Received August 9, 2021, accepted August 21, 2021, date of publication August 24, 2021, date of current version September 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3107296

Impact of Multi-Vendor Transponders Performance on Design Margin in Optical Networks

ANKUSH MAHAJAN[1], (Member, IEEE), KONSTANTINOS CHRISTODOULOPOULOS[2], RICARDO MARTÍNEZ[1], (Senior Member, IEEE), RAUL MUÑOZ[1], (Senior Member, IEEE), AND SALVATORE SPADARO[3], (Senior Member, IEEE)

[1]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Castelldefels, Spain
[2]Nokia Bell Labs, 70435 Stuttgart, Germany
[3]Optical Communications Group, Signal Theory and Communications Department, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain

Corresponding author: Ankush Mahajan (ankush.mahajan@cttc.cat)

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Actions.

ABSTRACT For reliable and efficient network planning and operation, accurate estimation of Quality of Transmission (QoT) is necessary. In optical networks, a physical layer model (PLM) is typically used as a QoT estimation tool (Qtool) including a design margin to account for modeling and parameter inaccuracies, to ensure acceptable performance. Such margin also covers the performance variations of the transponders (TPs) which are relatively low in a single vendor environment. However, for disaggregated networks that utilize TPs from multiple vendors, such as partial disaggregated networks with open line system (OLS), this traditional approach limits the Qtool estimation accuracy. Although higher TP performance variations can be covered with an additional margin, this approach would reduce the efficiency and consume the benefits of disaggregation. Therefore, we propose PLM extensions that capture the performance variations of multi-vendor TPs. In particular, we propose four TP vendor dependent performance factors and we also devise a Machine Learning (ML) scheme to learn these performance factors in offline and online network planning scenarios. The proposed extended PLM and ML training scheme are evaluated through realistic simulations. Results show a design margin reduction of greater than 1 dB for new connection requests in a disaggregated network with TPs from four vendors. On top of this, the results also show a ∼0.5 dB additional Signal to Noise Ratio (SNR) saving for new connection requests by proper selection of the TPs.

INDEX TERMS Design margin, optical networks, quality of transmission (QoT), transponders.

I. INTRODUCTION

The increasing popularity and rapid development of emerging services and cloud-based applications along with the latest networking paradigms (e.g., Internet of Things) require high capacity and improved transport infrastructure [1]–[3]. To fulfill these ever-increasing services and applications requirements, data traffic will experience a dramatic evolution over the next years [4]. Moreover, this substantial traffic growth will push network operators for a continuous investment in their optical transport infrastructure. To sustain this traffic growth, optical transmission systems do not only need to scale up in transported bits per fiber, but also lower the cost of transmission [5]. One of the key requirements to cope-up with the cost, is to develop highly interchangeable products to enable end-to-end vendor-diverse inter-operable coherent optical systems [6]. Typically, the design and operation of an optical network requires a Physical Layer Model (PLM). The PLM with the addition of design margin is typically referred to as the Quality of Transmission (QoT) estimation tool or Qtool. Such a Qtool is used by optimization algorithms to estimate connections' QoT while examining candidate optimization operations. The range of optimization problems varies from static use cases (e.g., incremental planning) to more dynamic use cases such as dynamic connection establishments, automatic network reconfiguration, defragmentation, virtual network reconfiguration etc. The accuracy of the PLM is quite crucial to achieve high efficiency in such optimizations, and plays an important role to bring forward the benefits of disaggregated optical networks, as discussed in the following.

The associate editor coordinating the review of this manuscript and approving it for publication was Qunbi Zhuge.

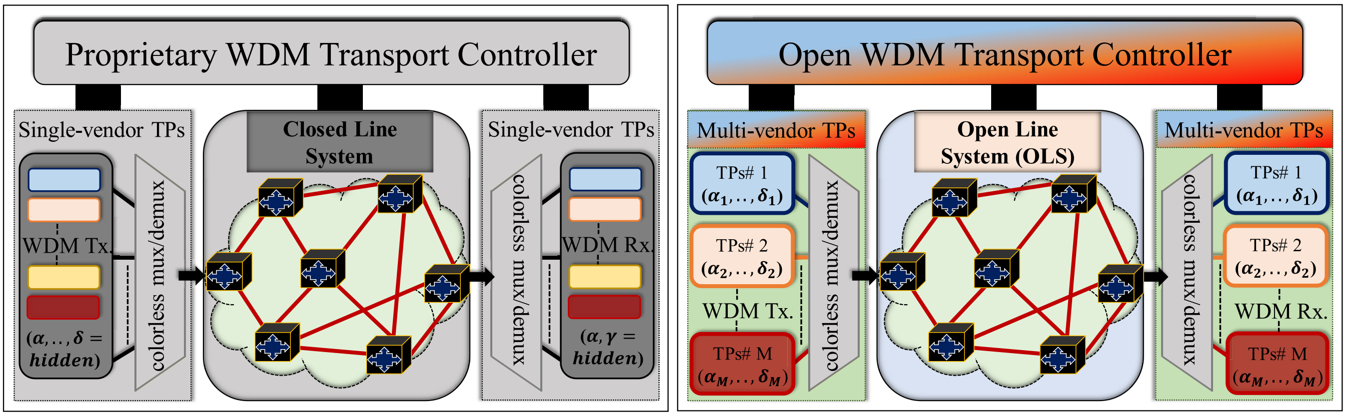


FIGURE 1. **(a) Traditional WDM transport system: line system and TPs with proprietary controller, and (b) OLS of multi-vendor TPs with open transport controller.**

The concept of disaggregation in optical networking is "inherited" from the datacenter's architecture. Datacenters are built upon interchangeable, highly flexible computing and network nodes [7]. This flexible datacenter architecture approach pushes optical networking to explore multi-vendor disaggregating hardware and software with a strong focus on the interoperability. Substantial efforts have been made in developing vendor-neutral software controllers, third party network orchestrators, and standardized northbound/ southbound interfaces. In this regard, some works demonstrated line systems that are open to multiple vendor network elements [7], [8]. Such developments have an ultimate objective to enable an open or disaggregated line system (OLS), where the optical hardware from multiple vendors can be interconnected and configured centrally through a common control plane. Thus, disaggregated optical transport system is considered as a means to accomplish higher flexibility and cost reduction. Different levels of disaggregation are being discussed, from partial to full disaggregation [7]. As expected, disaggregation trades off data/control plane complexity for cost, and the different disaggregation levels achieve different tradeoff levels. Partial disaggregation, where the line system is open (open line system OLS) to multiple vendor transponders (TPs), has gained significant attention due to the ease in the control plane implementation.

While the introduction of disaggregated optical platforms is expected to push for equipment commoditization and generate new business models, there are still some uncertainties regarding the performance of such systems and their applicability to backbone/core networks that have stringent optical performance requirements. One of the most basic requirement for such networks is to have a PLM/Qtool that accounts for vendor dependent performance factors of network elements (e.g. multi-vendor TPs) rather than relying on traditional closed source vendor Qtool estimator and planning tools.

In traditional optical transport networks, both the line system and the TPs are aggregated (*i.e.,* single vendor) and controlled via a proprietary Network Management System (NMS) or controller as shown in Fig. 1a. For such networks, network elements operating parameters are decided by the network proprietary controller. For instance, the TPs launch power is generally set to a fixed value by the vendor itself that is optimized to achieve the best Signal to Noise Ratio (SNR) for the targeted channel under high load, *i.e.,* worst conditions. While optimizing/deciding that, several factors need to be accounted, including statistical variations of TPs' components, even if they come from the same vendor [9]. The used margins, which account for impairments calculations, uncertainties and ageing, generally cover such TP performance variations [10], [11].

In the envisioned OLS scenario, the network infrastructure is expected to comprise of multi-vendor equipment such as TPs which are controlled via well-defined models and control interfaces (e.g., YANG, REST APIs, NETCONF etc.) [7], [12]. The reference architecture for such a multi-vendor network is shown in Fig. 1b. So, for multi-vendor TPs in an OLS, apart from the aforementioned TP statistical variances, vendor dependent factors play a crucial role in QoT/SNR performance and estimation. To be more specific, in such a network, performance variations arise from the different TP components, digital signal processing (DSP), forward error correction (FEC) coding techniques etc. used by each vendor [9], [13]. Consequently, relying on a typical single vendor PLM to estimate the performance of a multi-vendor TP network may result in huge deviations in the estimated and actual QoT values. A solution to mitigate this is to add/increase the used margin on top of the PLM estimations. However, adding such additional margin would lead to lower efficiency and underutilization of the network and thus diminish the benefits of disaggregation.

In light of the above, herein we propose extensions to the PLM/Qtool to accurately model the physical layer in multi-vendor TP environment that accounts for vendor dependent performance factors. Although we focus on multiple vendor TPs, the proposed model can be extended to cover other line system characteristics such as amplifier ripples, filtering penalties etc. Compared to our previous work in [14], two major novelties are brought: i) we devise both offline and online machine learning (ML) assisted training methods to identify the TP vendor dependent performance factors; and ii) we apply the proposed solution to a different use case, that is, the incremental planning problem, rather than the TPs launch power optimization problem investigated in [14].

The remainder of this paper is organized as follows. In Section II, we overview the related work and current limitation in existing PLMs. Then, in Section III, we provide the details about the proposed PLM extensions to model multi-vendor TPs in OLS network scenarios. In the same section, we also provide basic results that emphasize the need for the proposed extensions. In Section IV, we present the ML-assisted training schemes to find the proposed PLM parameters. Then, in Section V, we show the performance evaluation in terms of margin reduction achieved by the proposed modeling scheme. We also verify that additional SNR improvements for new connection requests can be achieved with proper selection of TPs. Finally, Section VI concludes the paper.

II. RELATED WORK

There are several factors which impact the accuracy of the Qtool and cause estimation uncertainty [15]. To mitigate such uncertainty, the design margin is used [10]. The reduction of the design margin during connections provisioning based on the actual network conditions and the actual capabilities of equipment has recently attracted attention from both the research and industrial community [16]–[19]. Monitoring and ML schemes have been adopted as key enablers to reduce the design margin for either planning or incremental planning of the optical networks [19], [20].

In the literature many numerical and non-ML (analytical) based models for QoT estimation exist [6], [17], [18]. For such models, QoT estimation is performed using an analytical PLM which is the key module of the Qtool. In such tools, the linear amplified spontaneous emission (ASE) noise injected by the optical amplifiers, the nonlinear noise caused by the fiber Kerr effects, the filtering penalties at reconfigurable optical add drop multiplexer (ROADM) nodes, the polarization dependent losses (PDL), etc. are the effects that are modelled in the connection's QoT calculation. Modeling simplifications or uncertainties in the parameters of these calculations define the accuracy of the models [15]. Thus, a common practice is to add a design margin to the Qtool to account for the modeling simplification assumptions and other uncertainty parameters. The ASE noise calculations contributed by each erbium doped fiber amplifier (EDFA) in the network is quite straightforward and depends on the average gain and noise figure (NF) values. Many models are available to estimate the contribution of the nonlinear interference (NLI) noise generated due to fiber nonlinear effects [21]–[23]. One of the major bottlenecks in choosing these models for NLI estimation is the computation speed. For example, the split step Fourier method is very accurate and versatile as it can address complex scenarios including the mix of non-coherent and coherent signals in networks with dispersion management. However, the trade-off is with the computation speed, as this method is very slow. The Gaussian Noise (GN) model has been introduced and shown to be quite accurate, while its approximated closed form analytical version combines both good accuracy and low computational complexity [21]. Since then, the GN model became the first choice as PLM for many research works [6], [16]–[18], [24]. Several works explored the estimation of the filtering penalty induced at the ROADM nodes [24], [26]. Some of these methods are analytical whilst others proposed to use monitoring and ML to estimate these penalties by correlating the information from the established connections.

The capability to continuously monitor the network and the QoT of the established connections is the key to lower the design margin. For example, we can use a feedback-based QoT estimation approach, which correlates the monitored QoT values of established connections to estimate the QoT of the unestablished ones. Based on similar fundamentals, the penalties due to the EDFA gain ripple and filtering at the ROADM nodes are estimated in [24], [25]. In many past works, monitoring along with ML is exploited to reduce the design margin by minimizing the uncertainty of the input parameters of the Qtool [15]–[19]. ML-based estimation techniques have also gained a lot of attention to improve the Qtool accuracy in the past [17], [27]. One of the major benefit of these full *black box* Qtool is to model all the physical layer effects together. However, the need of a huge training dataset is one of the biggest limitations of such approaches. Hybrid approaches (a mix of PLMs and real-world data) also came up recently with comparatively less data requirements [15]. All these trained models at the end are used to attain more accurate QoT estimations embracing diverse use cases. Almost every scheme in these mentioned research works utilized the traditional GN-based PLMs. This approach takes into account different attributes such as channels separations, symbol rate, transmitted wavelength, launch power etc. However, these PLMs lack the information about the vendor characteristics, which are crucial for multi-vendor network scenarios. The common approach is to increase the amount of design margin to account for these vendor dependent performance factors. However, this would result in inaccurate QoT estimation (as margin is increased) along with network underutilization.

In brief, to the best of our knowledge, the QoT estimation accuracy and design margin reduction in multi-vendor TPs with OLS network scenarios have not been investigated. For such networks, a proper PLM is needed to account for TP vendor dependent performance factors. Keeping this in mind, we first propose a PLM to capture the performance variations of multi-vendor TPs. Then, a ML-assisted training scheme to trace these TP vendor dependent performance factors is developed. We verify the benefits of our proposed modeling scheme in terms of substantial design margin reduction for new connection requests. We also highlight the idea of using the proposed PLM with basic resource allocation strategies to decide the best TPs at the time of setting up the new connection requests along with the additional SNR improvements.

III. PRELIMINARY STUDY AND MOTIVATION

In traditional aggregated optical network with proprietary controller (Fig. 1a), the TPs parameters are only known by the network or domain vendor. In such a network setting, several past works already proved that the QoT of connections can be estimated quite accurately using the GN model for fiber nonlinearities [15]–[18]. The GN model's main assumption is that, in dispersion uncompensated transmission systems where the NLI caused by the Kerr effect is relatively small. Consequently, the NLI can be modeled as an additive Gaussian noise that is statistically independent of signal and ASE noise [21]. The GN model is also a well-accepted PLM for single- and multi-vendor networks [6], [16]–[18], [28].

We assume a network with $n = 1, 2, \ldots, N$ established connections. Let $\boldsymbol{p} = p_1, p_2, .., p_N$ represents the TP launch power vector of those established connections. The GN PLM is a model that takes as input several parameters and calculates the generalized SNR values of the connections. Let $\boldsymbol{z}$ represents the *fixed* input parameters of PLM, such as routes, used wavelengths, span lengths, etc. Let $\boldsymbol{r}$ denote the set of GN model fitted parameters: i) fiber attenuation coefficients, ii) fiber non-linear coefficients, iii) fiber dispersion coefficients, and iv) a bias. According to the GN model, the impact of optical fiber transmission effects on the generalized SNR of connection $n \in N$ generated at span s can be modeled as

$$SNR_{n,s}\left(\boldsymbol{p}, \boldsymbol{r}, z\right) = \frac{P_{o,n,s}}{P_{ASE,n,s} + P_{NLI,n,s}} \qquad (1)$$

where $P_{o,n,s}$ is the optical signal average power level, $P_{ASE,n,s}$ is the ASE noise power, and $P_{NLI,n,s}$ is the NLI contribution to the noise of connection n generated at span s.

Assuming incoherent noise accumulation over the spans of the path for connection n, we can sum $P_{ASE,n,s}$ and $P_{NLI,n,s}$ over the spans that comprise the path of n, to obtain $P_{ASE,n}$ and $P_{NLI,n}$, respectively. Then the generalized SNR of connection n at the path end, $SNR_n\left(\boldsymbol{p}, \boldsymbol{r}, z\right)$, is given by

$$SNR_n\left(\boldsymbol{p}, \boldsymbol{r}, z\right) = \frac{P_{o,n}}{P_{ASE,n} + P_{NLI,n}} \qquad (2)$$

Then, the total SNR of connection n, calculated by the traditional single vendor Qtool Q_{SV} (using Eq. (2)) is given by

$$Q_{SV,n}\left(\boldsymbol{p}, \boldsymbol{r}, z\right) = \left[SNR_n\left(\boldsymbol{p}, \boldsymbol{r}, z\right)\right]_{dB} - SNR_{b2b} - DM_1 \quad (3)$$

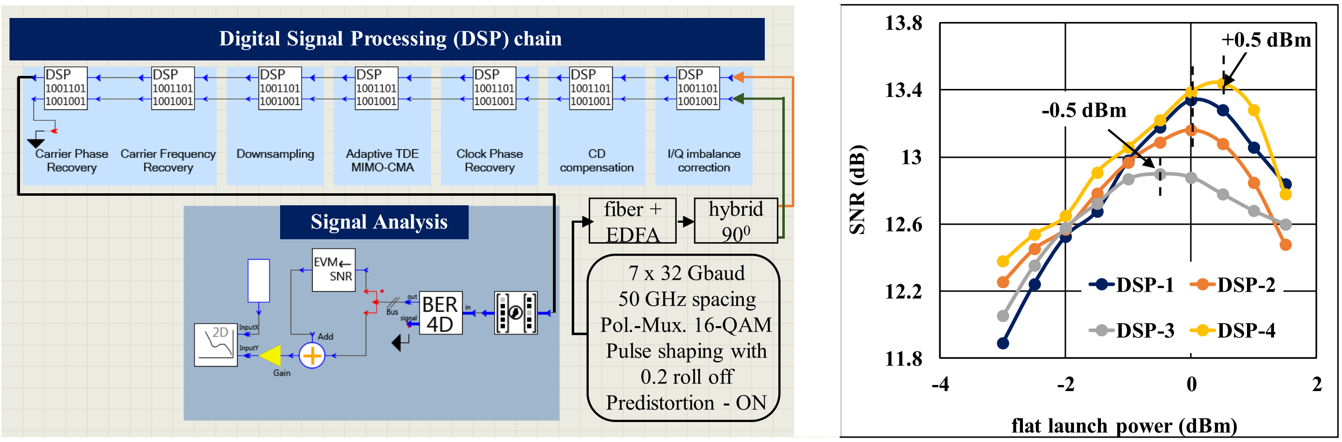
where SNR_{b2b} is the dB penalty in TP's back to back (b2b) configuration and DM_1 stands for the design margin, which is the additional margin (in dB) added on top of PLM calculations to cover modeling inaccuracies such as EDFA gain ripple penalties, TPs performance variations etc. Eq. (1) to Eq. (3) collectively form the traditional model/Qtool to estimate generalized SNR. We call it as a single vendor Qtool, and refer to it with Q_{SV}.

The above described Q_{SV} considers in a coarse manner the characteristics of the TP. In general, the average power of the output signal and NLI noise, that is $\{P_o, P_{NLI}\}$ terms, are affected by receivers' characteristics of the TPs such as the digital signal processing (DSP) implementation, performance variations of TP components (e.g., laser linewidth, photodiode's responsivity, etc.). The linear noise term, P_{ASE}, is mostly determined by the optical amplifiers. Parts of the TP characteristics are included in the SNR_{b2b} and then the uncertainties/variations of the TP performance are covered in the design margin DM_1. This can be acceptable for a single vendor TP with small variations.

A. MOTIVATION

Considering optical networks with line system open to TPs from multiple vendors it stands to reason to expect to have higher variation in TPs performance as compare to a single vendor environment [13]. The DSP chain, which is generally implemented by different vendors in different ways, such as different algorithms or the same algorithms with different parameters (e.g., number of digital filter taps), is one of the major source of variation in multi-vendor TPs performance. Furthermore, different vendors use different components (from different third-party vendors) such as balanced photodiodes, local oscillator (LO) lasers (drifts, linewidths, etc.), analog amplifiers, etc. As discussed, there are statistical variances within the TP components which cause performance (statistical) variations even in single-vendor TPs. However, in OLS scenarios with TPs from multiple vendors, the effects of these variations in the overall network performance become more critical.

Keeping this in mind, we simulated seven polarization-multiplexed (pol.-mux.) 16-Quadrature Amplitude Modulation (QAM) channels, spaced at 50 GHz and modulated at 32 Gbaud symbol rate with root raised cosine (RRC) pulse shaping (roll off-0.2) in VPI Transmission Maker [29]. We simulated the setup with total fiber/link length of 160 km and 480 km with 2 and 6 identical spans of length 80 km each respectively. The EDFAs (or the line system part) were assumed to be completely flat, to capture the performance of the TPs. The total impairments compensated at the receiver DSP block is shown in Fig. 2a. We also varied the balanced photodiodes responsivity at the coherent receiver front-end from 0.8 (worst) to 1.0 (ideal/best) and the LO laser linewidth to emulate the component statistical variation originating from different vendor components. In chromatic dispersion (CD) compensation module, we varied the effective fast Fourier transform (FFT) size and the phase noise component. For polarization demultiplexing algorithm, we varied the constant modulus algorithm (CMA) and the multi- modulus algorithm (MMA) with different initial taps and number of iterations. We also varied the number of samples during clock phase recovery module, in order to emulate performance variation within different vendors DSP chains. Inside the same module, we also varied the fourth power operation on the received samples, before estimating the frequency offset. Lastly, in the carrier phase recovery module, we only varied phase noise parameter from 0 to 20 radians. Based on the different combinations of the above DSP algorithms and parameter variations, we implemented four different DSP



(a). (b).

FIGURE 2. **(a) Simulated set up in VPI transmission maker in order to emulate different DSP chains for different TP vendors, and (b) different value of flat optimized launch power for different DSP chain (TP vendor) at a transmission distance of 480 km (6 × 80 Km spans).**

chains at the receiver side (DSP-1 to DSP- 4), to emulate TPs from four different vendors.

As expected, due to the maximum nonlinear interference noise from the neighboring channels, the central channel has the worst performance/lowest SNR in a wavelength division multiplexed (WDM) system for both transmission distances. As so, we measured the SNR at the central channel while varying the flat or uniform launch power vector for all channels. For a transmission distance of 480 km, from Fig. 2b, it can be seen that depending upon the DSP chain or TPs performance, the best optimized flat launch power is in the range of −0.5 dBm to +0.5 dBm for the different (vendor) TPs. Similar behaviour was also observed for shorter link length of 160 km. Note that each vendor, if it would be the sole vendor (aggregated network) would perform such optimization, considering the network specificities, operation load, etc., and select the corresponding optimized power. It is also worth noting that while some vendors' launch power (e.g., TP with DSP-4) is optimized, other TPs (DSP-1, DSP-2 and DSP-3) may be in their nonlinear range.

However, assuming an environment with TPs from multiple vendors, such an optimization would not be feasible, and related variations in the performance would need to be covered by a corresponding increased margin as discussed in the upcoming Section IV. Thereby, assuming that we use the GN model as the PLM, we need to increase the DM_1 in Eq. (3) to account for the performance differences caused by the TP's different characteristics. Increasing the margin results in underutilization of network capacity as certain TPs deployed in the network have better performance than others at certain conditions (e.g., network load). Our goal is thus to extend the GN model based PLM to capture the TP characteristics in a generic way, so as to reduce the margin required for multi-vendor TP environment.

B. MULTI-VENDOR PLM

To extend the GN model to capture TP characteristics, we introduce four performance factors, $\boldsymbol{v} = \{\alpha, \beta, \gamma, \delta\}$, where α and δ cover vendor specific TP components variations; β covers amplifier characteristics; and γ covers vendor specific DSP implementation variations [9], [13], [14]. In a multi- vendor TPs scenario, these performance terms $\{\alpha, \beta, \gamma, \delta\}$ would be different for the heterogeneous TPs. Though this model is also applicable for single vendor TPs (or networks with alien wavelength/TPs), its importance is more relevant in multi-vendor networks. To be more specific, we consider a scenario, where the line system is open and TPs from M vendors are deployed (or available for deployment for the new connections).

For a vendor i in M, we calculate the SNR of connection n, which uses TP i (denoted as $i = TP(n)$) at the end of the path, with Eq. (4), instead of Eq. (2):

$$\begin{aligned} SNR_n\left(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{v}, z\right) &= \frac{\alpha_i.P_{o,n}}{\beta_i.P_{ASE,n} + \gamma_i.P_{NLI,n}} \\ &= \frac{\left(\frac{\alpha}{\beta}\right)_i.P_{o,n}}{P_{ASE,n} + \left(\frac{\gamma}{\beta}\right)_i.P_{NLI,n}}, \quad i = TP\left(n\right) \end{aligned} \tag{4}$$

where $\boldsymbol{v}$ denote the TPs' parameters vectors that includes $\boldsymbol{v_i} = \{\alpha_i, \beta_i, \gamma_i, \delta_i\}$ the performance factors of transponder i.

The QoT/SNR of connection n, calculated by the Qtool that accounts for the proposed multi-vendor TPs dependent performance factors is given as

$$\begin{aligned} Q_{MV,n}\left(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{v}, z\right) &= \left[SNR_n\left(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{v}, z\right)\right]_{dB} - SNR_{b2b,i} \\ &\quad -DM_2, \quad i = TP\left(n\right), \\ SNR_{b2b,i} &= SNR_{b2b} + \delta_i \end{aligned} \tag{5}$$

where $SNR_{b2b,i}$ is the total dB penalty for i-th TP in back to back (b2b) configuration, and δ_i is its vendor dependent variation to some reference SNR_{b2b} value (similar to Eq. (3)).

The performance factor β corresponds to the amplifiers' performance and its contribution comes from the OLS. An important effect that can be captured in the β parameter is the wavelength dependent penalty (additional ASE noise) due to the EDFA gain ripple effect. Several works have been published targeting to estimate the penalties contributed by this effect [15], [24], [25]. However, this would be wavelength dependent and common for all TPs, and thus it will be a network (line system) rather than a TP vendor dependent factor. Thus, for the remaining of this paper we will assume $\beta = 1$ for all TPs.

Eq. (4) and Eq. (5) account for vendor specific performance factors. We call this as a multi-vendor Qtool and refer to it by Q_{MV}. Note that in the past, a vendor dependent bias term was added in the accumulation of linear/ASE noise to account for TP implementations [30]. Our model, on the other hand, is more generic and captures a broader range of implementation factors, including bias [14]. Furthermore, while we started with the GN model and described our extensions to it in the preceding definitions, the concept is generic, and our proposed extensions can be applied to other PLMs/Qtools.

Q_{MV} includes performance factors coming from different vendor TPs. Hence, when modeling a multi-vendor TP network, the margin DM would be higher, if we use a single vendor/traditional Q_{SV} compared to the margin used with the proposed multi-vendor Q_{MV}, that is, $DM_1 \gg DM_2$ (demonstrated in results, Section V). Note that in this work, we assume that for the connections, the transmitter – receiver (Tx. -Rx.) pairs are from the same vendor. In more diverse network disaggregation scenarios, where interoperability between Tx. - Rx. from different vendors is possible, such communication would follow a specific standardized configuration (modulation format, DSP, etc.). In such a case, we should have specific TP performance factors $\boldsymbol{v_i}$ for the standardized interoperable configurations, allowing the proposed Q_{MV} to achieve good accuracy also in such a network.

IV. CASE STUDY–FLAT POWER OPTIMIZATION USING Q_{SV} AND Q_{MV}

In this section, we discuss the discrepancies in the SNR values, if Q_{SV} is used in networks where TPs from multiple vendors are deployed (or available for deployment) and how this could affect the optimization of the launch power of the channels.

In traditional or single-vendor optical networks, all network elements are *aggregated* and under the control of the proprietary controller, as shown in Fig. 1a. Every network element configuration such as launch power of TPs, operating points of amplifiers etc. are decided by the vendor. One of the most crucial decisions by the vendor is to adjust the TPs launch power. There are several ways to decide the launch power of the TPs in a WDM system. The most common practice, and the one discussed in the previous section, is to set a flat launch power to all TPs which is found to be the optimum for the central (target) channel at full load (worst condition) on a multi-span link [14]. Such power optimization results in low efficiency or excess margins for side channels or diverse network paths and can be improved using more sophisticated techniques [31], [32]. Setting a conservative launch power is definitely a good strategy for networks with limited knowledge about the used elements such as the disaggregated scenario [7]. In general, due to these uncertainties,

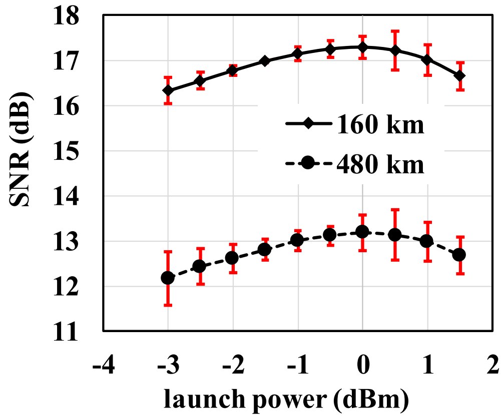
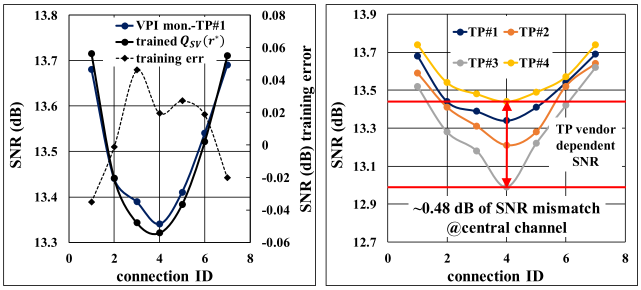


FIGURE 3. Average and minimum-maximum SNR (dB) variation on the central channel for different TP vendors, with respect to varying launch powers at different transmission lengths.

the impact of the network disaggregation is accounted for by considering an additional penalty or margin as discussed in [6], [10] and also at the end of the previous paragraph.

Fig. 3 shows the average SNR value of the central channel SNR (dB) of the four different (DSP) TP vendors at varying transmission distances of 160 km (2 spans) and 480 km (6 spans). Note that the number of channels (seven) was assumed to be fixed. The curves show in addition to the average SNR (dB) performance across all TP vendors, the minimum-maximum SNR (dB) in the error bars (red color) indicating the variation of the central channel for the different TP at each flat optimized launch power. Because of the various DSP implementations, Fig. 2b already showed that the optimal channel launch power for each TP vendor was within 1 dB of one another. As so, when setting a flat launch power for all TP, as done in this set of simulations, it will not be optimal for some TPs. We noticed a similar behavior at different transmission distances of 160 km and 480 km. In case of smaller distances (160 km), the maximum SNR (dB) variation was found to be ~0.4 dB. However, for longer transmission distance (480 km), we observe slightly higher variation of ~0.59 dB in maximum average SNR (dB) as shown in Fig. 3. Note that the channel count was low and fixed. The VPI-implemented DSP chains, which act as different vendor TPs, account for the majority of these SNR variations. It is possible that the peak SNR (dB) variation further increase to greater than 1 dB, when full network (high load, diverse paths, etc.) are considered, as presented in the results section.

Let us now evaluate the discrepancies in SNR values if we have a network, where TPs from $M = 4$ different vendors are deployed (or available for deployment). Note that the four different DSP chains implemented in VPI (described in Section III.A) are treated as four different TP vendors in this section. Fig. 4a shows the VPI measured SNR values for the seven pol.-mux. 16-QAM channels at 0 dBm of flat/uniform launch power, when the DSP chain of TP#1 is considered for a link length of 480 km. As can be seen, the minimum SNR is obtained at the central channel, which is to be expected.



(a). (b).

FIGURE 4. (a) SNR (dB) values for VPI measured and trained Q_{SV} for seven channels at 480 km of link length considering DSP-1 chain, and (b) TP vendor dependent (different DSP chains) SNR (dB) values at flat optimized launch power values of −0.5 dBm for TP#3, 0 dBm for TP#1 and TP#2 and 0.5 dBm for TP#4 as presented in Fig. 2b.

Fig. 4a also shows the SNR values for these seven channels estimated after training the parameter vector $\boldsymbol{r}$ (fiber coefficients and bias) of the standard Qtool Q_{SV} (training details in Section IV.A). We also indicated the SNR (dB) training error between trained single-vendor Qtool and VPI measured values in the same plot, with maximum SNR (dB) mean square error (MSE) of $\sim 5.8 \times 10^{-03}$. We also noticed similar training error ($\sim 6 \times 10^{-03}$) for the transmission distance of 160 km. Note that, to improve the training accuracy with more samples, we also used VPI monitored SNR values at different spans lengths (during training). This is the typical outcome we should expect when training the standard Q_{SV} for the single vendor environment. A key point that needs to be highlighted here is that the results would be the same, whether we assumed a single TP vendor or multiple TP vendors in Q_{SV}.

Now considering the scenario with $M = 4$ TP vendors (with different vendor dependent DSP chains). If we use the above trained single vendor Qtool Q_{SV}, since there is no TP/vendor dependent parameters in Q_{SV}, the estimated SNR of all TPs would be the same, as shown in Fig. 4a. Instead, as shown in Fig. 4b, the real SNR values of the four TPs vary (according to their flat optimized power of Fig. 2b), and in particular we observe a ~0.4 dB difference in the SNR value of the central channel. So, a ~0.4 dB of higher margin would be required to cover this estimation error. It is important to note that this is a special case, a linear network of two spans, where all the paths have the same length. For a more realistic network with diverse links and paths, this margin value is higher, as discussed in Section V.

By definition, no TP vendor dependent performance factors are accounted in Q_{SV} calculations, as opposed to Q_{MV}. The vendor dependent factors accounted in Q_{MV} can improve the QoT estimation accuracy and enable the use of a lower design margin. In our previous work [14], we tackled the problem of TPs launch power optimization, where it was assumed that we can choose for each channel a different launch power, according to its needs, being a more sophisticated solution than the flat power optimization we discussed above. We showed notable improvements into two objective functions using Q_{MV} instead of Q_{SV}. In this work, we will target a different optimization problem, *i.e.*, incremental planning, and thus showcase that the proposed PLM/Qtool modeling is generic and can benefit a variety of use cases.

A. MACHINE LEARNING ASSISTED MODEL TRAINING

As shown in Fig. 4b and discussed above, the SNR values vary for the different TPs. This cannot be captured by a single vendor PLM as in Q_{SV}. For this purpose, we proposed (in Section III.B) the multi-vendor PLM Q_{MV}, which includes additional – vendor or TP specific – parameters $\boldsymbol{v}$. We now discuss how we can identify these parameters, which was not explicitly provided in our previous work [14]. This can be done through ML-assisted training in a greenfield deployment or while the network operates (brownfield).

The idea in such training is to fit the parameters/coefficients of the Q_{MV} with measurements/monitoring data in a testbed or in the operating network (reality/ground truth), so that the proposed Q_{MV} behaves similarly to the real (multi-TP) world. In this work, we relied on ML based nonlinear fitting techniques to do this. Since we consider QoT estimation, the SNR and the bit error ratio (BER) are the typical estimation targets [15]. We choose the former, *i.e.*, the SNR, as the targeted value for our study. Regarding the input data/features, they include parameters such as connection's route, modulation format, symbol rate, launch power vector etc. However, the most important input parameter is the power level of the connection, since many effects (e.g., interference noise, amplifier gain values, etc.) depend on that. Thus, we should include variable power level measurements in our training scenario to identify the related PLM and TP coefficients.

To be more specific, in our proposed model, we assume that the input data includes the TPs launch power vector $\boldsymbol{p}$, whereas the target data is the monitored SNR vector (on input power vector $\boldsymbol{p}$). Let $Y_n(\boldsymbol{p})$ denote the monitored SNR value of connection n which uses transponder $i = TP(n)$, and $\boldsymbol{Y}_N(\boldsymbol{p})$ denote the vector for all the connections n in N. The monitoring is assumed to be done at the coherent receivers implementing the vendor dependent DSP chains. Note that in the specific example that we consider, seven connections are established whose monitored data is gathered from VPI setup as described in Section III.A. In VPI, we implement the DSP chain that is configured with four sets of parameters to emulate the effect of four different TPs/vendors.

According to the used PLM, and in particular whether it calculates NLI based on the actual utilization or with coarser calculations, a change in the launch power of a single connection impacts on the SNR values of several others (*i.e.*, those connections that interfere/cross the same connection). In our case the implemented PLM model is the analytical version of the GN model with detailed NLI calculations per span [21].

To make a multi-vendor PLM, we first identify the parameters of the PLM model, and in particular the Q_{MV}, that need to be fitted/trained. In this work, we select the following two sets of parameters to be trained based on the monitored information:

i) $\boldsymbol{r}$ – fiber coefficients- attenuation, non-linear coefficients, dispersion coefficients, and a bias (TP independent) [17], [28].
ii) $\boldsymbol{v_i}$ – i-th TP vendor dependent performance factors $[\{\alpha_i, \gamma_i, \delta_i\}$ for all $i = 1, .., M]$, where M is the number of different TP vendors (= 4 in our simulated scenarios).

The generic goal is to apply a ML-based non-linear fitting technique to fit the parameter vector $\boldsymbol{r}$ and/or $\boldsymbol{v}$ of the Q_{MV} with the monitored SNR information, that is $\boldsymbol{Y_N(p)}$. For this nonlinear fitting part, several approaches are available in the literature, however, we relied on the Levenberg- Marquardt (LM) algorithm which is suitable for solving nonlinear least squares fitting problems [33]. The LM algorithm finds

$$\boldsymbol{r}^*, \boldsymbol{v}^* = (\boldsymbol{r}, \boldsymbol{v})^* = argmin_{\boldsymbol{r},\boldsymbol{v}} \left(\boldsymbol{Q_{N,MV}}\left(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{v}, z\right) - \boldsymbol{Y_N}(\boldsymbol{p})\right)^2 \tag{6}$$

where the asterisk (*) symbol represents the corresponding trained parameter vector. Note that we assume known allocation of TPs to the connections, which could be included in z. If we focus on a particular transponder type i, then we can write the above Eq. (6) as:

$$\begin{aligned} \boldsymbol{r}_i^*, \boldsymbol{v}_i^* &= (\boldsymbol{r}, \boldsymbol{v})_i^* \\ &= argmin_{\boldsymbol{r}_i,\boldsymbol{v}_i} \left(\boldsymbol{Q_{N_i,MV}}\left(\boldsymbol{p}, \boldsymbol{r}_i, \boldsymbol{v}_i, z\right) - \boldsymbol{Y_{N_i}}(\boldsymbol{p})\right)^{\boldsymbol{2}} \end{aligned} \tag{7}$$

where N_i denotes the set of connections using transponder i, $\boldsymbol{Y_{N_i}}$ is the vector of monitored values and $\boldsymbol{v_i} = \{\alpha_i, \gamma_i, \delta_i\}$ for fixed $\beta_i = 1$, correspond to transponder i.

Note that we have two training options: i) *joint training,* where we combine all training sets for all TPs and train globally, as described in Eq. (6), and ii) *separate training,* where we train independently for each TP type i, as described in Eq. (7), in which case we obtain $\boldsymbol{v}_i^*$ and M different $\boldsymbol{r}_i^*$ vectors. There are various methods to combine these $\boldsymbol{r}_i^*$ vectors and improve this fitting. In particular, we average the $\boldsymbol{r}_i^*$ vectors, make them constant, and rerun the fitting of Eq. (7) to only obtain the TP performance factors $\boldsymbol{v}_i^*$ for each TP i.

We now focus on two different ML-assisted training schemes, depending upon the use case, for deriving the parameter vectors $\boldsymbol{r}^*, \boldsymbol{v}^*$.

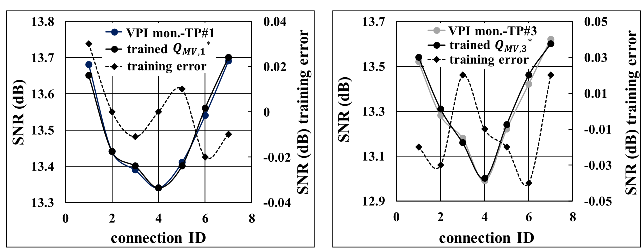
B. ML ASSISTED OFFLINE TRAINING

This approach is applicable to scenarios where the TPs from M different vendors are available to characterize prior to the field deployment, *i.e.,* the planning phase or greenfield deployment. Each vendor TP is characterized by a possible set of modulation formats, symbol rates, launch power range, etc. To plan the network, a PLM/Qtool would be available and used by the network planner to assess the connections' QoT prior to their establishment. In our case, this tool is the GN model based multi-vendor Qtool, *i.e.*, Q_{MV}.

We assume that we generate N_c connections with the central channel $\left(n_i = \frac{N_c+1}{2}\right)$ being of TP type i at power vector $\boldsymbol{p}$, and we repeat this for all the available TP types. We then measure the SNR value $Y_{ni}(\boldsymbol{p})$, of the central channel. The launch power vector $\boldsymbol{p}$ is then varied (for all channels) over a range of values, such as from p_{min} to p_{max} (e.g. −4 dBm to +4 dBm) at a step of p_{step} (e.g. 0.5 dB). The SNR is monitored for the central channel n_i over the tuned power range and stored in a vector denoted by $\boldsymbol{Y_{ni}}$. As discussed, we need to choose the parameters to train, which we identified to be the sets $\boldsymbol{r}$ and/or $\boldsymbol{v}$ in the previous subsection. Generally, we envision to perform this offline training phase in a lab environment, as so accurate information about parameter vector $\boldsymbol{r}$ of the real network would be hard to obtain and would not be needed. Hence the goal in such an offline training phase is to fit only parameter vector $\boldsymbol{v}$ for a fixed / known $\boldsymbol{r}$ vector which is obtained by knowing the specifications of the implemented setup (e.g., spans, attenuators, EDFAs, etc.) or by fitting these parameters independently from the TPs. To identify $\boldsymbol{v}$, we change the launch powers of the TPs and monitor the SNR vector. If available, we repeat with another TP of the same type in the center and/or move the TP to different locations apart from the center to enrich further our training set. After training with LM algorithm – Eq. (7), the fitted $\boldsymbol{v}_i^*$ is obtained for vendor i and used in the $Q_{MV,i}$. Then we repeat this for the other TP/vendors. Fig. 6a provides the pseudocode for offline training phase for TPs from M different vendors.

TABLE 1. Ground truth/reality (Q_{MV}^{GT}).

Q_{MV} model fitted parameters	fiber coefficients and bias ($\boldsymbol{r_i} = \boldsymbol{r} = fixed$)				TP performance factors ($\boldsymbol{v_i}$)		
	att. (dB/km)	Non-linear (W/km)	disp. ($ps/(nm.km)$)	bias (dB)	α_i	γ_i	δ_i
TP1	0.21	1.36	17.19	-2.6	0.81	0.78	0.85
TP2					0.82	0.96	0.72
TP3					0.96	0.86	0.91
TP4					0.82	0.84	0.94



(a). (b).

FIGURE 5. Standard and trained (a). $Q_{MV,1}$, and (b) $Q_{MV,3}$ with LM algorithm on VPI monitored dataset for TP#1 and TP#3 respectively along with training error.

We now discuss the results of applying this ML-assisted training method to fit the measurements obtained through VPI for the four TPs presented in Section III.A and Fig. 4b. For this specific example, Fig. 5a and Fig. 5b show the trained multi- vendor Qtool on TP#1 and TP#3 monitored data of VPI, respectively. The LM-based training clearly fitted the

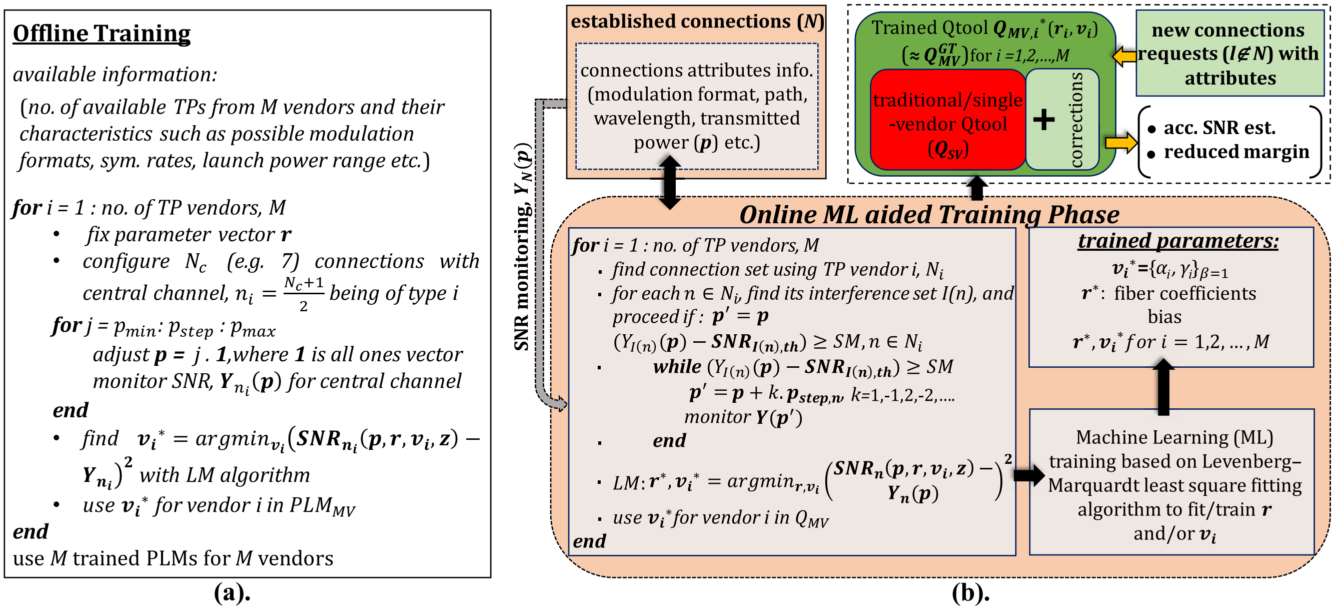


FIGURE 6. Pseudo code for (a). offline, and (b) online training phase to identify TP vendor dependent performance factors.

Q_{MV} parameters ($\boldsymbol{v_1}$ and $\boldsymbol{v_3}$ and other TPs are not presented in the figure for the sake of brevity) very well, with maximum absolute training error of less than 0.05 dB (for all TPs). Similar training error was also noticed for shorter links of length 160 km. It is worth noting that in this performed fitting, we trained GN- based Q_{MV} with data measured in VPI. This indicates that the proposed TP modeling and the GN-model extensions can capture the performance variations of another more realistic model and gives us confidence that they can capture the related effects in real networks/fields.

Table. 1 indicates the fitted set of $\boldsymbol{v_i} = \{\alpha_i, \gamma_i, \delta_i\}$ for the four TP vendors, after the training by the LM algorithm (VPI setup at a length of 160 km). Identical TP parameters are observed for link length of 480 km with slightly different GN model fitted parameters. Note that the $\boldsymbol{r}$ parameters vector was assumed to be known/fixed for all TPs, as discussed above and also indicated in Table. 1, so that the training focuses on the TP vendor dependent performance factors.

C. ML ASSISTED ONLINE TRAINING

The online training of the Qtool is applicable for more accurate QoT/SNR estimation in an operating network. The idea is to derive the TP performance factors from the established connections and use them in the Q_{MV} for any future optimization operation. One of the main assumption of this technique is to know the TPs' category for the already established connections from the pool of M vendor TPs. Let us assume N established connections with the power vector $\boldsymbol{p}$. Again the goal is to find $\boldsymbol{v_i} = [\{\alpha_i, \gamma_i, \delta_i\}, \text{for } i = 1, 2, .., M]$ similar to offline training. However, the network is operational and the training similar to that of an offline network is an issue since, changing powers may render some connections infeasible, resulting in the disruption of some services.

We denote by $i = TP(n)$ the transponder type of connection n. We also denote by N_i the set of connections using TP from vendor i (with total TP vendors M). For the connection n let us denote by $I(n)$ the set of connections that n interferes. We choose connection n so that

$$(\boldsymbol{Y_{I(n)}}(\boldsymbol{p}) - \boldsymbol{SNR_{I(n),th}}) \geq SM, \quad n \in N_i \tag{8}$$

where SM stands for safety margin (e.g. = 1 dB) and is chosen to avoid connections in $I(n)$ to reach infeasibility level. $\boldsymbol{SNR_{I(n),th}}$ is the SNR threshold vector for $I(n)$ set of connections, depending upon their modulation format. In this work, we used SNR threshold calculated at a bit error ratio of 1×10^{-2}, from the formulas indicated in the Appendix A of [34].

The next step is to vary $\boldsymbol{p}$ and collect the training dataset. We denote by $\boldsymbol{p_{step,n}}$ the vector that includes all zeros and a value of p_{step} (e.g. = 0.5 dB) for connection n. For each TP vendor i, we identify candidate connections n in N_i with Eq. (8) and then change the initial $\boldsymbol{p}$ vector to $\boldsymbol{p}+k.\boldsymbol{p_{step,n}}$, $k = -1, 1, -2, 2, \ldots$.. This change in launch power vector $\boldsymbol{p}$ is performed for values of k until Eq. (8) stops to be satisfied. So we start with small positive and negative values of k and increase/decrease it as long as the criterion of Eq. (8) is met. This is described in more detail in Fig. 6b. In this way we obtain a set of $\boldsymbol{Y_n}(\boldsymbol{p})$ for different $\boldsymbol{p}$ and the next step is to apply fitting as discussed in Section IV.A (Eq. (6) and Eq. (7)).

The evaluation of this online training is more complicated than the offline one described above and necessitates network level simulations. Since VPI simulations cannot achieve network wide simulations (computation time constrain), it cannot be used as the ground truth. Therefore, we replaced it with a faster model, the proposed Q_{MV}^{GT}. Then, the ground truth and the fitted model are both the same. The model used as the ground truth has several parameters (span fiber coefficients, bias, etc.) that are hidden/unknown and are fitted in the second model. Note that this might have lower level of realism than using two different models as done above, but we have verified in the previous subsection (Section IV.B) that the proposed model fits very well to VPI, which is considered to be a very realistic and accurate model.

V. RESULTS AND DISCUSSIONS

To quantify the benefits of the proposed Qtool extensions in multi-vendor TPs network scenarios, we performed network level simulations. For this analysis, we consider an Italian backbone topology with 27 nodes and 43 bidirectional links as shown in Fig. 7 and is similar to our past work [14]. The link lengths range from 80 to 480 km. We assumed standard single mode fiber (SSMF) spans of 80 km, a traffic load of 500 connections with uniformly chosen source-destination nodes.

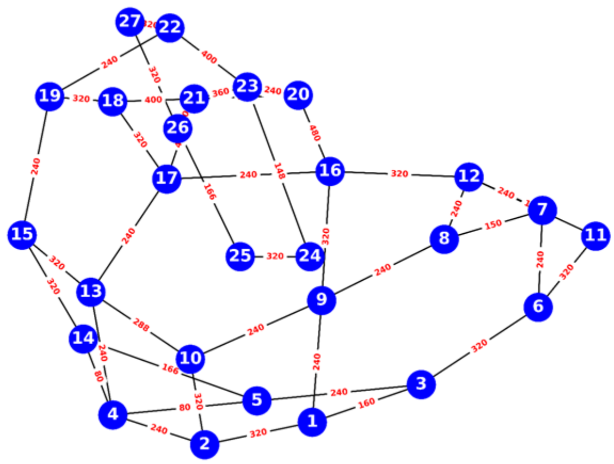


FIGURE 7. Simulated network topology with 27 nodes and 43 bidirectional links.

We assumed attenuation coefficient of 0.2 $dB.\ km^{-1}$, dispersion coefficient of 16.7 $ps.\ (nm.\ km)^{-1}$, and nonlinear coefficient of 1.3 $W.\ km^{-1}$ for SSMF. We considered $M = 4$ TP vendors with performance factors $\{\alpha_i, \gamma_i, \delta_i\} = [\{0.81, 0.78, 0.85\}, \{0.82, 0.96, 0.72\}, \{0.96, 0.86, 0.91\}, \{0.82, 0.84, 0.94\}]$ and fixed $\beta = 1$ obtained after proper training of multivendor Qtool on VPI monitored data (Fig. 4). Table. 1 shows the values for $\boldsymbol{r}$ and $\boldsymbol{v_i}$ that were set to Q_{MV}^{GT} which was used as the ground truth in these simulations. Each demand was served with one wavelength, assumed to be modulated at symbol rate of 32 Gbaud with uniformly chosen pol.-mux. transponders from $M = 4$ vendors. We assumed modulation-tunable TPs that could adapt to {QPSK, 8-QAM, 16-QAM} modulation formats, leading to {100, 150, 200} Gb/s of datarate, respectively. We assumed a frequency slot size of 12.5 GHz in the C-band and allocated 3 spectrum slots (so 37.5 GHz) for each 32 Gbaud connection. We consider a stable network state, where a specific set of connections are established and the goal is to establish a new set of connections, in a network upgrade/incremental planning phase.

We assumed a specific set of connections, which we routed using a routing and spectrum assignment (RSA) algorithm based on shortest path and first fit slots. Typically, routing and wavelength assignment (RWA) algorithms refers to a 50 GHz grid, although they might be used for cases where all connections have the same bandwidth. However, we would like to point out that the proposed solution can be applied for connections of any bandwidth. Then we generated monitored data using the ground truth Q_{MV}^{GT} (parameters indicated in Table. 1). Then these parameters were assumed unknown and estimated through the proposed ML training technique. So, we trained a second multivendor Qtool Q_{MV} on these established connections and their monitored dataset. We assumed that we have a new set of connection requests of 10% of the total load. For example, at a load of 100 established connections which were used for training, additional 10% (10 new connections) were generated and needed to be established. For these new connections, we estimated their QoT with our trained Qtool and calculated the testing error with respect to the Q_{MV}^{GT}. To verify the stability of the trained Qtools, we averaged the results over 100 iterations at each load.

It is also worth noting that, the problem at hand is a nonlinear fitting problem. As so the LM algorithm can find and be stuck in a local minimum, which is not necessarily the global minimum. Adding constraints to the model and train it with better/accurate data is one way to avoid this. Certain parameters in real networks, such as network measurements or equipment datasheets, have a range of error/inaccuracy (e.g. 5-10 percent) values, and these values can be used to add such constraints. In our work, we assumed such constraints on the GN model parameters such as, attenuation coefficient from 0.18-0.22 $dB.\ km^{-1}$, dispersion coefficient from 16.7-17.4 $ps.\ (nm.\ km)^{-1}$, and nonlinear coefficient from 1.28-1.42 $W.\ km^{-1}$.

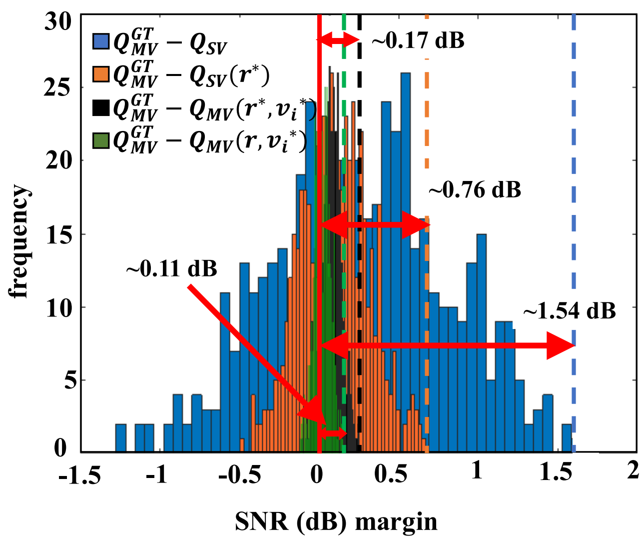


FIGURE 8. Penalty distribution for 500 connections, indicating min. required design margin to accommodate TP performance variations.

As discussed, we randomly assigned a TP, chosen uniformly among the pool of M vendors to each connection. We assumed a 0 dBm value of TP launch power independent of TP vendor category for fair comparison. Taking as reference the SNR obtained by using Q_{SV} from Eq. (3) with default parameters $\boldsymbol{r}$, Fig. 8 depicts the estimation error for 500 connections ($Q_{MV}^{GT} - Q_{SV}$), which pertains to the lack

TABLE 2. Training with only r^*.

Q_{SV} model fitted parameters	fiber coefficients ($\boldsymbol{r}^*$) and bias			
	attenuation $(dB/\,km)$	Non-linear $(W/\,km)$	dispersion $(ps/(nm.km))$	bias (dB)
Same for all M=4 TPs	0.19	1.39	16.94	-1.85

TABLE 3. Training with both r^* and v_i^*.

$Q_{MV}{}^*$ model fitted parameters	fiber coefficients and bias $(\boldsymbol{r_i}^*)$				TP performance factors $(\boldsymbol{v_i}^*)$		
	att. (dB/km)	Non-linear (W/km)	disp. $(ps/(nm.km))$	bias (dB)	$\alpha_i{}^*$	$\gamma_i{}^*$	$\delta_i{}^*$
TP1	0.19	1.34	17.28	-2.1	0.78 (3.7)	0.78 (1.1)	0.83 (2.4)
TP2					0.81 (1.2)	0.92 (4.2)	0.76 (4.6)
TP3					0.92 (3.8)	0.82 (4.1)	0.88 (3.3)
TP4					0.84 (2.4)	0.81 (3.5)	0.91 (3.2)

of knowledge about the network and the TPs performance factors. The penalties were distributed in positive and negative sides depending upon the TP performance factors and were ∼1.5 dB in total. Positive/negative penalties result in upper/lower bounds for the design margin, which we call as *high*/*low* margin. In standards ∼2 dB of design margin is typically used to accommodate all uncertainties in single vendor Qtool (Q_{SV}) [10]. Fig. 8 allows verifying that ∼1.5 dB of QoT tool design margin would be required to accommodate these TP dependent penalties. This indicates that an additional 1-2 dB of margin would be required for multivendor networks [11]. The remaining part of the design margin would cover the other uncertain effects such as EDFA gain ripple, fast time varying polarization effects etc. To improve the estimation accuracy, the very first step is to train the Qtool to obtain $(\boldsymbol{r}, \boldsymbol{v})^*$ parameters vector. We relied on LM algorithm to fit $(\boldsymbol{r}, \boldsymbol{v})^*$ at the load of 500 established connections (new/unestablished connections = 50) as presented in Section IV.A.

The parameter vector $\boldsymbol{r}^*$ plays a crucial role in attaining good estimation accuracy. In our previous work [17], [28], we trained the main parameters of GN model based PLM, such as fiber attenuation, nonlinear and dispersion coefficient ($\boldsymbol{r}$) along with a bias (TP independent). In the following it is shown and discussed how much accuracy can be attained by proper selection of these parameters ($\boldsymbol{r}$ and/or $\boldsymbol{v}$) in case of a partial disaggregated network scenario where additional uncertainties are present.

A. CASE 1 (REFERENCE) - TRAINING GN PARAMETERS VECTOR r^*

This case is used as a reference case in our work and is inspired from [17] and [28]. In this case, only the fiber coefficients and an additional bias (TP independent) are trained with the LM algorithm to minimize the nonlinear least square fitting error between $Q_{SV}(\boldsymbol{r}^*)$ and the ground truth Q_{MV}^{GT}. As so there is no vector $\boldsymbol{v}$ for training. Table. 2 shows the trained (green) parameters vector $\boldsymbol{r}^*$ (along with the bias) obtained at load of 500 connections.

B. CASE 2–TRAINING GN PARAMETERS VECTORS r^* AND TPs PARAMETERS VECTOR v^*

This case corresponds to our proposed solution. We propose to additionally train TP parameters vector $\boldsymbol{v}$ along with the GN model parameters vector $\boldsymbol{r}$, which was trained in the previous reference case. As discussed there are two options, to train either i) jointly for all TPs or ii) individually. In these simulations, we examined the fitting of individual training. In such case, we obtain a different fiber coefficients and bias vector ($\boldsymbol{r}_i^*$) for each TP i and then average those to obtain the $\boldsymbol{r}^*$ vector. We then keep this $\boldsymbol{r}^*$ vector constant and fit only $\boldsymbol{v}_i^*$ in a second round, to focus on the effect of TP vendor dependent performance factors.

Table. 3 shows all trained (green) parameters obtained during the online training phase. From the table one may observe that the obtained values of α_i^*, γ_i^* and δ_i^* do not exactly match to the ground truth parameters of Q_{MV}^{GT}, as shown in Table. 1. This happened because in addition to the fitting of those we need to fit the GN model coefficients ($\boldsymbol{r}^*$ vector and bias) which complicates the fitting. It should be noted that the results mentioned in the Table. 3 (and also in other tables) represent the average value of trained $\boldsymbol{r}^*$ and $\boldsymbol{v}_i^*$ over 100 iterations at a load of 500 connections. Within each iteration, uniformly distributed source-destination node pairs and TPs from $M = 4$ vendors are assigned to verify the working accuracy of the trained Qtools on different training (500 × 100) and testing sets (50 × 100). We also varied the launch powers of established connections and measured the SNR values, similar to the online training approach, to further enrich the training dataset. It is worth noting that as we varied the launch powers, we always double-checked that the criteria of Eq. (8) were met, ensuring that none of the connections went down.

The deviation percentage is a metric to measure the variation between the ground truth parameters (Table. 2) and the estimated ones, as shown in Table. 3. To be more specific, the deviation percentage was defined as the difference between the real/measured value and the estimated one and normalized over the real value ((*estimated* − *real* / *real*).100%). The deviation percentage (%) value for each estimated parameter of the $\boldsymbol{v}_i^*$ vector is indicated in brackets (red color) in Table. 3. We observed over the 100 runs, the maximum percent deviation of ∼5% for estimated $\boldsymbol{v}_i^*$ vector.

C. CASE 3–TRAINING TPs PARAMETERS VECTORS v^*

This case considers perfect knowledge of the fiber coefficients (that is $\boldsymbol{r}$ vector) and the goal is to trace optimum parameter vector $\boldsymbol{v}$ for each TP vendor. In a sense this is an unrealistic case (for online optimization) and is mainly used as an upper performance bound. To perform this training, we set to the fitted Qtool, the $\boldsymbol{r}^*$ parameters that were used in the ground truth Q_{MV}^{GT}, which were those shown in Table. 1. Then we trained individually for each TP and obtained the fitted parameters α_i^*, γ_i^* and δ_i^* which are shown in Table. 4. As shown the fitted $\boldsymbol{v}_i^*$ (green) parameters vector is almost similar to the one used in ground truth (Table. 1) and results in most accurate SNR estimation compare to the rest of the cases.

TABLE 4. Training with only v_i^*.

$Q_{MV}{}^*$ model fitted parameters	fiber coefficients and bias (r^*)				TP performance factors ($v_i{}^*$)		
	att. (dB/km)	Non-linear (W/km)	disp. (ps/(nm.km))	bias (dB)	$\alpha_i{}^*$	$\gamma_i{}^*$	$\delta_i{}^*$
TP1	0.21	1.36	17.19	-2.6	0.80 (1.2)	0.79 (1.3)	0.82 (3.5)
TP2					0.84 (2.4)	0.95 (1.0)	0.74 (2.8)
TP3					0.94 (2.1)	0.88 (2.3)	0.92 (1.1)
TP4					0.81 (1.2)	0.84 (0.4)	0.91 (3.2)

Note that the values of the parameters vector $\boldsymbol{v}_i^*$ in Table. 4 represent the average value over 100 iterations. The deviation percentage (%) value for each estimated parameter of the $\boldsymbol{v}_i^*$ vector is indicated in brackets (red color) in Table. 4. We observed that the maximum percent deviation for estimated $\boldsymbol{v}_i^*$ vector was ~3.5% (during 100 runs). However, the major limitation of this approach is the availability of precise information of other physical layer parameters (GN parameters vector $\boldsymbol{r}$ in this case), which make this case not feasible in practical/real network scenario.

D. COMPARISON OF STUDIED SOLUTIONS–MARGIN REDUCTION

We also calculated the mean square fitting error, MSE $(dB)^2$ on the SNR values for all the above discussed three cases with respect to the ground truth Q_{MV}^{GT}. As expected, we noticed the best performance for Case 3, with a minimum value of MSE $= 8.82 \times 10^{-04}$. For the Case 1 and Case 2, the MSE was found to be 1.8×10^{-02} and 3.8×10^{-03}, respectively. Note that this MSE is calculated at online training phase (500 established connections) and as so it represents how accurately the model fits the training dataset (averaged over 100 iterations). We then used the testing dataset (50 new connections) to obtain the testing error (again averaged over 100 iterations). Apart from the MSE, the maximum and minimum error are quite important in optical networks and QoT estimation, since they define the required margins.

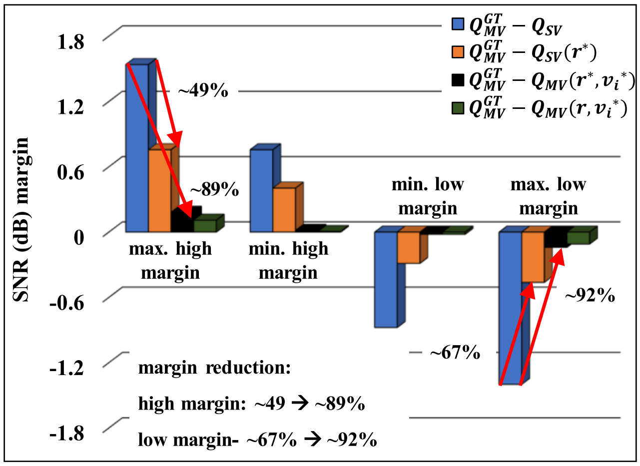


FIGURE 9. Minimum-maximum overestimation and underestimation error range for both high and low margin sides: for the standard-untrained Q_{SV}, for the trained $Q_{SV}(r^*)$, for the proposed $Q_{MV}(r^*, v_i^*)$, and the Q_{MV} with perfect physical layer knowledge (r, v_i^*) with respect to the ground truth Q_{MV}^{GT}, on testing dataset at a load of 500 connections.

Fig. 9 shows the maximum and minimum overestimation (positive) and underestimation (negative) SNR error in dB over the 100 random traffic simulations. It also shows the improvements relative to the traditional untrained Q_{SV}. Note that we define as maximum/minimum high margin the maximum/minimum overestimation and as maximum/minimum low margin the maximum/minimum underestimation error. Note that these maximum and minimum error values are calculated/chosen over 100 iterations. These maximum overestimation/underestimation error values would be the corresponding high/low margins for the Qtool implemented with training parameters vector of each case presented in the previous subsections.

The high margin for the case that we only trained the fiber coefficients and an additional bias (Case 1- reference) was found to be ~0.76 dB, yielding a ~0.78 dB margin reduction with respect to an untrained single vendor Qtool, *i.e.*, Q_{SV}. The low margin was reduced to ~0.46 dB ultimately resulting in margin reduction of ~0.94 dB. For Case 2, the proposed solution, where three additional parameters α_i^*, γ_i^* and δ_i^* (per TP vendor) are also trained, resulted in ~1.3 dB of both high and low margin savings as shown in Fig. 9. Further improvements are obtained with Case 3, but as we already discussed, such case is not very realistic, since it assumes that precise information of fiber coefficients are known with high accuracy. In total, with Case 2 (proposed solution), we achieved additional ~40% and ~25% more margin reduction compared to Case 1 for maximum high and maximum low margin respectively. Similar savings are also indicated in the Fig. 8 for better understanding of readers. Note that when we only used the initial launch power, we saved ~64-70% of the margin (both high and low), which was improved to additional ~20-25% by probing with different launch powers (discussed in Section IV) as shown in Fig. 8 and Fig. 9.

E. ROLE OF PROPOSED QTOOL IN TPs ASSIGNMENT

Besides Qtool training and margin reduction for unestablished connections, we also examined how we can use such information to decide the appropriateness of the available TPs from the M vendors for unestablished connections. For the SNR estimation, we used trained Qtools $Q^{*}_{MV,i}$ for the particular TP vendors (Table. 3). Out of those TPs, the one with the highest SNR was chosen over the others.

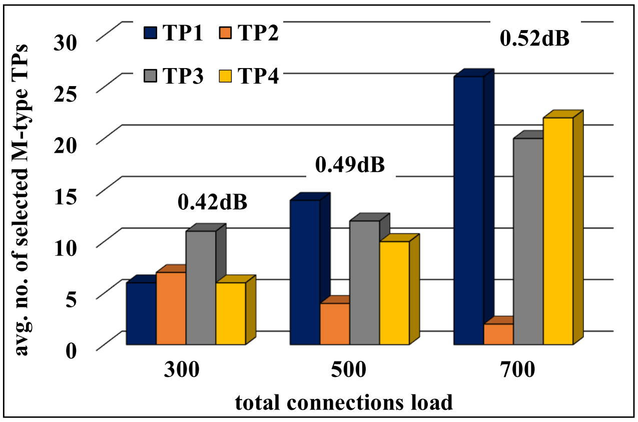


FIGURE 10. Additional SNR (dB) improvements by proper selection of TPs from $M = 4$ vendors at different traffic loads.

Fig. 10 represents the number of newly established connections (10% of the load) that used each available TP from $M = 4$ vendors at different loads. Note that the TPs are uniformly assigned for the total connection requests (established and unestablished). Hence, our interest here is to showcase the importance/role of vendor dependent TPs performance factors in TPs assignment of the new connection requests. For this, we uniformly generated and also averaged the results 100 times at each load.

At low loads, TPs from vendor 2 and vendor 3 were more utilized as the nonlinear noise contribution is low and thus TPs with better α^{*} parameter are selected, as indicated in Fig. 10. At high loads, the amount of TPs assignment form vendor 2 was substantially decreased compared to the low loads. This is because TPs from vendor 2 has the worst γ^{*} parameter, resulting in higher nonlinear noise/penalty, which plays a more important role at higher loads. Also, the utilization of TPs from vendor 3 and vendor 4 increased from medium (500) to high load (700) compared to vendor 2 for the considered network topology. It is because the links are highly occupied and the nonlinear noise becomes prominent. Fig. 10 also shows the additional maximum SNR (dB) value (averaged over 100 iterations) that was achieved by proper selection of TPs for the unestablished connection requests. We observed an additional SNR improvement of 0.42-0.52 dB by proper selection of TPs from $M = 4$ different vendors at different traffic loads on the considered network topology.

VI. CONCLUSION

We proposed PLM extensions that capture the performance variations of multi-vendor TPs. We also devised a monitoring and ML based training scheme (based on non-linear fitting) to estimate the TP vendor dependent performance factors for accurate Qtool modeling in partial disaggregated optical networks. The trained multi-vendor Qtool is then used for estimating the SNR values of new or unestablished connection requests, improving the estimation accuracy and thus reducing the required margin. With the proposed approach, we accomplished a design margin reduction from ∼1.54 dB to ∼0.18 dB for new connection requests with respect to a standard Qtool. We also showed additional SNR improvements of upto ∼0.5 dB that can be achieved by proper selection of TPs from different vendors on top of the previously reduced design margin. This indicates that we can reduce the overprovisioning of the emerging disaggregated optical networks by combining the proposed Qtool with the resource allocation algorithms to achieve such savings.

ACKNOWLEDGMENT

This work is a part of the Future Optical Networks for Innovation, Research and Experimentation (ONFIRE) Project (Project ID-765275).

REFERENCES

[1] V. Lopez and L. Velasco, Eds., "Elastic optical networks," in *Architectures, Technologies, and Control*. Berlin, Germany: Springer, 2016.

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

[3] P. Lu, L. Zhang, X. Liu, J. Yao, and Z. Zhu, "Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks," *IEEE Netw.*, vol. 29, no. 5, pp. 36–42, Sep./Oct. 2015.

[4] Cisco. *Cisco Visual Networking Index: Forecast and Methodology 2016–2021*. [Online]. Available: https://twiki.cern.ch/twiki/pub/HEPIX/TechwatchNetwork/HtwNetworkDocuments/white-paper-c11-741490.pdf

[5] M. Filer, H. Chaouch, and X. Wu, "Toward transport ecosystem interoperability enabled by vendor-diverse coherent optical sources over an open line system," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A216–A224, Feb. 2018.

[6] M. Filer, M. Cantono, A. Ferrari, G. Grammel, G. Galimberti, and V. Curri, "Multi-vendor experimental validation of an open source QoT estimator for optical networks," *J. Lightw. Technol.*, vol. 36, no. 15, pp. 3073–3082, Aug. 1, 2018.

[7] E. Riccardi, P. Gunning, O. G. de Dios, M. Quagliotti, V. Lãpez, and A. Lord, "An operator view on the introduction of white boxes into optical networks," *J. Lightw. Technol.*, vol. 36, no. 15, pp. 3062–3072, Aug. 1, 2018.

[8] J. Santos, N. Costa, and J. Pedro, "On the impact of deploying optical transport networks using disaggregated line systems," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 1, pp. A60–A68, Jan. 2018.

[9] Y. R. Zhou, K. Smith, J. Weatherhead, P. Weir, A. Lord, J. Chen, W. Pan, D. Tanasoiu, and S. Wu, "Demonstration of a novel technique for non-intrusive in-band OSNR derivation using flexible rate optical transponders over a live 727 km flexible grid optical link," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4399–4405, Oct. 15, 2017.

[10] Y. Pointurier, "Design of low-margin optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. A9–A17, Jan. 2017.

[11] M. P. Belanger, M. O'Sullivan, and P. Littlewood, "Margin requirement of disaggregating the DWDM transport system and its consequence on application economics," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3.

[12] L. Nadal, J. M. Fabrega, M. S. Moreolo, F. J. Vílchez, R. Casellas, R. Muñoz, R. Vilalta, and R. Martínez, "Programmable disaggregated multi-dimensional S-BVT as an enabler for high capacity optical metro networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 13, no. 6, pp. C31–C40, Jun. 2021.

[13] "Evolving the awareness of optical networks," White Paper, Infinera.

[14] A. Mahajan, K. Christodoulopoulos, R. Martinez, S. Spadaro, and R. Munoz, "Modelling multi-vendor transponders performance and optimizing launch power," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Dec. 2020, pp. 1–4.

[15] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B60–B71, Apr. 2021.

[16] K. Christodoulopoulos, C. Delezoide, N. Sambo, A. Kretsis, I. Sartzetakis, A. Sgambelluri, N. Argyris, G. Kanakis, P. Giardina, G. Bernini, D. Roccato, A. Percelsi, R. Morro, H. Avramopoulos, P. Castoldi, P. Layec, and S. Bigo, "Toward efficient, reliable, and autonomous optical networks: The ORCHESTRA solution [Invited]," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 11, no. 9, pp. C10–C24, Sep. 2019.

[17] I. Sartzetakis, K. K. Christodoulopoulos, and E. Manos Varvarigos, "Accurate quality of transmission estimation with machine learning," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 11, no. 3, pp. 140–150, Mar. 2019.

[18] E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A298–A306, Feb. 2018.

[19] D. Rafique and L. Velasco, "Machine learning for network automation: Overview, architecture, and applications [invited tutorial]," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, vol. 10, no. 10, pp. D126–D143, Oct. 2018.

[20] P. Soumplis, K. Christodoulopoulos, M. Quagliotti, A. Pagano, and E. Varvarigos, "Network planning with actual margins," *J. Lightw. Technol.*, vol. 35, no. 23, pp. 5105–5120, Dec. 1, 2017.

[21] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 15, 2014.

[22] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25685–25699, Oct. 2013.

[23] P. Johannisson and E. Agrell, "Modeling of nonlinear signal distortion in fiber-optic networks," *J. Lightw. Technol.*, vol. 32, no. 23, pp. 4544–4552, Dec. 1, 2014.

[24] A. Mahajan, K. Christodoulopoulos, R. Martinez, S. Spadaro, and R. Munoz, "Modeling EDFA gain ripple and filter penalties with machine learning for accurate QoT estimation," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2616–2629, May 1, 2020.

[25] A. Mahajan, K. Christodoulopoulos, R. Martinez, S. Spadaro, and R. Munoz, "Improving QoT estimation accuracy with DGE monitoring using machine learning," in *Proc. Int. Conf. Opt. Netw. Design Modeling (ONDM)*, May 2020, pp. 1–6.

[26] C. Delezoide, P. Layec, and S. Bigo, "Automated alignment between channel and filter cascade," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Mar. 2019, pp. 1–3.

[27] L. Barletta, A. Giusti, C. Rottondi, and M. Tornatore, "QoT estimation for unestablished lighpaths using machine learning," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Mar. 2017, pp. 1–3.

[28] A. Mahajan, K. Christodoulopoulos, R. Martínez, R. Muñoz, and S. Spadaro, "Quality of transmission estimator retraining for dynamic optimization in optical networks," *J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B45–B59, Apr. 2021.

[29] *VPIPhotonics*. [Online]. Available: https://www.vpiphotonics.com/index.php

[30] F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, and S. Bigo, "On nonlinear distortions of highly dispersive optical coherent systems," *Opt. Exp.*, vol. 20, no. 2, pp. 1022–1032, Jan. 2012.

[31] I. Roberts, J. M. Kahn, and D. Boertjes, "Convex channel power optimization in nonlinear WDM systems using Gaussian noise model," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3212–3222, Jul. 1, 2016.

[32] I. Roberts and J. M. Kahn, "Measurement-based optimization of channel powers with non-Gaussian nonlinear interference noise," *J. Lightw. Technol.*, vol. 36, no. 13, pp. 2746–2756, Jul. 1, 2018.

[33] H. P. Gavin, "The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems," Duke Univ., Durham, NC, USA, Notes CE281, 2020. [Online]. Available: http://people.duke.edu/~hpgavin/ce281/lm.pdf

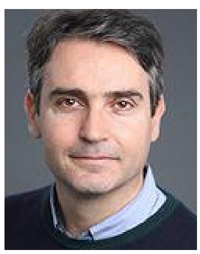
[34] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 15, 2012.

ANKUSH MAHAJAN (Member, IEEE) received the bachelor's degree in electronic and communication engineering from SVITS Indore and the master's degree in electronic and communication engineering from the Indian Institute of Technology (IIT-ISM), Dhanbad, India. He is currently pursuing the Ph.D. degree with the Department of Signal Theory and Communications, UPC, Barcelona, Spain, under the European Union funded H2020-MSCA-ITN Fellowship (ONFIRE Project). He is also currently working as a Research Assistant with the Optical Networks & Systems Department, Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona. His research interests include machine learning assisted QoT estimation, planning, and optimization of optical networks.

KONSTANTINOS (KOSTAS) CHRISTODOULOPOULOS has been working as a Researcher at Nokia Bell Labs, Stuttgart, Germany, since 2018. Before that, he worked as a Senior Researcher at the ECE-NTUA, Computer Technology Institute, Greece; the Trinity College Dublin, Ireland; and IBM Ireland, and as an Adjunct Assistant Professor at the CEID-UoP, Greece. He has coauthored over 150 peer-reviewed journals and conference papers. His research interests include optimization, algorithms and protocols for optical telecom, and datacom networks.

RICARDO MARTÍNEZ (Senior Member, IEEE) is currently working as a Senior Researcher with the Optical Networks & Systems Department, CTTC, Barcelona, Spain. He has participated in over 40 Research and Development Projects funded by the European Commission's Framework Programmes (H2020, FP7, FP6, FP5, CELTIC, EUREKA, and MSCA ITN). He has authored and coauthored over 200 peer-reviewed journals and conference papers. His research interests include orchestration, (GMPLS/PCE and software defined networking—SDN) control and service management architectures, protocols and traffic engineering mechanisms/solutions embracing heterogeneous transport technologies (i.e., packet and optical) within diverse network segments, such as edge, metro, and core interconnecting distributed NFV infrastructures.

RAUL MUÑOZ (Senior Member, IEEE) is currently working as the Research Director, the Head of the Optical Networks and Systems Department, and the Manager of the Communications Networks Division at CTTC, Barcelona, Spain. He has participated in over 50 Research and Development Projects funded by the European Commission's Framework Programmes (H2020, FP7, FP6, and FP5), and the Spanish National Research Programmes, as well as industrial contracts. He has published over 85 JCR journal articles, 275 international conference papers, and three patents. His research interests include control and service management architectures (SDN, NFV, and MEC) for future optical networking and communications.



SALVATORE SPADARO (Senior Member, IEEE) is currently a Full Professor with the Optical Communications Group, Signal Theory and Communications Department, UPC, Barcelona, Spain. He has contributed to more than 250 scientific publications in journals, magazines, and international conferences. His research interests include optical networks solutions/technologies in support of 5G and beyond, AI-based QoS/QoE-aware network control and management solutions, optical data centers/disaggregated data centers, AI- assisted QoT in optical networks, and SDM-based optical networks design.