

Closed-Form Statistics and Design of Mode-Division-Multiplexing Systems Employing Group-Delay Compensation and Mode Permutation

Anirudh Vijay¹, Member, IEEE, Nika Zahedi¹, Oleksiy Krutko¹, Member, IEEE, Rebecca Refaee¹, and Joseph M. Kahn¹, Life Fellow, IEEE

I. INTRODUCTION

Abstract—Excessive accumulation of group-delay spread increases computational complexity and affects tracking of receiver-based multi-input multi-output signal processing, posing challenges to long-haul mode-division multiplexing in multi-mode fiber. Group-delay compensation, which involves periodically exchanging propagating signals between modes with lower and higher group delays, can potentially reduce group-delay spread. In this work, we investigate two group-delay compensation schemes: conventional compensation, which alternates fiber types with opposite group delay orderings, and self-compensation, which employs a single fiber type with periodically inserted mode permuters. We provide analytical expressions for group-delay statistics in compensated systems with arbitrary multi-mode fiber types and mode permuters, accounting for random inter-group coupling, mode scrambling, and refractive index errors. To enhance the effectiveness of group-delay compensation over the C-band, we propose optimized graded-index depressed-cladding multi-mode fibers with index profiles tailored to control mode-dependent chromatic dispersion. We also analyze the impact of fiber fabrication errors and explore mitigation strategies based on group-delay characterization and sorting. Design examples and numerical simulations demonstrate improved system performance, highlighting the trade-offs between compensation effectiveness, system complexity, and transmission losses. For a 5000 km multi-mode fiber link, the design examples achieve a group-delay standard deviation of 438 ps (effectively 6.2 ps/ $\sqrt{\text{km}}$) and 257 ps (effectively 3.6 ps/ $\sqrt{\text{km}}$) for conventional and self-compensation schemes, respectively, under ideal conditions. In the presence of random inter-group coupling (characteristic coupling length = 500km), the group-delay standard deviation increases to 803 ps (11.4 ps/ $\sqrt{\text{km}}$) and 575 ps (8.1 ps/ $\sqrt{\text{km}}$), respectively. Additionally, with refractive-index errors (error STD = 10^{-5}), the group-delay standard deviation further increases to 1.28 ns (18.1 ps/ $\sqrt{\text{km}}$) for both schemes.

Index Terms—Group-delay compensation, long-haul multi-mode fiber systems, mode-division multiplexing, mode permutations, self-compensation.

Received 21 May 2025; revised 3 September 2025; accepted 26 September 2025. Date of publication 1 October 2025; date of current version 16 December 2025. This work was supported in part by Ciena Corporation, in part by Stanford Shoucheng Zhang Graduate Fellowship, and in part by Stanford Graduate Fellowship. (Corresponding author: Anirudh Vijay.)

The authors are with the E.L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: avijay@alumni.stanford.edu; nzahedi@stanford.edu; oleksiyk@alumni.stanford.edu; becca24@stanford.edu; jmk@ee.stanford.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JLT.2025.3616280>.

Digital Object Identifier 10.1109/JLT.2025.3616280

SPACE-DIVISION multiplexing (SDM) enhances capacity, integration, and power efficiency in long-haul optical coherent communication systems [1], [2], [3]. It can be implemented using parallel single-mode fibers (SMFs), uncoupled-core or coupled-core multi-core fibers (MCFs), or multi-mode fibers (MMFs). Among these options, mode-division multiplexing (MDM) in MMFs is particularly attractive for achieving the highest level of integration [4], enabling amplification with fewer pump modes than signal modes [5].

Graded-index (GI) MMFs are preferred for long-haul transmission when the number of spatial and polarization modes exceeds six ($D > 6$) due to their inherently low group-delay (GD) standard deviation (STD) [6], a key measure of modal dispersion. Reducing the end-to-end GD STD is desirable as it decreases the digital signal processing (DSP) complexity at the receiver.

GD management strategies in the literature can be broadly subdivided into two classes. The first class focuses on the design of fibers with favorable intrinsic GD properties. This includes engineering the refractive index profile to achieve low uncoupled GDs while promoting mode coupling and optimizing wavelength-dependent characteristics such as chromatic dispersion (CD) [7]. The second class involves system-level techniques to manage GD accumulation, such as mode scrambling to enhance mode coupling [8], [9] or GD compensation by concatenating fibers with complementary GD orderings. Recently, mode permutation to exchange power between slow and fast mode groups has been shown to reduce end-to-end GD STD, both theoretically [10] and experimentally [11], [12], [13].

This paper focuses on GD compensation, studying two approaches: concatenating multiple fiber types (referred to as *conventional compensation*) and using a single fiber type with mode permutation (referred to as *self-compensation*). Achieving a sufficiently low end-to-end GD STD requires the design of special fibers having modal GDs suitable for these compensation schemes. The systematic design of fibers for compensation involves optimizing the refractive index profile to achieve specific uncoupled GD values, as proposed in prior research [14], [15], [16], [17].

Several challenges impact the effectiveness of GD compensation, necessitating careful consideration in system design. One

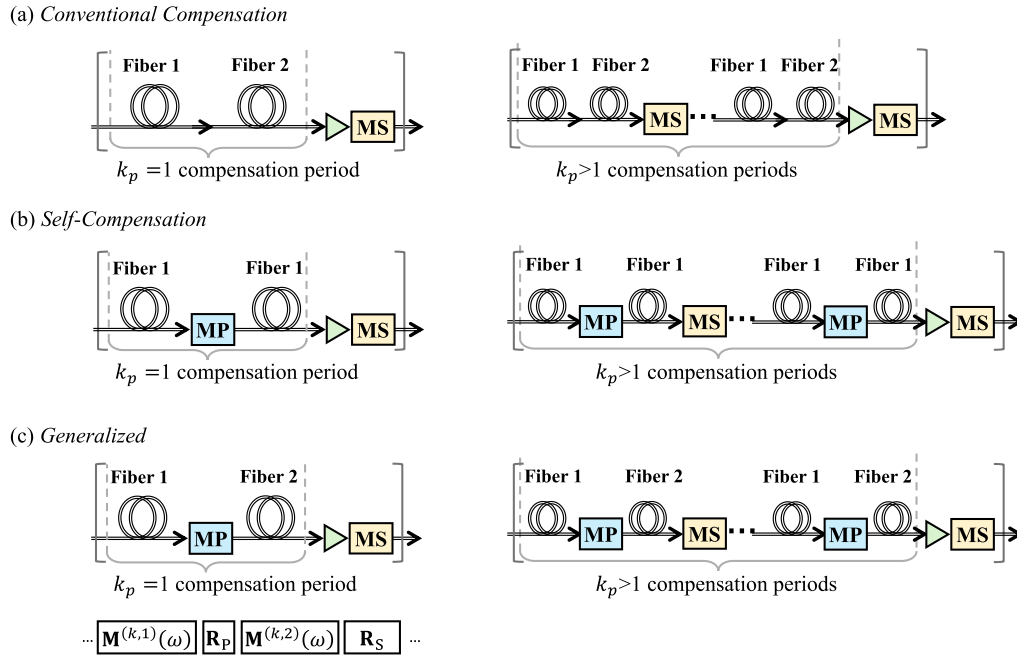


Fig. 1. Diagrams of one span of a long-haul MDM transmission link with periodic amplification and GD compensation. Each fiber span can have one or more compensation periods. (a) Conventional compensation scheme employing two fiber types and mode scramblers; (b) Self-compensation scheme employing one fiber type, mode permeters and mode scramblers; (c) Generalized compensation scheme employing up to two fiber types, mode permeters and mode scramblers. MP: Mode Permuter, MS: Mode Scrambler.

key challenge is the wavelength dependence of relevant fiber GD properties, especially mode-dependent chromatic dispersion (MDCD). A large MDCD can result in significant variations in effective GD STD across different wavelength channels. This is undesirable because receiver DSP complexity is dictated by the channel with the highest GD STD. Fiber designs should therefore aim to minimize MDCD, thus we propose incorporating pedestal-like features into the refractive index profile, similar to techniques proposed in the literature [18], [19]. Fiber fabrication errors also pose a challenge to GD compensation [15]. It is essential to assess their impact and design fibers with GD properties that maintain effective compensation despite imperfect fabrication. During system integration, manufactured fibers can be characterized, and fibers with compatible GD orderings can be selected for use in the same span to enhance compensation efficacy.

An important system-level challenge arises from random distributed inter-group coupling in MMFs. Since GD compensation relies on controlled and complete power exchanges between fiber segments, random coupling can undermine compensation efficacy [8]. To address this, it is crucial to quantify the impact of inter-group coupling on the effective GD STD and develop fiber and system designs that are resilient to this effect. One proposed mitigation strategy is to shorten the fiber segments used in GD compensation to reduce the influence of random inter-group coupling.

In this paper, we present a generalized approach to studying both conventional compensation and self-compensation of GD. We derive analytical expressions for end-to-end GD STD, incorporating the effects of mode scrambling and random inter-

group coupling. For both compensation schemes, we provide design examples of optimized GI-MMFs supporting 12 spatial and polarization modes and evaluate their GD statistics through extensive numerical simulations. We propose parameterized fiber designs and carefully select a cost function that accounts for the system-level effects of modal GDs and MDCDs. Using numerical optimization techniques such as particle swarm optimization, we systematically search for fiber designs that minimize the effective GD STD after compensation while ensuring low MDCD and robustness to random inter-group coupling and fabrication errors. Additionally, we analyze the impact of fabrication errors and introduce a semi-analytic penalty term in the GD STD expression. Our findings indicate that achieving GD STD values competitive with coupled-core MCFs requires stringent fabrication tolerances.

The rest of the paper is organized as follows: Section II presents the modeling of GD compensation systems, analytical expressions for GD STD in such systems in the presence of random inter-group coupling, and numerical simulations to validate the analytical expressions. Section III presents a fiber design strategy for MMF links supporting six spatial modes in the C-band, design examples for both conventional and self-compensation schemes, and numerical simulations to evaluate the end-to-end GD STD. Sections IV and V present discussion and conclusions, respectively.

II. GROUP-DELAY COMPENSATION IN MMF SYSTEMS

GD compensation schemes for long-haul MDM transmission systems employing MMFs with periodic amplification and mode

scrambling are shown in Fig. 1. We define a *span* as a portion of the link between successive amplifier subsystems, each including both an inline amplifier and a mode scrambler. Each span is of length L_S and has identical statistical properties. We define a *segment* as a piece of fiber of a given type. We define a *compensation period* as a pair of fiber segments whose group delays ideally cancel.¹ A span may contain one or more compensation periods, with the number of periods per span denoted by k_p .

The conventional compensation scheme (Fig. 1(a)) uses two fiber types per span whose modal GDs, relative to the mode-averaged GD, are ideally equal and opposite. The self-compensation scheme (Fig. 1(b)) uses one fiber type and a mode permuter, and ideally, the permuter exchanges power between the fastest and the slowest mode groups. We study both schemes within a generalized framework (Fig. 1(c)) that uses up to two fiber types and an optional mode permuter. In this section, we provide analytical expressions for the GD STD in systems employing GD compensation and present simulation results that validate these expressions.

As shown in Fig. 1(c), in the generalized Jones representation for D spatial and polarization modes, the matrices $\mathbf{M}_{k,1}(\omega)$, and $\mathbf{M}_{k,2}(\omega)$ denote the $D \times D$ frequency-dependent complex-baseband random electric field transfer matrices for the two MMF segments in the k th span (we are assuming $k_p = 1$ compensation periods per span for the time being). \mathbf{R}_P and \mathbf{R}_S denote the $D \times D$ frequency-independent deterministic transfer matrix of the mode permuters and scramblers, respectively. 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 i th mode group is denoted by \mathcal{M}_i and the degeneracy is given by $d_i = |\mathcal{M}_i|$, where $|\cdot|$ represents the cardinality.

A. Analysis of GD Spread in GD Compensation Systems

Following the analysis in [10], we define mode-group-averaged N_g -dimensional GD vectors (measured in seconds) of the two fiber types. For mode groups $i = 1, 2, \dots, N_g$,

$$\tau_{1,i} = \frac{L_1}{d_i} \sum_{l \in \mathcal{M}_i} \beta_1^{(1)}[l],$$

$$\tau_{2,i} = \frac{L_2}{d_i} \sum_{l \in \mathcal{M}_i} \beta_1^{(2)}[l],$$

where $\beta_1^{(m)}[l]$, $m = 1, 2$ is the uncoupled GD (measured in seconds per meter) of the mode l such that $\sum_{l=1}^D \beta_1^{(m)}[l] = 0$, and L_1 and L_2 are the lengths of the fiber segments in one compensation period. The $N_g \times N_g$ mode-group power coupling matrices of the mode permuter and mode scrambler are given by

$$\mathcal{P}[i, j] = \sum_{l \in \mathcal{M}_i} \sum_{m \in \mathcal{M}_j} |\mathbf{R}_P[l, m]|^2,$$

$$\mathcal{S}[i, j] = \sum_{l \in \mathcal{M}_i} \sum_{m \in \mathcal{M}_j} |\mathbf{R}_S[l, m]|^2,$$

¹In Section IV-D, we briefly consider a compensation scheme employing three fiber segments per compensation period.

respectively. If mode permuters are not used in the link, \mathcal{P} can be set to \mathbf{D} , the diagonal matrix of mode group degeneracies. The GD STD after K spans is given by

$$\sigma_{\text{GD}}(K) = \sqrt{\frac{\mathbb{E}\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\}}{D}}, \quad (1)$$

where $\mathbb{E}\{\|\boldsymbol{\tau}_{\text{tot}}\|^2\}$ denotes the expected squared norm of the coupled GDs and is given by (2), shown at the bottom of the next page. For fiber types $m = 1, 2$, $\sigma_{\text{intra},m,i} \propto \sqrt{L_m L_{\text{intra},i}}$ is the intra-group GD STD of the i th mode group with intra-group coupling length $L_{\text{intra},i}$ over length L_m . The factors

$$\eta_m = \frac{L_{\text{inter}}}{2L_m} \left(1 - e^{-2L_m/L_{\text{inter}}}\right), \quad (3)$$

$$\nu_m = \frac{L_{\text{inter}}}{L_m} - \frac{L_{\text{inter}}^2}{2L_m^2} \left(1 - e^{-2L_m/L_{\text{inter}}}\right) \quad (4)$$

are correction terms inspired by a Stokes-vector analysis incorporating finite characteristic length L_{inter} of random inter-group coupling [20]. L_{inter} is inversely related to the strength of random inter-group coupling. In graded-index MMFs, there is strong random intra-group coupling and weak random inter-group coupling: $L_{\text{intra},i} \ll L_m$, $i = 1, \dots, N_g$, and $L_m < L_{\text{inter}}$.

While (2) may appear complicated, a well-designed mode scrambler can make the terms in the square brackets negligible as K increases [10]. Among the remaining terms, the intra-group GD spread terms $\sigma_{\text{intra},1,i}^2 + \sigma_{\text{intra},2,i}^2$ have negligible contributions as well, since $L_{\text{intra}} \ll L_S$. The key terms in (2) are $(\nu_1 \boldsymbol{\tau}_1^T \mathbf{D} \boldsymbol{\tau}_1 + \nu_2 \boldsymbol{\tau}_2^T \mathbf{D} \boldsymbol{\tau}_2)$, the *GD accumulation* term corresponding to the fiber segments individually, and $(\eta_1 \eta_2 \boldsymbol{\tau}_1^T \mathcal{P} \boldsymbol{\tau}_2 + \eta_1 \eta_2 \boldsymbol{\tau}_2^T \mathcal{P} \boldsymbol{\tau}_1)$, the *GD compensation* term due to the correlated delays of the fiber segments. The sum of these terms corresponds to the net effective GD accumulation in one span. It is important to note that, owing to the inclusion of mode scramblers, a multiplicative factor K is involved. While the accumulation terms are inherently positive, the compensation terms may assume negative values if the fibers and/or mode permuters are designed intentionally to achieve this effect. Mathematically, we obtain the following inequalities:

$$|\boldsymbol{\tau}_1^T \mathbf{D} \boldsymbol{\tau}_1 + \boldsymbol{\tau}_2^T \mathbf{D} \boldsymbol{\tau}_2| \geq |\boldsymbol{\tau}_1^T \mathcal{P} \boldsymbol{\tau}_2 + \boldsymbol{\tau}_2^T \mathcal{P} \boldsymbol{\tau}_1|, \quad (5)$$

$$\nu_1 + \nu_2 \geq 2\eta_1 \eta_2. \quad (6)$$

Perfect compensation is possible only if both (5) and (6) are satisfied with equality. The first equality can be achieved if $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$ and (if applicable) \mathcal{P} are designed together. The second equality is achieved only if $L_{\text{inter}} \rightarrow \infty$. This means that random inter-group coupling from the fiber segments is always unfavorable to GD compensation.

B. Multi-Section Simulations

In this sub-section, we verify the analytical expression for GD STD through multi-section simulations similar to those discussed in [21], [22] for systems supporting three mode groups with a total of six spatial modes. We choose system configurations that achieve perfect GD compensation in the absence of random inter-group coupling in the fibers. However, we evaluate

TABLE I
SIMULATION PARAMETERS I

Parameter	Symbol	Value/Expression
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
Compensation periods per span	k_p	1
Span length	L_S	
Characteristic inter-group coupling length	L_{inter}	$L_{\text{inter}} \in \{2L_S, 20L_S, 200L_S\}$
Ideal Conventional Compensation		
Mode-averaged GDs over one span	(τ_1, τ_2)	$\tau_1 = -\tau_2$
Segment lengths	(L_1, L_2)	$L_1 \neq L_2$, e.g. $(0.6L_S, 0.4L_S)$
Ideal Self-Compensation		
Mode-averaged GDs over one span	(τ_1, τ_2)	$\tau_1 = \tau_2$, $\tau_{1,1} = \tau_{1,2} = -\tau_{1,3}$
Segment lengths	(L_1, L_2)	$(0.5L_S, 0.5L_S)$

the GD STD when random inter-group coupling is present. The simulation parameters are provided in Table I and the results are presented in Fig. 2.

Conventional compensation and self-compensation can achieve perfect compensation of inter-group GD spread when $L_{\text{inter}} \rightarrow \infty$ under the following conditions:

- *Ideal conventional compensation*: The delay vectors of the two fiber types are equal and opposite, $\tau_1 = -\tau_2$. A mode permuter is not required between the two fiber segments and as a result, $\mathcal{P} = \mathbf{D}$, as shown in Fig. 2(a).
- *Ideal self-compensation*: Only one fiber type is used, $\tau_2 = \tau_1$. For $N_g = 3$, the average delays of the first two mode groups are equal in magnitude and sign. The average delay of the third mode group is equal in magnitude but opposite in sign to that of the first two mode groups:

$$\tau_{1,1} = \tau_{1,2} = -\tau_{1,3}.$$

A mode permuter that transfers all the power from the first two mode groups to the third mode group, and

vice-versa, is needed between the two segments [10]. It has a power coupling matrix

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

as shown in Fig. 2(d). For this choice of τ_1 and \mathcal{P} , we find that $\tau_1^T \mathbf{D} \tau_1 + \tau_1^T \mathcal{P} \tau_1 = 0$. It can be shown that this is the only configuration that achieves zero GD STD for MDM systems with mode-group degeneracies $\mathbf{d} = [2, 4, 6]^T$.

We investigate the effect of random inter-group coupling on GD compensation for $L_{\text{inter}} \in \{2L_S, 20L_S, 200L_S\}$. Mode scramblers with dense power coupling matrices, as shown in Fig. 2(b) and (e), are used to reduce the accumulation of residual GD spread at the end of each span.

The GD STD is evaluated through multisection simulations after every fiber section and the analytical expression in (1) after every span. Fig. 2(c) and (f) show the normalized GD STD as a function of the propagation length normalized to the span length for ideal conventional compensation and self-compensation, respectively. The zig-zag pattern illustrates the processes of GD accumulation and compensation. The plots indicate that compensation remains effective even with moderate random inter-group coupling. The estimates obtained from the analytical expressions agree closely with those derived from numerical simulations. The overall trend follows a square-root relationship, which is expected owing to the use of mode scramblers. Notably, if the ratio L_{inter}/L_S decreases by a factor of 10, the asymptotic value of the GD STD increases by a factor of $\sqrt{10}$.

C. System- and Fiber-Design Considerations

From a system design perspective, to mitigate the effects of random inter-group coupling, the compensation scheme can be implemented multiple times within a single span. For k_p compensation periods per span, the GD STD will decrease by a factor of $\sqrt{k_p}$ in the absence of random inter-group coupling. In terms of fiber design for GD compensation, it is essential to reduce the magnitudes of the uncoupled delays in the fiber segments. This is because the residual GD spread caused by random inter-group

$$\begin{aligned} \mathbb{E} \left\{ \|\tau_{\text{tot}}\|^2 \right\} &= K k_p \sum_{i=1}^{N_g} (\sigma_{\text{intra},1,i}^2 + \sigma_{\text{intra},2,i}^2) + K k_p (\nu_1 \tau_1^T \mathbf{D} \tau_1 + \nu_2 \tau_2^T \mathbf{D} \tau_2 + \eta_1 \eta_2 \tau_1^T \mathcal{P} \tau_2 + \eta_1 \eta_2 \tau_2^T \mathcal{P} \tau_1) \\ &\quad + \left[2 \sum_{k=1}^{K k_p - 1} \left(\eta_1^2 \tau_1^T \left(\mathbf{D} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^k + \mathbf{D} (\mathbf{D}^{-1} \mathcal{P} \mathbf{D}^{-1} \mathbf{S})^k \right) \tau_1 e^{\frac{2L_1}{L_{\text{inter}}}} \right. \right. \\ &\quad + \eta_2^2 \tau_2^T \left(\mathbf{D} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^k + \mathbf{D} (\mathbf{D}^{-1} \mathcal{P} \mathbf{D}^{-1} \mathbf{S})^k \right) \tau_2 e^{\frac{2L_2}{L_{\text{inter}}}} \\ &\quad + \eta_1 \eta_2 \tau_1^T \left(\mathcal{P} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^k + \left(\mathcal{P} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^k \right)^T \right) \tau_2 \\ &\quad \left. \left. + \eta_1 \eta_2 \tau_1^T \left(\mathbf{D} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^{k-1} \mathbf{D}^{-1} \mathbf{S} + \left(\mathbf{D} (\mathbf{D}^{-1} \mathbf{S} \mathbf{D}^{-1} \mathcal{P})^{k-1} \mathbf{D}^{-1} \mathbf{S} \right)^T \right) \tau_2 e^{\frac{2L_S}{L_{\text{inter}}}} \right) \right]. \end{aligned} \quad (2)$$

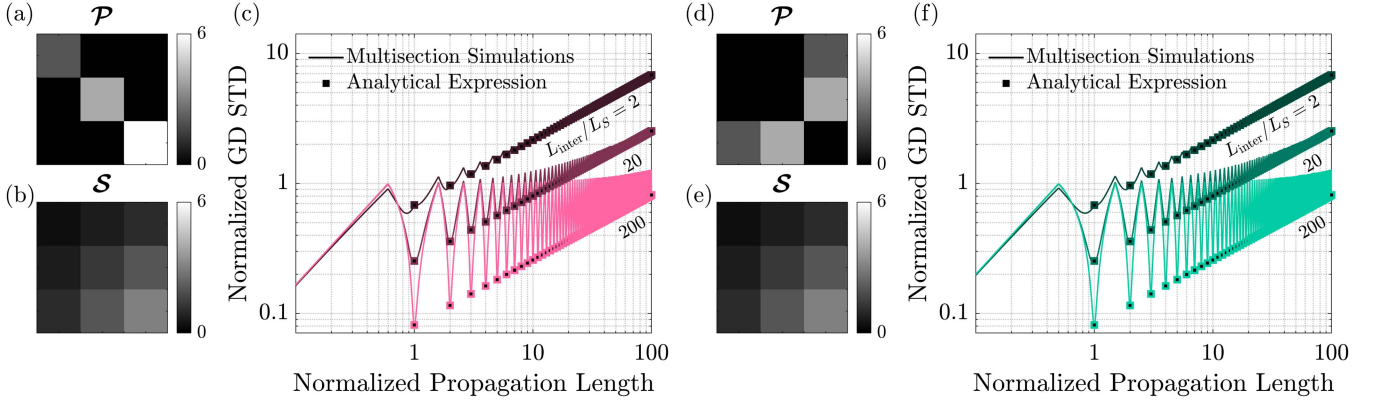


Fig. 2. Simulation of GD compensation for three values of inter-group coupling lengths for $k_p = 1$ compensation period per span. (a – c) Conventional compensation; (d – f) Self-compensation. (a, d) Mode permuter power coupling matrix; (b, e) Mode scrambler power coupling matrix; (c, f) GD STD as a function of normalized propagation length. The solid lines correspond to numerical estimates from multi-section simulations and the square markers correspond to analytical estimates from the formula (1) evaluated for propagation length equal to integer multiples of span length. The GD STD is normalized to the one-segment GD STD. The propagation length is normalized to the span length L_S .

coupling is proportional to the GD accumulation terms. If the uncoupled delays of the fiber segments are reduced by a factor κ , the GD STD will also decrease by the same factor of κ . It is important to note that even when the length of inter-group coupling is similar to the span length, GD compensation can still be more effective in managing GD than no compensation at all for specific GD orderings.

III. DESIGN OF COMPENSATING FIBERS SUPPORTING SIX SPATIAL MODES

This section presents a design strategy for fibers that support six spatial modes, which is applicable to both conventional and self-compensating schemes. We provide design examples and evaluate system performance across the C-band (1530 nm to 1565 nm). Our primary goal is to minimize GD STD, while also ensuring that the designs are tolerant to random inter-group coupling and errors in fiber fabrication. Additionally, the designs must meet specifications for macro-bending loss and cut-off wavelength [23]. The design strategy involves selecting a cost function and a search region that take all these factors into account. We parameterize the fiber design to facilitate numerical optimization techniques such as particle swarm optimization since the cost function is non-convex over the search space.

A. Cost Functions for Optimal GD Compensation

We formulate minimization problems for conventional and self-compensation. For conventional compensation, we minimize a wavelength-averaged cost function $J^{(\text{conv})}(\lambda_i)$, which is a function of $n^{(1)}, n^{(2)}$, the refractive index profiles of the two fiber types, and $f = L_1/(L_1 + L_2)$, the ratio of the length of the first fiber to the combined length:

$$\begin{aligned} & \underset{n^{(1)}, n^{(2)}, f}{\text{minimize}} \quad \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} J^{(\text{conv})}(\lambda_i), \\ & \text{subject to} \quad 1/3 \leq f \leq 2/3. \end{aligned} \quad (7)$$

We limit the ratio f to avoid large disparities in segment lengths. Similarly, for self-compensation, we minimize a wavelength-averaged cost function $J^{(\text{self})}(\lambda_i)$ which is a function of n , the refractive index profile of the self-compensation fiber:

$$\underset{n}{\text{minimize}} \quad \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} J^{(\text{self})}(\lambda_i), \quad (8)$$

The cost functions are given by

$$\begin{aligned} J^{(\text{conv})}(\lambda) = & \left\| f\bar{\beta}_1^{(1)} + (1-f)\bar{\beta}_1^{(2)} \right\| \\ & + \gamma_{\text{CD}} \left\| f\overline{\text{CD}}^{(1)} + (1-f)\overline{\text{CD}}^{(2)} \right\| \\ & + \gamma_{\text{GD}} \left(\left\| f\beta_1^{(1)} \right\| + \left\| f\beta_1^{(2)} \right\| \right), \end{aligned} \quad (9)$$

where, for fiber types $m = 1, 2$, $\beta_1^{(m)}$ is the vector of uncoupled modal GDs, $\bar{\beta}_1^{(m)}$, $m = 1, 2$ is the vector of mode-group-averaged uncoupled GDs, and $\overline{\text{CD}}^{(m)}$ is the vector of mode-group-averaged CD parameters, and

$$\begin{aligned} J^{(\text{self})}(\lambda) = & \left\{ 2(\beta_1[01] - \beta_1[11])^2 + (\beta_1[01] + \beta_1[02])^2 + \right. \\ & \left. 2(\beta_1[01] + \beta_1[21])^2 + 2(\beta_1[02] - \beta_1[21])^2 \right\}^{1/2} \\ & + \gamma_{\text{CD}} \left\| \overline{\text{CD}} \right\| \\ & + \gamma_{\text{GD}} (\|g_{\text{penalty}}(\beta_1, \beta_{th})\|), \end{aligned} \quad (10)$$

where β_1 is the vector of uncoupled modal GDs and $\beta_1[lm]$ represents the uncoupled GD of spatial mode LP_{lm} , $\overline{\text{CD}}$ is the vector of mode-group-averaged CD parameters, and g_{penalty} is a regularization term defined for a threshold β_{th} as

$$g_{\text{penalty}}(x, \beta_{th}) = \begin{cases} \beta_{th} \sqrt{1 - x/\beta_{th}}, & x < \beta_{th} \\ x - \beta_{th}, & x \geq \beta_{th}. \end{cases}$$

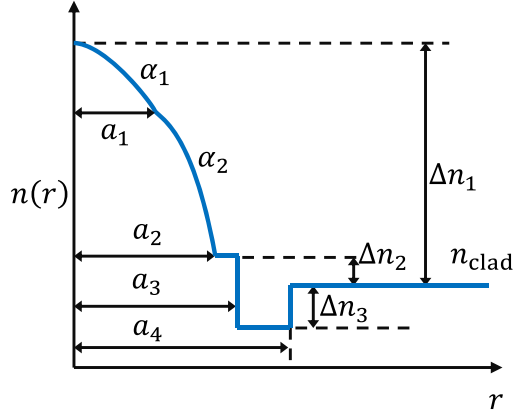


Fig. 3. Refractive index profile of a graded-index, trench-assisted compensating fiber. The inner and outer parts of the core have (possibly) different α -law gradations. There is a pedestal/cliff feature at the core-cladding boundary.

Both cost functions (9) and (10) have three components: the first captures the effective GD spread, the second captures the effective MDCD spread and is multiplied by a hyperparameter γ_{CD} , and the third captures the STD of the uncoupled GDs and is multiplied by a hyperparameter γ_{GD} .

The CD regularization term in each cost function is included in addition to the wavelength averaging in (7) and (8) to ensure wavelength-flat GD compensation performance. The GD regularization term in (9) is to ensure low uncoupled GD STD to reduce the effect of random inter-group coupling. The GD regularization term in (10) performs a similar function, except it also penalizes designs with very low uncoupled GDs, which can affect fabrication tolerances (discussed in Section IV-B).

B. Parameterized Fiber Profile

A parametric approach to fiber design facilitates the use of numerical optimization methods, such as particle-swarm optimization, which are preferred since the cost function is a non-convex function of the refractive index profile. Previous research has suggested parameterized fiber design for GD management and compensation [15], [17], [24]. Parameters defining the fiber profile, including core-to-cladding index contrast, core radius, and the α -parameter describing the core's gradation, significantly influence propagation constants and modal dispersion.

Fig. 3 plots the refractive index profile highlighting the important features and parameters. We define a refractive index profile with nine parameters as follows:

$$n(r) = \begin{cases} (n_{\text{clad}} + \Delta n_2) \sqrt{\frac{1-2\Delta_1 \left(\frac{r}{a_1}\right)^{\alpha_1}}{1-2\Delta_1}}, & 0 \leq r \leq a_1 \\ (n_{\text{clad}} + \Delta n_2) \sqrt{\frac{1-2\Delta_2 \left(\frac{r}{a_2}\right)^{\alpha_2}}{1-2\Delta_2}}, & a_1 < r \leq a_2 \\ n_{\text{clad}} + \Delta n_2, & a_2 < r \leq a_3 \\ n_{\text{clad}} - \Delta n_3, & a_3 < r \leq a_4 \\ n_{\text{clad}}, & a_4 < r \end{cases} \quad (11)$$

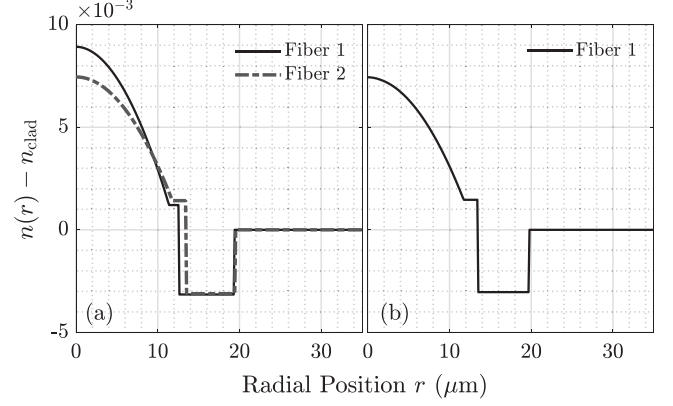


Fig. 4. Design example for compensation. Optimized fiber refractive index profiles relative to the cladding index for (a) conventional compensation and (b) self-compensation.

where r is the radial position, and

$$\Delta_1 = \frac{1}{2} \left(1 - \left(\frac{n_{\text{clad}} + \Delta n_2}{n_{\text{clad}} + \Delta n_1} \right)^2 \right), \quad (12)$$

$$\Delta_2 = \frac{\Delta_1 \left(1 - \left(\frac{a_1}{a_2} \right)^{\alpha_1} \right)}{1 - \left(\frac{a_1}{a_2} \right)^{\alpha_2} + 2\Delta_1 \left(\left(\frac{a_1}{a_2} \right)^{\alpha_2} - \left(\frac{a_1}{a_2} \right)^{\alpha_1} \right)}. \quad (13)$$

The pedestal/cliff feature at the core-cladding boundary can be adjusted to control the CD behavior. The different α parameters for the inner and outer core provide control of uncoupled GDs in different mode groups.

C. Design Examples and System Performance

We perform particle-swarm optimization [25] to search for an optimal pair of fibers for conventional compensation (7) and an optimal fiber for self-compensation (8). The hyperparameters for conventional compensation are chosen to be $\gamma_{CD} = 0.05$, $\gamma_{GD} = 0.005$, and those for self-compensation are chosen to be $\gamma_{CD} = 1$, $\gamma_{GD} = 0.1$, $\beta_{th} = 10$ ps/km. The parameter search space is chosen such that the fibers support six spatial modes in each polarization, and comply with leaky- and guided-mode bending loss standards over the C-band. The refractive index contrast between the center of the core and the trench is consistent with a germanium-free core, enabling low transmission loss [26], [27].

1) *Fiber Design Examples:* Table II provides the parameters of the optimized fibers. Fig. 4 shows the refractive index profiles relative to the cladding as a function of the radial position for the conventional compensation fibers and the self-compensation fiber. Fig. 5 shows the dispersion performance of the optimized fibers over the C-band. Here are the key observations.

Conventional Compensation:

- Fig. 5(a) shows the scaled uncoupled GDs of the spatial modes of the two fibers, $f\beta_1^{(1)}$ and $-(1-f)\beta_1^{(2)}$, as a function of wavelength. The scaled delays align closely, indicating good optimization over the entire C-band.

TABLE II
FIBER DESIGN PARAMETERS

	α_1	α_2	Δn_1	Δn_2	Δn_3	$a_1(\mu\text{m})$	$a_2(\mu\text{m})$	$a_3(\mu\text{m})$	$a_4(\mu\text{m})$	Notes
Conventional Compensation										
Fiber 1	1.962	1.962	$8.92 \cdot 10^{-3}$	$1.21 \cdot 10^{-3}$	$3.15 \cdot 10^{-3}$	5.7	11.38	12.52	19.41	$f = 1/3$
Fiber 2	1.989	1.989	$7.45 \cdot 10^{-3}$	$1.43 \cdot 10^{-3}$	$3.11 \cdot 10^{-3}$	5.89	11.77	13.4	19.54	$(1-f) = 2/3$
Self-Compensation										
Fiber 1	1.925	1.994	$7.43 \cdot 10^{-3}$	$1.46 \cdot 10^{-3}$	$3.03 \cdot 10^{-3}$	1.66	11.77	14.6	19.78	

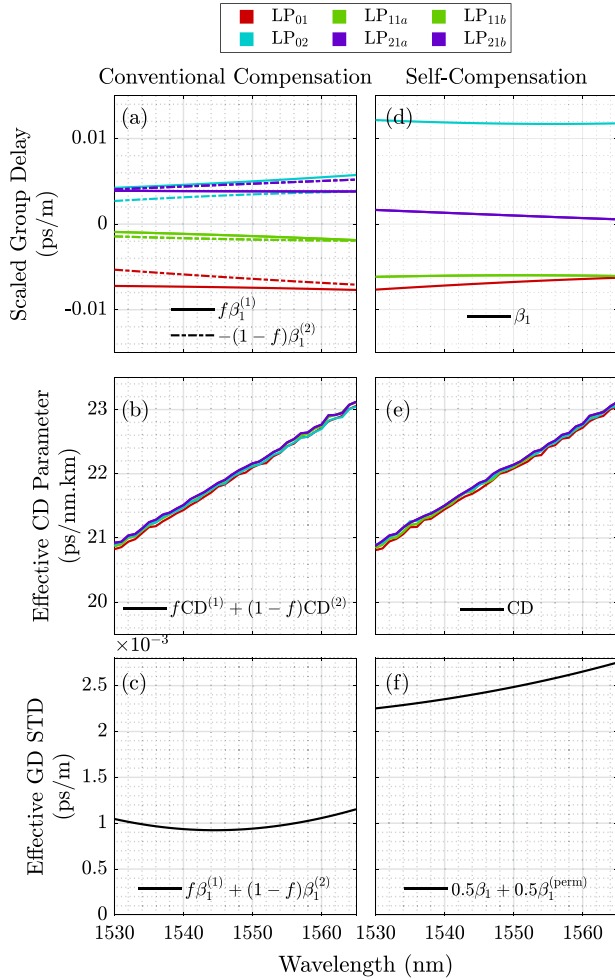


Fig. 5. Design example for compensation. (a - c) Conventional compensation; (d - f) Self-compensation. (a, d) Scaled uncoupled GDs as a function of wavelength over the C-band of the spatial modes of the optimized fiber(s); (b, e) Effective uncoupled CD parameter as a function of wavelength over the C-band; (c, f) Effective uncoupled GD STD as a function of wavelength over the C-band.

- Fig. 5(b) shows the effective CD parameters as a function of wavelength. The effective CD parameter $f\text{CD}^{(1)} + (1-f)\text{CD}^{(2)}$ is the same for each mode. Consequently, the effective uncoupled GD STD exhibits a flat wavelength trend in the C-band, as seen in Fig. 5(c).

Self-Compensation:

- Fig. 5(d) shows the uncoupled GDs of the spatial modes of the optimized fiber, β_1 as a function of wavelength. As desired, the delays of the first two mode groups with

spatial modes $\{\text{LP}_{01}\}, \{\text{LP}_{11a}, \text{LP}_{11b}\}$ are almost equal. Their average is equal to the negative of the average of the delays of the third mode group with spatial modes $\{\text{LP}_{02}, \text{LP}_{21a}, \text{LP}_{21b}\}$. We also observe that the mode-group-averaged GDs are not too close to zero, which would negatively impact fabrication error tolerances.

- Fig. 5(e) shows the uncoupled CD parameters of the spatial modes **CD** as a function of wavelength. The CD is the same for each mode.
- An effective uncoupled GD STD can be defined as $\text{std}(0.5\beta_1 + 0.5\beta_1^{(\text{perm})})$, where $\beta_1^{(\text{perm})}$ represents the uncoupled GD vector after mode permutations. In this case, the mode permutations exchange power between LP_{01} and LP_{02} and between LP_{11} and LP_{21} . As a result of low MDCD and effective optimization, the effective GD STD is small and exhibits a flat wavelength trend in the C-band, as seen in Fig. 5(f).

Fig. 5 is intended to characterize the uncoupled dispersion properties of the fibers, and does not take into account random intra- and inter-group coupling when calculating dispersion performance. The statistics of the coupled GDs are highly relevant to the complexity of DSP at the receiver, and are studied via multi-section simulations in the following subsection.

2) **Multi-Section Simulations:** We perform numerical multi-section simulations to calculate the end-to-end GD STD for the optimized fibers. We select a span length $L_S = 50$ km. We assume strong random intra-group coupling and weak random inter-group coupling in the fiber segments [21].

Random inter-group coupling is influenced by various factors, including effective refractive index differences between mode groups and manufacturing processes involved. The literature reports a wide range of coupling strengths. For instance, one study reports coupling strengths of less than -20 dB over 10.5 km of a GD compensation system [17] where the fibers support six spatial modes and have index contrast similar to the designs presented in this paper. The reported coupling strength corresponds to a characteristic coupling length exceeding 700 km. In our simulations, we set the value of the inter-group coupling length to $L_{\text{inter}} = 500$ km, which is equal to ten span lengths. Additionally, we compare our results to a scenario with no random inter-group coupling, represented by $L_{\text{inter}} = \infty$. For the conventional compensation system, we set the lengths of the fiber segments according to the optimized length fractions: $L_1 = fL_S/k_p$, $L_2 = (1-f)L_S/k_p$. For the self-compensation system, we set the lengths of the fiber segments to be equal and assume ideal mode permutations.

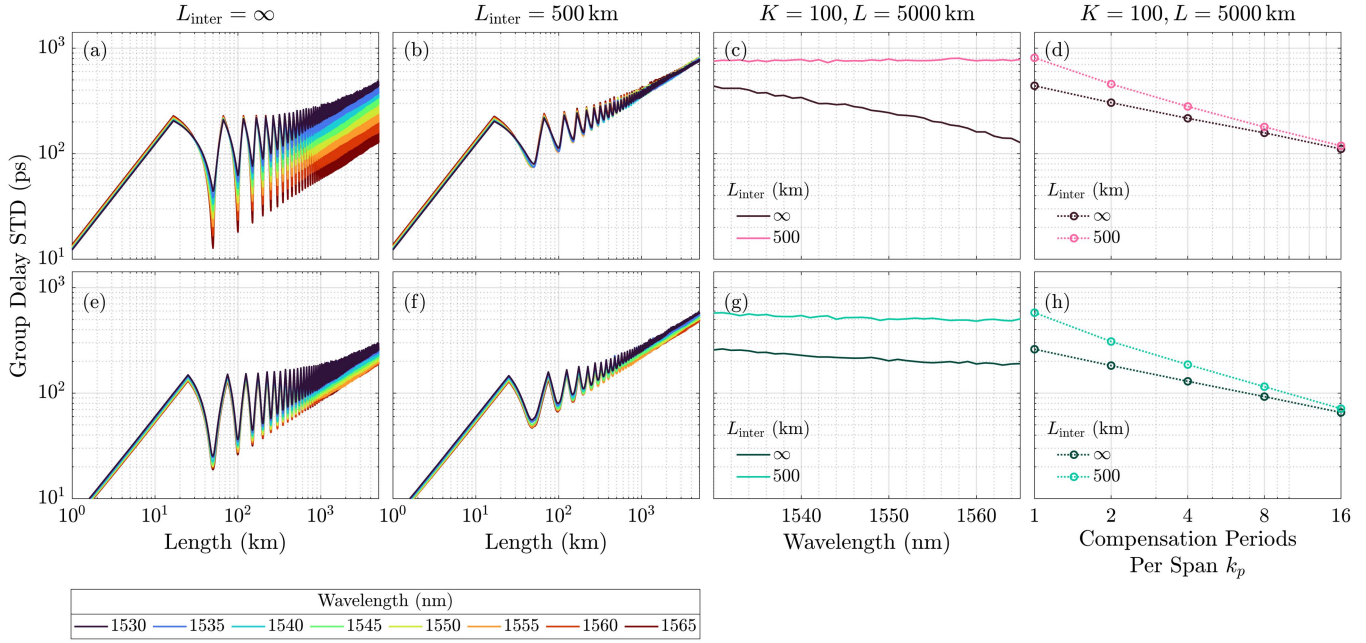


Fig. 6. Simulation of compensation. (a) - (d) Conventional compensation; (e) - (h) Self-compensation. (a), (e) GD STD as a function of length for various wavelengths in the C-band when $L_{\text{inter}} = \infty$ and $k_p = 1$; (b), (f) GD STD as a function of length for various wavelengths in the C-band when $L_{\text{inter}} = 500$ km and $k_p = 1$; (c), (g) GD STD as a function of wavelength after propagation over 5000 km for $L_{\text{inter}} = \infty, 500$ km and $k_p = 1$; (d), (h) GD STD as a function of number of compensation periods per span k_p after propagation over 5000 km for $L_{\text{inter}} = \infty, 500$ km.

Fig. 6 shows the GD STD as a function of the length of the link over $K = 100$ spans for several wavelength channels in the C-band. For the conventional compensation system, for $k_p = 1$, the worst-case GD STD after $K = 100$ spans is 438 ps, effectively $6.2 \text{ ps}/\sqrt{\text{km}}$, when $L_{\text{inter}} = \infty$, as seen in Fig. 6(a), and 803 ps, effectively $11.4 \text{ ps}/\sqrt{\text{km}}$, when $L_{\text{inter}} = 500$ km, as seen in Fig. 6(b). For the self-compensation system, for $k_p = 1$, the worst-case GD STD after $K = 100$ spans is 257 ps, effectively $3.6 \text{ ps}/\sqrt{\text{km}}$, when $L_{\text{inter}} = \infty$, as seen in Fig. 6(e), and 575 ps, effectively $8.1 \text{ ps}/\sqrt{\text{km}}$, when $L_{\text{inter}} = 500$ km, as seen in Fig. 6(f).

Fig. 6(d) and (h) illustrate the impact of increasing the number of compensation periods per span k_p . As discussed in Section II-C, GD STD decreases with the square root of k_p , provided there is no inter-group coupling. This reduction in GD STD is more pronounced in the presence of inter-group coupling. For the conventional compensation system, as k_p increases from 1 to 16, the worst-case GD STD reduces from 438 ps to 109 ps ($4.9 \text{ ps}/\sqrt{\text{km}}$) when $L_{\text{inter}} = \infty$, and from 803 ps to 115 ps ($5.1 \text{ ps}/\sqrt{\text{km}}$) when $L_{\text{inter}} = 500$ km. For the self-compensation system, the worst-case GD STD reduces from 257 ps to 65 ps ($2.9 \text{ ps}/\sqrt{\text{km}}$) when $L_{\text{inter}} = \infty$, and from 575 ps to 70 ps ($3.1 \text{ ps}/\sqrt{\text{km}}$) when $L_{\text{inter}} = 500$ km. When k_p is very large, the effect of inter-group coupling becomes negligible. However, this choice has its drawbacks. The number of splices, mode scramblers, and permuters required is proportional to k_p , resulting in higher losses. Consequently, the mode-averaged loss will increase linearly, and the MDL STD will increase with the square root of k_p . Overall, the results indicate that effective group delay compensation is achievable over the C-band with

thoughtful fiber and system design; however, there are trade-offs to consider.

The effective GD STD of the design examples in this paper is comparable to those reported for coupled MCFs supporting a similar number of modes [28], [29], [30]. Optimized three-core coupled MCFs have achieved GD STD as low as $4.2 \text{ ps}/\sqrt{\text{km}}$ [29]. Additionally, four-core coupled MCFs have reported a GD STD of $3.14 \text{ ps}/\sqrt{\text{km}}$ with an optimized fiber bend radius of 31 cm, and $10.6 \text{ ps}/\sqrt{\text{km}}$ in loosely bent cables [30]. One important issue in such coupled MCF systems is their sensitivity to the bend radius. Tight or excessively loose bends in the deployed fibers can lead to an increase in GD STD [30]. Additionally, fabrication errors can impact the core radius and refractive index, which also affects the GD STD [28]. Fabrication errors are also an issue in GD-compensating MMF systems, as discussed in Section IV-B.

IV. DISCUSSION

In this section, we discuss additional aspects of GD compensation.

A. Realization of Mode Permuters

GD self-compensation requires devices capable of performing precise mode permutations. One way to achieve these mode permutations is by utilizing a fan-out and fan-in configuration [11], [31]. In this configuration, a mode demultiplexer directs the signals from the spatial modes of the multimode fiber (MMF) into individual single-mode fibers (SMFs). Subsequently, a mode multiplexer combines the signals back into the MMF

after executing the desired permutation. However, this approach presents several challenges, including high mode-averaged and mode-dependent losses, as well as undesirable modal crosstalk.

An alternative mode permuter design uses a cascade of several long-period fiber grating (LPFG)-based mode couplers [32]. The design of each LPFG, including its grating period and symmetry, is intended to execute a subset of the necessary mode permutations. Undesired couplings by the LPFGs are suppressed by propagation constant engineering, in which the transverse refractive index profile is optimized to render the propagation constant spacings for the undesired couplings significantly different from those for the desired couplings. The LPFG-based permuter design promises to achieve much lower mode-averaged and mode-dependent losses and modal crosstalk than designs based on fan-out and fan-in.

B. Sensitivity to Fiber Fabrication Errors

Changes to the designed refractive index profile of fibers can impact the modal propagation constants, which in turn affect the modal GDs. Variations in the uncoupled modal GDs of the compensating fibers can result in an increased effective GD STD. In extreme cases, these changes can alter the order of the modal GDs, potentially rendering a compensation scheme completely ineffective. We study the sensitivity of GD compensation to fiber fabrication errors by introducing Gaussian random noise to the refractive index profile of the fiber, similar to previous studies [9], [33]. Research shows that different pieces of fiber drawn from the same preform can exhibit different refractive index errors [15]. Therefore, in our analysis, we assume that each fiber segment has an independent realization of refractive index error.

GD STD Penalty: In the presence of refractive index errors, the expression for the effective GD STD can be modified to include a penalty term:

$$\sigma_{\text{GD}}(K) = \sqrt{\sigma_{\text{GD,ideal}}^2(K) + \sigma_{\text{GD,err}}^2(K)}, \quad (14)$$

where $\sigma_{\text{GD,ideal}}$ is given by the expression in (2) and $\sigma_{\text{GD,err}}$ is the contribution of the refractive index errors. Using first-order perturbation theory, we can estimate the deviation in the GDs resulting from refractive index profile errors. Appendix A provides the derivation of this relationship. The penalty due to profile errors is given by:

$$\sigma_{\text{GD,err}}^2(K) = \left(K \frac{\nu_1 L_1 + \nu_2 L_2}{D} \right) \times \text{tr} \left(\left(\mathbf{D} - \frac{\mathbf{d}\mathbf{d}^T}{D} \right) \Sigma_n \right) \sigma_{\Delta n}^2, \quad (15)$$

where $\sigma_{\Delta n}$ is the STD of the refractive index errors, and $\Sigma_n \sigma_{\Delta n}^2$ is the $N_g \times N_g$ covariance matrix of mode-group-averaged GD deviations. For the design examples in this paper,

$$\Sigma_n = 10^{-18} \begin{bmatrix} 0.40 & 0.29 & 0.11 \\ 0.29 & 0.47 & 0.33 \\ 0.11 & 0.33 & 0.50 \end{bmatrix} (\text{s}^2/\text{m}). \quad (16)$$

We observe that Σ_n shows minimal variation with wavelength across the C-band.

Multi-Section Simulations: We perform multisection simulations for the design examples presented in Section III for $k_p = 1$. Fig. 7 plots the GD STD for conventional and self-compensation for inter-group coupling lengths, $L_{\text{inter}} = \infty$ and $L_{\text{inter}} = 500$ km for the following cases:

- No refractive index error
- With refractive index error of STD $\sigma_{\Delta n} = 10^{-5}$.
- With refractive index error of STD $\sigma_{\Delta n} = 10^{-5}$ and fiber sorting.

Fiber sorting involves characterizing the fabricated fibers. For each selected fiber from the set of fabricated fibers, another one that minimizes the effective GD STD of the combination is chosen as its pair. This process is repeated without replacement until all fibers are paired. The plots also include the estimates from the analytical expression in (14). The expectation in the GD variance is computed over the realizations of the random coupling matrices and refractive index errors.

From Fig. 7, we observe that the analytical expression accurately captures the GD STD penalty caused by refractive index errors. For the conventional compensation system, the worst-case GD STD after $K = 100$ spans is 1.33 ns (18.8 ps/ $\sqrt{\text{km}}$), when $L_{\text{inter}} = \infty$, as shown in Fig. 7(a), and increases to 1.50 ns (21.2 ps/ $\sqrt{\text{km}}$) when $L_{\text{inter}} = 500$ km, as seen in Fig. 7(b). Fiber sorting provides only marginal improvement, resulting in a GD STD of 1.09 ns (15.4 ps/ $\sqrt{\text{km}}$) for $L_{\text{inter}} = \infty$ and 1.28 ns (18.1 ps/ $\sqrt{\text{km}}$) for $L_{\text{inter}} = 500$ km.

For the self-compensation system, the worst-case GD STD after $K = 100$ spans is 1.20 ns (17 ps/ $\sqrt{\text{km}}$) when $L_{\text{inter}} = \infty$, as shown in Fig. 7(c), and 1.28 ns (18.1 ps/ $\sqrt{\text{km}}$) when $L_{\text{inter}} = 500$ km, as seen in Fig. 7(d). In this case, fiber sorting proves to be more effective, reducing the GD STD to 0.76 ns (10.7 ps/ $\sqrt{\text{km}}$) for $L_{\text{inter}} = \infty$ and 0.94 ns (13.3 ps/ $\sqrt{\text{km}}$) for $L_{\text{inter}} = 500$ km.

After just one span, the effective GD STD is significantly higher when index errors are present compared to when they are not. Additionally, when comparing scenarios with and without inter-group coupling, the effective GD STD remains nearly unchanged in the presence of index errors. These findings indicate that the GD STD is primarily influenced by the contribution from the error penalty term. It is crucial for fabrication processes to maintain tight tolerances in order for GD compensation in MMF to be effective. Experimental investigations are needed to improve the refractive index error models and verify the numerical estimates presented in this paper, making this an important area of future research.

C. Digital Signal Processing

GD STD is a measure of the memory of the MMF channel, which in turn affects the DSP complexity of the MIMO equalizer. Additionally, since the MMF channel is dynamic, the equalization process must be adaptive [34]. Submarine SMF channels have been observed to change on a time scale of seconds [35]. Fibers supporting more modes are expected to be more sensitive to environmental perturbations [36], [37]. A larger GD STD makes it more challenging to learn and adapt to a time-varying

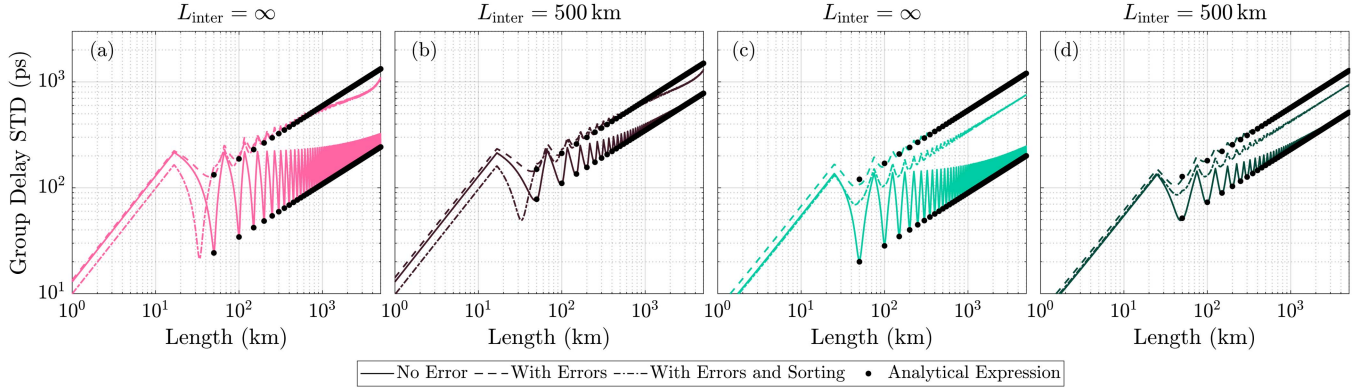


Fig. 7. Compensation in the presence of refractive index errors for $k_p = 1$. (a) - (b) Conventional compensation; (c) - (d) Self-compensation. (a), (c) GD STD as a function of length at 1550nm for $L_{\text{inter}} = \infty$; (b), (d) GD STD as a function of length at 1550nm for $L_{\text{inter}} = 500$ km; The solid lines correspond to GD performance with no refractive index errors, $\sigma_{\Delta n} = 0$; the dashed lines correspond to the GD performance with refractive index errors of STD $\sigma_{\Delta n} = 10^{-5}$; the dotted-dashed lines correspond to the GD performance with fiber sorting and refractive index errors of STD $\sigma_{\Delta n} = 10^{-5}$; and the solid circles correspond to analytical estimates.

channel in the presence of noise. The number of MIMO equalizer taps required at the receiver is given by [38]

$$N_{\text{MD}} = \lceil \sigma_{\text{GD}} u_D(p) r_{\text{os}} R_s \rceil, \quad (17)$$

where $u_D(p) \approx 4.6$ for $D = 12$ with probability $1 - p = 1 - 10^{-6}$, r_{os} is the oversampling ratio and R_s is the symbol rate. For example, for $\sigma_{\text{GD}} = 1$ ns, $r_{\text{os}} = 17/16$, and $R_s = 50$ GBaud, we need $N_{\text{MD}} = 245$ taps. An adaptive 12×12 MIMO frequency-domain equalization using a cyclic prefix (CP) of length N_{MD} taps and block length of $N_{\text{FFT}} = 2^{12}$ achieves a CP efficiency $N_{\text{FFT}}/(N_{\text{FFT}} + N_{\text{MD}}) = 94\%$.

Adaptive algorithms such as recursive least-squares (RLS) can train the equalizers using $n_{\text{tr}} = 50$ training blocks corresponding to an adaptation time of $t_{\text{adapt}} = (N_{\text{FFT}} + N_{\text{MD}})/r_{\text{os}} R_s = 4.09 \mu\text{s}$. The data-aided training stage can be followed by decision-directed tracking. For a signal-to-noise ratio of 20dB at the end of 100 spans and a MDL STD of 0.3 dB per span, the RLS algorithm is found to track a dynamic channel with a generalized stokes rotation rate of more than 314 krad/s for QPSK signaling and up to 100 krad/s for 16-QAM signaling. To determine the Stokes rotation rates and select an appropriate modulation and coding scheme, it is essential to experimentally measure the dynamics of the fabricated MMFs. If the observed dynamics are faster than the estimates made here, then the FFT block length N_{FFT} can be reduced and the channel estimates can be computed more frequently, although with a reduced CP efficiency. Alternatively, pilot symbols can be periodically transmitted to enable intermittent data-aided tracking. This method also leads to a reduced information rate.

D. Extension to More Than Three Mode Groups

The analytical expressions presented in Section II are applicable to all GD compensation systems characterized by strong random intra-group coupling and weak inter-group coupling in the fiber segments.

Fibers that support more than three mode groups can be designed using the parameterized refractive index profile (11). However, additional features may be necessary to control the

GDs and CDs of the mode groups. This is essential to creating a feature space sufficiently rich to enable numerical design optimization.

For the self-compensation scheme, the optimal mode permutation and the associated mode-group-averaged GD vector change as the number of mode groups increases. Combinatorially, there exist several possible ideal permutation schemes for any number of mode groups N_g . It is important to emphasize that realizing some of these ideal schemes may not be trivial, as each scheme requires a MMF with mode groups having specific GDs, as well as a mode permuter realizing a specific permutation pattern. We discuss two examples here.

The first ideal permutation scheme consists of partitioning the mode groups into three groups, each with a different mode-group-averaged GD from the set $\{-\tau, 0, \tau\}$. The number of mode groups with a GD of 0 is $m = N_g \pmod{4}$. For odd m , one of these mode groups has $(N_g + 1)/2$ modes. For $m = 2, 3$, there exists one pair of mode groups with a GD of 0 comprising mode groups MG_i and MG_j , where i, j are the mode group indices and $i + j = N_g + 1$.

The remaining number of mode groups is divisible by 4, so we can form an even number of pairs of mode groups such that for every pair comprising MG_i , MG_j , $i + j = N_g + 1$. If the transmission fiber is designed so that half of these pairs have a GD of τ and the other half have a GD of $-\tau$, then permuting these two groups at the midpoint of each span effects perfect self-compensation (more generally, if $k_p > 1$, we permute the two groups at the midpoint of each compensation period). For example, when $N_g = 5$, the ideal GD vector for this scheme is $[\tau, -\tau, 0, -\tau, \tau]$. Midway through each span (or compensation period), MG_1 and MG_5 are permuted with MG_2 and MG_4 , and MG_3 remains unchanged. For larger values of N_g , this scheme can be extended to allow different GD values for different pairs of mode groups, as long as for each pair of mode groups with a GD of τ , there exists a corresponding mode group pair with a GD of $-\tau$.

An alternative ideal permutation scheme involves partitioning the mode groups into four groups, each with a different mode-group-averaged GD from the set $\{0, \tau_1, \tau_2, \tau_3\}$. The number of

mode groups with a GD of 0 is $m = N_g \pmod{6}$. For odd m , one of these mode groups has $(N_g + 1)/2$ modes. For $m \geq 2$, the other pairs of mode groups with a GD of 0 are comprised of MG_i, MG_j , where $i + j = N_g + 1$.

Since the remaining number of mode groups with nonzero GD is divisible by 6, we can form M pairs of mode groups such that for every pair comprised of $MG_i, MG_j, i + j = N_g + 1$, where M is divisible by 3. If the transmission fiber is then designed so that a third of these pairs have a GD of τ_1 , a third have a GD of τ_2 , and a third have a GD of τ_3 , cyclically permuting these three groups twice per span effects perfect self-compensation (more generally, if $k_p > 1$, we permute these three groups twice per compensation period). For example, in the case where $N_g = 6$, the ideal GD vector for this scheme is $[\tau_1, \tau_2, \tau_3, \tau_3, \tau_2, \tau_1]$. We subdivide each span into three equal-length segments (more generally, when $k_p > 1$, we subdivide each compensation period into three segments), and at the two boundaries between the three segments, we apply an ideal permutation that cycles power from MG_1 and MG_6 to MG_2 and MG_5 , from MG_2 and MG_5 to MG_3 and MG_4 , and from MG_3 and MG_4 to MG_1 and MG_6 . Power can also be cycled amongst these three pairs of mode groups in the reverse direction. For larger values of N_g , multiple such 3-cycles can exist, each with potentially different values of τ_1, τ_2, τ_3 .

Although these ideal permutation schemes have been shown to exist for any N_g , their implementation becomes increasingly difficult as N_g increases. Designing a mode permuter for large N_g requires careful consideration, especially when opting for cascaded LPFG-based devices. The sheer number of necessary power transfers amongst all modes greatly increases the number of LPFGs the device requires. In the schemes described here, potentially one mode would need to be transferred to a mode group four levels away, and to prevent coupling into unguided modes, this transfer should take place over several steps, each requiring an LPFG [32]. There is likely a limit to how many different LPFGs can be inscribed in the same fiber without introducing significant undesired coupling, so multiple mode permutation fibers may be necessary [32]. These are topics for future research.

V. CONCLUSION

In this paper, we examined GD compensation for MMF links using two approaches: conventional compensation, which alternates fiber types with opposite GD orderings, and self-compensation, which utilizes a single fiber type with periodically inserted mode permuters. We provided analytical expressions for the GD STD in these systems, taking into account factors such as random inter-group coupling, mode scrambling, and refractive index errors. We proposed fiber designs based on graded-index MMFs with specialized index profiles to control the MDCD in order to achieve effective compensation over the entire C-band. We also addressed design strategies for both systems and fibers to address the challenges of random inter-group coupling and refractive index errors. These strategies include increasing the number of compensation periods per

span, using a regularized cost function in fiber design, and fiber sorting. The design examples we presented achieved an effective GD STD of $18.1 \text{ ps}/\sqrt{\text{km}}$ with conventional compensation and $13.3 \text{ ps}/\sqrt{\text{km}}$ with self-compensation for $L_{\text{inter}} = 500 \text{ km}$ and $\sigma_{\Delta n} = 10^{-5}$. We discussed the implications of GD compensation for DSP complexity and performance. Finally, we discussed methods for extending GD compensation to fibers that support more than three mode groups.

APPENDIX A

IMPACT OF REFRACTIVE INDEX ERRORS ON MODAL GROUP DELAYS

In this appendix, we derive a semi-analytic expression for the penalty in GD STD caused by errors in the fiber refractive index profile. Consider a fiber with a refractive index profile $n(r)$ and refractive index error $\Delta n(r)$ as a function of radial position r . From a first-order perturbation theory analysis, the change in the propagation constant of the i th mode is

$$\Delta\beta[i] = \frac{1}{2\beta[i]} \frac{\iint 2k_0^2 n \Delta n |\psi_i|^2 r dr d\theta}{\iint |\psi_i|^2 r dr d\theta}, \quad (18)$$

where ψ_i is the electric field amplitude of the i th mode. The error in its GD is

$$\Delta\beta_1[i] = \frac{d}{d\omega} \Delta\beta[i]. \quad (19)$$

Numerically, the integration can be replaced by a summation and the derivative can be replaced by a finite difference as follows:

$$\Delta\beta_1[i] = \frac{1}{2} \sum_j \frac{1}{\Delta\omega} \left(\frac{a_j^+[i]}{\beta^+[i]} - \frac{a_j^-[i]}{\beta^-[i]} \right) \Delta n[j], \quad (20)$$

where $j = 1, \dots, N_r$ is a radial position index, $\Delta\omega$ is a frequency difference, superscripts $+$ and $-$ indicate the values at $+\Delta\omega/2$ and $-\Delta\omega/2$, respectively, and

$$a_j[i] = 2k_0^2 n[j] r[j] \frac{|\psi_i[j]|^2}{\sum_l |\psi_i[l]|^2 r[l]}. \quad (21)$$

From the above expressions, the errors in the mode-group-averaged GDs can be obtained:

$$\Delta\bar{\beta}_1 = \bar{\mathbf{A}} \Delta \mathbf{n}, \quad (22)$$

where $\Delta \mathbf{n}$ is the N_r -dimensional error vector and $\bar{\mathbf{A}}$ is an $N_g \times N_r$ matrix describing the mode-group-averaged GD error coefficients and is given by

$$\bar{\mathbf{A}}[l, j] = \frac{1}{2d_l \Delta\omega} \sum_{i \in M_l} \left(\frac{a_j^+[i]}{\beta^+[i]} - \frac{a_j^-[i]}{\beta^-[i]} \right).$$

Assuming $\{\Delta n[j]\}$ are independent and identically distributed Gaussian random variables with STD $\sigma_{\Delta n}$ results in $\Delta\bar{\beta}_1$ being a zero-mean Gaussian random vector whose second-order statistics are related to those of $\Delta \mathbf{n}$ by

$$\Sigma = \mathbb{E} \left\{ \Delta\bar{\beta}_1^T \Delta\bar{\beta}_1 \right\} = \sigma_{\Delta n}^2 \bar{\mathbf{A}} \bar{\mathbf{A}}^H. \quad (23)$$

ACKNOWLEDGMENT

Most of the computing for this project was performed on the Sherlock cluster at Stanford University. The authors would like to thank the Stanford Research Computing Center for providing this cluster and technical support. The authors are grateful for helpful discussions with Dr. Ethan Liang.

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/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.)*, 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," *J. Lightw. Technol.*, vol. 41, no. 15, pp. 4989–5002, Aug. 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/10064021/>
- [6] 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," presented at Opt. Fiber Commun. Conf., Los Angeles, CA, USA, 2015, Paper W2A.34. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=OFC-2015-W2A.34>
- [7] P. Sillard et al., "Low-differential-mode-group-delay 9-LP-mode fiber," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 425–430, Jan. 2016. [Online]. Available: <https://ieeexplore.ieee.org/document/7174947>
- [8] 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/>
- [9] O. Krutko, R. Refaee, A. Vijay, and J. M. Kahn, "Ultra-low-Loss fiber bragg grating mode scrambler design exploiting propagation constant engineering," *J. Lightw. Technol.*, vol. 43, no. 6, pp. 2883–2896, Mar. 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10772334/>
- [10] A. Vijay, O. Krutko, R. Refaee, and J. M. Kahn, "Modal statistics in mode-division-multiplexed systems using mode scramblers," *J. Lightw. Technol.*, vol. 43, no. 2, pp. 845–856, Jan. 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10685071/?arnumber=10685071>
- [11] K. Shibahara, T. Mizuno, H. Ono, K. Nakajima, and Y. Miyamoto, "Long-haul DMD-Unmanaged 6-Mode-Multiplexed transmission employing cyclic mode-group permutation," presented at Opt. Fiber Commun. Conf., San Diego, CA, USA: 2020, Paper Th3H.3. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=OFC-2020-Th3H.3>
- [12] G. D. Sciuillo et al., "Enhancing long-haul 15-Mode fiber performance: Mode permutation for reduced modal dispersion," *J. Lightw. Technol.*, vol. 43, no. 2, pp. 481–491, Jan. 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10663941/?arnumber=10663941>
- [13] T. Gao et al., "Novel mirror-flipped mode permutation technique for long-haul 6-mode transmission," *Opt. Commun.*, vol. 570, Nov. 2024, Art. no. 130892. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0030401824006291>
- [14] R. Maruyama, N. Kuwaki, S. Matsuo, and M. Ohashi, "Two mode optical fibers with low and flattened differential modal delay suitable for WDM-MIMO combined system," *Opt. Exp.*, vol. 22, no. 12, pp. 14311–14321, Jun. 2014. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?uri=oe-22-12-14311>
- [15] L. Gruner-Nielsen et al., "Few mode transmission fiber with low DGD, low mode coupling, and low loss," *J. Lightw. Technol.*, vol. 30, no. 23, pp. 3693–3698, Dec. 2012. [Online]. Available: <https://ieeexplore.ieee.org/document/6353104/?arnumber=6353104>
- [16] F. Ferreira, D. Fonseca, and H. Silva, "Design of few-mode fibers with arbitrary and flattened differential mode delay," *IEEE Photon. Technol. Lett.*, vol. 25, no. 5, pp. 438–441, Mar. 2013. [Online]. Available: <https://ieeexplore.ieee.org/document/6417958/?arnumber=6417958>
- [17] T. Mori, T. Sakamoto, M. Wada, T. Yamamoto, and F. Yamamoto, "Few-mode fibers supporting more than two LP modes for mode-division-multiplexed transmission with MIMO DSP," *J. Lightw. Technol.*, vol. 32, no. 14, pp. 2468–2479, Jul. 2014. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?uri=jlt-32-14-2468>
- [18] F. Gerome, J.-L. Auguste, J. Maury, J.-M. Blondy, and J. Marcou, "Theoretical and experimental analysis of a chromatic dispersion compensating module using a dual concentric core fiber," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 442–448, Jan. 2006. [Online]. Available: <https://ieeexplore.ieee.org/document/1589076/?arnumber=1589076>
- [19] J. Riishede and O. Sigmund, "Inverse design of dispersion compensating optical fiber using topology optimization," *JOSA B*, vol. 25, no. 1, pp. 88–97, Jan. 2008. [Online]. Available: <https://opg.optica.org/josab/abstract.cfm?uri=josab-25-1-88>
- [20] A. Vijay, O. Krutko, R. Refaee, and J. M. Kahn, "Modal statistics in mode-division-multiplexed systems using mode scramblers," *J. Lightw. Technol.*, vol. 43, no. 2, pp. 845–856, Jan. 2025. [Online]. Available: <http://arxiv.org/abs/2409.06908>
- [21] K.-P. Ho and J. M. Kahn, "Linear propagation effects in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 614–628, Feb. 2014. [Online]. Available: <https://ieeexplore.ieee.org/document/6615972/>
- [22] 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>
- [23] *Transmission Media and Optical Systems Characteristics—Characteristics of a Cut-off Shifted Single-mode Optical Fibre and Cable*, ITU-T G.654, International Telecommunication Union, Geneva, Switzerland, 1997. [Online]. Available: <https://www.itu.int/rec/T-REC-G.654-202408-1/en>
- [24] T. Mori, T. Sakamoto, M. Wada, T. Yamamoto, and F. Yamamoto, "Low DMD four LP mode transmission fiber for wide-band WDM-MIMO system," presented at Opt. Fiber Commun. Conf./Nat. Fiber Optic Engineers Conf., Mar. 2013, Paper OTH3K.1. [Online]. Available: <https://opg.optica.org/abstract.cfm?uri=OFC-2013-OTH3K.1>
- [25] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw.*, Nov. 1995, pp. 1942–1948. [Online]. Available: <https://ieeexplore.ieee.org/document/488968>
- [26] P. Sillard, M. Bigot, F. Achten, M. Bsaibes, Y. Quiquempois, and L. Bigot, "Germanium-free graded-index-core few-mode fiber," in *Proc. IEEE 50th Eur. Conf. Opt. Commun.*, Sep. 2024, pp. 294–297. [Online]. Available: <https://ieeexplore.ieee.org/document/10926350>
- [27] P. Sillard et al., "Germanium-free graded-index-core few-mode fibers," *J. Lightw. Technol.*, vol. 43, no. 13, pp. 6278–6284, Jul. 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10859173/>
- [28] K. Saitoh, "Multi-core fiber technology for SDM: Coupling mechanisms and design," *J. Lightw. Technol.*, vol. 40, no. 5, pp. 1527–1543, Mar. 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9691859/>
- [29] T. Hayashi, T. Nagashima, T. Muramoto, F. Sato, and T. Nakanishi, "Spatial mode dispersion suppressed randomly-coupled multi-core fiber in straightened loose-tube cable," in *Proc. 2019 Opt. Fiber Commun. Conf. Exhib.*, Mar. 2019, pp. 1–3. [Online]. Available: <https://ieeexplore.ieee.org/document/8696706>
- [30] H. Sakuma et al., "Microbending behavior of randomly-coupled ultra-low-loss multi-core fiber," in *Proc. IEEE 45th Eur. Conf. Opt. Commun.*, Sep. 2019, pp. 1–3. [Online]. Available: <https://ieeexplore.ieee.org/document/9125478/>
- [31] G. Di Sciuillo et al., "Reduction of modal dispersion in a long-haul 15-mode fiber link by means of mode permutation," in *Proc. 49th Eur. Conf. Opt. Commun.*, Oct. 2023, pp. 250–253. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10484924>
- [32] O. Krutko, R. Refaee, A. Vijay, N. Zahedi, and J. M. Kahn, "Low-loss all-fiber mode permuter design exploiting propagation constant engineering and cascaded Bragg gratings," May 2025. [Online]. Available: <http://arxiv.org/abs/2505.06775>
- [33] V. M. Gololobov, P. S. Anisimov, V. V. Zemlyakov, and J. Gao, "Multimode optical fibers with harmonically-Tempered refractive index profile functions," *J. Lightw. Technol.*, vol. 41, no. 14, pp. 4840–4846, Jul. 2023. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10048542>

- [34] S. O. Arik, K.-P. Ho, and J. M. Kahn, "Group delay management and multiinput multioutput signal processing in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 34, no. 11, pp. 2867–2880, Jun. 2016. [Online]. Available: <https://ieeexplore.ieee.org/document/7409913/>
- [35] Z. Zhang, X. Bao, Q. Yu, and L. Chen, "Fast state of polarization and PMD drift in submarine fibers," *IEEE Photon. Technol. Lett.*, vol. 18, no. 9, pp. 1034–1036, May 2006. [Online]. Available: <https://ieeexplore.ieee.org/document/1621268/?arnumber=1621268>
- [36] K. Choutagunta, I. Roberts, and J. M. Kahn, "Efficient quantification and simulation of modal dynamics in multimode fiber links," *J. Lightw. Technol.*, vol. 37, no. 8, pp. 1813–1825, Apr. 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8588335/>
- [37] K. Choutagunta et al., "Modal dynamics in spatially multiplexed links," in *Proc. 2019 Opt. Fiber Commun. Conf. Exhib.*, Mar. 2019, pp. 1–3. [Online]. Available: <https://ieeexplore.ieee.org/document/8696593>
- [38] K.-P. Ho and J. M. Kahn, "Mode coupling and its impact on spatially multiplexed systems," in *Optical Fiber Telecommunications* 6th ed.. Boston, MA, USA: Academic Press, 2013, pp. 491–568, ch. 11. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123969606000110>

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, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2025. His research interests include optical communications, mode-division multiplexing, and data-center applications.

Nika Zahedi is currently working towards the B.S. and M.S. degrees in electrical engineering with Stanford University, Stanford, CA, USA. Her research interests include mode-division multiplexing, signal processing, and optimization techniques.

Oleksiy Krutko (Member, IEEE) received the B.S. degree in electrical engineering from the University of Texas at Austin, Austin, TX, USA, in 2020 and the Ph.D. degree from Stanford University, Stanford, CA, USA, in 2025. 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 current 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, Berkeley, CA, USA, in 1981, 1983 and 1986, respectively. From 1987 to 1990, he was with 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 with Berkeley. He demonstrated coherent detection of QPSK in 1992. In 1999, they have authored or coauthored the first work on probabilistic shaping for optical communications. In the 1990s and early 2000, he and collaborators performed seminal work on indoor and outdoor free-space optical communications and multi-input multi-output wireless communications. In 2000, they have 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, they applied to patent the first electronic compensation of fiber Kerr nonlinearity. StrataLight was acquired by Opnext in 2009. In 2003, he became a Professor of electrical engineering with the E. L. Ginzton Laboratory, Stanford University. 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, they independently invented simplified digital backpropagation for compensating fiber Kerr nonlinearity and dispersion. Since 2004, he and collaborators have studied 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 current research addresses optical frequency comb generators, coherent data center links, rate-adaptive access networks, fiber Kerr nonlinearity mitigation, ultra-long-haul submarine links, and optimal free-space transmission through atmospheric turbulence. He was the recipient of the National Science Foundation Presidential Young Investigator Award in 1991.