MDG and SNR Estimation in SDM Transmission Based on Artificial Neural Networks

MDG and SNR Estimation in SDM Transmission Based on Artificial Neural Networks

Ruby S. B. Ospina, Menno van den Hout, *Student Member, IEEE*, Sjoerd van der Heide, *Student Member, IEEE*, John van Weerdenburg, *Student Member, IEEE*, Roland Ryf, *Fellow, IEEE*, Nicolas K. Fontaine, *Senior Member, IEEE*, Haoshuo Chen, Rodrigo Amezcua-Correa, *Member, IEEE*, Chigo Okonkwo, *Senior Member, IEEE*, and Darli A. A. Mello, *Member, IEEE*

***Abstract*—The increase in capacity provided by coupled space division multiplexing (SDM) systems is fundamentally limited by mode-dependent gain (MDG) and amplified spontaneous emission (ASE) noise. Therefore, monitoring MDG and optical signal-to-noise ratio (SNR) is essential for accurate performance evaluation and troubleshooting. Recent works show that the conventional MDG estimation method based on the transfer matrix of multiple-input multiple-output (MIMO) equalizers optimizing the minimum mean square error (MMSE) underestimates the actual value at low SNRs. Besides, estimating the optical SNR itself is not a trivial task in SDM systems, as MDG strongly influences the electrical SNR after the equalizer. In a recent work we propose an MDG and SNR estimation method using artificial neural networks (ANNs). The proposed ANN-based method processes features extracted at the receiver after digital signal processing (DSP). In this paper, we discuss the ANN-based method in detail, and validate it in an experimental 73-km 3-mode transmission link with controlled MDG and SNR. After validation, we apply the method in a case study consisting of an experimental long-haul 6-mode link. The results show that the ANN estimates both MDG and SNR with high accuracy, outperforming conventional methods.**

***Index Terms*—Mode-dependent gain, mode-dependent loss, optical fiber communications, space division multiplexing.**

Manuscript received 8 November 2021; revised 2 February 2022 and 31 March 2022; accepted 9 May 2022. Date of publication 12 May 2022; date of current version 2 August 2022. This work was supported in part by FAPESP under Grants 2018/25414-6, 2018/14026-5, 2017/25537-8, 2015/24341-7, and 2015/ 24517-8, and in part by the TU/e-KPN Smart Two Project. This article was presented in part at the Optical Fiber Communications Conference, San Diego, CA, USA, June. 2021 [DOI: 10.1364/ofc.2021.th1a.20]. *(Corresponding author: Ruby S. B. Ospina.)*

Ruby S. B. Ospina and Darli A. A. Mello are with the School of Electrical and Computer Engineering, State University of Campinas, Campinas 13083-970, Brazil (e-mail: ruby@decom.fee.unicamp.br; darli@unicamp.br).

Menno van den Hout, Sjoerd van der Heide, and Chigo Okonkwo are with the High Capacity Optical Transmission Laboratory, Electro-Optical Communications Group, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: m.v.d.hout@tue.nl; s.p.v.d.heide@tue.nl; cokonkwo@tue.nl).

John van Weerdenburg was with the Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. He is now with the Infinera, San Jose, CA 95119 USA (e-mail: J.J.A.v.weerdenburg@tue.nl).

Roland Ryf, Nicolas K. Fontaine, and Haoshuo Chen are with Nokia-Bell Labs, Holmdel, NJ 07733 USA (e-mail: roland.ryf@nokia.com; nicolas.fontaine@nokia-bell-labs.com; haoshuo.chen@nokia-bell-labs.com).

Rodrigo Amezcua-Correa is with CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL 32816 USA (e-mail: r.amezcua@creol.ucf.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2022.3174778.

Digital Object Identifier 10.1109/JLT.2022.3174778

I. Introduction

Space division multiplexing (SDM) with coupled channels is a promising solution to scale the fiber capacity in future optical system generations. Coupled SDM has been effectively demonstrated in laboratory experiments over coupled-core multi-core fibers (MCFs) [1], multi-mode fibers (MMFs) [2], [3], few-mode fibers (FMFs) [4], [5], and few-mode multi-core fibers (FM-MCFs) [6]. Among the impairments that affect coupled SDM transmission, the interaction of additive noise and mode-dependent gain (MDG)[1] fundamentally limits the system capacity. The random power variations of guided modes induced by MDG turn the channel capacity into a random variable, reducing the average capacity and generating outages [7]–[9]. Therefore, assessing the accumulated MDG and the optical signal-to-noise ratio (SNR) at the receiver is essential for performance evaluation and troubleshooting.

In the recent literature, MDG has been estimated by digital signal processing (DSP) using the transfer function of multiple-input multiple-output (MIMO) [4], [5], [10], [11] equalizers. However, we show in [12] that, as adaptive MIMO equalizers typically use the minimum mean square error (MMSE) criterion [13], the DSP-based estimation accuracy is affected by noise. We show that the accumulated MDG is underestimated for high levels of MDG and low optical SNRs[2]. To circumvent this limitation, we propose in [14] a correction factor to partially compensate for MDG estimation errors. The validity of the correction factor is verified experimentally in [14], [15]. One drawback of the correction factor is that it requires a known optical SNR that may not be readily available.

Estimating the optical SNR in coupled SDM receivers is also not trivial. In single-mode transmission, polarization-dependent gain (PDG) is not a limiting effect, and the optical SNR can be estimated from the electrical SNR by a simple direct equation [16]. In coupled SDM transmission, however, the electrical SNR may be strongly affected by MDG. In this case, estimating the optical SNR directly from the electrical SNR would underestimate the actual value.

[1]The results of this paper apply to the combined effects of MDG and mode-dependent loss (MDL). However, for the sake of simplicity, we refer simply to MDG.

[2]We avoid using the OSNR acronym because we evaluate the optical SNR at the signal bandwidth, instead of the usual 12.5 GHz bandwidth.

Currently, machine learning (ML) techniques are being extensively investigated for optical performance monitoring in both single-mode [17] and mode-multiplexed systems [18]. In [19], we propose an artificial neural network (ANN)-based solution to estimate both MDG and optical SNR in coupled SDM transmission. The results are validated in a back-to-back 32.5m 3-mode FMF link. This paper extends the results in [19], discussing the method in detail, and validating it in an experimental short-haul 73km 3-mode FMF link with controlled MDG and optical SNR. In addition, we apply the ANN estimator in a case study of an experimental long-haul 6-mode transmission setup with unknown MDG and SNR.

The remainder of this paper is structured as follows. Section II reviews the conventional methods used to estimate MDG and SNR in coupled SDM transmission. Section III presents the ANN-based solution. Section IV presents validation results in an experimental short-reach transmission setup. Section V applies the method in a long-haul case study. Lastly, Section VI concludes the paper.

II. Conventional Methods for MDG and SNR Estimation

A. MDG Estimation

The MDG of a link with transfer matrix $\mathbf{H}$ can be computed from the eigenvalues λ_i^2 of $\mathbf{H}\mathbf{H}^H$, where $(.)^H$ is the Hermitian transpose operator [7], [8]. The accumulated MDG can be quantified by two metrics. The first one is the peak-to-peak MDG given by the subtraction of the largest and the lowest eigenvalues in dB ($10\log_{10}(\lambda_i^{\max})^2 - 10\log_{10}(\lambda_i^{\min})^2$) [8]. The second one is the standard deviation of the eigenvalues in logarithmic scale ($\sigma_{\text{mdg}} = \text{std}(\log(\lambda_i^2))$). An interesting advantage of the standard deviation metric is that, in long-haul links with strong mode coupling, it allows to estimate the impact of MDG on capacity using analytic formulas [7], [9]. Therefore, in this paper, we use the standard deviation metric.

In DSP-based MDG estimation, $\mathbf{H}$ is unknown and needs to be estimated from the data. Classic channel estimation in narrowband and subcarrier transmission systems use the least-squares (LS) or the linear MMSE algorithms[3] [20]. These techniques transmit pilot sequences and estimate $\mathbf{H}$ based on block operations on the transmit and receive signals. Both methods have their limitations, as the LS algorithm is severely impaired by noise, and the MMSE strategy required previous knowledge of the channel covariance matrix and the noise variance [21]. These classic channel estimation algorithms cannot be directly applied in SDM optical transmission systems owing to the severe conditions of phase noise and frequency drift generated in semiconductor lasers typically used in fiber optic communications.

Alternatively, the inverse of the equalizer transfer function, $\mathbf{W}_{\text{EQ}}^{-1}$, is conventionally used as an estimate of $\mathbf{H}$ [4], [5], [11]. MIMO receivers are usually implemented by MMSE equalizers, whose transfer function can be expressed as [22], [23]

$$\mathbf{W}_{\text{MMSE}} = \left(\frac{\mathbf{I}}{\text{SNR}} + \mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H, \quad (1)$$

[3]Please note the distinction between MMSE channel estimation and MMSE equalization.

where SNR is calculated in optical domain using the signal bandwidth as reference noise bandwidth. The standard deviation σ_{mdg} is then computed from the eigenvalues $\lambda_{i,\text{MMSE}}^2$ of $\mathbf{W}_{\text{MMSE}}^{-1}(\mathbf{W}_{\text{MMSE}}^{-1})^H$. As optical communications systems are typically wideband, a narrowband per-frequency $\mathbf{W}_{\text{MMSE}}$ is calculated through the Fourier transform of the equalizer impulse response.

From the eigendecomposition of $\mathbf{W}_{\text{MMSE}}^{-1}(\mathbf{W}_{\text{MMSE}}^{-1})^H$, the relationship between the actual eigenvalues, λ_i^2, and the eigenvalues obtained by DSP, $\lambda_{i,\text{MMSE}}^2$, is given by [12]

$$\lambda_{i,\text{MMSE}}^2 = \left[\frac{\left(\lambda_i^2\right)^{-1}}{\text{SNR}^2} + \frac{2}{\text{SNR}} + \lambda_i^2\right]. \quad (2)$$

The standard deviation metric is then estimated as

$$\widehat{\sigma}_{\text{mdg}} = \text{std}\left(\log\left(\lambda_{i,\text{MMSE}}^2\right)\right). \quad (3)$$

Equations (2) and (3) indicate that the accuracy of the conventional method that estimates σ_{mdg} based on the DSP-estimated eigenvalues, $\lambda_{i,\text{MMSE}}^2$, is clearly affected by the optical SNR [12]. In coupled SDM transmission, the MMSE equalizer is usually implemented by means of semi-supervised or supervised adaptive schemes, such as the well-known least mean square (LMS) algorithm, intertwined with a phase recovery loop [24]. The results of this paper are obtained by a fully-supervised LMS algorithm. Although the LMS algorithm reaches the MMSE for mild channel conditions, it can suffer from implementation issues in extreme channel conditions, such as in pathological levels of MDG.

Fig. 1 shows the maximum and minimum eigenvalues for coupled SDM transmission, for three different SNR values. For each transmission distance, matrices $\mathbf{H}$ are generated using the semi-analytical multisection model presented in [7]. Details of the multisection model employed to simulate the SDM channel are found in Appendix A. The model simulates the coupled transmission of $2\text{M} = 12$ spatial and polarization modes over 50-km fiber spans. The per-amplifier MDG standard deviation, σ_g, is set to 1 dB. Eigenvalues λ_i^2 are calculated directly from $\mathbf{H}$. Eigenvalues $\lambda_{i,\text{MMSE}}^2$ are computed by inverting $\mathbf{W}_{\text{MMSE}}$ calculated in (1). Eigenvalues $\lambda_{i,\text{LMS}}^2$ are obtained by Monte-Carlo simulation of a complete coupled SDM transmission system. The transmitter generates 12 independent sequences of 460,000 16-QAM symbols at 30 GBd. The complex signals are shaped by root-raised-cosine (RRC) filters and converted to the optical domain by a Mach-Zehnder modulator (MZM) model. The simulated channel consists of 1,000 frequency bins spread over 240 GHz (note that the simulation bandwidth is 30 GHz times 8 samples per symbol, yielding 240 GHz). The resolution of the channel in frequency domain is adjusted by replicating channel matrices between simulated frequency bins. The modal dispersion per span is 21.9 ps, corresponding to 50-km of a fiber with group delay standard deviation of 3.1 ps/$\sqrt{\text{km}}$ [25]. Additive white Gaussian noise is added with equal power to all received channel streams to set the desired receiver SNR. At the receiver, the signals are converted to the electric domain by a coherent receiver front-end model. The electrical signals are then filtered, digitized, and processed by the DSP chain, including a

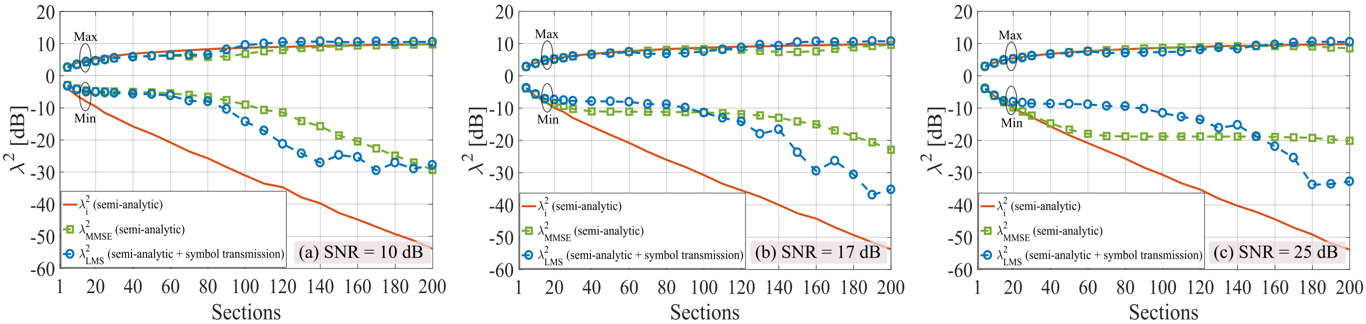

Fig. 1. Evolution of the maximum and minimum actual eigenvalues, λ_i^2, and DSP-estimated eigenvalues, $\lambda_{\mathrm{MMSE}}^2$, and λ_{LMS}^2, with the number of sections. (a) At low SNR = 10 dB. (b) At medium SNR = 17 dB. (c) At high SNR = 25 dB.

fully-supervised LMS algorithm. The detailed description of the simulation setup is presented in [14].

In Figs. 1(a)–(c), three regimes of $\lambda_{i,\mathrm{MMSE}}^2$ can be identified. In the first regime, the absolute values of both maximum and minimum $\lambda_{i,\mathrm{MMSE}}^2$ simultaneously increase, tracking the actual eigenvalues λ_i^2 . In the second regime, both maximum and minimum $\lambda_{i,\mathrm{MMSE}}^2$ remain approximately constant, leading the minimum $\lambda_{i,\mathrm{MMSE}}^2$ to deviate considerably from λ_i^2 . In the third regime, the absolute value of the minimum $\lambda_{i,\mathrm{MMSE}}^2$ increases again. The LMS MIMO equalizer results in a maximum $\lambda_{i,\mathrm{LMS}}^2$ that tracks the maximum λ_i^2 and $\lambda_{i,\mathrm{MMSE}}^2$ with high accuracy over the entire link. On the other hand, the minimum $\lambda_{i,\mathrm{LMS}}^2$ diverges from the minimum $\lambda_{i,\mathrm{MMSE}}^2$ for high values of accumulated MDG.

The results in Fig. 1 indicate that eigenvalues derived directly from the equalizer coefficients, such as $\lambda_{i,\mathrm{LMS}}^2$, track the actual eigenvalues λ_i^2 only over low-MDG links. For long distances and high MDG, conventional estimation methods largely underestimate the link MDG. The correction factor proposed in [12] can partially compensate for this mismatch in scenarios of moderate MDG and low SNR, where $\lambda_{i,\mathrm{MMSE}}^2$ tracks $\lambda_{i,\mathrm{LMS}}^2$, however, its correction capability is limited in pathological scenarios of extremely high MDG. Although these pathological scenarios may seem unlike in a first glance, it is possible to reach these levels in weakly coupled transmission, for which the MDG increases linearly with the link length.

B. Optical SNR Estimation

Estimating the optical SNR is also not trivial in coupled SDM systems affected by MDG. In systems with coherent detection, the optical SNR can be estimated from the so-called electrical SNR. In systems with MMSE equalization, the electrical SNR in stream i is actually a signal-to-noise-plus interference ratio (SINR_i) [23]

$$\mathrm{SINR}_i = \frac{1}{\left[\left(\mathbf{I} + \mathrm{SNR}\ \mathbf{H}^H \mathbf{H}\right)^{-1}\right]_{i,i}} - 1, \qquad (4)$$

where $[\ ]_{i,i}$ indicates the i-th element in the main diagonal. The optical SNR is then estimated as

$$\widehat{\mathrm{SNR}} = \frac{1}{2\,\mathrm{M}} \sum_{i=1}^{2\,\mathrm{M}} \mathrm{SINR}_i. \qquad (5)$$

In single-mode transmission with low PDG, $\mathbf{H}$ is approximately unitary, such that, in (4) and (5), $\widehat{\mathrm{SNR}} \approx \mathrm{SNR}$. Therefore, $\widehat{\mathrm{SNR}}$ is usually obtained from SINR_i, which is calculated using the least-squares (LS) method [26], [27]. In MDG-impaired SDM systems, however, $\mathbf{H}$ is non-unitary, turning $\widehat{\mathrm{SNR}}$ dependent on $\mathbf{H}$. In this case, estimating the SNR from the SINR_i would underestimate the actual value.

C. Implementation Penalty

Another issue that may be taken into account in (1) to (5) is the fact that implementation imperfections also affect the interplay of noise and MDG. These imperfections can be modeled as a contribution added to the optical noise. In this case, the SNR can be redefined as SNR', expressed as

$$\mathrm{SNR}' = \left(\frac{1}{\mathrm{SNR}} + \frac{1}{\mathrm{SNR_{imp}}}\right)^{-1}, \qquad (6)$$

where $\mathrm{SNR_{imp}}$ is an implementation penalty computed as the average SINR_i estimated from the equalized data streams in back-to-back, i.e., without any MDG and optical noise.

To improve the accuracy of the conventional SNR estimation technique that employs the LS method, the implementation penalty contribution can be removed. In this case, the estimated SNR is redefined as

$$\widehat{\mathrm{SNR}} = \left(\frac{1}{\frac{1}{2\,\mathrm{M}} \sum_{i=1}^{2\,\mathrm{M}} \mathrm{SINR}_i} - \frac{1}{\mathrm{SNR_{imp}}}\right)^{-1}. \qquad (7)$$

III. ANN-Based Method for MDG and SNR Estimation

To circumvent the limitations of conventional methods, we propose in [19] an ANN-based method to estimate σ_{mdg} and SNR from features extracted after DSP. The block diagram of the proposed solution is depicted in Fig. 2.

The training dataset is generated according to Fig. 2(a). Using the multisection model, $2\mathrm{M} \times 2\mathrm{M}$ matrices $\mathbf{H}$ are generated to

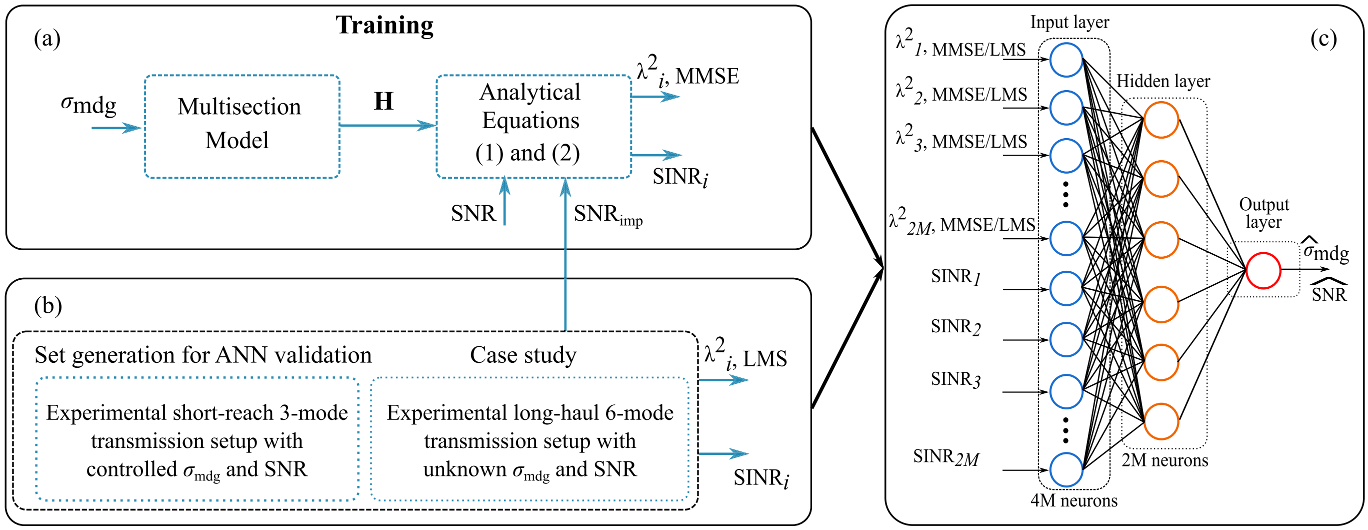

Fig. 2. ANN-based MDG and SNR estimator. (a) The training set is generated by numerical multisection simulation and analytic formulas. (b) The validation set is generated by a short-reach 3-mode transmission setup and the case study data is generated by a long-haul 6-mode transmission setup. (c) Proposed ANN. The algorithm applies two separate networks for σ_{mdg} and SNR estimation.

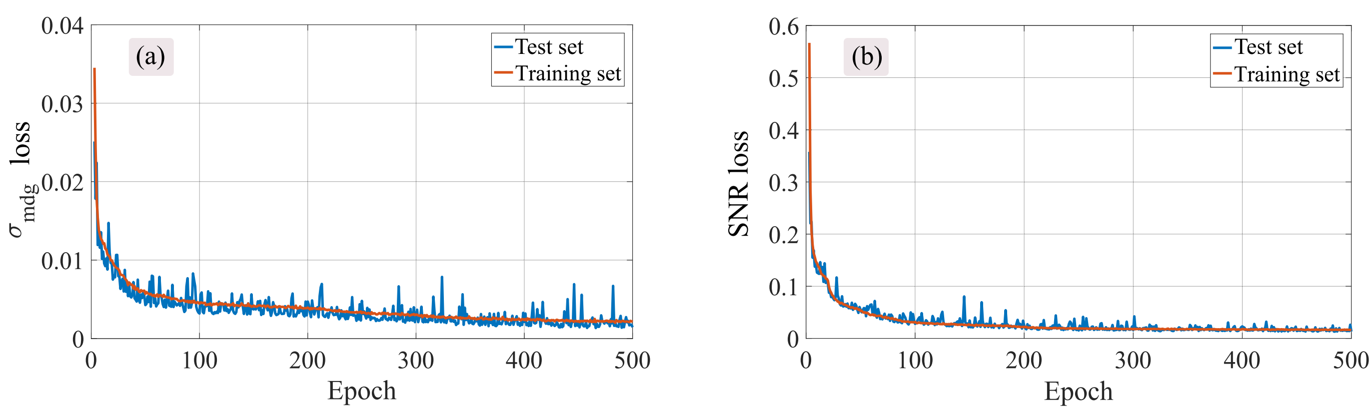

Fig. 3. ANN loss, calculated as the MSE, as a function of the number of epochs. (a) σ_{mdg} estimation. (b) SNR estimation. The curves indicate no overfitting and a good generalization ability.

simulate an M-mode transmission with polarization multiplexing over a link with $\mathrm{K}=50$ spans of 50 km each, yielding a total length of 2,500 km. The overall MDG of $\mathbf{H}$ is controlled by the per-amplifier MDG standard deviation, σ_g, that is employed to define the uncoupled gains vector per span of (11). The standard deviation of the overall MDG is given by [7]

$$\sigma_{\text{mdg}} = \xi\sqrt{1+\frac{\xi^2}{12\,(1-(2\,\mathrm{M})^{-2})}}, \tag{8}$$

where the accumulated MDG standard deviation, ξ, increases with the number of spans, K, as $\xi = \sigma_{\text{g}}\ \sqrt{\mathrm{K}}$ [7]. The MDG of the simulated 2,500 km-FMF link is adjusted to result in 0.2 dB $<$ $\sigma_{\text{mdg}} < 7.9$ dB[4]. As the actual σ_{mdg} parameter is known, the σ_{mdg} range is defined based on the minimum and maximum expected σ_{mdg} to be estimated. The modal dispersion per span obeys to a group delay standard deviation of 3.1 ps/$\sqrt{\text{km}}$ [25].

For each $\mathbf{H}$, the SNR is swept from 10 dB to 25 dB to generate $2\mathrm{M}$ $\lambda^2_{i,\text{MMSE}}$ values and $2\mathrm{M}$ SINR_i values using (1) and (4). Input SINR_{imp}, is a parameter measured experimentally in the transmission setup where the technique is being applied. The labelled set of $\lambda^2_{i,\text{MMSE}}$ and SINR_i is fed into the ANN shown in Fig. 2(c) as input training features. The ANN receives $2\mathrm{M}$ $\lambda^2_{i,\text{MMSE}}$ values and $2\mathrm{M}$ SINR_i values, and provides an estimate of σ_{mdg} or SNR. A hidden layer with $2\mathrm{M}$ neurons, and an output layer with 1 neuron, learn the relation between the input features and the output. The ANN is trained using the Adam optimizer [28] using batches of 5 samples.

Figs. 3(a) and 3(b) show the ANN convergence curves for MDG and SNR estimation, respectively, considering $2\mathrm{M}=6$. The ANN loss, calculated as the MSE, is depicted as a function

[4]Note that, according to [9], $\sigma_{\text{mdg}} > 7$ dB is sufficient to reduce the channel capacity below 80% with respect to a system without MDG. Therefore, we expect this range to cover most practical applications.

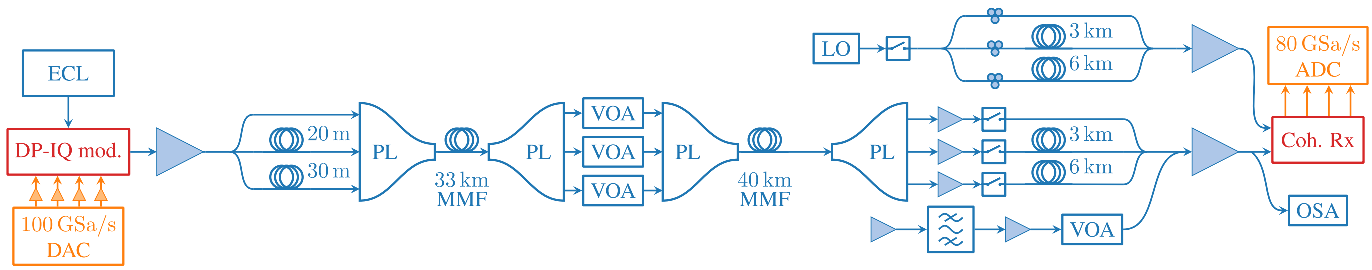

Fig. 4. Experimental setup for short-reach 3-mode transmission with polarization multiplexing [15]. The transmitter generates 16-QAM symbols at 25 GBd, which are subsequently split and delayed to create the input tributaries for the PL. The multi-mode signal is transmitted over 73 km of MMF [29]. VOAs are used to control the σ_{mdg} of the link. At the receiver, a TDM-SDM scheme is employed, and a noise loading stage is used to vary the OSNR. After DSP, the LMS eigenvalues are computed from $\mathbf{W}_{\mathrm{LMS}}^{-1}(\mathbf{W}_{\mathrm{LMS}}^{-1})^H$. The SINR_i is computed from each of the 6 equalized data streams.

of the number of epochs. Both training and test sample sets are evaluated. The results indicate a substantial reduction in the MSE after 100 epochs, and still a small improvement up to 500 epochs for σ_{mdg} estimation. Therefore, we use 500 epochs for training. The loss of the test set tracks the loss of the training set for the entire figure. We expect, therefore, no overfitting and a good ANN ability to generalize over unseen samples. After training, the ANN-based method is validated using data captured from a short-reach experimental setup with controlled parameters, and tested in a case study of a long-haul experimental link.

IV. Experimental Short-Reach Transmission With Known SNR and MDG

A. Experimental Short-Reach 3-Mode Validation Setup

The ANN-based estimator is validated using the short-reach 3-mode transmission setup presented in [15] and depicted in Fig. 4. The transmitter generates 2^{16} polarization-multiplexed 16-QAM symbols at a transmission rate of 25 GBd, based on a PCG-64 pseudo-random number generator. A RRC filter with 0.01 roll-off factor is used for pulse shaping. The pulse-shaped signal is converted to the analog domain by a 100 GSa/s digital-to-analog converter (DAC) followed by the optical modulator. After optical modulation, the signal is amplified by an erbium-doped fibre amplifier (EDFA), split and delayed to generate three decorrelated data streams to be launched through the LP_{01}, $\mathrm{LP}_{11\mathrm{a}}$ and $\mathrm{LP}_{11\mathrm{b}}$ spatial modes. Considering polarization modes, the setup supports the transmission of 6 orthogonal modes. The three polarization-multiplexed data streams are then multiplexed in space by a mode-selective photonic lantern (PL) [30].

The output of the PL is connected to a 50 μm core diameter graded-index MMF of 73 km [29]. The deployed multi-mode fiber supports up to 36 spatial modes, so that transmission can be eventually scaled to more spatial modes. To control the overall MDG, two photonic lanterns and three VOAs are placed after the first 33 km fiber segment. The three VOAs allow to sweep the MDG of the link by modifying the power in the three spatial modes. At the receive side, a fourth PL is used as mode demultiplexer.

The receiver employs a time-domain-multiplexed (TDM)-SDM receiver [31] to reduce the required amount of the coherent receivers. The ASE noise is varied at the coherent receiver input by a noise loading stage composed of two EDFAs, a wavelength selective switch (WSS) and a VOA. The SNR is computed as $\mathrm{SNR} = \mathrm{OSNR}\,(\mathrm{T_s} \times 12.5\,\mathrm{GHz})$, where $\mathrm{T_s} = 40\,\mathrm{ps}$ is the symbol time, and the OSNR is the traditional optical signal-to-noise ratio measured by an optical spectrum analyser (OSA) at the 12.5 GHz bandwidth [16]. The noisy signal is amplified and converted from the optical to the electrical domain by the receiver front-end. The TDM electric signals are fed into 80 GSa/s analog-to-digital converters (ADC) to be digitized.

In the DSP block, the TDM streams are parallelized and down-sampled to two samples per symbol. To compensate for modal dispersion and linear coupling, 6×6 MIMO equalization is carried out using a widely linear complex-valued adaptive equalizer, updated by a fully supervised LMS algorithm [32]. After DSP, the eigenvalues $\lambda^2_{i,\mathrm{LMS}}$ are computed at each frequency of $\mathbf{W}_{\mathrm{LMS}}$ and averaged across the signal band. The SINR_i is computed from each of the 6 equalized data streams using a single-coefficient LS estimator [26]. The implementation penalty is computed in back-to-back as $\mathrm{SNR_{imp}} = 18.8$ dB.

The ANN in Fig. 2(c) is fed with 9,610 analytical labelled samples generated using (1) and (4) as indicated in Fig. 2(a). In a first stage, 8,649 samples are used for model training and the remainder 970 samples for model testing. After training, model validation is performed by 520 experimental samples generated by the short-reach 3-mode transmission setup. Using the VOAs located in middle of the span, σ_{mdg} is varied from 4.5 dB to 6.5 dB. At the receiver, the noise loading stage sweeps the SNR from 11 dB to 22 dB.

B. Experimental Short-Reach 3-Mode Validation Results

The validation results of the conventional and ANN-based estimators in a short-reach transmission setup are depicted in Fig. 5. Figs. 5(a) and 5(d) compare the actual and estimated σ_{mdg} and SNR parameters generated by the conventional and ANN-based methods. The estimated values track the actual values within a small deviation over the entire range of parameters evaluated, resulting in a MSE of 0.11 for σ_{mdg} and 0.53 for SNR, both computed in dB.

Figs. 5(b) and 5(e) show the estimation error provided by the conventional method in dB, computed as the difference between

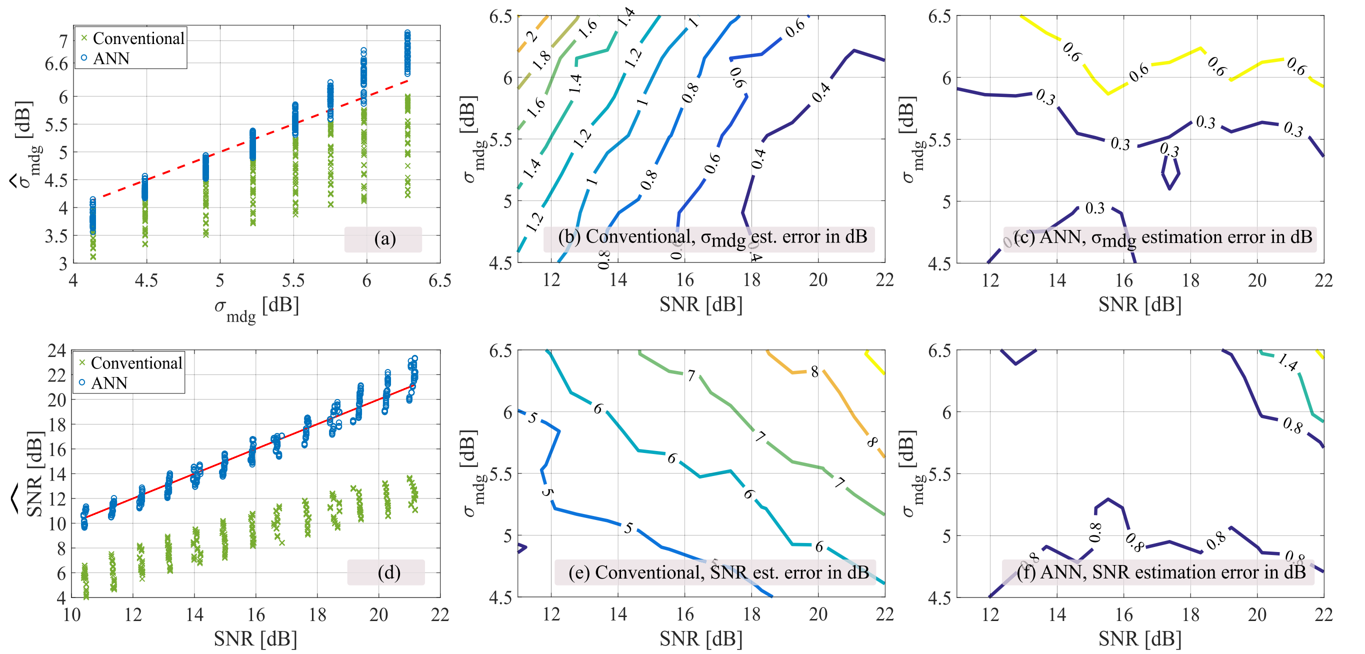

Fig. 5. Experimental short-reach 3-mode validation results. (a) Estimated σ_{mdg} as a function of the actual σ_{mdg}. (b) σ_{mdg} estimation error in dB generated by the conventional method as a function of the actual σ_{mdg} and SNR. (c) σ_{mdg} estimation error in dB generated by the ANN as a function of the actual σ_{mdg} and SNR. (d) Estimated SNR as a function of the actual SNR. (e) SNR estimation error in dB generated by the conventional method as a function of the actual σ_{mdg} and SNR. (f) SNR estimation error in dB generated by the ANN as a function of the actual σ_{mdg} and SNR.

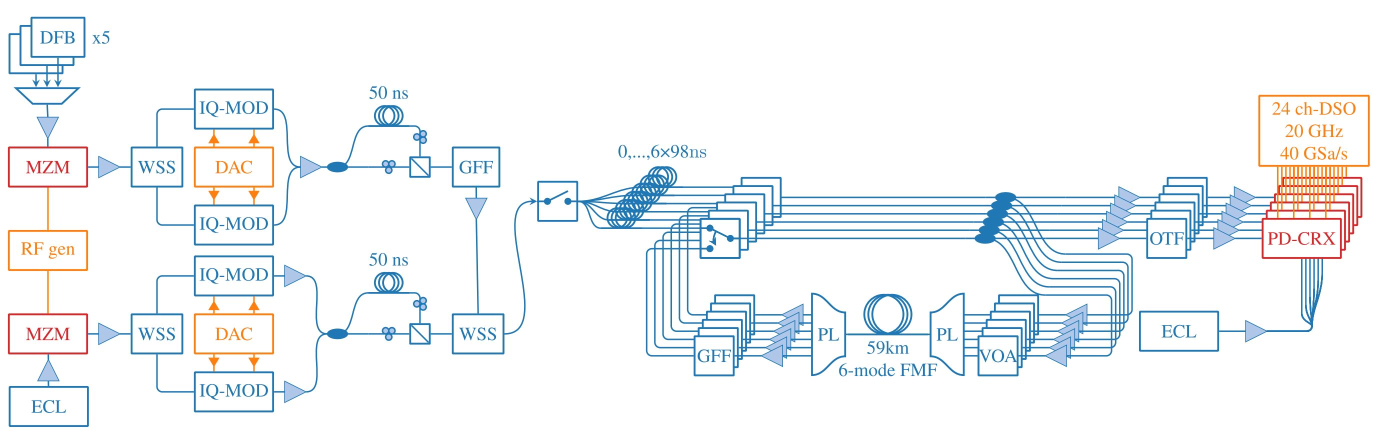

Fig. 6. Experimental long-haul 6-mode transmission setup with recirculating loop of 59 km. 12 spatial and polarization modes are supported, each one carry 15 WDM channels centered around 1550 nm. Triangles represent EDFAs. See [4] for more details. After DSP, the LMS eigenvalues are computed from $\mathbf{W}_{\mathrm{LMS}}^{-1}(\mathbf{W}_{\mathrm{LMS}}^{-1})^H$. The SINR_i is computed from each one of the 12 equalized data.

the actual value and the estimated value. The conventional method provides a σ_{mdg} estimation error up to 2 dB at high MDG and low SNR. In the case of SNR, the estimation error achieves up to 8 dB at high levels of MDG and high SNR. Figs. 5(c) and 5(f) show the estimation error in decibels for σ_{mdg} and SNR, respectively, for the ANN solution. The ANN estimator provides a highest residual σ_{mdg} estimation error of 0.6 dB in the region of high MDG, exhibiting a low dependence on the evaluated SNR. On most of the grid, the σ_{mdg} estimation error is lower than 0.3 dB. For the SNR, an estimation error higher than 1.4 dB is observed at high values of σ_{mdg} and SNR. Over most of the evaluated range, the SNR estimation error is lower than 0.8 dB.

V. Experimental Long-Haul Transmission With Unknown SNR and MDG

A. Experimental Long-Haul 6-Mode Case Study Setup

We also apply the ANN-based estimator to the long-haul 6-mode transmission with polarization multiplexing setup presented in [4] and depicted in Fig. 6. The transmission setup includes 15 WDM channels transmitted over 4 linearly polarised (LP) spatial modes ($\mathrm{LP}_{01}, \mathrm{LP}_{11}, \mathrm{LP}_{21}$, and LP_{02}). Including polarization and degenerate modes (LP_{11a}, LP_{11b}, LP_{21a}, and LP_{21b}) the setup supports 12 propagation modes. The 15-channel comb is generated using five distributed feedback

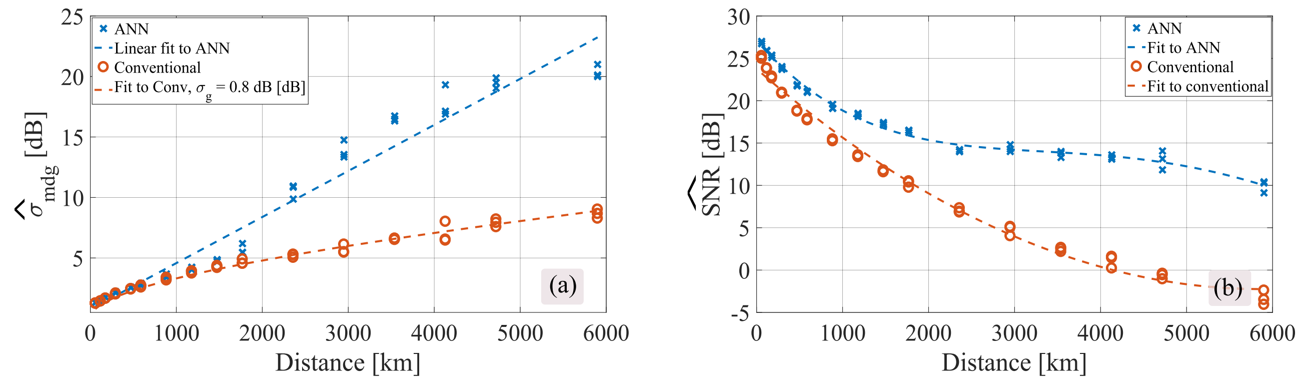

Fig. 7. Experimental long-haul 6-mode case study. (a) Estimated σ_{mdg} as a function of the transmission distance. (b) Estimated SNR as a function of the transmission distance.

lasers (DFB) and one phase-modulated Mach-Zehnder modulator (MZM). Odd and even channels are separately modulated with 120 Gb/s 16-QAM using IQ-modulators. Polarization-multiplexing is generated by splitting, delaying and combining the transmitted signals. The channel under test (CUT) is generated separately using a similar scheme. Six conventional single-mode recirculating loops are combined with PLs and a 59 km long 6-mode FMF. The output of the loop setup is amplified and forwarded to a coherent receiver array. The produced electrical signals are digitized by a 24 channel oscilloscope, followed by offline DSP. 12×12 equalization is carried out using a MIMO equalizer updated by a fully supervised LMS algorithm. After equalization, the eigenvalues $\lambda^2_{i,\mathrm{LMS}}$ are computed at each frequency of $\mathbf{W}_{\mathrm{LMS}}$, and averaged across the signal band. The SINR_i is computed for each of the 12 equalized data streams. The implementation penalty is computed in back-to-back as $\mathrm{SNR}_{\mathrm{imp}} = 18.6$ dB.

As the parameters σ_{mdg} and SNR are unknown, the simulated link used to generate the training set is adjusted to present $0.5\text{ dB} < \sigma_{\mathrm{mdg}} < 23.9$ dB and $10\text{ dB} < \mathrm{SNR} < 25$ dB, expecting that the minimum and maximum values to be estimated are close to or within these intervals. After ANN training by 46,035 labelled samples, the ANN-based method is applied to experimental traces corresponding to transmission distances between 59 km and 5,900 km.

B. *Experimental Long-Haul 6-Mode Case Study Results*

Fig. 7 shows σ_{mdg} and SNR estimated in the long-haul case study. Fig. 7(a) shows the estimated σ_{mdg} as a function of the transmission distance. The orange circles correspond to σ_{mdg} estimated by the conventional method. The dashed orange line fits the experimental data to (8) with a per-span MDG of $\sigma_{\mathrm{g}} = 0.8$ dB. The ANN-based estimates are shown by the blue crosses, indicating a large deviation with respect to the conventional method. The approximately linear increase of the σ_{mdg} estimated by the ANN suggests a possible weakly-coupled SDM transmission [33]. Fig. 7(b) shows the estimated SNR as a function of the transmission distance. As expected, the conventional technique results are substantially lower than those obtained by the ANN-based method. As the conventional method neglects the detrimental effects of MDG, it tends to underestimate the actual SNR. The dashed curves are polynomial fitting functions indicating the trend of the SNR with the transmission distance.

To further investigate the weak coupling hypothesis, we also evaluate the channel delay spread. In general, weakly-coupled transmission leads to a linear increase of the channel delay spread [34], [35]. Fig. 8(a) shows the averaged impulse response computed as the average of the 144 intensity matrices obtained from the MIMO equalizer after 1,770 km and 4,130 km. The dashed red curve is a Gaussian fit whose standard deviation provides a metric for evaluating the total delay spread. Fig. 8(b) shows the standard deviation of the averaged impulse response as a function of the transmission distance. The approximately linear increase in the equalizer impulse response corroborates the hypothesis of weak coupling [34].

VI. Discussion

In Section V, we observed considerable differences between ANN-based and conventional estimation methods for σ_{mdg} and SNR in long-haul transmission. We conjectured that the very high values of MDG estimated by the ANN appeared because of a potential linear accumulation of MDG in the recirculation loop. To further understand the problem we attempt in this section to reproduce by simulation the results observed in Section V.

The simulated transmitter generates 12 16-QAM symbol sequences at 30 GBd. The sequences are processed by RRC shaping filters and converted to the optical domain by an MZM model. The channel model generates 12×12 channel transfer matrices $\mathbf{H}$ using the analytical multisection model presented in [7]. The per-span MDG σ_g is set to 1.5 dB. The transmission distance is varied from 1 to 100 59-km spans, yielding σ_{mdg} from 0.5 dB to 22 dB. Such high MDG would severely impair the transmission capacity (according to [9], effective SNR losses higher than 1 dB are expected for $\sigma_{\mathrm{mdg}} >$ 3-4 dB). The SNR after the first span is set to 26.83 dB, and then decreased considering noise accumulation generated by amplifiers with 9-dB noise figure. The received sequence is fed into a coherent receiver model. The digital signals are processed by a DSP chain composed of an static equalizer and a 12×12 MIMO equalizer updated by the fully-supervised LMS algorithm.

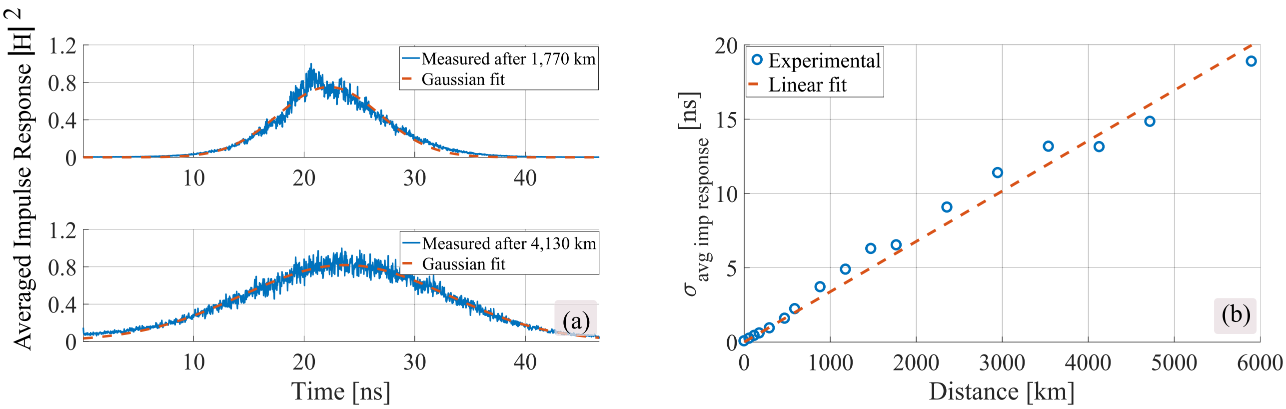

Fig. 8. (a) Averaged impulse response after 1,770 km (Top) and 4,130 km (Bottom). (b) Standard deviation of the averaged impulse response as a function of the transmission distance.

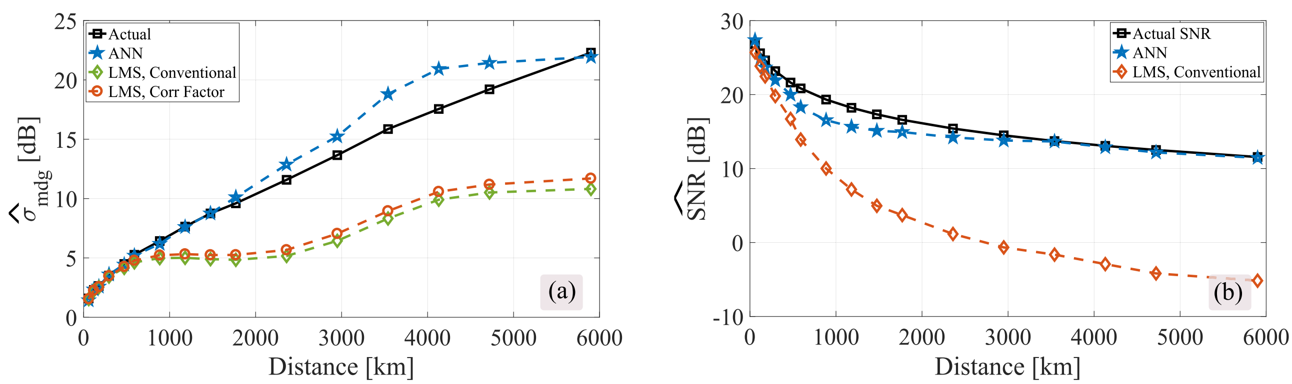

Fig. 9. (a) Estimated σ_{mdg} as a function of the transmission distance with conventional, correction factor, and ANN methods. (b) Estimated SNR as a function of the transmission distance with conventional and ANN methods.

Three σ_{mdg} estimation methods are evaluated. The LMS-based conventional method transmits symbols over the channel matrix $\mathbf{H}$. At the receiver, the transfer matrix of the dynamic LMS MIMO equalizer is used to estimate the eigenvalues $\lambda^2_{i,\text{LMS}}$ and σ_{mdg}. The LMS-based correction factor methods applies the correction factor proposed and validated in [12], [14], [15] to the DSP-estimated eigenvalues before estimating σ_{mdg}. The ANN estimator uses the same structure and training data as in the long-haul case study in Section V.

The results are shown in Fig. 9(a). The solid black line corresponds to the actual σ_{mdg} estimated from $\mathbf{H}$. The LMS-based conventional method provides accurate estimates with an estimation error less than 0.5 dB up to $\sigma_{\text{mdg}} = 4.5$ dB. After this point, the method starts to significantly underestimate the MDG, reaching two plateaus. The LMS-based technique with correction factor slightly improves the estimation quality. The correction factor provides a low correction capability at the beginning of the link because the SNR is relatively high. At the end of the link, the accumulated MDG is so high that the equalizer coefficients diverge from the MMSE coefficients. The estimates provided by the ANN-based estimator accurately track the actual MDG up to $\sigma_{\text{mdg}} = 10$ dB. For $\sigma_{\text{mdg}} > 10$ dB, the ANN estimator slightly overestimates σ_{mdg}. Although the two plateaus are not observed in the experimental data, we believe the trends are fairly reproduced.

We also estimate the SNR using the conventional (after LMS equalization) and ANN-based methods. The results are shown in Fig. 9(b). As expected, the SNR is underestimated by the conventional method because of the strong MDG added to the link. The ANN-based method offers more accurate estimates, also in reasonable agreement with the experimental results in Section V.

Finally, it should be noted that the entire simulation and experimental study was carried out under the assumption of balanced and spatially white noise. This assumption should hold in real-life long-haul links, which are in fact the systems that suffer most from MDG [7]. The effectiveness of the ANN in scenarios with noise correlation or SNR imbalances is left for a further study.

VII. Conclusion

In space division multiplexing (SDM) systems with coupled channels, the interaction of mode-dependent gain (MDG) and amplified spontaneous emission (ASE) fundamentally constrain the channel capacity and transmission distance. In these systems, accurate MDG and signal-to-noise ratio (SNR) estimation is mandatory for an adequate link assessment and troubleshooting. Conventional estimation methods present performance limitations in certain conditions of MDG and SNR. In this paper, we investigate an artificial neural network (ANN)-based method to estimate MDG and SNR in SDM systems with coupled channels based on features extracted after digital signal processing (DSP). The proposed method is validated in an experimental short-reach 3-mode transmission setup with polarization multiplexing. After validation, the ANN-based method is applied to a case study consisting of an experimental long-haul 6-mode transmission link with polarization multiplexing. The results suggest that the ANN-based method can largely exceed the performance provided by conventional methods in scenarios of high accumulated MDG, as in long-haul links with weak mode coupling.

APPENDIX A
MULTISECTION MODEL FOR STRONGLY-COUPLED SDM TRANSMISSION

The multisection model for strongly-coupled SDM transmission [7] splits the fiber link into K sections. At an angular frequency ω, the transfer matrix of the k^{th} section, $\mathbf{H}^{(\mathrm{k})}(\omega)$, can be modeled as the product of 3 $2\,\mathrm{M} \times 2\,\mathrm{M}$ matrices, with $2\,\mathrm{M}$ being the number of orthogonal modes (spatial and polarization modes)

$$\mathbf{H}^{(\mathrm{k})}(\omega) = \mathbf{V}^{(\mathrm{k})}\mathbf{\Lambda}^{(\mathrm{k})}(\omega)\mathbf{U}^{(\mathrm{k})H}, \quad \text{with } \mathrm{k} = 1, \ldots, \mathrm{K} \tag{9}$$

where $(.)^H$ denotes the Hermitian transpose operator. The random unitary matrices $\mathbf{V}$ and $\mathbf{U}$ are frequency-independent and describe mode coupling at the input and output of the section, respectively. The diagonal matrix $\mathbf{\Lambda}^{(\mathrm{k})}(\omega)$, accounting for MDG and mode dispersion, is given by

$$\mathbf{\Lambda}^{(\mathrm{k})}(\omega) = \mathrm{diag}\left[\mathrm{e}^{\frac{1}{2}\mathrm{g}_1^{(\mathrm{k})} - \mathrm{j}\omega\tau_1^{(\mathrm{k})}}, \mathrm{e}^{\frac{1}{2}\mathrm{g}_2^{(\mathrm{k})} - \mathrm{j}\omega\tau_2^{(\mathrm{k})}}, \ldots, \mathrm{e}^{\frac{1}{2}\mathrm{g}_{2\mathrm{M}}^{(\mathrm{k})} - \mathrm{j}\omega\tau_{2\mathrm{M}}^{(\mathrm{k})}}\right], \tag{10}$$

where

$$\mathbf{g}^{(\mathrm{k})} = \left(\mathrm{g}_1^{(\mathrm{k})}, \mathrm{g}_2^{(\mathrm{k})}, \ldots, \mathrm{g}_{2\mathrm{M}}^{(\mathrm{k})}\right), \tag{11}$$

and

$$\boldsymbol{\tau}^{(\mathrm{k})} = \left(\tau_1^{(\mathrm{k})}, \tau_2^{(\mathrm{k})}, \ldots, \tau_{2\mathrm{M}}^{(\mathrm{k})}\right), \tag{12}$$

are the uncoupled log-power gains and the uncoupled modal groups delays, respectively. The overall transfer matrix $\mathbf{H}(\omega)$ of a fiber of K sections is given by

$$\mathbf{H}(\omega) = \prod_{\mathrm{k}=1}^{\mathrm{K}} \mathbf{H}^{(\mathrm{k})}(\omega). \tag{13}$$

The transfer matrix $\mathbf{H}(\omega)$ can be also expressed, after singular-value decomposition, as the product of three matrices

$$\mathbf{H}(\omega) = \mathbf{V}(\omega)\mathbf{\Lambda}(\omega)\mathbf{U}^H(\omega), \tag{14}$$

where $\mathbf{V}$ and $\mathbf{U}$ are unitary matrices, and $\mathbf{\Lambda}$ is a diagonal matrix that characterizes the system MDG.

In this paper, the multisection model is implemented in Matlab assuming one section per fiber span, following the steps below:

1) Define the number of modes 2 M, the number of fiber spans K, the span length L (in km), the group delay standard deviation μ (in ps/$\sqrt{\text{km}}$), and the target accumulated MDG standard deviation ξ (in dB).
2) From ξ, compute the per-amplifier MDG standard deviation, σ_g, in log power units, $\sigma_{g,\,\text{logpw}}$.
3) Compute the uncoupled log-power gains vector of (11) as $[\text{ones}(1, \mathrm{M}) \; - \text{ones}(1, \mathrm{M})] \times \sigma_{g,\,\text{logpw}}$ [7].
4) Compute the group delay standard deviation per span through $\mu_{\text{per-span}} = \mu \times \sqrt{\mathrm{L}}$.
5) Compute the uncoupled modal groups delays vector of (12) as $\mathbf{GD} \times \mu_{\text{per-span}}$, where $\mathbf{GD}$ is a Gaussian distributed vector that sums to zero.
6) For each span, generate two frequency-independent random unitary matrices $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$ to describe mode coupling at the input and output.
7) At each frequency, ω, and for each k^{th} span, generate the diagonal matrix $\mathbf{\Lambda}^{(\mathrm{k})}(\omega)$ using the uncoupled log-power gains and the uncoupled modal groups delays, obtained in steps 3 and 5, respectively.
8) At each frequency, ω, and each k^{th} span, multiply the three matrices, $\mathbf{V}^{(k)}$, $\mathbf{\Lambda}^{(\mathrm{k})}(\omega)$, and $\mathbf{U}^{(k)}$ to generate $\mathbf{H}^{(\mathrm{k})}(\omega)$.
9) For each frequency, ω, multiply the K matrices $\mathbf{H}^{(\mathrm{k})}(\omega)$ to generate the overall transfer matrix $\mathbf{H}(\omega)$.
10) Normalize $\mathbf{H}(\omega)$ by the square root of the mean of the overall coupled modal power gains at all frequencies, to ensure equal input and output powers.

REFERENCES

[1] R. Ryf *et al.*, “Coupled-core transmission over 7-core fiber,” in *Proc. Opt. Fiber Commun. Conf.*, 2019, Paper Th4B–3.
[2] N. K. Fontaine *et al.*, “30 × 30 MIMO transmission over 15 spatial modes,” in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper Th5C–1.
[3] M. van den Hout *et al.*, “1 Tbit/s/λ transmission over a 130 km link consisting of graded-index 50 μm core multi-mode fiber and 6LP few-mode fiber,” in *Proc. Eur. Conf. Opt. Commun.*, 2020, pp. 1–4.
[4] J. van Weerdenburg *et al.*, “138-Tb/s mode-and wavelength-multiplexed transmission over six-mode graded-index fiber,” *J. Lightw. Technol.*, vol. 36, no. 6, pp. 1369–1374, 2018.
[5] G. Rademacher *et al.*, “Long-haul transmission over few-mode fibers with space-division multiplexing,” *J. Lightw. Technol.*, vol. 36, no. 6, pp. 1382–1388, 2018.
[6] D. Soma *et al.*, “10.16 peta-bit/s dense SDM/WDM transmission over low-DMD 6-mode 19-core fibre across C + L band,” in *Proc. Eur. Conf. Opt. Commun.*, 2017, pp. 1–3.
[7] K.-P. Ho and J. M. Kahn, “Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing,” *Opt. Exp.*, vol. 19, no. 17, pp. 16612–16635, 2011.
[8] P. J. Winzer and G. J. Foschini, “MIMO capacities and outage probabilities in spatially multiplexed optical transport systems,” *Opt. Exp.*, vol. 19, no. 17, pp. 16680–16696, 2011.
[9] D. A. A. Mello *et al.*, “Impact of polarization- and mode-dependent gain on the capacity of ultra-long-haul systems,” *J. Lightw. Technol.*, vol. 38, no. 2, pp. 303–318, 2020.
[10] J. van Weerdenburg *et al.*, “138 tbit/s transmission over 650 km graded-index 6-mode fiber,” in *Proc. Eur. Conf. Opt. Commun.*, 2017, pp. 1–3.
[11] G. Rademacher *et al.*, “10.66 peta-bit/s transmission over a 38-core-three-mode fiber,” in *Proc. Opt. Fiber Commun. Conf.*, 2020, Paper Th 3H–1.
[12] R. S. Ospina, C. Okonkwo, and D. A. Mello, “DSP-based mode-dependent loss and gain estimation in coupled SDM transmission,” in *Proc. Opt. Fiber Commun. Conf.*, 2020, pp. W2A–47.

[13] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1125–1141, 2017.

[14] R. S. Ospina *et al.*, "Mode-dependent loss and gain estimation in SDM transmission based on MMSE equalizers," *J. Lightw. Technol.*, vol. 39, no. 7, pp. 1968–1975, 2020.

[15] M. van den Hout *et al.*, "Experimental validation of MDL emulation and estimation techniques for SDM transmission systems," in *Proc. Eur. Conf. Opt. Commun.*, 2020, pp. 1–4.

[16] R.-J. Essiambre *et al.*, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, 2010.

[17] W. Saif *et al.*, "Machine learning techniques for optical performance monitoring and modulation format identification: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2839–2882, Oct.–Dec. 2020.

[18] W. S. Saif *et al.*, "Optical performance monitoring in mode division multiplexed optical networks," *J. Lightw. Technol.*, vol. 39, no. 2, pp. 491–504, 2020.

[19] R. S. B. Ospina, M. Van Den Hout, S. Van Der Heide, C. Okonkwo, and D. A. A. Mello, "Neural-network-based MDG and optical SNR estimation in SDM transmission," in *Proc. Opt. Fiber Commun. Conf.*, 2021, Paper Th1A–20.

[20] J.-J. Van De *et al.*, "On channel estimation in OFDM systems," in *Proc. IEEE 45th Veh. Technol. Conf.*, 1995, pp. 815–819.

[21] O. Edfors *et al.*, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[22] N. Kim, Y. Lee, and H. Park, "Performance analysis of MIMO system with linear MMSE receiver," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4474–4478, Nov. 2008.

[23] M. R. McKay, I. B. Collings, and A. M. Tulino, "Achievable sum rate of MIMO MMSE receivers: A general analytic framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 396–410, Jan. 2010.

[24] Y. Mori, C. Zhang, and K. Kikuchi, "Novel configuration of finite-impulse-response filters tolerant to carrier-phase fluctuations in digital coherent optical receivers for higher-order quadrature amplitude modulation signals," *Opt. Exp.*, vol. 20, no. 24, pp. 26236–26251, 2012.

[25] T. Hayashi *et al.*, "Record-low spatial mode dispersion and ultra-low loss coupled multi-core fiber for ultra-long-haul transmission," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 450–457, 2017.

[26] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe, "Comparison of EM-based algorithms for MIMO channel estimation," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 216–226, Jan. 2007.

[27] A. Das and B. D. Rao, "SNR and noise variance estimation for MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3929–3941, Aug. 2012.

[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[29] P. Sillard *et al.*, "50 μm multimode fibers for mode division multiplexing," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1672–1677, 2016.

[30] A. M. Velázquez Benítez *et al.*, "Scaling photonic lanterns for space-division multiplexing," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, 2018.

[31] R. Van Uden *et al.*, "Ultra-high-density spatial division multiplexing with a few-mode multicore fibre," *Nature Photon.*, vol. 8, no. 11, 2014, Art. no. 865.

[32] E. P. da Silva and D. Zibar, "Widely linear equalization for IQ imbalance and skew compensation in optical coherent receivers," *J. Lightw. Technol.*, vol. 34, no. 15, pp. 3577–3586, 2016.

[33] A. Lobato *et al.*, "Impact of mode coupling on the mode-dependent loss tolerance in few-mode fiber transmission," *Opt. Exp.*, vol. 20, no. 28, pp. 2977–29783, 2012.

[34] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *J. Lightw. Technol.*, vol. 29, no. 21, pp. 3119–3128, 2011.

[35] C. Antonelli, A. Mecozzi, and M. Shtaif, "The delay spread in fibers for SDM transmission: Dependence on fiber parameters and perturbations," *Opt. Exp.*, vol. 23, no. 3, pp. 2196–2202, 2015.