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

Ruby S. B. Ospina, Student Member, OSA, 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 loss, mode-dependent gain, space division multiplexing, optical fiber communications.

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) 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. 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

1The 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.

2We avoid using the OSNR acronym because we evaluate the optical SNR at the signal bandwidth, instead of the usual 12.5 GHz bandwidth.
MIMO receivers are usually implemented by MMSE equalizers, whose transfer function can be expressed as [20], [21]

\[
W_{\text{MMSE}} = \left( \frac{I}{\text{SNR}} + H^H H \right)^{-1} H^H,
\]

where SNR is calculated in optical domain using the signal bandwidth as reference noise bandwidth. The standard deviation \( \sigma_{\text{mdg}} \) is then computed from the eigenvalues \( \lambda^2_{\text{MMSE}} \) of \( W_{\text{MMSE}}^{-1} (W_{\text{MMSE}}^{-1})^H \).

From the eigendecomposition of \( W_{\text{MMSE}}^{-1} (W_{\text{MMSE}}^{-1})^H \), the relationship between the actual eigenvalues, \( \lambda^2_i \), and the eigenvalues obtained by DSP, \( \lambda^2_{\text{MMSE}} \), is given by [12]

\[
\lambda^2_{i,\text{MMSE}} = \left[ \left( \lambda^2_i \right)^{-1} + \frac{2}{\text{SNR}} + \lambda^2_i \right].
\]

The standard deviation metric is then estimated as

\[
\hat{\sigma}_{\text{mdg}} = \text{std}(\log(\lambda^2_{i,\text{MMSE}})).
\]

Eqs. (2) and (3) indicate that the accuracy of the conventional method that estimates \( \sigma_{\text{mdg}} \) based on the DSP-estimated eigenvalues, \( \lambda^2_{i,\text{MMSE}} \), 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. 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 \( H \) are generated using the semi-analytical multisecton model presented in [7]. The model simulates the coupled transmission of 2M = 12 spatial and polarization modes over 50-km fiber spans. The per-amplifier MDG standard deviation, \( \sigma_g \), is set to 1 dB. Eigenvalues \( \lambda^2_i \) are calculated directly from \( H \). Eigenvalues \( \lambda^2_{i,\text{MMSE}} \) are computed by inverting \( W_{\text{MMSE}} \) calculated in Eq. (1). Eigenvalues \( \lambda^2_{i,\text{LMS}} \) 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 Gb/s. 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
The detailed description of the simulation setup is presented in the DSP chain, including a fully-supervised LMS algorithm. Electrical signals are then filtered, digitized, and processed by a coherent receiver front-end model. The receiver SNR is calculated using the least-squares (LS) method [23], [24]. In MDG-impaired SDM systems, however, $H$ is non-unitary, turning $\text{SNR}$ dependent on $H$. In this case, estimating the SNR from the $\text{SNR}_i$ would underestimate the actual value.

C. Implementation penalty

Another issue that may be taken into account in Eqs. (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 $\text{SNR}'$, expressed as

$$\text{SNR}' = \left( \frac{1}{\text{SNR}} + \frac{1}{\text{SNR}_{\text{imp}}} \right)^{-1},$$

where $\text{SNR}_{\text{imp}}$ is an implementation penalty computed as the average $\text{SNR}_i$ 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

$$\hat{\text{SNR}} = \left( \frac{1}{2M} \sum_{i=1}^{2M} \text{SNR}_i - \frac{1}{\text{SNR}_{\text{imp}}} \right)^{-1}.$$  

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 $\sigma_{\text{mdg}}$ and $\text{SNR}$ 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 presented in [7], $2M \times 2M$ matrices $H$ are generated to simulate an $M$-mode transmission with polarization multiplexing over a link with $K = 50$ spans of 50 km each, yielding a total length of 2,500 km.

The overall MDG of the link is controlled by the per-amplifier MDG standard deviation, $\sigma_g$. The standard deviation of the overall MDG is given by [7]

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

where the accumulated MDG standard deviation, $\xi$, increases with the number of spans, $K$, as $\xi = \sigma_g \sqrt{K}$ [7]. The MDG of the simulated 2,500 km-FMF link is adjusted to result in $0.2$ dB < $\sigma_{\text{mdg}}$ < $6.2$ dB.

For each $H$, the SNR is swept from 10 dB to 25 dB to generate $2M \lambda_{\text{MMSE}}^2$ values and $2M \text{SNR}_i$ values using Eqs. (1) and (4). Input $\text{SNR}_{\text{imp}}$, is a parameter measured experimentally in the transmission setup where the technique is being applied. The labelled set of $\lambda_{\text{MMSE}}^2$ and $\text{SNR}_i$ is fed into the ANN shown in Fig. 2(c) as input training features. The ANN receives $2M \lambda_{\text{MMSE}}^2$ values and $2M \text{SNR}_i$, values,
and provides an estimate of $\sigma_{\text{mdg}}$ or SNR. A hidden layer with 2M 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 [25] using batches of 5 samples.

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

IV. EXPERIMENTAL SHORT-REACH TRANSMISSION

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 polarization-multiplexed 16-QAM symbols at a transmission rate of 25 Gb/s. 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 fiber amplifier (EDFA), split and delayed to generate three decorrelated data streams to be launched through the LP$_{01}$, LP$_{11a}$ and LP$_{11b}$ 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) [27].

The output of the PL is connected to a 50 μm core diameter graded-index MMF of 73 km [26]. 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 de-multiplexer.

The receiver employs a time-domain-multiplexed (TDM)-SDM receiver [28] 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 $SNR = OSNR (T_s \times 12.5 \, \text{GHz})$, where $T_s = 40 \, \text{ps}$ is the symbol time, and the OSNR is the traditional optical signal-to-noise ratio measured by an optical spectrum analyzer (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, $6 \times 6$ MIMO equalization is carried out using a widely linear complex-valued adaptive equalizer, updated by a fully supervised LMS algorithm [29]. After DSP, the eigenvalues $\lambda_{\text{LMS}}^2$ are computed at each frequency of $W_{\text{LMS}}$ and averaged across the signal band. The SINR, is computed from each of the 6 equalized data streams using a single-coefficient LS estimator [23]. The implementation penalty is computed in back-to-back as $\text{SNR}_{\text{imp}} = 18.8 \, \text{dB}$.

The ANN in Fig. 2(e) is fed with 9,610 analytical labeled samples generated using Eqs. (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, $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ 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 the actual value and the estimated value. The conventional method provides a $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ and SNR, respectively, for the ANN solution. The ANN estimator provides a highest residual $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ and SNR. Over most of the evaluated range, the SNR estimation error is lower than 0.8 dB.

**V. EXPERIMENTAL LONG-HAUL TRANSMISSION**

**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 polarized (LP) spatial modes ($LP_{11}$, $LP_{12}$, $LP_{21}$, and $LP_{22}$). Including polarization and degenerate modes ($LP_{12a}$, $LP_{12b}$, $LP_{21a}$, and $LP_{21b}$) the setup supports 12 propagation modes. The 15-channel comb is generated using five distributed feedback 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 50 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 \times 12 equalization is carried out using a MIMO equalizer updated.
Fig. 5. Experimental short-reach 3-mode validation results. (a) Estimated $\sigma_{\text{mdg}}$ as a function of the actual $\sigma_{\text{mdg}}$. (b) $\sigma_{\text{mdg}}$ estimation error in dB generated by the conventional method as a function of the actual $\sigma_{\text{mdg}}$ and SNR. (c) $\sigma_{\text{mdg}}$ estimation error in dB generated by the ANN as a function of the actual $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$ and SNR. (f) SNR estimation error in dB generated by the ANN as a function of the actual $\sigma_{\text{mdg}}$ 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 $W_{LMS}^{-1}(W_{LMS})^H$. The $\text{SINR}_i$ is computed from each one of the 12 equalized data streams. The implementation penalty is computed in back-to-back as $\text{SNR}_{\text{imp}} = 18.6 \text{ dB}$. 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 $\sigma_{\text{mdg}}$ and SNR estimated in the long-haul case study. Fig. 7(a) shows the estimated $\sigma_{\text{mdg}}$ as a function of the transmission distance. The orange circles correspond to $\sigma_{\text{mdg}}$ estimated by the conventional method. The dashed orange line fits the experimental data to Eq. (8) with a per-span MDG of $\sigma_{\text{g}} = 0.8 \text{ dB}$. 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 $\sigma_{\text{mdg}}$ estimated by the ANN suggests a possible weakly-coupled SDM transmission [30]. 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 [31], [32]. 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 [31].

VI. DISCUSSION

In Section V, we observed considerable differences between ANN-based and conventional estimation methods for $\sigma_{\text{mdg}}$ 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 $12 \times 12$ channel transfer matrices $H$ using the analytical multisection model presented in [7]. The per-span MDG $\sigma_g$ is set to 1.5 dB. The transmission distance is varied from 1 to 100 59-km spans, yielding $\sigma_{\text{mdg}}$ from 0.5 dB to 22 dB. Such high MDG would severely impair the transmission capacity (according to [33], effective SNR losses higher than 1 dB are expected for $\sigma_{\text{mdg}} > 3-4$ 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 a static equalizer and a $12 \times 12$ MIMO equalizer updated by the fully-supervised LMS algorithm. Three $\sigma_{\text{mdg}}$ estimation methods are evaluated. The LMS-based conventional method transmits symbols over the channel matrix $H$. At the receiver, the transfer matrix of the dynamic LMS MIMO equalizer is used to estimate the eigenvalues $\lambda^2_{\text{LMS}}$ and $\sigma_{\text{mdg}}$. The LMS-based correction factor methods applies the correction factor proposed and validated in [12], [14], [15] to the DSP-estimated eigenvalues before estimating $\sigma_{\text{mdg}}$. 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 $\sigma_{\text{mdg}}$ 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 $\sigma_{\text{mdg}}$. 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.

REFERENCES

[1] R. Ryf, J. C. Alvarado-Zacarias, S. Wittek et al., “Coupled-core transmission over 7-core fiber,” in Optical Fiber Communication Conference. Optical Society of America, 2019, pp. Th4B–3.
[2] N. K. Fontaine, R. Ryf, H. Chen et al., “30 x 30 MIMO transmission over 15 spatial modes,” in Optical Fiber Communication Conference. Optical Society of America, 2015, pp. Th5C–1.
[3] M. van den Hout, S. van der Heide, J. van Weerdenburg et al., “1 Tb/s/λ transmission over a 130 km link consisting of graded-index 50 μm core multi-mode fiber and 6LP few-mode fiber,” in 2020 European Conference on Optical Communications (ECOC). IEEE, 2020, pp. 1–4.
[4] J. van Weerdenburg, R. Ryf, J. C. Alvarado-Zacarias et al., “138-Tb/s mode-and wavelength-multiplexed transmission over six-mode graded-index fiber,” Journal of Lightwave Technology, vol. 36, no. 6, pp. 1369–1374, 2018.
[5] G. Rademacher, R. Ryf, N. K. Fontaine et al., “Long-haul transmission over few-mode fibers with space-division multiplexing,” Journal of Lightwave Technology, vol. 36, no. 6, pp. 1382–1388, 2018.
[6] D. Soma, Y. Wakayama, S. Beppu et al., “10.16 peta-bit/s dense SDM/WDM transmission over low-DMD 6-mode 19-core fibre across C+L band,” in 2017 European Conference on Optical Communication (ECOC). IEEE, 2017, pp. 1–3.
[7] K.-P. Ho and J. M. Kahn, “Mode-dependent loss and gain: statistics and effect on mode-division multiplexing,” Optics express, 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,” Optics express, vol. 19, no. 17, pp. 16680–16696, 2011.
[9] D. A. A. Mello, H. Srinivas, K. Choutagunta et al., “Impact of polarization- and mode-dependent gain on the capacity of ultra-long-haul systems,” Journal of Lightwave Technology, vol. 38, no. 2, pp. 303–318, 2020.
[10] J. van Weerdenburg, R. Ryf, J. C. Alvarado-Zacarias et al., “138 Tb/s transmission over 650 km graded-index 6-mode fiber,” in 2017 European Conference on Optical Communication (ECOC). IEEE, 2017, pp. 1–3.
[11] G. Rademacher, B. J. Puttnam, R. S. Luís et al., “10.66 Peta-Bit/s transmission over a 38-core-three-mode fiber,” in Optical Fiber Communication Conference. Optical Society of America, 2020, pp. Th3H–1.
R. S. Ospina, C. Okonkwo, and D. A. Mello, “DSP-based mode-dependent loss and gain estimation in coupled SDM transmission,” in *Optical Fiber Communication Conference*. Optical Society of America, 2020, pp. W2A–47.

M. S. Faruk and S. J. Savory, “Digital signal processing for coherent transceivers employing multilevel formats,” *Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1125–1141, 2017.

R. S. Ospina, M. Van den Hout, J. C. Alvarado-Zacarias et al., “Mode-dependent loss and gain estimation in SDM transmission based on MMSE equalizers,” *Journal of Lightwave Technology*, vol. 39, no. 7, pp. 1968–1975, 2020.

M. van den Hout, R. S. Ospina, S. van der Heide et al., “Experimental validation of MDL emulation and estimation techniques for SDM transmission systems,” in *2020 European Conference on Optical Communications (ECOC)*. IEEE, 2020, pp. 1–4.

R.-J. Essiambre, G. Kramer, P. J. Winzer et al., “Capacity limits of optical fiber networks,” *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, 2010.

W. Saif, M. A. Esmail, A. Ragheb et al., “Machine learning techniques for optical performance monitoring and modulation format identification: A survey,” *IEEE Communications Surveys & Tutorials*, 2020.

W. Saif, A. M. Ragheb, T. A. Alshawi et al., “Optical performance monitoring in mode division multiplexed optical networks,” *Journal of Lightwave Technology*, vol. 39, no. 2, pp. 491–504, 2020.

R. S. B. Ospina, M. van den Hout, S. van der Heide et al., “Neural-network-based MDG and Optical SNR Estimation in SDM Transmission,” 2021, arXiv preprint arXiv:2104.06803, 2021, accepted for presentation at OFC 2021.

N. Kim, Y. Lee, and H. Park, “Performance analysis of MIMO system with linear MMSE receiver,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4474–4478, 2008.

M. R. McKay, I. B. Collings, and A. M. Tulino, “Achievable sum rate of MIMO MMSE receivers: A general analytic framework,” *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 396–410, 2009.

T. Hayashi, Y. Tamura, T. Hasegawa et al., “Record-low spatial mode dispersion and ultra-low loss coupled multi-core fiber for ultra-long-haul transmission,” *Journal of Lightwave Technology*, vol. 35, no. 3, pp. 450–457, 2017.

X. Wautelet, C. Herzet, A. Dejonghe et al., “Comparison of embased algorithms for MIMO channel estimation,” *IEEE Transactions on communications*, vol. 55, no. 1, pp. 216–226, 2007.

A. Das and B. D. Rao, “SNR and noise variance estimation for MIMO systems,” *IEEE Transactions on Signal processing*, vol. 60, no. 8, pp. 3929–3941, 2012.

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.

P. Sillard, D. Molin, M. Bigot-Astruc et al., “50 μm multimode fibers for node division multiplexing,” *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1672–1677, 2016.