Hybrid Learning Assisted Impairments Abstraction for Service Planning and Provisioning Over Multi-Domain Optical Networks

RUI WANG[1,*], REZA NEJABATI[1], AND DIMITRA SIMEONIDOU[1]

[1] High Performance Networks Group, Department of Electrical and Electronic Engineering, University of Bristol, BS8 1UB, UK
[*] Corresponding author: rui.wang@bristol.ac.uk

Compiled November 6, 2020

This article proposes and demonstrates a hybrid learning assisted impairments abstraction framework for planning and provisioning intra-domain and inter-domain services in a field-trial multi-domain optical networks testbed. The proposed abstraction strategy consists of a parametric and a non-parametric machine learning technique to allow control plane implementing impairments abstraction with different accessible data or monitoring technologies in the data plane. The hybrid learning assisted abstraction framework aims to abstract the property of segmental links along the lightpath and combine them for end-to-end performance evaluation. By deploying the proposed abstraction framework, network providers or operators are able to exchange the abstracted information for end-to-end impairments abstraction without revealing detailed information within each network. We experimentally demonstrate the proposed solution over a three-network field-trial testbed with real monitored data. The hybrid learning assisted impairments abstraction proves to be an accurate abstraction tool, with an average of 0.33 dB end-to-end signal-to-noise ratio estimation error for services across three networks.

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

1. INTRODUCTION

The emerging new applications such as autonomous vehicles, mobile edge computing, and internet of things pose great new challenges to heterogeneous networks to provide end-to-end connectivity. Service orchestration over multiple network domains can realise flexible and efficient resource management for end-to-end connections [1, 2]. Cross-domain orchestration requires the support from local software-defined networking (SDN) controllers of each network domain to provide network knowledge and information to enable autonomous networks. Software-defined and programmable optical networking [1, 3] have attracted a wide range of attention in recent years as the technology enabler for the realisation of flexible, dynamic and autonomous optical network [4, 5]. However, orchestration over multiple optical networks to implement end-to-end intelligent algorithms require complex analytical models to calculate and manage available network resources and link impairments. An abstraction layer to hide these complexities and to enable network dynamicity in the control plane will pave the way for the implementation of intelligent and complex autonomous networking algorithms.

Most of the research activities for abstraction in optical networks focus on the hardware level abstraction. This includes extension of Openflow [6–8] or deploy open source-based YANG data model [9, 10] and OpenROADM [11–13] to manage and support non-SDN enabled devices in the optical networks. However, these research works focus on control and management layer while ignoring the physical layer characteristics of optical networks, which leads to inefficient network utilisation. As a result, an abstraction layer for monitoring physical layer information and mapping these parameters to a common entity or indicator is necessary and essential. This will allow network providers or operators to implement complex algorithms with simple abstracted information in their own networks. Furthermore, for services across multiple optical network domains, each network domain can deploy such a physical characteristics abstraction layer and exchange the abstracted information to avoid sharing detailed knowledge of their networks for confidential and security reason. Such an abstraction layer should capture the physical layer information, such as gain of erbium-doped fiber amplifier (EDFA), amplified spontaneous emission (ASE) noise, dispersion and fibre nonlinear impairments, extract and map them to a quality of transmission (QoT) metric. Authors in [14] proposed an abstraction strategy over multiple locations based on Gaussian noise (GN) model [15, 16] as the nonlinear impairments evaluation tool. Similarly, resource allocation schemes in [17–19] aiming to reduce network blocking ratio and improve network utilisation rely on theoretical model or GN model for

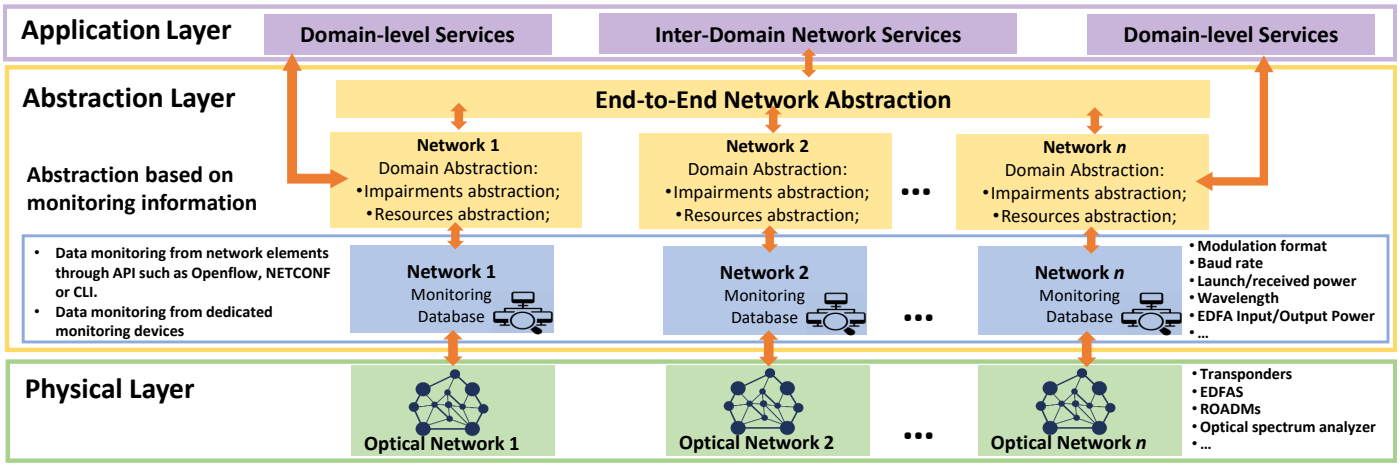

Fig. 1. Architecture of abstraction layer in the optical networks

QoT assessment. However, these models cannot efficiently and accurately capture the dynamicity and uncertainty of optical system like power fluctuation, the ripple of EDFA amplification curve by assuming only certain types of impairments presented in the optical networks. Due to the complex mechanism of impairments and uncertainty of the practical system, it is impossible to derive a closed-form formula to take account of all the factors.

Recently, machine learning (ML) has attracted huge interest in optical networks [20], especially for lightpath QoT prediction to enable such an abstraction layer. In [21], authors present a self-learning network to predict the QoT of lightpaths based on Gaussian process regression (GPR), but it only performs prediction on one link and lacks accuracy for the link with few training data without considering the impact of transponders. Effect of EDFA gain ripple and filtering have been studied in [22] using support vector machine regression for QoT Estimation. The strategy mainly focuses on capturing the effects of optical nodes than fibre links, thus making the end-to-end QoT estimation less effective. Artificial neural network (ANN)-based QoT prediction over field-trial testbed or multiple optical networks are presented in [23–25]. The ANN-based solution performs QoT prediction of one channel in one link in [23], which cannot be extended to end-to-end QoT learning. One-hot encoding of assigned wavelengths has been adopted in [24], this strategy cannot achieve accurate results in case a wavelength never established in the past or lacking historical operational data. Wavelength dependency of a lightpath for QoT prediction is not considered in [25], thus lacking QoT estimation accuracy due to not considering the effect of EDFA gain fluctuation and nonlinear impairments. Transfer learning is adopted in [26] to predict QoT, which reduces the requirement for available amount of training data set. However, it is based on simulation which may lack practice in the real optical network.

In this paper, we first propose an impairments abstraction framework - a simple signal-to-noise ratio (SNR) degradation factor, which network providers/operators can share without revealing details of their networks. Unlike previous research activities focusing on direct lightpath QoT learning, the proposed SNR degradation factor presents how the signal quality degrades within one or several links in the presence of various impairments. The end-to-end QoT of the network services can be formed by aggregating the SNR degradation factors of segmental links within the lightpath. A segmental link can be formed by a fiber link or several fiber links along the selected lightpath. This allows the information shared between network domains but increasing the confidentiality of each network. Based on this concept, we further propose a hybrid learning framework combining deep learning technique and GPR model to learn the abstracted SNR degradation factors of the segmental links. More specifically, the proposed framework uses information of intra-domain services for deep learning as it requires more specific data for learning. In case of absence of training data or services across multiple networks, the proposed framework adopt GPR as it does not require a huge amount of features and training sets like deep learning. This paper aims to investigate the end-to-end performance combining different machine learning techniques for the proposed abstraction scheme. We extend our previous work in [27] by providing a detailed description of the abstraction strategy, the experimental implementation, more scenarios and results analysis from the field-trial experiment. We demonstrate the hybrid learning assisted abstraction strategy over 3 optical networks. The results reveal the proposed strategy achieves high accuracy for SNR prediction with an average of 0.33 dB end-to-end estimation error. In Section 2, the impairments abstraction model is introduced. Section 3 describes the back-to-back characterisation of transponders (TRx). Section 4 presents the ML techniques in the hybrid learning assisted abstraction model. Experimental testbed setup is explained in Section 5, followed by results analysis in Section 6.

2. IMPAIRMENTS ABSTRACTION MODEL

Optical networks suffer from various impairments such as ASE noise, fibre non-linear impairments due to the Kerr effect, and dispersion. These impairments degrade the quality of the signal, which reduces system reach distance, and the overall throughput of the optical networks. Therefore, it is essential to capture these features to allow high layer applications to evaluate the performance of the established services and future services in the optical networks. One the other hand, detailed physical layer information is difficult to provide a direct interpretation of network status to the high layer applications for decision making. Therefore, an abstraction layer hiding detailed knowledge of optical networks while mapping or translating the physical layer information into certain QoT indicator is necessary for the control plane architecture. Such an abstraction layer along with

the QoT indicator can provide a simple and unified abstract entity across links in single or multiple optical network domains. Further, the abstraction mechanism can increase confidentiality of detailed information in the network such as link lengths, node or EDFA location and settings to the network application layer and other optical network domains.

In Fig. 1, we illustrate such a layer to enable the abstraction mechanism in optical networks. The abstraction layer employs a monitoring plane to monitor the network operational information and status from optical network physical infrastructures such as EDFAs, transmitters, receivers, and reconfigurable optical add-drop multiplexers (ROADMs). Besides the data from optical elements, the abstraction layer can also collect information from dedicated monitoring devices deployed in the optical networks like optical spectrum analyzers. For SDN-enabled network elements, the information can be exchanged through southbound application programming interface (API) like Openflow, NETCONF, and OPENROADM. An agent will be needed between physical layer and the monitoring layer for non-SDN supported optical devices. The agent then can facilitate information exchange through the vendor-specific command line, convert it to southbound API message/protocol and vice versa. Monitoring plane can monitor information from TRx like wavelength, modulation format, forward error correction (FEC) coding schemes, launch and received power. It can also collect corresponding EDFA data, such as input and output power, amplification gain, etc. The monitored information of an optical network domain can be fed into its domain abstraction layer to enable single network domain level abstraction (impairments abstraction and resources abstraction). As detailed knowledge of an optical network is not permitted to be shared between network providers for provisioning inter-domain services, each network provider can thus advertise and share the abstracted information for service provisioning.

To achieve the above purpose, we propose an impairments abstraction framework to capture the general effects of all types of impairments in the network. SNR degradation factor is introduced as the impairments abstracted entity, which can indicate how the quality of the optical signal degrades in a single link in the presence of various impairments. Combining the SNR degradation factors over the path, the end-to-end performance can be formulated as:

$$SNR_{end-end} = 10 \cdot log_{10}\left[\left(\sum_{n}\sum_{l}\frac{1}{10^{\frac{SNR_{n,l}}{10}}} + \frac{1}{10^{\frac{SNR_{TRx}}{10}}}\right)^{-1}\right] \quad \textbf{(1)}$$

where $SNR_{n,l}$ is the SNR degradation factor of link l within network n and the SNR_{TRx} represent the performance of the TRx, which are usually quantified through the back-to-back measurement. According to the equation, the end-to-end performance can be splitted into two parts: the performance of the optical links and the performance of TRx. The back-to-back TRx performance measurement can be carried out off-line before deployed into the network and the performance of optical link/network can be obtained through ML techniques, which will both be shown in the following sections.

3. BACK-TO-BACK PERFORMANCE MEASUREMENT

The back-to-back performance of a pair of TRx presents its ideal performance without the interference of ASE noise, dispersion and non-linear impairments from fibre links. It is often neglected in impairments-aware optical networking, however, playing an important role in impairments management, especially for short-distance transmission. Its performance varies against detection technique (direct detection or coherent detection), modulation scheme, digital signal processing techniques and actual implementing hardware. The experimental setup of back-to-back performance measurement includes the transmitter connected to the receiver via a 3 dB coupler. An ASE noise source is also connected to the coupler to provide a different level of ASE noise by adjusting the variable optical attenuator. This would allow the control of signal optical signal-to-noise ratio (OSNR). In the back-to-back performance measurement, we utilise dual-polarisation (DP)-16 quadrature amplitude modulation (QAM) as the modulation format with 15% FEC overhead for the transmitter and compatible settings of coherent detection for the receivers. Due to space limit and visual convenience, the performance of four TRx out of in total 24 TRx (to be deployed in this work) are illustrated in Fig. 2. The received SNR of four TRx is plotted as the function of measured OSNR, which is calculated as the ratio of signal power to the average power of 0.1 nm out-band ASE noise from both sides of the signal spectrum. The figure indicates a linear relationship between received SNR and OSNR for all four TRx when OSNR is less than 25 dB. Increasing OSNR beyond this value does not lead to proportional growth of received SNR, mainly due to the limitation of the electronic components inside the TRx like shot noise or thermal noise of digital-to-analog converter, analog-to-digital converter, filter and electrical amplifier. The received SNR of the TRx saturates approximate between 19.3 and 19.8 dB while decreasing the power of coupled ASE noise. TRx adopting DP-quadrature phase shift keying (QPSK) shows a similar SNR versus OSNR performance compared to using DP-16QAM scheme. Therefore, for simplicity, we denote the SNR degradation of all the TRx with 15% FEC overhead scheme to be 19.4 dB in this work regardless of their modulation formats. The impact of optical fibre links in terms of SNR degradation can be obtained by subtracting the performance of TRx from the end-to-end monitored SNR, which can be expressed as:

$$SNR_{link} = 10 \cdot log_{10}\left[\left(\frac{1}{10^{\frac{SNR_{measured}}{10}}} - \frac{1}{10^{\frac{SNR_{TRx}}{10}}}\right)^{-1}\right] \quad \textbf{(2)}$$

where $SNR_{measured}$ is the SNR performance (dB) reported by receivers. In the following section, the hybrid learning assisted impairments abstraction model will be introduced and discussed.

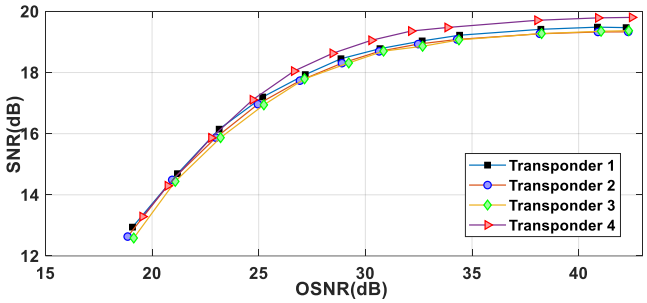

Fig. 2. Back-to-back characterisation of the transponders.

4. HYBRID LEARNING ASSISTED IMPAIRMENTS ABSTRACTION MODEL

Compared to impairments of the TRx, impairments from optical links are the dominant factor to limit the performance of the

lightpaths. The impairments change dynamically with varying temperature, the assigned wavelength, link loading condition, fibre vibration, etc. In this work, we adopt the ML methods to learn and predict the proposed SNR degradation entity as described in Section 2. Conventional ML techniques for QoT estimation aims to predict the end-to-end performance of the target lightpath. However, end-to-end QoT prediction which requires information monitoring along the lightpath is not feasible for service learning across multi-network domains due to privacy violation as explained in Section 2. Furthermore, monitoring functions may not be applicable even for the single optical network which lacks sufficient monitoring techniques for all the devices. Hence, we propose a hybrid learning assisted impairments abstraction model to overcome the issue.

A. Hybrid learning model overview

The proposed hybrid learning assisted impairments abstraction model in this work consists of two learning modules: a deep learning module (parametric ML method) and a GPR module (non-parametric ML method). Both modules aim to learn the SNR degradation factor of an optical link or several links as proposed in Section 2 according to the routing and wavelength assignment decision. As deep learning module requires the detailed monitoring features to feed into the learning model, it is suitable to learn the SNR degradation factor from the intra-domain services. However, it fails to perform learning from inter-domain services as exchanging and exposing complex and detailed knowledge between networks violates private and confidential policy. As a result, we adopt GPR to learn the SNR degradation factor from the inter-domain services. This is because as a non-parametric ML technique, GPR only requires SNR information reported from the receivers while the SNR itself is the proposed abstraction entity. It worth noting that GPR can also perform as an alternative abstraction tool for learning from intra-domain services. However, we only consider deep learning for intra-domain services in this work due to its high accuracy.

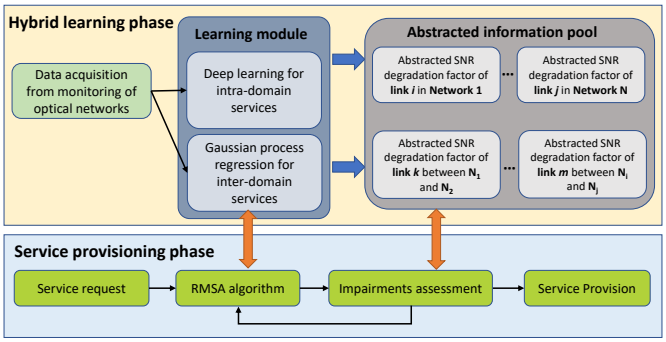

Fig. 3. Hybrid learning assisted impairments abstraction learning phase and service provisioning phase.

The hybrid abstraction learning phase is illustrated as the light yellow block in Fig. 3. The data or the features in the optical network are acquired and fed into the two learning modules according to their service types for training purpose. The SNR degradation factors learned from intra-domain and inter-domain services are routing, modulation and spectrum assignment (RMSA) algorithm dependent. Different routing and spectrum assignments will affect the network performance, including the existing embedded services and service to be provisioned. Therefore, the abstracted information pool with the shared SNR degradation factors is valid only one time for the particular new service. Any service setup or tear down, lightpath re-routing or EDFA adjustment will require the training modules providing a new abstracted information pool even for the same RMSA decision.

In the service provisioning phase shown as the light blue block, the RMSA algorithm is able to provide routing and spectrum allocation decision for the new service. The learning modules will calculate the SNR degradation factors based on the selected path and spectrum to form the abstracted information pool for the new service. Then, the impairments assessment function retrieves the relevant SNR degradation factors from the abstracted information pool and combines the performance of TRx for end-to-end QoT evaluation according to Eq. 1. The information in the pool might be redundant for the selected path and wavelength. For example, RMSA calculates a wavelength using path A-B-C-D for the new request between A and D. The abstracted information pool might contain the SNR degradation factors of that wavelength of segmental link A-B, B-C, C-D, A-B-C, B-C-D or even A-B-C-D, as different SNR degradation factors are evaluated based on different data sets and possibly different ML techniques. The criteria for selecting the most suitable SNR degradation factors from the abstracted information pool for the end-to-end QoT evaluation is outside the scope of this paper. We follow a simple and basic rule: selecting the SNR degradation factors trained from the deep learning module if possible, and then selecting the combination for QoT evaluation with a minimum number of SNR degradation factors if possible. In the end, the new service will be provisioned providing end-to-end QoT requirement are met for the selected path, spectrum and modulation formats, otherwise, RMSA algorithm needs to calculate a new assignment result for another round of abstracted information pool generation and end-to-end QoT evaluation.

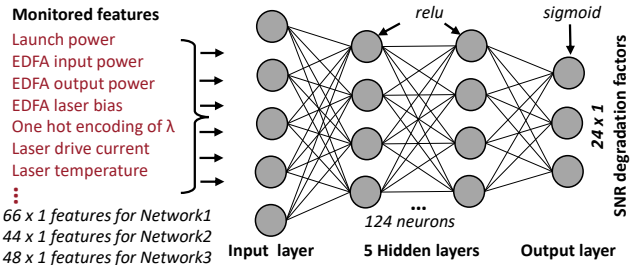

Fig. 4. DNN architecture implemented for intra-domain service learning.

B. Deep learning for impairments abstraction

Deep learning technique has recently received great attention in natural language processing and computer vision [28]. Deep learning also known as deep neural networks DNN, are neural networks equipped with the multi-layer feed-forward architecture [29]. Activation functions in the multi-layer are capable of nonlinear transformation, thus achieving universal approximation and probabilistic inference [30]. As a result, it becomes an ideal candidate to capture the linear and nonlinear characteristics of fibre links in the optical networks and map them to the proposed SNR degradation factor. Input features of DNN vary against different optical network domains and optical links with different monitored data. Every network provider/operator can design and adapt their own DNN for impairments abstrac-

tion according to the scale of their networks, topology, monitoring techniques, accessible data from the physical layer devices. However, they need to follow the common strategy to learn and abstract the impairments as the SNR degradation factor for lightpaths within their network domains. The deep learning module will provide the SNR degradation factors of the links in one network domain according to the result of the deployed RMSA algorithm. The obtained abstraction entities will then be uploaded to the abstracted information pool, which can be used to calculate end-to-end QoT prediction for the new service. The DNN abstraction process can be formulated as:

$$\begin{bmatrix} D_{SNR}^{\lambda^m} \\ \vdots \\ D_{SNR}^{\lambda^q} \end{bmatrix} = F\left(EDFA_{attr_r}^{i}, ..., TRx_{attr_s}^{j}, ..., freq_{one-hot}^{l} \right) \quad (3)$$

where F represents the function of DNN to compute SNR degradation factors of the specific segmental link of intra-domain services. $D_{SNR}^{\lambda^m}$, as the output of DNN, is the SNR degradation factor for the channel of interest λ^m in the segmental link. The input features of DNN, $EDFA_{attr_r}^{i}$ and $TRx_{attr_s}^{j}$, is the attribute r and s of $EDFA^i$ and TRx^j in the network respectively. $freq_{one-hot}^{l}$ is the one-hot encoding of wavelength assignment condition of link l in the network.

In this work, we adopt DNN with 5 hidden layers despite the different size and features of the input layers for different optical networks, as illustrated in Fig. 4. The input layers contains the monitored features including but not limited to signal launch power, the input and output power of each EDFA along with the link, laser drive current and temperature of EDFA, one-hot encoding to convert wavelength to 0 or 1 according to its 'on-off' status. However, the sizes of input layer of DNN vary against different monitoring schemes, topology and the scales in different networks. '*Min-max*' normalisation technique is adopted for all the training data set to prevent gradient-descent optimiser oscillating around global/local optimum. Each hidden layer consists of 124 neurons and '*relu*' serves as the activation function. Ideally, the output layer should have a dimension of 80×1, which corresponds to 80 wavelengths of C-band in DWDM system. For simplicity of training, the output layer is of size 24×1 with 24 TRx deployed. As the SNR degradation factors are normalised, the '*sigmoid*' is selected as the activation function at the output layer for better performance. Details of implementation of DNN are presented in the following section. The proposed deep learning module thus can capture the network status and provide the intra-domain SNR degradation factors according to the RMSA algorithm of the new service request.

C. Gaussian process regression for impairments abstraction

Although deep learning is an excellent candidate for impairments abstraction, it relies on adequate amount of training data set. However, the training samples might not be available for a variety of reasons, such as for inter-domain services. In this case, we adopt the GPR for SNR degradation factor learning. Gaussian process can be regarded as an infinite dimensional Gaussian distribution where any finite subset range follows a multivariate Gaussian distribution [31]. We subtract performance of TRx from the SNR measured by receivers of the existing lightpaths to calculate pure optical link performance in form of SNR degradation factor (in dB), as shown in Eq. 2. Then the calculated SNR degradation factors of optical links for the existing lightpaths act as the training data for the GPR model. As the measured SNR degradation factor for a one scenario is noisy for varying monitoring time and techniques, we assume the function of SNR degradation factors follow Gaussian process:

$$f_{SNR}(\boldsymbol{\lambda}) \sim GP(m_{SNR}(\boldsymbol{\lambda}), K_{SNR}(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{'})) \quad (4)$$

where the $m_{SNR}(\boldsymbol{\lambda})$ is the mean of all the functions in the distribution at wavelength $\boldsymbol{\lambda}$ and the $k_{SNR}(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{'})$ is the covariance matrix, often known as kernel [32]. We adopt the squared exponential function as the kernel, which is expressed as:

$$\begin{aligned} K_{SNR}(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{'}) &= \sigma_f^2 \cdot exp\Big(- \frac{\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^{'}\|^2}{2l^2} \Big) + \sigma_n^2 \cdot \delta(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{'}) \\ &= \begin{bmatrix} k(\lambda_1, \lambda_1) & k(\lambda_1, \lambda_2) & \cdots & k(\lambda_1, \lambda_n) \\ k(\lambda_2, \lambda_1) & k(\lambda_2, \lambda_2) & \cdots & k(\lambda_2, \lambda_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\lambda_n, \lambda_1) & k(\lambda_n, \lambda_2) & \cdots & k(\lambda_n, \lambda_n) \end{bmatrix} \end{aligned} \quad (5)$$

where the $\delta(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{'})$ is the Kronecker delta function. The hyper parameters length-scale l, 'noise variance' σ_n and 'signal variance' σ_f vary for different levels of prior correlation between training points (λ_s). When the mean function (from training data set) and kernel function in Eq. 4 is selected, we can use the Gaussian process for posterior inference (estimation) based on the observations. The SNR degradation factor for the channel of interest λ^* is denoted as $f_{SNR}(\lambda^*)$, where the joint distribution can be expressed as:

$$\begin{bmatrix} \boldsymbol{f_{SNR}(\lambda)} \\ f_{SNR}(\lambda^*) \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \boldsymbol{K_{SNR}(\lambda, \lambda)} & \boldsymbol{K_{SNR}(\lambda, \lambda^*)} \\ \boldsymbol{K_{SNR}(\lambda^*, \lambda)} & K_{SNR}(\lambda^*, \lambda^*) \end{bmatrix} \right) \quad (6)$$

The conditional distribution $p(f_{SNR}(\lambda^*) | \boldsymbol{\lambda}, \boldsymbol{f_{SNR}(\lambda)}, \lambda^*)$ as a multivariate Gaussian distribution gives how likely the SNR degradation prediction is for a new channel λ^* given the performance of existing channels. Its mean function is shown as:

$$m_{SNR}(\lambda^*) = \boldsymbol{K_{SNR}(\lambda^*, \lambda)} \cdot \Big[\boldsymbol{K_{SNR}(\lambda, \lambda)} \Big]^{-1} \cdot \boldsymbol{f_{SNR}(\lambda)} \quad (7)$$

and its variance is estimated as:

$$\begin{aligned} var(f_{SNR}(\lambda^*)) &= K_{SNR}(\lambda^*, \lambda^*) - \\ & \boldsymbol{K_{SNR}(\lambda^*, \lambda)} \cdot \Big[\boldsymbol{K_{SNR}(\lambda, \lambda)} \Big]^{-1} \cdot \boldsymbol{K_{SNR}(\lambda, \lambda^*)} \end{aligned} \quad (8)$$

Usually, the kernel has unknown hyper-parameters ($\boldsymbol{\theta} = (l, \sigma_f^2, \sigma_n^2)$) which can be inferred and tuned from the training data set. The hyper-parameters can be estimated by maximising its log marginal likelihood with the given observation data set $(\boldsymbol{X}, \boldsymbol{y})$, which is expressed as:

$$log\ p(\boldsymbol{y} | (\boldsymbol{X}, \boldsymbol{\theta}) = -\frac{1}{2} \boldsymbol{y}^T \boldsymbol{K}_y^{-1} \boldsymbol{y} - \frac{1}{2} log |\boldsymbol{K}_y| - \frac{n}{2} log 2\pi \quad (9)$$

where $\boldsymbol{K}_y$ is the covariance matrix of the training data. The marginal likelihood can be optimised by applying gradient-decent algorithm to its reciprocal. With hyper-parameters finely tuned based on the observations, the GPR can act as a abstraction tool to offer inter-domain SNR degradation factors abstraction for provisioning new services.

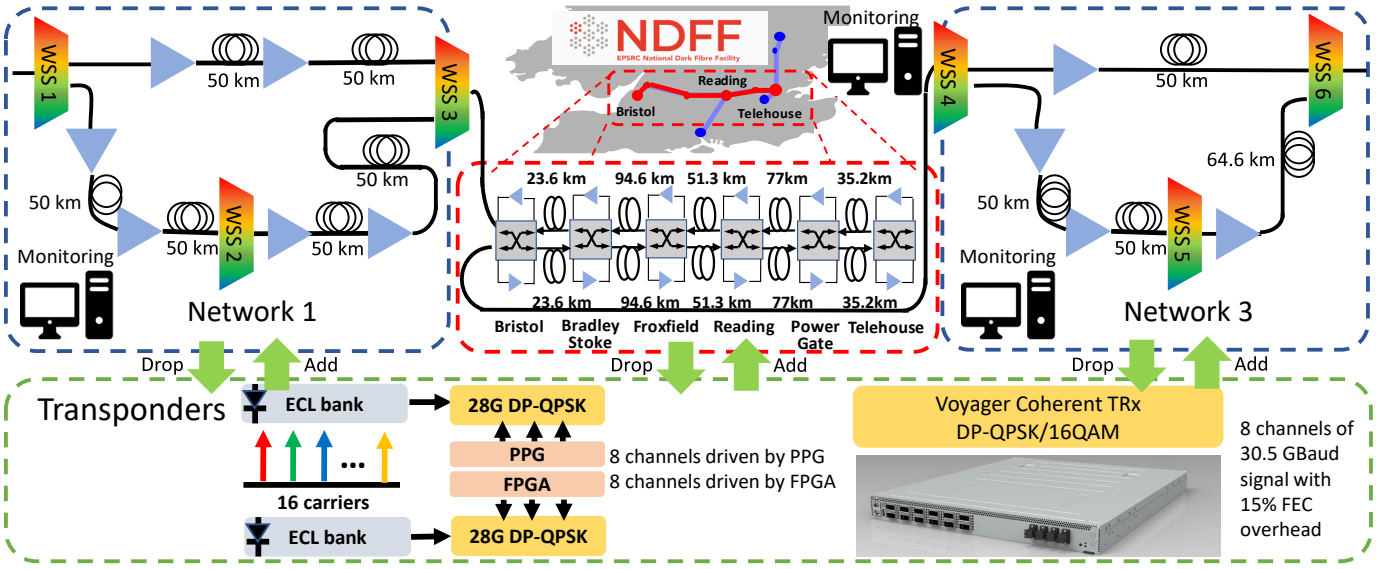

Fig. 5. Field-trial experimental testbed setup includes 3 optical network domains with network monitoring: Network 1 (N1), NDFF (N2) and Network 3 (N3), TRx arbitrarily add/drop to any of the network


```
"_id" : ObjectId("5da5ab9d7b3765411c2d0d78"),
    "PolatisConfigDynamic" :
          "RecordID" : 1,
          "EDFA_ID" : "22",
          "Device_Name" : "EDFA_PGT_2",
          "Link_ID" : 7
    "EDFA_Data" :
          "Device_Output_Port" : "2",
          "Device_IP" : "195.194.3.227",
          "Temperature" : 20,
          "Laser_Bias" : "149",
          "Device_Input_Port" : "2",
          "Output_Power" : 12.8,
          "Timestamp" : 1571138461.377551,
          "Laser_Temperature": 25,
          "Gain" : 13.5,
          "Input_Power" : -0.7
```


Fig. 6. Mongo database information of one EDFA at Powergate of NDFF.

5. EXPERIMENTAL SETUP

In this section, the experimental setup of field-trial testbed is introduced. The data monitoring for the testbed optical networks is then demonstrated. We present how the proposed hybrid learning assisted impairments abstraction can be obtained, tuned and trained with the monitored data.

A. Field-Trial Experimental Testbed

Field-trial experimental setup to demonstrate the proposed hybrid learning assisted abstraction is illustrated in Fig. 5. The testbed consists of three optical networks domains. Network 1 (N1) and Network 3 (N3) are set up in the lab and the middle one (N2) is National Dark Fibre Facility (NDFF) in the UK, a real operating infrastructure, which connects the University of Bristol, University of Cambridge, University College London and University of Southampton. The N1 and N3 are both three-node networks consisting of three optical links, where their lengths vary from 50 km to 100 km, as depicted in Fig. 5. In this experiment, we use the NDFF from the University of Bristol to Telehouse and then loop back to our lab. Each intermediate node is associated with a Polatis switch, connected with EDFAs to provide amplification for bi-directional communications. The NDFF in this experiment consists of 10 fibre spans with overall 564 km. The whole experimental setup is depicted in Fig. 5. Wavelength-selective switch (WSS) is emulated as the reconfigurable optical add-drop multiplexer in this experiment with add, drop, switching and power equalisation functionality. WSS3 and WSS4 are regarded as the border nodes, both with the capability to serve service for N1 and N2 or N2 and N3. 24 coherent TRx are shown in the lower part of the figure, which can be arbitrarily added or dropped through any WSS as long as colour, direction and contention (CDC) constraint is satisfied. The power of different signal will be adjusted to the same level by WSS providing they are routed to the same fibre link. Amid 24 coherent TRx, we have 16 coherent signals generated by modulating carriers from two external cavity laser (ECL) banks. One ECL bank is driven and modulated by field programmable gate array (FPGA) and the other one is driven by pulse-pattern generator (PPG) while both can provide DP-28 Gbaud QPSK signal. Besides, two Facebook Voyager switches are deployed, each with 4 TRx. The TRx modules inside the Voyager switches are capable of adapting the modulation formats to be DP-QPSK, DP-8QAM, DP-16QAM, FEC schemes with different portions of overhead, operating frequency and the launch power.

For quick measurement and fair comparison, we utilise the coherent receiver module in Voyager switches to detect the pre-FEC bit error rate (BER) and then convert it to SNR for SNR degradation factor calculation. When measuring the BER of channels generated by PPG and FPGA, we swap their frequency with the frequency of Voyager TRx. In this case, the signal from ECL banks are considered as partial-dumb channels but the performance of the occupied wavelengths in the optical network can be measured and calculated by subtracting the SNR performance of TRx from the measured SNR. This would enable fast and fair comparison of channels launched with different frequencies compared to utilising oscilloscope with offline digital

signal processing.

B. Data monitoring and training for abstraction

We first focus on the data collection of the intra-domain services, which can enable the training of the DNN to provide the proposed impairments abstraction of links in each individual network domain. As N1 and N3 are the networks being set up in the lab while the N2 is a real deployed network, the amount and dimension of accessible information is different. However, we mainly focus on the data collection of EDFAs for all three networks, since ASE and nonlinearity play a critical role in the era of coherent detection. The information is collected via vendor's specific API in this work for every new event in the networks. Mongo database is implemented in N2 to monitor the EDFA information, as depicted in Fig. 6. Data collection in N2 not only includes input/output power of EDFA, device temperature, EDFA laser temperature, laser bias of EDFA, but also contains signal launch/received power, the one-hot encoding of the assigned wavelength of TRx and end-to-end BER of the lightpaths. While in N1 and N3, laser drive current/power of each stages in EDFA and their temperature are also included besides the above features.

We first launch all the channels via WSS1 and drop them at WSS3 to emulate the intra-domain services in N1 for training its deep learning module. All the wavelengths are able to select either the path WSS1-WSS3 or WSS1-WSS2-WSS3 to study the effect of different routing decisions. Different channels are then randomly manipulated to be 'live' or 'off' to generate different loading conditions of the optical network. In this experiment, the 24 channels are equally positioned with 100 GHz spacing ranging from 1541 nm to 1560 nm. The total spectrum spanning over 2.4 THz, is capable to explore the impact of EDFA gain ripple or fluctuation to the signal quality. The one-hot encoding translates the loading condition of 24 wavelengths occupation information to a binary 1×24 vector [0, 1, 0, 1, 1, ... , 0, 1]. The value of i_{th} element in the vector is 1 when the i_{th} wavelength is launched and 0 for i_{th} wavelength to be off. As generating all the 2^{24} channel loading scenarios is neither experimentally feasible nor meaningful for ML, we randomly select 8 TRx which can be turned on and off while keeping the rest 16 channels always launching as the background traffic. As a result, 255 scenarios (apart from all 8 TRx being off) are experimentally implemented with each scenario executed and measured 5 times. In total, this leads to 1275 training samples for a selected route in N1. The input features not only consist of one-hot encoding of wavelength loading conditions of the assessed path, but also contains the vectorised input/output power, bias of EDFAs, laser current of each stage of EDFA, the signal transmission and received power. It can be represented as a vector $[wavelength\ one-hot\ encoding, EDFA^i_{in}, EDFA^i_{out}, EDFA^i_{bias}, EDFA^{i,j}_{ls}..., P^m_{Tx}, P^m_{Rx}, ...]$. The $EDFA^i_{in}$ and $EDFA^i_{out}$ are the input and output power of i_{th} EDFA respectively, $EDFA^i_{bias}$ presents the laser bias of i_{th} EDFA, $EDFA^{i,j}_{ls}$ is the laser current of j_{th} stage of i_{th} EDFA, P^m_{Tx} and P^m_{Rx} denotes the transmission power and received power of m_{th} pair of TRx. The size of the feature vector may vary against different networks with different monitoring techniques and availble information.

Similar settings are applied to enable the intra-domain services training for deep learning modules in N2 and N3 with 1275 training samples respectively for a selected path. The input features in N1 is of dimension 66×1, while N2 and N3 have 44×1 and 48×1 input features for their own DNN respectively. The output layer is a 24×1 vector $[0, 0, SNR^{n,\lambda_m}_{dg}, 0, ...]$, where the SNR^{n,λ_m}_{dg} is the SNR degradation factor of λ_m of the selected path in a particular network domain and 0 represents signal of the particular wavelength not launching. Apart from the number of hidden layers, the number of neurons in each hidden layer and the activation functions, we also set other hyper-parameters in the deep learning module. Adam optimiser [33] is adopted with the learning rate $\alpha = 0.0005$, the first exponential decay rate $\beta 1 = 0.9$ and the second exponential decay rate $\beta 2 = 0.999$ after several attempts of manually hyper-parameters tuning. Mini batch gradient descent with the batch size of 16 is applied to enable efficient training process. In this experiment, we also adopt L2 regularisation to prevent overfitting issue with the value set to be 0.00001. For each network domain, 70% of 1275 scenarios and their corresponding information are randomly selected for training (50% as training data set and 20% as validation data set), while the rest 30% are used as test data set.

Compared to deep learning, GPR is more straightforward which does not require complex monitoring techniques. It relies on the SNR degradation of co-propagation channels in the network to predict the SNR degradation factor of the unestablished channels. As mentioned in the previous section, SNR degradation factor of existing co-propagation channels for the optical links can be obtained via subtracting the performance of TRx from the end-to-end SNR performance measured by receivers. The SNR degradation factors as the input of Gaussian process model can be presented as: $[SNR^{n,\lambda_o}_{dg}, SNR^{n,\lambda_p}_{dg}, SNR^{n,\lambda_q}_{dg}, ...]$ and the corresponding wavelength presented as $[\lambda_o, \lambda_p, \lambda_q, ...]$. The SNR degradation performance of the new channels $[\lambda_r, \lambda_s, ...]$ expected to establish along the co-propagation links can be inferred using the GPR proposed in Section 4 as $[SNR^{n,\lambda_r}_{dg}, SNR^{n,\lambda_s}_{dg}, ...]$. Due to this characteristic, GPR can also apply to SNR degradation factor prediction of multi-links or inter-domain lightpath, such as path *WSS1-WSS2-WSS3-WSS4*. We consider using GPR in this work to predict the SNR degradation factor of the new channel for inter-domain services. Similar to the deep learning, the training data is generated assuming the channels can be turned on and off. Besides, 3 different inter-domain scenarios are considered including: 1) launching signal at N1 using WSS1 and drop at N2 using WSS4 (path *WSS1-WSS2-WSS3-WSS4*); 2) launching signal at N2 using WSS3 and drop at N3 using WSS6 (path *WSS3-WSS4-WSS5-WSS6*) and 3) launching signal at N1 using WSS1 and drop at N3 using WSS6 (path *WSS1-WSS2-WSS3-WSS4-WSS5-WSS6*). The training process of GPR mainly targets on hyper-parameters optimisation to achieve the maximum marginal likelihood. To keep training process consistency, the same scenarios for the DNN will also be adopted for GPR.

According to our evaluation of GPR, a typical execution time varies from a few seconds to tens of seconds to optimise the corresponding hyper-parameters of the selected kernel. With 15 training samples, the optimiser converged after 1000-3000 iterations compared to 5000 - 8000 iterations with 7 training samples. DNN require significant more time compared to GPR, with manually tuning different combinations of hyper-parameters. The processing time for DNN increases with the increasing number of training data sets, the number of features, the number of hidden layers, and the number of neurons. For each combination of hyper-parameters trial, it takes between 1 and 3 minutes to complete the training process with reasonable accuracy.

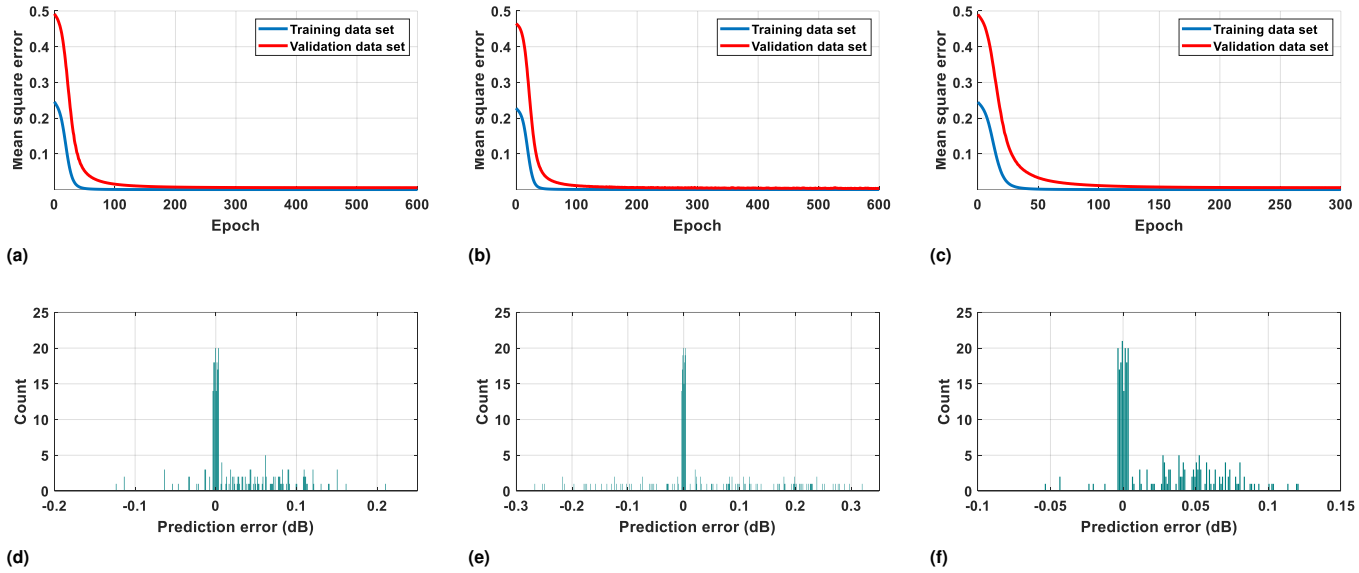

Fig. 7. Performance of intra-domain service abstracion using deep learning. (a) Normalised MSE vs Epoch for intra-domain services in N1; (b) Normalised MSE vs Epoch in N2; (c) Normalised MSE vs Epoch in N3; (d) Histogram of SNR prediction error distribution in N1; (e) Histogram of SNR prediction error distribution in N2; (f) Histogram of SNR prediction error distribution in N3.

6. RESULTS ANALYSIS

In this section, the experimental results of DNN for intra-domain services are presented first, then followed by the results of GPR for inter-domain services. Finally, the end-to-end performance of hybrid learning assisted abstraction are demonstrated and compared with different combinations of DNN and GPR results of segmental links along the end-to-end lightpath.

A. Results of deep learning for intra-domain services

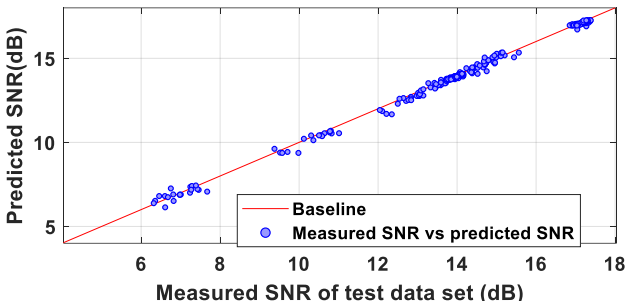

Fig. 8. Predicted SNR using deep learning versus measured SNR for test intra-domain services in three optical networks.

The overall performance of applying DNN for intra-domain service learning is illustrated in Fig. 7. Specifically, Fig. 7a, 7b and 7c show the normalised mean square error (MSE) versus the increasing number of epoch during the training process for intra-domain services with routing path of *WSS1-WSS2-WSS3* in N1, *WSS3-WSS4* in N2 and *WSS4-WSS5-WSS6* in N3 respectively. The figures indicate the normalised MSE decreases against the increasing number of training iterations or epoch. As the training process starts, the normalised MSE of predicted SNR degradation factors first reduce significantly both for training data set and validation data set. Then the decreasing trend becomes slow and finally, the normalised MSE of training data set and validation data set converge and become stable after 200 epoch for N1 and 150 epoch for N2 and N3. Once the training process completed, the rest 30% of scenarios are tested and evaluated. The histogram of predicted SNR errors of the corresponding path in N1, N2 and N3 are depicted in Fig. 7d, 7e and 7f respectively. It can be observed that the majority of error counts are distributed around 0 dB for all three networks. It also proves DNN are capable to provide high prediction accuracy of SNR with 92.4 %, 80.4% and 98.9% errors falling within 0.1 dB estimation error for N1, N2 and N3 respectively. Fig. 8 demonstrates the predicted SNR against the measured performance of all intra-domain service scenarios of N1, N2 and N3 in the test data set. The predicted SNR is calculated by combining the SNR degradation factors of optical link obtained from deep learning module and the SNR degradation factors of TRx. The figure indicates the predicted SNR using the proposed abstraction framework achieves high estimation accuracy for intra-domain services in different optical networks, with all the scattered points close to the red baseline.

Fig. 9 shows the performance of two intra-domain service scenarios in the test data set for N1, N2 and N3 using the deep learning for impairments abstraction. We launch the same signal spectrum into all the three optical networks separately following the path *WSS1-WSS2-WSS3* in N1, *WSS3-WSS4* in N2 and *WSS4-WSS5-WSS6* in N3 for individual scenario. Due to the number of available TRx and the CDC constraint of WSS, the experiments of the same scenario for intra-domain services learning in different networks are carried out over a different period. There are 16 background intra-domain services in each network shown as the purely blue or black spectrum in Fig. 9a and 9e for two scenarios respectively. Then we emulate that the RMSA algorithm would assign the same route for the new services compared to the existing 16 services, for example, *WSS1-WSS2-WSS3* in N1. Given the

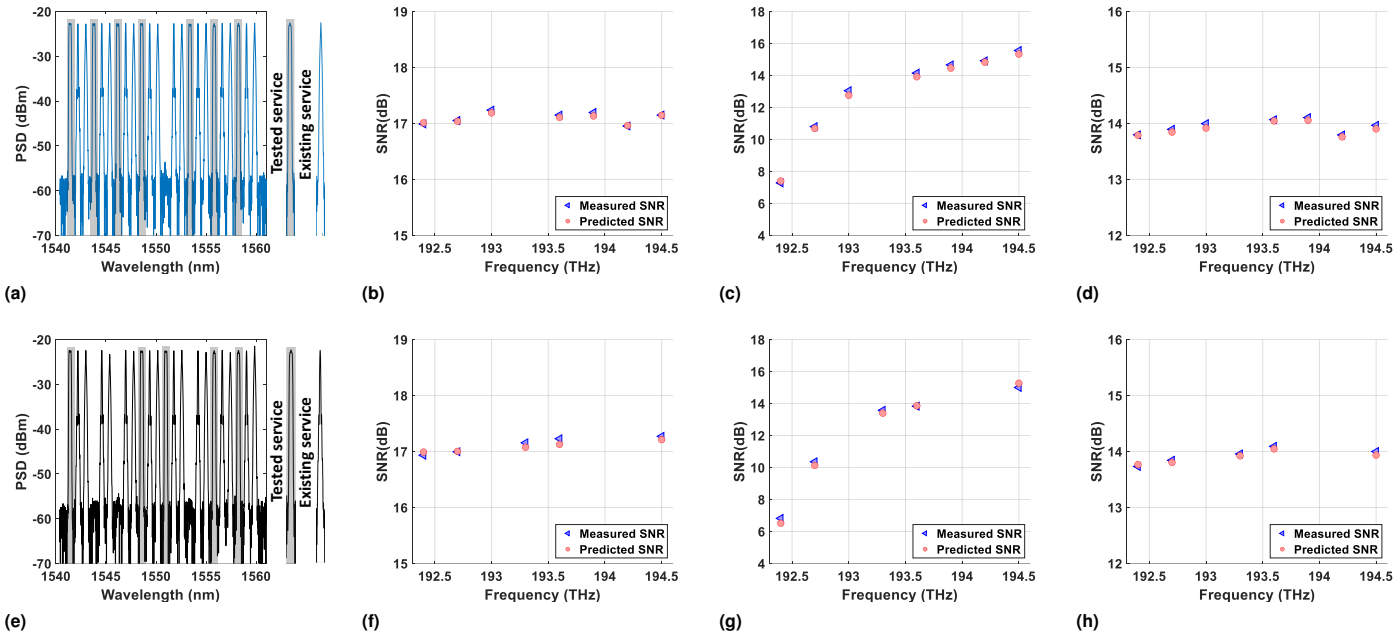

Fig. 9. Two intra-domain service scenarios and their performance in three optical networks using deep learning. The coloured (blue or black) spectrum is the existing channels and the grey covered spectrum represents the tested channels. (a) Launching signal spectrum of the first scenario; (b) The performance of the first scenario for services in N1; (c) The performance of the first scenario in N2; (d) The performance of the first scenario in N3; (e) Launching signal spectrum of the second scenario; (f) The performance of the second scenario for services in N1; (g) The performance of the second scenario in N2; (h) The performance of the second scenario in N3.

results of RMSA, the evaluation of end-to-end SNR is performed using the proposed impairments abstraction framework. The new services to be established are shown as the grey block in Fig. 9a and Fig. 9e for the first and second scenario separately. In the first scenario, the end-to-end SNR performance of seven new intra-domain services in N1, N2 and N3 are evaluated and are depicted as the red circle markers in Fig. 9b, 9c and 9d respectively. Once the predicted end-to-end SNR satisfies the SNR constraint for the assigned modulation format, the seven new services will then be provisioned accordingly. Their spectrum are measured as the grey block covered blue spectrum shown in Fig. 9a. The end-to-end SNR performance of these new services (tested services) is measured and demonstrated as the blue triangle markers in Fig. 9b for services in N1, Fig. 9c for services in N2 and Fig. 9d for services in N3. Similarly, the predicted QoT performance of the second scenario with five new intra-domain services in N1, N2 and N3 are shown as the red circle markers in Fig. 9f, 9g, and 9h individually. After provisioning new services in scenario 2, their performance is plotted as the blue triangle markers in Fig. 9f - 9h for three networks. From the figure, it is observed that the proposed impairments abstraction framework deployed with deep learning can achieve high abstraction accuracy, thus can provide a valuable tool for intra-domain service learning.

B. Results of GPR for inter-domain services

Similar to intra-domain service learning, RMSA algorithm calculates and assigns the wavelengths and routes for the new inter-domain services. Based on the RMSA results, the proposed impairments abstraction framework executes the GPR to fill in the abstracted information pool based on the performance of the existing co-propagated channels, as shown in Fig. 3. Co-propagated channels are also known as training channels, which share the same routes of the new services expected to be established. Their SNR degradation performance is measured from corresponding receivers as the training data set. We emulate the RMSA assigns the routes of two types of inter-domain services in the following way: 1) service across N1 and N2 using path *WSS1(add)-WSS2(switch)-WSS3(switch)-WSS4(drop)*, and 2) service across N2 and N3 using path *WSS3(add)-WSS4(switch)-WSS5(switch)-WSS6(drop)*. We focus on the analysis of the results for the same network condition as shown in Fig.9a for GPR assisted inter-domain service learning.

Fig. 10 demonstrates the performance of using GPR for two inter-domain services learning with a different number of training samples for the proposed scenario. The red square markers represent the performance of the existing co-propagated channels or training channels. Based on the performance of these channels, the GPR calculates the estimation mean of unestablished channels of different wavelengths shown as the red line, and 95% confidence interval as the grey area. The blue circle markers in the figures denote measured QoT performance when the new inter-domain services are provisioned after successful impairments evaluation using the proposed abstraction framework. Fig. 10a, 10b and 10c depict the performance of GPR for inter-domain services between N1 and N2 with 15, 10 and 7 training samples respectively. From the results, the GPR is able to provide high estimation accuracy when there is a sufficient training data set. The performance of all the test channels falls within 95% confidence interval with 15 samples achieving an average of 0.21 dB estimation error. The estimation mean values also perform very closed to the measured SNR of newly provisioned inter-domain services. GPR can achieve a smaller

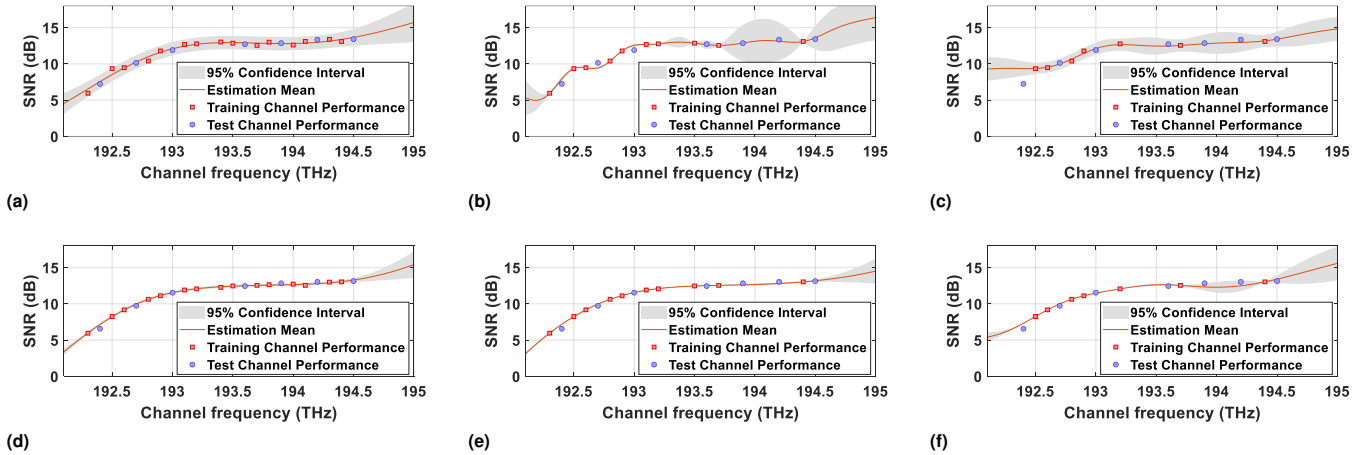

Fig. 10. The performance of GPR for inter-domain service learning. (a)N1-N2 with 15 training samples; (b)N1-N2 with 10 training samples; (c)N1-N2 with 7 training samples; (d)N2-N3 with 15 training samples; (e)N2-N3 with 10 training samples; (f)N1-N2 with 7 training samples.

area of 95% confidence interval with more training samples thus have better abstraction accuracy. However, the area of 95% confidence interval increases in the region with fewer training data for GPR as shown in Fig. 10b, thus yielding less confidence in the estimation. Especially, it fails to capture the performance of the new channel of frequency 192.4 THz due to EDFA gain ripple with 7 training samples. The GPR assisted learning results for inter-domain services between N2 and N3 are shown in Fig. 10d, 10e and 10f with 15, 10 and 7 training samples respectively. Similarly, the GPR is able to perform accurate impairments abstraction with a sufficient amount of training data. However, the measurement of training channels itself is less noisy for services between N2 and N3 compared to service between N1 and N2. As a result, GPR calculates a smaller area of 95% confidence interval. The area of 95% confidence interval grows as the available training samples reduce. With 7 training samples, the real QoT performance of the test channels with frequency 193.9 THz and 194.2 THz is around the boundary of the 95% confidence interval while the channel with the frequency of 192.4 THz falls out of the boundary. However, due to the less noisy observation data from training channels, the SNR prediction for services across N2-N3 achieves around 0.37 dB estimation error with 7 training samples.

C. Results of hybrid learning for end-to-end performance

In the field trial experiment, we aim to establish seven new inter-domain services across N1, N2 and N3. It is emulated that RMSA algorithm assigns the routing path for the new services as *WSS1(add)-WSS2(switch)-WSS3(switch)-WSS4(switch)-WSS5(switch)-WSS6(drop)*. The network condition along the selected path is identical to the scenario shown in Fig. 9a with 16 existing co-propagated channels. The seven new inter-domain services are assigned the frequency of 192.4 THz, 192.7 THz, 193.0 THz, 193.6 THz, 193.9 THz, 194.2 THz and 194.5 THz respectively, corresponding to the grey covered spectrum. To predict their performance, the learning module in the abstraction framework adapt, learn and yield SNR degradation factors to form the abstracted information pool based on current network condition, RMSA assignment of new services and historical monitoring data of both intra-domain and inter-domain services in all three networks. The impairments assessment module will then retrieve the SNR degradation factors from the abstracted information pool for end-to-end QoT evaluation for the new services. Under the emulated case, the abstracted information pool may contain the full or part of SNR degradation factor set: SNR degradation factor of *N1 (WSS1-WSS2-WSS3)* learned from deep learning, of *N2 (WSS3-WSS4)* learned from deep learning, of *N3 (WSS4-WSS5-WSS6)* learned from deep learning, of *N1-N2 (WSS1-WSS2-WSS3-WSS4)* learned from GPR, of *N2-N3 (WSS3-WSS4-WSS5-WSS6)* learned from GPR, and of *N1-N2-N3 (WSS1-WSS2-WSS3-WSS4-WSS5-WSS6)* learned from GPR. We will showcase the performance of the proposed abstraction framework with hybrid learning techniques and will also compare the end-to-end predicted performance utilising a different combination of learning schemes.

Fig. 11 illustrates the end-to-end QoT performance prediction of the proposed multi-domain hybrid learning assisted impairments abstraction strategy with different combinations of ML techniques. The red circle markers represent the predicted SNR based on the proposed impairments abstraction scheme before the channels are provisioned and blue triangle markers are the measured performance after the new services have been set up. Specifically, Fig. 11a shows the QoT prediction of the new channels using SNR degradation factors generated by deep learning using the historical data of intra-domain services in N1, N2 and N3 for the end-to-end performance evaluation. From the figure, it is observed that all deep learning information based end-to-end abstraction achieves high accuracy with only average 0.19 dB estimation error compared to their measured performance. But domain level abstracted information from deep learning may not be available and sufficient in all circumstance, or network orchestrator may not use all the deep learning generated information for end-to-end abstraction due to its own policy or algorithm. In this case, we compare the prediction performance of using the SNR degradation factors learned from deep learning in N1 and those learned from inter-domain services between N2 and N3 using GPR, against their real QoT performance as depicted in Fig. 11b. The average estimation error is around 0.5 dB, which is significantly higher than using SNR degradation factors only formed by deep learning. The similar concept applies to the third case as illustrated in Fig. 11c with the abstraction is calculated by aggregating the SNR degradation factors learned

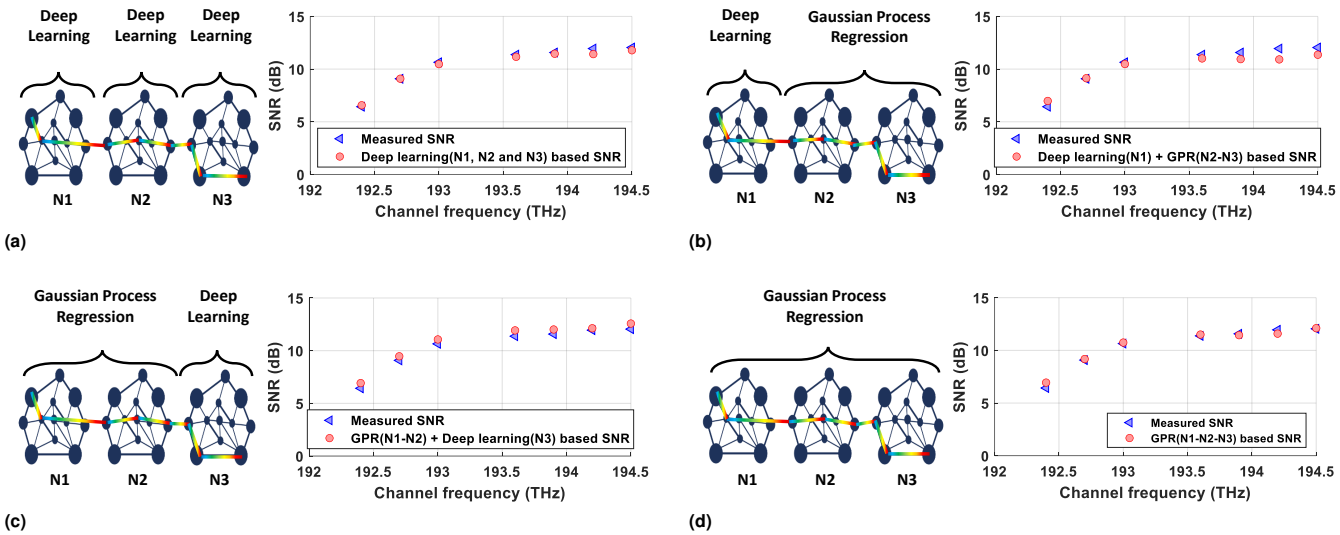

Fig. 11. End-to-end QoT performance of the proposed multi-domain hybrid-learning assisted impairments abstraction for network services across three networks with path *WSS1-WSS2-WSS3-WSS4-WSS5-WSS6*. (a) QoT prediction with SNR degradation factors learned from intra-domain services using deep learning in N1, N2 and N3; (b) QoT prediction with SNR degradation factors learned from intra-domain services using deep learning in N1 and from inter-domain services using GPR across N2-N3; (c) QoT prediction with SNR degradation factors learned from inter-domain services using GPR across N1-N2 and from intra-domain services using deep learning in N1; (d) QoT performance prediction with SNR degradation factors learned directly from inter-domain services using GPR across N1-N2-N3.

from GPR of N1-N2 and deep learning of N3. Such abstracted information combination for end-to-end performance assessment provides average 0.43 dB abstraction inaccuracy. In case of deep learning-based abstracted information unavailable, the network orchestrator can also utilise the information of 16 existing co-propagated channels across N1, N2 and N3 for GPR only learning. The performance is shown in Fig. 11d with 16 training samples thus also achieves high abstraction accuracy with 0.21 dB abstraction error. The experimental results of hybrid learning also confirm the simple rule of selecting the suitable SNR degradation factors proposed in Section 4. Considering the above four hybrid learning cases, the proposed abstraction framework is able to perform accurate impairments abstraction with an average of 0.33 dB inaccuracy without SNR degradation factors selection criteria.

7. CONCLUSION

In this paper, we first propose a multi-domain hybrid learning assisted impairments abstraction framework in optical networks for the performance evaluation of new network services. The proposed abstraction framework consists of a parametric ML technique and a non-parametric ML technique, both relying on the monitoring of network operational data and status. The dimensions, categories and variety of required data for two ML solutions are different, thus makes the abstraction framework suitable for networks with different monitoring strategies. The hybrid learning mechanism enables the abstraction of QoT performance of both intra-domain and inter-domain network services while increasing the information confidentiality of individual network. The proposed solution is investigated over a field-trial testbed consisting of three optical networks with real monitored data. The results show individual ML technique can perform accurate abstraction. Further, the experiments also verify the hybrid learning assisted abstraction by combining the results of different ML strategies to form the end-to-end abstraction. The proposed abstraction framework is proved to be an accurate tool for implementing QoT evaluation of intra-domain and inter-domain services in dynamic and cognitive optical networks.

FUNDING

This work was funded by UK EPSRC project TOUCAN EP/L020009/1 and EU funded project Metro-Haul (761727).

REFERENCES

1. R. Muñoz, R. Vilalta, R. Casellas, R. Martínez, F. Francois, M. Channegowda, A. Hammad, S. Peng, R. Nejabati, D. Simeonidou et al., "Transport network orchestration for end-to-end multilayer provisioning across heterogeneous sdn/openflow and gmpls/pce control domains," J. Light. Technol. **33**, 1540–1548 (2015).
2. N. F. S. de Sousa, D. A. L. Perez, R. V. Rosa, M. A. Santos, and C. E. Rothenberg, "Network service orchestration: A survey," Comput. Commun. (2019).
3. D. McDysan, "Software defined networking opportunities for transport," IEEE Commun. Mag. **51**, 28–31 (2013).
4. 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. **11**, C10–C24 (2019).
5. R. Munoz, R. Vilalta, R. Casellas, R. Martínez, T. Szyrkowiec, A. Autenrieth, V. López, and D. López, "Sdn/nfv orchestration for dynamic deployment of virtual sdn controllers as vnf for multi-tenant optical networks," in Optical Fiber Communication Conference, (Optical Society of America, 2015), pp. W4J–5.
6. M. Channegowda, R. Nejabati, and D. Simeonidou, "Software-defined optical networks technology and infrastructure: Enabling software-

defined optical network operations," J. Opt. Commun. Netw. **5**, A274–A282 (2013).
7. B. Guo, Y. Shang, Y. Zhang, W. Li, S. Yin, Y. Zhang, and S. Huang, "Timeslot switching-based optical bypass in data center for intrarack elephant flow with an ultrafast dpdk-enabled timeslot allocator," J. Light. Technol. **37**, 2253–2260 (2019).
8. R. Casellas, R. Martínez, R. Muñoz, R. Vilalta, L. Liu, T. Tsuritani, and I. Morita, "Control and management of flexi-grid optical networks with an integrated stateful path computation element and openflow controller," J. Opt. Commun. Netw. **5**, A57–A65 (2013).
9. Y. Lee, D. Dhody, S. Karunanithi, R. Vilalta, D. King, and D. Ceccarelli, "Yang models for actn te performance monitoring telemetry and network autonomics," IETF, TEAS Work. Group (2017).
10. M. Dallaglio, N. Sambo, F. Cugini, and P. Castoldi, "Control and management of transponders with netconf and yang," IEEE/OSA J. Opt. Commun. Netw. **9**, B43–B52 (2017).
11. "Open roadm msa device white paper for release 2.2, v1.1 08/17/18," .
12. F. Paolucci, A. Sgambelluri, F. Cugini, and P. Castoldi, "Network telemetry streaming services in sdn-based disaggregated optical networks," J. Light. Technol. **36**, 3142–3149 (2018).
13. R. Casellas, R. Martínez, R. Vilalta, and R. Muñoz, "Abstraction and control of multi-domain disaggregated optical networks with openroadm device models," J. Light. Technol. **38**, 2606–2615 (2020).
14. D. J. Ives, S. Yan, L. Galdino, D. J. Elson, F. J. Vaquero-Caballero, G. Saavedra®, R. Wang, D. Lavery, R. Nejabati, P. Bayvel, D. Simeonidou, and S. J. Savory, "A comparison of impairment abstractions by multiple users of an installed fiber infrastructure," in 2019 Optical Fiber Communications Conference and Exhibition (OFC), (2019), pp. 1–3.
15. 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. lightwave technology **32**, 694–721 (2013).
16. P. Poggiolini, "The gn model of non-linear propagation in uncompensated coherent optical systems," J. Light. Technol. **30**, 3857–3879 (2012).
17. R. Wang, S. Bidkar, F. Meng, R. Nejabati, and D. Simeonidou, "Load-aware nonlinearity estimation for elastic optical network resource optimization and management," IEEE/OSA J. Opt. Commun. Netw. **11**, 164–178 (2019).
18. I. Sartzetakis, K. Christodoulopoulos, C. P. Tsekrekos, D. Syvridis, and E. Varvarigos, "Quality of transmission estimation in wdm and elastic optical networks accounting for space–spectrum dependencies," IEEE/OSA J. Opt. Commun. Netw. **8**, 676–688 (2016).
19. N. Sambo, Y. Pointurier, F. Cugini, L. Valcarenghi, P. Castoldi, and I. Tomkos, "Lightpath establishment assisted by offline qot estimation in transparent optical networks," IEEE/OSA J. Opt. Commun. Netw. **2**, 928–937 (2010).
20. F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," J. Light. Technol. **37**, 493–516 (2019).
21. F. Meng, S. Yan, K. Nikolovgenis, Y. Ou, R. Wang, Y. Bi, E. Hugues-Salas, R. Nejabati, and D. Simeonidou, "Field trial of gaussian process learning of function-agnostic channel performance under uncertainty," in Optical Fiber Communication Conference, (Optical Society of America, 2018), pp. W4F–5.
22. A. Mahajan, K. Christodoulopoulos, R. Martínez, S. Spadaro, and R. Muñoz, "Modeling edfa gain ripple and filter penalties with machine learning for accurate qot estimation," J. Light. Technol. **38**, 2616–2629 (2020).
23. S. Yan, F. N. Khan, A. Mavromatis, D. Gkounis, Q. Fan, F. Ntavou, K. Nikolovgenis, F. Meng, E. H. Salas, C. Guo et al., "Field trial of machine-learning-assisted and sdn-based optical network planning with network-scale monitoring database," in 2017 European Conference on Optical Communication (ECOC), (IEEE, 2017), pp. 1–3.
24. Z. Gao, S. Yan, J. Zhang, M. Mascarenhas, R. Nejabati, Y. Ji, and D. Simeonidou, "Ann-based multi-channel qot-prediction over a 563.4-km field-trial testbed," J. Light. Technol. **38**, 2646–2655 (2020).
25. G. Liu, K. Zhang, X. Chen, H. Lu, J. Guo, J. Yin, R. Proietti, Z. Zhu, and S. B. Yoo, "Hierarchical learning for cognitive end-to-end service provisioning in multi-domain autonomous optical networks," J. Light. Technol. **37**, 218–225 (2019).
26. C. Rottondi, R. di Marino, M. Nava, A. Giusti, and A. Bianco, "On the benefits of domain adaptation techniques for quality of transmission estimation in optical networks," IEEE/OSA J. Opt. Commun. Netw. **13**, A34–A43 (2021).
27. R. Wang, X. Chen, Z. Gao, S. Yan, R. Nejabati, and D. Simeonidou, "Hybrid learning assisted abstraction for service performance assessment over multi-domain optical networks," in 2020 Optical Fiber Communications Conference and Exhibition (OFC), (IEEE, 2020), pp. 1–3.
28. P. Wiriyathammabhum, D. Summers-Stay, C. Fermüller, and Y. Aloimonos, "Computer vision and natural language processing: Recent approaches in multimedia and robotics," ACM Comput. Surv. **49** (2016).
29. G. Cybenko, "Approximation by superpositions of a sigmoidal function," Math. control, signals systems **2**, 303–314 (1989).
30. J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks **61**, 85–117 (2015).
31. M. Ebden, "Gaussian processes for regression: A quick introduction.[online] available at:< http://www. robots. ox. ac. uk/˜ mebden/reports," GPtutorial. pdf (2008).
32. F. Jäkel, B. Schölkopf, and F. A. Wichmann, "A tutorial on kernel methods for categorization," J. Math. Psychol. **51**, 343–358 (2007).
33. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).