Impact of Polarization- and Mode-Dependent Gain on the Capacity of Ultra-Long-Haul Systems

Darli A. A. Mello, *Member, IEEE*, Hrishikesh Srinivas, *Student Member, IEEE*, Karthik Choutagunta, *Student Member, IEEE*, and Joseph M. Kahn, *Fellow, IEEE*

(Invited Paper)

***Abstract*—Motivated by recent interest in single-mode semiconductor optical amplifiers and multimode erbium-doped fiber amplifiers, we present a unified, comprehensive treatment of the effect of polarization- and mode-dependent gain (PDG and MDG) on the capacity of ultra-long-haul transmission systems. We study the problem using simulations of a multisection model including the effects of PDG or MDG and polarization mode dispersion (PMD) or modal dispersion. We also analytically derive exact expressions for the capacity distribution of PDG-impaired single-mode systems. In agreement with previous work, we find that PDG and MDG cause fluctuations in capacity, which reduce average capacity and may cause outage. We show that the multimode systems studied, with at least $D = 14$ spatial/polarization modes, have sufficient modal diversity and frequency diversity to strongly suppress capacity fluctuations and reduce outage probability, so that the outage capacity approaches the average capacity. We show that single-mode systems, by contrast, inherently provide low modal and frequency diversity, making them more prone to outage. To alleviate this problem, frequency diversity can be increased by artificially inserting PMD. Finally, we quantify the PDG/MDG requirements of optical amplifiers to ensure that the average capacity is close (within a 1-2 dB effective SNR loss) to the theoretical optimum. We show that these PDG/MDG requirements are stringent, especially considering the minimum-mean-square error linear equalizers implemented in typical multiple-input multiple-output receivers.**

I. Introduction

In optical communications systems, the multiple polarization and spatial modes guided by fibers can be subject to different gains in optical amplifiers. This effect is known as polarization-dependent gain (PDG) in single-mode fibers (SMFs), and mode-dependent gain (MDG) in multimode and coupled-core multicore fibers (CC-MCFs) [1]. PDG and MDG not only decrease the average channel capacity, but also can cause outages, in an effect similar to fading in wireless communications systems. The cumulative nature of PDG and MDG makes them especially critical in the scope of ultra-long-haul (ULH) optical systems that operate at transoceanic distances.

Current erbium-doped fiber amplifiers (EDFAs) for SMF achieve a very low gain difference between the two transmitted polarization modes, typically below 0.1 dB [2]. However, there is a renewed interest in semiconductor optical amplifiers (SOAs) in SMF systems, particularly at submarine distances [3], as a way to increase bandwidth and power efficiency [4][5][6]. One of the main challenges in deploying SOAs is their non-negligible polarization dependence, requiring careful PDG management across the link. This can be accomplished in several ways, e.g., by making use of polarization-diverse architectures, or by amplifying the signal in two passes within the same device, in forward and backward directions, with an added polarization rotation. Nevertheless, sometimes even so-called polarization-insensitive architectures exhibit MDG values (e.g. 0.5 dB in [7]) that can be excessively high for ULH optical links.

In the scope of space division multiplexing (SDM), CC-MCFs have demonstrated improved nonlinear tolerance [8] and a smaller group delay (GD) spread than conventional multimode fibers [9][8], reducing the MIMO digital signal processing (DSP) complexity. SDM has been proposed for use in submarine systems, which require a high level of integration in fibers and devices. In these systems, the performance and power consumption of optical amplifiers are of critical importance, and a major challenge for CC-MCF amplifiers is to minimize their MDG [1].

PDG and MDG have been extensively studied from a theoretical and experimental viewpoint. Winzer et al. demonstrated in [10] that MDG and mode dependent loss (MDL) have different impacts on the capacity of short-reach systems, but equivalent impacts on long-haul systems. Mecozzi and Shtaif investigated in [11] the distribution of polarization dependent loss (PDL) in optical systems. Shtaif addressed the impact of PDL on the system performance in [12], using analytical formulas. Nafta et al. quantified the capacity of systems affected by MDL using Monte-Carlo simulations [13]. Nelson et al. presented in [2] extensive field measurements of PDL. Winzer and Foschini investigated PDL and MDL in [14]. They performed numerical simulations of a narrowband channel, to which they fitted analytic expressions relating average and outage capacities to the number of modes and fiber spans. Their results illustrated the effect of modal diversity, in which relative losses in capacity decrease as the number of modes increases. The same effect was observed by Antonelli et al. in [15], which investigated the performance of different amplification schemes in the presence of MDL using simulations and analytical derivations. They showed that the capacity losses due to MDL depend on the operating regimes of optical

The authors are with the E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: darli@unicamp.br, hrishsri@stanford.edu, kchoutag@stanford.edu, jmk@ee.stanford.edu). D. A. A. Mello is also with the School of Electrical and Computer Engineering, State University of Campinas (UNICAMP), Campinas 13083-970, Brazil.

Darli Mello was supported by FAPESP grants 2018/14026-5, 2015/24341-7 and 2015/24517-8. This research was supported by Corning, Inc. and a Stanford Graduate Fellowship.

amplifiers, and that this dependency is particularly relevant in low-mode-count systems. Ho and Kahn provided in [1] analytic expressions for the channel capacity of MDL-impaired systems, and a proof that additive noise in mode-multiplexed signals is spatially white. The same authors investigated in [16] the effect of frequency diversity generated by modal dispersion (MD) in systems impaired by MDG. Owing to random, time-varying polarization or spatial mode coupling, PDG and MDG are random, time-varying effects, so the channel capacity varies randomly with time. When PMD or MD are present, PDG and MDG become frequency-dependent. If the channel bandwidth is sufficiently wide, frequency diversity causes the outage capacity to approach the average capacity, and causes the outage probability to approach zero [16]. Interestingly, the lower variability observed in the capacity of high-mode-count systems with high MD is also observed in their intensity impulse response, as demonstrated by Mecozzi et al. in [17]. Antonelli et al. assumed in [15] a collection of subcarriers whose MDG realizations are uncorrelated because of frequency diversity, but did not specifically address the values of modal dispersion required to induce a specific frequency diversity order over a specific WDM channel bandwidth. Experimental measurements of MDL in ULH multimode propagation (in multimode fibers and CC-MCFs) have been presented in [8][18][19][20].

In this paper, we review the concepts of channel capacity, modal diversity and frequency diversity, and present the following contributions to the understanding of the impact of PDG and MDG in ULH optical systems:

- We derive exact expressions for the cumulative distribution function (CDF) and probability density function (PDF) of capacity in narrowband PDG-impaired SMF systems.
- The traditional capacity expressions for MDG-impaired systems assume an optimal receiver whose performance can only be attained by nonlinear structures. However, practical systems usually resort to linear structures that attempt to minimize the mean-square error (MSE) [21][22]. To take this limitation into account, we include in our analysis the performance of minimum-MSE (MMSE) receivers.
- We show that a typical CC-MCF ULH link with a large number of propagating spatial and polarization modes has sufficient intrinsic modal diversity and frequency diversity to ensure that the outage capacity approaches the average capacity.
- In SMF systems with SOA amplification, however, there is low modal diversity, and the PMD of typical SMF does not provide sufficient frequency diversity to ensure the outage capacity approaches the average capacity. In agreement with [23], [24], we show that artificially inserting extra PMD in optical amplifier nodes increases frequency diversity and increases outage capacity, at the cost of increased complexity in the receiver adaptive equalizer.
- Finally, we use an average capacity metric and a related effective SNR loss metric to quantify the maximum tolerable per-amplifier PDG and MDG.

The remainder of this paper is structured as follows. Section II presents the multisection matrix model used to study the effects of PDG or MDG, PMD or MD and random mode coupling. Section III discusses the impact of PDG and MDG on the system capacity using optimal and MMSE receivers. Section IV describes the contributions of modal and frequency diversities in mitigating outage. Section V presents analytic and simulation results of two case studies, namely an SOA-based SMF link, and a seven-core CC-MCF link. Finally, Section VI concludes the paper.

II. Multisection Channel Model

In this section, we briefly review the multisection model to study the effects of PDG and MDG on optical transmission [1][25][26]. Consider a transmission link supporting D spatial and polarization modes (e.g., an SMF has $D = 2$, corresponding to one spatial mode with two polarization modes). Throughout this paper, we assume that the link comprises K independent spans, where each span has transmission fiber of length L_{span} and contains a lumped amplifier at the end of the span. The total length of the link is $L = KL_{\text{span}}$.

The transfer matrix of the k^{th} span, where $1 \leq k \leq K$, at angular frequency ω, can be expressed as a singular value decomposition of a $D \times D$ matrix [27]

$$\mathbf{H}^{(k)}(\omega) = \mathbf{V}^{(k)}\mathbf{\Lambda}^{(k)}(\omega)\mathbf{U}^{(k)H}, \tag{1}$$

where $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$ are $D \times D$ random unitary matrices that describe random mode coupling. The superscript $(\cdot)^H$ denotes matrix Hermitian conjugate. Assuming the channel of interest is centered at $\omega = 0$, without loss of generality, the diagonal matrix $\mathbf{\Lambda}^{(k)}(\omega)$ accounting for PDG or MDG and PMD or MD becomes

$$\mathbf{\Lambda}^{(k)}(\omega) = \text{diag}\left[\exp\left(\mathbf{g}^{(k)}/2 - j\omega\boldsymbol{\tau}^{(k)}\right)\right], \tag{2}$$

where $\mathbf{g}^{(k)}$ is a vector of uncoupled modal gains in log-power-gain units and $\boldsymbol{\tau}^{(k)}$ is a vector of uncoupled modal GDs. For notational convenience, $\exp(\cdot)$ denotes element-wise exponential, and $\text{diag}[\cdot]$ denotes a diagonal matrix formed by placing the argument vector on the main diagonal of a matrix of zeros.

In modeling each span using (1) and (2), we assume that GD accumulates in a distributed fashion and all modes are strongly coupled within each span. We model the uncoupled modal delay vector $\boldsymbol{\tau}^{(k)}$ of each span as an independent random vector whose elements have standard deviation (STD) $\sigma_\tau \propto \sqrt{L_{\text{span}}}$ (in seconds). We assume that MDG is a lumped effect that originates in the amplifier at the end of each span, and so we choose the uncoupled modal gain vectors $\mathbf{g}^{(k)}$ to be independent random vectors whose elements have STD σ_g (in log-power-gain units). This is consistent with the physical layout of the link, as typical SMFs and CC-MCFs are long dispersive waveguides with strong mode coupling and negligible MDG, whereas typical amplifiers have negligible GD dispersion and non-negligible MDG. The uncoupled modal gain and modal delay vectors satisfy the mode-averaged

constraints $\sum_i \mathbf{g}_i^{(k)} = 0$ and $\sum_i \boldsymbol{\tau}_i^{(k)} = 0$. Finally, in the SMFs and CC-MCFs we model, all of the modes are coupled after each span, and so $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$ are random $D \times D$ unitary matrices. In simulations, they can be generated by applying the Gram-Schmidt process to matrices constructed from D random complex Gaussian vectors [1].

The overall transfer matrix of the link is given by

$$\mathbf{H}^{(t)}(\omega) = \prod_{k=1}^{K} \mathbf{H}^{(k)}(\omega), \tag{3}$$

$$= \mathbf{V}(\omega)\boldsymbol{\Lambda}(\omega)\mathbf{U}^{H}(\omega), \tag{4}$$

where the superscript $(\cdot)^{(t)}$ denotes the total effect formed by multiplying together the individual transfer matrices of each span. The number of spans K is very large in ULH systems, which causes the overall PMD or MD and the overall PDG or MDG to be in the strong coupling regime. The columns of the unitary matrices $\mathbf{U}(\omega)$ and $\mathbf{V}(\omega)$ in (4) are the input and output Schmidt modes at frequency ω, respectively. The diagonal matrix elements of $\boldsymbol{\Lambda}(\omega)$ contain information about the overall coupled GDs and coupled modal gains. The overall coupled GDs $\boldsymbol{\tau}^{(t)} = \left[\boldsymbol{\tau}_1^{(t)}, \ldots, \boldsymbol{\tau}_D^{(t)}\right]$ are given by the eigenvalues of a GD operator defined as

$$-j\mathbf{H}^{(t)H}(\omega)\partial\mathbf{H}^{(t)}(\omega)/\partial\omega, \tag{5}$$

and have STD σ_{gd}. The overall coupled modal gains $\mathbf{g}^{(t)} = \left[\mathbf{g}_1^{(t)}, \ldots, \mathbf{g}_D^{(t)}\right]$ are given (in log-power-gain units) by the logarithm of the eigenvalues of an MDG operator defined as

$$\mathbf{H}^{(t)}(\omega)\mathbf{H}^{(t)}(\omega)^{H}, \tag{6}$$

and have STD σ_{mdg}. In the strong coupling regime, the statistics of PMD or MD and PDG or MDG are completely characterized by the STD of the overall coupled GDs σ_{gd}, and the STD of the overall coupled modal gains σ_{mdg}, respectively. For a detailed treatment of the statistics of PMD or MD and PDG or MDG, we refer the reader to [1], [25], [28].

A fair comparison among channels corrupted by PDG or MDG requires the total optical power to be conserved from the channel input to its output, which is consistent with the physics of amplified optical links. In the model, this corresponds to normalizing the transfer matrix by the square root of the mean of the overall coupled modal power gains, considering the cascade of K spans:

$$\mathbf{H}(\omega) = \frac{\mathbf{H}^{(t)}(\omega)}{\sqrt{\frac{1}{D}\sum_{i=1}^{D}\mathbb{E}\{e^{g_i^{(t)}}\}}}, \tag{7}$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation operator over ensembles of realizations of channel $\mathbf{H}^{(t)}(\omega)$. The normalization factor in (7) is discussed in more detail in Appendix A[1]. This normalization factor is compatible with amplification schemes 1 and 4 discussed in [15], where amplifiers are operated respectively to maintain the overall output power at the end of the link, or constant average power gain per amplifier, considering the entire WDM spectrum. As pointed out in [15], this assumption is justified in ULH links, where MD or PMD provides sufficient frequency diversity across multiple WDM channels[2].

The capacity of MDG-limited channels is studied in the following section after normalizing the channels using (7).

III. Capacity of mode-dependent-gain-limited channels

A. Instantaneous capacity

The information-theoretical capacity per unit bandwidth of an ideal communication channel with D dimensions (in our case, modes), in the absence of PDG or MDG is given in units of b/s/Hz by

$$C = D\log_2\left(1 + \frac{\mathrm{SNR}}{D}\right), \tag{8}$$

where the signal-to-noise ratio (SNR) is defined as the ratio of total signal power (considering all modes), to the noise power per mode.

Before calculating the capacity of a system with PDG or MDG, it is necessary to determine if the transmitter has channel state information (CSI). The availability of transmitter CSI requires channel feedback, which is difficult to implement in ULH systems because of their high latency. For this reason, all derivations in this paper assume no transmitter CSI. In long-distance optical transmission, the PDG or MDG generated in optical amplifiers accumulates along the link and reduces capacity. Assuming a narrowband signal (and omitting the dependence on ω), the capacity of a channel realization with transfer matrix $\mathbf{H}$ can be derived from the canonical expression for MIMO channels [29]

$$C_{\mathrm{mdg}} = \log_2\left[\det\left(\mathbf{I} + \frac{\mathrm{SNR}}{D}\mathbf{H}\mathbf{H}^{H}\right)\right], \tag{9}$$

where $\mathbf{I}$ is a $D \times D$ identity matrix.

Alternatively, C_{mdg} can also be obtained from the D eigenvalues λ_i of $\mathbf{H}\mathbf{H}^{H}$ as [29]

$$C_{\mathrm{mdg}} = \sum_{i=1}^{D}\log_2\left(1 + \frac{\mathrm{SNR}}{D}\lambda_i\right). \tag{10}$$

Equation (10) enables the calculation of the MIMO channel capacity as the sum of the capacities of single input single output (SISO) channels with $\mathrm{SNR}_i = \mathrm{SNR}\cdot\lambda_i, 1 \le i \le D$.

B. Average capacity and effective SNR loss

For a cascade of K identical spans with per-amplifier STD MDG σ_g, the accumulated MDG ξ scales as $\xi = \sqrt{K}\sigma_g$ [1]. The STD of the modal gains at the end of the link (equal to

[1]Notice that, differently from [1], in this work we include the normalization factor inside the multisection model, rather than in the formulas for capacity calculation.

[2]Notice that amplification schemes 2 and 3 studied in [15], which do not involve ensemble averages, imply per-channel power equalization schemes with tracking speeds faster than polarization and mode-coupling perturbations. The implementation of these schemes may become challenging in systems with rapid perturbations.

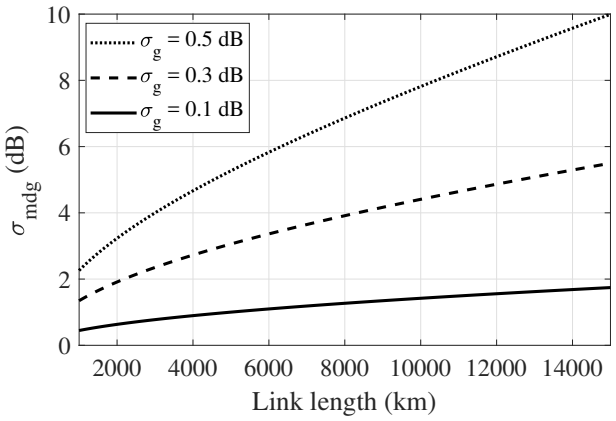

Fig. 1. Evolution of the STD of the overall link MDG σ_{mdg} with distance, for various values of per-amplifier MDG STD σ_g. The span length is assumed to be 50 km.

the overall RMS MDG since the modal gains sum to zero) is given by [1]

$$\sigma_{\mathrm{mdg}} = \xi\sqrt{1+\frac{\xi^2}{12(1-D^{-2})}} \approx \xi\sqrt{1+\frac{\xi^2}{12}}. \quad (11)$$

The approximation in (11), where σ_{mdg} is independent of D, is valid in the low-to-medium MDG regime [1], as in the practical scenarios considered in this paper. Fig. 1 shows the approximate σ_{mdg} accumulated over ULH links for three representative values of per-amplifier MDG STD σ_g, calculated according to (11). Typical SMF EDFAs have PDG STD σ_g below 0.1 dB [2]. The so-called polarization insensitive SOA presented in [7] has PDG STD σ_g below 0.5 dB. Recent realizations of SDM transmission with parallel SMF amplification demonstrate a wide range of per-span MDG STD σ_g, such as 0.36 dB (7-core CC-MCF transmission, [8]), 0.62 dB (6-mode few-mode fiber transmission, [20]) and 0.7 dB (3-mode few mode fiber transmission, [18]).

MDG of amplifiers is frequently characterized either in terms of the peak-to-peak or STD of the modal gains σ_g. While the peak-to-peak MDG is relevant for predicting the transmission performance of weakly coupled MDM links, σ_g and the number of amplifiers K are sufficient to calculate σ_{mdg} and the average channel capacity in ULH systems. Conversely, σ_{mdg} can be estimated from the receiver's adaptive equalizer and can be used to estimate σ_g, as in [8], [18], [20][3].

In channels with MDG and strong mode coupling, the average capacity assuming no transmitter CSI, $\bar{C}_{\mathrm{mdg}} = \mathbb{E}\{C_{\mathrm{mdg}}\}$, can be expressed analytically by [1]

$$\bar{C}_{\mathrm{mdg}} = D\int_{-\infty}^{\infty} \log_2\left[1+\frac{\chi}{D}\exp(\sigma_{\mathrm{mdg}}x)\right] p_D(x)dx, \quad (12)$$

where $p_D(x)$ is the probability density function (PDF) of the normalized log power gains[4]. The constant χ equals SNR/α^2, where α^2 is the expected value of the unordered linear power gains (see Appendix A)

$$\alpha^2 = \int_{-\infty}^{\infty} \exp(\sigma_{\mathrm{mdg}}x)p_D(x)dx. \quad (13)$$

Fig. 2(a) shows the average capacity of a system with MDG, normalized by the capacity of a system without MDG ($\bar{C}_{\mathrm{mdg}}/C$), as a function of the overall link PDG or MDG STD σ_{mdg}, for two extreme SNR values covering a wide range of ULH systems. As observed by Winzer and Foschini in [14], there is a weak dependence of the normalized capacity metric on the number of modes. Nevertheless, there is a noticeable dependence on the link SNR.

In this paper, we wish to find the requirements on σ_g to ensure that the impact of MDG on the system performance is below a certain threshold. We use an effective SNR loss as a metric, defined as

$$\Delta_{\mathrm{mdg}} = 10\log_{10}(\mathrm{SNR}) - 10\log_{10}\left(\mathrm{SNR}_{\mathrm{mdg}}\right), \quad (14)$$

where $\mathrm{SNR}_{\mathrm{mdg}}$ is the equivalent SNR of a system degraded by MDG, obtained by the inverse of (8):

$$\mathrm{SNR}_{\mathrm{mdg}} = D\left(2^{\bar{C}_{\mathrm{mdg}}/D}-1\right). \quad (15)$$

Other performance metrics could also be used, e.g., the normalized capacity ($\bar{C}_{\mathrm{mdg}}/C$) or an SNR penalty. While the normalized capacity provides a direct assessment of the system performance [14], optical system design usually relies on a reference SNR calculated based on the signal power, noise power, and nonlinear interference in the signal bandwidth. This reference SNR is the basis of capacity calculations. The effective SNR loss Δ_{mdg} is the decrease of the reference SNR that produces the same loss of capacity as the MDG causes. By contrast, the usual SNR penalty is the increase of the reference SNR in a system with MDG required to produce the same capacity as a system without MDG. Typical ULH and submarine constraints make it difficult to increase the SNR arbitrarily. The effective SNR loss is more appropriate than the usual SNR penalty in such systems, in which it is not possible to increase SNR arbitrarily, but it is possible to lower the code rate, in order to overcome performance degradation [30]. Thus, in the presence of MDG, $\mathrm{SNR}_{\mathrm{mdg}}$ should be used as the effective reference SNR.

In order for a system design to be advantageous, the PDG or MDG effective SNR loss should be constrained. Otherwise, eventual capacity gains obtained, for example, by increasing the nonlinear tolerance using CC-MCF transmission fibers, or increasing the system bandwidth using SOA-based amplification, will be diminished. If we assume that the effective SNR loss is limited to 1 dB, Fig. 2(b) indicates a maximum tolerable σ_{mdg} for the link between 3 dB (at SNR = 20 dB) and 4 dB (at SNR = 5 dB). According to Fig. 2(a), such a range of σ_{mdg} values corresponds to normalized capacities between 0.95 (at SNR = 20 dB) and 0.86 (at SNR = 5 dB). Referring to Fig. 1, this implies that the link length would be limited by MDG to approximately 3,000 km for $\sigma_g = 0.5$ dB and to 8,000 km for $\sigma_g = 0.3$ dB. Assuming that the effective SNR loss is limited to 2 dB would indicate a maximum tolerable σ_{mdg} for the link

[3]For a further discussion about the differences between peak-to-peak and STD MDG, including a treatment of low-complexity MDG characterization methods, please refer to [26].

[4]Analytic expressions for $p_D(x)$ are provided in [1].

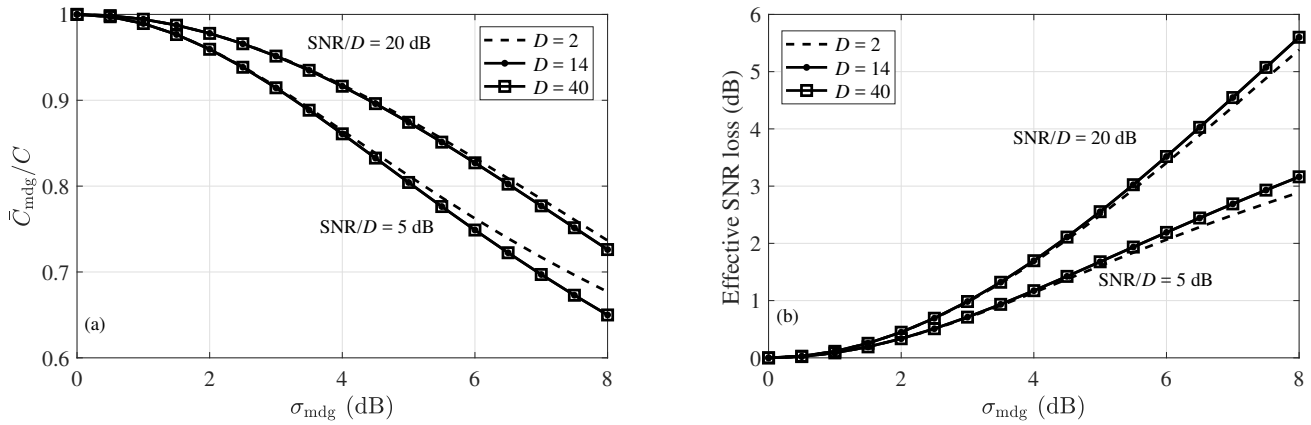

Fig. 2. (a) Normalized capacity and (b) effective SNR loss (14) of systems using optimal receivers, for $D = 2, 14$ or 40 modes, as a function of the overall link PDG or MDG STD σ_{mdg}.

between 4.5 (at SNR = 20 dB) and 6 dB (at SNR = 5 dB), corresponding to normalized capacities between 0.9 (at SNR = 20 dB) and 0.75 (at SNR = 5 dB).

C. Outage capacity

The variability of the instantaneous channel capacity is addressed in communications theory using an outage capacity metric. The outage capacity $C_{\mathrm{out},q}$ is defined for outage probability q as the information rate that is exceeded in a fraction $(1-q)$ of the channel realizations, i.e., $P(C \leq C_{\mathrm{out},q}) = q$ [29]. For a given outage probability, the lower the spread of the instantaneous capacity around its mean, the greater the outage capacity. For a fixed outage probability, as the variability of the instantaneous channel capacity is reduced, the outage capacity more closely approaches the average capacity.

A relevant issue is to determine how the outage capacity scales as the number of modes is increased. The calculation of the outage capacity requires the CDF of the instantaneous capacity, which can be obtained by analysis or by simulation. To the best of our knowledge, all previous publications address this problem using simulations [16][13], or by fitting simulated data to parameterized functions [14]. Indeed, finding the exact analytical forms of the instantaneous capacity PDF or CDF is difficult because the capacities of the different modes are statistically dependent random variables. We present in Appendix E an analytic derivation of the narrowband capacity CDF and PDF for $D = 2$ modes. For $D > 2$ modes, we resort to Monte-Carlo simulations.

D. Minimum capacity

The instantaneous capacity of a system impaired by a certain amount of PDG or MDG is lower-bounded by a minimum capacity value given by (see Appendix F)

$$C_{\mathrm{mdg}}^{\min} = D \log_2 \left(1 + \frac{\mathrm{SNR}}{D\alpha^2}\right). \tag{16}$$

Since the minimum capacity is a lower bound on the instantaneous capacity, (16) provides an analytic lower bound on the outage capacity.

Fig. 3(a) presents the PDFs of the instantaneous capacity C_{mdg} for 2-, 6-, and 14-mode systems using optimal receivers. The analytic PDF for $D = 2$ is derived in Appendix E, while the 6- and 14-mode empirical PDFs correspond to 100,000 realizations of 300 spans, with per-amplifier MDG STD $\sigma_g = 0.3$ dB. The SNR at the end of the link per mode SNR/D is set to 6.2 dB. The average ($\bar{C}_{\mathrm{mdg}}$) and minimum ($C_{\mathrm{mdg}}^{\min}$) capacities are shown as dot-dashed and dashed vertical lines, respectively.

In PDG-limited single-mode systems ($D = 2$), the minimum capacity provides a rule-of-thumb outage capacity, as it practically coincides with the outage capacity for outage probabilities $q \leq 0.001$. For increasing $D > 2$, the minimum capacity is seen from Fig. 3(a) to be an increasingly pessimistic lower bound for any practical outage probability, as it is only approached when all log power gains tend to zero.

E. MMSE receiver capacity

Equations (9) and (10) provide the capacity achievable by an optimal receiver for a certain channel realization described by transfer matrix $\mathbf{H}$. Practical systems, however, typically implement linear equalizers that attempt to minimize the MSE at their output. The constrained capacity, considering an MMSE receiver, can be obtained from the signal-to-noise-plus-interference ratio SINR_i for signal stream i [21][29][22]

$$\mathrm{SINR}_i = \frac{1}{\mathrm{MMSE}_i} - 1 = \frac{1}{\left[\left(\mathbf{I} + \frac{\mathrm{SNR}}{D}\mathbf{H}^H\mathbf{H}\right)^{-1}\right]_{i,i}} - 1, \tag{17}$$

where $[\cdot]_{i,i}$ indicates the i^{th} diagonal element. Then, the capacity of a narrowband system using an MMSE receiver, assuming no transmitter CSI, is given by

$$C_{\mathrm{mmse}} = \sum_{i=1}^{D} \log_2 \left(1 + \mathrm{SINR}_i\right). \tag{18}$$

The average constrained capacity, assuming an MMSE receiver, can then be calculated as $\bar{C}_{\mathrm{mmse}} = \mathbb{E}\{C_{\mathrm{mmse}}\}$.

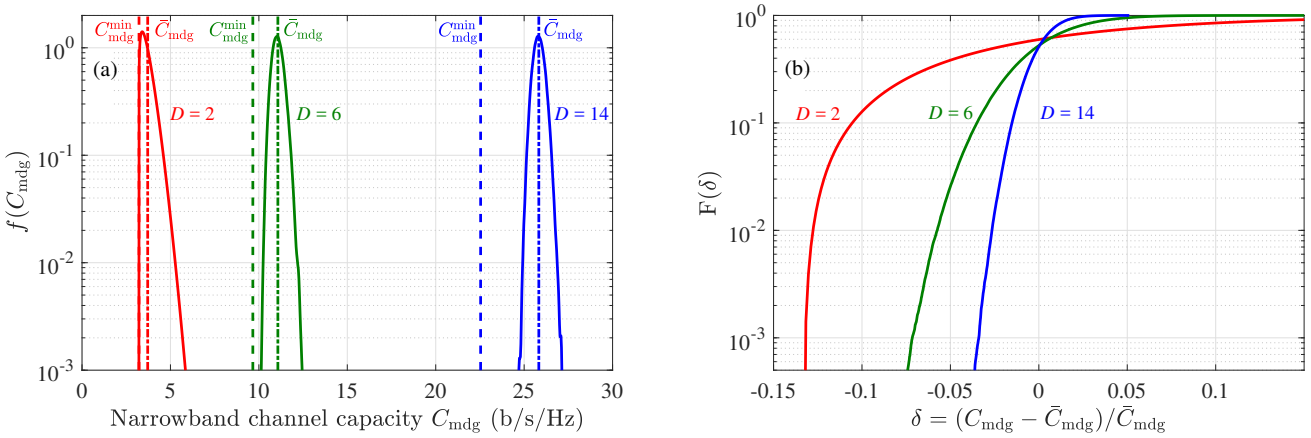

Fig. 3. PDFs (a) and CDFs (b) of the instantaneous capacity C_{mdg} of narrowband systems employing optimal receivers with $D = 2$ (red), $D = 6$ (green) and $D = 14$ (blue) modes. The distributions for $D = 2$ are obtained by the analytic expressions derived in Appendix E, in agreement with multisection simulation. The distributions for $D = 6$ and $D = 14$ are obtained from simulations with 100,000 realizations of the multisection model, each with 300 spans, with 0.3 dB per-amplifier MDG STD σ_g. The dot-dashed vertical lines indicate mean capacities $\bar{C}_{\mathrm{mdg}}$, while the dashed vertical lines indicate minimum capacities $C_{\mathrm{mdg}}^{\mathrm{min}}$, as labelled. The SNR per mode (SNR/D) at the end of the link is set to 6.2 dB.

IV. Modal Diversity and Frequency Diversity

While the average capacity, shown by the dot-dashed vertical lines in Fig. 3(a), scales quasi-linearly with the number of modes, the STD of the instantaneous capacity remains almost unchanged [31]. As a consequence, for a given STD of the overall link MDG σ_{mdg}, the ratio of the STD of capacity to the average capacity scales inversely with the number of modes. Accordingly, high-mode-count systems offer an intrinsic *modal diversity* that helps to minimize the probability of outage as indicated by the empirical CDF of the normalized instantaneous capacity loss $\delta = (C_{\mathrm{mdg}} - \bar{C}_{\mathrm{mdg}})/(\bar{C}_{\mathrm{mdg}})$, shown in Fig. 3(b). This effect was observed in [14] and modelled analytically in [15].

Thus far we have assumed narrowband signals for which the modal gains are constant over the channel bandwidth. In fiber transmission, however, the combined effect of MDG and MD generates a frequency-selective channel. Assuming that the channel bandwidth can be divided into N frequency bins, each having a constant gain, the instantaneous capacity is given by

$$C_{\mathrm{mdg}} = \frac{1}{N}\sum_{i=0}^{N-1} C_{\mathrm{mdg}}\left(\frac{2\pi(2i+1)B_{\mathrm{sig}}}{2N}\right), \qquad (19)$$

where B_{sig} is the total signal bandwidth, and $C_{\mathrm{mdg}}(\omega)$ is the capacity of a narrowband channel of bandwidth B_{sig}/N centered at ω. In systems with high MD or PMD, the channel frequency response varies rapidly between regions of low and deep fading, producing an averaging effect that reduces capacity fluctuations. This phenomenon is called *frequency diversity*, and was studied by Ho and Kahn in [16].

As an example, consider the frequency response of an SMF system with PDG, as shown in Fig. 4 [16]. Figs. 4(a,b) show the Schmidt mode gains computed from the MDG operator, in the regimes of low PMD (Fig. 4(a)) and high PMD (Fig. 4(b)). Over the frequency range plotted, the modal gains in linear units $e^{g_1(\omega)}$ and $e^{g_2(\omega)}$ (equivalent to $g_1(\omega)$ and $g_2(\omega)$ in log-power-gain units) vary slowly over frequency ω when the STD

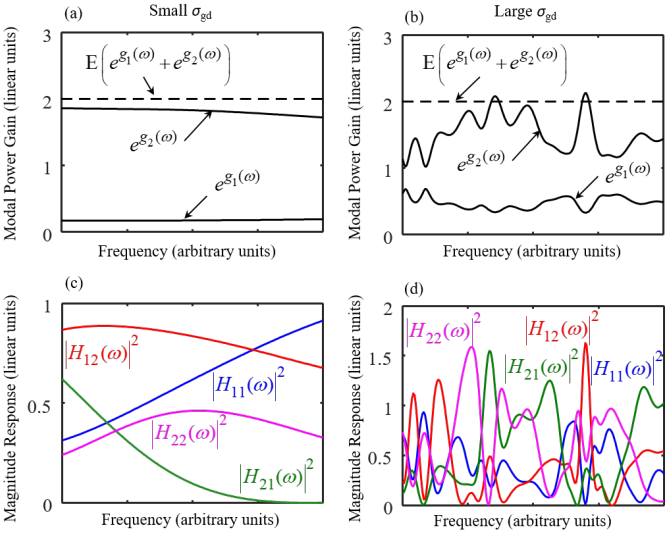

Fig. 4. Frequency dependence of the gains of the Schmidt modes in SMF for (a) small σ_{gd} and (b) large σ_{gd}. The Schmidt modes and their gains at each frequency are computed from the MDG operator after normalizing the transfer matrix, as in (7). The channel magnitude response of the SMF as a function of frequency for (c) small σ_{gd} and (d) large σ_{gd}.

of the overall coupled GDs σ_{gd} is small, and vary rapidly over ω when σ_{gd} is large. Figs. 4(c,d) show the frequency response of the four elements of a Jones matrix corresponding to an SMF link with PDG, also in the regimes of low PMD (Fig. 4(c)) and high PMD (Fig. 4(d)). It is interesting to note that, in the low-PMD regime, even though the modal gains are relatively flat over frequency, the magnitude responses of the Jones matrix elements exhibit larger fluctuations and show regions of deep fading. This effect is even more pronounced in the high-PMD regime.

As shown in [16], in a link affected by PDG or MDG the correlation between the channel gains at two frequency components spaced by $\Delta\omega$ depends on a normalized frequency

separation $\vartheta = \Delta\omega\sigma_{\rm gd}/(2\pi)$. As the frequency dependence of PDG or MDG is caused by PMD or MD, the coherence bandwidth of PDG or MDG should also be similar to that of PMD or MD, which is of the order of $1/\sigma_{\rm gd}$ [16]. Accordingly, it was observed in [16] that $\vartheta > 1$ is sufficient to guarantee a high decorrelation level for $D = 10$ spatial modes and an accumulated MDG of $\xi = 10$ dB, corresponding to an overall link MDG STD of $\sigma_{\rm mdg} = 12$ dB. Thus, a normalized bandwidth $b = B_{\rm sig}\sigma_{\rm gd} \gg 1$ ensures sufficient independent realizations of the channel capacity such that, as a consequence of the law of large numbers, the instantaneous capacity approaches the average capacity. On the other hand, if $b \ll 1$ the signal bandwidth lies within the MDG or PDG coherence bandwidth, and no frequency diversity gains are achieved.

Fig. 5 shows the normalized bandwidth b for a Nyquist-shaped 64-Gbaud channel with 0.1 roll-off factor, corresponding to a total bandwidth $B_{\rm sig} = 70.4$ GHz. The figure illustrates how b scales with link length, for an SMF and for a CC-MCF. The SMF has a GD STD $\sigma_\tau/\sqrt{L_{\rm span}}$ (equal to approximately 0.42 times the mean differential GD) of 0.05 $\mathrm{ps}/\sqrt{\mathrm{km}}$, slightly higher than commercial SMFs for ULH submarine transmission [32]. The CC-MCF has GD STD $\sigma_\tau/\sqrt{L_{\rm span}}$ of 3.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$, which is a low value for CC-MCFs [33][8].

Fig. 5 shows that the ULH SDM system using CC-MCFs guarantees that $b \gg 1$ for link lengths over 2,000 km, indicating an intrinsic frequency diversity due to MD. This observation suggests a note of caution for SDM fiber design. While one of the major goals of CC-MCF design is to reduce MD and therefore mitigate the required MIMO signal processing complexity at the receiver, excessively reducing the MD of CC-MCFs would lead to reduced frequency diversity, thereby increasing the time-variability of channel capacity and thus the probability of outage.

Fig. 5 also shows that the PMD intrinsic to standard SMFs is often insufficient to provide adequate frequency diversity, as $b < 1$ for all link lengths considered. Even so, it is possible to artificially insert additional PMD in fiber amplifiers to achieve diversity and thereby increase the outage capacity, as proposed in [23], [24]. In the example of Fig. 5, a per-span GD STD $\sigma_\tau = 3.1\sqrt{50} \approx 22$ ps would be required to achieve the same level of frequency diversity as that of CC-MCFs. The extra PMD can be obtained, e.g., by adding differential path length delays between the two polarization modes in the amplifiers. The differential delays should vary between amplifiers to avoid frequency response replication within the channel bandwidth.

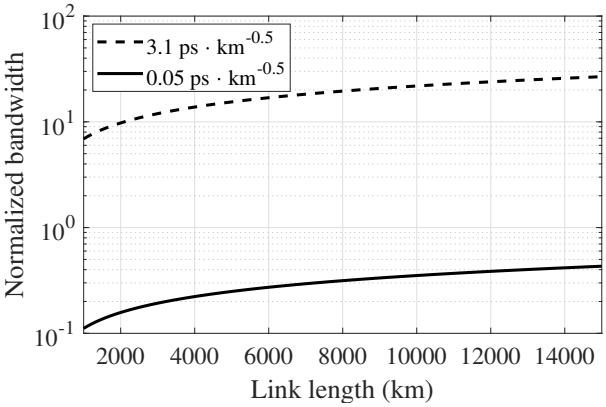

Fig. 5. Normalized bandwidth $b = B_{\rm sig}\sigma_{\rm gd}$ for a Nyquist-shaped 64-Gbaud signal with 0.1 roll-off factor, corresponding to a total bandwidth $B_{\rm sig} = 70.4$ GHz. The GD STD $\sigma_\tau/\sqrt{L_{\rm span}}$ was set to 3.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$ (corresponding to a CC-MCF), and 0.05 $\mathrm{ps}/\sqrt{\mathrm{km}}$ (corresponding to an SMF). The span length was assumed to be 50 km.

TABLE I
SIMULATION PARAMETERS

Parameter	Symbol/Description	Value
Span length	$L_{\rm span}$	50 km
Symbol rate	R_s or $1/T_s$	64 GBaud
Roll off factor	α	0.1
Signal Bandwidth	$B_{\rm sig}$	70.4 GHz
Number of spatial and polarization modes	D	2 (SMF)
Polarization-mode dispersion	$\sigma_\tau/\sqrt{L_{\rm span}}$	0.05 $\mathrm{ps}/\sqrt{\mathrm{km}}$
Number of spatial and polarization modes	D	14 (CC-MCF)
Modal dispersion	$\sigma_\tau/\sqrt{L_{\rm span}}$	3.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$

V. Simulation Results

A. Simulation setup

The simulation setup used in this paper is based on two illustrative scenarios of ULH transmission, in which PDG or MDG is likely to become a limiting effect. The first one is the transmission over a seven-core CC-MCF supporting 14 modes ($D = 14$), corresponding to 7 spatial modes having 2 polarization modes each. The second one is the transmission over an SMF, in which the optical signal is amplified by SOAs, leading to PDG accumulation. The main parameter of the transmission fibers used in the simulations is the STD of GD of the CC-MCF and SMF (see Table I).

Assessing the impact of PDG and MDG on the system performance also requires a scaling rule for obtaining the SNR at the receiver. In this paper, we evaluate the PDG and MDG amplifier tolerance using the SNR per mode (SNR/D) curves shown in Fig. 6, which correspond to different propagation regimes.

Using the GN-model for nonlinear propagation [34][35] (see Appendix B), the SNR at the end of a link of K spans[5] is given by

$$\mathrm{SNR}_{\rm nli}(K) = \frac{\mathrm{SNR}_{\rm nli}(1)}{K^{(1+\epsilon/3)}}, \tag{20}$$

where $\mathrm{SNR}_{\rm nli}(1)$ is the NLI-limited SNR at the end of the first span, and ϵ is a parameter related to noise accumulation over multiple spans. In the incoherent noise accumulation regime, the nonlinear noise generated is statistically independent from span to span, leading to $\epsilon = 0$. In the coherent noise accumulation regime $\epsilon \neq 0$. In an SMF uncompensated WDM optical system with a fully loaded spectrum, ϵ has

[5]Note that SNR(K) is the receiver SNR after a link of K spans, and not the SNR after the $K^{\rm th}$ span of a longer link.

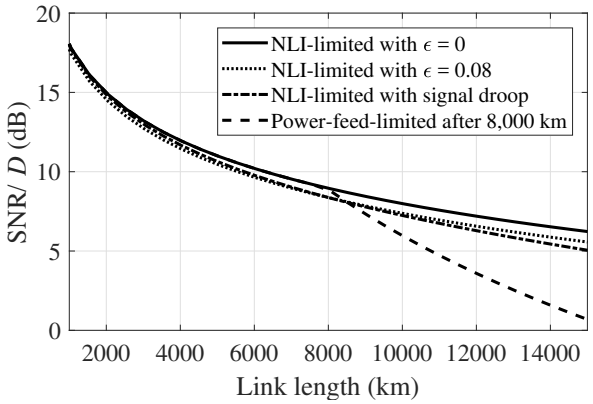

Fig. 6. Receiver SNR per mode (SNR/D) for a system limited by nonlinear interference (NLI) or power feed (PF) constraints. The span length is assumed to be 50 km. In the NLI-limited cases, $\mathrm{SNR}_{\mathrm{nli}}(1)/D$ is set to achieve $\mathrm{SNR}/D = 14$ dB at 2,500 km, in accordance with the Q^2 values for QPSK obtained in [8], and $\mathrm{SNR}_{\mathrm{nli,sd}}(1)$ is set equal to $\mathrm{SNR}_{\mathrm{nli}}(1)$. In the power-feed-limited case, $\mathrm{SNR}_{\mathrm{pf}}(1)/D$ is adjusted to make PF limitations the dominant effect after 8,000 km.

been shown to be quite small, with values ranging from 0.03 to 0.08 [35]. The modeling of nonlinearities in CC-MCF systems, including their accumulation with the number of spans, is still the subject of active research [36]. However, Fig. 2 indicates that, for relatively low values of MDG, the effective SNR loss exhibits a weak dependence on the link SNR. Therefore, in this paper, instead of delving into the intricacies of multimode nonlinear fiber propagation, we make the simplifying assumption that the NLI accumulates as in SMF systems, and evaluate the results for two values of ϵ corresponding to incoherent noise accumulation ($\epsilon = 0$) and coherent noise accumulation ($\epsilon = 0.08$). Uncompensated ULH multimode fiber propagation of a fully loaded spectrum, for instance, should lead to a low degree of coherence.

The second propagation regime considered in this paper assumes the effect of signal droop. In ULH transmission with saturated amplifiers, the accumulated ASE noise power saturates the amplifier gain, leading to a decrease in signal power, and a faster decay in SNR (see Appendix C). In this case, the SNR at the end of the link is given by

$$\mathrm{SNR}_{\mathrm{nli,sd}}(K) = \frac{\mathrm{SNR}_{\mathrm{nli,sd}}(1)+1}{K^{(1+\epsilon/3)}} - 1, \qquad (21)$$

where $\mathrm{SNR}_{\mathrm{nli,sd}}(1)$ is the NLI-limited SNR with signal droop at the end of the first span. Finally, the third regime assumes that links longer than a certain distance become limited by power-feed (PF) constraints (see Appendix D). In these systems nonlinearities are neglected, and the SNR at the end of the link is approximated by

$$\mathrm{SNR}_{\mathrm{pf}}(K) = \frac{\mathrm{SNR}_{\mathrm{pf}}(1)}{K^3}, \qquad (22)$$

where $\mathrm{SNR}_{\mathrm{pf}}(1)$ is the power-feed-limited SNR at the end of the first span. In practice, $\mathrm{SNR}_{\mathrm{nli}}(1) \ll \mathrm{SNR}_{\mathrm{pf}}(1)$, and only transoceanic submarine links with high fiber counts become power-feed-limited [3][37][38][39].

In Fig 6, for the NLI-limited cases, the SNR per mode $\mathrm{SNR}_{\mathrm{nli}}(1)/D$ is set to achieve $\mathrm{SNR}/D = 14$ dB at 2,500 km, in accordance with the Q^2 values for QPSK obtained in [8], and $\mathrm{SNR}_{\mathrm{nli,sd}}(1)$ is set equal to $\mathrm{SNR}_{\mathrm{nli}}(1)$. In the power-feed-limited case, $\mathrm{SNR}_{\mathrm{pf}}(1)/D$ is adjusted to make PF limitations the dominant effect after 8,000 km. In practice, the crossover point at which PF constraints become the limiting effect depends on several parameters, such as power-feed voltage, number of fibers, cable resistivity, and electrical-to-optical conversion efficiency.

B. Average and outage capacity

In this section, we study the scaling of average and outage channel capacities with link length for the cases of $D = 2$ (SMF) and $D = 14$ (CC-MCF in [8]). The STD of per-amplifier PDG or MDG σ_g is set to a modest 0.3 dB. The SNR at the end of the link is modeled assuming that the optical power is upper-limited by nonlinearities with incoherent accumulation (NLI with $\epsilon = 0$ in Fig. 6). As we will see later, this assumption provides a simple rule to make the analysis tractable, and has little impact on determination of the PDG or MDG amplifier tolerance.

Fig. 7 shows the scaling of the average and outage channel capacities measured in terms of spectral efficiency per spatial mode, considering both polarization modes. The capacity of the channel without MDG or PDG (dotted blue line) is calculated from the capacity formula in (8). This is the ideal maximum capacity that can be achieved assuming optimal coding at the transmitter and perfect equalization at the receiver. The average capacity of the PDG- or MDG-impaired channel using an optimal receiver is calculated analytically from (12) and (13) (dashed red line). In addition, the multisection model is applied to generate 10,000 realizations of 200 frequency bins equally spaced over the channel bandwidth. The generated channel realizations are used to obtain: (a) using (9), the channel average capacity assuming an optimal receiver (red squares); (b) using (18), the MMSE average capacity (solid black line with stars); (c) using empirical CDFs, the outage capacities corresponding to optimal (solid red line with open circles) and MMSE receivers (dashed black line with open squares), for an outage probability $q = 0.001$. For both $D = 2$ modes and $D = 14$ modes, the analytic results for the optimal receiver (dashed red line) show good agreement with those produced by the multisection model (red squares).

The results of the case study that uses the parameters of the seven-core CC-MCF ($D = 14$ modes) are shown in Fig. 7(a). We observe that this system configuration provides high modal and frequency diversity, such that the outage capacities approach the average capacities. At distances longer than 6,000 km, MDG reduces the achievable capacity by approximately 1 b/s/Hz/spatial mode. An MMSE receiver further lowers the capacity by roughly the same amount. For a given required capacity, e.g. 4 b/s/Hz, an MMSE receiver reduces the link length by approximately 3,000 km with respect to the optimal receiver. These results motivate interest in nonlinear equalizer structures, such as successive interference cancellation schemes [40][41].

The case study of an SMF link with SOA-based amplification ($D = 2$ modes), shown in Fig. 7(b), differs substantially

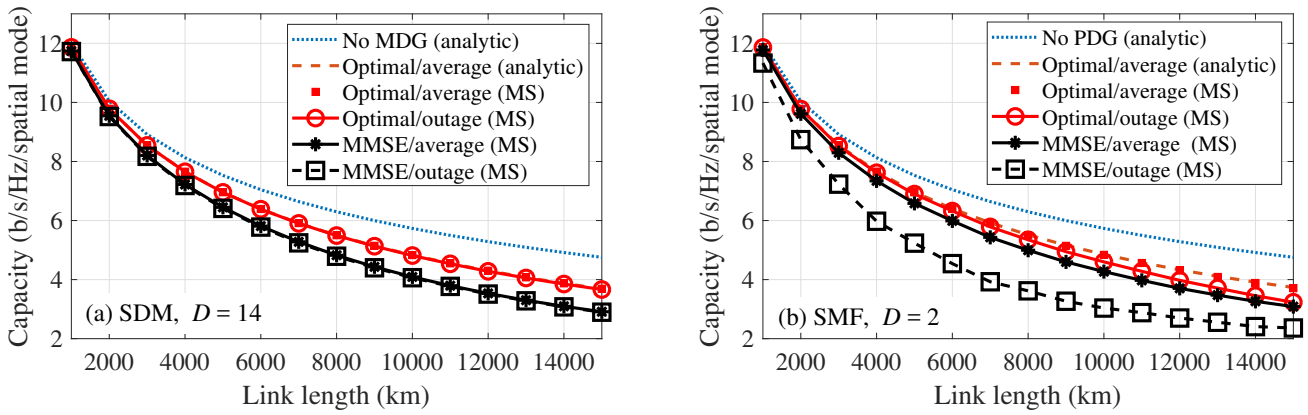

Fig. 7. Average and outage capacities per spatial mode, considering optimal and MMSE receivers for (a) SDM link with D = 14 modes and (b) SMF link with D = 2 modes. Outage capacity corresponds to outage probability $q = 0.001$. MS indicates results obtained using the multisection model with 10,000 realizations of 200 frequency bins. The per-amplifier MDG STD σ_g is assumed to be 0.3 dB, and the span length 50 km. The SNR at the end of the link is modeled assuming that the optical power is upper-limited by nonlinearities with incoherent accumulation (NLI-limited system with $\epsilon = 0$ in Fig. 6).

from the CC-MCF ($D = 14$ modes) case in terms of outage capacity. Here, the system configuration does not provide sufficient modal or frequency diversity, and the outage capacity is considerably lower than the average capacity. Let us, for example, set the required capacity at 4 b/s/Hz. In this case, using an optimal receiver, the maximum link length is reduced by approximately 1,000 km when the outage capacity is taken into account. The impact of outages are even more severe considering the use of MMSE receivers, for which the maximum length is reduced from approximately 11,000 km to only 7,000 km, as a consequence of a heavy tail of the MMSE receiver capacity distribution towards lower capacities. Setting the outage capacity at 4 b/s/Hz, the MMSE receiver reduces the link length by approximately 5,000 km with respect to the optimal receiver. These results indicate that in SOA-based ULH systems affected by PDG, advanced schemes such as successive interference cancellation can provide substantial reach and capacity gains. Furthermore, although PMD is intrinsically low, frequency diversity can be improved through the insertion of artificial PMD [23], [24].

Fig. 8(a) shows the empirical CDF $F(\delta)$ of the normalized capacity δ for links of length 1,000 km, 10,000 km and 15,000 km, for seven-core CC-MCF transmission, assuming an optimal receiver. The multisection model is applied to generate 10,000 channel realizations with 200 frequency bins each. It can be observed that for ULH transmission, the outage capacity for $q = 0.001$ deviates by a modest 1% from the average channel capacity. A comparison with the narrowband scenario for $D = 14$ and 15,000 km in Fig. 3(b) shows that frequency diversity contributes to reduce the capacity loss at $q = 0.001$ from 4% to 1%. The results change considerably for the SMF, as shown in Fig. 8(b). Here, the outage capacity for $q = 0.001$ and 15,000 km deviates 13% from the average capacity. A comparison with the narrowband channel results for $D = 2$ and 15,000 km in Fig. 3(b) reveals no frequency diversity gains in outage capacity. Fortunately, frequency diversity can be improved by the artificial insertion of PMD. A possible way to insert PMD in an SOA-based link is shown in Fig. 8(c). The lengths of the two delay elements in each span can be randomly chosen such that one polarization mode is delayed or advanced with respect to another. In this example, the delay elements in each span are chosen such that the differential delay $\tau_1 - \tau_2$ is uniformly distributed between -60 ps and 60 ps[6]. With this simple setup the capacity CDF improves considerably, and the outage capacity deviates just 6% from the average capacity, as shown in Fig. 8(d).

One disadvantage of the artificial insertion of PMD is the required increase in DSP equalizer length. Consider a 10,000 km SMF link (corresponding to $K = 200$ spans, $L_{\text{span}} = 50$ km) with an intrinsic GD STD $\sigma_\tau/\sqrt{L_{\text{span}}}$ of 0.05 ps/$\sqrt{\text{km}}$. The STD of per-span GD using this fiber is $\sigma_\tau = 0.35$ ps, and the STD of the overall link GD is $\sigma_{\text{gd}} = 5$ ps (0.35 ps $\times\sqrt{200}$), which is proportional to the peak-to-peak overall GD spread. PMD equalizers typically have length N_{PMD} that is some multiple of the peak-to-peak overall GD spread in order to guarantee having sufficient memory length. Considering a $T_s/2$ sampling time of approximately 7 ps, and assuming N_{PMD} is 5 times longer than σ_{gd} [42], the number of required equalizer taps without artificially added PMD is

$$N_{\text{PMD}} \text{ (no added PMD)} = \lceil 5 \times \sigma_{\text{gd}}/7 \text{ ps} \rceil = 4 \text{ taps.} \quad (23)$$

Now suppose that random additional PMD, with differential delay $\tau_1 - \tau_2$ uniformly distributed between -60 ps and 60 ps, is artificially added to each span of the SMF link. Then, the STD of per-span GD becomes $\sigma_\tau = 17.3$ ps, and the STD of the overall link GD becomes $\sigma_{\text{gd}} = 245$ ps (17.3 ps $\times\sqrt{200}$). Therefore, the number of required equalizer taps with artificially added PMD becomes

$$N_{\text{PMD}} \text{ (with added PMD)} = \lceil 5 \times \sigma_{\text{gd}}/7 \text{ ps} \rceil = 175 \text{ taps,} \quad (24)$$

which is a significant increase in equalizer complexity compared to the case without artificially added PMD.

[6]Note that inserting an identical delay in each amplifier node would generate a periodic channel frequency response, which leads to a lower frequency diversity order than a random channel frequency response.

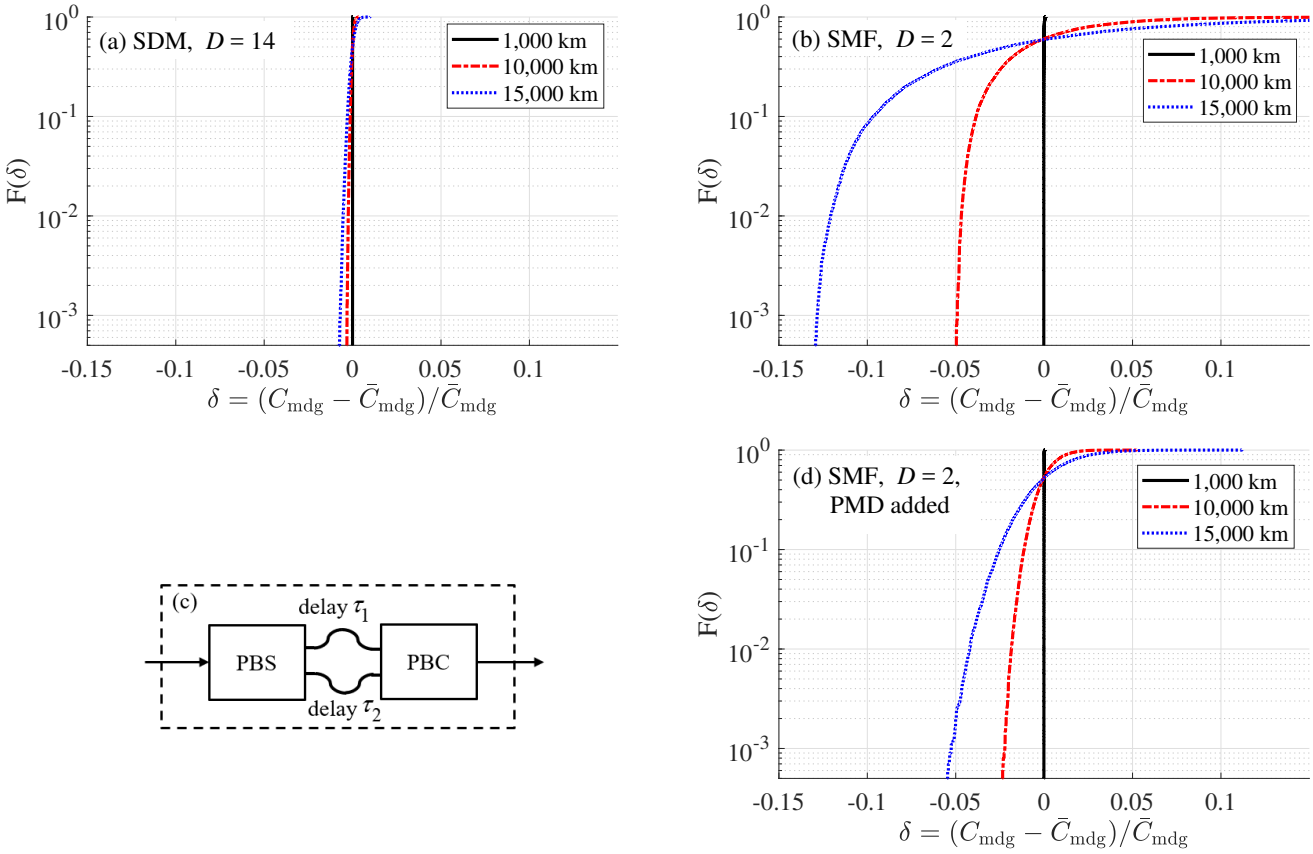

Fig. 8. Impact of PDG/MDG and PMD/MD on capacity of ULH links illustrating combined effects of modal diversity and frequency diversity. Optimal receivers are assumed. (a) CDF of the normalized instantaneous capacity of a CC-MCF link with $D = 14$ modes, as a function of δ, which is the difference between the instantaneous and average capacity, normalized to the average capacity. (b) CDF of the normalized instantaneous capacity of an SOA-based SMF link with $D = 2$ modes as a function of δ. (c) Artificial insertion of additional PMD using a polarizing beam splitter (PBS), random delay elements which introduce a uniformly distributed differential delay of $\tau_1 - \tau_2$ between the two polarization states, and a polarizing beam combiner (PBC). (d) CDF of the normalized instantaneous capacity as a function of δ using the same conditions as (b), and a delay element as in (c) in each span with differential delay $\tau_1 - \tau_2$ uniformly distributed between -60 ps and 60 ps. Parts (a), (b), (d) assume an outage probability $q = 0.001$. The CDFs are obtained from simulations with 10,000 realizations of the multisection model with 200 frequency bins each. The STD of per-amplifier MDG σ_g is assumed to be 0.3 dB, and the span length 50 km.

C. *Effective SNR loss*

In this section we discuss the effective SNR loss and its dependence on the SNR scaling rule. These results are based on the average capacity. We observed in the previous section that, although the outage capacity is an important concept when the channel is random, it actually approaches the average capacity for systems with sufficient modal and frequency diversity. In CC-MCF transmission, the outage capacity naturally approaches the average capacity because of the intrinsic modal and frequency diversity. In SMF systems with SOA amplification, frequency diversity can be achieved by the artificial insertion of PMD.

Fig. 9 shows the effective SNR loss caused by MDG on ULH links with $D = 14$ spatial and polarization modes, assuming optimal receivers. The results for other values of D are omitted, given the weak dependence of the effective SNR loss on the number of modes observed in Fig. 2. The simulations assume 50-km spans, and $\sigma_g = 0.2$ dB or $\sigma_g = 0.3$ dB. The capacity calculations are carried out using the analytic expression (12), and the effective SNR loss is calculated using (14) and (15). We used the four SNR scaling rules depicted in Fig. 6. All four curves agree for link lengths shorter than 8,000 km. For longer links, the effective SNR loss corresponding to the power-feed-limited case lies up to 0.5 dB below the other curves. This can be explained by observing Fig. 2, which, indeed, indicates that the effective SNR loss caused by MDG is smaller for links with low SNRs.

A relevant question is how much effective SNR loss caused by MDG can be tolerated in the system. Excessively high penalties may make the deployment of certain system designs impracticable. For example, PDG or MDG can eliminate potential capacity gains offered by the large bandwidth of SOAs, or by the increased NLI tolerance of multimode propagation. Suppose that the per-amplifier PDG or MDG STD is $\sigma_g = 0.3$ dB, and that the transmission system can tolerate a maximum of 1 dB effective SNR loss. In this case, Fig. 9 shows that the maximum reach of the SDM link is only approximately 6,000 km. This reach can be increased to approximately 14,000 km by reducing σ_g to 0.2 dB. Alternatively, a similar increase in reach can be obtained by keeping $\sigma_g = 0.3$ dB, and allowing an effective SNR loss higher than 2 dB. Analyses of this kind are useful to device and system design engineers because they

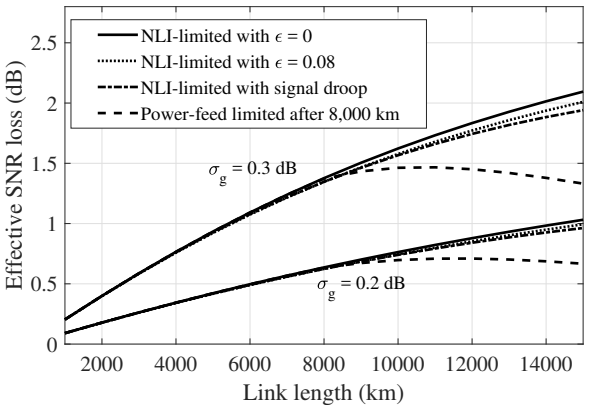

Fig. 9. Effective SNR loss obtained analytically using (12) and (14), for optimal receivers, assuming different SNR scaling rules, as explained in the text. The span length is assumed to be 50 km, and the number of modes D = 14. The results for other values of D are omitted, given the weak dependence of the effective SNR loss on the number of modes.

can provide the MDG requirements of optical amplifiers for given target link distances and tolerable effective SNR losses.

D. MDG tolerance

We calculate the per-amplifier MDG STD σ_g tolerance for different transmission configurations and for $D = 14$ modes. The results for $D = 2$ modes are similar, and are omitted from the figures for the sake of clarity. Fig. 10(a) shows the tolerable value of per-amplifier MDG STD σ_g for a maximum effective SNR loss of 1 dB, for optimal receivers, assuming different scaling rules, obtained by analytic calculation of (12). Clearly, the dependence of the σ_g tolerance on the SNR scaling rule is minimal. The worst-case variation among all rules is about 0.05 dB. For transoceanic distances, a value of σ_g below 0.2 dB is required. This requirement can increase to 0.3 dB if an effective SNR loss of 2 dB is allowed, as indicated in Fig. 10(b).

Given the weak dependence of the σ_g tolerance on the SNR scaling rule, Figs. 10(c) (for a maximum 1 dB effective SNR loss) and 10(d) (for a maximum 2 dB effective SNR loss) evaluate the impact of the receiver architecture (optimal or MMSE) for the scaling rule that assumes incoherent NLI limitation ($\epsilon = 0$). The optimal receiver results (filled circles) are obtained using effective SNR losses based on the analytic average capacity. The MMSE results (filled squares) are obtained by applying (17) and (18) to channel transfer functions obtained by multisection simulation. As a general rule, the use of an MMSE receiver instead of an optimal receiver further reduces the MDG and PDG tolerance by approximately 0.08 to 0.15 dB, depending on the link length. The results indicate that nonlinear receiver architectures based, e.g., on successive interference cancellation, may be used to help alleviate the PDG and MDG requirements on inline optical amplifiers. In addition, attaining low values of per-amplifier MDG STD may require adaptive control of the modal gains at each amplifier [43].

Finally, we investigate the sensitivity of calculating the σ_g tolerance based on the average capacity of the optimum receiver, instead of the outage capacity. We obtain the MDG tolerance based on an effective SNR loss calculated using the minimum capacity, providing a strict lower bound on the MDG tolerance for any outage capacity. With sufficient frequency and modal diversity, tolerances are well approximated by those computed on the basis of average capacity (filled circles), but at any outage probability, never decrease below those computed on the basis of the minimum capacity (open circles). A similar analysis can be carried out in the case of an MMSE receiver, but an analytical bound may not be available.

VI. Conclusion

This paper evaluates the impact of PDG and MDG on ULH optical communications. We address two representative case studies, namely SMF transmission with SOA-based amplification, and a seven-core CC-MCF transmission. We review the concepts of average and outage capacities, and their implications on the system design. We show that the CC-MCF setup exhibits sufficient modal and frequency diversity to ensure stable system operation. In this case, the outage capacity approaches the average capacity. In the SMF link, however, modal diversity is weak and frequency diversity is low. As a consequence, the outage capacity can be noticeably lower than the average capacity. Confirming the results in [23], [24], we show that this problem can be alleviated by the artificial insertion of PMD in the system, generated, e.g., using polarization beam splitters and combiners, and random delay elements. This additional PMD improves frequency diversity and increases the outage capacity. As a disadvantage, this solution increases the complexity of the adaptive equalizer in a coherent receiver. We show that the maximum tolerable per-amplifier PDG or MDG STD σ_g in typical ULH and submarine links using MMSE receivers is between 0.1 and 0.3 dB, for effective SNR losses between 1 and 2 dB. These stringent tolerance requirements may be relaxed by up to 0.1 dB using more advanced nonlinear receiver structures, such as those implementing successive interference cancellation.

Appendix A
Normalization factor in channel capacity calculations

Here we derive the normalization factor used in capacity calculations for a fair comparison between channels with varying numbers of modes and varying levels of MDG. The derivation is based on the assumption that, for a single fiber and amplifier span, the total output power across all modes and frequencies is equal to the total input power across all modes and frequencies. Therefore, this assumption is also valid for the overall fiber link, which is a cascade of such spans. The input-output relationship of a fiber link supporting D modes at each frequency ω can be modeled as

$$\mathbf{y}(\omega) = \frac{1}{\alpha(\omega)}\mathbf{V}(\omega)\mathbf{\Lambda}(\omega)\mathbf{U}^H(\omega)\mathbf{x}(\omega), \tag{25}$$

where $\mathbf{x}(\omega)$ and $\mathbf{y}(\omega)$ are D-dimensional vectors of input and output complex baseband electric field amplitudes, respectively, $\mathbf{V}(\omega)$ and $\mathbf{U}(\omega)$ are random unitary matrices that model

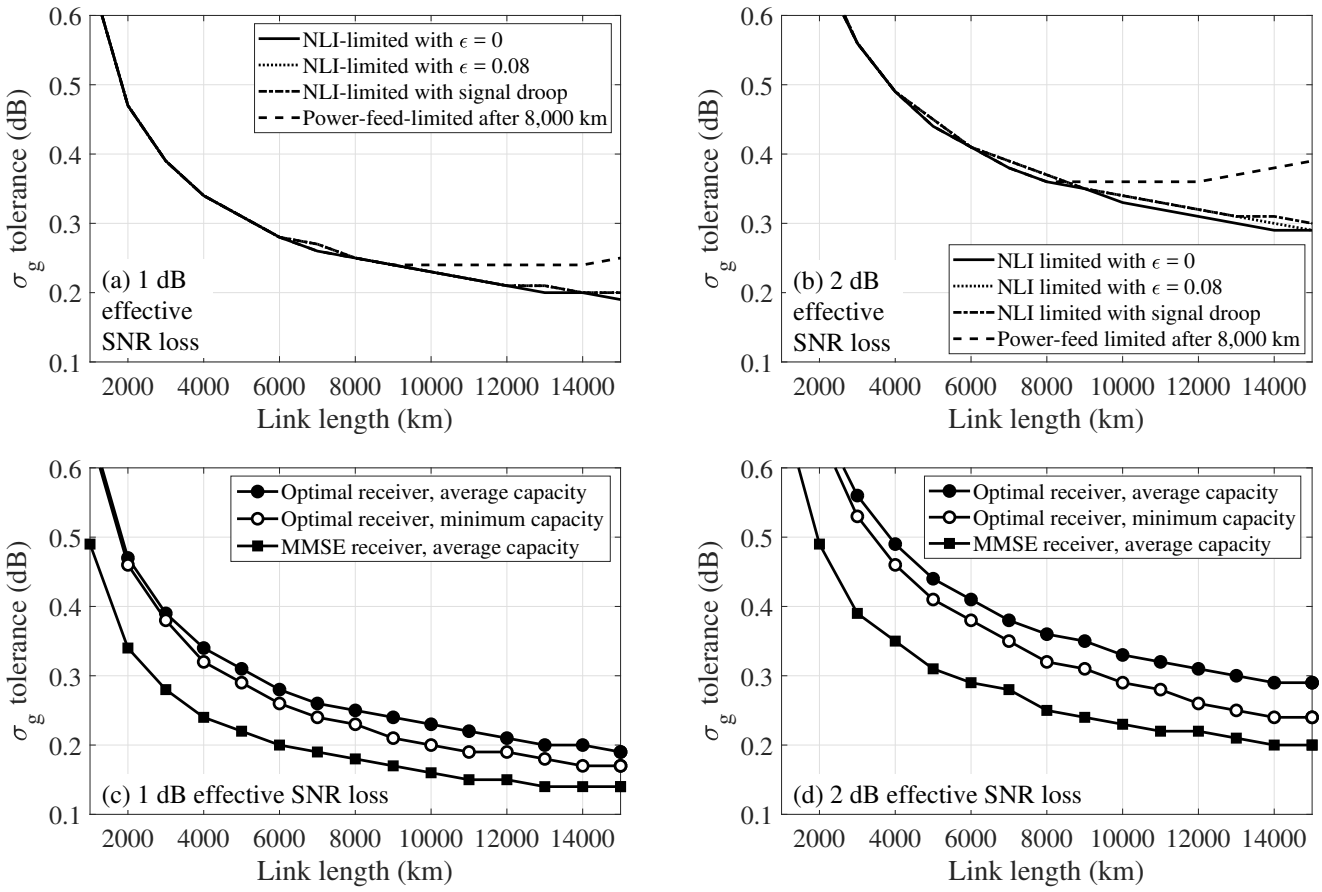

Fig. 10. Tolerable values of per-amplifier MDG STD σ_g assuming optimal receivers, considering various SNR scaling rules, for maximum effective SNR losses of (a) 1 dB and (b) 2 dB. The curves in (a) and (b) are obtained analytically using (12) and (14) for $D = 14$ modes, assuming a span length of 50 km. Tolerable values of σ_g for optimal and MMSE receivers, considering maximum effective SNR losses of (c) 1 dB and (d) 2 dB. The optimal receiver performances in (c) and (d) are obtained by the analytic formulas in Sections III-B and III-D for average and minimum capacities, respectively. The MMSE receiver performances in (c) and (d) are obtained by transfer matrices simulated using the multisection model in Section II, with a single realization of 20 frequency bins, applied to the capacity equation (18). The SNR scaling rule is chosen as NLI-limited with incoherent noise accumulation $\epsilon = 0$. The results are obtained for $D = 14$ modes, assuming a span length of 50 km. The results for other values of D are omitted, given the weak dependence of the effective SNR loss on the number of modes.

mode coupling in the channel, and $\mathbf{\Lambda}(\omega)$ is a diagonal matrix whose diagonal elements are the modal field gains $e^{g_i(\omega)/2}$ (in linear units). In (25), $\alpha(\omega)$ is the normalization factor, whose formula we will derive here.

The total output power of the amplifier at a fixed frequency ω with fixed transmitted signals $\mathbf{x}(\omega)$ across an ensemble average of the channel realizations is given by

$$\begin{aligned}\left\|\mathbf{y}(\omega)\right\|^2 &= \mathbb{E}\left[\frac{1}{\alpha^2(\omega)}\mathbf{x}(\omega)^H\mathbf{U}(\omega)\mathbf{\Lambda}^2(\omega)\mathbf{U}^H(\omega)\mathbf{x}(\omega)\right] \\ &= \frac{1}{\alpha^2(\omega)}\left\|\mathbf{x}(\omega)\right\|^2\left(\frac{1}{D}\sum_{i=1}^{D}\mathbb{E}\left\{e^{g_i(\omega)}\right\}\right), \quad (26)\end{aligned}$$

where the expectation is taken over random unitary matrices $\mathbf{U}$. Expression (26) uses the fact that $\mathbf{x}(\omega)$ is projected onto each Schmidt mode of the amplifier with equal probability, and therefore its squared norm is scaled by the expected mean of the squared singular values $e^{g_i(\omega)}$ given by the diagonal elements of $\mathbb{E}\left\{\mathbf{\Lambda}^2(\omega)\right\}$.

Suppose that the total input power across all frequencies in a bandwidth Ω is fixed, i.e. $\int_\Omega \left\|\mathbf{x}(\omega)\right\|^2 d\omega = P_{in}$. Integrating the output power (26) over Ω yields

$$\begin{aligned}P_{out} &= \int_\Omega \left\|\mathbf{y}(\omega)\right\|^2 d\omega \\ &= \int_\Omega \frac{1}{\alpha^2(\omega)}\left\|\mathbf{x}(\omega)\right\|^2\left(\frac{1}{D}\sum_{i=1}^{D}\mathbb{E}\left\{e^{g_i(\omega)}\right\}\right)d\omega, \quad (27)\end{aligned}$$

which means that we require the normalization factor to be

$$\alpha(\omega) = \sqrt{\frac{1}{D}\sum_{i=1}^{D}\mathbb{E}\left\{e^{g_i(\omega)}\right\}} \quad (28)$$

in order to satisfy the condition $P_{out} = P_{in}$. Assuming that the PDG and MDG statistics are independent of frequency, the normalization factor can be given as the expected value of the unordered linear power gains, e^g, as

$$\alpha^2(\omega) = \alpha^2 = \int_{-\infty}^{\infty}\exp(g)p(g)dg \quad (29)$$

where $p(g)$ is the probability density function of the log power gains g. Analytic expressions for $p(g)$ for channels with strong

mode coupling can be found in [1]. It is interesting to note that after normalizing the channel using (28), as in (7), the expected sum of the modal power gains equals the number of modes

$$\sum_{i=1}^{D} \mathbb{E}\left\{e^{g_i(\omega)}\right\} = D. \tag{30}$$

APPENDIX B
SNR EVOLUTION IN SYSTEMS LIMITED BY NONLINEAR INTERFERENCE

The capacity penalties imposed by MDG are related to the SNR at the end of the link. Optical systems typically operate in a regime in which the launched signal power is high enough to overcome additive noise, but low enough to ensure that nonlinear interference generated by the Kerr effect is kept below a certain threshold. Recent works have shown that this nonlinear interference can be well-modeled by additive white Gaussian noise in long dispersion-uncompensated links [34][35]. Thus, the total noise power can be decomposed into two main contributions, namely the noise generated by amplified spontaneous emission (ASE) in optical amplifiers, and the nonlinear interference (NLI) produced during fiber propagation. In this case, the optical SNR (OSNR) at the end of the link becomes [34][35]

$$\mathrm{OSNR} = \frac{P_{\mathrm{ch}}}{P_{\mathrm{ase}} + P_{\mathrm{nli}}}, \tag{31}$$

where P_{ch} is the optical signal launch power and P_{ase} and P_{nli} are the ASE and NLI noise powers in a specific reference bandwidth respectively. In general, for a polarization-multiplexed signal, the OSNR and the electrical SNR are related by [44]

$$\mathrm{OSNR} = \frac{R_s}{B_{\mathrm{ref}}}\mathrm{SNR}, \tag{32}$$

where R_s is the symbol rate and B_{ref} is the reference bandwidth used to quantify the OSNR. For the sake of simplicity, and without loss of generality, we assume in this paper that $R_s = B_{\mathrm{ref}}$, such that $\mathrm{OSNR} = \mathrm{SNR}$. Henceforth we will use SNR to designate both the electrical and optical signal-to-noise ratios.

The ASE noise power is known to accumulate linearly with the number of spans, i.e., $P_{\mathrm{ase}}(K) = KP_{\mathrm{ase-ss}}$, $P_{\mathrm{ase-ss}}$ being the ASE noise power added in a single span. The NLI power, in turn, increases quasi-linearly with the number of spans in SMF systems, according to $P_{\mathrm{nli}}(K) = K^{(1+\epsilon)}P_{\mathrm{nli-ss}}$, $P_{\mathrm{nli-ss}}$ being the NLI noise power generated in a single span [34]. Thus, the SNR at the end of a link of K spans is given by

$$\begin{aligned}\mathrm{SNR}_{\mathrm{nli}}(K, P_{\mathrm{ch}}) &= \frac{P_{\mathrm{ch}}}{KP_{\mathrm{ase-ss}} + K^{(1+\epsilon)}P_{\text{nli-ss}}} \\ &= \frac{P_{\mathrm{ch}}}{KP_{\mathrm{ase-ss}} + K^{(1+\epsilon)}\eta P_{\mathrm{ch}}^3},\end{aligned} \tag{33}$$

where η is a proportionality coefficient independent of K or P_{ch}. The optimal launch power per channel P_{opt} can be obtained by differentiating (33) with respect to P_{ch}, yielding [35]

$$P_{\mathrm{opt}} = \left(\frac{P_{\mathrm{ase-ss}}}{2\eta K^{\epsilon}}\right)^{\frac{1}{3}}. \tag{34}$$

Under the optimal launch power condition, the SNR at the end of a link of K spans is given by

$$\mathrm{SNR}_{\mathrm{nli}}(K) = \frac{1}{K}\frac{P_{\mathrm{opt}}}{1.5P_{\mathrm{ase-ss}}} = \frac{\mathrm{SNR}_{\mathrm{nli}}(1)}{K^{(1+\epsilon/3)}}, \tag{35}$$

where $\mathrm{SNR}_{\mathrm{nli}}(1)$ is the signal-to-noise ratio at the end of the first span.

APPENDIX C
SNR EVOLUTION IN SYSTEMS LIMITED BY NONLINEAR INTERFERENCE WITH SIGNAL DROOP

In certain ULH systems, in particular in submarine environments, the noise accumulated along the link contributes to saturation of the amplifiers, causing signal power depletion. This phenomenon, known as signal droop, leads to an evolution of SNR along the link as

$$\begin{aligned}\mathrm{SNR}_{\mathrm{nli,sd}}(K, P_{\mathrm{ch}}) &= \frac{P_{\mathrm{ch}} - (KP_{\mathrm{ase-ss}} + K^{(1+\epsilon)}\eta P_{\mathrm{ch}}^3)}{KP_{\mathrm{ase-ss}} + K^{(1+\epsilon)}\eta P_{\mathrm{ch}}^3} \\ &= \frac{P_{\mathrm{ch}}}{KP_{\mathrm{ase-ss}} + K^{(1+\epsilon)}\eta P_{\mathrm{ch}}^3} - 1,\end{aligned} \tag{36}$$

where, here, P_{ch} is the total launched optical power, including contributions of signal, nonlinear interference, and additive noise. The numerator of (36) finds the actual signal power by subtracting the components of nonlinear interference and noise. Equation (36) differs only by a constant offset from (33) and, therefore, the optimal launch power is still given by (34). Thus, in the optimal launch power condition

$$\mathrm{SNR}_{\mathrm{nli,sd}}(K) = \frac{1}{K}\frac{P_{\mathrm{opt}}}{1.5P_{\mathrm{ase-ss}}} - 1 = \frac{\mathrm{SNR}_{\mathrm{nli,sd}}(1) + 1}{K^{(1+\epsilon/3)}} - 1. \tag{37}$$

APPENDIX D
SNR EVOLUTION IN SYSTEMS LIMITED BY POWER-FEED CONSTRAINTS

Long-haul transoceanic submarine systems are limited by power-feed equipment constraints [45][3][38][37], producing a different scaling rule for the SNR (note again that we are referring to the SNR at the end of the link, and not the SNR evolution within a link). In power-limited submarine systems, the total power P available to feed all amplifiers is limited by the power-feed voltage V as [46]

$$P = \frac{V^2}{4\rho K L_{\mathrm{span}}}, \tag{38}$$

where L_{span} is the span length and ρ is the cable resistance per unit length. Neglecting the power overhead term that accounts for electrical power spent in operations not directly related to optical amplification, and assuming that the total available electrical power is equally divided among all amplifiers, yields

$$P_{\mathrm{ch}} = \gamma'\frac{P}{K}, \tag{39}$$

where γ' is a proportionality constant related to the number of fibers and the electrical-to-optical power conversion efficiency

[46]. Under these assumptions, the SNR at the end of the link becomes

$$\mathrm{SNR}_{\mathrm{pf}}(K) = \frac{P_{\mathrm{ch}}}{KP_{\mathrm{ase-ss}}} = \frac{\gamma/K^2}{KP_{\mathrm{ase-ss}}} = \frac{\mathrm{SNR}_{\mathrm{pf}}(1)}{K^3}, \quad (40)$$

where $\mathrm{SNR}_{\mathrm{pf}}(1)$ is the SNR after the first span, and γ is another proportionality constant. In practice, as $\mathrm{SNR}_{\mathrm{nli}}(1) \ll \mathrm{SNR}_{\mathrm{pf}}(1)$, NLI dominates in shorter links, while power-feed limitations constrain the available power for transoceanic distances. Therefore, it is reasonable to approximate the SNR at the end of the link by the minimum of the SNRs computed for NLI-limited or power-feed-limited systems.

Appendix E
Capacity distribution for an SMF channel impaired by PDG

It is possible to analytically derive the instantaneous capacity distribution for narrowband SMF channels ($D = 2$) that have been impaired by PDG. The following analysis is not valid for wideband channels with PMD (and therefore those with high frequency diversity). In an SMF channel affected by PDG, the log power gains follow a two-sided non-central chi distribution [47][48]

$$p(g) = 3\sqrt{\frac{3}{2\pi}}\frac{g\sinh g}{\xi^3}\exp\left(-\frac{3}{2}\frac{g^2}{\xi^2} - \frac{\xi^2}{6}\right), \quad (41)$$

which tends to the well-known two-sided Maxwellian distribution in the limit of low PDG [11][1]. Supposing the log power gains of the overall channel have STD σ_{mdg}, the realization of a certain polarization channel will experience a linear power gain e^g/α^2, whereas the other polarization will experience a linear power gain of e^{-g}/α^2. For simplicity, we have directly absorbed the effect of the normalization factor α (see Appendix A) into the polarization power gains. By the properties of the distribution in (41) we can show that

$$\begin{aligned}\alpha^2 &= \left(\mathbb{E}\{e^g\} + \mathbb{E}\{e^{-g}\}\right)/2 \\ &= \mathbb{E}\left\{e^g\right\} \\ &= \exp\left(\xi^2/2\right). \end{aligned} \quad (42)$$

Using (10), the capacity of this channel realization is given by

$$C_{\mathrm{pdg}} = \log_2\left(1 + \frac{\mathrm{SNR}}{2}\frac{e^g}{\alpha^2}\right) + \log_2\left(1 + \frac{\mathrm{SNR}}{2}\frac{e^{-g}}{\alpha^2}\right), \quad (43)$$

which can be simplified to

$$C_{\mathrm{pdg}} = \log_2\left(1 + \frac{\mathrm{SNR}^2}{4\alpha^4} + \frac{\mathrm{SNR}}{\alpha^2}\cosh(g)\right). \quad (44)$$

As the $\cosh(x)$ function has its minimum equal to 1 when $x = 0$, there is a minimum capacity given by

$$C_{\mathrm{pdg}}^{\mathrm{min}} = 2\log_2\left(1 + \frac{\mathrm{SNR}}{2\alpha^2}\right). \quad (45)$$

Since (44) is an even function and monotonic when restricted to either positive or negative real values of g, the capacity CDF can then be expressed as

$$\begin{aligned}F(C_{\mathrm{pdg}}) &= 2\int_0^{g(C_{\mathrm{pdg}})} p(g')dg' \quad (46)\\ &= 1 - Q_{3/2}\left(\frac{\xi}{\sqrt{3}}, \frac{\sqrt{3}}{\xi}g(C_{\mathrm{pdg}})\right), \quad (47)\end{aligned}$$

where $Q_M(\cdot,\cdot)$ is the generalized Marcum Q-function of order M, and

$$g(C_{\mathrm{pdg}}) = \mathrm{acosh}\left[\frac{\alpha^2}{\mathrm{SNR}}\left(2^{C_{\mathrm{pdg}}} - 1 - \frac{\mathrm{SNR}^2}{4\alpha^4}\right)\right]. \quad (48)$$

The capacity PDF is then given by

$$f(C_{\mathrm{pdg}}) = \frac{p\left(g(C_{\mathrm{pdg}})\right)\cdot 2^{C_{\mathrm{pdg}}+1}\ln 2}{\left(\left(2^{C_{\mathrm{pdg}}} - 1 - \frac{\mathrm{SNR}^2}{4\alpha^4}\right)^2 - \frac{\mathrm{SNR}^2}{\alpha^4}\right)^{1/2}}. \quad (49)$$

Fig. 11 compares the results obtained analytically and using the multisection model. Note that the capacity distribution is asymmetric: the maximum capacity is unbounded, whereas the minimum capacity is given by $C_{\mathrm{pdg}}^{\mathrm{min}}$, which depends on the values of α (equivalently, σ_{mdg}) and SNR. For given values of σ_{mdg} and SNR, the outage capacity is equal to $C_{\mathrm{pdg}}^{\mathrm{min}}$ for outage probabilities below $F\left(C_{\mathrm{pdg}}^{\mathrm{min}}\right)$. Even though C_{pdg} is an increasing function of $|g|$ in (44), its average value is still lower than that in the case without PDG due to the power normalization factor α^2.

Appendix F
Minimum channel capacity

The instantaneous capacity for a given channel realization of D modes can be written as

$$C_{\mathrm{mdg}} = \sum_{i=1}^{D}\log_2\left(1 + \beta e^{g_i}\right), \quad (50)$$

where $\beta = \mathrm{SNR}/(D\alpha^2)$. As a function of the log power gains $C_{\mathrm{mdg}} = C_{\mathrm{mdg}}\left(\{g_i\}_{i=1}^{D}\right)$, a total differential element of capacity is written as the sum

$$dC_{\mathrm{mdg}} = \sum_{i=1}^{D}\left(\frac{\partial C_{\mathrm{mdg}}}{\partial g_i}\right)dg_i, \quad (51)$$

where $\left(\frac{\partial C_{\mathrm{mdg}}}{\partial g_i}\right)$ is the partial derivative of capacity C_{mdg} with respect to the i^{th} gain holding the remaining gains constant. We impose the single constraint that the gains sum to zero,

$$\sum_{i=1}^{D} g_i = 0 \implies \sum_{i=1}^{D} dg_i = 0, \quad (52)$$

which results in one of the gain variables becoming dependent on the remaining $D-1$ variables. Without loss of generality, assume g_D is the dependent variable, so (51) becomes

$$dC_{\mathrm{mdg}} = \sum_{i=1}^{D-1}\left(\frac{\partial C_{\mathrm{mdg}}}{\partial g_i} - \frac{\partial C_{\mathrm{mdg}}}{\partial g_D}\right)dg_i. \quad (53)$$

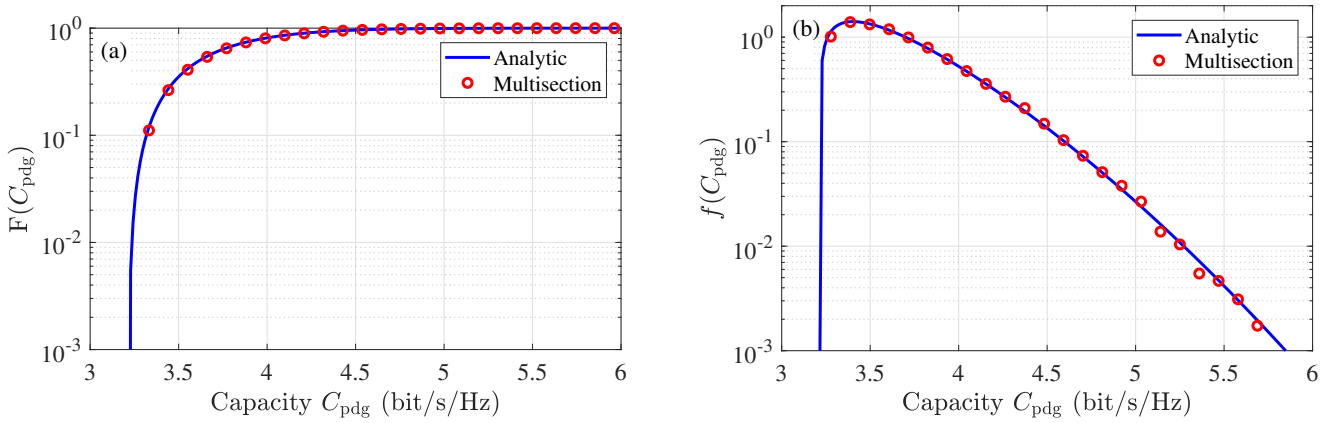

Fig. 11. CDF (a) and PDF (b) of the instantaneous capacity of a narrowband channel with $D = 2$ (SMF). The analytic CDF is given by (47), and the PDF by (49). The simulations correspond to 100,000 realizations of 300 spans, with per-amplifier STD of PDG $\sigma_g = 0.3$ dB. The SNR per mode (SNR/D) at the end of the link is set to 6.2 dB.

$dC_{\mathrm{mdg}} = 0$, i.e., C_{mdg} attains an extremum with respect to the gains, only if $\frac{\partial C_{\mathrm{mdg}}}{\partial g_i} = \frac{\partial C_{\mathrm{mdg}}}{\partial g_D}$ for $i = 1, \ldots, D-1$.

Now, C_{mdg} is a convex function of the log power gains, since its Hessian is a positive semi-definite diagonal matrix whose k^{th} diagonal entry is:

$$\frac{1}{\ln 2} \frac{\beta e^{g_k}}{(1+\beta e^{g_k})^2} > 0. \tag{54}$$

Hence, the condition for minimum capacity from (53) is

$$\frac{\beta e^{g_i}}{1+\beta e^{g_i}} = \frac{\beta e^{g_D}}{1+\beta e^{g_D}} \qquad i = 1, \ldots, D-1 \tag{55}$$

which occurs only when $g_1 = \ldots = g_{D-1} = g_D = 0$, i.e., the gains are equal and sum to zero, hence each gain must be zero. This gives the minimum channel capacity as

$$C_{\mathrm{mdg}}^{\min} = D \log_2 \left(1 + \frac{\mathrm{SNR}}{D\alpha^2}\right). \tag{56}$$

REFERENCES

[1] 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. 16 612–16 635, August 2011.

[2] L. E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, and L. Rapp, "Statistics of polarization dependent loss in an installed long-haul WDM system," *Optics Express*, vol. 19, no. 7, pp. 6790–6796, March 2011.

[3] R. Dar, P. J. Winzer, A. R. Chraplyvy, S. Zsigmond, K. . Huang, H. Fevrier, and S. Grubb, "Cost-optimized submarine cables using massive spatial parallelism," *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 3855–3865, September 2018.

[4] A. Ghazisaeidi, "Theory of coherent WDM systems using in-line semiconductor optical amplifiers," *Journal of Lightwave Technology*, pp. 1–1, 2019.

[5] J. Renaudier and A. Ghazisaeidi, "Scaling capacity growth of fiber-optic transmission systems using 100+nm ultra-wideband semiconductor optical amplifiers," *Journal of Lightwave Technology*, vol. 37, no. 8, pp. 1831–1838, April 2019.

[6] J. Renaudier, A. Arnould, D. Le Gac, A. Ghazisaeidi, P. Brindel, M. Makhsiyan, A. Verdier, K. Mekhazni, F. Blache, H. Debregeas, A. Boutin, N. Fontaine, D. Neilson, R. Ryf, H. Chen, M. Achouche, and G. Charlet, "107 Tb/s transmission of 103-nm bandwidth over 3×100 km SSMF using ultra-wideband hybrid Raman/SOA repeaters," in *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, March 2019.

[7] Z. Zhu, X. Li, and Y. Xi, "A polarization insensitive semiconductor optical amplifier," *IEEE Photonics Technology Letters*, vol. 28, no. 17, pp. 1831–1834, September 2016.

[8] R. Ryf, J. C. Alvarado-Zacarias, S. Wittek, N. K. Fontaine, R.-J. Essiambre, H. Chen, R. Amezcua-Correa, H. Sakuma, T. Hayashi, and T. Hasegawa, "Coupled-core transmission over 7-core fiber," in *Optical Fiber Communication Conference Postdeadline Papers 2019*, 2019, p. Th4B.3.

[9] R. Ryf, N. K. Fontaine, B. Guan, R. . Essiambre, S. Randel, A. H. Gnauck, S. Chandrasekhar, A. Adamiecki, G. Raybon, B. Ercan, R. P. Scott, S. J. Ben Yoo, T. Hayashi, T. Nagashima, and T. Sasaki, "1705-km transmission over coupled-core fibre supporting 6 spatial modes," in *2014 The European Conference on Optical Communication (ECOC)*, September 2014, pp. 1–3.

[10] P. J. Winzer, H. Chen, R. Ryf, K. Guan, and S. Randel, "Mode-dependent loss, gain, and noise in MIMO-SDM systems," in *2014 The European Conference on Optical Communication (ECOC)*, Sep. 2014, pp. 1–3.

[11] A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 3, pp. 313–315, March 2002.

[12] M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," *Optics Express*, vol. 16, no. 18, pp. 13 918–13 932, September 2008.

[13] A. Nafta, E. Meron, and M. Shtaif, "Capacity limitations in fiber-optic communication systems as a result of polarization-dependent loss," *Optics Letters*, vol. 34, no. 23, pp. 3613–3615, December 2009.

[14] 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. 16 680–16 696, August 2011.

[15] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Modeling and performance metrics of MIMO-SDM systems with different amplification schemes in the presence of mode-dependent loss," *Opt. Express*, vol. 23, no. 3, pp. 2203–2219, Feb 2015.

[16] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *Journal of Lightwave Technology*, vol. 29, no. 24, pp. 3719–3726, December 2011.

[17] A. Mecozzi, C. Antonelli, and M. Shtaif, "Intensity impulse response of SDM links," *Opt. Express*, vol. 23, no. 5, pp. 5738–5743, Mar 2015.

[18] G. Rademacher, R. S. Luís, B. J. Puttnam, T. A. Eriksson, R. Ryf, E. Agrell, R. Maruyama, K. Aikawa, Y. Awaji, H. Furukawa, and N. Wada, "High capacity transmission with few-mode fibers," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 425–432, Jan 2019.

[19] J. van Weerdenburg, R. Ryf, R. Alvarez-Aguirre, N. K. Fontaine, R. Essiambre, H. Chen, J. C. Alvarado-Zacarias, R. Amezcua-Correa, S. Gross, N. Riesen, M. Withford, D. W. Peckham, A. McCurdy, R. Lingle, T. Koonen, and C. Okonkwo, "Mode-multiplexed 16-QAM transmission over 2400-km large-effective-area depressed-cladding 3-mode fiber," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, March 2018.

[20] J. van Weerdenburg, R. Ryf, J. C. Alvarado-Zacarias, R. A. Alvarez-Aguirre, N. K. Fontaine, H. Chen, R. Amezcua-Correa, Y. Sun,

L. Grüner-Nielsen, R. V. Jensen, R. Lingle, T. Koonen, and C. Okonkwo, "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, March 2018.

[21] 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, January 2010.

[22] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 271–286, January 2006.

[23] W. Shieh, "PMD-supported coherent optical OFDM systems," *IEEE Photonics Technology Letters*, vol. 19, no. 3, pp. 134–136, Feb 2007.

[24] A. Andrusier, M. Shtaif, E. Meron, and M. Feder, "Equalization performance in the presence of linear polarization impairments," in *Optical Fiber Communication Conference*. Optical Society of America, 2012, p. OTu1A.1.

[25] K.-P. Ho and J. Kahn, "Linear propagation effects in mode-division multiplexing systems," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 614–628, February 2014.

[26] K. Choutagunta, S. O. Arik, K.-P. Ho, and J. M. Kahn, "Characterizing mode-dependent loss and gain in multimode components," *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 3815–3823, September 2018.

[27] K. Choutagunta, I. Roberts, and J. M. Kahn, "Efficient quantification and simulation of modal dynamics in multimode fiber links," *Journal of Lightwave Technology*, vol. 37, no. 8, pp. 1813–1825, April 2019.

[28] K.-P. Ho and J. M. Kahn, "Statistics of Group Delays in Multimode Fiber With Strong Mode Coupling," *Journal of Lightwave Technology*, vol. 29, no. 21, pp. 3119–3128, November 2011.

[29] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[30] A. Ghazisaeidi, I. Fernandez de Jauregui Ruiz, L. Schmalen, P. Tran, C. Simonneau, E. Awwad, B. Uscumlic, P. Brindel, and G. Charlet, "Submarine transmission systems using digital nonlinear compensation and adaptive rate forward error correction," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1886–1895, April 2016.

[31] S. O. Arik and J. M. Kahn, "Diversity-multiplexing tradeoff in mode-division multiplexing," *Opt. Lett.*, vol. 39, no. 11, pp. 3258–3261, Jun 2014.

[32] *Corning Vascade Optical Fiber*, Corning, Inc., March 2017.

[33] T. Hayashi, Y. Tamura, T. Hasegawa, and T. Taru, "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, February 2017.

[34] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857–3879, December 2012.

[35] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, February 2014.

[36] C. Antonelli, M. Shtaif, and A. Mecozzi, "Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 36–54, January 2016.

[37] S. Desbruslais, "Maximizing the capacity of ultra-long haul submarine systems," in *Proceedings of NOC'15*, vol. 25, 2015, pp. 1–6.

[38] O. D. Domingues, D. A. A. Mello, R. da Silva, S. O. Arik, and J. M. Kahn, "Achievable rates of space-division multiplexed submarine links subject to nonlinearities and power feed constraints," *Journal of Lightwave Technology*, vol. 35, no. 18, pp. 4004–4010, September 2017.

[39] O. V. Sinkin, A. V. Turukhin, Y. Sun, H. G. Batshon, M. V. Mazurczyk, C. R. Davidson, J. Cai, W. W. Patterson, M. A. Bolshtyansky, D. G. Foursa, and A. N. Pilipetskii, "SDM for power-efficient undersea transmission," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 361–371, January 2018.

[40] K. Shibahara, T. Mizuno, and Y. Miyamoto, "LDPC-coded FMF transmission employing unreplicated successive interference cancellation for MDL-impact mitigation," in *Proceedings of ECOC*, 2017, pp. 1–3.

[41] K. Shibahara, T. Mizuno, D. Lee, Y. Miyamoto, H. Ono, K. Nakajima, Y. Amma, K. Takenaga, and K. Saitoh, "DMD-Unmanaged Long-Haul SDM Transmission Over 2500-km 12-Core × 3-Mode MC-FMF and 6300-km 3-Mode FMF Employing Intermodal Interference Canceling Technique," *Journal of Lightwave Technology*, vol. 37, no. 1, pp. 138–147, January 2019.

[42] S. O. Arik, J. M. Kahn, and K. P. Ho, "MIMO signal processing for mode-division multiplexing: an overview of channel models and signal processing architectures," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 25–34, mar. 2014.

[43] R. Nasiri Mahalati, D. Askarov, and J. M. Kahn, "Adaptive Modal Gain Equalization Techniques in Multi-Mode Erbium-Doped Fiber Amplifiers," *Journal of Lightwave Technology*, vol. 32, no. 11, pp. 2133–2143, June 2014.

[44] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, 2010.

[45] J. D. Downie, "Maximum capacities in submarine cables With fixed power constraints for C-band, C+L-band, and multicore fiber systems," *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 4025–4032, September 2018.

[46] J. Krause Perin, J. M. Kahn, J. D. Downie, J. Hurley, and K. Bennett, "Importance of amplifier physics in maximizing the capacity of submarine links," *Journal of Lightwave Technology*, vol. 37, no. 9, pp. 2076–2085, May 2019.

[47] A. Galtarossa and L. Palmieri, "The exact statistics of polarization-dependent loss in fiber-optic links," *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 57–59, Jan 2003.

[48] K.-P. Ho, "Exact model for mode-dependent gains and losses in multimode fiber," *Journal of Lightwave Technology*, vol. 30, no. 23, pp. 3603–3609, Dec 2012.

Darli A. A. Mello (M'08) received the graduate degree in electrical engineering from the University of Campinas (UNICAMP), Campinas, Brazil, in 2000, the M.Sc. degree from the Institute for Communications Engineering, Munich University of Technology (TUM), Munchen, Germany, in 2002. During his Master's studies, he carried out both experimental and theoretical work with the Siemens Research Labs, Munich, Germany. In 2006, he received the Ph.D. degree from UNICAMP after research stays at the TUM and at the California Institute of Technology (Caltech). After his Ph.D. studies, he spent one year with Padtec S.A. as a Senior Technology Engineer. From August 2008 to March 2014, he was an Assistant Professor with the Department of Electrical Engineering, University of Brasilia. He is currently an Assistant Professor with the Department of Communications, School of Electrical and Computer Engineering, UNICAMP. His main research interests include optical transmission and networking. He was a TPC Member of several conferences, including The Optical Fiber Communication Conference, Asia Communications and Photonics Conference, International Conferences on Communications, and IEEE Global Communications Conference. He was the TPC Co-Chair of SBrT 2011, Imoc 2017, and of the Optical Communications Subcommittee of LAOP 2016.

Hrishikesh Srinivas (S'16) received the B.E. degree in photonic engineering and the B.Sc. degree in mathematics and physical science from the University of New South Wales, Sydney, NSW, Australia, in 2017, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2019, where he is working toward the Ph.D. degree. His current research interests include optical fiber communications, optical amplifier physics and photonics.

Karthik Choutagunta (S'11) received the B.S. degree from The University of Texas at Austin in 2014, and the M.S. degree from Stanford University in 2016, both in electrical engineering. He is currently working towards the Ph.D. degree at Stanford University. He worked at Coherent Logix (Austin, TX, USA) in summer 2012, at Intel Corp. (Austin, TX, USA) from 2012-2014, at Apple Inc. (Cupertino, CA, USA) in fall 2015 and summer 2016, at Facebook Inc. (Menlo Park, CA, USA) in summer 2017, and at Nokia Bell Laboratories, Crawford Hill Laboratory (Holmdel, NJ, USA) in summer 2018. His main current research interests include various topics in communications, signal processing, machine learning, and optimization.

Joseph M. Kahn (M'90-SM'98-F'00) received A.B., M.A. and Ph.D. degrees in Physics from the University of California, Berkeley in 1981, 1983 and 1986. In 1987-1990, Kahn was at AT&T Bell Laboratories. In 1989, he demonstrated the first successful synchronous (i.e., coherent) detection in optical fiber systems, achieving record receiver sensitivity. In 1990-2003, Kahn was on the Electrical Engineering and Computer Sciences faculty at Berkeley. He demonstrated coherent detection of QPSK in 1992. In 1999, D.-S. Shiu and Kahn published the first work on shaping and nonequiprobable signaling for optical communications. In the 1990s and early 2000s, Kahn and collaborators performed seminal work on indoor and outdoor free-space optical communications and multi-input multi-output wireless communications. In 2000, Kahn and K.-P. Ho founded StrataLight Communications, whose 40 Gb/s-per-wavelength long-haul fiber transmission systems were deployed widely by AT&T, Deutsche Telekom, and other carriers. In 2002, Ho and Kahn applied to patent the first electronic compensation of fiber Kerr nonlinearity. StrataLight was acquired by Opnext in 2009. In 2003, Kahn became a Professor of Electrical Engineering in the E. L. Ginzton Laboratory at Stanford University. Kahn and collaborators have extensively studied rate-adaptive coding and modulation, as well as digital signal processing for mitigating linear and nonlinear impairments in coherent systems. In 2008, E. Ip and Kahn (and G. Li independently) invented simplified digital backpropagation for compensating fiber Kerr nonlinearity and dispersion. Since 2004, Kahn and collaborators have studied propagation, modal statistics, spatial multiplexing and imaging in multi-mode fibers, elucidating principal modes and demonstrating transmission beyond the traditional bandwidth-distance limit in 2005, deriving the statistics of coupled modal group delays and gains in 2011, and deriving resolution limits for imaging in 2013. Kahn's current research addresses optical frequency comb generators, coherent data center links, rate-adaptive access networks, fiber Kerr nonlinearity mitigation, ultra-long-haul submarine links, and optimal free-space transmission through atmospheric turbulence. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. In 2000, he became a Fellow of the IEEE.