

# Learning from the Optical Spectrum: Failure Detection and Identification [Invited]

Behnam Shariati, Marc Ruiz, Jaume Comellas, and Luis Velasco

**Abstract**—The availability of coarse-resolution cost-effective Optical Spectrum Analyzers (OSA) allows its widespread deployment in operators' networks. In this paper, we explore several machine learning approaches for soft-failure detection, identification and localization that take advantage of OSAs. In particular, we present three different solutions for the two most common filter-related soft-failures; filter shift and tight filtering which noticeably deform the expected shape of the optical spectrum. However, filter cascading is a key challenge as it affects the shape of the optical spectrum similarly to tight filtering; the approaches are specifically designed to avoid the misclassification of properly operating signals when normal filter cascading effects are present. The proposed solutions are: *i) multi-classifier* approach, which uses features extracted directly from the optical spectrum, *ii) single-classifier* approach, which uses pre-processed features to compensate for filter cascading, and *iii) residual-based* approach, which uses a residual signal computed from subtracting the signal acquired by OSAs from an expected signal synthetically generated. Extensive numerical results are ultimately presented to compare the performance of the proposed approaches in terms of accuracy and robustness.

**Index Terms**—Soft-Failure Detection and identification, Optical Performance Monitoring, Elastic Optical Networks.

## I. INTRODUCTION

NETWORK performance monitoring [1] is a key enabler for failure identification and localization, which can greatly bring down both, the repair time and operational cost of optical networks [2]. Many research efforts have been dedicated to develop failure localization techniques for *hard failures*, i.e., unexpected events that suddenly interrupt the established connections (see e.g., [3]). Nonetheless, although some works can be found in the literature focused on the identification and localization of *soft failures*, i.e., events that progressively degrade the quality of transmission (QoT) (see e.g., [4]), this topic remains rather unexplored. Owing to the fact that soft failures might eventually evolve to hard failures, it is of paramount importance not only to detect them a priori before connections disruption, but also to localize their cause in order to take proper action, e.g., finding a restoration path for the affected connections avoiding the failed element [5].

Such performance monitoring is enabled by the ability of optical components to take measurements. In fact, one of the key features to be exploited in next generation optical networks is the availability of monitoring data that now can be used by data analytics applications, especially those based on

machine learning (ML) [6]. This paves the way to future autonomic networking [7], including self-protection and self-healing.

Aiming at detecting traffic anomalies in packet networks, the authors in [8] proposed bringing data analytics toward the network nodes to reduce the amount of monitoring data to be conveyed to the control and management plane, while improving detection times. Following such idea, the authors in [9] and [10] proposed a distributed Monitoring and Data Analytics (MDA) framework that includes MDA agents running close to the observation points in the network nodes, as well as a centralized MDA controller running in the control and management plane besides the Software Defined Networking (SDN) controller. Such MDA framework is the base to build autonomic optical networks, especially in the case of utilizing white boxes which might include specific optical monitoring devices [11].

Recently, the authors in [12] proposed several solutions to monitor the performance of lightpaths at the transponders side to verify their proper operation, as well as to detect bit-error-rate (BER) degradations thus, anticipating connection disruptions. The authors studied several soft failure causes affecting signal QoT, such as *laser drift*, *Filter Shift(FS)*, and *Filter Tightening(FT)*, and proposed algorithms to detect and identify the most probable failure. Some of these failures happen in the optical switching intermediate nodes, so monitoring the signal solely at the egress node (or even ingress) does not allow their localization. Hence, monitoring techniques to analyze and evaluate QoT in-line are required.

As the abovementioned failures noticeably affect the optical spectrum of the lightpaths, Optical Spectrum Analyzers (OSA) can be used to monitor the spectrum along the transmission line aiming at detecting and localizing that type of failures. Practically speaking, the realization of such solutions become possible with the emergence of a new generation of compact cost-effective OSAs with sub-GHz resolution in the form of optical components [13] allowing real-time monitoring of the optical spectrum of the lightpaths and their corresponding optical signal to noise ratio (OSNR).

Considering the optical spectrum of a lightpath, when a signal is properly configured, its central frequency should be around the center of the assigned frequency slot to avoid filtering effects, and it should be symmetrical with respect to its central frequency. The authors in [14], presented several descriptive features to characterize the optical spectrum of a lightpath. Such features were used to train a set of ML algorithms to detect and identify failures. The most common

Manuscript received June 5th, 2018.

Behnam Shariati, Marc Ruiz, Jaume Comellas, and Luis Velasco (lvelasco@ac.upc.edu), are with the Optical Communications Group (GCO) at Universitat Politècnica de Catalunya (UPC), Barcelona, Spain.

filtering related failures are FS and FT; the optical spectrum becomes asymmetrical in the case of FS, and its edges get noticeably rounded in the case of FT. These irregularities allow distinguishing optical spectra suffering from such failures from the properly configured ones.

However, one of the challenging tasks in [14] was to distinguish whether a signal is affected by a FT failure or it is just degraded due to filter cascading effects, as both phenomena show similar effects. Therefore, aiming at developing sufficiently robust solutions, it is essential to cope with this issue thus, preventing the misclassification of a properly configured signal as a failed one.

In this paper, we extend our previous work in [15] and study three different approaches to detect filter related failures thus, improving the results obtained by the approach followed in [14]. Each of the approaches deals with filter cascading effects differently allowing the development of even more robust solutions when they are combinedly exploited. The approaches are based on a set of classifiers that make predictions; more specifically, the approaches can be categorized in two groups: *i)* *feature-based* approaches that use optical spectrum features for classification. Two different approaches can be considered to deal with filter cascading effects: *i.a)* the *multi-classifier approach*, in which different classifiers are employed for signals experiencing different levels of filter cascading and *i.b)* the *single-classifier approach*, in which the lightpaths' features are pre-processed to compensate for the filter cascading effect allowing the use of a single classifier for lightpaths disregarding the level of filter cascading, and *ii)* the *residual-based approach*, in which the received signal is pre-processed using a theoretically-calculated expected signal allowing the use of a single classifier as in *i.b)*. Ultimately, the optical spectrum analysis can be used by sophisticated algorithms able to localize failures.

The classifiers used in the above-mentioned approaches are developed with two different ML-based algorithms: Decision Trees (DT) and Support Vector Machines (SVM) [6]. The differences between feature-based and residual-based approaches lie in the inputs that those ML algorithms require for predicting failure classes. Assuming the distributed MDA framework discussed above, the proposed approaches for failure detection and identification can be deployed in the MDA agents, close to the devices generating measurements, whereas other algorithms, including the one for failure localization, need be deployed in the MDA controller, so as to provide the global network vision required for their purposes.

The rest of the paper is organized as follows. Section II explores the opportunities that OSAs provide for failure detection, identification and localization and proposes options for feature-based approaches. Section III is exclusively devoted to the residual-based approach that includes new ML-based classifiers. Simulation results showing the benefits and drawbacks of different approaches are presented in Section IV. Finally, Section V concludes the paper.

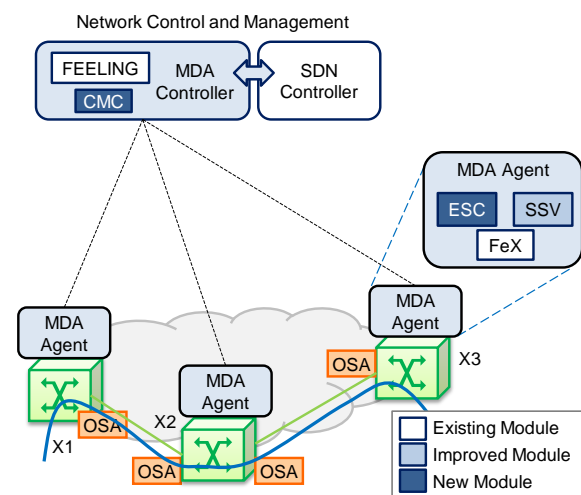


Fig. 1. Evolved architecture for failure identification and localization.

## II. APPLICATION OF OSA FOR SOFT-FAILURE DETECTION, IDENTIFICATION AND LOCALIZATION

Real-time optical spectrum monitoring provides opportunities for soft-failure detection and identification; particularly, those failures significantly deforming the optical spectrum of a lightpath. For precise detection and identification, algorithms need to be capable of classifying a properly operating lightpath from a failed one, which entails that a set of descriptive features should be identified for classification purposes; building upon such features, sophisticated classifiers can be trained to perform the classification task. For the sake of completeness, we review the soft-failure detection, identification, and localization processes utilizing the optical spectrum proposed in [14] and describe how such process is improved in the rest of the paper.

### A. Soft-failure detection, identification, and localization

The failure detection, identification, and localization process involves modules running in the MDA agents and modules running in the MDA controller, as shown in Fig. 1; this follows principle of data analytics distribution proposed in [8]. In the MDA controller, the Failure Cause Localization for optical NetworkG (FEELING) algorithm is primarily responsible for supervising the failure detection and identification modules running in the MDA agents. Ultimately, it performs the failure localization task.

Optical spectrum acquired by OSAs is collected and becomes available in MDA agents, where it is used to feed the Feature Extraction (FeX) module. An example of a 30-GBaud QPSK modulated optical spectrum acquired by an OSA of 312.5 MHz resolution is shown in Fig. 2. In general, QPSK and 16QAM -modulated optical signals present, once filtered, a flat spectral region around the central frequency, sharp edges, and a round region between the edges and the central one. In contrast, the impact of FS and FT on the properly operating lightpaths makes that some or all these features change, as presented in Fig. 3 for the specified *failure magnitude*, where the impact of filter cascading after crossing several wavelength selective switches (WSS) is also depicted.

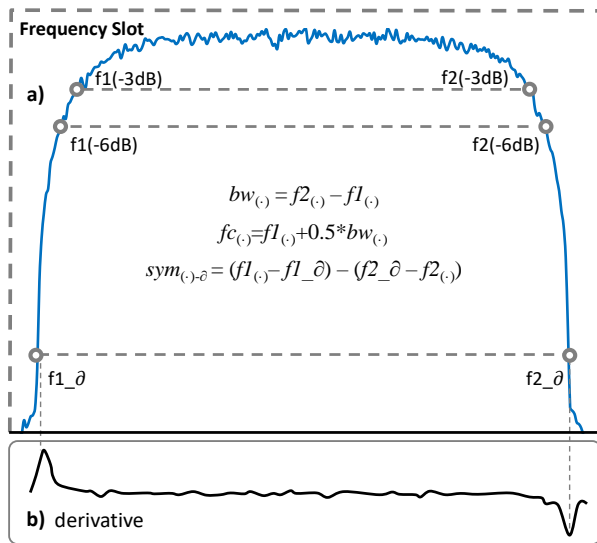


Fig. 2. Relevant points of a QPSK modulated signal

The FeX module processes the acquired optical spectrum for a given lightpath, which consists of an ordered list of frequency-power ( $\langle f, p \rangle$ ) pairs in the allocated frequency slot. After equalizing power, so the maximum power is set to be 0 dBm, the derivative of the power with respect to the frequency is computed; Fig. 2b illustrates the derivative of the example optical signal, where sharp convexity can be observed close to the edges. Next, the FeX module characterizes the mean ( $\mu$ ) and the standard deviation ( $\sigma$ ) of the power around the central frequency ( $fc \pm \Delta f$ ), as well as a set of primary features computed as cut-off points of the signal with the following power levels: *i*) edges of the signal computed using the derivative, denoted as  $\delta$ ; *ii*) a family of power levels computed w.r.t.  $\mu - k\sigma$ , denoted as  $k\sigma$ ; and *iii*) a family of power levels computed with respect to  $\mu - mdB$ , denoted as  $-mdB$ . Each of these power levels generates a couple of cut-off points denoted as  $f1_{(c)}$  and  $f2_{(c)}$ . In addition, the assigned frequency slot is denoted as  $f1_{slot}$ ,  $f2_{slot}$ . Other features, which are computed as linear combinations of the relevant points, focus on characterizing a given optical signal (see embedded equations in Fig. 2a); they include: bandwidth ( $bw$ ), central frequency ( $fc$ ), and symmetry ( $sym$ ) with respect to a reference (frequency slot or derivatives). Some features are more appropriate for filter-related failure detection and identification, such as bandwidth and symmetry, whereas other features, such as the central frequency, are more appropriate for laser drift identification.

When the extracted features from the measured signal are available, a classification module, named Signal Spectrum Verification (SSV), running also in the MDA agents analyzes them to detect a soft-failure. The SSV module was implemented as a multiclass classifier in the form of a DT that produces a diagnosis which consists of: *i*) a predicted class among ‘Normal’, ‘LaserDrift’, and ‘FilterFailure’; and *ii*) a subset of relevant signal points for the predicted class. In the case that a filter failure is detected, another classifier is used to predict whether the failure is due to FS or to FT.

As discussed before, one of the key challenges in the identification of filter related failures is the misclassification of a normal signal that has passed through several filters, i.e., affected by filter cascading, as a signal that has suffered from filter failure. Therefore, to improve failure identification accuracy, the FEELING algorithm must be able to distinguish between actual failures and normal effects arising from filter cascading. In the next section, we propose and study three alternatives to prevent such misclassification. Furthermore, since filter failures have impact on the OSNR, and ultimately in the BER, in [14] we took advantage of OSNR computation in intermediate nodes from the optically acquired signal and compare such measurements against the expected value using analytic formulae. As such method could also be used for failure localization, it can be used to complement and enhance FEELING.

### B. Options for Classification using FeX relevant points

As a result of filter cascading, signal features change in a similar way as when a tight filtering failure takes place; this increases the likelihood of misclassify a properly operating lightpath as a failed one. In the following, we propose two different strategies preventing such misclassification. The strategies, built in the SSV module, are based on processing the features extracted by the FeX module; strategies are summarized in Fig. 4a-b. Selected features for classification are:  $bw_{\delta}$ ,  $bw_{5\sigma}$ ,  $bw_{-3dB}$ ,  $bw_{-6dB}$ ,  $sym_{5\sigma-\delta}$ ,  $sym_{-3dB-\delta}$ , and  $sym_{-6dB-\delta}$ .

#### 1) Multi-Classifer Approach

The most straightforward solution is to use different classifiers as a function of the number of WSSs that a given lightpath has passed through. As shown in Fig. 4a, a set of classifiers are required in every intermediate node and the appropriate one is used when an optical spectrum is acquired.

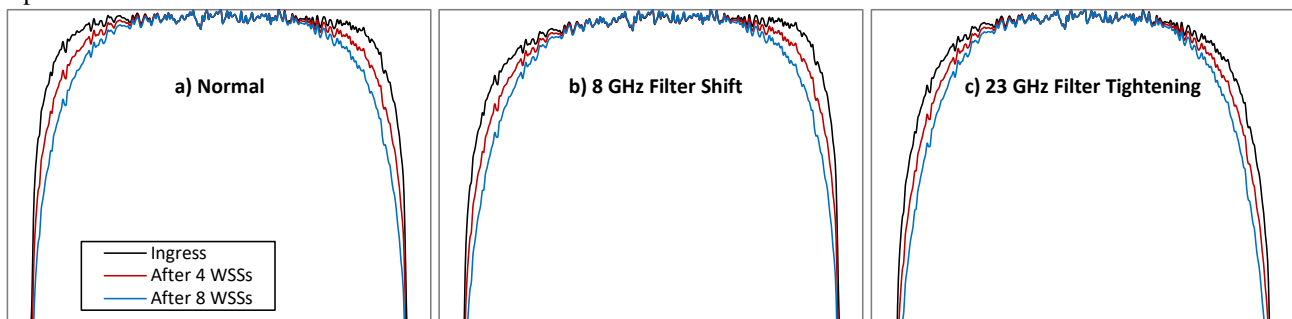


Fig. 3. The shape of the optical spectrum for a normal signal (a), a signal experiencing FS (b), and a signal experiencing FT (c).

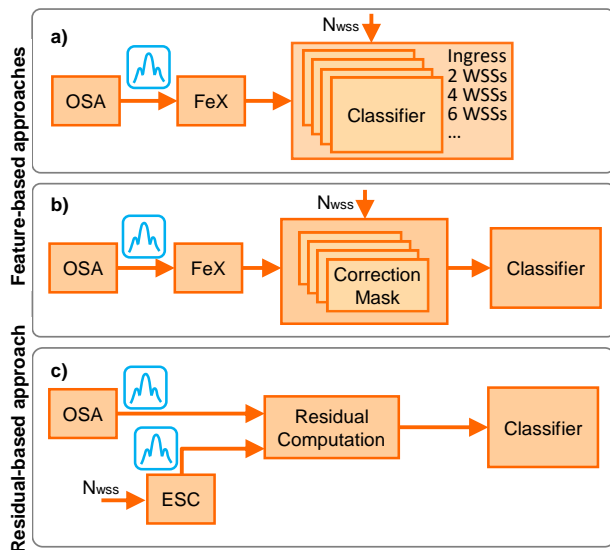


Fig. 4. Approaches to solve the filter cascading problem: a) multi-classifier, b) single-classifier, and c) residual computation.

This approach can be considered as the baseline, as the selected classifier decides based on the features extracted directly from the acquired spectrum and do not need any kind of feature pre-processing. However, a very large dataset of optical spectra with different levels of filter cascading is required for training all the classifiers, which is the main drawback of this approach.

To avoid using multiple classifiers, some pre-processing needs to be done so one single classifier can be used despite the level of filter cascading. The second approach proposes a strategy to pre-process the extracted features.

## 2) Single-Classifier Approach

Filter cascading strongly affect some of the features that a classifier uses for prediction. Therefore, if the alteration of those features due to filter cascading would be compensated, a single classifier could be considered regardless of the number of filters a signal passes through. The features of a signal acquired after passing  $N$  filters can be compensated by adding/subtracting the differences between the values of a properly configured signal at that node w.r.t. those just after the transponder. These differences are stored in a vector called *correction mask*; note that, different levels of filter cascading require different correction masks to be used.

Correction masks can be computed a priori, assuming the effects that the spectrum of a normal signal experiences while passing through different number of filters. It is worth mentioning that the calculation of the correction masks requires just the spectrum of a single properly configured lightpath passing through the desired number of filters, from zero to the maximum allowed cascaded filters in a network; this is in contrast to the previous approach, where the training phase requires that spectral data with different failures and with various magnitudes to be captured after every filter up to the maximum allowed number of filters. The Correction Mask Calculator (CMC) module placed in the MDA controller (see

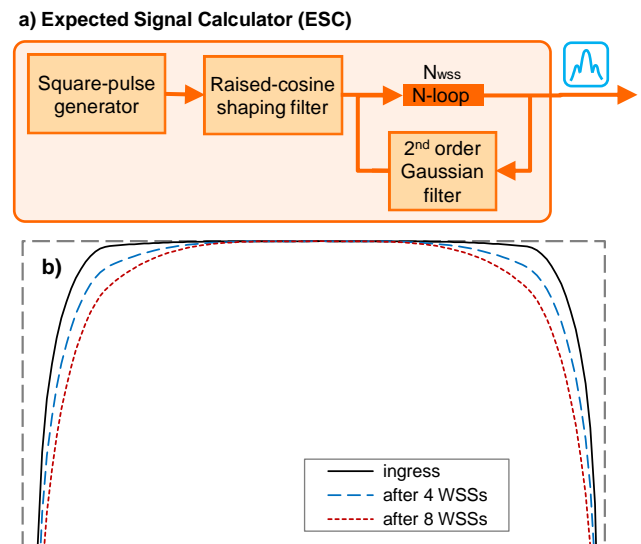


Fig. 5. ESC module (a) and generated signals (b)

Fig. 1) is responsible for generating the correction masks to be sent to the MDA agents. Note that all the correction masks need to be available in the MDA agents, so the proper one can be selected. Following this approach, the classifier can be trained based on the observations of a signal passing through a single filter after the ingress node, where the impact of filter cascading is negligible, making the training phase less data-hungry by far compared to the previous approach.

These two approaches use features extracted from the optical spectrum to decide whether a lightpath experiences a soft-failure; the next section presents, a totally new approach that analyzes the optical spectrum in a very different way.

## III. RESIDUAL-BASED APPROACH FOR OPTICAL SPECTRUM ANALYSIS AND SIGNAL CLASSIFICATION

An alternative approach compared to the ones in the previous section is to pre-process the acquired optical spectrum by comparing it to the one that would be expected after passing the same number of filters than the signal (shown in Fig. 4c). This comparison / computation produces a *residual signal* representing the differential deformation in its shape that might be due to a failure. Note that this approach does not use the FeX module. In order to compute the residual signal, we consider two new modules, as shown in Fig. 4c: *i*) the Expected Signal Calculation (ESC) and *ii*) the residual computation module.

The ESC module generates a theoretically-calculated optical spectrum emulating a properly operating lightpath. The aim of ESC module is to synthetically reproduce an averaged noise-free version of the optical signal. In order to do so, the signal is modelled as an ideal square pulse, with  $bw_{.3dB}$  equal to the baud rate of the optically modulated signal, shaped by a raised-cosine shaping filter with 0.15 roll-off factor [16]. Then, in order to model different levels of filter cascading a 2<sup>nd</sup> order Gaussian filter, emulating a WSS, is used (Fig. 5a). This results in an emulated noise-free spectrum, similar to a noise-free 100G DP-QPSK (or 200G DP-16QSPK) modulated



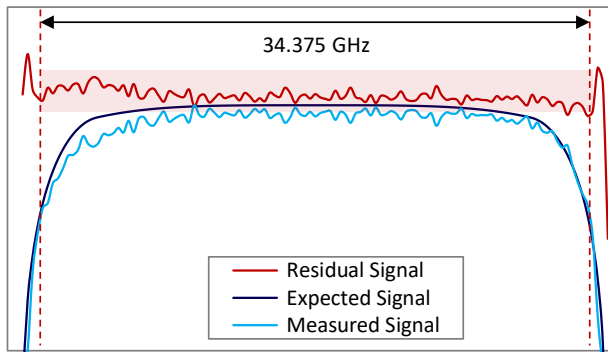


Fig. 6. Residual signal calculation (a) and residual based classification approach (b).

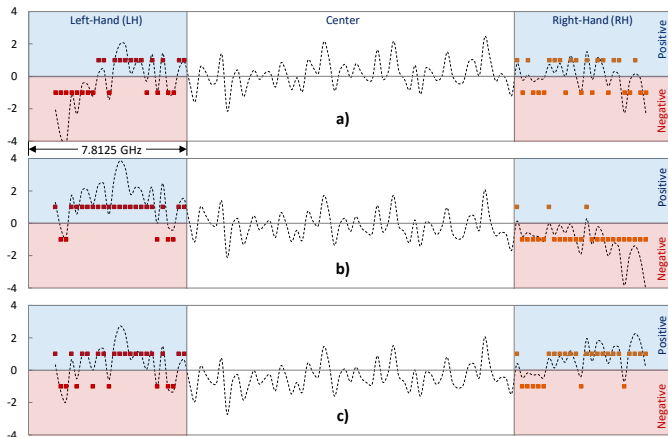


Fig. 7. Example of residual based approach workflow considering a normal signal (a), filter shift (b), and filter tightening (c).

signal (Fig. 5b). Every time a new spectrum is acquired for a given lightpath, the residual computation module subtracts it from the expected signal produced by the ESC module.

Fig. 6 presents an illustrative case of the computed residual signal from the expected and the measured signals. As the residual signal experiences undesired changes at the two ends of the acquisition window of the OSA (see Fig. 6), we just consider a central spectral window of size 34.375 GHz out of 37.5 GHz. Note that in case of a normal signal, the residual value fluctuates around a mean value along the whole signal spectrum range. To analyze the residual signal, we normalize the values, so the mean equals 0; in such case, the most likely situation is to have as many positive as negative values. However, in the event of a filter failure, that similar proportion between positive and negative residuals will be altered. Fig. 7 shows an example of how residuals behave in all considered cases. Out of the whole range of the residuals, the left and the right hand-sides capture the effects of soft-failures and are the operational regions for the analysis presented next. The spectral window of the left (LH) and the right (RH) hand-sides are set to 7.8125 GHz, as they contain the sufficient number of points to capture the effects taking place in the edges. In the normal case (Fig. 7a), the residuals oscillate uniformly between positive and negative values (dots in Fig. 7 are computed from the normalized residuals applying the  $sign()$  function). In the case of filter shift (Fig. 7b), the residuals show a clear distortion toward positive and negative parts in

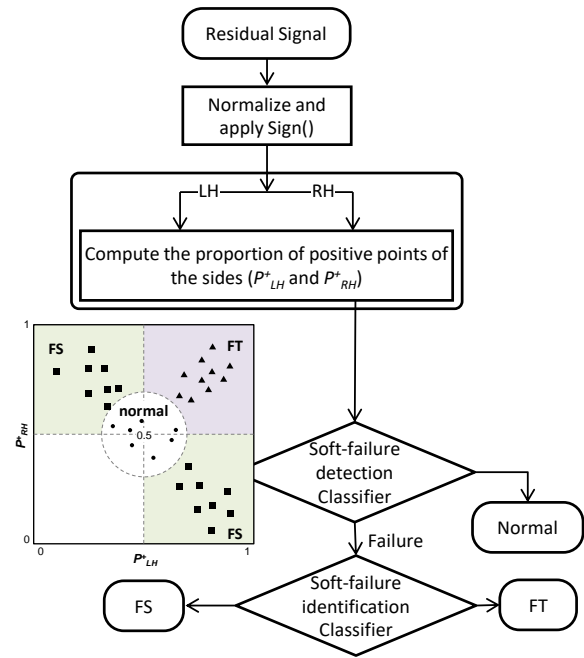


Fig. 8. Residual-based classification procedure

the LH and RH side, respectively. Finally, in the case of filter tightening (Fig. 7c), the residuals move toward the positive part in both the LH and RH sides.

In conclusion, comparing LH and RH sides, one can predict whether the signal is normal (symmetric sides and unbiased distribution) or either it is affected by filter shift (asymmetric sides) or a filter tightening (symmetric and biased distribution).

In light of this analysis, we propose the residual-based procedure presented in Fig. 8, where the residuals are first normalized with respect to their mean value and centered in zero; then, the  $sign()$  function is applied to convert normalized residuals into points of amplitude +1 or -1. After selecting the sides, the proportion of positive points in the LH and the RH sides ( $P_{LH}^+$  and  $P_{RH}^+$ , respectively) is computed. For illustrative purposes, Fig. 8 plots possible observations in the semi-plane  $P_{LH}^+, P_{RH}^+$ . According to the rationale previously presented, normal observations should be kept within an area centered around (0.5,0.5), FS observations should be in the quadrants ( $>0.5, <0.5$ ) or ( $<0.5, >0.5$ ), while FT observations should be in the quadrant ( $>0.5, >0.5$ ), both FS and FT outside the normal area. Therefore, the coordinates of the observations in the semi-plane  $P_{LH}^+, P_{RH}^+$  can be used as features for DT and SVM -based classifiers. In the proposed procedure, two classifiers are trained; the first one for soft-failure detection and the second one for its identification.

#### IV. ILLUSTRATIVE RESULTS

In this section, we numerically compare the performance of different approaches described in the previous sections. Firstly, the description of the transmission set-up modeled in VPI Photonics is described; the set-up is used to generate the optical spectrum database required for training and testing the

proposed algorithms. Next, the two feature-based approaches implemented as DT and SVM classifiers are evaluated thus, revealing their benefits/drawbacks for filter failure identification tasks. Finally, the performance of the three proposed approaches for failure detection and localization are compared.

### A. VPI set-up for data collection

The VPI set-up is shown in Fig. 9a. In the transmitter side (N1), a 30 GBd DP-QPSK signal is generated, passes through 7 intermediate nodes performing optical switching and ends in a coherent receiver (N9) that compensates for the impairments introduced throughout the transmission. Nodes are interconnected by single mode fiber spans; after each span, Erbium-Doped Fiber Amplifiers (EDFA) compensate for the accumulated attenuation of the fiber. The transmitter and receiver are assumed to be installed in optical nodes, which are modeled with two 2<sup>nd</sup> order Gaussian filters emulating optical switching functionality for add/drop and pass-through performed by WSSs [17], [18]; filters bandwidth is set to 37.5 GHz, leaving 7.5 GHz as a guard band.

One OSA per outgoing link, configured with 312.5 MHz resolution, is considered in every node to monitor the optical spectrum. As previously discussed, a correction mask should be considered for the features affected by filter cascading, as features get modified while passing through WSSs. Fig. 9b shows an example of the amount of reduction in the  $bw_{-3dB}$  feature of a lightpath in the set-up and the corresponding correction mask, obtained by fitting a 2<sup>nd</sup> order polynomial.

Aiming at emulating failure scenarios, we modify the characteristics of the 2<sup>nd</sup> WSS of each node (from N1 to N8) in the set-up; its bandwidth and central frequency are modified to model FT and FS failures, respectively. A large dataset of failures was collected by inducing failures of magnitude in the range [1-8] GHz for FS and in the range [1-15] GHz for FT, both with 0.25 GHz step-size, where the magnitude of FT is defined as the difference between the ideal bandwidth of the filter (37.5 GHz) and its actual bandwidth during the failure.

### B. ML-based classification comparison

We compare the performance of the feature-based approaches in terms of its *accuracy*, defined as the number of correctly detected failures over the total failures.

Fig. 10a-b show the accuracy of detecting FS and FT, respectively at node N1 in terms of the magnitude of the failure. Note that in N1 both multi-classifier and single-classifier are the same as no filter mask is required. Every point in Fig. 10a-b aggregates failure scenarios with different magnitude by considering all the observations belonging to a particular failure magnitude and above. As shown, the accuracy of detecting FS larger than 1 GHz is around ~96% when classifiers are based on SVMs, while it hardly approaches 89% when they are based on DTs. On the other hand, the accuracy of SVMs reaches 100% for failures larger than 5 GHz, while this level of accuracy for DTs is achieved for failures larger than 6 GHz. Regarding FT detection, the

best accuracy of the proposed classifiers for low magnitudes (below 6 GHz) is around 80% (achieved for SVMs), which is due to the fact that the shape of the optical spectrum is quite similar to the normal scenario, making it very challenging for the classifier to distinguish. This is in contrast to the case of FS, whose effect is more evident even for low magnitudes due to its asymmetric impact on the optical spectrum. For the magnitudes above 7 GHz, the SVM-based classifier perfectly detects the failure. Note that DT-based classifiers achieve perfect accuracy for magnitudes above 10.5 GHz.

Let us now compare feature-based approaches implemented with DT and SVM -based classifiers for detecting failures in all 8 nodes of the set-up. Recall that multiple classifiers are needed for the first approach and several filter masks are required for the second approach. The results are shown in Fig. 10c-d for FS and FT, respectively, where every point aggregates the results for all the nodes. As observed, SVM-based classifiers significantly outperform DT-based ones in both approaches and failures. As a result, SVM-based classifiers can be selected as the preferred option for feature-based approaches.

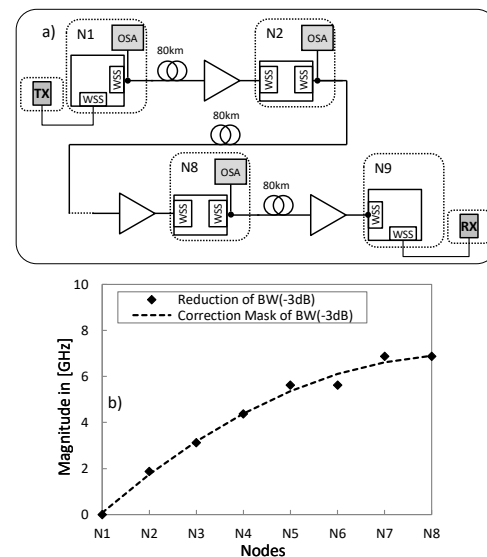


Fig. 9. VPI setup (a) and correction mask of  $bw_{-3dB}$  of the setup (b)

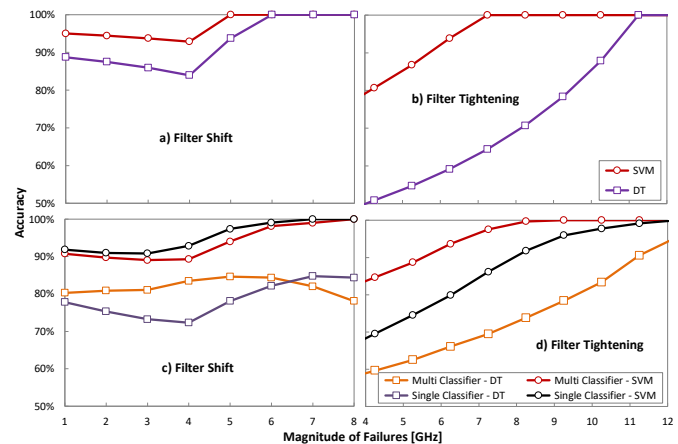


Fig. 10. Accuracy of DT and SVM for FS (a, c) and FT (b, d). (a, b) show the accuracy at N1, while (c, d) show the average over N1-N8.

Comparing the different SVM-based approaches, the single-classifier performs slightly better in the case of FS detection while the performance of the multi-classifier approach is much better than the single-classifier one in the case of FT failure. Therefore, we can conclude that training multiple classifiers with the data collected at nodes experiencing different levels of filter cascading performs better than correcting the features with the purpose of using a single classifier, as the impact of filter cascading is similar to the effect of FT on the shape of optical spectrum.

### C. Comparisons of different approaches of signal analysis

Let us first focus on the residual-based approach and analyze the potential of using  $P^+_{LH}$  and  $P^+_{RH}$  as discriminatory features for training the classifiers. Fig. 11 plots the samples of the dataset used for training the residual-based classifiers at N1. It is clear, in view of Fig. 11, that the selected features allow classifying the different cases easily. Although not shown in detail, we first compared DT and SVMs classifiers and concluded that SVMs perform better also for this approach. Therefore, the results presented next compare the performance of all three approaches using SVM-based classifiers.

Fig. 12a-b show the accuracy of detecting FS and FT solely at node N1. As illustrated, multi-classifier and single-classifier approaches show identical performance as they are essentially the same at node N1, where no filter mask is necessary to be used. However, the residual-based approach outperforms feature-based approaches for the detection of FS and FT. The performance of the residual-based approach is noticeable for detecting soft-failures, as it reaches 100% accuracy for 3 GHz FS magnitude and 5 GHz FT magnitude, 2 GHz smaller than the other two approaches.

Let us now look at the average accuracy of the approaches over all the nodes; the results are shown in Fig. 12c-d for FS and FT failures, respectively. The residual-based approach remains the best solution by far for detecting and identifying

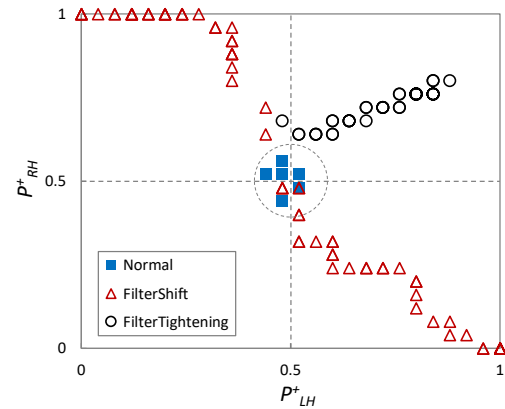


Fig. 11.  $P^+_{LH}$  and  $P^+_{RH}$  features for failure detection and identification

both, FS and FT failures. In contrast, the multi-classifier-based approach shows the worst overall accuracy for FS failures, as the features selected for classification get very close to each other and it becomes very difficult for the support vectors to distinguish them perfectly. Interestingly, the situation is different for FT failures, where the multi-classifier approach performs better than the single-classifier.

In addition to classification accuracy, the *robustness* of the approaches with respect to the number of traversed nodes is of paramount importance for practical implementations. Hence, let us compare the robustness of different approaches in terms of the smallest failure magnitude after which the classification accuracy reaches 100%. Plots in Fig. 13a-b represent such robustness in terms of the location of FS and FT failures, respectively. Note that, points for N1 in Fig. 13a-b correspond to the first points in Fig. 12a-b where 100% accuracy is achieved. As observed, the residual-based approach shows the highest level of robustness compared to the feature-based ones for both FS and FT failures. It can be understood that the residual-based approach is robust regardless of the location of the failure as it perfectly detects and identifies failures with magnitude above the values in Fig. 13.

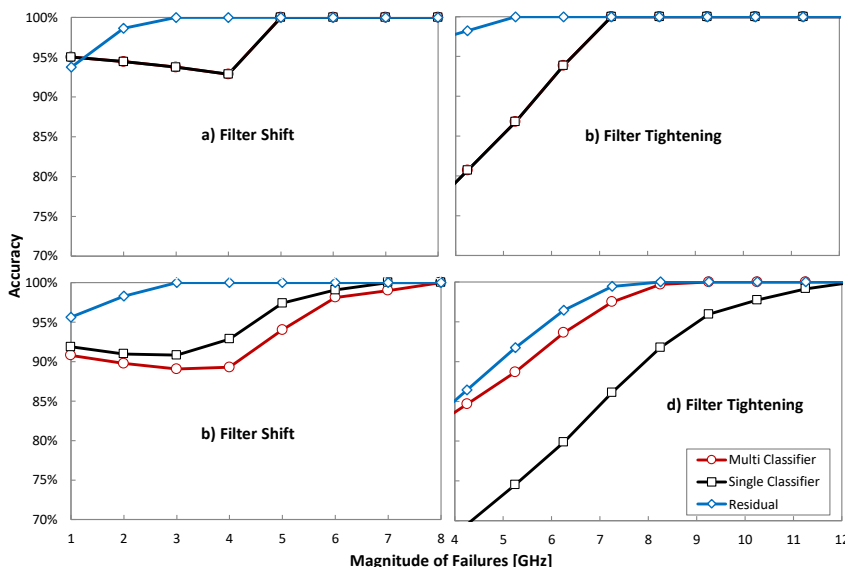


Fig. 12. Accuracy of the different approaches

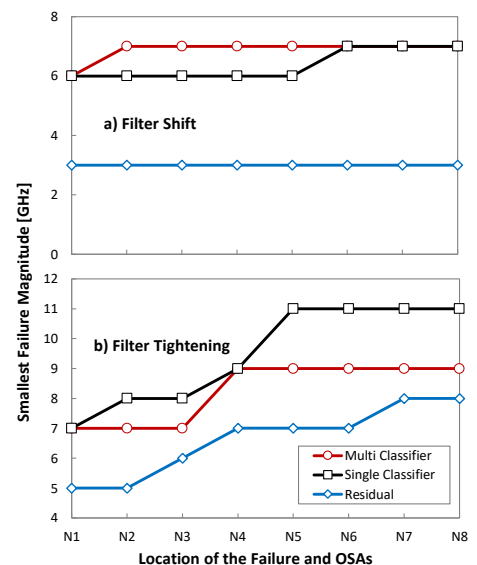


Fig. 13. Robustness of the different approaches

Table I. Key characteristics and results for the considered approaches for failure detection and identification

	Pre-processing	Classification method	Training phase	Number of classifiers	Availability at the node level	Accuracy	Robustness w.r.t # of nodes
<b>Multi Classifier</b>	not required	SVM	requires observations of every level of filter cascading	# of nodes to support	all classifiers	good	good
<b>Single Classifier</b>	pre-processing of the features	SVM	requires observations of just a single level of filter cascading	1	one classifier + correction masks	good	good
<b>Residual Based</b>	pre-processing of the optical spectrum	SVM	requires observations of just a single level of filter cascading	1	one classifier + ESC and residual signal computations modules	very good	very good

### V. CONCLUDING REMARKS

Three different approaches for filter-related soft-failures detection and identification, have been proposed and their performance was compared in terms of accuracy and robustness. The key characteristics are summarized in Table I.

On the one hand, feature-based approaches (i.e., multi-classifier and single-classifier), even though their performance is comparable, have notable differences in their implementation complexity. While the multi-classifier approach requires a huge dataset for the training phase, single-classifier approach requires  $N$  times less data, being  $N$  the maximum number of nodes an optical connection might pass-through. However, this is at the cost of pre-processing optical spectrum features, which requires the calculation of the correction masks. On the other hand, the residual-based approach, which is based on a single classifier strategy, significantly outperforms the feature-based ones and brings down the complexity of training phase compared to multi-classifier approach. However, it requires two additional modules to be available in the MDA agent.

To conclude, it is beneficial to bring data analytics as close as possible to the source of the monitoring data (MDA agents) as the complexity of the proposed modules is low enough to be integrated in programmable units that are expected to be available in the future whitebox-based technologies for switching and transponder optical nodes.

### ACKNOWLEDGMENT

This work was partially supported by the EC through the METRO-HAUL (G.A. n° 761727) project, from the Spanish AEI/FEDER TWINS (TEC2017-90097-R) project and from the Catalan Institution for Research and Advanced Studies (ICREA).

### REFERENCES

- [1] Z. Dong *et al.*, "Optical performance monitoring: a review of current and future technologies," *IEEE/OSA Journal of Lightwave Technology (JLT)*, vol. 34, pp. 525-543, 2015.
- [2] Victor López and Luis Velasco, *Elastic Optical Networks: Architectures, Technologies, and Control*, Springer, 2016.
- [3] C. Mas, I. Tomkos, O. K. Tonguz, "Failure location algorithm for transparent optical networks," *IEEE Journal Sel. Areas Commun.* vol. 23, pp. 1508-1519, 2005.
- [4] C. Delezoide *et al.*, "Pre-emptive Detection and Localization of Failures Towards Marginless Operations of Optical Networks," in *Proc. ICTON 2018*.
- [5] L. Velasco, A. P. Vela, F. Morales, and M. Ruiz, "Designing, Operating and Re-Optimizing Elastic Optical Networks," (Invited Tutorial) *IEEE/OSA Journal of Lightwave Technology (JLT)*, vol. 35, pp. 513-526, 2017.
- [6] D. Rafique and L. Velasco, "Machine Learning for Optical Network Automation: Overview, Architecture and Applications," (Invited Tutorial) *IEEE/OSA Journal of Optical Communications and Networking (JOCN)*, vol. 10, 2018.
- [7] M. Behringer, M. Pritikin, S. Bjarnason, A. Clemm, B. Carpenter, S. Jiang, L. Ciavaglia, "Autonomic Networking: Definitions and Design Goals," *IETF RFC 7575*, 2015.
- [8] A. P. Vela, M. Ruiz, and L. Velasco, "Distributing Data Analytics for Efficient Multiple Traffic Anomalies Detection," *Elsevier Computer Communications*, vol. 107, pp. 1-12, 2017.
- [9] L. Velasco *et al.*, "An Architecture to Support Autonomic Slice Networking [Invited]," *IEEE/OSA Journal of Lightwave Technology (JLT)*, vol. 36, pp. 135-141, 2018.
- [10] Ll. Gifre *et al.*, "Autonomic Disaggregated Multilayer Networking," *IEEE/OSA Journal of Optical Communications and Networking (JOCN)*, vol. 10, pp. 482-492, 2018.
- [11] L. Velasco *et al.*, "Building Autonomic Optical Whitebox-based Networks," *IEEE/OSA Journal of Lightwave Technology (JLT)*, 2018.
- [12] A. P. Vela *et al.*, "BER degradation detection and failure identification in elastic optical networks," *IEEE/OSA Journal of Lightwave Technology (JLT)*, vol. 35, pp. 4595-4604, 2017.
- [13] Flexgrid High Resolution Optical Channel Monitor (OCM) [On-line] [www.finisar.com](http://www.finisar.com), accessed June 2018.
- [14] A. P. Vela *et al.*, "Soft Failure Localization During Commissioning Testing and Lightpath Operation [Invited]," *IEEE/OSA Journal of Optical Communications and Networking (JOCN)*, vol. 10, pp. A27-A36, 2018.
- [15] L. Velasco, B. Shariati, A. P. Vela, M. Ruiz, J. Comellas, "Learning from the optical spectrum: soft-failure identification and localization [Invited]," in *Proc. OFC*, 2018.
- [16] J. G. Proakis, M. Salehi, *Digital Communications*, McGraw-Hill, 2008.
- [17] C. Pulikkaseril *et al.*, "Spectral modeling of channel band shapes in wavelength selective switches," *OSA Optics Express*, vol. 19, pp. 8458-8470, 2011.
- [18] Finisar Whitepaper, "Filter bandwidth definition of the waveshaper S-series programmable optical processor," [On-line] [www.finisar.com](http://www.finisar.com), accessed June 2018.