Stokes-space analysis of Modal Dispersion of SDM Fibers with Mode-Dependent Loss: Theory and Experiments

Cristian Antonelli, Antonio Mecozzi, Mark Shtaif, Nicolas Fontaine, Haoshuo Chen, and Roland Ryf

(Invited Paper)

***Abstract*—Signal propagation in Space-Division Multiplexed (SDM) systems in the linear regime is dominated by the effects of modal dispersion (MD) and mode-dependent loss (MDL). While multiple models have been proposed for characterizing these phenomena separately or to study the effect of MD on MDL, the effect of MDL on the system MD has never been analyzed. In this work we report such an analysis, where the inclusion of MDL is accounted for by introducing a complex MD vector $\vec{\tau}$. We show that the signal delay spread, quantified by the duration of the intensity impulse response function, is not affected by the presence of MDL, and its functional dependence on $\vec{\tau}$ remains the same as in the absence of MDL (in which case $\vec{\tau}$ is a real-valued vector). The model, which represents SDM systems operating in the regime of strong coupling between modes, is validated by comparison with experimental data.**

***Index Terms*—Optical fiber communications, Space-Division Multiplexing, Modal Dispersion, Mode-Dependent Loss.**

I. Introduction

Modal dispersion (MD) and mode-dependent loss (MDL) are two of the most relevant propagation effects in space-division multiplexed (SDM) transmission systems, and a number of models assessing their impact on system performance have been reported. Models of MD study the distortion of the propagating waveform as a result of the accumulated delay spread [1], [2]. The characteristics of MD are very different in the two regimes of partial [3]–[7] and strong [1], [8], [9] coupling between modes, and they have been studied in most cases by neglecting the effect of MDL. On the other hand, the study of MDL and its effect on system capacity has been conducted by focusing on the MDL-induced fluctuations of the signal-to-noise ratio [10]–[13]. The frequency dependence of MDL and MDL-induced capacity fluctuations as a consequence of the fiber MD has been also studied in [14]. Nonetheless, the way in which MDL affects the system MD, and most importantly the link intensity impulse response, has not been investigated.

In this work we propose a unified model for MD and MDL. We consider the transmission regime where all propagating modes undergo random coupling, as is the case in coupled-core multi-core fibers [8], [15], [16]. The proposed model studies the accumulation of MD and MDL along the link in the generalized Stoke-space representation of multiple-mode propagation [2], [6], [17]–[21], and describes the combined effect of MD and MDL on the system intensity impulse response (IIR), which as has been shown in [8], is a deterministic property of SDM propagation and its duration determines the memory that is required for the MIMO-SDM receiver. In the process, we define a complex MD vector $\vec{\tau}$ that generalizes the MD vector introduced in [2] so as to account for the presence of MDL. We show that the duration of the IIR is simply related to the mean square value of this vector, and so is the system mean MDL. A complex vector $\vec{\tau}$ was also introduced in the context of single-mode systems [22] so as to account for the joint effect of polarization-mode dispersion (PMD) and polarization-dependent loss (PDL), and a physical interpretation of its real and imaginary parts was given in the regime of small PDL [23]. Here the definition of the complex MD vector allows to construct a general theory of the IIR that goes beyond the derivation in [8] and its precursor [24]. The proposed model relies on two scalar parameters only which can be extracted from experiments. Finally, the proposed model is validated by comparison with experimental data collected on a SDM system employing a coupled-core three-core fiber. To the the best of our knowledge, this is the first direct comparison between theory and experiments in the context of MD and MDL. It should be noted that the experiment relied on a spooled fiber, whereas experiments on deployed SDM fibers are planned to take place in the recently established SDM fiber infrastructure in the Italian city of L'Aquila [25].

The paper is organized as follows. In Sec. II we introduce the model and the formalism that is involved, define the complex MD vector and derive the equation governing its evolution. In Sect. III we characterize the statistics of the MD vector in terms of its autocorrelation functions. In Sect. IV we derive the expression for the IIR in the presence of MDL and its relation with the complex MD vector. In Sec. V we present a comparison between theory and experimental results.

II. The complex MD vector

We express the electric field propagating in a fiber that supports $2N$ modes, where N is the number of spatial modes and the factor of two accounts for polarization degeneracy, as

$$\vec{\mathrm{E}}(x,y,z,t) = \mathrm{Re}\left[\sum_{n=1}^{2N} \vec{f}_n(x,y) E_n(z,t) e^{-i\Omega_0 t}\right]. \quad (1)$$

Manuscript received February 25, 2020

C. Antonelli and A. Mecozzi are with the Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila 67100, Italy. M. Shtaif is with the Department of Physical Electronics, Tel Aviv University, Tel Aviv 69978, Israel., N. Fontaine, H. Chen, and R. Ryf are with Nokia Bell Labs, Crawford Hill Laboratory, Holmdel, New Jersey 07733, USA.

Here z is the propagation coordinate, Ω_0 is the central optical frequency, n is the mode index, $\vec{f}_n(x,y)$ and $E_n(z,t)$ are the corresponding lateral profile and complex field envelope, respectively, normalized as in [26]. We construct the field vector $\vec{E}(z,t)$ by stacking the $2N$ complex envelopes on top of one another, and we denote its Fourier transform by $\tilde{\vec{E}}(z,\omega)=\int_{-\infty}^{+\infty}\vec{E}(z,t)\exp(i\omega t)\mathrm{d}t$, so that linear propagation is described by the relation

$$\tilde{\vec{E}}(z,\omega)=\mathbf{T}(z,\omega)\tilde{\vec{E}}(0,\omega), \tag{2}$$

where $\mathbf{T}(z,\omega)$ is the transfer matrix from the link input to a generic point z along the fiber, and where ω is the offset from the carrier frequency Ω_0. The evolution of $\mathbf{T}(z,\omega)$ obeys the following equation,

$$\begin{aligned}\frac{\partial\mathbf{T}(z,\omega)}{\partial z} &= \frac{1}{2}(g_0-\alpha_0)\mathbf{T}+i\phi\mathbf{T}\\ &\quad+\left[-\frac{\vec{\alpha}\cdot\vec{\Lambda}}{2}+i\frac{\vec{\beta}\cdot\vec{\Lambda}}{2N}\right]\mathbf{T}(z,\omega),\end{aligned} \tag{3}$$

where g_0 and α_0 are the mode-averaged gain loss coefficients, respectively, and ϕ accounts for the mode-averaged phase accumulation. The quantities $\vec{\alpha}\cdot\vec{\Lambda}$ and $\vec{\beta}\cdot\vec{\Lambda}$ represent $2N\times 2N$ Hermitian matrices accounting for the local effects of MDL and mode coupling,[1] respectively. Following the notation of [2] and [21], we denote by $\vec{\alpha}$ and $\vec{\beta}$ the local MDL vector [21] and the generalized birefringence vector [2], respectively, whereas $\vec{\Lambda}$ is a vector whose elements Λ_n are generalized Pauli matrices forming a basis for $2N\times 2N$ traceless Hermitian matrices, so that

$$\vec{\alpha}\cdot\vec{\Lambda} = \sum_{n=1}^{D}\alpha_n\Lambda_n \tag{4}$$

$$\vec{\beta}\cdot\vec{\Lambda} = \sum_{n=1}^{D}\beta_n\Lambda_n \tag{5}$$

where $D=4N^2-1$ is the dimensionality of the generalized Stokes space [2] in which $\vec{\alpha}$ and $\vec{\beta}$ are defined. In what follows we assume that the dependence of the local MDL vector $\vec{\alpha}$ on frequency can be neglected within the bandwidth values that are relevant to this work (the rational of this assumptions and the implications of frequency-dependent MDL are discussed in Sec. IV-A). For the generalized birefringence vector we adopt the usual first-order expansion around the carrier frequency Ω_0, $\vec{\beta}_0+\omega\vec{\beta}_\omega$, which renders

$$\begin{aligned}\frac{\partial\mathbf{T}(z,\omega)}{\partial z} &= \frac{1}{2}(g_0-\alpha_0)\mathbf{T}+i\phi\mathbf{T}\\ &\quad+\left[-\frac{\vec{\alpha}\cdot\vec{\Lambda}}{2}+i\frac{(\vec{\beta}_0+\omega\vec{\beta}_\omega)\cdot\vec{\Lambda}}{2N}\right]\mathbf{T}(z,\omega)\end{aligned} \tag{6}$$

The contribution of $\vec{\beta}_0$ can be eliminated from (6) by shifting to a rotating reference frame that follows the evolution of $\vec{\beta}_0$. Namely, instead of considering $\mathbf{T}(z,\omega)$ we consider $\mathbf{U}_0^\dagger(z)\mathbf{T}(z,\omega)$ where $\mathbf{U}_0(z)$ is a frequency independent unitary (i.e. basis rotating) matrix obtained from

$$\frac{\mathrm{d}\mathbf{U}_0}{\mathrm{d}z}=i\frac{\vec{\beta}_0\cdot\vec{\Lambda}}{2N}\mathbf{U}_0. \tag{7}$$

In what follows $\mathbf{T}$ will denote the transfer matrix in the rotating reference frame (i.e. it will replace $\mathbf{U}_0^\dagger\mathbf{T}$), so that its evolution is governed by

$$\begin{aligned}\frac{\partial\mathbf{T}(z,\omega)}{\partial z} &= \frac{1}{2}(g_0-\alpha_0)\mathbf{T}+i\phi\mathbf{T}\\ &\quad+\left[-\frac{\vec{\alpha}\cdot\vec{\Lambda}}{2}+i\omega\frac{\vec{\beta}_\omega\cdot\vec{\Lambda}}{2N}\right]\mathbf{T}(z,\omega).\end{aligned} \tag{8}$$

Although we continue using the same notation, the vectors $\vec{\alpha}$ and $\vec{\beta}_\omega$ in Eq. (8) are rotated with respect to $\vec{\alpha}$ and $\vec{\beta}_\omega$ appearing in Eq. (6). Yet, since in the relevant limits, these vectors are modeled as white noise processes that are isotropic in the generalized Stokes space, these rotations do not affect their statistical properties. This explains why the transition to a rotating reference frame has no effect on the statistics of $\mathbf{T}$ or on any physically measurable results reported in this paper. This reasoning also implies that the distribution of $\mathbf{T}$ is independent of ω,[2] and that the statistics of the two-frequency product $\mathbf{T}^\dagger(z,\omega_1)\mathbf{T}(z,\omega_2)$ depends only on the frequency difference $\omega_2-\omega_1$. These properties can be viewed as demonstrating the wide-sense stationarity of the matrix $\mathbf{T}$ with respect to frequency, which has been well known in the context of PMD studies in single-mode fiber systems [27], [28].

The expression for the frequency derivative of $\mathbf{T}$ is similar to Eq. (8),

$$\frac{\partial\mathbf{T}(z,\omega)}{\partial\omega}=\left(i\tau_0-\frac{\gamma}{2}\right)\mathbf{T}(z,\omega)+i\frac{\vec{\tau}\cdot\vec{\Lambda}}{2N}\mathbf{T}(z,\omega), \tag{9}$$

where, as will be clarified later in this section, τ_0 and γ are related to mode-averaged group delay and loss, respectively, and where $\vec{\tau}$ is a complex-valued Stokes vector whose components, as follows from Eq. (9), are given by

$$\tau_n=-i\mathrm{Trace}\left\{\Lambda_n\frac{\partial\mathbf{T}(z,\omega)}{\partial\omega}\mathbf{T}^{-1}(z,\omega)\right\}. \tag{10}$$

In the absence of MDL, in which case the matrix $\mathbf{T}$ is unitary up to a proportionality coefficient that accounts for mode-independent loss, the vector $\vec{\tau}$ becomes real-valued and coincides with the MD vector introduced in [2]. In the presence of MDL, it generalizes the complex PMD vector of [22], [23] to the multi-mode case, and hence we refer to it as the *complex MD vector*. This vector provides a complete characterization of MD and MDL. In the single-mode case, the complex PMD vector has been shown to be simply related to the distortion of the propagated signal within the familiar first-order approximation [23]. This interpretation can be trivially extended to the multi-mode case, except that the first-order approximation is never valid in the case of practical SDM fibers [29]. Instead, in such fibers, where the transfer matrix changes wildly within the typical bandwidth of the information

[1]The frequency dependence of $\vec{\beta}$ is responsible for modal dispersion, which thereby is also accounted for by the term $\vec{\beta}\cdot\vec{\Lambda}$.

[2]That is because for any fixed frequency ω omitting the term $\omega\vec{\beta}_\omega$ is equivalent to modifying the matrix $\mathbf{U}_0$ that only rotates the reference frame.

carrying signal, linear propagation effects are better captured in terms of the IIR which was introduced in [8]. The relation between the MD vector and this entity is described in Sec. IV.

The evolution equations for τ_0, γ, and $\vec{\tau}$, are extracted from the equality

$$\frac{\partial^2 \mathbf{T}(z,\omega)}{\partial\omega\partial z} = \frac{\partial^2 \mathbf{T}(z,\omega)}{\partial z\partial\omega}, \tag{11}$$

where the left- and right-hand sides of the equation are obtained by differentiating Eq. (3) and (8) with respect to ω and z, respectively. Equating the terms proportional to the identity matrix gives the evolution equations for the mode-averaged group delay and loss,

$$\frac{\partial\tau_0}{\partial z} = \frac{\partial\phi}{\partial\omega} \tag{12}$$

$$\frac{\partial\gamma}{\partial z} = -\frac{\partial(g_0-\alpha_0)}{\partial\omega}, \tag{13}$$

which can be integrated with the result

$$\tau_0(z,\omega) = \int_0^z \frac{\partial\beta_0(z',\omega)}{\partial\omega}\mathrm{d}z' \tag{14}$$

$$\gamma(z,\omega) = -\int_0^z \left[\frac{\partial g_0(z',\omega)}{\partial\omega} - \frac{\partial\alpha_0(z',\omega))}{\partial\omega}\right]\mathrm{d}z'. \tag{15}$$

Here the term $\partial\phi/\partial\omega$ reflects the fiber design and perturbations of the mode-averaged propagation constant, whereas g_0 and α_0 depend on the amplification scheme [13], and on the mode-averaged link loss, respectively. Equating the traceless terms yields

$$\frac{\partial\vec{\tau}}{\partial z}\cdot\vec{\Lambda} = \vec{\beta}_\omega\cdot\vec{\Lambda} + \frac{i}{2N}\left[(\omega\vec{\beta}_\omega + iN\vec{\alpha})\cdot\vec{\Lambda}, \vec{\tau}\cdot\vec{\Lambda}\right] \tag{16}$$

where by $[\mathbf{A},\mathbf{B}] = \mathbf{AB}-\mathbf{BA}$ we denote the commutator between matrices $\mathbf{A}$ and $\mathbf{B}$. As shown in [2], the above can be expressed as

$$\frac{\partial\vec{\tau}}{\partial z} = \vec{\beta}_\omega + (\omega\vec{\beta}_\omega + iN\vec{\alpha})\times\vec{\tau}, \tag{17}$$

where by the symbol $\times$ we denote the generalized cross product between Stokes-space vectors [2]. This equation generalizes the evolution equation of the complex PMD vector Eq. (32) of [23] to the multi-mode case. It is worth noting that the validity of Eq. (17) is not limited to any specific regime of operation.

We conclude this section by pointing out that, similarly to the matrix $\mathbf{T}$, the complex MD vector $\vec{\tau}$ is also stationarity with respect to frequency. To see why this is the case, one can define an orthogonal matrix R_0 in the generalized Stokes space, which evolves according to

$$\frac{\mathrm{dR}_0}{\mathrm{d}z} = \omega_0\vec{\beta}_\omega\times\mathrm{R}_0. \tag{18}$$

Then the vector $\mathrm{R}_0^{-1}\vec{\tau}(\omega)$ satisfies Eq. (17) with ω replaced by $\omega-\omega_0$, and since $\vec{\beta}_\omega$ is delta-correlated in z, it has the same distribution as $\vec{\tau}(\omega-\omega_0)$. Similarly to our previous discussion, this implies that the average of products of the form $\vec{\tau}(z,\omega_1)\cdot\vec{\tau}(z,\omega_2)$ or $\vec{\tau}^*(z,\omega_1)\cdot\vec{\tau}(z,\omega_2)$ depends only on the frequency difference $\omega_1-\omega_2$.

III. Two-frequency correlation functions of the complex MD vector

The main goal of this section is to characterize the statistical properties of the vector $\vec{\tau}$. We aim to relate the two averages $\langle\vec{\tau}\cdot\vec{\tau}\rangle$ and $\langle\vec{\tau}\cdot\vec{\tau}^*\rangle$ to the model parameters κ_α and κ_{β_ω}, so as to enable their extraction in the comparison with experimental data. To this end we fist derive the two correlation functions of $\vec{\tau}$, which can be used to characterize the bandwidth of MD and MDL. The physical significance of $\langle\vec{\tau}\cdot\vec{\tau}\rangle$ and $\langle\vec{\tau}\cdot\vec{\tau}^*\rangle$ is discussed in Sec. VII.

As is evident from the evolution equation (17), the statistical properties of the complex MD vector follow from those of the generalized birefringence and local MDL vectors. While in general these may depend on the specific fiber type, in the regime of strong mode mixing that is considered in this work the properties of the accumulated channel transfer matrix do not depend on the details of the local coupling and mode-dependent loss statistics. In this regime, consistent with previous related work [2], [6], [8], [21]–[23], $\vec{\alpha}$ and $\vec{\beta}_\omega$ can be legitimately modeled as two independent D-dimensional Gaussian processes such that $\langle\alpha_n(z)\alpha_m(z')\rangle = \delta_{n,m}\kappa_\alpha^2\delta(z-z')/D$ and $\langle\beta_{\omega,n}(z)\beta_{\omega,m}(z')\rangle = \delta_{n,m}(2N)^2\kappa_{\beta_\omega}^2\delta(z-z')/D$, where angled brackets are used to denote ensemble averaging. A formal treatment of the dynamics requires that $\vec{\alpha}$ and $\vec{\beta}_\omega$ are associates with two corresponding Wiener processes $\vec{W}_a$ and $\vec{W}_b$, whose increments are given by $\mathrm{d}\vec{W}_a = \vec{\alpha}\mathrm{d}z$ and $\mathrm{d}\vec{W}_b = \vec{\beta}_\omega\mathrm{d}z$, such that

$$\mathrm{d}W_{a,n}(z)\mathrm{d}W_{a,m}(z) = \delta_{n,m}\frac{\kappa_\alpha^2}{D}\mathrm{d}z, \tag{19}$$

$$\mathrm{d}W_{b,n}(z)\mathrm{d}W_{b,m}(z) = \delta_{n,m}\frac{(2N)^2\kappa_{\beta_\omega}^2}{D}\mathrm{d}z, \tag{20}$$

in which case Eq. (17) can be rewritten as

$$\mathrm{d}\vec{\tau} = \mathrm{d}\vec{W}_b + (\omega\mathrm{d}\vec{W}_b + iN\mathrm{d}\vec{W}_a)\times\vec{\tau} - \frac{4N^2\kappa_{\beta_\omega}^2\omega^2 - N^2\kappa_\alpha^2}{D}\vec{\tau}\,\mathrm{d}z \tag{21}$$

which is a stochastic differential equation in the Ito sense [30]. The last term on the right-hand side of the equality is the so-called Ito term, which follows from the formulation of the physical differential equation (17) as a stochastic differential equation in which the Wiener process increments $\mathrm{d}\vec{W}_a(z)$ and $\mathrm{d}\vec{W}_b(z)$ are statistically independent when evaluated at the same position z. Its derivation is given in the appendix.

We now proceed to the derivation of the two-frequency correlation function $f_{\vec{\tau},\vec{\tau}}(z,\omega) = \langle\vec{\tau}(z,\omega)\cdot\vec{\tau}(z,0)\rangle$, which follows from the same path as [2], [27]. Defining for short $\vec{\tau} = \vec{\tau}(z,\omega)$ and $\vec{\tau}_0 = \vec{\tau}(z,0)$, the differential of f is expressed as

$$\mathrm{d}f_{\vec{\tau},\vec{\tau}} = \langle\mathrm{d}\vec{\tau}\cdot\vec{\tau}_0\rangle + \langle\vec{\tau}\cdot\mathrm{d}\vec{\tau}_0\rangle + \langle\mathrm{d}\vec{\tau}\cdot\mathrm{d}\vec{\tau}_0\rangle, \tag{22}$$

which, using the statistical independence of all quantities involved, yields

$$\mathrm{d}f_{\vec{\tau},\vec{\tau}} = -\frac{4N^2\kappa_{\beta_\omega}^2\omega^2 - 2N^2\kappa_\alpha^2}{D}f_{\vec{\tau},\vec{\tau}}\mathrm{d}z + 4N^2\kappa_{\beta_\omega}^2\mathrm{d}z - N^2\left\langle\left(\mathrm{d}\vec{W}_a\times\vec{\tau}\right)\cdot\left(\mathrm{d}\vec{W}_a\times\vec{\tau}_0\right)\right\rangle. \tag{23}$$

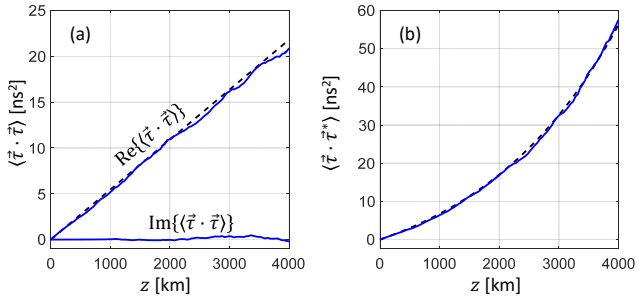

Fig. 1. Real and imaginary part of $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ (a) and $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ (b) versus z, as obtained from Monte Carlo simulations. Dashed lines show the corresponding analytical expressions in Eqs. (28) and (29), respectively.

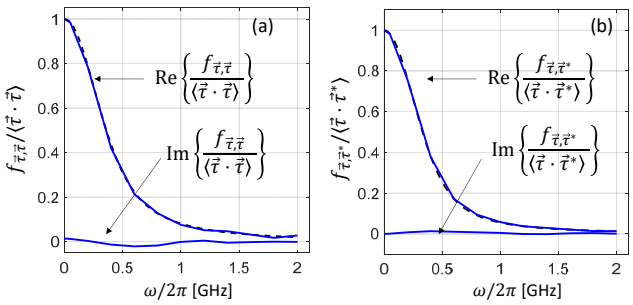

Fig. 2. Real and imaginary part of the autocorrelation functions $\langle \vec{\tau}_1 \cdot \vec{\tau}_2 \rangle$ (left) and $\langle \vec{\tau}_1 \cdot \vec{\tau}_2^* \rangle$ versus $(\omega_1 - \omega_2)/2\pi$. Equations (25) and (27) are plotted by dashed lines for comparison.

The second average on the right-hand side of the equality, as we show in the appendix, reduces to $-2N^2\kappa_\alpha^2 f \mathrm{d}z/D$, with the result

$$\frac{\mathrm{d}f_{\vec{\tau},\vec{\tau}}}{\mathrm{d}z} = 4N^2\kappa_{\beta_\omega}^2 - \frac{4N^2\kappa_{\beta_\omega}^2\omega^2}{D} f_{\vec{\tau},\vec{\tau}}, \tag{24}$$

which yields

$$f_{\vec{\tau},\vec{\tau}}(z,\omega) = \frac{D}{\omega^2}\left\{1 - \exp\left[-\frac{4N^2\kappa_{\beta_\omega}^2\omega^2 z}{D}\right]\right\}. \tag{25}$$

The frequency correlation function $f_{\vec{\tau},\vec{\tau}*}(z,\omega) = \langle \vec{\tau}(z,\omega) \cdot \vec{\tau}^*(z,0) \rangle$, follows from the same procedure. Its evolution equation is found to be

$$\frac{\mathrm{d}f_{\vec{\tau},\vec{\tau}*}}{\mathrm{d}z} = 4N^2\kappa_{\beta_\omega}^2 - 4N^2\frac{\kappa_{\beta_\omega}^2\omega^2 - \kappa_\alpha^2}{D} f_{\vec{\tau},\vec{\tau}*}, \tag{26}$$

with the result

$$f_{\vec{\tau},\vec{\tau}*}(z,\omega) = \frac{D}{\omega^2 - \kappa_\alpha^2/\kappa_{\beta_\omega}^2} \times \left\{1 - \exp\left[-4N^2\kappa_{\beta_\omega}^2\frac{\omega^2 - \kappa_\alpha^2/\kappa_{\beta_\omega}^2}{D} z\right]\right\}. \tag{27}$$

It is worth noticing that only Eq. (27) is affected by the presence of MDL, through the parameter κ_α, whereas Eq. (25) has the same expression that it would have in the absence of MDL [2] (that is, by setting $\kappa_\alpha = 0$). Of course Eq. (27) reduces to (25) as κ_α vanishes, in which case $\vec{\tau}$ becomes real-valued. Note also that both correlation functions are real-valued, a result that lacks any intuitive justification.

As follows from stationarity with respect to frequency, the averages $\langle \vec{\tau} \cdot \vec{\tau} \rangle = f_{\vec{\tau},\vec{\tau}}(z,0)$ and $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle = f_{\vec{\tau},\vec{\tau}*}(z,0)$, are frequency independent and their expressions are

$$\langle \vec{\tau} \cdot \vec{\tau} \rangle = 4N^2\kappa_{\beta_\omega}^2 z, \tag{28}$$

$$\langle \vec{\tau} \cdot \vec{\tau}^* \rangle = \frac{D\kappa_{\beta_\omega}^2}{\kappa_\alpha^2}\left[\exp\left(\frac{4N^2\kappa_\alpha^2}{D} z\right) - 1\right]. \tag{29}$$

We tested the derived quantities by comparison with Monte Carlo simulations. For this purpose we numerically solved Eq. (21) for 1000 independent instantiations of $\vec{W}_a(z)$ and $\vec{W}_b(z)$ and for a number of frequency values, while setting $\kappa_{\beta_\omega} \simeq 12.31$ ps/km$^{1/2}$ and $\kappa_\alpha \simeq 0.02$ km$^{-1/2}$. The numerical integration is based on a wave-plate approach, where the k-th fiber section of length Δz is characterized by a generalized birefringence vector $\vec{b}_k$ and a local MDL vector $\vec{a}_k$, which are selected independently for each section, each having D statistically independent zero-mean Gaussian components of variance $(2N)^2\kappa_{\beta_\omega}^2\Delta z/D$ and $\kappa_\alpha^2\Delta z/D$, respectively.

In Fig. 1a we plot $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ versus propagation distance. For the numerically obtained average we plot by solid lines its real and imaginary parts. The dashed line is a plot of the theoretical expression in Eq. (28). In Fig. 1b the solid line refers to the numerically obtained average $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$, and the dashed line to the theoretical expression in Eq. (29). The good agreement between theory and simulations is self-evident.

In Fig. 2a we plot the normalized correlation function $f_{\vec{\tau},\vec{\tau}}/\langle \vec{\tau} \cdot \vec{\tau} \rangle$ versus ω at $z = 2000$ km, for the same settings as in Fig. 1. The real and imaginary parts of the numerically obtained correlation function are plotted by solid lines,[3] whereas the dashed line is a plot of the theoretical expression in Eq. (25). In Fig. 2b we plot the normalized correlation function $f_{\vec{\tau},\vec{\tau}*}/\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$. The good agreement between theory and simulations is self-evident also in this case. Note that in this example the term $\kappa_\alpha/\kappa_{\beta_\omega}$ is of about 260 MHz, and therefore it introduces a small effect in the correlation function $f_{\vec{\tau},\vec{\tau}*}$ as compared with $f_{\vec{\tau},\vec{\tau}}$. However this correction could be greater depending on the actual MD and MDL of the studied fiber.

IV. The IIR and MDL

The IIR of a fiber in the regime of strong mode mixing is defined as follows. At the fiber input we excite the k-th mode with a signal whose complex envelope is $\psi_0(t)$, and look at the total output power, which is the sum of the output powers in the various modes. This quantity can be expressed as

$$I_k(t) = \sum_{j=1}^{2N} |H_{j,k}(t)|^2, \tag{30}$$

where by $H_{j,k}(t)$ we denote the elements of the matrix

$$\mathbf{H}(t) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \exp(-i\omega t)\, \mathbf{T}(L,\omega)\tilde{\psi}_0(\omega), \tag{31}$$

[3]We note that the numerically obtained correlation function was normalized with respect to the real part of the numerically obtained average $\langle \vec{\tau} \cdot \vec{\tau} \rangle$.

with $\tilde{\psi}_0(\omega) = \int \mathrm{d}t \exp(i\omega t)\,\psi_0(t)$. By exciting the various fiber modes one by one and by averaging the received power signals, we obtained the quantity $I(t)$ in which we are interested,

$$I(t) = \frac{1}{2N}\sum_{k=1}^{2N} I_k(t) = \frac{1}{2N}\mathrm{Trace}\left[\mathbf{H}(t)\mathbf{H}^{\dagger}(t)\right]. \quad (32)$$

The Fourier transform of $I(t)$ is readily found to be

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\Omega}{2\pi}\tilde{\psi}_0(\omega+\Omega)\psi_0^*(\Omega)\tilde{R}(L,\omega,\Omega), \quad (33)$$

where

$$\tilde{R}(L,\omega,\Omega) = \frac{1}{2N}\mathrm{Trace}\left[\mathbf{T}(L,\omega+\Omega)\mathbf{T}^{\dagger}(L,\Omega)\right]. \quad (34)$$

As argued in [8], the function $\tilde{R}(L,\omega,\Omega)$ can be replaced with its statistical average, provided that the bandwidth of $\psi_0(t)$ is much larger than the modal dispersion bandwidth defined below.[4] By defining

$$\tilde{r}(L,\omega) = \langle \tilde{R}(L,\omega,\Omega)\rangle, \quad (35)$$

where the dependence on Ω disappears owing to the stationarity with respect to frequency, the expression for $I(t)$ simplifies to

$$I(t) = r(L,t) \otimes I_0(t), \quad (36)$$

with $I_0(t) = |\psi_0(t)|^2$, and where the symbol $\otimes$ denotes convolution. The above shows a linear relation between the input and output powers, and hence $r(L,t)$ represents the IIR. We now proceed to derive $\tilde{r}(L,\omega)$ – the Fourier transform of the IIR – and to this end we look at the evolution of the average $\mathbf{A}(z,\omega) = \langle \mathbf{T}(z,\omega)\mathbf{T}^{\dagger}(z,0)\rangle$. To this end we re-express Eq. (8) as follows,

$$\begin{aligned}\mathrm{d}\mathbf{T}(z,\omega) &= \left[\frac{1}{2}(g_0-\alpha_0)+i\phi\right]\mathbf{T}(z,\omega)\mathrm{d}z \\ &+\left(-\frac{\mathrm{d}\vec{W}_a\cdot\vec{\Lambda}}{2}+i\omega\frac{\mathrm{d}\vec{W}_b\cdot\vec{\Lambda}}{2N}\right)\mathbf{T}(z,\omega) \\ &+\left(\frac{N^2}{8}\kappa_{\alpha}^2-\frac{\omega^2}{2}\kappa_{\beta_\omega}^2\right)\mathbf{T}(z,\omega)\mathrm{d}z, \quad (37)\end{aligned}$$

where the last term at the right-hand side of the equality is the Ito term (whose derivation is presented in the appendix). The evolution of $\mathbf{A}(z,\omega)$ is therefore governed by

$$\begin{aligned}\mathrm{d}\mathbf{A}(z,\omega) &= \langle \mathrm{d}\mathbf{T}(z,\omega)\mathbf{T}^{\dagger}(z,0)\rangle + \langle \mathbf{T}(z,\omega)\mathrm{d}\mathbf{T}^{\dagger}(z,0)\rangle \\ &+\langle \mathrm{d}\mathbf{T}(z,\omega)\mathrm{d}\mathbf{T}^{\dagger}(z,0)\rangle \\ &= \left(g_0-\alpha_0+\frac{N\kappa_{\alpha}^2}{8}-\kappa_{\beta_\omega}^2\frac{\omega^2}{2}\right)\mathbf{A}(z,\omega), \quad (38)\end{aligned}$$

which after taking the trace of both sides of the equation and integrating from $z=0$ to $z=L$ yields

$$\tilde{r}(L,\omega) = \tilde{r}_0(L)\exp\left(-\frac{\kappa_{\beta_\omega}^2 z\omega^2}{2}\right), \quad (39)$$

where $\tilde{r}_0 = \int_0^L [g_0(z)-\alpha_0(z)]\,\mathrm{d}z + N\kappa_{\alpha}^2 L/8$ is a normalization constant accounting for gain and loss. The inverse Fourier transform of the above finally gives

$$r(L,t) = r_0\exp\left(-\frac{t^2}{2\kappa_{\beta_\omega}^2 L^2}\right), \quad (40)$$

or equivalently

$$r(L,t) = r_0\exp\left[-\frac{t^2}{2T^2(L)}\right], \quad (41)$$

$$T^2(L) = \kappa_{\beta_\omega}^2 L = \frac{\langle\vec{\tau}\cdot\vec{\tau}\rangle}{4N^2}, \quad (42)$$

where $T(z)$ is the mean square duration of the IIR at point z. Equations (40)–(42) show that the expression for the IIR is the same as in the absence of MDL [8], provided that by $\vec{\tau}$ one denotes the complex MD vector introduced in the present work. In particular, Eq. 42 gives a clear physical interpretation of the model parameter $\kappa_{\beta_\omega}^2$, beyond simply being the mere average intensity of the vector $\vec{\beta}_\omega$. The invariance of the IIR expression with respect to the lossless case legitimates also extending the definition of the MD bandwidth introduced[5] in [8] – the frequency for which the power spectrum of the IIR $|\tilde{r}(L,\omega)|^2$ reduces from its peak value by the factor $1/e$ – to the general case in which MDL is present, namely $B_{\mathrm{MD}} = 2N/(2\pi\langle\vec{\tau}\cdot\vec{\tau}\rangle^{1/2})$.

The broadening of the IIR with the link length is accompanied by a corresponding increase of the link MDL. As shown in [21], the accumulation of the mean square link MDL in decibels in the regime of strong mode mixing obeys the equation

$$\langle\rho_{\mathrm{dB}}^2\rangle = \frac{10^2}{\ln^2(10)}f(N)\kappa_{\alpha}^2 z, \quad (43)$$

where $f(N)$ is the empirical relation

$$f(N) = 4\frac{(N-1)^2+24.7(N-1)+16.14}{0.2532(N-1)^2+7.401(N-1)+16.14}, \quad (44)$$

and where ρ_{dB} is given by the familiar expression $\rho_{\mathrm{dB}} = 10\log_{10}(\lambda_{\max}/\lambda_{\min})$, with $\lambda_{\max}$ and $\lambda_{\min}$ being the largest and smallest eigenvalue $\mathbf{T}\mathbf{T}^{\dagger}$, respectively. Similarly, the more familiar mean MDL can be expressed as

$$\langle\rho_{\mathrm{dB}}\rangle = \frac{10}{\ln(10)}\frac{\Gamma\left(\frac{K_N+1}{2}\right)\sqrt{2}}{\Gamma\left(\frac{K_N}{2}\right)\sqrt{K_N}}\sqrt{f(N)\kappa_{\alpha}^2 z}, \quad (45)$$

where $K_N = 3+\lceil 10.39(N-1)^{1.36}\rceil$, and where by $\Gamma(a)$ we denote the Gamma function of argument a.

In Fig. 3a we plot the function $I(t)$ defined in Eq. (33) at various propagation distances along the fiber, for an input excitation $\psi_0(t)$ which was set to be a Nyquist signal with bandwidth $B = 20$GHz around the carrier frequency, $\psi_0(t) = \sqrt{B}\sin(\pi Bt)/\pi Bt$. The numerically generated fiber transfer matrix $\mathbf{T}(z,\omega)$ was obtained using a wave-plate model. The k-th section of length Δz is characterized by the transfer function $\mathbf{T}_k(\omega_h) = \exp\left[\left(-\vec{a}_k/2+i\omega_h\vec{b}_k/2N\right)\cdot\vec{\Lambda}\right]$, where $\vec{a}_k$ and $\vec{b}_k$ are generated as in Figs. 1 and 2, and where the angular

[4]Note that the presence of MDL does not affect the validity of the frequency averaging because the MDL bandwidth is identical to the bandwidth of MD, and hence much smaller than the bandwidth of $\psi_0(t)$.

[5]A related definition can be found in [17].

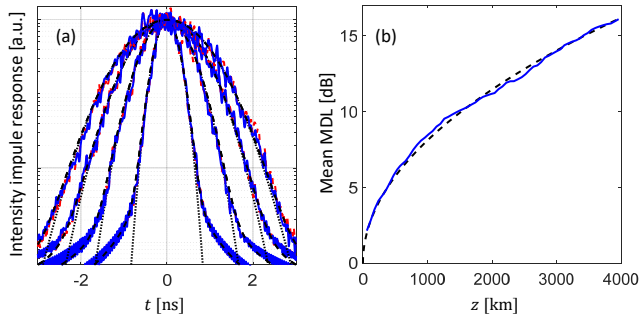

Fig. 3. (a) The mode-average intensity waveform $I(t)$ defined in Eq. (33) at various propagation distances along the fiber, for an input Nyquist signal $\psi_0(t)$ with a bandwidth of 20 GHz. The red curves were obtained in the absence of MDL, whereas the blue curves were obtained by setting κ_α to the same value as in Figs. 1 and 2. The value of κ_{β_ω} was also set as is Figs. 1 and 2. The dashed curves (almost completely hidden by the simulation results) refer to the theoretical expression from Eqs. (36) and (40), and the dotted curves are a plot of the IIR in Eq. (40). All plotted quantities are normalized so as to have unit area. (b) The frequency-average of the accumulating link MDL as a function of propagation distance. The dashed curve refers to the theoretical expression of the link MDL in Eq. (45).

frequencies ω_h are equally spaced between $-B/2$ and $B/2$. The function $I(t)$ is then extracted according to the definition in Eq. (33). The red curves in the figure were obtained in the absence of MDL, whereas the blue curves were obtained by setting κ_α to the same value as in Figs. 1 and 2. The value of κ_{β_ω} was also set as in Figs. 1 and 2. The dashed curves are the plot of the theoretical expressions in Eqs. (36) and (40). The agreement between theory and numerical results is excellent in all cases. Also plotted by dotted lines is the expression of the IIR. As can be seen in the figure, there is perfect overlap between the IIR $r(z,t)$ and the mode-averaged intensity waveform in the main lobe $I(t) = r(z,t) \otimes I_0(t)$, whereas the two differ in the tails, owing to the finite bandwidth of the input excitation signal $\psi_0(t)$. In Fig. 3b we plot the link MDL versus propagation distance. The plotted quantity is the mean MDL in decibels, where the numerical averaging is performed over frequency using the frequency-dependent MDL values $\rho_{\mathrm{dB}}(\omega_h)$. Since the bandwidth over which the averaging is performed is multiple times larger than the MD bandwidth at all propagation distances, the frequency average coincides with the ensemble average, as implied by the excellent agreement with the theoretical result of Eq. (43) that is plotted by the dashed curve. Unlike in single-mode systems where the signal is affected by a particular random instantiation of PDL at its frequency, here the typical signal bandwidth is much larger than the MD bandwidth and hence the effect of MDL is averaged across the signal spectrum [29]. This implies that the mean MDL is the only relevant quantity in single-carrier SDM transmission in the strong coupling regime.

We conclude this section by stressing that the theory of the IIR relies on the fact that the MD bandwidth be multiple times smaller than the bandwidth of a WDM channel. Were this requirement not fulfilled, a different approach to characterizing MD and its consequences on the complexity of the MIMO-receiver equalization scheme would be necessary.

A. Dependence of MDL on frequency

The expression of the IIR has been derived under the assumption that the local MDL vector $\vec{\alpha}$ does not depend on frequency within the bandwidth of the propagating signal. This assumption, which is customary in the PDL/MDL literature [14], [22], [23], is related to the fact that the frequency dependence of the loss itself is negligible over bandwidths on the order of a single communication channel in a wavelength-division multiplexed system. In order to see that most easily, consider scalar propagation in a single-mode fiber in which the propagation constant is given by $k(\tilde{\omega}) = \tilde{\omega} n/c$, where $n = n' + in''$, so that n' is the refractive index whereas n'' accounts for the effect of loss, and where $\tilde{\omega} = \Omega_0 + \omega$ is the optical frequency. The loss coefficient at the center frequency of the signal is $n''\Omega_0/c$ and its relative variation over the signal bandwidth B is given by $B/\Omega_0 + B(\mathrm{d}n''/\mathrm{d}\omega)/n''$. Obviously the first term B/Ω_0 is much smaller than unity. Since fiber-optic systems operate far from material resonances, the variation $B(\mathrm{d}n''/\mathrm{d}\omega)$ is very small relative to n'' and so the second term is much smaller than unity as well. The extension of this argument to the case of polarization-multiplexed and mode-multiplexed systems is straightforward.

It is interesting to note that if the frequency dependence of $\vec{\alpha}$ had to be taken into account, it would have a major effect on the system model characteristics. In particular, the propagation model would loose its frequency stationarity property and its analytical tractability would reduce significantly.

Finally, note that the independence of $\vec{\alpha}$ on frequency does not imply that MDL is frequency-independent. Modal dispersion is responsible for the frequency dependence of the accumulated MDL, consistently with previous MDL studies [14]. In particular, the MDL correlation bandwidth is closely related to the MD correlation bandwidth [14], [31], which in medium-to-long reach links is much narrower than the typical transmission bandwidth. Under the assumptions of frequency stationarity and ergodicity, this implies that MDL-induced impairments are averaged over the signal spectrum [29].

V. Comparison between theory and experiments

In this section we validate the proposed model through a comparison with experimental data taken on a coupled-core three-core fiber in a loop configuration. A detailed description of the experimental setup as well as of the technique that was used to acquire the data can be found in [15]. The estimated impulse response of the link was used to evaluate the transfer matrix $\mathbf{T}$ at the end of each loop over a frequency range of 20 GHz with a resolution of $\Delta f = 0.1$ GHz. The frequency-dependent vector $\vec{\tau}$ was extracted from $\mathbf{T}$ using Eq. (10), where the numerical frequency derivative at the angular frequency ω_h was computed according to the following formula,

$$\frac{\mathrm{d}\mathbf{T}(z_k,\omega_h)}{\mathrm{d}\omega} = \frac{\mathbf{T}(z_k,\omega_h + n_k\Delta\omega) - \mathbf{T}(z_k,\omega_h - n_k\Delta\omega)}{2n_k\Delta\omega} \tag{46}$$

where $z_k = kL_{\mathrm{F}}$ is the link length at the end of the k-th loop recirculation, and n_k is an integer-valued number ranging between $n_k = 3$ for small values of k and $n_k = 1$ for large

values of k. The use of a symmetric expression improves the estimation of the numerical derivative by averaging out additive noise that is always present in experimental data. This averaging is more critical when the modal dispersion bandwidth is larger, in which case measurement noise may be dominant over the frequency dependence of the transfer matrix. This explains the use of larger values of n_h for small values of k. The frequency-dependent MDL parameter $\rho_{\rm dB}$ was also extracted from the transfer matrix $\mathbf{T}$, as described in the previous section. Finally, ensemble averages were estimated by averaging the quantities of interest with respect to frequency.[6]

In Figs. 4a and 4b we plot the two averages $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ (real and imaginary part) and $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$, respectively. Figure 4a shows that the real part of $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ grows almost linearly with distance, whereas its imaginary part is practically vanishing, as predicted by theory. The solid curve is the plot of the theoretical expression $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ in Eq. (28), where $N = 3$ and κ_{β_ω} was used as a fitting parameter. The solid curve in Fig. 4b is the plot of the theoretical expression for $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ in Eq. (29), where the parameter κ_α was used as a second fitting parameter. The fitting procedure returned the values $\kappa_{\beta_\omega} \simeq 12.12$ ps/km$^{1/2}$ and $\kappa_\alpha \simeq 0.019$ km$^{-1/2}$, and the agreement between theory and data is self-evident. In Fig. 4c we plot the mean MDL (in decibels) versus propagation distance, as well as the theoretical expression for $\langle \rho_{\rm dB} \rangle$ given by Eq. (45), where we set κ_α to the value obtained from the fitting performed in Fig. 4b.

Notice that the measured mean MDL amounts to $\sim$ 3dB already after a single loop of 60 km. Such large MDL is to be attributed to the SDM transmitter and amplifier, where the contribution of the first is present only at the system input, while the contribution of the second accumulates at each loop. We can therefore express the overall system transfer matrix as $\mathbf{T} = \mathbf{T}_s \mathbf{T}_{\rm in}$, where $\mathbf{T}_{\rm in}$ accounts for the transmission and amplification stages preceding fiber propagation and $\mathbf{T}_s$ accounts for the link itself. Notice that when the frequency dependence of $\mathbf{T}_{\rm in}$ is negligible, its presence does not affect the accumulation of $\vec{\tau}$. This follows from the fact that

$$\begin{aligned} i\frac{\vec{\tau}\cdot\vec{\Lambda}}{2N} &= \frac{\partial(\mathbf{T}_s\mathbf{T}_{\rm in})}{\partial\omega}(\mathbf{T}_s\mathbf{T}_{\rm in})^{-1} \\ &= \frac{\partial\mathbf{T}_s}{\partial\omega}\mathbf{T}_s^{-1} + \mathbf{T}_s\frac{\partial\mathbf{T}_{\rm in}}{\partial\omega}(\mathbf{T}_s\mathbf{T}_{\rm in})^{-1}, \end{aligned} \quad (47)$$

where the second term on the right-hand side vanishes when $\partial\mathbf{T}_{\rm in}/\partial\omega = 0$. The fact that data points of Figs. 4a and 4b tend to zero at the link input $z = 0$, in spite of the fact that MDL does not vanish at that point, is consistent with this interpretation.

It is interesting to note that, using Eq. (45), the ratio between $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ and $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ can be expressed in terms of $\langle \rho_{\rm dB} \rangle$ as

$$\frac{\langle \vec{\tau} \cdot \vec{\tau}^* \rangle}{\langle \vec{\tau} \cdot \vec{\tau} \rangle} = \frac{\exp\left(a_N \langle \rho_{\rm dB} \rangle^2\right) - 1}{a_N \langle \rho_{\rm dB} \rangle^2}, \quad (48)$$

where

$$a_N = \frac{4N^2}{D}\frac{\ln^2(10)}{10^2 f(N)}\frac{K_N \Gamma^2\left(\frac{K_N}{2}\right)}{2\Gamma^2\left(\frac{K_N+1}{2}\right)}. \quad (49)$$

[6]Data points close to the edges of the input signal spectrum $\tilde{\psi}_0(\omega)$ may suffer from system bandwidth limitation-related effects and hence were discarded.

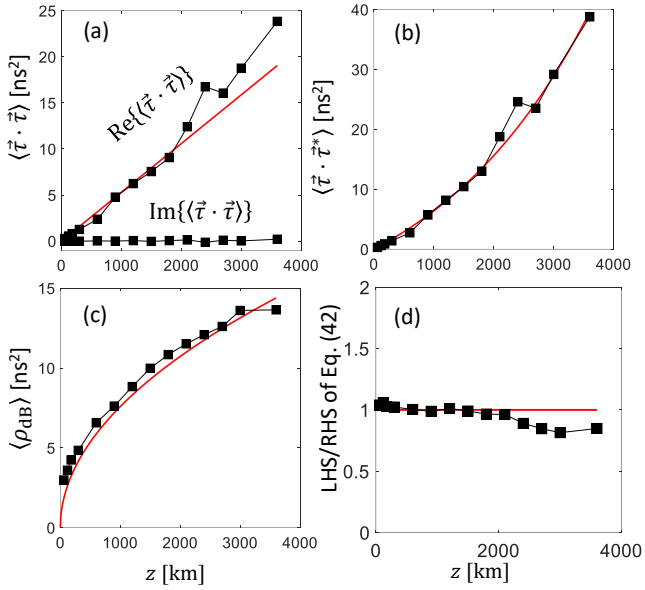

Fig. 4. (a) Experimentally extracted average of $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ (real and imaginary part) versus propagation distance. The solid line is a plot of the theoretical expression in Eq. (28), where $N = 3$ and κ_{β_ω} is used as a fit parameter. (b) Experimentally extracted average of $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ versus propagation distance. The solid line is a plot of the theoretical expression in Eq. (29), where κ_α is used as a second fit parameter. (c) Experimental mean MDL $\langle \rho_{\rm dB} \rangle$ versus propagation distance. The solid curve refers to the theoretical expression in Eq. (45), where κ_α is set to the same value as in (b). (d) Ratio between the left-hand side and the right-hand side of the Eq. (48).

When computed on the experimentally extracted quantities, Eq. (48) provides a simple global test for the accuracy of the proposed model, and it is particularly useful for the regime of small MDL, where it is practically exact. The ratio between the left- and right-hand sides of Eq. (48) is plotted in Fig. 4d. The small deviation from unity for all data points shows the good accuracy of the model.

Finally, in the left and center panels of Fig. 5 we plot the experimentally extracted autocorrelation functions (real and imaginary parts) $\langle \vec{\tau}(\omega + \Omega) \cdot \vec{\tau}(\omega) \rangle$ and $\langle \vec{\tau}(\omega + \Omega) \cdot \vec{\tau}^*(\omega) \rangle$, respectively, versus Ω, for a number of propagation distances. Also plotted by thin lines are the theoretical expressions in Eqs. (25) and (27). In the right panels we plot the corresponding mode-averaged intensity waveforms $I(t)$ and by thin lines the theoretical expression for the IIR from Eqs. (41) and (42), where the value of κ_{β_ω} is the one resulting from the fitting performed in Fig. 4a.

VI. MDL-INDUCED CAPACITY REDUCTION

Mode-dependent loss is a non-unitary effect and therefore it reduces the SDM system capacity. While this is a well known fact [10], [21], in general it is not trivial to quantify the impact of MDL in terms of capacity based on experimental data. For instance, many studies use the capacity ratio as a figure of merit, i.e. the ratio between the system capacity and the capacity that the same system would achieve in the absence of MDL. This ratio cannot be extracted from transfer matrix measurements of the kind that are available in this work owing to the fact that it depends on the specific value

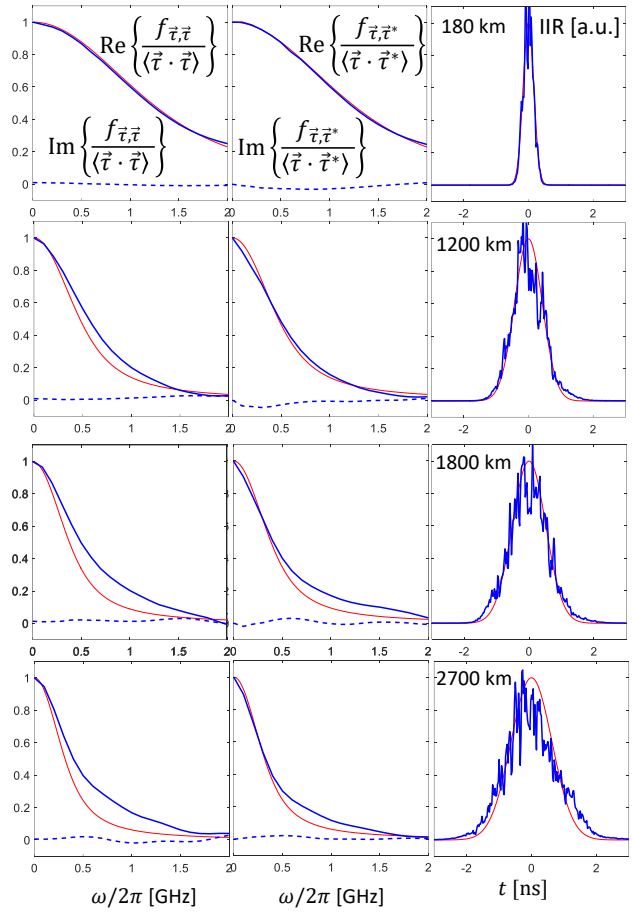

Fig. 5. The thick lines in left and center panels are a plot of the experimentally extracted autocorrelation functions (real and imaginary parts) $\langle \vec{\tau}(\omega+\Omega)\cdot\vec{\tau}(\omega)\rangle$ and $\langle \vec{\tau}(\omega+\Omega)\cdot\vec{\tau}^{*}(\omega)\rangle$, respectively, versus Ω, for the displayed propagation distances. The thin lines refer to the theoretical expressions in Eqs (25) and (27). In the right panels the thick curves are a plot of the corresponding mode-averaged intensity waveforms $I(t)$, while the thin lines refer to the theoretical expression for the IIR from Eqs. (41) and (42).

of the mode-averaged signal-to-noise ratio, as well as on the noise coherency matrix [32], which have to be measured separately. A convenient figure of merit that can be extracted from the system transfer matrix is the capacity[7] loss per mode ℓ introduced in [21], which is defined as the difference between some reference capacity and the capacity of the system under study, divided by the number of scalar modes. When the reference capacity is the capacity that the system under study would achieve in the absence of MDL, the average value of this quantity in the regime of large signal-to-noise ratio (SNR) is independent of the SNR itself and is given by Eq. (3) of [21]

$$\langle \ell \rangle = \frac{\langle \vec{\Gamma}\cdot\vec{\Gamma}\rangle}{3\log(2)}, \tag{50}$$

where $\vec{\Gamma}$ is the MDL vector defined through the equality

$$\mathbf{T}\mathbf{T}^{\dagger} = \gamma\left(\mathbf{I} + \vec{\Gamma}\cdot\vec{\Lambda}\right), \tag{51}$$

[7]Just as in [21], by the term capacity we refer to the spectral efficiency, which is obtained by dividing the actual capacity by the channel bandwidth

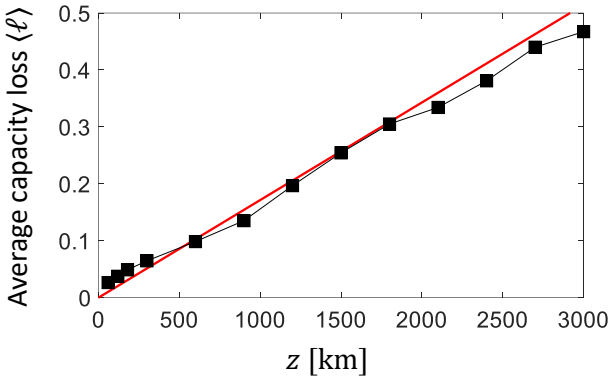

Fig. 6. Average capacity loss per mode versus propagation distance. The symbols show the result obtained by frequency-averaging the scalar product $\vec{\Gamma}\cdot\vec{\Gamma}$ obtained from the experimental data using Eq. (52). The solid line is the plot of theoretical expression in Eq. (53).

with γ being the mode-averaged gain/loss. Owing to the frequency-averaging effect produced by MD [8], the average value of the capacity loss per mode is the relevant quantity to be evaluated, and its extraction from the experimental data is straightforward, using

$$\vec{\Gamma}\cdot\vec{\Gamma} = \frac{\mathrm{Trace}\left[\left(\mathbf{T}\mathbf{T}^{\dagger}\right)^{2}\right]}{\mathrm{Trace}\left[\mathbf{T}\mathbf{T}^{\dagger}\right]^{2}} - 1 \tag{52}$$

and by averaging this quantity with respect to frequency. The average capacity loss is also independent of how the inline amplifiers are operated [13], and is related to the model parameters through the simple expression [21]

$$\langle \ell \rangle = \frac{\kappa_{\alpha}^{2} z}{3\log(2)}, \tag{53}$$

which follows from $\langle \vec{\Gamma}\cdot\vec{\Gamma}\rangle = \kappa_{\alpha}^{2} z$. In Fig. 6 we plot the experimentally extracted capacity loss per mode Eq. (50) versus propagation distance. The solid line is a plot of the theoretical expression in Eq. (53), where we set κ_{α} to the value obtained from the fitting performed in Fig. 4b. The figure shows that the effect of MDL is to reduce the system capacity by about 0.5 bits/s/Hz after 3000 km.

We conclude this section by noting that the vector $\vec{\Gamma}$ is also related to an alternative metric for MDL, the rms MDL, which was investigated in [33]. This metric is proportional to the square-root of the quantity $g^2 = (2N)^{-1}\sum_n g_n^2$, where we denote the n-th eigenvalue of $\mathbf{T}\mathbf{T}^{\dagger}$ by $\exp(g_0+g_n)$, and assume that $\sum_n g_n = 0$. It is straightforward to show that

$$\begin{aligned} g^2 &= \frac{1}{2N}\mathrm{Trace}\left[\log^2\left(\mathbf{I}+\vec{\Gamma}\cdot\vec{\Lambda}\right)\right] - \log^2\left(\gamma e^{-g_0}\right) \\ &\simeq \vec{\Gamma}\cdot\vec{\Gamma} - \log^2\left(\gamma e^{-g_0}\right) \end{aligned} \tag{54}$$

where the second equality holds in the regime of small-to-moderate MDL. The term by which g^2 and $\vec{\Gamma}\cdot\vec{\Gamma}$ differ depends on the way in which the amplifiers are operated (see [13]) and the characterization of its statistics goes beyond the scope of the present work.

VII. Discussion: Model Key-Parameters and Their Physical Significance

Prior to the conclusions, we summarize in this section the parameters of the proposed model and their physical meaning. As should be clear at this point, the key quantity of our model is the vector $\vec{\tau}$. This is a complex-valued vector and can be expressed as $\vec{\tau} = \vec{\tau}_R + i\vec{\tau}_I$. The relation $T^2 = \langle \vec{\tau} \cdot \vec{\tau} \rangle / 4N^2$ provides a simple meaning for the scalar product $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ – the duration of the IIR. Changes in the system MDL do not affect the value of T, in spite of the fact that both MD and MDL affect the accumulation of $\vec{\tau}$. The dependence on MDL emerges in the average $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$. The significance of $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ stems Eq. (48), which shows that the ratio of $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ to $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ depends only on MDL. In the regime of low-to-moderate MDL the exponential function at the right-hand side of the equality can be expanded to the second order with respect to $a_N \langle \rho_{\mathrm{dB}} \rangle^2$, yielding

$$\frac{\langle \vec{\tau} \cdot \vec{\tau}^* \rangle - \langle \vec{\tau} \cdot \vec{\tau} \rangle}{\langle \vec{\tau} \cdot \vec{\tau} \rangle} = \frac{1}{2} a_N \langle \rho_{\mathrm{dB}} \rangle^2, \tag{55}$$

which shows that the relative difference between $\langle \vec{\tau} \cdot \vec{\tau}^* \rangle$ and $\langle \vec{\tau} \cdot \vec{\tau} \rangle$ is proportional to the link square mean MDL. The only two model parameters κ_{β_ω} and κ_α, can therefore be equivalently extracted from space- and frequency-resolved measurements of the system transfer matrix $\mathbf{T}(z, \omega)$ (with the help of Eqs. (28) and (29)), or from space-resolved measurements of the system IIR durations and mean MDL (with the help of Eqs. (42) and (45)).

VIII. Conclusions

We proposed a unified model for modal dispersion and mode-dependent loss in the generalized Stokes space representation of multiple-mode propagation in SDM fibers. The key-quantity of our model is the complex-valued extension of the mode dispersion vector $\vec{\tau}$ that was introduced in [2] as a generalization of the famous polarization-mode dispersion vector [34] to the multi-mode case. We used the proposed model to develop the theory of the intensity impulse response of SDM links affected by mode-dependent loss in the regime of strong mode mixing, previously developed for unitary SDM links [8]. We found that the time duration of the intensity impulse response is determined by $\langle \vec{\tau} \cdot \vec{\tau} \rangle$, just like in the unitary-channel case, except that now $\vec{\tau}$ is a complex-valued vector, whose imaginary part originates from the presence of mode-dependent loss. We validated the model by comparison with experimental data collected on a SDM system based on a coupled-core three-core fiber. To the best of our knowledge, this is the first detailed comparison between SDM theory and experimental data.

Appendix

Ito term of Eq. (21)

The Ito term of Eq. (21) is given by [30]

$$\frac{1}{2}(\omega \mathrm{d}\vec{W}_b + iN\mathrm{d}\vec{W}_a) \times \mathrm{d}\vec{\tau}. \tag{56}$$

After replacing $\mathrm{d}\vec{\tau}$ with the right-hand side of Eq. (21) and retaining only the terms of the order of $\mathrm{d}z$ (while neglecting those of higher order) and those that involve both $\mathrm{d}\vec{W}_b$ and $\mathrm{d}\vec{W}_a$ (which average to zero) it reduces to

$$\frac{\omega^2}{2}\mathrm{d}\vec{W}_b \times (\mathrm{d}\vec{W}_b \times \vec{\tau}) - \frac{N^2}{2}\mathrm{d}\vec{W}_a \times (\mathrm{d}\vec{W}_a \times \vec{\tau}). \tag{57}$$

The final result follows from [2]

$$\mathrm{d}\vec{W}_{a/b} \times (\mathrm{d}\vec{W}_{a/b} \times \vec{\tau}) = \mathrm{d}W^2_{a/b,n}\vec{\tau}, \tag{58}$$

where $\mathrm{d}W^2_{a/b,n}$ is given by Eqs. (19) and (20).

Average in Eq. (23)

Using the definition of generalized vector product, the average that appears in Eq. (23) can be expressed as

$$\begin{aligned} & \left\langle \left(\mathrm{d}\vec{W}_a \times \vec{\tau}\right) \cdot \left(\mathrm{d}\vec{W}_a \times \vec{\tau}_0\right) \right\rangle \\ = & \sum_n \sum_{i,j} \sum_{l,m} \langle f_{i,j,n} f_{l,m,n} \mathrm{d}W_{a,i} \tau_j \mathrm{d}W_{a,l} \tau_{0,m} \rangle \\ = & \ \mathrm{d}W^2_{a,k} \sum_{j,m} \left(\sum_{n,i} f_{i,j,n} f_{i,m,n} \right) \langle \tau_j \tau_{0,m} \rangle \end{aligned} \tag{59}$$

Using $\sum_{n,i} f_{i,j,n} f_{i,m,n} = 2\delta_{j,m}$ [32] and Eq. (19) yields the final result.

Ito term of Eq. (37)

The Ito term of Eq. (37) is given by [30]

$$\frac{1}{2}\left[-\frac{\mathrm{d}\vec{W}_a \cdot \vec{\Lambda}}{2} + i\omega \frac{\mathrm{d}\vec{W}_b \cdot \vec{\Lambda}}{2N} \right] \mathrm{d}\mathbf{T}. \tag{60}$$

Replacing $\mathrm{d}\mathbf{T}$ with the right-hand side of Eq. (21) and neglecting here too the terms proportional to $\mathrm{d}z$ as well as those that involve both $\mathrm{d}\vec{W}_b$ and $\mathrm{d}\vec{W}_a$, it simplifies to

$$\frac{1}{2}\left[\frac{(\mathrm{d}\vec{W}_a \cdot \vec{\Lambda})^2}{4} - \omega^2 \frac{(\mathrm{d}\vec{W}_b \cdot \vec{\Lambda})^2}{4N^2} \right] \mathbf{T}. \tag{61}$$

The final result follows from [2], [32]

$$(\mathrm{d}\vec{W}_{a/b} \cdot \vec{\Lambda})^2 = \mathrm{d}W^2_{a/b,n} \sum_k \Lambda_k^2 = \mathrm{d}W^2_{a/b,n} D\, \mathbf{I}, \tag{62}$$

and Eqs. (19) and (20).

Ackowledgement

C. Antonelli and A. Mecozzi acknowledge financial support from the Italian Government under Cipe resolution n. 135 (Dec. 21, 2012), project INnovating City Planning through Information and Communication Technologies (INCIPICT), as well as under PRIN 2017 project Fiber Infrastructure for Research on Space-division multipled Transmission (FIRST). Mark Shtaif acknowledges the support of the Israel Science Foundation Grant No. 1401/16.

REFERENCES

[1] S. Arik, D. Askarov, and J. M. Kahn, "Effect of mode coupling on signal processing complexity in mode-division multiplexing," *J. Lightwave Technol.*, vol. 31, no. 3, pp. 423–431, Feb 2013. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-31-3-423

[2] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Stokes-space analysis of modal dispersion in fibers with multiple mode transmission," *Opt. Express*, vol. 20, no. 11, pp. 11 718–11 733, May 2012. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-11-11718

[3] L. Palmieri and A. Galtarossa, "Coupling effects among degenerate modes in multimode optical fibers," *IEEE Photonics Journal*, vol. 6, no. 6, pp. 1–8, Dec 2014.

[4] F. M. Ferreira, C. S. Costa, S. Sygletos, and A. D. Ellis, "Semi-analytical modelling of linear mode coupling in few-mode fibers," *Journal of Lightwave Technology*, vol. 35, no. 18, pp. 4011–4022, Sept 2017.

[5] F. M. Ferreira, D. Fonseca, and H. J. A. da Silva, "Design of few-mode fibers with m-modes and low differential mode delay," *J. Lightwave Technol.*, vol. 32, no. 3, pp. 353–360, Feb 2014. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-32-3-353

[6] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Random coupling between groups of degenerate fiber modes in mode multiplexed transmission," *Opt. Express*, vol. 21, no. 8, pp. 9484–9490, Apr 2013. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-8-9484

[7] H. Kogelnik and P. J. Winzer, "Modal birefringence in weakly guiding fibers," *J. Lightwave Technol.*, vol. 30, no. 14, pp. 2240–2245, Jul 2012. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-30-14-2240

[8] A. Mecozzi, C. Antonelli, and M. Shtaif, "Intensity impulse response of SDM links," *Opt. Express*, vol. 23, no. 5, pp. 5738–5743, Mar 2015. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-23-5-5738

[9] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *J. Lightwave Technol.*, vol. 29, no. 21, pp. 3119–3128, Nov 2011. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-29-21-3119

[10] ——, "Mode-dependent loss and gain: statistics and effect on mode-division multiplexing," *Opt. Express*, vol. 19, no. 17, pp. 16 612–16 635, Aug 2011. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-17-16612

[11] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, no. 17, pp. 16 680–16 696, Aug 2011. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-17-16680

[12] S. Warm and K. Petermann, "Splice loss requirements in multi-mode fiber mode-division-multiplex transmission links," *Opt. Express*, vol. 21, no. 1, pp. 519–532, Jan 2013. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-1-519

[13] 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. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-23-3-2203

[14] 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, Dec 2011.

[15] R. Ryf, R. Essiambre, A. Gnauck, S. Randel, M. A. Mestre, C. Schmidt, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, T. Hayashi, T. Taru, and T. Sasaki, "Space-division multiplexed transmission over 4200 km 3-core microstructured fiber," in *Optical Fiber Communication Conference*. Optical Society of America, 2012, p. PDP5C.2. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2012-PDP5C.2

[16] C. Antonelli, A. Mecozzi, and M. Shtaif, "The delay spread in fibers for SDM transmission: dependence on fiber parameters and perturbations," *Opt. Express*, vol. 23, no. 3, pp. 2196–2202, Feb 2015. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-23-3-2196

[17] Q. Hu and W. Shieh, "Autocorrelation function of channel matrix in few-mode fibers with strong mode coupling," *Opt. Express*, vol. 21, no. 19, pp. 22 153–22 165, Sep 2013. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-19-22153

[18] I. Roudas and J. Kwapisz, "Stokes space representation of modal dispersion," *IEEE Photonics Journal*, vol. 9, no. 5, pp. 1–15, Oct 2017.

[19] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Delay spread reduction in mode-division multiplexing: Mode coupling versus delay compensation," *J. Lightwave Technol.*, vol. 33, no. 21, pp. 4504–4512, Nov 2015. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-33-21-4504

[20] ——, "Group delay management and multiinput multioutput signal processing in mode-division multiplexing systems," *J. Lightwave Technol.*, vol. 34, no. 11, pp. 2867–2880, Jun 2016. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-34-11-2867

[21] A. Andrusier, M. Shtaif, C. Antonelli, and A. Mecozzi, "Assessing the effects of mode-dependent loss in space-division multiplexed systems," *J. Lightwave Technol.*, vol. 32, no. 7, pp. 1317–1322, Apr 2014. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-32-7-1317

[22] B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 2, pp. 317–329, March 2000.

[23] M. Shtaif and O. Rosenberg, "Polarization-dependent loss as a waveform-distorting mechanism and its effect on fiber-optic systems," *J. Lightwave Technol.*, vol. 23, no. 2, p. 923, Feb 2005. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-23-2-923

[24] S. D. Personick, "Time dispersion in dielectric waveguides," *Bell System Technical Journal*, vol. 50, no. 3, pp. 843–859, 1971. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1971.tb01885.x

[25] T. Hayashi, T. Nagashima, T. Nakanishi, T. Morishima, R. Kawawada, A. Mecozzi, and C. Antonelli, "Field-deployed multi-core fiber testbed," in *2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC)*, July 2019, pp. 1–3.

[26] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Nonlinear propagation equations in fibers with multiple modes—transitions between representation bases," *APL Photonics*, vol. 4, no. 2, p. 022806, 2019. [Online]. Available: https://doi.org/10.1063/1.5084118

[27] M. Shtaif and A. Mecozzi, "Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion," *Opt. Lett.*, vol. 25, no. 10, pp. 707–709, May 2000. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-25-10-707

[28] M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," *Opt. Lett.*, vol. 24, no. 14, pp. 939–941, Jul 1999. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-24-14-939

[29] C. Antonelli, O. Golani, M. Shtaif, and A. Mecozzi, "Propagation effects in few-mode fibers," in *2017 European Conference on Optical Communication (ECOC)*, Sep. 2017, pp. 1–3.

[30] C. Gardiner, *Handbook of stochastic methods for physics, chemistry, and the natural sciences*, ser. Springer series in synergetics. Springer-Verlag, 1985. [Online]. Available: https://books.google.it/books?id=cRfvAAAAMAAJ

[31] C. Antonelli, A. Mecozzi, L. E. Nelson, and P. Magill, "Autocorrelation of the polarization-dependent loss in fiber routes," *Opt. Lett.*, vol. 36, no. 20, pp. 4005–4007, Oct 2011. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-36-20-4005

[32] A. Mecozzi and C. Antonelli, "Degree of coherence in space-division multiplexed transmission," *J. Lightwave Technol.*, vol. 32, no. 1, pp. 63–69, Jan 2014. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-32-1-63

[33] K. Choutagunta, S. O. Arik, K.-P. Ho, and J. M. Kahn, "Characterizing mode-dependent loss and gain in multimode components," *J. Lightwave Technol.*, vol. 36, no. 18, pp. 3815–3823, Sep 2018. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-36-18-3815

[34] J. P. Gordon and H. Kogelnik, "Pmd fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of the National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, 2000. [Online]. Available: https://www.pnas.org/content/97/9/4541