Modal Statistics in Mode-Division-Multiplexed Systems Using Mode Scramblers

Anirudh Vijay , *Member, IEEE*, Oleksiy Krutko , *Student Member, IEEE*, Rebecca Refaee ,
and Joseph M. Kahn , *Life Fellow, IEEE*

Abstract—**Typical multi-mode fibers exhibit strong intra-group mode coupling and weak inter-group mode coupling. Mode scramblers can be inserted at periodic intervals to enhance inter-group coupling. The deterministic mode coupling of the mode scramblers, in concert with the random mode coupling of the fiber spans, can effect strong random mode coupling between all modes. This reduces both modal dispersion and mode-dependent loss, thereby decreasing receiver complexity and increasing link capacity. In this paper, we analyze the effect of mode scramblers on end-to-end group-delay and mode-dependent loss standard deviations in long-haul multi-mode fiber links. We develop analytical tools in the generalized Jones and Stokes representations. We propose design criteria for mode scramblers that ensure strong end-to-end coupling: the mode-group-averaged power coupling matrix should be primitive and its non-dominant eigenvalues should be near zero. We argue that when the mode scramblers satisfy these criteria, the probability distribution of the system transfer matrix asymptotically approaches that of a system with strong random mode coupling between all modes. Consequently, group-delay and mode-dependent loss standard deviations become sufficient statistics of the eigenvalues of the group-delay operator and the modal gains operator, respectively. We also show that under certain conditions on the uncoupled group delays, it is possible to design self-compensating mode scramblers to reduce group delay accumulation below that of standard strong random coupling.**

Index Terms—**Group-delay spread, long-haul multi-mode fiber systems, mode-dependent loss, mode-division multiplexing, mode scramblers, strong mode coupling.**

I. Introduction

SPACE-DIVISION multiplexing (SDM) can increase capacity, integration, and power efficiency in long-haul optical coherent communication systems [1], [2], [3]. SDM can be implemented using parallel single-mode fibers (SMFs), uncoupled-core or coupled-core multi-core fibers (MCFs), or multi-mode fibers (MMFs). While each of these options has its pros and cons, mode-division multiplexing (MDM) in MMFs is considered attractive in achieving the highest level of integration [4], for

example, by enabling amplification using fewer pump modes than signal modes [5].

Long-haul MDM systems in MMFs need strong mode coupling to manage modal dispersion and mode-dependent gain and loss (collectively referred to as MDL) [6]. Strong random coupling reduces the overall MDL standard deviation (STD), thereby increasing the average capacity and reducing the outage probability, and reduces the group-delay (GD) STD (also known as root-mean-squared GD spread), thereby reducing the digital signal processing (DSP) complexity at the receiver. Systems with strong random coupling exhibit GD and MDL STDs proportional to the square root of the length of propagation. Additionally, strong mode coupling improves frequency diversity, which further reduces the outage probability [7], [8].

Graded-index MMFs are often proposed for long-haul MDM links, owing to their relatively low uncoupled GD STD [9], [10]. In these fibers, the spatial and polarization modes form mode groups. The propagation constants of modes in the same mode group are nearly equal and those in different mode groups are significantly different. Random perturbations in these fibers cause strong intra-group mode coupling and weak inter-group mode coupling [11]. One way of improving inter-group coupling is to periodically insert mode scramblers along the link [12]. Mode scrambling can be achieved in several ways, for example, using offset fiber launches [13], distributed multiple point-loads [14], long-period fiber Bragg gratings (LPFGs) [15], [16], multi-plane light converters (MPLCs) [17], and using mode demultiplexer-multiplexer combinations [18], [19], [20]. Mode scramblers must have low intrinsic MDL STD, because signals in long-haul links may pass through hundreds of them. The mode scramblers with low MDL tend to be highly engineered and have deterministic transfer functions [16]. The analysis in this paper focuses on deterministic mode scramblers, such as those based on LPFGs.

Previous studies to design and analyze mode scramblers have employed heuristic objectives based on off-diagonal entries of the coupling matrix [16]. However, these criteria are restrictive, and evaluating mode scrambler performance requires brute-force multi-section simulations of the end-to-end system to quantify GD spread and overall MDL. These multi-section simulations can significantly slow down iterative design procedures employing performance objectives of end-to-end GD or MDL STD. There is a need for analytical tools that impose less restrictive design criteria and enable faster evaluation of performance objectives.

In this work, we study systems that interleave deterministic mode scramblers with GI-MMFs from first principles and develop analytical tools in the generalized Jones and the generalized Stokes representations. We derive analytical expressions for end-to-end GD STD and MDL STD for such systems, pessimistically assuming no random inter-group coupling in the fiber. We find that a mode-group power coupling matrix is a sufficient description of the mode scrambling induced by a mode scrambler. In particular, the structure and eigenvalues of the power coupling matrix emerge as the design criteria relevant for achieving strong coupling. We also provide a rigorous analysis to incorporate the effects of random inter-group coupling in the fiber and provide closed-form expressions for the GD STD. We present simulations of systems with 6 spatial modes (12 spatial and polarization modes) over 100 spans of propagation, representative of long-haul and ultra-long-haul links. We argue that when the mode scramblers satisfy the aforementioned design criteria, the combination of deterministic mode scrambling and strong random intra-group coupling is asymptotically equivalent to strong random coupling between all modes. Consequently, the GD and MDL STDs become sufficient statistics of the group-delay operator and the modal gains operator, respectively, as the number of spans grows large. We also show that the design space of mode scramblers includes a self-compensating regime wherein the GD STD periodically alternates between increasing and decreasing trends, similar to systems with group-delay compensating fibers. Finally, our GD STD formulae are particularly useful in designing and optimizing mode scrambler devices, and can dramatically speed up evaluations of this objective in iterative design procedures.

The remainder of the paper is organized as follows: Section II presents the study of GD and MDL STDs in the presence of mode scramblers in MDM-MMF systems using the generalized Jones representation under the assumption of strong intra-group coupling and no inter-group coupling in the fiber. Sections II-A and II-B provide analytical expressions for GD STD and MDL STD, respectively. Section II-C includes simulation results that verify the analytical expressions. Section III provides design criteria for a strong mode scrambler. Section IV presents the study of GD STD in the presence of non-zero inter-group coupling in the MMF using the generalized Stokes representation. Section V presents the equivalence between mode-scrambler-aided coupling combined with strong random intra-group coupling and strong random coupling between all modes. Section VI discusses some aspects of long-haul system design with mode scramblers. Section VII presents the conclusion.

An extended version of this paper [21] includes eight appendices providing derivations of key analytical results.

II. GD SPREAD AND MDL IN THE PRESENCE OF MODE SCRAMBLERS

A long-haul MDM transmission system using MMF with periodic amplification and mode scrambling is shown in Fig. 1. We define a *span* as a section of MMF between successive mode scramblers. Each span is of length L_S and each span has identical statistical properties. In the generalized Jones
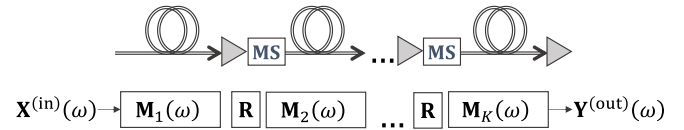

Fig. 1. Diagram of long-haul MDM transmission link over K spans of propagation with periodic amplification and mode scrambling. MS: Mode scrambler.

representation, $\mathbf{X}^{(\text{in})}(\omega)$ and $\mathbf{Y}^{(\text{out})}(\omega)$ denote the complex baseband input and output electric field vectors, respectively, in D spatial and polarization modes at an angular frequency ω. The matrices $\mathbf{M}_1(\omega), \mathbf{M}_2(\omega), \ldots, \mathbf{M}_K(\omega)$ denote the $D \times D$ frequency-dependent random transfer matrices for the MMF spans including the amplifiers, and $\mathbf{R}$ denotes the $D \times D$ frequency-independent deterministic transfer matrix of each of the mode scramblers. The D modes are grouped into N_g mode groups; modes in a mode group have nearly equal propagation constants. The set of modes in the ith mode group is denoted by $\mathcal{M}_i$ and the degeneracy is given by $d_i = |\mathcal{M}_i|$, where $|\cdot|$ represents the cardinality. The $D \times D$ transfer matrix of the system can be written as a product of the transfer matrices:

$$\mathbf{M}_{\text{tot}}(\omega) = \mathbf{M}_K(\omega)\mathbf{R}\mathbf{M}_{K-1}(\omega)\mathbf{R}\ldots\mathbf{M}_2(\omega)\mathbf{R}\mathbf{M}_1(\omega). \quad (1)$$

In the remainder of this paper, the ω dependence of the transfer matrices will not be explicitly shown. For simplicity in analysis, we assume the mode-averaged gain of each amplifier perfectly compensates for the attenuation in the preceding MMF span, and any MDL caused by fiber splices, amplifiers, and mode scramblers is captured in $\mathbf{R}$. Therefore, we can model $\mathbf{M}_1(\omega), \mathbf{M}_2(\omega), \ldots, \mathbf{M}_K(\omega)$ by unitary matrices.

In typical GI-MMFs, strong intra-group coupling occurs over length scales of $\sim$1-500 m [22], [23]. During fiber manufacturing, processes such as spinning can also increase intra-group coupling strength. On the other hand, inter-group coupling is weak and typically occurs over larger length scales. In published experiments, inter-group coupling has been observed over a wide range of lengths. Some report significant inter-group coupling after a few tens of kilometers [9], [10], [24], [25], whereas others report weak inter-group coupling even after $\sim$ 100 km of propagation [26].

In this section, we assume strong intra-group and no inter-group coupling in the fiber to obtain conservative estimates of how the GD and MDL STDs depend on the mode scrambler transfer matrix $\mathbf{R}$. Therefore, the matrices $\mathbf{M}_1(\omega), \mathbf{M}_2(\omega), \ldots, \mathbf{M}_K(\omega)$ can be modeled by block unitary matrices [6]. The assumption of no inter-group coupling yields robust design criteria for mode scramblers in Section III. Later, in Section IV, we discuss how inter-group coupling in the MMF affects the GD STD.

A. Analysis of GD Spread in the Generalized Jones Space

In this subsection, we provide an analysis of GD STD for the system shown in Fig. 1 using the generalized Jones representation. The group-delay operator (GDO) of the end-to-end link is

defined as [27]:

$$\mathbf{G}_{\text{tot}} = -\jmath \mathbf{M}_{\text{tot}}^{-1} \frac{\partial \mathbf{M}_{\text{tot}}}{\partial \omega}, \quad (2)$$

where $\jmath = \sqrt{-1}$. The GDs $\boldsymbol{\tau}_{\text{tot}}$ are the eigenvalues of the GDO. Since we are only interested in the GD STD, we assume that the eigenvalues sum to zero $\sum_{l=1}^{D} \tau_{\text{tot},l} = 0$. The GD STD can be written as

$$\sigma_{\text{GD}}(K) = \sqrt{\frac{\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}}{D}} = \sqrt{\frac{\mathbb{E}\left\{\text{tr}\left(\mathbf{G}_{\text{tot}}^H \mathbf{G}_{\text{tot}}\right)\right\}}{D}}, \quad (3)$$

where $\mathbb{E}\{\cdot\}$ denotes the expected value, $\|\cdot\|$ represents the Eulidean or l^2-norm, $\text{tr}(\cdot)$ denotes the matrix trace operation, and $(\cdot)^H$ denotes the conjugate transpose.

We can derive an analytical expression for $\mathbb{E}\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\}$ as a function of the number of spans K:

$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K) \approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0$$
$$+ \boldsymbol{\tau}_0^T \left(2 \sum_{k=1}^{K-1} (K-k) \mathbf{D} \left(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}\right)^k\right) \boldsymbol{\tau}_0, \quad (4)$$

where $\tau_{0,i} = L_S \sum_{l \in \mathcal{M}_i} \beta_{1,l}/d_i$, $i = 1, 2, \ldots N_g$ is the average GD of the ith mode group over a span of propagation, $\beta_{1,l}$ is the uncoupled GD per unit length of the lth mode such that $\sum_{l=1}^{D} \beta_{1,l} = 0$, $\sigma_{\text{intra},i} \propto \sqrt{L_S L_{\text{intra},i}}$ is the intra-group GD STD of the ith mode group with intra-group coupling length $L_{\text{intra},i}$ over a span of propagation, $\mathbf{D}$ is the $N_g \times N_g$ diagonal matrix of mode group degeneracy, $\mathbf{D}[i,i] = d_i$, and $\boldsymbol{\mathcal{P}}$ is the $N_g \times N_g$ mode-group power coupling matrix of the mode scrambler. The i,j element of $\boldsymbol{\mathcal{P}}$ is defined as the power transferred by the mode scrambler from mode group j to mode group i:

$$\boldsymbol{\mathcal{P}}[i,j] = \sum_{l \in \mathcal{M}_i} \sum_{m \in \mathcal{M}_j} |\mathbf{R}[l,m]|^2. \quad (5)$$

The diagonal elements of $\boldsymbol{\mathcal{P}}$ indicate the power retained by each mode group after transmission through the mode scrambler. In the case without mode scramblers, $\mathbf{R}$ and $\boldsymbol{\mathcal{P}}$ become diagonal matrices.

The analytical expression in (4) can be derived by expanding (2) using the definition of $\mathbf{M}_{\text{tot}}$ in (1) and using results from random matrix theory. The derivation assumes MDL is negligible, in which case, $\mathbf{R}$ and $\mathbf{M}_{\text{tot}}$ are unitary. This approximation is valid when the overall MDL is low. A detailed derivation is provided in [21, Appendix B].

The expression (4) for GD STD indirectly depends on the mode scrambler transfer matrix $\mathbf{R}$ through the power coupling matrix $\boldsymbol{\mathcal{P}}$. Eq. (4) simplifies the expression (3) containing $D \times D$-dimensional matrices to a sum containing real-valued scalars, N_g-dimensional vectors, and $N_g \times N_g$-dimensional matrices. The first two terms are proportional to K and are independent of the mode scrambler. The first term is the intra-group coupling term. The second term is the sum of the squares of the average

delays of the mode groups. The rest of the terms are dependent on the mode scrambler and form an arithmetico-geometric progression with ratio $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$.

In the absence of a mode scrambler, $\mathbf{R}$ is a diagonal matrix and differs from the $D \times D$ identity matrix owing to MDL from the amplifier and from splicing. As a result, $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ is diagonal and close to the $N_g \times N_g$ identity matrix, and we get

$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K) \approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0$$
$$+ 2 \sum_{k=1}^{K-1} (K-k) \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0$$
$$= K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K^2 \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0,$$

which is proportional to K^2 for large K. Consequently, σ_{GD} is proportional to K. In the presence of a strong mode scrambler, we expect the sum of the arithmetico-geometric progression to be much smaller than the first two terms. Therefore, we can write $\mathbb{E}\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\}(K) \approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0$, which is proportional to K. Consequently, σ_{GD} is proportional to $\sqrt{K}$.

B. Analysis of MDL in the Generalized Jones Space

In this subsection, we provide an analysis of MDL STD for the system shown in Fig. 1 using the generalized Jones representation. The overall MDL of the system is described by the eigenvalues of the modal gain operator (MGO) defined as [28]:

$$\mathbf{F}_{\text{tot}} = \mathbf{M}_{\text{tot}} \mathbf{M}_{\text{tot}}^H. \quad (6)$$

The MGO is Hermitian-symmetric and can be written as $\mathbf{F}_{\text{tot}} = \mathbf{V} \boldsymbol{\Lambda}_{\text{tot}}^{(g)} \mathbf{V}^H$, where $\boldsymbol{\Lambda}_{\text{tot}}^{(g)} = \text{diag}\{e^{g_{\text{tot},1}}, \ldots, e^{g_{\text{tot},D}}\}$ is a diagonal matrix of positive eigenvalues representing the optical power gains and $\mathbf{V}$ is a unitary output beam-forming matrix [28]. Without loss of generality, the real-valued logarithmic power gains $\mathbf{g}_{\text{tot}} = [g_{\text{tot},1}, \ldots, g_{\text{tot},D}]^T$ sum to zero, $g_{\text{tot},1} + \cdots + g_{\text{tot},D} = 0$. The standard deviation of the overall MDL σ_{MDL} is a useful quantity for characterizing MDL [28] and is given by

$$\sigma_{\text{MDL}}(K) = \sqrt{\frac{1}{D} \mathbb{E}\left\{\|\mathbf{g}_{\text{tot}}\|^2\right\}}. \quad (7)$$

σ_{MDL} and $\mathbf{g}_{\text{tot}}$ are measured in log-power-gain units and can be converted to decibels by multiplying by $\gamma = 10/\ln 10 \approx 4.34$, i.e., $\sigma_{\text{MDL}}(\text{dB}) = \gamma \sigma_{\text{MDL}}(\text{log power gain})$. $\mathbb{E}\{\|\mathbf{g}_{\text{tot}}\|^2\}$ can be written in terms of $\mathbf{M}_{\text{tot}}$ as

$$\mathbb{E}\left\{\|\mathbf{g}_{\text{tot}}\|^2\right\}(K) = \mathbb{E}\left\{\text{tr}\left(\left(\log \mathbf{M}_{\text{tot}} \mathbf{M}_{\text{tot}}^H\right)^2\right)\right\}. \quad (8)$$

In practical systems, wherein the overall MDL is low, the logarithm can be approximated by the first-order term. Under this approximation, we can derive an analytical expression for $\mathbb{E}\{\|\mathbf{g}_{\text{tot}}\|^2\}(K)$:

$$\mathbb{E}\left\{\|\mathbf{g}_{\text{tot}}\|^2\right\}(K) \approx 2 \left(\mathbf{d}^T \left(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}\right)^{K-1} \mathbf{1}_{N_g} - D\right), \quad (9)$$

TABLE I
SIMULATION PARAMETERS

Number of spatial and polarization modes	D	12
Number of mode groups	N_g	3
Mode group degeneracy	$\mathbf{d}$	$[2, 4, 6]^T$
Number of spans	K	100
Uncoupled GDs per unit length	$\boldsymbol{\beta}_1$	$[0.149, 0.149,$ $0.042, 0.042,$ $0.042, 0.042,$ $-0.080, -0.080,$ $-0.076, -0.076,$ $-0.076, -0.076]^T$ ps/km
Span length	L_S	50 km
Mode-averaged GDs over one span	$\boldsymbol{\tau}_0$	$[25.2, 7.05, -13.1]^T$ ns
Number of sections per span in the multi-section model	N	50
Number of fiber realizations		10,000
Per-span MDL STD	$\sigma_{\text{MDL},0}$	0.2 dB

where $\mathbf{d}$ is the column vector of mode-group degeneracies, and $\mathbf{1}_{N_g}$ is the all-ones column vector of dimension N_g. Similar to the expression for GD STD, (9) is also an $N_g \times N_g$-dimensional expression and depends on the mode scrambler transfer matrix $\mathbf{R}$ indirectly through the power coupling matrix $\mathcal{P}$. A detailed derivation of (9) is provided in [21, Appendix C].

C. Multi-Section Simulations

In this subsection, we verify the analytical expressions for GD and MDL STDs in Sections II-A and II-B through multi-section simulations similar to those discussed in [6], [29]. The simulation parameters are provided in Table I. The mode indexing follows decreasing order of propagation constants: $[\{LP_{01,x},\ LP_{01,y}\}, \{LP_{11a,x},\ LP_{11a,y},\ LP_{11b,x},\ LP_{11b,y}\}, \{LP_{02,x}, LP_{02,y}, LP_{21a,x}, LP_{21a,y}, LP_{21b,x}, LP_{21b,y}\}]$. Numerous fibers are generated using the parameters provided in Table I. Each of the K spans is modeled as a concatenation of $N = 50$ smaller sections with random block unitary coupling matrices to simulate strong intra-group coupling and no inter-group coupling. For $k = 1, 2, \ldots, K$, we have

$$\mathbf{M}_k = \mathbf{T}_{k,N+1}\boldsymbol{\Lambda}_k\mathbf{T}_{k,N} \ldots \mathbf{T}_{k,2}\boldsymbol{\Lambda}_k\mathbf{T}_{k,1},$$

where $\boldsymbol{\Lambda}_k$ is a frequency-dependent diagonal matrix given by

$$\boldsymbol{\Lambda}_k = \text{diag}\left\{e^{(\jmath\omega\beta_{1,1}L_S/N)}, \ldots, e^{(\jmath\omega\beta_{1,D}L_S/N)}\right\},$$

and $\mathbf{T}_{k,1}, \ldots, \mathbf{T}_{k,N+1}$ are frequency-independent and independent and identically distributed (IID) random block unitary coupling matrices. Each of their block-diagonal components is independently distributed according to the Haar measure.

To verify the GD and MDL STD formulae, we compare the GD spread and MDL performance in three scrambling scenarios.
1) *No mode scrambling:* In this case, the mode scrambler transfer matrix is diagonal, $\mathbf{R}_0 = \boldsymbol{\Lambda}_0^{(g)}$, where $\boldsymbol{\Lambda}_0^{(g)} =$

$\text{diag}\{e^{g_{0,1}/2}, \ldots, e^{g_{0,D}/2}\}$ is the diagonal matrix of optical field gains, such that $g_{0,1} + \cdots + g_{0,D} = 0$.
2) *Moderate mode scrambling:* In this case, we generate a $D \times D$ unitary matrix $\mathcal{V}$ that is almost diagonal and incorporate MDL, $\mathbf{R}_1 = \mathcal{V}^{0.5}\boldsymbol{\Lambda}_0^{(g)}\mathcal{V}^{0.5}$.
3) *Strong mode scrambling:* Intuitively, any mode scrambler with a dense transfer matrix is a suitable candidate. Thus, we pick the normalized DFT matrix[1] $\mathbf{W}_D$ of size $D \times D$ and incorporate MDL, $\mathbf{R}_2 = \mathbf{W}_D^{0.5}\boldsymbol{\Lambda}_0^{(g)}\mathbf{W}_D^{0.5}$.

In all three cases, $\mathbf{g}_0$ is a real-valued vector chosen such that $g_{0,1} + \cdots + g_{0,D} = 0$ and the per-span MDL STD, $\gamma\sigma_{\text{MDL},0} = \gamma\sqrt{\text{tr}((\log \mathbf{RR}^H)^2)/D} = 0.2\,\text{dB}$. It is to be noted that the sources of MDL, namely, mode scrambler devices, amplifiers, and fiber slices, are assumed to be lumped at the end of a span. Amplifier and fiber splice transfer matrices can be modeled by appropriate non-unitary block-diagonal matrices. Owing to the assumption of strong intra-group coupling, this modeling is redundant. We have verified through simulations that in the absence of mode scramblers, the GD and MDL STDs caused by such devices are the same as those obtained with a diagonal transfer matrix $\mathbf{R}_0$ when the per-span MDL STD is scaled appropriately. The power coupling matrices $\mathcal{P}_0$, $\mathcal{P}_1$ and $\mathcal{P}_2$ corresponding to $\mathbf{R}_0$, $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively, are shown in Fig. 2(a).

For the three scrambling scenarios, the GDO in (2) and the MGO in (6) are computed after every fiber section and for each of the fiber realizations. The eigenvalues are then extracted to obtain numerical estimates of the GD and MDL STDs. Analytical estimates of GD STD (3) and MDL STD (7) are computed after every span of propagation.

Fig. 2(b) shows the normalized GD STD $\sigma_{\text{GD}}(K)/\sigma_{\text{GD}}(1)$ plotted as a function of the number of spans K. As indicated by the slope of the log-log plot, in the case of no mode scrambling, the GD STD increases linearly with the number of spans. In the case of strong mode scrambling, the GD STD increases with the square root of the number of spans. These observations are consistent with the behavior of GD STD in weak and strong random coupling regimes, respectively. In all cases with low overall MDL, the estimates from the analytical expression in (3) closely match the numerical estimates from the simulation. The discrepancy between the estimates for large K in the moderately scrambled scenario can be attributed to a significant overall MDL, so the assumption that $\mathbf{M}_{\text{tot}}$ is unitary does not hold.

Fig. 2(c) shows the normalized MDL STD $\sigma_{\text{MDL}}(K)/\sigma_{\text{MDL},0}$ plotted as a function of the number of spans K. $\sigma_{\text{MDL},0} = \sqrt{\text{tr}((\log \mathbf{RR}^H)^2)/D}$ is the per-span MDL STD. The step-like pattern of the multi-section simulations results from the MDL sources being localized at the ends of the spans, unlike modal dispersion, which is distributed. Overall, in the case of no mode scrambling, the MDL STD increases linearly with the number of spans. In the case of strong mode scrambling, the MDL STD increases with the square root of the number of spans. These observations are consistent with the behavior of the MDL STD

---

[1]The l, m element of the DFT matrix is $\mathbf{W}_D[l, m] = (1/\sqrt{D}) * \exp(2\pi\jmath(l-1)(m-1)/D)$.
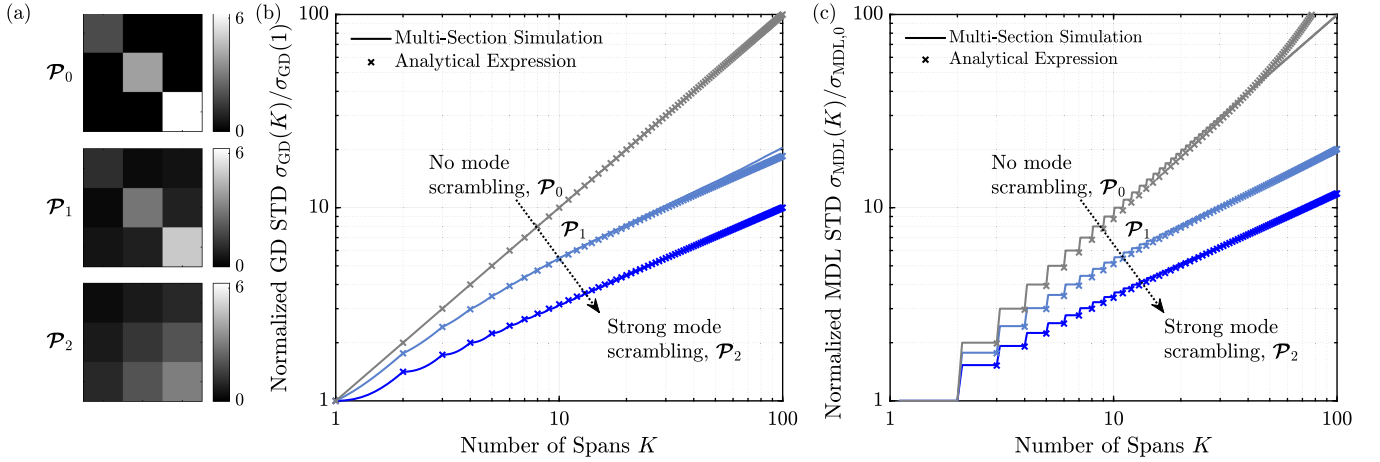
Fig. 2. Simulations of three mode scrambling scenarios. (a) Power coupling matrices; (b) GD STD as a function of the number of spans for three scrambling scenarios. The solid lines correspond to numerical estimates from multi-section simulations and the x-markers correspond to analytical estimates from the formula (3). The GD STD $\sigma_{\mathrm{GD}}(K)$ is normalized by the one-span GD STD $\sigma_{\mathrm{GD}}(1)$; (c) MDL STD as a function of the number of spans for three scrambling scenarios. The solid lines correspond to numerical estimates from multi-section simulations and the x-markers correspond to analytical estimates from the formula in (7). The MDL STD $\sigma_{\mathrm{MDL}}(K)$ is normalized by the per-span MDL STD $\sigma_{\mathrm{MDL},0}$.

in the weak and strong random coupling regimes, respectively, in links with low overall MDL. In all cases with low overall MDL, the estimates from the analytical expression (7) closely match the numerical estimates from the simulation. The difference between the estimates for large K in the case with no mode scrambling can be attributed to large overall MDL ($> 13\,\mathrm{dB}$), where the first-order approximation of the logarithm in (8) is not accurate.

While we show the comparison of simulation and analytical results for only one fiber type, we have verified that the formulae are applicable to fibers with different numbers of modes or mode groups as long as they show strong intra-group coupling and weak inter-group coupling.

III. DESIGN CRITERIA FOR A STRONG MODE SCRAMBLER

In this section, we derive sufficient conditions on $\mathbf{R}$ to obtain GD and MDL STDs similar to those of systems with strong random mode coupling, wherein the STDs are proportional to the square root of the number of spans [6]. Owing to our pessimistic assumption of no inter-group coupling in the MMFs, the estimates of GD and MDL STDs obtained here are typically higher than in practical scenarios with non-zero inter-group coupling in the MMFs. In view of the conservative assumption made, the obtained design criteria are robust.

As the closed-form expressions of GD STD (4) and MDL STD (7) discussed in Section II depend on higher powers of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$, we are interested in the spectral properties of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ and the behavior of $(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}})^k$ as $k \to \infty$. In this regard, the Perron-Frobenius theory for non-negative matrices [30] is quite useful.

Properties of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$: Consider the eigenvalue decomposition $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{-1}$, where $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_{N_g}\}$ is the diagonal matrix of eigenvalues and $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_g}]$ is the matrix of right eigenvectors. Without loss of generality, $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{N_g}|$. The spectral radius ρ of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ is defined as the

maximum absolute value of the eigenvalues, $\rho(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}) = |\lambda_1|$. The following are the properties of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$:

1) $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ is a non-negative matrix, as is evident from the definition of $\boldsymbol{\mathcal{P}}$ in (5). This means that the dominant eigenvalue is equal to the spectral radius; $\lambda_1 = \rho(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}})$.

2) If the mode scrambler is power conserving, i.e., $\mathbf{R}$ is unitary, then:

a) $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ is a right stochastic matrix, defined as a non-negative square matrix, with each row summing to 1. $\boldsymbol{\mathcal{P}}\mathbf{D}^{-1}$ is a left stochastic matrix, defined as a non-negative square matrix, with each column summing to 1.

b) $\mathbf{1}_{N_g}$ is a right eigenvector of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ with eigenvalue $\lambda_1 = 1$. $\mathbf{d}^T$ is the corresponding left eigenvector.

$$\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}\mathbf{1}_{N_g} = \mathbf{1}_{N_g}$$

$$\mathbf{d}^T\mathbf{D}^{-1}\boldsymbol{\mathcal{P}} = \mathbf{d}^T.$$

Similarly, $\mathbf{1}_{N_g}^T$ is a left eigenvector of $\boldsymbol{\mathcal{P}}\mathbf{D}^{-1}$ with eigenvalue $\lambda_1 = 1$. $\mathbf{d}$ is the corresponding right eigenvector:

$$\mathbf{1}_{N_g}^T\boldsymbol{\mathcal{P}}\mathbf{D}^{-1} = \mathbf{1}_{N_g}^T$$

$$\boldsymbol{\mathcal{P}}\mathbf{D}^{-1}\mathbf{d} = \mathbf{d}.$$

c) $\lambda_1 = 1$ is also the dominant eigenvalue of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$. The spectral radius, $\rho(\mathbf{D}^{-1}\boldsymbol{\mathcal{P}})$ is therefore also equal to 1.

Design Criteria for $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$: A practical mode scrambler device is not power-conserving and will exhibit MDL. However, a well-designed mode scrambler should have low intrinsic MDL. Hence, the dominant eigenvalue of $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ will be close to 1. The conditions on $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ for a mode scrambler such that the GD and MDL STDs are proportional to the square root of K are:

1) $\mathbf{D}^{-1}\boldsymbol{\mathcal{P}}$ is a primitive matrix.

A matrix $\mathbf{B} \in \mathbb{C}^{n \times n}$ is said to be *irreducible* if and only if for any two distinct indices $1 \leq i, j \leq n$, there is a sequence of nonzero elements of $\mathbf{B}$ of the form

$\{\mathbf{B}[i, i_1], \mathbf{B}[i_1, i_2], \dots, \mathbf{B}[i_m, j]\}$. $\mathbf{B}$ is said to be *primitive* if it has real non-negative entries and there exists a positive integer k for which every entry in $\mathbf{B}^k$ is positive. A primitive matrix is a special case of a non-negative irreducible matrix such that only one eigenvalue has an absolute value equal to its spectral radius. It should be noted that:

$$\mathcal{P} \text{ is primitive} \iff \mathbf{D}^{-1}\mathcal{P} \text{ is primitive.}$$

2) The non-dominant eigenvalues of $\mathbf{D}^{-1}\mathcal{P}$ should be close to zero.

The conditions can be understood as follows. In the first condition, the power coupling matrix implies all mode groups interact with each other in one or more passes through the mode scrambler. For example, in the simulations presented in Section II-C, $\mathcal{P}_0$ is a reducible matrix but $\mathcal{P}_1$ and $\mathcal{P}_2$ are irreducible matrices. As a result, the GD STD and the MDL STD corresponding to $\mathcal{P}_0$ are higher than those corresponding to $\mathcal{P}_1$ and $\mathcal{P}_2$.

Reducibility is a fundamental issue in the problem we are trying to solve. In the simulations in Section II-C, weak inter-group coupling in the fiber spans leads to a reducible representation of the GDO and the MGO; each mode group can be considered an independent subsystem, and the mode-group-averaged delays and gains would drift apart as the number of spans increases. As a result, the $D \times D$ GDO and MGO lead to expressions for the GD and MDL STDs involving $N_g \times N_g$ matrices. In the case of strong random coupling between all modes, the expressions could have been simplified to use 1×1 scalar representations.

$(\mathbf{D}^{-1}\mathcal{P})^k$ can be expanded as $\lambda_1^k \mathbf{u}_1 \acute{\mathbf{u}}_1^T + \dots + \lambda_{N_g}^k \mathbf{u}_{N_g} \acute{\mathbf{u}}_{N_g}^T$, where $\acute{\mathbf{u}}_1^T, \dots, \acute{\mathbf{u}}_{N_g}^T$ are the rows of $\mathbf{U}^{-1}$. If $\mathbf{D}^{-1}\mathcal{P}$ is primitive, then $(\mathbf{D}^{-1}\mathcal{P})^k$ tends to the rank-one matrix $\lambda_1^K \mathbf{u}_1 \acute{\mathbf{u}}_1^T$ as k tends to infinity. Therefore, the contribution from the terms with $(\mathbf{D}^{-1}\mathcal{P})^k$ reduces as k increases.

The second condition is related to the convergence rate. If the non-dominant eigenvalues are close to zero, then $(\mathbf{D}^{-1}\mathcal{P})^k \approx \lambda_1^K \mathbf{u}_1 \acute{\mathbf{u}}_1^T$ is valid even for smaller values of k. The non-dominant eigenvalues of $\mathbf{D}^{-1}\mathcal{P}_2$ are smaller in magnitude than those of $\mathbf{D}^{-1}\mathcal{P}_1$. Hence, the GD and MDL STDs corresponding to $\mathcal{P}_1$ are higher than those of $\mathcal{P}_2$.

GD STD and MDL STD for a Strong Mode Scrambler: We can show that GD and MDL STDs are proportional to $\sqrt{K}$ when $\mathbf{D}^{-1}\mathcal{P}$ satisfies the above-mentioned conditions. In the low-MDL regime, $\lambda_1 \approx 1$, $\mathbf{u}_1 \approx \mathbf{1}_{N_g}$, $\acute{\mathbf{u}}_1 \approx \mathbf{d}/D$. If $\mathbf{D}^{-1}\mathcal{P}$ is primitive with small non-dominant eigenvalues,

$$(\mathbf{D}^{-1}\mathcal{P})^k \approx \lambda_1^k \mathbf{u}_1 \acute{\mathbf{u}}_1^T \approx \frac{\lambda_1^k}{D} \mathbf{1}_{N_g} \mathbf{d}^T.$$

From (4) we obtain

$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K) \approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K\boldsymbol{\tau}_0^T \mathbf{D}\boldsymbol{\tau}_0$$

$$+ \frac{1}{D}\boldsymbol{\tau}_0^T \left(2 \sum_{k=1}^{K-1} (K-k)\lambda_1^k \mathbf{1}_{N_g} \mathbf{d}^T\right)\boldsymbol{\tau}_0.$$
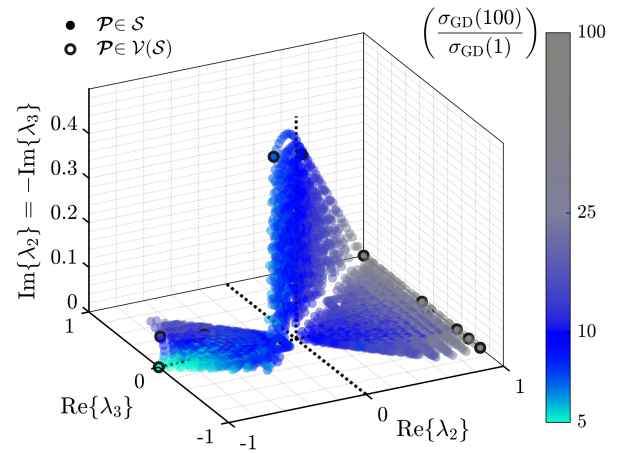

Fig. 3. Non-dominant eigenvalues of $\mathbf{D}^{-1}\mathcal{P}$ for various $\mathcal{P}$ and $\mathbf{D} = \text{diag}\{2, 4, 6\}$. Each circle represents a different $\mathcal{P}$ of size 3×3 from the set $\mathcal{S}$ such that $\mathbf{D}^{-1}\mathcal{P}$ is right-stochastic and $\mathcal{P}\mathbf{D}^{-1}$ is left-stochastic. The darkened circles represent $\mathcal{P}$ matrices from the set of vertices $\mathcal{V}(\mathcal{S})$ of the convex hull $\mathcal{S}$. The color indicates the GD STD after $K = 100$ spans of propagation normalized to $\sigma_{\text{GD}}(1)$. The $\boldsymbol{\tau}_0$ parameter is taken from Table I.

Noting that $\mathbf{d}^T \boldsymbol{\tau}_0 = \sum_{l=1}^D \beta_{1,l} = 0$, we get

$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K) \approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K\boldsymbol{\tau}_0^T \mathbf{D}\boldsymbol{\tau}_0, \qquad (10)$$

which is proportional to K. Therefore, σ_{GD} is proportional to $\sqrt{K}$. Similarly, (9) becomes

$$\mathbb{E}\left\{\|\mathbf{g}_{\text{tot}}\|^2\right\}(K) \approx 2\left(\frac{1}{D}\lambda_1^{K-1}\mathbf{d}^T \mathbf{1}_{N_g} \mathbf{d}^T \mathbf{1}_{N_g} - D\right)$$

$$= 2D\left(\lambda_1^{K-1} - 1\right)$$

$$\approx 2D\left(K-1\right)\left(\lambda_1 - 1\right), \qquad (11)$$

which is proportional to K. Therefore, σ_{MDL} is proportional to $\sqrt{K}$.

Eigenvalue Space of $\mathbf{D}^{-1}\mathcal{P}$: Fig. 3 shows the dependence of the GD STD on the eigenvalues of $\mathbf{D}^{-1}\mathcal{P}$ for $N_g = 3$, $\mathbf{d} = [2, 4, 6]^T$, and negligible MDL. The other simulation parameters have been taken from Table I. We generate numerous $\mathcal{P}$ matrices such that $\mathbf{D}^{-1}\mathcal{P}$ is right stochastic and $\mathcal{P}\mathbf{D}^{-1}$ is left stochastic, evaluate the formula in (4) at $K = 100$, and study the dependence of the normalized GD STD on the eigenvalues of $\mathbf{D}^{-1}\mathcal{P}$. The set of all such matrices

$$\mathcal{S} = \left\{\mathcal{P} \middle| \mathcal{P} \in \mathbb{R}^{N_g \times N_g}, \mathcal{P} > 0, \right.$$

$$\left. \mathbf{D}^{-1}\mathcal{P}\mathbf{1}_{N_g} = \mathbf{1}_{N_g}, \mathbf{1}_{N_g}^T \mathcal{P}\mathbf{D}^{-1} = \mathbf{1}_{N_g}^T\right\}$$

forms a convex hull. $\mathcal{V}(\mathcal{S})$ denotes the vertices of the convex hull. The eigenvalues of the generated $\mathbf{D}^{-1}\mathcal{P}$ are of the form $[\lambda_1 = 1, \lambda_2, \lambda_3]$ such that the non-dominant eigenvalues λ_2 and λ_3 are either complex conjugates of one another or both real. Fig. 3 is a scatter plot of λ_2 and λ_3. Each solid circle represents a $\mathcal{P}$ from $\mathcal{S}$ and the color of the circle indicates the normalized GD STD $\sigma_{\text{GD}}(K = 100)/\sigma_{\text{GD}}(1)$.

The figure shows that the blue region represents strong scrambling with normalized GD STD close to $\sqrt{K} = 10$. In this region, λ_2 and λ_3 are close to zero and the corresponding $\mathbf{D}^{-1}\mathcal{P}$ matrices are dense and irreducible, making them suitable for scrambling. An example would be the all-ones matrix with the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 0$, and $\lambda_3 = 0$. On the other hand, the gray region represents weak scrambling with normalized GD STD close to $K = 100$. In this region, λ_2 and/or λ_3 are close to 1 and the corresponding $\mathbf{D}^{-1}\mathcal{P}$ matrices are not sufficiently dense and are not suitable for scrambling. We also observe a cyan region around $(\lambda_2, \lambda_3) = (-1, 0)$ that exhibits normalized GD STD values less than $\sqrt{K} = 10$. This region corresponds to special $\mathcal{P}$ matrices that satisfy $\boldsymbol{\tau}_0^T \mathcal{P} \boldsymbol{\tau}_0 < 0$, or in other words, $\mathcal{P}\boldsymbol{\tau}_0$ is anti-correlated with $\boldsymbol{\tau}_0$. This anti-correlation causes the arithmetico-geometric progression terms in (4) to alternate in sign. These mode scramblers are "self-compensating", and are discussed in Section VI-C.

IV. EFFECT OF INTER-GROUP COUPLING ON GD STD

In Section II, the inter-group coupling length was assumed to be infinite to obtain a closed-form expression of GD STD (4) in the generalized Jones representation. While this assumption yields pessimistic estimates of GD STD and robust design criteria, it is important to quantify the impact of finite inter-group coupling that occurs in physical MMFs. Hence, in this section, we remove this assumption and show that we can obtain a similar closed-form expression incorporating a finite inter-group coupling length for the case of $D = 12$ modes and $N_g = 3$ mode groups. The mathematical framework of the generalized Stokes space is preferred here as it conveniently incorporates random distributed coupling between any pair of modes, making it suitable for the analytical treatment of accumulated GD STD.

Analysis in the generalized Stokes space involves representing the $D \times D$ GDO in the basis of $D^2 - 1$ generalized Pauli matrices [31], [32]. This leads to stochastic differential equations governing the evolution of the $(D^2 - 1) \times 1$ modal dispersion vector in that basis. Subsequently, the expected modal dispersion vector and GD STD can be described by deterministic ordinary differential equations (ODEs). Although these equations contain $(D^2 - 1)$-dimensional vectors and matrices, the GD STD can be well approximated by equations containing only $(D/2 - 1)$-dimensional vectors and matrices by assuming that polarization mode dispersion is negligible and that all mode coupling interactions are captured by appropriately chosen inter- and intra-group coupling lengths. Detailed analysis is presented in [21, Appendix D] and the closed-form expression for GD STD obtained from this analysis is similar in structure to the one derived in Section II-A. This enables us to obtain simple correction factors to include the effect of inter-group coupling in (4). We verify all the resulting expressions using numerical modeling.

From the generalized Stokes analysis, we include simple correction factors in (4) to capture finite inter-group coupling effects:

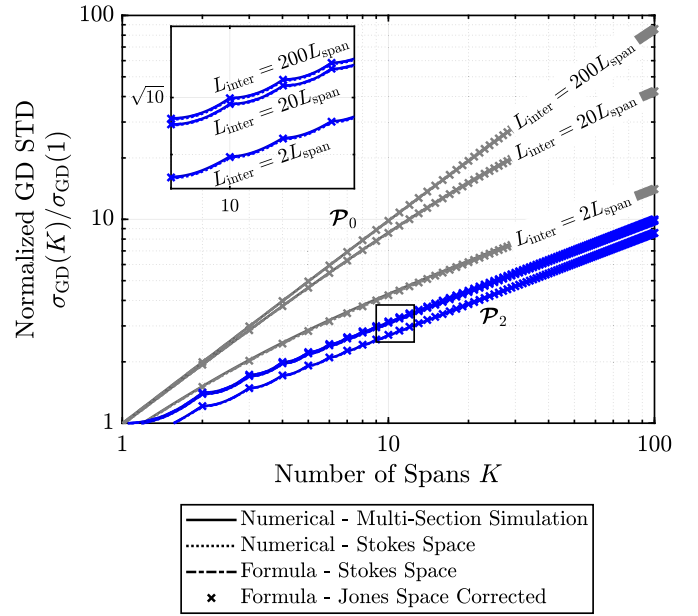$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K)$$

Fig. 4. GD STD as a function of the number of spans for various inter-group coupling lengths and two scrambling scenarios. The solid lines correspond to numerical estimates from multi-section simulations, the dotted lines correspond to the numerical evaluation of the two ODEs [21, eq. S29] and [21, eq. S30] from the Stokes analysis, the dashed lines correspond to the analytical formula from the Stokes analysis [21, eq. S41], and the crosses correspond to the analytical formula from the Jones analysis with corrections for a finite inter-group coupling length (12). The GD STD $\sigma_{\text{GD}}(K)$ is normalized by the one-span GD STD $\sigma_{\text{GD}}(1)$.

$$\approx K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K \frac{2\zeta_2\left(L_S, L_{\text{inter}}/2\right)}{L_S^2} \boldsymbol{\tau}_0^T \mathbf{D} \boldsymbol{\tau}_0$$

$$+ \boldsymbol{\tau}_0^T \left(2 \sum_{k=1}^{K-1} (K-k) \mathbf{D} \left(\mathbf{D}^{-1}\mathcal{P}\right)^k e^{\left(-\frac{2(k-1)L_S}{L_{\text{inter}}}\right)} \right) \boldsymbol{\tau}_0$$

$$\times \left(\frac{\zeta_1\left(L_S, L_{\text{inter}}/2\right)}{L_S} \right)^2, \tag{12}$$

where L_{inter} is the inter-group coupling length and

$$\zeta_1(z, L) = L\left(1 - \exp\left(-\frac{z}{L}\right)\right),$$

$$\zeta_2(z, L) = Lz - L^2\left(1 - \exp\left(-\frac{z}{L}\right)\right).$$

Fig. 4 shows the normalized GD STD $\sigma_{\text{GD}}(K)/\sigma_{\text{GD}}(1)$ plotted as a function of the number of spans K for $L_{\text{inter}} \in \{2L_S, 20L_S, 200L_S\}$. We consider power coupling matrices $\mathcal{P}_0$ and $\mathcal{P}_2$ corresponding to the cases with no scrambling and strong scrambling, respectively, as in Section II-C. To isolate the effects of inter-group coupling, we consider no MDL in the mode scrambler transfer matrices. The plot compares the GD STDs obtained by four methods: the multi-section model, numerical evaluation of the two ODEs [21, eq. S29] and [21, eq. S30], and the analytical estimates from the Stokes formula [21, eq. S41] and from the Jones formula with corrections for finite inter-group coupling length (12). We see good agreement between the analytical formulae and the estimates from numerical simulation.

V. Equivalence to Strong Random Coupling

In deriving the conditions for a strong mode scrambler in Section III, we assumed that the mode scramblers with deterministic transfer matrices are the only source of inter-group coupling and that the fiber sections are a source of random intra-group coupling. We showed that such a design can replicate the effect of strong random coupling between all modes by analyzing GD and MDL STDs. The question of whether GD and MDL STDs expressions in Section II are definitive indicators of system performance is yet unanswered in this paper. However, as a result of the Central Limit Theorem (CLT), by combining strong deterministic mode scrambling with strong intra-group coupling, we expect the probability distributions of modal delays and modal gains to approach those in systems with strong random coupling between all modes.

In this section, we (i) demonstrate close agreement between the tail probabilities of the peak-to-peak GD spread and the cumulative distribution function (CDF) of the capacity of two systems: one with strong random coupling between all modes and the other with mode-scrambler-aided coupling satisfying the design criteria in Section III, and (ii) discuss the asymptotic convergence of the distributions of the system transfer matrices in both cases.

A. Group-Delay Tail Probabilities and Distribution of Capacity

In terms of system complexity and performance, two distributions are particularly useful, namely, the tail probabilities of the peak-to-peak GD spread and the cumulative distribution function (CDF) of the capacity [33]. The DSP complexity of the receiver equalizers depends on the statistics of the peak-to-peak GD spread. The average and outage capacities can be computed from the capacity distribution, which depends on the distribution of modal power gains. In the presence of strong random coupling between all modes, the distributions of the group delays and the modal gains are well known, and the GD STD and the MDL STD fully describe the respective distributions [6]. Moreover, the peak-to-peak GD spread and the GD STD are equivalent metrics. Using multi-section simulations, we verify here that these distributions are similar for both strong random coupling and strong deterministic mode scrambling with strong intra-group coupling, thereby validating the applicability of the formulae in Section II.

Fig. 5 shows the complementary CDF of the normalized peak-to-peak GD spread after $K = 100$ spans of propagation for mode-scrambler-aided coupling and strong random coupling. The figure highlights the tail probabilities of the peak-to-peak GD spread. The mode scrambler transfer matrix is chosen according to the design criteria in Section III. To obtain the CDF, random fiber realizations are generated and the peak-to-peak GD spread is computed as the difference between the maximum and the minimum GDs, $\Delta\tau = \max_l(\tau_{\text{tot},l}) - \min_l(\tau_{\text{tot},l})$. The quantity is then normalized by σ_{GD}.

Fig. 6 shows the CDF of the capacity after $K = 100$ spans of propagation for mode-scrambler-aided coupling and strong random coupling for a signal-to-noise ratio (SNR) of 10 dB. The
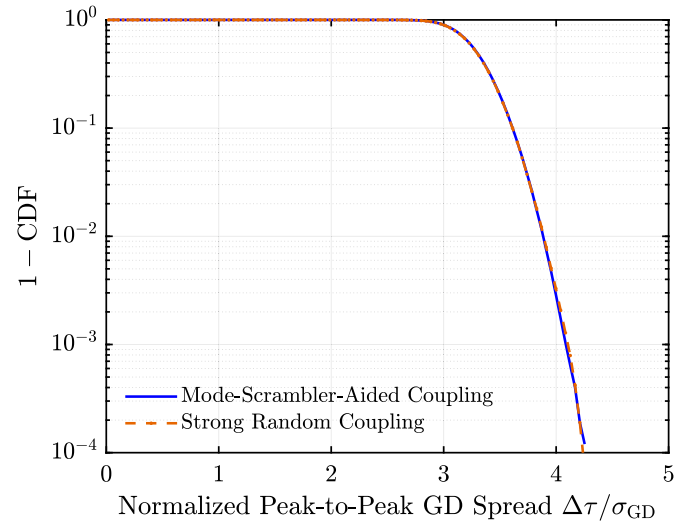

Fig. 5. Complementary CDF of the peak-to-peak GD spread normalized to the GD STD after $K = 100$ spans of propagation. The blue solid line corresponds to mode-scrambler-aided deterministic coupling for $D = 12$, $N_g = 3$, $\mathbf{D} = \text{diag}\{2, 4, 6\}$ and the orange dashed line corresponds to strong random coupling for $D = 12$.
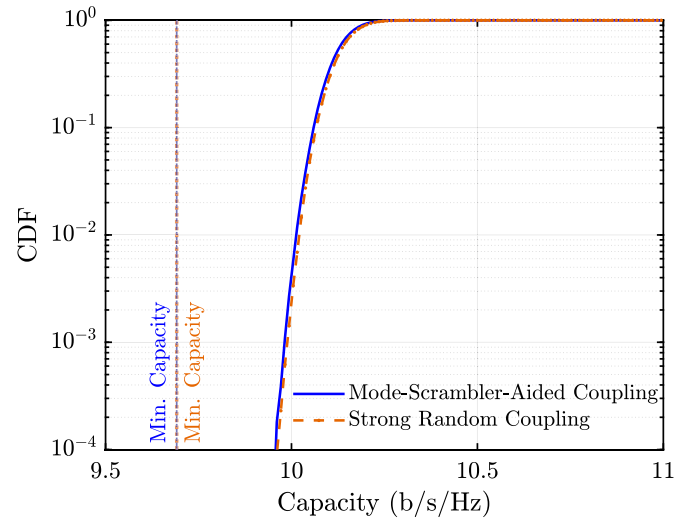

Fig. 6. CDF of capacity after $K = 100$ spans of propagation. The blue solid line corresponds to mode-scrambler-aided deterministic coupling for $D = 12$, $N_g = 3$, $\mathbf{D} = \text{diag}\{2, 4, 6\}$ and the orange dashed line corresponds to strong random coupling for $D = 12$, for an SNR of 10 dB. The vertical dotted lines indicate the minimum capacities.

per-span MDL STD $\sigma_{\text{MDL},0}$ is set to 0.2 dB. The channel capacity, measured in b/s/Hz, assuming channel-state information is not available, is defined as [8]

$$C = \log_2 \det\left(\mathbf{I}_D - \frac{\text{SNR}}{D}\widehat{\mathbf{M}}_{\text{tot}}\widehat{\mathbf{M}}_{\text{tot}}^H\right),$$

where $\widehat{\mathbf{M}}_{\text{tot}} = \mathbf{M}_{\text{tot}}/\alpha_0$, and the SNR is in linear units. α_0 is the ensemble-averaged mode-averaged gain, and $\alpha_0 = \sqrt{\sum_{l=1}^{D} \mathbb{E}\{e^{g_{\text{tot},l}}\}/D}$. Capacity is lower bounded by $C_{\min} = D\log_2(1 + \text{SNR}/\alpha_0^2 D)$ [8].

Figs. 5 and 6 are consistent with the assumption that combining periodic deterministic scrambling with random intra-group coupling results in the same distribution of modal delays and modal gains as a system with strong random coupling between all modes.

B. Asymptotic Convergence of the Distribution of the System Transfer Matrix

Transfer matrices of fiber sections with strong random coupling can be modeled by random unitary matrices distributed according to the Haar measure [33]. In a system with strong random coupling between all modes, wherein each fiber section effects an independent random unitary transformation, the GDO can be written as a sum of section-wise GDOs with IID eigenvectors. As a result of the CLT, the overall GDO asymptotically approaches a Gaussian Unitary Ensemble (GUE) with zero trace.

In a system with deterministic mode scrambling and strong random intra-group coupling, we expect that the system transfer matrix asymptotically approaches, in distribution, a random unitary matrix distributed according to the Haar measure. As a result, the GDO can be written as a sum of the span-wise GDOs with asymptotically IID eigenvectors. Here too, the CLT holds and the overall GDO asymptotically approaches a GUE with zero trace.

Arguments supporting the convergence of the distribution of $\mathbf{M}_{\text{tot}}$ according to the Haar measure on the unitary group in the presence of strong mode scrambling, and the convergence of $\mathbf{G}_{\text{tot}}$ to a GUE is provided in [21, Appendix G]. We pessimistically assume no inter-group coupling in the fiber. For simplicity, we also assume that the mode scrambler transfer matrix $\mathbf{R}$ is unitary. We define a Markov Chain on the system transfer matrix with the number of spans K as the Markov "time-step". In this framework, the system transfer matrix is a continuous-valued Markov "state" on the compact topological group of $D \times D$ unitary matrices. We argue that this Markov chain has key properties like irreducibility, aperiodicity, and recurrence when $\mathbf{R}$ satisfies the design criteria in Section III, and consequently establishes that the Law of Large Numbers (LLN) and the CLT hold for a series summation of certain classes of functions of the Markov state, including the GDO.

The eigenvalue distribution of a GUE with zero trace is well known [33]. The asymptotic convergence of the GDO therefore establishes that the mean and variance of the eigenvalues are sufficient statistics The mean is zero owing to the zero-trace property and the variance is given by (4). While we assumed that MDL is not present, it can be argued that when the overall MDL is low, the distribution of the eigenvalues of the MGO (measured in decibels) converges to a GUE with zero trace. Although the analysis in [21, Appendix G] indicates asymptotic convergence, our simulations demonstrate a rapid rate of convergence to the limiting distributions. This is supported by the close agreement shown in Figs. 5 and 6 for $K = 100$ spans.

VI. DISCUSSION

In this section, we discuss some aspects of long-haul system design with mode scramblers.

A. Non-Identical Mode Scramblers

The derivation of the analytical expressions in Section II assumed that all mode scramblers have identical transfer matrices. While this scenario is pessimistic in terms of mode-scrambling performance, it leads to robust design criteria. In a practical system, the mode scramblers could have different transfer matrices by design or due to manufacturing variations. In such cases, the formulae can be modified to incorporate the different power coupling matrices. In both cases, if individual mode scrambler devices satisfy the design criteria, then the GD and MDL STDs will resemble those of a system with strong mode coupling. A system design can be verified by analytical modeling using the measured power coupling matrices of fabricated mode scramblers.

B. Spatial Whiteness of Amplifier Noise

Spatial whiteness of the total amplified spontaneous emission (ASE) noise is another consequence of strong mode scrambling. The ASE noises from the amplifiers are not individually spatially white owing to MDL. The $D-$dimensional noise vectors from each amplifier experience mode scrambling and transformation by the fiber matrices before they reach the receiver. Following the analysis in [28, Discussion], the noise correlation matrices can be written in terms of the mode-dependent noise variances of each amplifier, the fiber transfer matrices $\{\mathbf{M}_k\}$ and $\mathbf{R}$, which has a form very similar to the terms in the expansion of the GDO [21, eq. S3]. The equations are provided in [21, Appendix H]. Equivalently, we can write the noise correlation matrices as a function of the Markov state. Their expected values are diagonal matrices and can be computed similarly to the analysis in [21, Appendix B]. When the number of spans is large, owing to the LLN, we can argue that the overall noise correlation matrix converges to a constant times the identity matrix; in other words, the noise is spatially white.

C. Group-Delay Self-Compensation

A good design for the mode scrambler does not necessarily solve all system-level problems. Strong mode scrambling is equivalent to strong random coupling with an effective inter-group coupling length of L_S: $\sigma_{\text{GD}}(K) \approx \sigma_{\beta_1}\sqrt{K L_S L_{\text{inter,eff}}}$, where $\sigma_{\beta_1} = \sqrt{(\beta_{1,1}^2 + \cdots + \beta_{1,D}^2)/D}$ and $L_{\text{inter,eff}} = L_S$. For optimum performance, the transmission fiber must have low uncoupled GD STD σ_{β_1}, and the optical devices must have low intrinsic MDL. Since modal dispersion is a distributed effect, placing multiple mode scramblers between successive amplifiers can be beneficial as it reduces the multiplicative factor $\sqrt{L_{\text{inter,eff}}}$ in σ_{GD}. However, there is a trade-off: the overall MDL will increase as more mode scramblers are used. Another approach is to co-design mode scramblers with transmission fibers. This can potentially result in $L_{\text{inter,eff}}$ being less than L_S. This corresponds to the cyan region in Fig. 3, where the power coupling matrix is such that $\mathcal{P}\boldsymbol{\tau}_0$ is anti-correlated with $\boldsymbol{\tau}_0$ and exhibits GD compensating behavior.

Scrambling to reduce the GD spread involves frequently splitting and redirecting data signals to different modes such that no
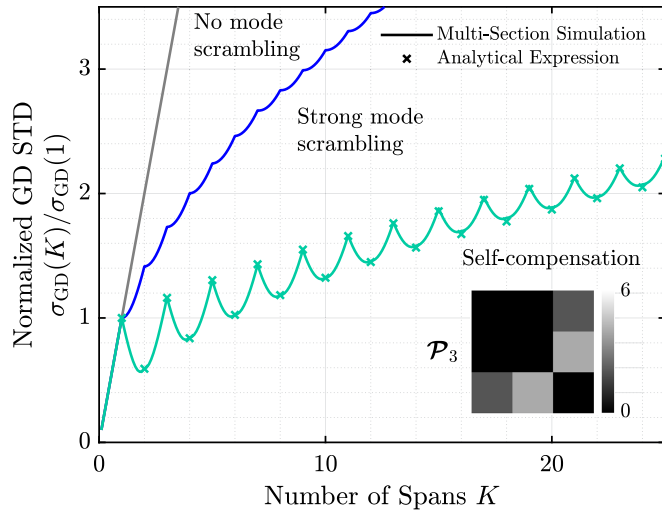
Fig. 7. Self-compensation with a mode permutator for $\mathbf{D} = \text{diag}\{2, 4, 6\}$. The solid lines correspond to numerical estimates from multi-section simulations and the x-markers correspond to analytical estimates from the formula in (3). GD STD $\sigma_{\text{GD}}(K)$ is normalized by one-span GD STD $\sigma_{\text{GD}}(1)$. The inset shows the power coupling matrix $\mathcal{P}_3$ used in obtaining the self-compensation curve.

data signal consistently travels in the fastest or the slowest mode. An optimal solution can be obtained by carefully redirecting data signals such that they transfer between propagation in fast and slow modes. This results in a net delay spread that is periodically minimal. This forms the basis of GD compensation using multiple transmission fiber types [32], [34], [35]. GD compensation can also be performed by careful mode permutations. In Fig. 3, the power coupling matrix corresponding to the minimum GD STD is

$$\mathcal{P}_3 = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & 0 \end{bmatrix},$$

which routes all the power from the first and second mode groups to the third mode group and vice-versa. We can see intuitively that this is the best option because the average group delays of the first and second mode groups are positive, while the average group delay of the third mode group is negative, $\boldsymbol{\tau}_0 = [25.2, 7.05, -13.1]^T$ ns. A matrix $\mathbf{R}_3$ with the above power coupling matrix is the anti-diagonal Exchange Matrix $\mathbf{J}_{12}$ whose elements are 1 on the anti-diagonal and 0 otherwise. The appropriate term for these devices is "mode permutators", as they do not scramble the modes. The technique of mode permutation has been proposed in the literature for the management of GD spread [18], [19], [20] and MDL caused by splices [36].

Fig. 7 shows the normalized GD STD plotted as a function of the number of spans K for a self-compensating mode permutator with the above power coupling matrix. Similarly to strong scrambling, the GD STD for the self-compensating mode permutator increases with the square root of the number of spans but $L_{\text{inter,eff}}$ is less than L_S. The alternating pattern in the GD STD is expected as a result of the anti-correlation of $\mathcal{P}_3\boldsymbol{\tau}_0$ and $\boldsymbol{\tau}_0$. $\mathbf{D}^{-1}\mathcal{P}_3$ has the eigenvalues $[\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 0]$, and satisfies the following property: $(\mathbf{D}^{-1}\mathcal{P}_3)^k = \mathbf{D}^{-1}\mathcal{P}_3$ when k is odd and $(\mathbf{D}^{-1}\mathcal{P}_3)^k = (\mathbf{D}^{-1}\mathcal{P}_3)^2$ when k is even. Therefore,

(4) for even K becomes,

$$\mathbb{E}\left\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\right\}(K) \approx$$
$$\left(K \sum_{i=1}^{N_g} \sigma_{\text{intra},i}^2 + K\boldsymbol{\tau}_0^T \mathbf{D}\left(\mathbf{I}_{N_g} - (\mathbf{D}^{-1}\mathcal{P}_3)^2\right)\boldsymbol{\tau}_0\right),$$

which is also proportional to K, similar to the expression for GD STD for a strong mode scrambler in (10). However, there is an extra factor of $\mathbf{I}_{N_g} - (\mathbf{D}^{-1}\mathcal{P})^2$ compared to (10); therefore $L_{\text{inter,eff}} < L_S$.

It should be noted that $\mathbf{D}^{-1}\mathcal{P}_3$ is an irreducible matrix but not primitive. Hence, it does not satisfy the conditions of a strong mode scrambler for arbitrary $\boldsymbol{\tau}_0$. For a different transmission fiber, the optimum power coupling matrix may differ from $\mathcal{P}_3$, especially if the signs of $\boldsymbol{\tau}_0$ change. We can define an optimization problem to obtain the best power coupling matrix for any given $\boldsymbol{\tau}_0$ by minimizing (4) for even and large K. We start by minimizing (4) for $K = 2$ spans by formulating the following optimization problem

$$\begin{aligned} \underset{\mathcal{P}}{\text{minimize}} \quad & \boldsymbol{\tau}_0^T \mathcal{P} \boldsymbol{\tau}_0 \\ \text{subject to} \quad & \mathcal{P}[i,j] \geq 0, \, 1 \leq i, j \leq N_g \\ & \mathbf{D}^{-1}\mathcal{P}\mathbf{1}_{N_g} = \mathbf{1}_{N_g} \\ & \mathbf{d}^T\mathbf{D}^{-1}\mathcal{P} = \mathbf{d}^T, \end{aligned} \quad (13)$$

where the constraints enforce that $\mathcal{P}$ is a non-negative matrix and $\mathbf{R}$ is a unitary matrix. The constraints are convex and the cost function is linear in $\mathcal{P}$. Therefore, we have a linear programming problem. The constraint set is feasible. It is well-known that the optimum value in a linear programming problem is found on the boundary of the constraint set [37]. To obtain an optimum $\mathbf{D}^{-1}\mathcal{P}$, we need to only search over the set of vertices of the convex hull $\mathcal{S}$. If multiple vertices attain the same minimum value of the cost function, then any convex combination of those vertices is a valid optimum.

An optimum $\mathcal{P}$ obtained for the minimization problem in (13) does not necessarily minimize the GD STD for $K > 2$ spans, wherein the cost function is non-convex in $\mathcal{P}$. However, we have verified through gradient descent searches in the constraint set that the optimum $\mathcal{P}$ for $K = 2$ and $K > 2$ are identical for $\boldsymbol{\tau}_0$ values similar to the one in Table I.

An important challenge for this approach lies in realizing a device with an optimal power coupling matrix. Possible device solutions include LPFGs and back-to-back multiplexer-demultiplexer using photonic lanterns, which require further study. To obtain maximum benefit, it is also necessary to carefully design the transmission fiber to yield a favorable $\boldsymbol{\tau}_0$. Using mode permutators for self-compensation has limitations caused by random inter-group mode coupling, similar to GD compensation using different fibers [12]. The GD STD derived from the Stokes space analysis [21, eq. S41] can be useful in evaluating the performance of systems designed with this technique. While this technique optimizes GD STD, its effect on MDL STD needs further study. These are topics for future research.

VII. Conclusion

In this paper, we studied the effects of the mode scrambler transfer matrix on modal dispersion and mode-dependent gain/loss. Using a generalized Jones representation, we derived analytical expressions for GD STD and overall MDL STD in terms of the mode-group power coupling matrix for systems with strong random intra-group coupling and weak inter-group coupling in the MMF. Through multi-section simulations, we verified the analytical expressions in the low-MDL regime. We also proposed the following design criteria for a mode scrambler to obtain GD and MDL STDs proportional to the square root of the number of spans of propagation: the power coupling matrix $\mathcal{P}$ should be primitive and the non-dominant eigenvalues of $\mathbf{D}^{-1}\mathcal{P}$ should be close to zero. Systems using such mode scramblers exhibit inter-group coupling lengths equal to the span length L_S. Using a generalized Stokes representation, we re-derived analytical expressions for GD STD to incorporate the effects of inter-group coupling in the MMF.

We argued that deterministic mode scrambling in the presence of strong intra-group coupling is asymptotically equivalent to strong random coupling. We also verified that the tail probabilities of peak-to-peak GD spread and channel capacity match in the two scenarios. Finally, we discussed the self-compensation regime wherein the GD STD periodically increases and decreases and has an effective inter-group coupling length less than L_S.

Acknowledgment

References

[1] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photon.*, vol. 7, no. 5, pp. 354–362, May 2013. [Online]. Available: https://www.nature.com/articles/nphoton.2013.94

[2] R.-J. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks," *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, May 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6170861

[3] B. J. Puttnam, G. Rademacher, and R. S. Luís, "Space-division multiplexing for optical fiber communications," *Optica*, vol. 8, no. 9, pp. 1186–1203, Sep. 2021. [Online]. Available: https://opg.optica.org/optica/abstract.cfm?uri=optica-8-9-1186

[4] P. J. Winzer, R. Ryf, and S. Randel, "Spatial multiplexing using multiple-input multiple-output signal processing," in *Optical Fiber Telecommunications*, 6th ed., in Series Optics and Photonics, I. P. Kaminow, T. Li, and A. E. Willner, Eds., Boston, MA, USA: Academic Press, Jan. 2013, ch. 10, pp. 433–490. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123969606000109

[5] H. Srinivas, O. Krutko, and J. M. Kahn, "Efficient integrated multimode amplifiers for scalable long-haul SDM transmission," *IEEE J. Lightw. Technol.*, vol. 41, no. 15, pp. 4989–5002, Aug. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10064021/

[6] K.-P. Ho and J. M. Kahn, "Linear propagation effects in mode-division multiplexing systems," *IEEE J. Lightw. Technol.*, vol. 32, no. 4, pp. 614–628, Feb. 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6615972/

[7] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *IEEE J. Lightw. Technol.*, vol. 29, no. 24, pp. 3719–3726, Dec. 2011. [Online]. Available: https://ieeexplore.ieee.org/document/6060841/

[8] D. A. A. Mello, H. Srinivas, K. Choutagunta, and J. M. Kahn, "Impact of polarization- and mode-dependent gain on the capacity of ultra-long-haul systems," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 303–318, Jan. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8918466/

[9] R. V. Jensen, L. Grüner-Nielsen, N. H. L. Wong, Y. Sun, Y. Jung, and D. J. Richardson, "Demonstration of a 9 LP-mode transmission fiber with low DMD and loss," in *Proc. Opt. Fiber Commun. Conf.*. Los Angeles, CA, USA: OSA, 2015, pp. W2A–34. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=OFC-2015-W2A.34

[10] R. Ryf et al., "Mode-multiplexed transmission over conventional graded-index multimode fibers," *Opt. Exp.*, vol. 23, no. 1, pp. 235–246, Jan. 2015. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-23-1-235

[11] N. K. Fontaine et al., "Characterization of space-division multiplexing systems using a swept-wavelength interferometer," in *Proc. 2013 IEEE Opt. Fiber Commun. Conf. Expo. Nat. Fiber Optic Engineers Conf.*, Mar. 2013, pp. 1–3. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6533121

[12] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Group delay management and multiinput multioutput signal processing in mode-division multiplexing systems," *IEEE J. Lightw. Technol.*, vol. 34, no. 11, pp. 2867–2880, Jun. 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7409913/

[13] L. Raddatz, I. White, D. Cunningham, and M. Nowell, "An experimental and theoretical study of the offset launch technique for the enhancement of the bandwidth of multimode fiber links," *J. Lightw. Technol.*, vol. 16, no. 3, pp. 324–331, Mar. 1998. [Online]. Available: https://ieeexplore.ieee.org/document/661357

[14] Y. Liu, Y. Jung, Z. Yang, L. Zhang, and D. Richardson, "Wideband and low-loss mode scrambler for few-mode fibers based on distributed multiple point-loads," *IEEE Photon. J.*, vol. 13, no. 3, pp. 1–7, Jun. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9448421/

[15] Y. Zhao, H. Chen, N. K. Fontaine, J. Li, R. Ryf, and Y. Liu, "Broadband and low-loss mode scramblers using CO_2-laser inscribed long-period gratings," *Opt. Lett.*, vol. 43, no. 12, pp. 2868–2871, Jun. 2018. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?uri=ol-43-12-2868

[16] D. Askarov and J. M. Kahn, "Long-period fiber gratings for mode coupling in mode-division-multiplexing systems," *IEEE J. Lightw. Technol.*, vol. 33, no. 19, pp. 4032–4038, Oct. 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7163511/

[17] J. Li et al., "Design and demonstration of mode scrambler supporting 10 modes using multiplane light conversion," in *Proc. 2018 IEEE Eur. Conf. Opt. Commun.*, Sep. 2018, pp. 1–3. [Online]. Available: https://ieeexplore.ieee.org/document/8535381

[18] K. Shibahara, T. Mizuno, H. Ono, K. Nakajima, and Y. Miyamoto, "Long-haul DMD-unmanaged 6-mode-multiplexed transmission employing cyclic mode-group permutation," in *Proc. IEEE Opt. Fiber Commun. Conf. Exhib. 2020*. San Diego, CA, USA: 2020, pp. Th3H.3. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=OFC-2020-Th3H.3

[19] G. D. Sciullo et al., "Reduction of modal dispersion in a long-haul 15-mode fiber link by means of mode permutation," in *Proc. IEEE 49th Eur. Conf. Opt. Commun. (2023)*, vol. 2023, Oct. 2023, pp. 250–253. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10484924

[20] M. v. d. Hout et al., "Transmission of 273.6 tb/s over 1001 km of 15-mode multi-mode fiber using C-band only 16-QAM signals," *J. Lightw. Technol.*, vol. 42, no. 3, pp. 1136–1142, Feb. 2024. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?uri=jlt-42-3-1136

[21] A. Vijay, O. Krutko, R. Refaee, and J. M. Kahn, "Modal statistics in mode-division-multiplexed systems using mode scramblers," Sep. 2024, *arXiv:2409.06908v1*. [Online]. Available: https://arxiv.org/abs/2409.06908v1

[22] R. Gabet et al., "Complete dispersion characterization of few mode fibers by OLCI technique," *J. Lightw. Technol.*, vol. 33, no. 6, pp. 1155–1160, Mar. 2015. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?uri=jlt-33-6-1155

[23] L. Grüner-Nielsen et al., "Measuring distributed mode scattering in few mode fibers with high and low differential group delay," in *2012 IEEE Photon. Soc. Summer Topical Meeting Ser.*, Jul. 2012, pp. 193–194, iSSN: 2376-8614. [Online]. Available: https://ieeexplore.ieee.org/document/6280768

[24] P. Genevaux, M. Salsi, A. Boutin, F. Verluise, P. Sillard, and G. Charlet, "Comparison of QPSK and 8-QAM in a three spatial modes transmission," *IEEE Photon. Technol. Lett.*, vol. 26, no. 4, pp. 414–417, Feb. 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6691930/

[25] N. Bai et al., "Mode-division multiplexed transmission with inline few-mode fiber amplifier," *Opt. Exp.*, vol. 20, no. 3, pp. 2668–2680, Jan. 2012. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-20-3-2668

[26] R. Ryf et al., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 x 6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6074912/

[27] S. Fan and J. M. Kahn, "Principal modes in multimode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, Jan. 2005. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=ol-30-2-135

[28] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Exp.*, vol. 19, no. 17, p. 16612–16635, Aug. 2011. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-19-17-16612

[29] A. Vijay and J. M. Kahn, "Effect of higher-order modal dispersion in direct-detection mode-division-multiplexed links," *J. Lightw. Technol.*, vol. 41, no. 6, pp. 1670–1683, Mar. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9969895/citations?tabFilter=papers#citations

[30] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: SIAM, May 2010.

[31] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Stokes-space analysis of modal dispersion in fibers with multiple mode transmission," *Opt. Exp.*, vol. 20, no. 11, pp. 11718–11733, May 2012. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-20-11-11718

[32] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Delay spread reduction in mode-division multiplexing: Mode coupling versus delay compensation," *J. Lightw. Technol.*, vol. 33, no. 21, pp. 4504–4512, Nov. 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7234843/

[33] K. P. Ho, J. M. Kahn, I. P. Kaminow, T. Li, and A. E. Willner, "Mode coupling and its impact on spatially multiplexed systems," in *Optical Fiber Telecommunications*, 6th ed., in Series Optics and Photonics, I. P. Kaminow, T. Li, and A. E. Willner, Eds. Boston, MA, USA: Academic Press, 2013, ch. 11, pp. 491–568, iSSN: 15575837. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123969606000110

[34] S. Randel et al., "Mode-multiplexed 6x20-GBd QPSK transmission over 1200-km DGD-compensated few-mode fiber," in *Proc. OFC/NFOEC*, 2012, pp. 1–3.

[35] N. Bai and G. Li, "Equalizer tap length requirement for mode group delay-compensated fiber link with weakly random mode coupling," *Opt. Exp.*, vol. 22, no. 4, p. 4247–4255, Feb. 2014. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-22-4-4247

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

[37] V. Chvátal, *Linear Programming*, in Series of books in the mathematical sciences. New York, NY, USA: W. H. Freeman, 1983. [Online]. Available: https://books.google.com/books?id=DN20_tW_BV0C

Anirudh Vijay (Member, IEEE) received the B.Tech. and M.Tech. degrees in electrical engineering from the Indian Institute of Technology Madras, Chennai, India, in 2019. He is currently working toward the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA. His research interests include optical communications, mode-division multiplexing, and data-center applications.

Oleksiy Krutko (Student Member, IEEE) received the B.S. degree in electrical engineering from the University of Texas at Austin, Austin, TX, USA, in 2020. He is currently working toward the Ph.D. degree from Stanford University, Stanford, CA, USA. His research interests include optical fiber communications and photonic devices.

Rebecca Refaee received the B.S. degree in mathematics and the M.S. degree in electrical engineering in 2024 from Stanford University, Stanford, CA, USA, where she is currently working towards the Ph.D. degree in electrical engineering. Her research interests include optical communications and mode-division multiplexing.

Joseph M. Kahn (Life Fellow, IEEE) received the A.B., M.A. and Ph.D. degrees in physics from the University of California, Berkeley, CA, USA, in 1981, 1983, and 1986 respectively. From 1987 to 1990, he was with the AT&T Bell Laboratories. In 1989, he demonstrated the first successful synchronous i.e., coherent detection using semiconductor lasers, achieving record receiver sensitivity. From 1990 to 2003, he was with the Electrical Engineering and Computer Sciences Faculty, Berkeley. In 1992, he demonstrated coherent detection of QPSK. In 1999, D. S. Shiu along with himself authored or coauthored the first work on probabilistic shaping for optical communications. In 1990s and early 2000s, he along with his collaborators performed seminal work on indoor and outdoor free-space optical communications and multi-input multi-output wireless communications. In 2000, K. P. Ho along with himself 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 he applied to patent the first electronic compensation of fiber Kerr nonlinearity. In 2009, StrataLight was acquired by Opnext. In 2003, he became a Professor of electrical engineering with the E. L. Ginzton Laboratory, Stanford University, Stanford, CA. He 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 himsely (and G. Li independently) invented simplified digital backpropagation for compensating fiber Kerr nonlinearity and dispersion. Since 2004, he and collaborators has been studying propagation, modal statistics, spatial multiplexing and imaging in multimode 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. His research interests include 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. He was the recipient of the National Science Foundation Presidential Young Investigator Award in 1991.