where each service uses an arbitrary number of subcarriers, i.e., the bandwidth of a channel may vary, and the channel needs to meet the spectrum continuity and contiguity constraints.
- Routing, core, modulation format, and spectrum assignment (RCMSA): Environment that models a multi-core EON where the fiber core also needs to be selected for the provisioning service, and such core may be fixed or variable along the path, respecting the spectrum continuity and contiguity constraints.
- DeepRMSA: Similar to the RMSA environment, but adopting the state, action, and reward representation defined in [22].

B. Interfaces

The ONG defines the interface meant for facilitating the interaction between the environment and the algorithm (e.g., heuristic, DRL) that will solve the problem. Listing 1 shows a simplified view of the interface of an ONG environment using a simplified Python syntax. The environment follows the same interface and inherits from the generic environment as in the *gymnasium* library [23], which is a maintained successor to the initial OpenAI Gym [24]. While the Open AI Gym interface was initially developed for the training of RL agents, it has proven to be suitable in non-RL simulations too. The environment divides a sequence of interactions as episodes. The *initialization* method receives the environment configuration parameters as an argument, where the topology, modulation formats, and traffic generation variables are commonly defined.

The *step* method is the main method, where an action is received from the solution algorithm, implemented in the optical network representation, and its result is returned. For instance, in the case of an RMSA problem, the action represents a particular route, modulation format, and spectrum to provision the path. In the case of a spectrum defragmentation problem, the action determines which lightpath to defragment and how to do it. Once the action is received, the *step* method prepares the environment for the next action by advancing the network state until the point where an action is needed. Then, the method returns a tuple with five elements. The first element is the observation, which exposes the current network state to the algorithm used to solve the problem. The reward is a numerical quantification of the quality of the received action, helpful in DRL algorithms, but that can be useful for other algorithms. The next element specifies whether or not the current episode has terminated, e.g., if a pre-defined number of requests have been processed. The fourth element specifies whether or not the current episode has been truncated, i.e., it cannot continue and has not reached the desired number of interactions. This is useful, for instance, in scenarios where incremental traffic is generated, i.e., optical network services do not leave the network and the episode finishes when traffic can no longer be accommodated [11]. Finally, the last element is a dictionary containing custom information defined by the environment, providing flexibility for various scenarios. For instance, in the RMSA scenario described in the next section, relevant information are the request and bit rate blocking ratios, overall and in the current episode. In general, metrics such as spectrum usage, spectrum efficiency, among others can also be returned from the environment through this dictionary.

The *reset* method is used when starting to run the simulations or when running a new episode. It allows the specification of a (new) seed for the random number generator. The method returns the current environment state and the dictionary with custom information. The *render* method generates a graphical representation of the current state of the environment. In the current implementation, it generates a matrix visualization for the spectrum usage of the spectral resources versus the links in the network. Finally, the *close* method releases any resources that might be used by the environment, such as a file to write intermediate or final results. Note that the environment is meant to be executed in a single central processing unit (CPU). However, multiple environment instances can be executed in parallel using various CPUs with near-to-linear performance scaling.

Listing 2 shows a simple use of the environment interface to execute a series of actions governed by a hypothetical algorithm. The environment is initialized with a set of keyword arguments (line 1). Usual arguments that can be mentioned are: network topology, traffic generator parameters, per-step results file, and number of steps per episode. The algorithm to be used to solve the environment is initialized with a set of keyword arguments (line 2). Simple heuristics such as the ones discussed in Sec. 3.B do not have configuration parameters, but more elaborated ones may require for instance the setting of thresholds. The environment has its seed for random number generation reset (line 3). Then, the environment is executed for a pre-defined number of episodes (line 4). Episodes are a well-known concept in the RL field, where the RL agent is trained at the end of each episode [25]. However, using episodes to divide the simulation into blocks is also interesting in optical networking simulations. At the end of each episode, preliminary/intermediate statistics can be extracted from the environment through, e.g., the *info* dictionary, and used to build a time-series of relevant metrics. One example is the analysis of how the blocking ratio evolves over time. Another example enabled by episodes is the implementation of changes in the traffic characteristics (e.g., uniform vs. non-uniform) during a simulation [15, 22].

Within each episode, the algorithm is invoked, receiving the current environment state, and returning the action to be taken (line 6). The environment is invoked receiving the action selected by the algorithm, and returning a tuple with the five elements discussed before (line 7–8). The steps in lines 6–8 are repeated until the episode is terminated or truncated (line 5). Once an episode is finished, the environment is reset (line 9). The environment is closed once the pre-defined number of episodes is reached (line 10). Note that listing 2 defines a clear interface for

Listing 1. Interface of an Optical Networking Gym (ONG) environment.

```
from typing import Any, SupportsFloat

import gymnasium as gym

class OpticalEnv(gym.Env[ObsType, ActType]):

    def __init__(self, **kwargs):
        # initialization

    def step(self, action: ActType) -> tuple[ObsType, SupportsFloat, bool, bool, dict[str, Any]]:
        # implementation
        return observation, reward, terminated, truncated, info

    def reset(self,
        *,
        seed: int | None = None,
        options: dict[str, Any] | None = None,
    ) -> tuple[ObsType, dict[str, Any]]:
        # implementation
        return observation, info

    def render(self) -> RenderFrame | list[RenderFrame] | None:
        # implementation

    def close(self) -> None:
        # implementation
```

Listing 2. Illustrative use of the adopted environment interface.

```
env = OpticalEnv(env_kwargs)
algo = Algorithm(alg_kwargs)
obs, info = env.reset(seed=42)
for episode in range(1000):
    while not (term or trunc):
        action = algo.solution(obs)
        obs, rew, term, trunc, info \
            = env.step(action)
    obs, info = env.reset()
env.close()
```

the algorithms to be used, which receive the environment state and return the action to be taken.

C. Functionalities

The ONG implements a set of functionalities that simplify the execution of optical resource assignment problems. For instance, the ONG implements a topology creation script that enables researchers to load network topologies from popular datasets such as SNDlib [26], Topology Zoo [27], or simple ASCII files. Such script also allows the user to control how fiber spans are calculated within an optical link, and their individual physical layer parameters. Finally, a set of scripts shows examples of popular use cases for which the ONG is suitable.

The ONG toolkit provides several other artifacts that are not detailed in this section due to space constraints, but are worth being briefly mentioned in the following. The ONG toolkit:

- takes advantage of a modern Python implementation using the latest developments and best practices in the language;
- integrates a set of relevant libraries relevant to the tasks involved in simulating optical networks, taking advantage of developments by the open-source community;
- provides comprehensive documentation, not only through detailed comments in the open-source code, but also with tutorials;
- provides a set of scripts that can be used to execute simulations using the provided environments and algorithms, including the ones shown in Sec. 4;
- provides a set of scripts for performing statistical analysis over the results obtained from simulations;
- establishes good practices on how to perform simulations, extract statistics in the network, and persist them in interoperable file formats;
- accepts contributions from the community through the usual open-source process, i.e., discussion about potential functionalities followed by pull request.

3. QOT-AWARE DYNAMIC RMSA ENVIRONMENT

This section describes the QoT-aware dynamic RMSA environment available within ONG. The environment models the QoT-aware dynamic provisioning of optical network service requests, hereinafter called *requests*. Upon the request provisioning, the associated optical network service is realized by a lightpath in an EON. Before provisioning a request, the RMSA problem needs to be solved.

Fig. 1 illustrates the architecture of the QoT-aware dynamic RMSA environment. In this environment, a network topology comprises an optical line system composed of optical

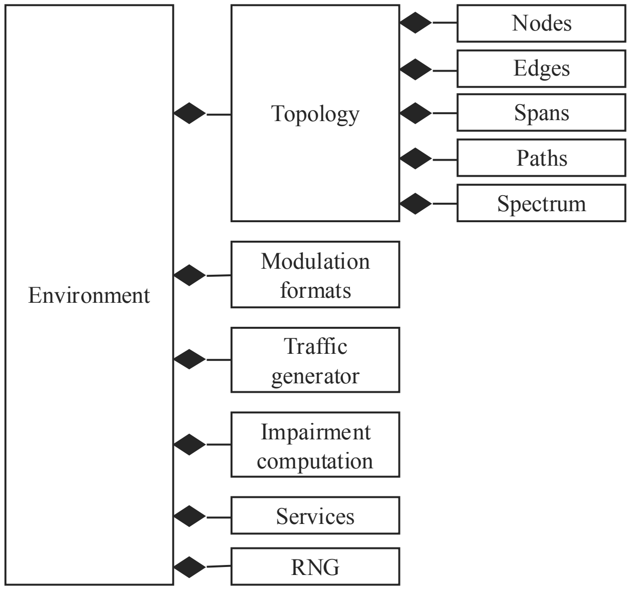


Fig. 1. General architecture of the QoT-aware dynamic RMSA environment.

nodes equipped with reconfigurable optical add-drop multiplexers (ROADMs) and optical terminals, and optical fiber links equipped with in-line amplifiers. The optical spectrum is divided into frequency slots. Optical terminals are bandwidth-variable transceivers (BVTs) used at the communicating node pairs, enabling the establishment of lightpaths that use multiple frequency slots. Each lightpath uses a modulation format characterized by two main parameters: spectral efficiency and minimum QoT level. The spectral efficiency, measured in b/Hz/s, defines how much data a lightpath using a modulation format can transmit per unit of bandwidth and time. The minimum QoT level guarantees that the modulation format will work under a pre-defined bit error rate, given a forward error correction scheme, usually measured in terms of generalized signal-to-noise ratio (GSNR) [dB].

In the dynamic RMSA problem considered in this environment, an optical service request specifies the node pair that needs to be connected and the bit rate of the lightpath to be created. The intensity of the traffic is governed by the load, measured in Erlangs. This load, in turn, defines the inter-arrival time and holding time of the requests, following a Poisson process. The node pairs related to the requests can be selected uniformly (i.e., each node has the same probability of originating/terminating traffic) or non-uniformly (each node has an independent probability of originating/terminating traffic).

Then, the RMSA solution needs to find a route (sequence of fiber links), a modulation format, and a spectrum slice that accommodates the requested bit rate under two constraints. Firstly, the selected modulation format requires a minimum QoT level, and its spectral efficiency will define the necessary bandwidth (i.e., the number of frequency slots) given the bit rate requested. Depending on the technology used by the optical devices, the number of frequency slots may be bound by a minimum and maximum value. The QoT level of an unestablished optical service can be estimated, for instance, by using analytical models such as the one detailed in Sec. 3.A. Secondly, the spectrum slice must be continuous (along all fiber links), contiguous (across the spectrum), and not overlapping with any other lightpath.

In the following, we discuss the physical layer model adopted to compute the QoT of lightpaths in the environment, characterized by their GSNR. Then, we detail heuristic solutions to the problem that can be used as a benchmark in future studies. In the remainder of this paper, we adopt the following variable definitions:

$G(\mathcal{V},\mathcal{E})$	Undirected network graph where $\mathcal{V}$ is a set of optical nodes, and $\mathcal{E}$ is a set of optical fibers that interconnect optical nodes.
$\mathcal{P}^e$	Set of fiber spans of link $e \in \mathcal{E}$.
$\mathcal{M}$	Set of modulation formats, where each $m \in \mathcal{M}$ have m^s [b/Hz/s] representing its spectral efficiency, and m^{th} representing the minimum QoT required for the modulation format to work.
$\mathcal{K}$	Set of k-shortest-paths between all node pairs, where each $k \in \mathcal{K}$ contains its length k^{len} and set of links k^{links}.
$\mathcal{K}_{n1,n2}$	represent the k-shortest-paths between nodes $n1$ and $n2$, with $n1, n2 \in V$ and $\lvert\mathcal{K}_{n1,n2}\rvert = k$.
$\mathcal{L}$	Set of lightpaths currently running in the network, where each $l \in \mathcal{L}$ specifies its source l^{src} and destination l^{dst} nodes, its bit rate l^{br}, and the route $l^r : l^r \in \mathcal{K}_{l^{src},l^{dst}}$.
o	service request between nodes $o^{src} \in \mathcal{V}$ and $o^{dst} \in \mathcal{V}$ requiring bit rate o^{br}.

A. Physical Layer Model

The ONG enables the use of physical layer models that consider the linear and non-linear impairments of lightpaths and their mutual impact. Generally, a hypothetical function implements a physical layer model to estimate the QoT of an unestablished lightpath under test (LUT) l traversing route r and occupying w frequency slots. Such models require the knowledge of the network graph $G(\mathcal{V},\mathcal{E})$. Moreover, the co-propagating lightpaths can be found within the set of lightpaths currently running in the network $\mathcal{L}$. Finally, the function returns the QoT of the LUT.

In the specific environment analyzed in this paper, we adopt the enhanced Gaussian noise (EGN) analytical model described in [28]. This analytical model was selected due to its wide acceptance in the literature, its ability to consider partially-loaded networks, and its arbitrary lightpath bandwidth. Note that the ONG provides interfaces for implementing virtually any physical layer model, including closed-form analytical models such as those in [28], as well as semi-analytical models like the one implemented in GNPy [8], or machine learning (ML). The ONG also has the potential for customization in multi-band and/or space division multiplexing scenarios.

In the EGN model, the GSNR of a LUT that uses path k can be calculated as a function of the set of spans $\mathcal{P}$ along the path, where $\mathcal{P} = \bigcup_{e \in k^{links}} \mathcal{P}^e$:

$$GSNR_{\text{LUT}}{}^{-1} = \sum_{p \in \mathcal{P}} \left(\frac{P^p_{\text{LUT}}}{P^p_{\text{ASE}} + P^p_{\text{NLI}}} \right)^{-1}, \quad (1)$$

where, for each span $p \in \mathcal{P}$, P^p_{CUT} is the launch power, P^p_{ASE} is the amplified spontaneous emission (ASE) noise power incurred by doped fiber amplifiers, and P^p_{NLI} is the noise incurred by non-linearity. The complete formulation of the P^p_{ASE} and P^p_{NLI} terms can be found in [28]. In the environment, we also measure the signal-to-noise ratio (SNR) considering only the ASE and the SNR considering only the non-linear (NLI), enabling the analysis of the trade-offs between these two components of the GSNR.

B. Heuristic Solutions

This section discusses frequently used algorithms for solving the dynamic RMSA problem. The algorithms follow the common interface defined in Listing 2. These algorithms are included within the ONG, can serve as a benchmark for new proposals, and can be easily extended to include newly-published solutions. In this paper, we focus on four heuristic solutions that are frequently used as benchmarks, i.e., the ones prioritizing the best possible modulation format, the lowest frequency slot, or load balancing. We refer the reader to [29] for a more in-depth discussion of the algorithms. Given the open-source nature of ONG, other heuristics and DRL-based solutions can be easily added by the research community as contribution to the project.

All the heuristics in this section have access to the current request o, the set of pre-computed k-shortest paths $\mathcal{K}$ (sorted from the shortest to the longest), the set of currently running services $\mathcal{L}$, and the set of modulation formats $\mathcal{M}$ (sorted from the highest to the lowest spectral efficiency). If the request is accepted, the algorithm outputs the tuple $< r, m, s >$ representing the route, modulation format, and initial frequency slot to be used by the request, respectively. If the request cannot be accommodated, the heuristic returns $\emptyset$. In general, we adopted a naming convention that specifies the sequence of priorities of each heuristic.

Algorithm 1. Pseudocode of the k-shortest-path best-modulation first-fit (KSP-BM-FF) heuristic.

1: **for** $r \in \mathcal{K}_{o^{src},o^{dst}}$ **do**
2: $\quad available\_slots \leftarrow get\_available\_slots(r)$
3: $\quad$ **for** $m \in \mathcal{M}$ **do**
4: $\quad\quad number\_slots \leftarrow get\_number\_slots(o^{br}, m)$
5: $\quad\quad s \leftarrow first\_fit(available\_slots, number\_slots)$
6: $\quad\quad$ **if** $s = null$ **then**
7: $\quad\quad\quad$ **continue** to next route
8: $\quad\quad qot \leftarrow calculate\_qot(\mathcal{L}, o, r, s, number\_slots)$
9: $\quad\quad$ **if** $qot \geq m^{th}$ **then**
10: $\quad\quad\quad$ **return** r, m, s
11: **return** $\emptyset$

The k-shortest-path best-modulation first-fit (KSP-BM-FF) heuristic is illustrated in Alg. 1. This heuristic finds the shortest route, most spectrally efficient modulation format, and lowest spectrum that meets the problem's constraints. We start by iterating over the set of pre-computed shortest-paths (line 1), and obtaining the continuous available slots in such a route (line 2). Then, for each modulation format (line 3), we obtain the number of frequency slots necessary to accommodate the requested bit rate (line 4). The heuristic finds the first frequency slot of a contiguous spectrum block able to accommodate the necessary number of frequency slots (line 5), returning a *null* value if such a block cannot be found. If a free spectrum block cannot be found (line 6), we can continue to the next route since evaluating the next (less efficient) modulation format would require a larger spectrum block than the current one (line 7). The QoT of the current solution is computed (line 8). If the QoT meets the requirements of the modulation format (line 9), the solution tuple is returned (line 10). If no solution is found, the heuristic returns an empty tuple (line 11).

Algorithm 2. Pseudocode of the best-modulation lowest-spectrum k-shortest-path (BM-LS-KSP) heuristic.

1: $solution \leftarrow \emptyset$ $\qquad\triangleright$ solution tuple
2: $lfs \leftarrow \infty$ $\qquad\triangleright$ current lowest initial frequency slot
3: **for** $m \in \mathcal{M}$ **do**
4: $\quad number\_slots \leftarrow get\_number\_slots(o^{br}, m)$
5: $\quad$ **for** $r \in \mathcal{K}_{o^{src},o^{dst}}$ **do**
6: $\quad\quad available\_slots \leftarrow get\_available\_slots(r)$
7: $\quad\quad s \leftarrow first\_fit(available\_slots, number\_slots)$
8: $\quad\quad$ **if** $s \neq null \wedge s < lfs$ **then**
9: $\quad\quad\quad qot \leftarrow calculate\_qot(\mathcal{L}, o, r, s, number\_slots)$
10: $\quad\quad\quad$ **if** $qot \geq m^{th}$ **then**
11: $\quad\quad\quad\quad lfs \leftarrow s$
12: $\quad\quad\quad\quad solution \leftarrow r, m, s$
13: $\quad$ **if** $solution \neq \emptyset$ **then**
14: $\quad\quad$ **return** $solution$
15: **return** $\emptyset$

Alg. 2 illustrates the best-modulation lowest-spectrum k-shortest-path (BM-LS-KSP) heuristic. This heuristic finds the best modulation format possible for the current request, selecting the route that allows the use of the lowest spectrum, i.e., the lowest initial frequency slot. The heuristic starts by initializing its state variables (lines 1–2). For each modulation format (line 3), we find the number of frequency slots necessary to accommodate the requested bit rate (line 4). Then, we evaluate all routes (lines 5–7) and store the one that has the lowest initial frequency slot (lines 8) and that meets the QoT requirements of the modulation format (lines 9–12). Suppose a solution is found for the current (most spectrally efficient) modulation format (line 13). In that case, the heuristic returns the solution (line 14) since evaluating other (less efficient) modulation formats is unnecessary. If no solution is found, the heuristic returns an empty tuple (line 15).

Algorithm 3. Pseudocode of the best-modulation load-balancing k-shortest-path (BM-LB-KSP) heuristic.

1: $solution \leftarrow \emptyset$ $\qquad\triangleright$ solution tuple
2: $war \leftarrow 0$ $\qquad\triangleright$ current highest weighted available resources
3: **for** $m \in \mathcal{M}$ **do**
4: $\quad number\_slots \leftarrow get\_number\_slots(o^{br}, m)$
5: $\quad$ **for** $r \in \mathcal{K}_{o^{src},o^{dst}}$ **do**
6: $\quad\quad available\_slots \leftarrow get\_available\_slots(r)$
7: $\quad\quad$ **if** $sum(available\_slots)/|r^{links}|^2 > war$ **then**
8: $\quad\quad\quad s \leftarrow first\_fit(available\_slots, number\_slots)$
9: $\quad\quad\quad$ **if** $s = null$ **then**
10: $\quad\quad\quad\quad$ **continue**
11: $\quad\quad\quad qot \leftarrow calculate\_qot(\mathcal{L}, o, r, s, number\_slots)$
12: $\quad\quad\quad$ **if** $qot \geq m^{th}$ **then**
13: $\quad\quad\quad\quad war \leftarrow sum(available\_slots)/|r^{links}|^2$
14: $\quad\quad\quad\quad solution \leftarrow r, m, s$
15: $\quad$ **if** $solution \neq \emptyset$ **then**
16: $\quad\quad$ **return** $solution$
17: **return** $\emptyset$

Alg. 3 details the best-modulation load-balancing k-shortest-path (BM-LB-KSP) heuristic. This heuristic finds the best modulation format possible for the current request, selecting the route with the highest weighted available resources and, conversely, the lowest weighted resource usage. The weighted available resources, which follows the inverse intuition of the traffic load balancing (TLB) factor in [30], is computed as the number of continuous available frequency slots in a given route divided by the squared number of hops in the route. The heuristic defines two state variables (lines 1–2), iterates over the modulation formats (line 3), and computes the number of frequency slots required to fulfill the requested bit rate (line 4). For each route (line 5), the available slots are computed (line 6), and the weighted available resources are compared to the previous best one found (line 7). The heuristic only evaluates the route if it has a higher weighted factor. If a spectrum block is not found for the current route, the heuristic continues to the next route (lines 8–10). Otherwise, the QoT of the current solution is evaluated (lines 11–12) and stored in the state variables if the modulation format QoT threshold is met (lines 12–14). Suppose a solution is found for the current (most spectrally efficient) modulation format (line 15). In that case, the heuristic returns the solution (line 16) since evaluating other (less efficient) modulation formats is unnecessary. If no solution is found, the heuristic returns an empty tuple (line 17).

Algorithm 4. Pseudocode of the load-balancing best-modulation k-shortest-path (LB-BM-KSP) heuristic.

1: $solution \leftarrow \emptyset$ ▷ solution tuple
2: $war \leftarrow 0$ ▷ current highest weighted available resources
3: **for** $r \in \mathcal{K}_{o^{src},o^{dst}}$ **do**
4: $\quad available\_slots \leftarrow get\_available\_slots(r)$
5: $\quad$ **if** $sum(available\_slots)/|r^{links}|^2 > war$ **then**
6: $\quad\quad$ **for** $m \in \mathcal{M}$ **do**
7: $\quad\quad\quad number\_slots \leftarrow get\_number\_slots(o^{br}, m)$
8: $\quad\quad\quad s \leftarrow first\_fit(available\_slots, number\_slots)$
9: $\quad\quad\quad$ **if** $s = null$ **then**
10: $\quad\quad\quad\quad$ **continue** to next route
11: $\quad\quad\quad qot \leftarrow calculate\_qot(\mathcal{L}, o, r, s, number\_slots)$
12: $\quad\quad\quad$ **if** $qot \geq m^{th}$ **then**
13: $\quad\quad\quad\quad war \leftarrow sum(available\_slots)/|r^{links}|^2$
14: $\quad\quad\quad\quad solution \leftarrow r, m, s$
15: $\quad\quad\quad\quad$ **break**
16: **return** $solution$

Alg. 4 illustrates the load-balancing best-modulation k-shortest-path (LB-BM-KSP) heuristic. This heuristic follows a similar intuition to the previous one. Still, unlike the previous one, it selects the route with the highest weighted available resources regardless of the modulation format to be used. State variables are initialized (lines 1–2). For each route, the available slots are computed (lines 3–4). If the weighted resources available are higher than the previous best one (line 5), the modulation formats are evaluated (lines 6–15). When a modulation format in the current route has its QoT threshold met, it is saved as the best one (lines 12–14), and the heuristic follows to evaluate the next route. Finally, the heuristic returns the solution found (line 16), an empty tuple in case no feasible solution was found.

4. USE CASES

This section demonstrates three potential use cases for the dynamic RMSA environment described in the previous section, highlighting the analysis that can be performed using its functionalities. Since use cases related to DRL have already been explored in the literature, we focus on the ones that highlight other aspects of the analysis enabled by the ONG. In particular, the three use cases presented take advantage of the EGN model and heuristic implementation described in the previous section.

For the results in this section, we use the European network topology shown in Fig. 2, identified as *nobel-eu* in SNDlib [26]. This topology was selected because it has a wide range of link lengths, resulting in a good ground for analysis of the dynamic RMSA problem under realistic modeling of physical layer impairments. Fig. 2 also shows that the path length distribution contains a wide range of values, being a representative topology for core networks where the dynamic QoT-aware RMSA is more relevant. The topology has 28 nodes and 41 links. Spans within a link have equal length with a maximum of 80 km. We compute the 5 shortest paths for each node pair. Source and destination nodes are uniformly selected, and the bit rate is chosen uniformly among {10, 40, 100, 400} Gbps [31]. Request arrivals follow a Poisson process, where the service holding time is exponentially distributed with an average of 60 time units. The holding time is exponentially distributed, whose mean is defined by the specific load value under exam. The environment is configured to run episodes of 1,000 request arrivals.

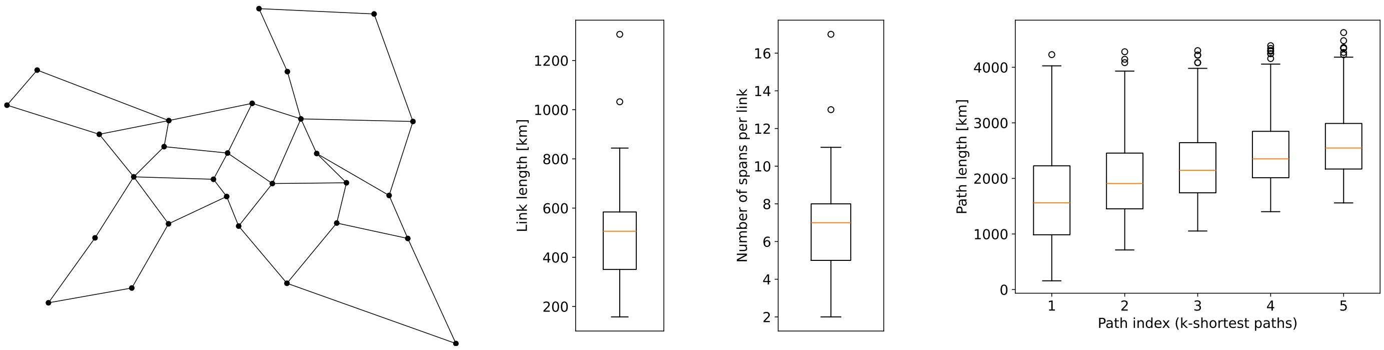
In the simulations, we consider the C-band with 4 THz of bandwidth divided into 320 frequency slots of 12.5 GHz. A guardband of one frequency slot is required between two neighboring lightpaths. All spans have an attenuation coefficient equal to 0.2 dB/km, equipped with an Erbium-dopped fiber amplifier (EDFA) with a noise figure of 4.5 dB. We consider six modulation formats: BPSK, QPSK, 8-, 16-, 32-, and 64-QAM, with their GSNR threshold set to {3.71, 6.72, 10.84, 13.24, 16.16, 19.01} dB, respectively [32].

In the following use cases, we assess the performance of a solution using the following metrics. The request blocking ratio represents the ratio between blocked requests and the total number of requests. Similarly, the bit rate blocking ratio is the ratio between the total bit rate of the blocked requests and the total bit rate of all requests. The SNR is measured for the linear (ASE) and non-linear (NLI) components, as well as the GSNR. Finally, the request processing time measures the total time taken to process, on average, each optical service request in the ONG environment. The time measurements were collected using a workstation equipped with an AMD Ryzen Threadripper 3960X 24-core processor with 128 GB of RAM running Ubuntu 22.04 and Python 3.12.

A. Margin optimization

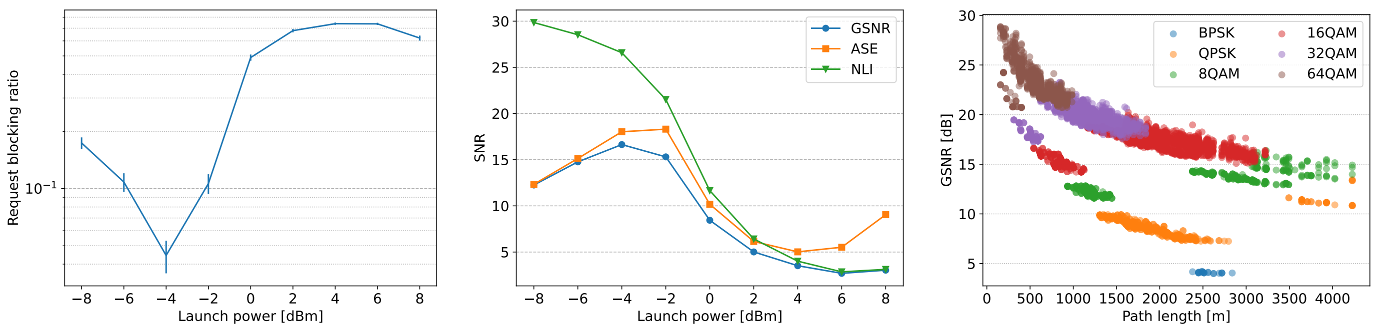
The margin optimization use case studied in this paper focuses on finding one single flat launch power that, when used by all lightpaths in the network, yields the lowest request blocking ratio. This is an essential step when benchmarking RMSA algorithms due to the impact that the launch power has on the GSNR profile obtained in the network. This is also a relevant problem in the literature, analyzed in various contexts, and that needs to be revisited for new optical architectures and technologies [33–35]. Moreover, the launch power selection is impacted by several physical layer properties such as the span length, attenuation, and noise figure. Therefore, a launch power optimization campaign is necessary whenever QoT-aware algorithms are to be benchmarked in a given network. We analyze the blocking probability in different scenarios by varying the launch power in the range of (-8, 8) dBm, with steps of 2 dBm. For each scenario, we run 1 million request arrivals, i.e., 1,000 episodes of 1,000 arrivals, resulting in an average standard deviation that represents less than 1% of the average request blocking ratio value.

Fig. 3 shows the average simulation results. Fig. 3a shows



(a) Topology **(b)** Link length **(c)** Spans per link **(d)** Path length per shortest path

Fig. 2. The European network topology with link and path length distributions.



(a) Request blocking ratio **(b)** Generalized signal-to-noise ratio (GSNR) **(c)** GSNR vs. path length for launch power of -4 dBm

Fig. 3. Impact of the launch power on the overall request blocking rate for the European topology considering a load of 210 Erlang.

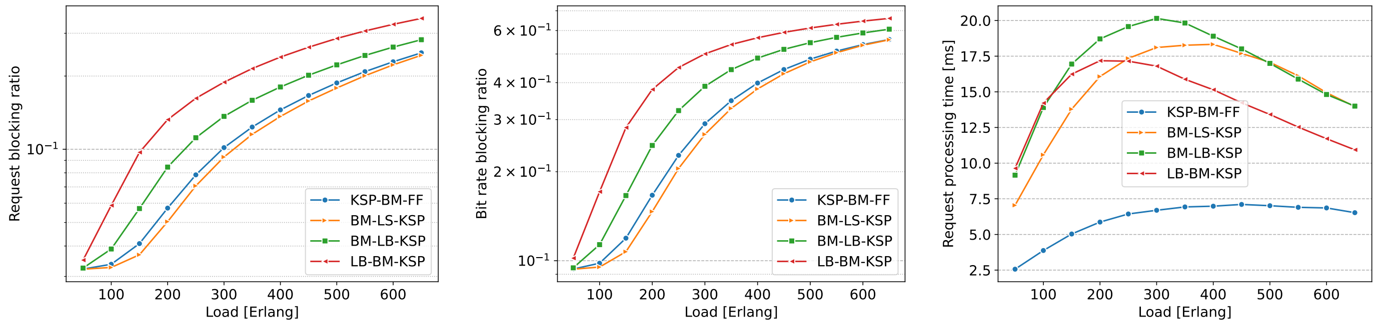
the average request blocking ratio, which indicates that -4 dBm yields the lowest blocking over the simulated launch power values. Fig. 3b shows the average SNR of all the accepted requests in terms of the overall GSNR and its linear and non-linear components. We can see that the best GSNR is achieved when adopting -4 dBm as the launch power, which coincides with and explains the blocking ratio results in Fig. 3a. Moreover, we can confirm and measure the intuition that in the lower launch power range, the GSNR is limited by the ASE, while in the higher launch power, the GSNR is limited by the NLI. Fig. 3c shows the GSNR achieved by lightpaths according to their path length for the first 100,000 accepted requests, assuming the best launch power found (-4 dBm). We can observe that there is a large range within which the same modulation format is selected, indicating that distance-based modulation format selection is not ideal. ONG makes the analysis in Fig. 3 accessible, since it can be performed for any topology and physical layer parameter configuration without the need of any modifications to the code. Finally, the simulations generating the results reported in Fig. 3 took approximately 2 hours to complete, using 9 CPU cores.

B. Benchmarking QoT-Aware Dynamic RMSA Algorithms

Evaluating and benchmarking algorithms that solve the QoT-aware dynamic RMSA problem is one of the most common use cases in the literature [1, 5, 22]. In this use case, the algorithm needs to find a solution for the RMSA problem upon the arrival of an optical service request. This evaluation is usually performed by assessing the performance of various algorithms while varying the load in the network. Although common in the literature, these analysis usually include only a few algorithms, since implementing previous algorithms from the literature and integrating them into custom simulators is laborious. Moreover, if QoT is considered, this assessment needs to be preceded by a launch power optimization as the one previously presented.

Fig. 4 shows the simulation results for the heuristic algorithms presented in Sec. 3.B. Each data point in the figure is obtained by averaging the results over 1 million arrivals, using -4 dBm launch power as obtained in the previous use case. Fig. 4a shows that the best-modulation lowest-spectrum k-shortest-path (BM-LS-KSP) achieves the lowest request blocking ratio. Moreover, we can observe that the k-shortest-path best-modulation first-fit (KSP-BM-FF), commonly used in the literature as the benchmark for new algorithms, closely follows the performance of the best strategy with only a 5% gap. The load-balancing heuristics achieve the highest blocking ratio. Fig. 4b shows that the bit rate blocking ratio follows a similar trend as the request blocking ratio.

Fig. 4c shows the average processing time taken to simulate the processing of each optical service request. The KSP-BM-FF has the lowest processing time, with a maximum of 7.5 ms, due to its simple implementation that considers the first feasible solution found. This means that one million request arrivals can be processed in around 2 hours, similar to the results obtained in [16] for a different scenario. Note that this time represents the time taken by one CPU core, i.e., no parallelism is used. More importantly, it shows that the code optimizations in ONG make its GSNR estimation execution quite fast. Using multiple CPU cores

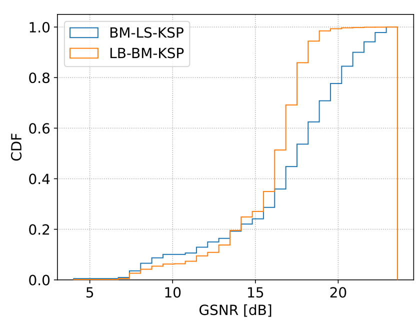


(a) Request blocking rate **(b)** Bit rate blocking ratio **(c)** Average step time

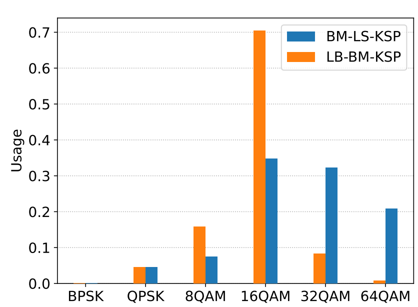
Fig. 4. Analysis of the RMSA heuristics over the offered load for the European topology with -8 dBm as the launch power.

yields close-to-linear performance scaling. The other heuristics require a similar processing time and have a typical behavior of increasing time with the load up to a point at which further increasing the load decreases the processing time. This behavior is explained by the fact that as the network load increases, so does the number of co-propagating channels, slightly increasing the processing time of the physical layer model. Further increases in load result in less available resources, which results in the heuristics performing an early stop due to the lack of resources, not requiring the computation of the GSNR.

C. Generation of datasets for AI/ML-based QoT Estimation



(a) Cummulative distribution function (CDF) of the GSNR



(b) Modulation format usage

Fig. 5. Statistical analysis of the dataset generated by the ONG for the European topology with 210 Erlang load and -8 dBm as the launch power.

This use case illustrates the possibility of leveraging ONG to collect QoT datasets suitable for use in AI/ML QoT estimation studies [17, 18]. Previous works in the literature sharing analytical or experimental datasets have been essential for the progress of AI/ML QoT estimation studies [19, 20]. However, these datasets contain a fixed information set related to each lightpath, which may not always be relevant to the problem [36], or not always contain all the relevant information. For instance, the [37] considers link-level formulation where a vector specifies which links are traversed by a lightpath. Another example is [38], where the GSNR of all lightpaths is calculated for every new network state. Naturally, pre-defined datasets are not able to include any possible relevant information about the network and lightpaths. ONG adopts a different approach by enabling researchers to generate their own dataset, including any information that is deemed relevant, or whose relevance will be studied. Examples of such information are the specific links and/or spans traversed by the lightpath, and their details such as length, attenuation coefficient, and amplifier noise. Another example are the detailed information of co-propagating channels, such as bandwidth, bit rate, and modulation format. This flexibility enables researchers to generate accurate datasets for any network topology and physical layer parameters using new (never-used-before) information, as well as compare with previous setups, datasets, and dataset information.

To illustrate the dataset generation capabilities of ONG, we select two heuristics with the most divergent request blocking ratio performance, i.e., BM-LS-KSP and load-balancing best-modulation k-shortest-path (LB-BM-KSP) from Fig. 4. For each heuristic, we process 100,000 service request arrivals and save the information of the accepted lightpaths in the dataset. Fig. 5 shows a statistical analysis of the GSNR and modulation format usage of the two datasets generated by ONG. Fig. 5a shows a significant difference in the GSNR of the accepted lightpaths between the two heuristics. Nearly 20% of the lightpaths provisioned using BM-LS-KSP achieve a GSNR higher than 20 dB. Meanwhile, most lightpaths provisioned using LB-BM-KSP achieve a maximum of 20 dB. The impact of the GSNR difference can be observed in Fig. 5b. While more than 50% of the lightpaths provisioned using BM-LS-KSP have modulation formats 32- or 64-QAM, 70% of the lightpaths provisioned using LB-BM-KSP use 16-QAM. Besides enabling the generation of custom datasets for training AI/ML models, these details improve the understanding of the differences of the heuristics in terms of blocking ratio.

One important consideration when generating datasets is related to bias and fairness. The adopted RMSA solution will define the distribution of several features in the dataset, such as the modulation format illustrated in Fig. 5a. A few works have shown that biased datasets may impact the accuracy distribution of trained ML models [39, 40]. However, in real-world network it is unlikely that datasets will be completely balanced and fair, so techniques to tackle such imbalances need to be developed [40].

It is important to also highlight that, although we only show two pieces of information of the dataset, a wide range of information can be included on the dataset if needed. Finally, the ONG script made available to generate the results in Fig. 5 can be used to save the collected dataset in an interoperable format, ready to be shared with the community.

5. OPEN CHALLENGES AND FUTURE WORK

The ONG aims to represent a reference architecture for evaluating optical resource assignment problems. This objective is approached by defining interfaces that enable the toolkit to be easily extended, and a set of scripts that exemplify its key functionalities. With the rapid research and development of new technologies, several studies can be performed by extending the ONG toolkit.

One of these areas is the study of the dynamic operation of multi-band optical networks. This problem has already been studied in the literature [13], but with limited modeling of the physical layer impairments. More recently, studies using the generalized Gaussian model (GGN) model (implemented in GNPy [8]) have analyzed the performance of RL and DRL [14, 41]. Moreover, [42] studies how DRL can be used for multi-band networks using the EGN model assuming a fixed-grid scenario. However, considering a QoT-aware flexi-grid scenario, which has been considered for the C+L-band in [43], is essential to fully understand the potential of beyond C+L-band optical networks, and how ML and DRL can be leveraged to take the most advantage of it.

Among the future works is implementing physical layer models that consider the specific characteristics of these networks [44–46], which will enable a more realistic performance assessment of resource assignment algorithms for multi-band networks.

The extension of the environments with multi-core fiber technologies is another future work. This scenario is already a good fit for DRL [47, 48], but could also benefit from the realistic modeling of physical layer impairments [49]. More interestingly, combining multi-core and multi-band optical network transmission technologies represents a futuristic and relevant use case whose study can benefit from the artifacts provided by ONG.

Another area that can benefit from the ONG toolkit is the study of spectrum defragmentation algorithms. There are already works in the literature leveraging DRL to solve the spectrum defragmentation problems [50] including one using the previous version of ONG [15]. However, ONG enables the study of QoT-aware defragmentation algorithms, which pose new and more realistic challenges for their implementation in real deployments.

In terms of more realistic modeling of the physical layer impairments in the QoT-aware dynamic RMSA environment discussed in this paper, including the amplifier tilt can represent another interesting step [51]. The analytical model implemented in the ONG also allows per-span specification of physical layer parameters, making it possible to generate datasets that contain anomalies such as soft failures, which are an important use case for the use of AI/ML [52]. Analyzing which heuristics to use to generate balanced and fair datasets can be an interesting way to mitigate the impact of these imbalances in future ML-based QoT estimators. Extending ONG to support multicasting scenarios also represents an interesting future direction. Finally, developing environments for quantum key distribution scenarios can also enable further studies of how to operate networks with this technology [53].

6. FINAL REMARKS

This paper introduced the ONGs, a new open-source toolkit that builds upon a previous toolkit, improving and expanding its functionalities. The toolkit extends well-known interfaces from the RL research, making it suitable for a large range of use cases due to the modularity of the interfaces. The toolkit proposes a series of interfaces that simplify and streamline the development of new optical networking environments and/or respective solution algorithms. One of the available environments that model the QoT-aware dynamic RMSA problem is discussed in detail. This environment is particularly important for combining the resource assignment problem and an analytical model physical layer impairment estimation into a single environment. Three use cases demonstrate the toolkit's versatility, flexibility, and broad applicability. In the first use case, launch power optimization was performed, indicating which launch power is most suitable for the network scenario at hand. In the second use case, four different heuristics were analyzed, revealing their request blocking and processing time performance. In the third use case, two datasets were generated. The datasets are suitable for studies related to the AI/ML-based QoT estimation. We believe this toolkit represents a suitable benchmarking tool, enabling researchers to compare their proposed algorithms more easily with previous ones from the literature. In this regard, a few relevant and timely open challenges and potential future work were discussed. The toolkit also lowers the barrier of entry for new researchers in the field of optical networking resource assignment.

FUNDING

This work was supported by the ECO-eNET project with funding from the Smart Networks and Services Joint Undertaking (SNS JU) under grant agreement No. 10113933. The JU receives support from the European Union's Horizon Europe research and innovation programme. The authors from UFPA would like to acknowledge the support from the Brazilian National Council for Scientific and Technological Development (CNPq). The authors from UC3M would like to acknowledge the support of the EU-funded SEASON project (grant No.101096120) and Spanish-funded Fun4date-Redes project (grant No.PID2022-136684OB-C21). Farhad Arpanaei acknowledges support from the CONEX-Plus programme funded by Universidad Carlos III de Madrid and the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Actions.

REFERENCES

1. A. Mokhtar and M. Azizoglu, "Dynamic route selection and wavelength assignment in all-optical networks," in *LEOS '95. IEEE Lasers and Electro-Optics Society 1995 Annual Meeting. 8th Annual Meeting. Conference Proceedings*, vol. 1 (1995), pp. 220–221 vol.1.

2. L. Tang, S. Billenahalli, W. Huang, M. Razo, A. Sivasankaran, H. Vardhan, P. Monti, M. Tacca, and A. Fumagalli, “The PlaNet-OTN module: A double layer design tool for optical transport networks,” in *International Conference on Transparent Optical Networks,* (2009).
3. P. Pavon-Marino and J.-L. Izquierdo-Zaragoza, “Net2plan: an open source network planning tool for bridging the gap between academia and industry,” IEEE Netw. **29**, 90–96 (2015).
4. S. Azodolmolky, J. Perelló, M. Angelou, F. Agraz, L. Velasco, S. Spadaro, Y. Pointurier, A. Francescon, C. V. Saradhi, P. Kokkinos, E. Varvarigos, S. A. Zahr, M. Gagnaire, M. Gunkel, D. Klonidis, and I. Tomkos, “Experimental demonstration of an impairment aware network planning and operation tool for transparent/translucent optical networks,” J. Light. Technol. **29**, 439–448 (2011).
5. R. S. Tessinari, B. Puype, D. Colle, and A. S. Garcia, “Elastico++: An elastic optical network simulation framework for omnet++,” Opt. Switch. Netw. **22**, 95–104 (2016).
6. D. Wang, Y. Song, Y. Zhang, X. Jiang, J. Dong, F. N. Khan, T. Sasai, S. Huang, A. P. T. Lau, M. Tornatore, and M. Zhang, “Digital twin of optical networks: A review of recent advances and future trends,” J. Light. Technol. pp. 1–28 (2024).
7. D. Semrau, R. I. Killey, and P. Bayvel, “A closed-form approximation of the gaussian noise model in the presence of inter-channel stimulated raman scattering,” J. Light. Technol. **37**, 1924–1936 (2019).
8. A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrat, G. Grammel, G. Galimberti, and V. Curri, “GNPy: an open source application for physical layer aware open optical networks,” J. Opt. Commun. Netw. **12**, C31–C40 (2020).
9. A. A. Díaz-Montiel, B. Lantz, J. Yu, D. Kilper, and M. Ruffini, “Real-time QoT estimation through SDN control plane monitoring evaluated in Mininet-Optical,” IEEE Photonics Technol. Lett. **33**, 1050–1053 (2021).
10. C. Natalino and P. Monti, “The Optical RL-Gym: an open-source toolkit for applying reinforcement learning in optical networks,” in *International Conference on Transparent Optical Networks (ICTON),* (2020), p. Mo.C1.1.
11. 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,” J. Opt. Commun. Netw. **14**, 733–748 (2022).
12. N. Jara, H. Pempelfort, E. Viera, J. Sanchez, G. España, and D. Borquez-Paredes, “DREAM-ON GYM: A deep reinforcement learning environment for next-gen optical networks,” in *International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH),* INSTICC (SciTePress, 2024), pp. 215–222.
13. P. Morales, P. Franco, A. Lozada, N. Jara, F. Calderón, J. Pinto-Ríos, and A. Leiva, “Multi-band environments for optical reinforcement learning gym for resource allocation in elastic optical networks,” in *International Conference on Optical Network Design and Modeling (ONDM),* (2021), pp. 1–6.
14. A. B. Terki, J. Pedro, A. Eira, A. Napoli, and N. Sambo, “Deep reinforcement learning for resource allocation in multi-band optical networks,” in *International Conference on Optical Network Design and Modeling (ONDM),* (2024), pp. 1–4.
15. E. Etezadi, C. Natalino, R. Diaz, A. Lindgren, S. Melin, L. Wosinska, P. Monti, and M. Furdek, “Deep reinforcement learning for proactive spectrum defragmentation in elastic optical networks,” J. Opt. Commun. Netw. **15**, E86–E96 (2023).
16. M. Doherty and A. Beghelli, “XLRON: Accelerated reinforcement learning environments for optical networks,” in *Optical Fiber Communications Conference and Exhibition (OFC),* (2024), p. Th2A.21.
17. T. Panayiotou, S. P. Chatzis, and G. Ellinas, “Performance analysis of a data-driven quality-of-transmission decision approach on a dynamic multicast- capable metro optical network,” J. Opt. Commun. Netw. **9**, 98–108 (2017).
18. C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, “Machine-learning method for quality of transmission prediction of unestablished lightpaths,” J. Opt. Commun. Netw. **10**, A286–A297 (2018).
19. G. Bergk, B. Shariati, P. Safari, and J. K. Fischer, “ML-assisted QoT estimation: a dataset collection and data visualization for dataset quality evaluation,” J. Opt. Commun. Netw. **14**, 43–55 (2022).
20. Z. Zhai, L. Dou, Y. He, A. P. T. Lau, and C. Xie, “Open-source data for QoT estimation in optical networks from Alibaba,” J. Opt. Commun. Netw. **16**, 1–3 (2024).
21. “Optical networking gym,” https://github.com/carlosnatalino/optical-networking-gym (2024). DOI: 10.5281/zenodo.13906325.
22. X. Chen, B. Li, R. Proietti, H. Lu, Z. Zhu, and S. J. B. Yoo, “Deeprmsa: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks,” J. Light. Technol. **37**, 4155–4163 (2019).
23. M. Towers, J. K. Terry, A. Kwiatkowski, J. U. Balis, G. d. Cola, T. Deleu, M. Goulão, A. Kallinteris, A. KG, M. Krimmel, R. Perez-Vicente, A. Pierré, S. Schulhoff, J. J. Tai, A. T. J. Shen, and O. G. Younis, “Gymnasium,” (2023).
24. G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, “OpenAI Gym,” (2016).
25. R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction* (MIT Press, 2018).
26. S. Orlowski, R. Wessäly, M. Pióro, and A. Tomaszewski, “SNDlib 1.0–survivable network design library,” Networks. **55**, 276–286 (2010).
27. S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, “The internet topology zoo,” IEEE J. on Sel. Areas Commun. **29**, 1765–1775 (2011).
28. M. Ranjbar Zefreh, F. Forghieri, S. Piciaccia, and P. Poggiolini, “Accurate closed-form real-time EGN model formula leveraging machine-learning over 8500 thoroughly randomized full C-band systems,” J. Light. Technol. **38**, 4987–4999 (2020).
29. B. C. Chatterjee, N. Sarma, and E. Oki, “Routing and spectrum allocation in elastic optical networks: A tutorial,” IEEE Commun. Surv. & Tutorials **17**, 1776–1800 (2015).
30. F. Arpanaei, N. Ardalani, H. Beyranvand, and B. Shariati, “QoT-aware performance evaluation of spectrally–spatially flexible optical networks over FM-MCFs,” J. Opt. Commun. Netw. **12**, 288–300 (2020).
31. ITU.T, “Interfaces for the optical transport network (otn),” https://www.itu.int/rec/T-REC-G.709-200912-S (30 June 2024).
32. F. Arpanaei, J. M. Rivas-Moscoso, I. De Francesca, J. A. Hernandez, A. Sanchez-Macian, M. R. Zefreh, D. Larrabeiti, and J. P. Fernandez-Palacios, “Enabling seamless migration of optical metro-urban networks to the multi-band: unveiling a cutting-edge 6D planning tool for the 6G era,” J. Opt. Commun. Netw. **16**, 463–480 (2024).
33. D. J. Ives, P. Bayvel, and S. J. Savory, “Adapting transmitter power and modulation format to improve optical network performance utilizing the Gaussian noise model of nonlinear impairments,” J. Light. Technol. **32**, 4087–4096 (2014).
34. I. Roberts and J. M. Kahn, “Efficient discrete rate assignment and power optimization in optical communication systems following the Gaussian noise model,” J. Light. Technol. **35**, 4425–4437 (2017).
35. J. W. Nevin, F. J. Vaquero-Caballero, D. J. Ives, and S. J. Savory, “Physics-informed Gaussian process regression for optical fiber communication systems,” J. Light. Technol. **39**, 6833–6844 (2021).
36. O. Ayoub, S. Troia, D. Andreoletti, A. Bianco, M. Tornatore, S. Giordano, and C. Rottondi, “Towards explainable artificial intelligence in optical networks: the use case of lightpath qot estimation,” J. Opt. Commun. Netw. **15**, A26–A38 (2023).
37. I. Sartzetakis, K. K. Christodoulopoulos, and E. M. Varvarigos, “Accurate quality of transmission estimation with machine learning,” J. Opt. Commun. Netw. **11**, 140–150 (2019).
38. P. Safari, B. Shariati, W. Baltzer, and J. K. Fischer, “QoT violation in low-margin optical networks,” in *Optical Fiber Communications Conference and Exhibition (OFC),* (2023), p. W4G.2.
39. O. Ayoub, D. Andreoletti, S. Troia, S. Giordano, A. Bianco, and C. Rottondi, “Quantifying features’ contribution for ML-based quality-of-transmission estimation using explainable AI,” in *European Conference on Optical Communication (ECOC),* (2022), p. We3B.4.
40. C. Natalino, B. Shariati, P. Safari, J. K. Fischer, and P. Monti, “Analysis and mitigation of unwanted biases in ML-based QoT classification tasks,” in *Optical Fiber Communications Conference and Exhibition*

(OFC), (2024), p. M1H.3.
41. A. B. Terki, J. Pedro, A. Eira, A. Napoli, and N. Sambo, “Routing and spectrum assignment based on reinforcement learning in multi-band optical networks,” in *International Conference on Photonics in Switching and Computing (PSC),* (2023), p. Tu5.63.
42. Y. Teng, C. Natalino, F. Arpanaei, A. Sánchez-Macián, P. Monti, S. Yan, and D. Simeonidou, “DRL-assisted dynamic QoT-aware service provisioning in multi-band elastic optical networks,” in *European Conference on Optical Communication (ECOC),* (2024), p. Th1E.4.
43. M. Mehrabi, H. Beyranvand, M. J. Emadi, and F. Arpanaei, “Efficient statistical QoT-aware resource allocation in EONs over the C+L-band: a multi-period and low-margin perspective,” J. Opt. Commun. Netw. **16**, 577–592 (2024).
44. P. Poggiolini and M. Ranjbar-Zefreh, “Closed form expressions of the nonlinear interference for UWB systems,” in *European Conference on Optical Communication (ECOC),* (2022), p. Tu1D.1.
45. H. Buglia, M. Jarmolovičius, L. Galdino, R. I. Killey, and P. Bayvel, “A closed-form expression for the Gaussian noise model in the presence of Raman amplification,” J. Light. Technol. **42**, 636–648 (2024).
46. M. Jarmolovičius, D. Semrau, H. Buglia, M. Shevchenko, F. M. Ferreira, E. Sillekens, P. Bayvel, and R. I. Killey, “Optimising O-to-U band transmission using fast ISRS Gaussian noise numerical integral model,” J. Light. Technol. pp. 1–8 (2024).
47. J. Pinto-Ríos, F. Calderón, A. Leiva, G. Hermosilla, A. Beghelli, D. Bórquez-Paredes, A. Lozada, N. Jara, R. Olivares, and G. Saavedra, “Resource allocation in multicore elastic optical networks: A deep reinforcement learning approach,” Complexity. **2023**, 4140594 (2023).
48. Y. Teng, C. Natalino, H. Li, R. Yang, J. Majeed, S. Shen, P. Monti, R. Nejabati, S. Yan, and D. Simeonidou, “Deep-reinforcement-learning-based RMSCA for space division multiplexing networks with multi-core fibers,” J. Opt. Commun. Netw. **16**, C76–C87 (2024).
49. F. R. Lobato, A. Jacob, J. Rodrigues, A. V. Cartaxo, and J. Costa, “Inter-core crosstalk aware greedy algorithm for spectrum and core assignment in space division multiplexed elastic optical networks,” Opt. Switch. Netw. **33**, 61–73 (2019).
50. R. Li, R. Gu, W. Jin, and Y. Ji, “Learning-based cognitive hitless spectrum defragmentation for dynamic provisioning in elastic optical networks,” IEEE Commun. Lett. **25**, 1600–1604 (2021).
51. F. Arpanaei, K. Ghodsifar, H. Beyranvand, J. A. Hernández, J. M. Rivas-Moscoso, C. Natalino, M. R. Zefreh, A. Napoli, J. P. Fernández-Palacios, and D. Larrabeiti, “Hyperaccelerated power optimization in multi-band elastic optical networks,” in *Optical Fiber Communications Conference and Exhibition (OFC),* (2024), p. Th1I.6.
52. S. Barzegar, M. Ruiz, A. Sgambelluri, F. Cugini, A. Napoli, and L. Velasco, “Soft-failure detection, localization, identification, and severity prediction by estimating QoT model input parameters,” IEEE Transactions on Netw. Serv. Manag. **18**, 2627–2640 (2021).
53. Y. Cao, Y. Zhao, J. Li, R. Lin, J. Zhang, and J. Chen, “Multi-tenant provisioning for quantum key distribution networks with heuristics and reinforcement learning: A comparative study,” IEEE Transactions on Netw. Serv. Manag. **17**, 946–957 (2020).