# CompQoTE: Generalizing QoT Estimation with Composable ML and End-to-End Learning

**Hanyu Gao[1], Xiaoliang Chen[1], Lu Sun[3], and Zhaohui Li[1,2]**

*1. Guangdong Provincial Key Labratory of Optoelectronic Information Processing Chips and Systems, School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China*
*2. Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China*
*3. State Grid Information & Telecommunication Group Co., Ltd., Beijing, China*
*xlichen@ieee.org*

**Abstract:** This paper proposes CompQoTE, a composable QoT estimation design with end-to-end learning capability. Results show CompQoT can generalize arbitrary lightpaths while achieving $>90\%$ estimation accuracy for unseen lightpaths. 

## 1. Introduction

Quality-of-transmission (QoT) estimation is essential for assuring the correct operation of optical networks. Current optical networks mainly rely on domain-specific knowledge (i.e., analytical models) for lightpath QoT estimation, which suffers from poor estimation accuracy and thereby yields high operation margins. While the past few years witness significant advances in machine learning (ML), research attentions on QoT estimation have shifted to developing ML-based cognitive tools [1]. Thanks to the data-driven nature of ML and its powerful capability of feature extraction from high-dimensional data, ML-based QoT estimation has been shown to achieve higher accuracy [2] and been seen as a key enabler for automating and intelligentizing optical networks [3]. Nevertheless, existing works mostly build separate QoT models for lightpaths to account for heterogeneous network configurations (e.g., different numbers of links/spans traversed by signals, multi-vendor equipment deployed) [4]. This introduces nonnegligible overheads of data collection to the control plane and can lead to scalability issues. Although the application of transfer learning can ease such limitations [5], acquisition of extra data is still needed when training new QoT estimators, creating a major obstacle to actuating fast service delivery.

In this paper, we pursue the generalization of QoT estimation in optical networks by a composable QoT estimation design dubbed CompQoTE. CompQoTE derives QoT models for arbitrary lightpaths by composition of three basic modules, i.e., the *Launch*, *Propagation*, and *Readout* modules, to mimic the processing and propagation of optical signals. We apply an end-to-end learning scheme for training CompQoTE, which allows a unified tuning of the three modules with data collected from different lightpaths. Performance studies show that CompQoTE can achieve reasonably good estimation accuracy for unseen lightpaths traversing even different numbers of spans.

## 2. CompQoTE Design

Fig. 1(b) illustrates a conventional QoT estimation mechanism for four lightpaths sampled from the network in Fig. 1(a). Separate QoT models [e.g., neural networks (NNs)] are built and trained with data collected from the corresponding lightpaths. The structures of the models can differ depending on the scales of their input features. For instance, we parameterize $f_{\theta_4}(\cdot)$ by a larger NN with respect to $f_{\theta_1}(\cdot)$, where $\theta_1$ and $\theta_4$ are the sets of NN parameters. This way, the control and management system may need to maintain a large number of QoT models with the scale-up of network topologies. The problem becomes more intractable when constant model adaptions are necessary due to the evolution of network conditions.

*Architecture of CompQoTE*: CompQoTE tackles the aforementioned limitations with a composable QoT modeling shown in Fig. 1(c). It is built on three basic modules, namely, the *Launch*, *Propagation*, and *Readout* modules, each of which can be parameterized by an NN. We denote the three modules by $f_{\theta_l}(\cdot)$, $f_{\theta_p}(\cdot)$ and $f_{\theta_r}(\cdot)$, respectively. $f_{\theta_l}(\cdot)$ takes as input data featuring transceiver side configurations, such as launch power, central frequency, baud rate, and modulation format. The output of $f_{\theta_l}(\cdot)$ serves as partial input of $f_{\theta_p}(\cdot)$, representing higher-level features without an explicit physical significance. CompQoTE makes use of a concatenation of $n$ *Propagation* modules, one for every link/span. $f_{\theta_p}(\cdot)$ reads local features, such as link loads and gain profiles of the amplifiers, as additional input data, while its output feeds the module of the next stage. Finally, the output of the last $f_{\theta_p}(\cdot)$ of chain is fed to $f_{\theta_r}(\cdot)$ that produces the estimation of QoT end-to-end. The above procedures mimic the processing and propagation of optical signals. By composition of the three modules, CompQoTE is able to generalize arbitrary (and unseen) lightpaths and to remarkably reduce the number of models and parameters (i.e., only $\theta_l$, $\theta_p$, and $\theta_r$) to be maintained. In other words, CompQoTE offers a one-for-all solution. Note that, CompQoTE can augment the basic modules for more generic settings where heterogeneous technology or vendor domains coexist.

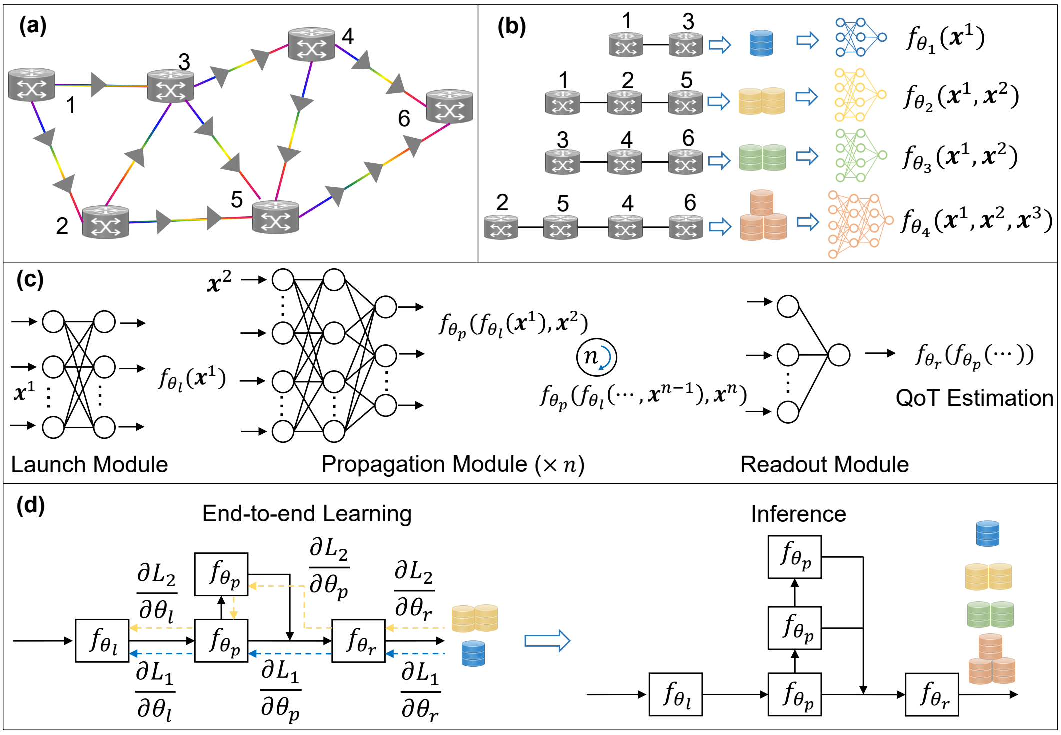



Fig. 1: Schematic of CompQoTE: (a) a six-node example topology; (b) existing QoT estimation schemes; (c) architecture of CompQoTE; and (d) end-to-end learning and inference of CompQoTE.

*Training*: we train CompQoTE with an end-to-end learning mechanism other than tuning each of the modules individually. Specifically, we aim at minimizing the overall estimation error on a compounded data set $\mathbb{D} = \{\mathbb{D}_k|_k\}$, where $\mathbb{D}_k = \{(x_{k,i}, y_{k,i})|_i\}$ represents the data from lightpath $k$. We define the loss function as follows,

$$L(\theta_l, \theta_p, \theta_r, \mathbb{D}) = \sum_k \alpha_k L_k(\theta_l, \theta_p, \theta_r, \mathbb{D}_k) = \sum_k \frac{\alpha_k}{|\mathbb{D}_k|} \sum_i (y_{k,i} - f_{[\theta_l, \theta_p, \theta_r],k}(x_{k,i}))^2, \tag{1}$$

where $\alpha_k$ is the weighting coefficient of the loss on $\mathbb{D}_k$, and $f_{[\theta_l, \theta_p, \theta_r],k}(\cdot)$ signifies the composed QoT estimation model for lightpath $k$. Any existing training algorithm can be applied, for instance, by iteratively computing and applying the gradients of the loss function regarding $\theta_l$, $\theta_p$, and $\theta_r$. While computing $\frac{\partial L_k}{\partial \theta_r}$ is straightforward, $\frac{\partial L_k}{\partial \theta_p}$ and $\frac{\partial L_k}{\partial \theta_l}$ can be obtained following the chain rule. Without loss of generality, consider a specific case where $f_{[\theta_l, \theta_p, \theta_r],k}(x) = f_{\theta_r}(f_{\theta_p}(f_{\theta_l}(x^1), x^2))$, $x = [x^1, x^2]$, we have,

$$\frac{\partial L_k}{\partial \theta_p} = \frac{1}{|\mathbb{D}_k|} \sum_i \frac{\partial L_k}{\partial f_{\theta_p}(f_{\theta_l}(x^1), x^2)} \bigg|_{f_{\theta_p}(f_{\theta_l}(x^1), x^2) = f_{\theta_p}(f_{\theta_l}(x^1_{k,i}), x^2_{k,i})} \frac{\partial f_{\theta_p}(f_{\theta_l}(x^1), x^2)}{\partial \theta_p} \bigg|_{x = x_{k,i}}, \tag{2}$$

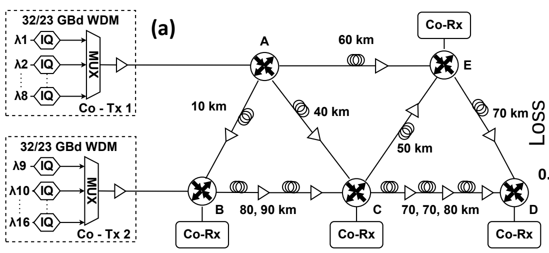
$$\frac{\partial L_k}{\partial \theta_l} = \frac{1}{|\mathbb{D}_k|} \sum_i \frac{\partial L_k}{\partial f_{\theta_l}(x^1)} \bigg|_{f_{\theta_l}(x^1) = f_{\theta_l}(x^1_{k,i})} \frac{\partial f_{\theta_l}(x^1)}{\partial \theta_l} \bigg|_{x^1 = x^1_{k,i}}. \tag{3}$$

Fig. 1(d) illustrates the principle of training CompQoTE with data collected from lightpaths 1-3 and 1-2-5. In each round of module update, gradients of $L_1$ and $L_2$ back propagate through different chains. Afterward, CompQoTE can be used to infer the QoT of lightpaths 3-4-6 and 2-5-4-6 even though they do not contribute to the training.

## 3. Performance Evaluation

We evaluated the performance of CompQoTE with data collected using the OptiSystem simulator implementing a five-node network (see Fig. 2(a)). Fig. 2(b) summarizes the parameter setup. The two I/Q modulator blocks (Co-Tx), each fed by eight lasers, adopt 16-QAM formats operated at different symbol rates (i.e., 23 Gbaud or 32 Gbaud). Nodes are connected through single-mode fibers. The length of fiber links ranges from 10 to 90 km. The fiber losses are compensated using gain-controlled amplifiers with noise figure ranging from 1 to 5 dB. We set up ten lightpaths composed one to five fiber spans. The bit-error rate (BER) of each testing lightpath is measured by a coherent receiver (Co-Rx). In Co-Rx, the sampled signals go through universal digital signal processing including dispersion compensation, constant modulus algorithm plus radius-detected equalizer, 4-th power algorithm for frequency offset estimation, and blind phase search algorithm for carrier phase estimation. We chose the total power of each Co-Tx ranging from $2.5-25$ mW to introduce nonlinear effects. In total, $16,450$ data instances were collected, with BER ranging from $10^{-7}$ to $10^{-1}$. We implemented the *Launch*, *Propagation* and *Readout* modules with neural networks of 4, 4, 3 layers, respectively. We compared CompQoTE with a baseline that trains a separate model for each lightpath and made them adopt the same model architectures for fair comparisons.

We first trained CompQoTE with 80% of the entire data set and Fig. 2(c) shows the evolution of training and validation losses. Both of the losses converge, indicating a good fitting of the model. Then, we counted the prediction errors of CompQoTE and the baseline on the testing data set and plotted the results in Fig. 2(d). Here, we divided the testing data based on the number of spans the lightpaths traverse. We can see that the two approaches achieve comparable performance, verifying the feasibility of CompQoTE. The average prediction errors are $\sim 4.5\%$, as labeled by the dashed lines. Next, we evaluated the generalization ability of CompQoTE by each time removing one lightpath from training data and testing the obtained model with data from the removed lightpath. For the baseline, we applied the model trained for one lightpath to another lightpath traversing the same number of spans. Fig. 2(e) shows the results for all the cases. It can be seen that CompQoTE is still able to predict the QoT of lightpaths not seen in the training phase with reasonably good accuracy ($> 90\%$ on average) and generalizes better than the baseline. Finally, we masked lightpaths with the same number of spans from training, which means that the training and testing data for CompQoTE involve lightpaths with different numbers of spans. In this scenario, the baseline does not apply because its model architectures differ when the length of lightpaths varies. Fig. 2(f) shows the comparison between the results from this scenario and those in Fig. 2(d) (yellow bars). We can observe that CompQoTE can still achieve $\sim 7.56\%$ prediction error on average, even with data from lightpaths with the number of spans completely missing from training, which further validate the generalization ability of CompQoTE.




(b)

| A | Path | Number of spans | Fiber Length /km | EDFA Gain /dB | EDFA NF /dB | Launch Power /W | BER |
|---|---|---|---|---|---|---|---|
| 1 | A-B | 1 | 10 | 2 | 1 | | |
| 2 | A-B-C | 3 | 180 | 2-18 | 1-5 | | |
| 3 | A-C | 1 | 40 | 8 | 2 | | |
| 4 | A-C-D | 4 | 260 | 8-16 | 2-5 | | |
| 5 | A-E | 1 | 60 | 12 | 3 | 2.5e-3-2.5e-2 | $10^{-7}$-$10^{-1}$ |
| 6 | A-E-D | 2 | 130 | 12-14 | 3-4 | | |
| 7 | B-C | 2 | 170 | 16-18 | 4-5 | | |
| 8 | B-C-D | 5 | 390 | 14-18 | 4-5 | | |
| 9 | B-C-E | 3 | 220 | 10-18 | 3-5 | | |
| 10 | B-C-E-D | 4 | 290 | 10-18 | 3-5 | | |

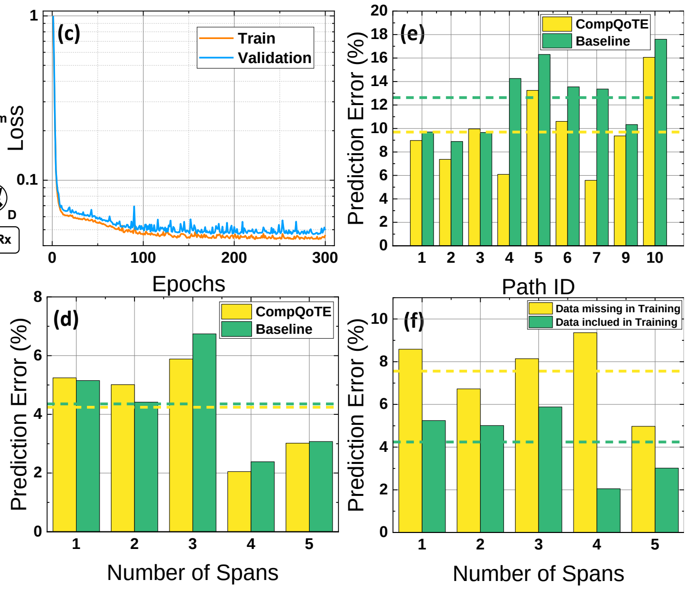



Fig. 2: (a) Five-node network topology; (b) parameter setup in the simulations; (c) training process of CompQoTE; (d)-(e) comparison between CompQoT and the baseline: (d) with random data partitioning (0.8 : 0.2) for training and testing, and (e) when applied to estimating the QoT of lightpaths missing from training; (f) performance of CompQoTE with lightpaths of the same number of spans masked from training.

## 4. Conclusion

This paper proposes a composable QoT estimation design with end-to-end learning dubbed CompQoTE. Results demonstrate excellent generalization ability of CompQoTE.

## References

1. F. Khan *et al.*, *J. Lightwave Technol.* 37, 493–516 (2019).
2. L. Barletta *et al.*, *Proc. of Opt. Fiber Commun. Conf.*, paper Th1J.1 (2017).
3. D. Rafique and L. Velasco, *J. Opt.https://www.overleaf.com/project/6342c3cb0089c40d260179d2 Commun. Netw.* 10, D126–D143 (2018).
4. G. Liu *et al.*, *J. Lightwave Technol.* 37, 218–225 (2019).
5. J. Yu *et al.*, *J. Opt. Commun. Netw.* 11, C48–C57 (2019).

*This work was supported in part by NSFC under Project 62201627.*