POLITECNICO DI TORINO
Repository ISTITUZIONALE

Experimental Demonstration of Linear Inter-Channel Interference Estimation Based on Neural Networks

Original
Experimental Demonstration of Linear Inter-Channel Interference Estimation Based on Neural Networks / Hraghi, A., Minelli, L., Nespola, A., Piciaccia, S., Bosco, G.. - In: IEEE PHOTONICS JOURNAL. - ISSN 1943-0655. - ELETTRONICO. - 15:2(2023), pp. 1-6. [10.1109/JPHOT.2023.3259009]

Availability:
This version is available at: 11583/2977568 since: 2023-03-29T08:56:40Z

Publisher:
IEEE

Published
DOI:10.1109/JPHOT.2023.3259009

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

11 September 2026

Experimental Demonstration of Linear Inter-Channel Interference Estimation Based on Neural Networks

A. Hraghi, L. Minelli, A. Nespola, S. Piciaccia, and G. Bosco, *Fellow, IEEE*

***Abstract*—In this paper, an algorithm for the estimation of the linear inter-channel crosstalk in a dense-WDM polarization-multiplexed 16-QAM transmission scenario is proposed and demonstrated. The algorithm is based on the use of a feed-forward neural network (FFNN) inside the coherent digital receiver. Two types of FFNNs were considered, the first based on a regression algorithm and the second based on a classification algorithm. Both FFNN algorithms are applied to features extracted from the histograms of the in-phase and quadrature components of the equalized digital samples. After a simulative investigation, the performance of the channel spacing estimation algorithms was experimentally validated in a 3 × 52 Gbaud 16-QAM WDM system scenario.**

***Index Terms*—Coherent optical systems, optical performance monitoring (OPM), artificial neural networks.**

I. Introduction

TODAY, the exponential growth in Internet traffic demands is due to the rapid proliferation of new emerging applications (Internet of Things (IoT), live games video, smart TV,...) [1]. This tremendous growth requires a significant upgrade in the data rate that should be guaranteed by optical networks. The Nyquist-WDM approach [2] was proposed to provide high network capacity and flexibility [3]. In ideal Nyquist-WDM systems, optical sub-channels with rectangular spectra are tightly packed at a channel spacing equal to the symbol rate, maximizing the transmitted data rate while avoiding the linear inter-channel interference (ICI) among adjacent sub-channels. However, in practical dense-WDM (DWDM) systems, some amount of linear ICI is present, due to a non-ideal spectral shaping of the transmitted pulses (with square-root cosine shape and roll-off values between 0.15 and 0.3), a non-perfect control of the channel spacing, and the presence of a frequency offset between the local oscillator and the transmitter laser [4]. Several solutions have been proposed to mitigate the linear ICI effect arising from these impairments [5], [6], [7] using digital signal processing techniques at the receiver side. Another approach consists in minimizing the impact of linear ICI by adapting the characteristics of the transmitted signal (modulation format, symbol rate, spectral shaping, etc.) to the conditions of the link by monitoring the frequency spacing among adjacent channels [8], [9], [10], a parameter that is strictly linked to the linear ICI. Machine learning (ML)-based methods have recently received considerable attention for linear ICI and channel spacing estimation. Among the most frequently used ML algorithms are k-nearest neighbors (KNN) [11], density-based spatial clustering of applications with noise (DBSCAN) [12], and support vector machine (SVM) [13]. However, while these methods have been employed to evaluate the amount of ICI, they are limited to providing qualitative assessments [12] or treating the estimation problem as a mere classification task [11], [13].

In this paper, we propose an alternative method, based on Feed Forward Neural Networks (FFNNs) applied to features extracted from the histograms of the in-phase and quadrature components of the equalized samples in a coherent receiver, demonstrating that it is a valid technique for monitoring the linear ICI in a DWDM optical communication system through the estimation of the frequency spacing between optical sub-channels. Two types of FFNNs were considered: a Cross-Entropy Loss (CEL) -based Neural Network, which uses a classification algorithm, and a Mean Square Error (MSE)-based Neural Network, which employs a regression algorithm. The regression-based FFNN slightly outperforms the classification-based FFNN, achieving a channel spacing estimation Root Mean Square Error equal to 0.32 GHz in simulations and 1.26 GHz in the experiments, in a 3x52Gbaud polarization-multiplexed (PM) 16-QAM scenario.

The paper is organized as follows. Section II introduces and describes the methodology, while Section III describes the simulation and the experimental setup. The obtained simulation results are then shown in Section IV, while the experimental results are described in Section V, comparing them with the simulations of the previous section. Finally, Section VI draws the conclusions.

II. Methodology

The proposed method is based on the generation of histograms using the in-phase (I) and quadrature (Q) signal components of 16-QAM modulated symbols after equalization in a coherent

Manuscript received 29 November 2022; revised 27 February 2023; accepted 15 March 2023. Date of publication 20 March 2023; date of current version 29 March 2023. This work was supported in part by the European Union through the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART") and in part by a CISCO Sponsored Research Agreement (SRA). *(Corresponding author: L. Minelli.)*

A. Hraghi was with Politecnico di Torino, 10129 Torino, Italy. She is now with Huawei Technologies France, Optical Communication Technology Lab, 92100 Boulogne-Billancourt, France (e-mail: abir.hraghi@polito.it).

L. Minelli and G. Bosco are with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy (e-mail: leonardo.minelli@polito.it; gabriella.bosco@polito.it).

A. Nespola is with the PhotonLab, LINKS, 10138 Torino, Italy (e-mail: antonino.nespola@linksfoundation.com).

S. Piciaccia is with the CISCO Photonics, 20871 Vimercate, Italy (e-mail: spiciacc@cisco.com).

Digital Object Identifier 10.1109/JPHOT.2023.3259009

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

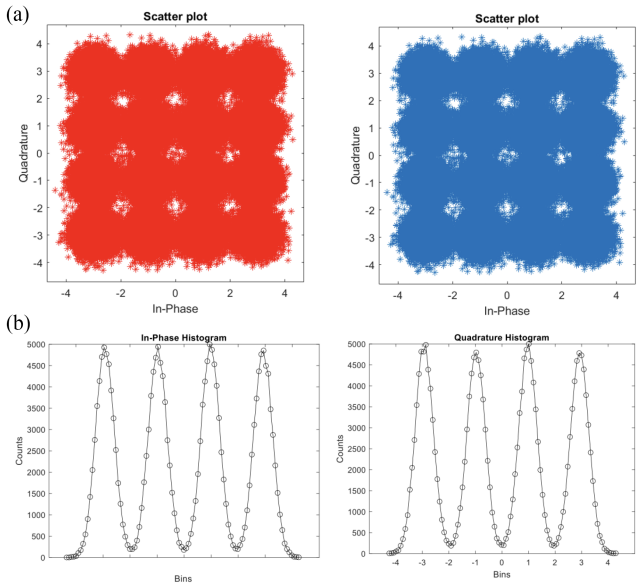

Fig. 1. a) Example of 16-QAM scattering diagrams for the two polarizations, with a channel spacing equal to 53.6 GHz at 52 Gbaud, with an OSNR equal to 20 dB. b) Corresponding I and Q histograms, built using 2^{15} symbols.

receiver, as in [11], [12], [13]. As shown in Fig. 1, four local maxima and three local minima can be indeed extracted from both the I and Q histograms. The resulting 14 features are concatenated in a vector, which is given as input to an FFNN. By denoting the input features as the vector $\mathbf{y_0}$, the output of a L-layered FFNN is the vector $\mathbf{y_L} = f_{FFNN}(\mathbf{y_0})$, that is obtained through the following recursive computation:

$$\mathbf{y_l} = a_l\left(\mathbf{W_l} \cdot \mathbf{y_{l-1}} + \mathbf{b_l}\right), \quad l = 1, \ldots L \tag{1}$$

where, for each l-th layer, $\mathbf{y_l}$ is the output vector, $\mathbf{W_l}$ and $\mathbf{b_l}$ are the weight matrix and the bias vector (i.e. the learnable parameters of structure), and a_l is the activation function.

This nonlinear mathematical model, whose coefficients are optimized through gradient-descent-based algorithms, can be used as a universal function approximator [14], able to solve nonlinear regression problems as well as classification tasks. In this paper, two alternative criteria are used to train the FFNN:

- **Mean Square Error (MSE) criterion (regression-based FFNN)**: the FFNN, having a linear output activation function and a scalar output, estimates the ICI in terms of sub-channel frequency spacing. In this situation, the network is trained to minimize the difference between the sub-channel spacing Δf, relative to given inputs, and the model predicted output $\hat{\Delta f}$, thus solving a nonlinear regression problem.
- **Cross-Entropy Loss (CEL) criterion (classification-based FFNN)**: Assigning N classes to N different values of frequency spacing, the FFNN exploits a $Softmax()$ output activation function to produce an N-dimensional output vector whose entries are an estimation of the probability that a given input belongs to a certain class [14]. In this situation, the FFNN is trained to minimize the cross-entropy between the distribution of the input samples and the distribution estimated by the model, translating the problem into a classification task.

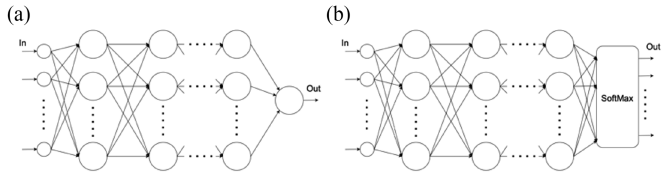

Fig. 2. Adopted FFNN structures a) MSE-based structure, terminated by a linear activation output; b) CEL-based structure, terminated by a Softmax() activation output.

The FFNN structures related to the MSE and CEL criteria are illustrated in Fig. 2

The proposed algorithms are able to estimate the amount of linear ICI that falls in the channel under test. Each amount of linear ICI corresponds to a single value of channel spacing, under the assumption that the two neighboring channels are equally spaced. For simplicity, and for a direct comparison with Refs. [11], [12], [13], the results in the following sections are expressed in terms of a single value of channel spacing Δf, thus assuming that the frequency spacing is the same between the channel under study and its two neighbors. Note that, if the neighboring channels are not equally spaced, the two values of channel spacing can still be estimated, for instance by complementing the ML algorithm results with spectral estimation methods, such as those described in [8], [9], able to identify the ratio between the linear ICI induced by the left and right channel, respectively.

The FFNN performances are evaluated by computing the Root Mean Square Error (RMSE) between the predictions and the actual outputs on a test dataset, as follows:

$$RMSE = \sqrt{\mathbb{E}\left[\left|\Delta f - \hat{\Delta f}\right|^2\right]} \tag{2}$$

To make the results achieved by the two approaches comparable, when testing the trained CEL-based FFNN a scalar product is performed between its SoftMax output $\mathbf{s} = [s_1, s_2 \ldots, s_N]^\top$ and the vector $\mathbf{\Delta f} = [\Delta f_1, \Delta f_2 \ldots, \Delta f_N]^\top$, where Δf_i is the frequency spacing related to the i-th class among the N selected. The resulting scalar output $\mathbf{o} = \mathbf{s}^\top \cdot \mathbf{\Delta f}$ becomes thus equivalent to the MSE-based FFNN output, and the RMSE can be computed as well. Using both criteria, we trained the FFNN for 30 epochs over the simulation training dataset, using an Adam optimizer with a learning rate $\epsilon = 0.001$, momentum $\alpha = 0.9$ and decay rate $\rho = 0.999$. In the hidden layers, the $ReLU()$ activation function has been used, since it is the most commonly adopted in Deep Learning Literature [14]. The remaining FFNN hyper-parameters are discussed in Section IV.

III. Experimental and Simulation Setup

To evaluate the performance of the proposed method in a flexible-grid system, a 52-Gbaud 16-QAM channel (the channel under test) and two similar interfering channels were generated with three CISCO-NCS1004 line cards. All 16 QAM transmitted

symbols were shaped by a Root Raised Cosine (RRC) filter with a roll-off factor equal to 0.25. The frequency spacing between channels varied from a minimum of 52 GHz (equivalent to the symbol rate) to a maximum of 61 GHz (at which the effect of linear ICI becomes negligible compared to the ASE noise). The precise and continuous grid-less tunability supported by the card made the frequency detuning effect negligible and allowed us to monitor the cross-talk only. Moreover, a large bandwidth filter (125 GHz) was used before the receive port of the line cards to avoid any filtering issues. Amplified Spontaneous Emission (ASE) noise was added at the receiver side to test the channel spacing monitoring algorithm at an initial signal-to-noise ratio (OSNR) equal to 17 dB for creating the training and test dataset, then varied up to 19 dB to assess the performance of the algorithm in presence of different OSNRs. The line cards allowed the downloading of symbols after equalization and before the decision.

A twin copy of the experimental setup was also simulated with a training and a validation dataset at an OSNR equal to 16 dB (corresponding to the same back-to-back performance achieved in the experimental set-up), in order to find the best FFNN structures to be exploited for the experiments. The IQ histograms have been generated with a resolution of 100 bins, using sequences of 2^{18} symbols. In the simulations, the training dataset consisted of 12440 labeled sample sequences (uniformly distributed among the 9 training sub-channel spacings), while the validation dataset consisted of 7000 labeled sample sequences (uniformly distributed as well). The IQ histograms, from which the 14 features were extracted, were generated using sequences originating from different Monte-Carlo simulations. In the experimental analysis, both the training and test dataset consisted of 4416 labeled sample sequences (uniformly distributed among the 7 sub-channel spacings). In this scenario, the sequences with the same sub-channel frequency spacing originated from a unique long experimental acquisition. Therefore, to extend the number of available sequences, IQ histograms were computed over partially-overlapped observation windows: this technique, which takes inspiration to window slicing data-augmentation [15], was adopted in order to extend the number of available training and test data. In the experiments, from 7 sequences of 1228800 symbols with fixed sub-channel spacing, 552 samples were retrieved using observation windows of 122880 symbols and a window shift size of 2000 symbols. In both experimental and simulation setups, a non-uniform distribution of the frequency spacing was used when generating the FFNN training sequences to take into account the non-linear dependence between the frequency spacing and the SNR penalty, as shown in Fig. 3.

The above figure clearly shows that higher channel spacings (i.e., from 58 GHz to 62 GHz) tend to not affect significantly the SNR with respect to lower ones: therefore, the related histograms tend to be more difficult to be discriminated.

According to this, the following values of frequency spacing were used in the simulation and experiment:

- Training dataset (simulations): [52 52.8 53.6 54.4 55.2 56 57 58.5 61] GHz
- Validation dataset (simulations): from 52 GHz to 58 GHz, with a step of 0.1 GHz
- Training dataset (experiments): [52 53 54 55 56 58.5 61] GHz
- Test dataset (experiments): [52.5 53.5 54.5 55.5 56.5 58 59] GHz

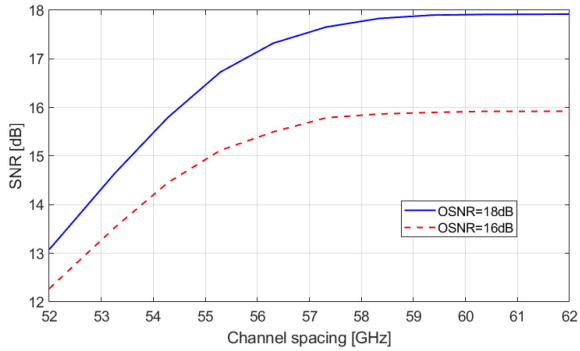

Fig. 3. SNR value, evaluated after the equalizer, vs. channel spacing for two different values of OSNR.

Moreover, in order to prevent overfitting during simulations, the related datasets were generated separately using different instances of 16-QAM symbols sequences and ASE noise: training and validation sequences having the same sub-channel spacing are thus uncorrelated between them.

IV. Simulation Results

In this section, we report the results achieved using the simulation setup, which we exploited to find the best FFNN architecture to be adopted for perform the experiments using real data. We leveraged the simulation training dataset to optimize the FFNN coefficients using Stochastic Gradient Descent optimization, while we used the validation dataset to evaluate the performance of the trained architecture (in order to prevent overfitting issues) and tune the model hyper-parameters. To accomplish this task, we performed a grid-search analysis, varying the number of hidden layers and the per-layer number of hidden units (i.e., the depth and the width of the FFNN). For each depth-width pair, we trained and validated the FFNN using different mini-batch sizes ($B = [1, 5, 25, 50, 100, 200]$), each one for 10 different runs, averaging then the RMSE results and selecting those with the best performances (i.e., lowest RMSE). The performance over the training and validation dataset are reported as contour line plots in Fig. 4

As it can be observed from the figures, configurations having lower RMSE on the training dataset do not necessarily have a good performance also in the validation dataset (thus indicating the possible occurrence of overfitting). This is especially evident in the case of the CEL-based FFNN: as the complexity (i.e., FFNN depth and width) increases, the validation RMSE (Fig. 4(c)) increases while the training RMSE (Fig. 4(a)) slightly decreases. Therefore, the best chosen architectures have respectively 1 hidden layer with 42 hidden neurons in the CEL-based FFNN case (RMSE = 0.36 GHz) and 2 hidden layers with

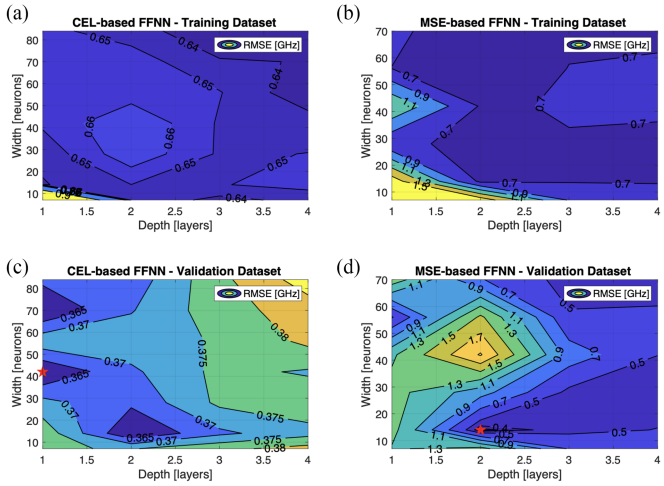

Fig. 4. Root Mean Square Error evolution over the grid-search space training (a,b) and validation (c,d) simulation datasets, related to the CEL-based FFNN (left) and MSE-based FFNN (right). Best selected FFNN architecture are marked down by a red star.

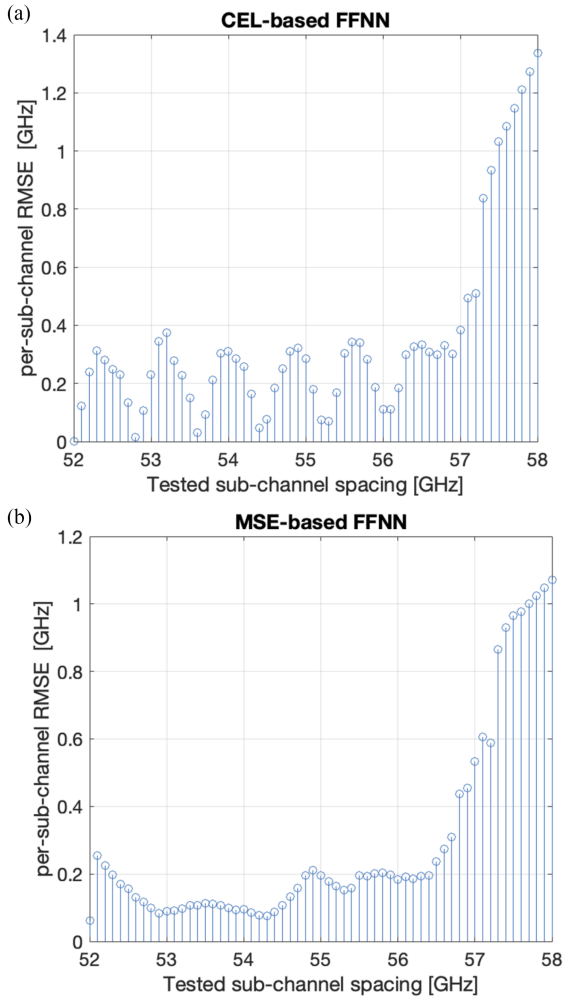

Fig. 5. Root Mean Square Error per tested sub-channel spacing, using simulation test dataset, using CEL-based FFNN (a) and MSE-based FFNN (b).

14 neurons each in the MSE-based FFNN case (RMSE = 0.32 GHz). It should be noted that the CEL-based FFNN, while achieving a slightly higher RMSE at best than the MSE-based FFNN, exhibits more stable performance as the structure varies (i.e. the RMSE never exceeds 0.45 GHz). In contrast, MSE-based FFNN performance varies strongly across the depth-width search space (with a maximum validation RMSE = 2 GHz). To better investigate the performance of the 2 best CEL-based and MSE-based FFNN architectures, the RMSE relative to each tested sub-channel spacing in the validation dataset has been then evaluated and reported in Fig. 5(a) and (b).

As expected by the SNR vs Channel Spacing analysis illustrated in Fig. 3, in both cases the RMSE tends to be higher when the overlap between the channels is small (i.e., sub-channel spacing higher than 57 GHz), as a consequence of the lower impact of the ICI on the IQ histograms. The extracted features indeed do not contain enough information for an accurate estimation of the channel spacing. Moreover, it can be observed that the error dependency on the spacing is smoother in case of the regression-based FFNN, as shown in Fig. 5(b), while the classification-based FFNN tends to predict the output as one of its main classes, as shown in Fig. 5(a).

As a further confirmation of the previous observations, Fig. 6 shows the simulation results in terms of the joint probability density function between the predicted and tested sub-channel spacing values for both the CEL-based and MSE-based FFNN. Consistently to Fig. 5(a), the CEL-based FFNN estimations in Fig. 6(a) tend to be distributed in a stepped fashion, according to the sub-channel frequency training classes. Moreover, as the tested sub-channel spacing is higher, the distribution points are more scattered in both the MSE and CEL cases.

In conclusion, the simulation results show that the FFNNs are able to effectively estimate the sub-channel spacing, with the MSE-based FFNN slightly outperforming the CEL-based one. In terms of computational complexity, both cases have a similar number of learnable parameters (i.e., weights and biases): 673 coefficients for the MSE-based FFNN, and 555 coefficients for the CEL-based FFNN. The optimal performance achieved by both FFNN architectures with a limited number of parameters (i.e., less than 1000) suggests that using Artificial Neural Networks for ICI estimation would not eventually impose an excessive computational burden in practical scenarios. Moreover, for both algorithms, the accuracy of the estimation tends to be lower in the region where the tested channel spacing is larger, i.e., where the impact of ICI on the system performance is less relevant and thus an accurate estimation of the amount of overlap is less important.

V. Experimental Results

This section shows the experimental validation of the channel spacing estimation algorithms, in a similar setup to that used in simulations (see Section II for details). In particular, only the results obtained using the regression-based FFNN are shown in the following, since it gives better performance than the

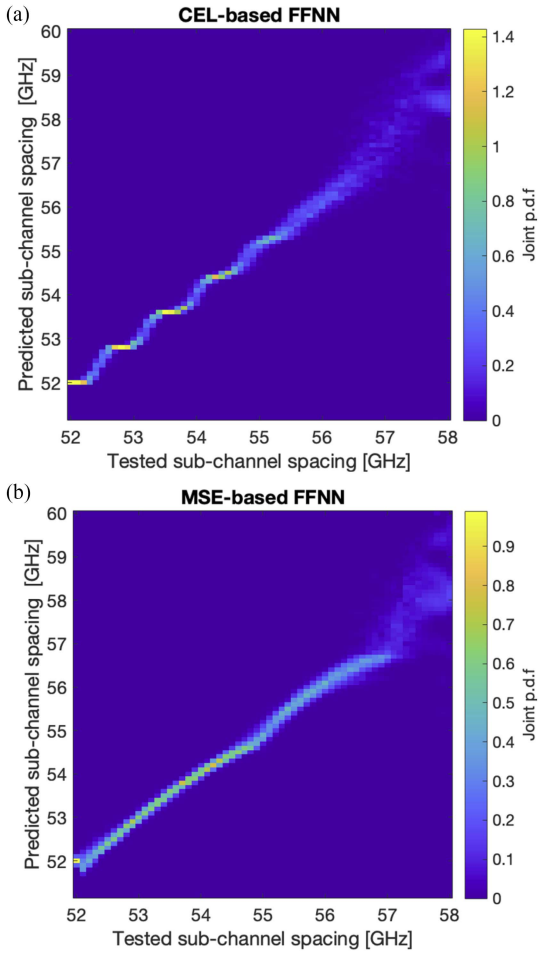

Fig. 6. Estimated joint probability density function of predicted and tested sub-channel spacing on the simulation test dataset, using CEL-based FFNN (left) and MSE-based FFNN (right).

TABLE I
PERFORMANCES OF THE EXPERIMENTAL TEST DATASET

Sub-channel spacing [GHz]	52.5	53.5	54.5	55.5	56.5	58	59
RMSE [GHz]	1.1	1.1	0.8	1.0	1.0	1.6	1.7

TABLE II
PERFORMANCES OF THE SIMULATION VALIDATION DATASET

Sub-channel spacing [GHz]	52.5	53.5	54.5	55.5	56.5	58
RMSE [GHz]	0.16	0.11	0.11	0.19	0.24	1.08

classification-based FFNN, but similar results were achieved with the classification-based FFNN.

The obtained results in terms of RMSE per each sub-channel spacing in the experimental test dataset are shown in Table I. Moreover, to make a comparison with the results illustrated in Section IV, Table II shows the RMSE performance on the simulation validation dataset testing the same sub-channel spacings as in the experiments.

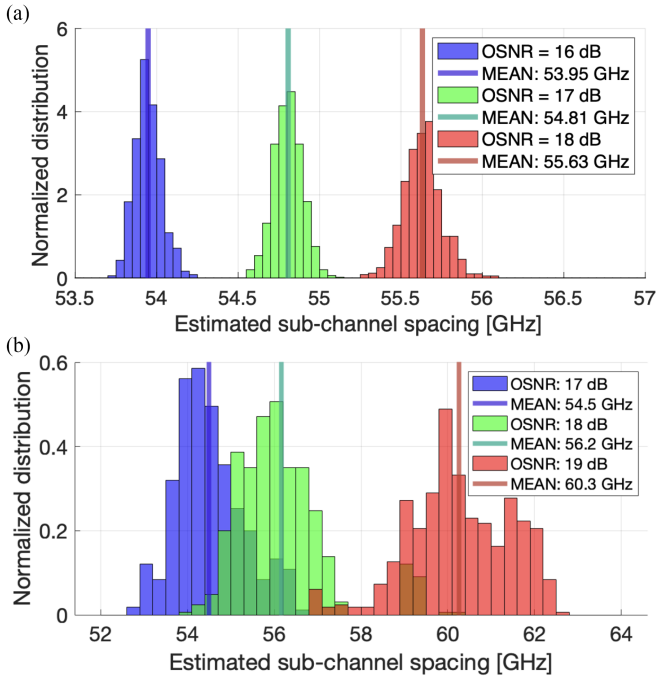

Fig. 7. MSE-based FFNN estimation of the sub-channel spacing from a tested 54 GHZ spacing measured using different OSNR (a) in simulation (b) in experiments. The FFNN has been trained with samples at OSNR=16 dB during the simulations and at OSNR=17 dB during experiments (using training datasets not including the 54 GHz sub-channel spacing).

While using the simulated dataset the overall RMSE is 0.32 GHz using the MSE-based FFNN, in the experiments the same Neural Network brings to an overall RMSE equal to 1.26 GHz. Furthermore, as observed in simulation, the RMSE relative to higher sub-channel frequency spacings (58-59 GHz) is higher with respect to the more overlapped scenarios. In case where ICI is more relevant, the RMSE is around 1 GHz, proving that the algorithm is able to effectively estimate the sub-channel spacing in a real experimental setup.

Finally, we assessed the impact of a non-perfect match between the OSNR value used in training and testing. In this analysis, the FFNN initially trained at a fixed OSNR=17 dB (16 dB in the simulation setup), has been tested using sequences generated at three different OSNR values, with only the first value matched to the OSNR used in the training phase:

- OSNR = [17,18,19] dB in the experimental setup
- OSNR = [16,17,18] dB in the simulation setup

As an example, we report here the results obtained processing a test signal with frequency spacing equal to 54 GHz, an intermediate value among those previously tested. Similar results, bringing to the same conclusions, are obtained for other values of sub-channel spacing. Fig. 7 shows the results in terms of distribution of the estimated frequency spacing in all the considered OSNRs for both simulation and experimental setup.

It's clear that, when the OSNR at which the FFNN is tested is different from the one used during training, a bias error in the estimation is present. If the OSNR values are known, this estimation error can be easily estimated and compensated for. However, if the difference in the OSNR values is high (e.g. 2 dB, in Fig. 7) the variance of the distribution increases and the RMSE will thus be higher than in the case in which the OSNR values in training and testing are similar.

VI. Conclusion

The FFNN-based estimator proposed and analyzed in this paper is able to estimate the frequency spacing between adjacent channels with an RMSE around 0.32 GHz in simulations and around 1.26 GHz in experiments. It was also shown that the estimation error is lower when the overlap between channels is higher, i.e. when the impact of ICI on the system performance is more relevant and thus an accurate estimation of the amount of overlap is more important. Moreover, our FFNN-based method allows for a much finer estimation of frequency spacing compared to other classification-based ML algorithms proposed in the literature [11], [13], as their maximum frequency resolution is limited to the spacing step among the adopted training classes.

In conclusion, while the use of Feed-Forward Neural Networks presents promising results, there is still room for further optimization. For example, 1-D Convolution Layers could be integrated before the FFNN to improve the feature extraction from the IQ histograms (e.g., going beyond the search for local extreme points). Nevertheless, the results presented in this paper strongly support the efficacy of Artificial Neural Networks as a valuable ICI monitoring tool for optimizing future grid-less optical networks. Further research in this direction is warranted to fully exploit the potential of this approach.

Acknowledgment

The authors would like to acknowledge the PhotoNext initiative at Politecnico di Torino (http://www.photonext.polito.it/) and its laboratory, where all experiments have been performed.

References

[1] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Opt. Exp.*, vol. 26, no. 18, pp. 24190–24239, Sep. 2018.

[2] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53–61, Jan. 2011.

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

[4] Z. Xiao, S. Fu, S. Yao, M. Tang, P. Shum, and D. Liu, "ICI mitigation for dual-carrier superchannel transmission based on m-PSK and m-QAM formats," *J. Lightw. Technol.*, vol. 34, no. 23, pp. 5526–5533, Dec. 2016.

[5] J. Pan, C. Liu, T. Detwiler, A. J. Stark, Y.-T. Hsueh, and S. E. Ralph, "Inter-channel crosstalk cancellation for nyquist-wdm superchannel applications," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3993–3999, Dec. 2012.

[6] K. Shibahara, A. Masuda, S. Kawai, and M. Fukutoku, "Multi-stage successive interference cancellation for spectrally-efficient super-nyquist transmission," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.

[7] C. Liu et al., "Joint digital signal processing for superchannel coherent optical communication systems," *Opt. Exp.*, vol. 21, no. 7, pp. 8342–8356, 2013.

[8] Y. Zhao et al., "Frequency domain DSP based channel spacing monitor in denser Nyquist-WDM system," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.

[9] B. Shariati, M. Ruiz, F. Fresi, A. Sgambelluri, F. Cugini, and L. Velasco, "Real-time optical spectrum monitoring in filterless optical metro networks," *Photon. Netw. Commun.*, vol. 13, no. 1, pp. 1–13, 2020.

[10] Y. Zhao et al., "Channel spacing monitor based on periodic training sequence in DWDM system," *J. Lightw. Technol.*, vol. 35, no. 8, pp. 1422–1428, Apr. 2017.

[11] A. E. Pérez, N. G. González, and J. J. G. Torres, "Inter-Channel interference estimation based on IQ histograms including machine learning," in *Proc. Signal Process. Photon. Commun*, 2020, Paper SpTh2I.6.

[12] J. J. G. Torres and N. G. González, "Overlapping estimation based on DBSCAN algorithm in Nyquist-WDM systems," in *Proc. Signal Process. Photon. Commun*, 2019, Paper SpTh1E.2.

[13] A. E. Pérez, O. D. V. Bonett, S. E. Ralph, and J. J. G. Torres, "Spectral spacing estimation in gridless Nyquist-WDM systems using local binary patterns," in *Proc. IEEE Photon. Conf.*, 2021, pp. 1–2.

[14] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[15] B. K. Iwana and S. Uchida, "An empirical survey of data augmentation for time series classification with neural networks," *PLoS One*, vol. 16, no. 7, 2021, Art. no. e0254841.

Open Access funding provided by 'Politecnico di Torino' within the CRUI CARE Agreement