Theory of intermodal four-wave mixing with random linear mode coupling in few-mode fibers

Yuzhe Xiao,^{1,2} René-Jean Essiambre,² Marc Desgroseilliers,³ Antonia M. Tulino,² Roland Ryf,² Sami Mumtaz,⁴ and Govind P. Agrawal¹

¹The Institute of Optics, University of Rochester, Rochester, NY, 14627, USA
 ²Bell Laboratories, Alcatel-Lucent, Holmdel, NJ, 07733, USA
 ³School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
 ⁴ Formerly with (1) and (2), now with Altera Europe Limited, Holmers Farm Way, High Wycombe, HP12 4XF, UK

Abstract: We study intermodal four-wave mixing (FWM) in few-mode fibers in the presence of birefringence fluctuations and random linear mode coupling. Two different intermodal FWM processes are investigated by including all nonlinear contributions to the phase-matching condition and FWM bandwidth. We find that one of the FWM processes has a much larger bandwidth than the other. We include random linear mode coupling among fiber modes using three different models based on an analysis of the impact of random coupling on differences of propagation constants between modes. We find that random coupling always reduces the FWM efficiency relative to its vale in the absence of linear coupling. The reduction factor is relatively small (about 3 dB) when only a few modes are linearly coupled but can become very large (> 40 dB) when all modes couple strongly. In the limit of a coupling length much shorter than the nonlinear length, intermodal FWM efficiency becomes vanishingly small. These results should prove useful in the context of space-division multiplexing with few-mode and multimode fibers.

© 2014 Optical Society of America

OCIS codes: (060.2330) Fiber optics communications; (060.4370) Nonlinear optics, fibers; (190.4380) Nonlinear optics, four-wave mixing

References and links

- R. H. Stolen, "Phase-matched-stimulated four-photon mixing in silica-fiber waveguides," IEEE J. Quantum Electron. 11, 100–103 (1975).
- 2. G. P. Agrawal, Nonlinear Fiber Optics, 5th ed. (Academic Press, New York, 2012).
- R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle Jr., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 × 6 MIMO processing," J. Lightwave Technol. 30, 521–531 (2012).
- 4. C. Koebele, M. Salsi, G. Charlet, and S. Bigo, "Nonlinear effects in mode division multiplexed transmission over few-mode optical fiber," IEEE Photon. Technol. Lett. 23, 1316–1318 (2011).
- A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multi-mode fibers in the strong coupling regime," Opt. Express 20, 11673–11678 (2012).
- S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Reduction of nonlinear penalties due to linear coupling in multicore optical fibers," IEEE Photon. Technol. Lett. 24, 1574–1576 (2012).
- A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," Opt. Express 20, 23436–23441 (2012).

- R.-J. Essiambre, R. Ryf, M. A. Mestre, A. H. Gnauck, R. W. Tkach, A. R. Chraplyvy, Y. Sun, X. Jiang, and R. Lingle Jr, "Experimental investigation of inter-modal four-wave mixing in few-mode fibers," IEEE Photon. Technol. Lett. 25, 539–541 (2013).
- J. Demas, P. Steinvurzel, B. Tai, Y. Chen, and S. Ramachandran, "Two octaves of frequency generation by cascaded intermodal nonlinear mixing in solid optical fiber," Proc. Conf. on Lasers and Electro-optics (2013). Paper CTu2E.5.
- A. D. Ellis, N. Mac Suibhne, F. C. Gunning, and S. Sygletos, "Expressions for the nonlinear transmission performance of multi-mode optical fiber," Opt. Express 21, 22834–22846 (2013).
- S. Rashleigh, "Origins and control of polarization effects in single-mode fibers," J. Lightwave Technol. 1, 312– 331 (1983).
- 12. D. Marcuse, Theory of dielectric optical waveguides (Academic Press, 1991).
- J. M. Fini, B. Zhu, T. F. Taunay, M. F. Yan, and K. S. Abedin, "Crosstalk in multicore fibers with randomness: gradual drift vs. short-length variations," Opt. Express 20, 949–959 (2012).
- 14. S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: generalization of the Manakov equations," J. Lightwave Technol. **31**, 398–406 (2013).
- 15. R. W. Boyd, Nonlinear optics (Academic Press, 1992).
- Z. Su, X. Zhu, and W. Sibbett, "Conversion of femtosecond pulses from the 1.5-to the 1.3-μm region by selfphase-modulation-mediated four-wave mixing," J. Opt. Soc. Am. B 10, 1050–1053 (1993).
- 17. R. Stolen, M. Bösch, and C. Lin, "Phase matching in birefringent fibers," Opt. Lett. 6, 213–215 (1981).
- R. Stolen, J. Bjorkholm, and A. Ashkin, "Phase-matched three-wave mixing in silica fiber optical waveguides," Appl. Phys. Lett. 24, 308–310 (1974).
- 19. D. Gloge et al., "Weakly guiding fibers," Appl. Opt 10, 2252–2258 (1971).
- R.-J. Essiambre, R. W. Tkach, and R. Ryf, "Fiber nonlinearity and capacity: Single-mode and multimode fibers," in "Optical Fiber Telecommunications VI B,", I. Kaminow, T. Li, and A. E. Willner, eds. (Academic Press, 2013), chap. 1, pp. 1–37.
- F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," J. Opt. Soc. Am. B 25, 1645–1654 (2008).
- 22. J. M. Fini, B. Zhu, T. F. Taunay, M. F. Yan, and K. S. Abedin, "Statistical models of multicore fiber crosstalk including time delays," J. Lightwave Technol. **30**, 2003–2010 (2012).
- C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Random coupling between groups of degenerate fiber modes in mode multiplexed transmission," Opt. Express 21, 9484–9490 (2013).
- M. E. Marhic, Fiber optical parametric amplifiers, oscillators and related devices (Cambridge university press, 2008).
- M. Wuilpart, P. Mégret, M. Blondel, A. J. Rogers, and Y. Defosse, "Measurement of the spatial distribution of birefringence in optical fibers," IEEE Photon. Technol. Lett. 13, 836–838 (2001).
- A. Galtarossa and L. Palmieri, "Spatially resolved PMD measurements," J. Lightwave Technol. 22, 1103–1115 (2004).
- 27. R. Olshansky, "Mode coupling effects in graded-index optical fibers," Appl. Opt. 14, 935–945 (1975).
- L. Palmieri, "Coupling mechanism in multimode fibers," in Proc. SPIE 9009, Next-Generation Optical Communication: Components, Sub-Systems, and Systems III (2013). Paper 90090G.
- C. Poole and R. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," Electronics Letters 22, 1029–1030 (1986).
- 30. P. Harremoës, "Maximum entropy on compact groups," Entropy 11, 222–237 (2009).
- 31. E. Breuillard, "Random walks on lie groups," Survey, http://www.math.u-psud.fr/ breuilla/part0gb.pdf.
- G. W. Stewart and J. guang Sun, *Matrix Perturbation Theory*, Computer Science and Scientific computing (Academic Press, 1990).
- X. S. Chen, W. Li, and W. W. Xu, "Perturbation analysis of the eigenvector matrix and singular vector matrices," Taiwanese J. Math. 16, 179–194 (2012).
- T. Kato, Perturbation theory for linear operators, vol. 132 of Grundlehren der mathematischen Wissenschaften (Springer, 1995).

1. Introduction

Four-wave mixing (FWM) is an important nonlinear process in the context of silica fibers. Although it was studied in few-mode fibers (FMFs) as early as 1975 [1], until recently most studies on FWM focused on single-mode fibers (SMFs) exclusively [2]. From the recent advances in space-division multiplexing (SDM), few-mode, multimode and multicore fibers are seen as candidates for increasing the capacity of telecommunication systems [3]. In the last few years, considerable work has been performed, both experimental and theoretical, to study the impact of fiber nonlinearity on SDM systems [4–10]. In particular, Essiambre *et al.* have observed two

types of FWM processes [8] in an experimental investigation of non-degenerate intermodal FWM (IM-FWM) in FMFs many kilometers in length. Many mechanisms exist through which random coupling among fiber modes can occur [11, 12]. Long FMFs typically exhibit random linear coupling between modes [13]. To fully understand the IM-FWM processes in such fibers, one should consider a model that incorporates the impact of random linear mode coupling between different groups of modes on IM-FWM.

In this paper, we present a theoretical model and perform detailed numerical simulations to understand the impact of random linear mode coupling on the IM-FWM process. In Sec. 2, the two non-degenerate IM-FWM configurations are discussed and the corresponding phasematching conditions are introduced. Section 3 presents the nonlinear propagation equations used in the numerical simulations. More specifically, the effects of random linear mode coupling are incorporated through a random coupling matrix that includes both polarization rotation and linear coupling between different fiber spatial modes. The results obtained in the absence of linear mode coupling are presented in Sec. 4. Under proper phase-matching conditions, both non-degenerate IM-FWM processes exhibit high efficiency. In this uncoupled mode limit, a simple analytical expression is used to estimate the power of the generated idler wave. The bandwidth of the two IM-FWM processes is analyzed in Sec. 5, again in the absence of linear mode coupling. The impact of nonlinear phase shifts of the pumps and the probe on the phase-matching condition of IM-FWM processes is also considered. Numerical results are compared with theoretical predictions and it is found that a simple theoretical model provides a good estimate of the idler power for both IM-FWM processes. Section 6 focuses on the impact of random linear mode coupling and introduces three different coupling models ranging from weak coupling to very strong coupling among all fiber modes. These models are justified by an analysis presented in the appendix that considers the impact of differences of propagation constants between modes on linear random coupling in the limit of infinite fiber lengths. Changes in the FWM efficiency occurring in three linear coupling regimes can be related to the Manakov equations derived for multimode fibers [5-7, 14].

2. IM-FWM in FMFs and phase-matching condition

FWM is a process involving up to three optical fields at different frequencies to produce a fourth wave through the third-order nonlinearity [15]. In this paper we focus on the so-called non-degenerate IM-FWM and study the generation of an idler wave using three waves distinct in frequencies (see Fig. 1). More specifically, we consider two pumps (labeled pump1 and pump2) at frequencies ω_1 and ω_3 that copropagate inside a FMF with a signal at frequency ω_2 . The idler wave is generated at a frequency ω_4 determined by the general energy conservation relation:

$$\omega_l = \omega_i + \omega_j - \omega_k , \qquad (1)$$

where the indices i, j, k, l label the waves. As is well known [2], for an IM-FWM process to become efficient, the following phase-matching condition also needs to be satisfied:

$$\Delta \boldsymbol{\beta} = \boldsymbol{\beta}^{(m)}(\boldsymbol{\omega}_l) + \boldsymbol{\beta}^{(n)}(\boldsymbol{\omega}_j) - \boldsymbol{\beta}^{(o)}(\boldsymbol{\omega}_k) - \boldsymbol{\beta}^{(p)}(\boldsymbol{\omega}_l) \approx 0 , \qquad (2)$$

where $\beta^{(r)}(\omega_q)$ is the propagation constant of the *q*th wave propagating in the *r*th fiber mode.

In practice, the frequencies of the four waves participating in the FWM process are close enough to each other that we can expand all propagation constants in a Taylor series around a common reference frequency ω_0 , often chosen somewhere near the middle of the frequency range involved. For a given mode *r*, expansion to third order provides

$$\beta^{(r)}(\boldsymbol{\omega}) \simeq \beta_0^{(r)} + \beta_1^{(r)} \Delta \boldsymbol{\omega} + \frac{\beta_2^{(r)}}{2} (\Delta \boldsymbol{\omega})^2 + \frac{\beta_3^{(r)}}{6} (\Delta \boldsymbol{\omega})^3 + \dots,$$
(3)



Fig. 1. Wave configurations and notations for two non-degenerate IM-FWM processes: (a) PROC1 and (b) PROC2.

where $\Delta \omega = \omega - \omega_0$ and $\beta_0^{(r)} = \beta^{(r)}(\omega_0)$. The first, second and third derivatives, β_1 , β_2 and β_3 , evaluated at ω_0 , represent the inverse of the group velocity, its dispersion and third-order dispersion, respectively [2]. For simplicity, we only consider terms up to second order in this section. For a SMF, all waves involved in the FWM process share the same spatial mode and hence have the same values of β_n to all orders. It is easy to deduce that $\Delta\beta$ cannot vanish in SMFs for any non-degenerate FWM process unless one operates at wavelengths near the zero group-velocity dispersion ($\beta_2 \sim 0$). Indeed, FWM in standard SMFs typically requires other techniques to realize phase matching, such as the use of nonlinear phase shifts [16], or the use of birefringent fibers [17].

The situation is different in FMFs whose spatial modes have in general different propagation constants, making it possible to make $\Delta\beta = 0$ by simply propagating different waves in different fiber modes [18]. Here we consider the specific fiber used in the 2013 experiment reported in Ref. [8]. It is a graded-index fiber and supports 3 spatial modes, LP01, LP11a, and LP11b in the linearly-polarized (LP) mode approximation [19]; the last two among these are degenerate and share the same propagation constant and are referred together as LP11 modes. Such a FMF is capable of transmitting 6 sets of wavelength-division multiplexing (WDM) channels, because each spatial fiber mode has two orthogonal states of polarization (again in the LP-mode approximation). In order to generate an idler wave through IM-FWM, we consider the configuration shown in Fig. 1 where two waves (probe and pump2) are injected into the LP01 mode, and another wave (pump1) into the LP11a mode. In the experiment, frequencies of the probe and pump2 waves were swept over the entire C-band, and power of the idler wave generated in the LP11 mode was measured. We are interested in two IM-FWM processes, denoted as PROC1 and PROC2 following Ref. [8, 20] and shown schematically in Fig. 1. The two processes differ in whether the probe or pump2 is conjugated in the IM-FWM process. The energy relations for PROC1 and PROC2 are described by Eq. (1) and, after identifying the waves from Fig. 1, become

$$\omega_4 = \omega_1 - \omega_2 + \omega_3 , \quad (PROC1) \tag{4}$$

$$\omega_4 = \omega_1 + \omega_2 - \omega_3 . \quad (PROC2) \tag{5}$$

The phase matching conditions have been discussed in Ref. [8], and they turn out to be identical for both processes:

$$\beta_1^{11} + \frac{\beta_2^{11}}{2} \left(\Delta \omega_1 + \Delta \omega_4 \right) = \beta_1^{01} + \frac{\beta_2^{01}}{2} \left(\Delta \omega_2 + \Delta \omega_3 \right) , \qquad (6)$$

where the superscripts 01 and 11 correspond to LP01 and LP11 modes, respectively. The contribution of β_3 is ignored here to simplify the discussion but is included later in (19). Equation (6) has a simple interpretation: phase matching of IM-FWM in a FMF is achieved when the group velocities at the average frequencies of the two waves present in each spatial mode are equal. For the 6-mode FMF considered, based on the measured group velocities [8], equation (6) can be satisfied when the waves in the LP11 spatial modes are separated from the two waves in the LP01 spatial mode by about 16 nm. In contrast to SMFs, where effective phase matching can only be achieved at certain pump power levels, full phase matching in a FMF can be achieved at arbitrarily low power levels for a given wavelength configuration.

3. Nonlinear modal propagation equations

To predict the power of the idler wave generated through IM-FWM, we need to solve a set of coupled nonlinear propagation equations, developed previously by several authors [5,21]. The set that we use for this purpose was given earlier as Eq. (29) in Ref. [14]:

$$\frac{\partial \mathbf{A}}{\partial z} = -\frac{\alpha}{2}\mathbf{A} + i\mathbf{B}_{0}\mathbf{A} - \mathbf{B}_{1}\frac{\partial \mathbf{A}}{\partial t} - i\frac{\mathbf{B}_{2}}{2}\frac{\partial^{2}\mathbf{A}}{\partial t^{2}} + \frac{\mathbf{B}_{3}}{6}\frac{\partial^{3}\mathbf{A}}{\partial t^{3}} + i\mathbf{Q}\mathbf{A} + i\frac{\gamma}{3}\iint [(\mathbf{A}^{\mathrm{T}}\mathbf{G}^{(2)}\mathbf{A})\mathbf{G}^{(2)*}\mathbf{A}^{*} + 2(\mathbf{A}^{\mathrm{H}}\mathbf{G}^{(1)}\mathbf{A})\mathbf{G}^{(1)}\mathbf{A}]\mathrm{d}x\mathrm{d}y.$$
(7)

In this matrix equation, **A** is a column vector with N = 2M elements, where *M* is the number of spatial modes. The *N* modes are arranged in a descending order with respect to their propagation constants, i.e.,

$$\mathbf{A} = [A_{\text{LP01x}}, A_{\text{LP01y}}, A_{\text{LP11ax}}, A_{\text{LP11bx}}, A_{\text{LP11by}}, \dots]^{\text{T}}.$$
(8)

The superscripts T and H denote the transpose and the Hermitian conjugate, respectively. **B**₀, **B**₁, **B**₂, and **B**₃ are $N \times N$ diagonal matrices containing the values of β_0 , β_1 , β_2 , and β_3 for each mode, and all modes are assumed to have the same linear absorption coefficient α .

The nonlinear effects are included through the last term containing the nonlinear parameter γ , defined relative to the fundamental spatial mode [14]. Their contributions depend on the mode profiles entering in Eq. (7) through two matrices $\mathbf{G}^{(1)}$ and $\mathbf{G}^{(2)}$, whose elements are given by

$$\mathbf{G}_{ij}^{(1)} = \Gamma_{ij} F_i^* F_j, \qquad \mathbf{G}_{ij}^{(2)} = \Gamma_{ij} F_i F_j . \tag{9}$$

The transverse profile $F_i(x,y)$ of the *i*th mode is normalized such that $\iint |F_i(x,y)|^2 dx dy = 1$, and Γ_{ij} is defined as

$$\Gamma_{ij} = \delta[\operatorname{mod}(i+j+1,2),1]. \tag{10}$$

Here the Kronecker delta function and the modular arithmetic operation lead to $\Gamma_{ij} = 1$ when both *i* and *j* are odd or even numbers, but $\Gamma_{ij} = 0$ in all remaining cases.

As can be seen from the form of the nonlinear term in Eq. (7), the nonlinear effects depend on the integration of the product of two matrices $\mathbf{G}^{(1)}$ and $\mathbf{G}^{(2)}$. Therefore, all nonlinear effects, including IM-FWM, involve spatial profiles of up to 4 modes participating in a specific nonlinear process [14], and are governed by the nonlinear overlap factors:

$$f_{lmnp} = \iint F_l^* F_m F_n F_p^* dx dy.$$
⁽¹¹⁾

However, only a few of these factors are nonzero in practice. Their values for our 6-mode FMF are found to be [20, Tab. II]:

$$f_{lmnp} = \begin{cases} 1 & \text{for } 4 \text{ waves in LP01 mode} \\ 0.747 & \text{for } 4 \text{ waves in LP11a or LP11b} \\ 0.496 & \text{for } 2 \text{ in LP01 and } 2 \text{ in LP11a or LP11b} \\ 0.249 & \text{for } 2 \text{ in LP11a and } 2 \text{ in LP11b} \\ 0 & \text{for } \text{ all other cases} \end{cases}$$
(12)

Random linear mode coupling [6, 22, 23] is included in Eq. (7) through a random matrix **Q**. Considering fluctuations in the refractive index profile, represented in the form of a scalar field $\Delta \varepsilon(x, y, z)$ as the source of random coupling, the elements $q_{ij}(z)$ of **Q**, with i, j = 1, 2, ..., N, can be calculated from the coupled-mode theory using (see [12, Sec. 4.7] and [14, Eq. (40)])

$$q_{ij}(z) = \frac{\Gamma_{ij}k_0}{2n_{\rm eff}} \iint \Delta \varepsilon(x, y, z) F_i(x, y) F_j^*(x, y) dx dy .$$
(13)

For an ideal fiber, $\Delta \varepsilon(x, y, z) = 0$, and there is no linear coupling between different spatial modes since they represent eigenmodes of an ideal fiber. However, random linear mode coupling may need to be included in practice, and its impact is discussed in Sec. 6.

4. IM-FWM in the absence of linear mode coupling

In this section, we ignore linear mode coupling by setting $\mathbf{Q} = 0$ in Eq. (7) and solve this matrix equation numerically for the two IM-FWM processes shown in Fig. 1. In all cases, we assume that both pumps and the probe are launched using continuous-wave (CW) lasers operating in the wavelength range of 1530 to 1570 nm (C-band amplification).

We first focus on the PROC1 process and assume that the pump 1 at $\lambda_{pump1} = 1530$ nm is injected into the LP11a mode with 16 dBm (40 mW) power, the pump 2 at $\lambda_{pump2} = 1554$ nm is launched into the LP01 mode with 18-dBm power, and a probe at $\lambda_{probe} = 1546$ nm is injected into LP01 mode with 6-dBm power. We choose the reference frequency $\omega_0 = 2\pi c/\lambda_0$ to correspond to $\lambda_0 = 1540$ nm. The dispersion parameters of our fiber are deduced from measurements using Eq. (3) and are summarized in Tab. 1. The LP11 values apply to both LP11a and LP11b modes. When expressed in terms of $\beta_2 = -\lambda^2 D/(2\pi c)$, we obtain $\beta_2^{01} = -24.30 \text{ ps}^2/\text{km}$ and $\beta_2^{11} = -23.04 \text{ ps}^2/\text{km}$. The differential group delay (DGD) between LP01 and LP11a (or LP11b) spatial modes is $d = \beta_1^{01} - \beta_1^{11} \approx 300 \text{ ps/km}$. All spatial modes are assumed to share the same value of dispersion slope $S = 0.055 \text{ ps/(nm^2-km)}$. The nonlinear parameter for our FMF is $\gamma = 1.77 \text{ W}^{-1}/\text{km}$, and we include losses in our calculations using $\alpha = 0.226 \text{ dB/km}$. We assume that all three waves are polarized in the same direction, along the *x* axis.

We solve Eq. (7) numerically over a distance of 4.7 km and plot the spectrum of the total field at the fiber output. Figure 2 compares the input (solid green lines) and output (dotted red lines) spectra. As seen in the figure, an idler wave is generated in the LP11a mode. Its wavelength of 1538 nm corresponds to what is expected from the PROC1 IM-FWM process. The powers

Table 1. Dispersion parameters for the LP01 and LP11 spatial modes for the 6 mode FMF.

	β_1 - β_1^{01}	D	S
	ps/km	ps/(nm-km)	ps/(nm ² -km)
LP01	0	19.3	0.055
LP11	300	18.3	0.055



Fig. 2. Generation of the idler through PROC1 in a 4.7-km-long FMF supporting three spatial modes. Comparison of input (solid green) and output (dotted red) spectra indicates that the idler is created at the expected wavelength of 1538 nm. The idler is generated in the LP11a mode.

of pump1 and pump2 are reduced by 1.18 and 1.09 dBm, respectively. Most of this power reduction is due to linear losses since the total loss for our 4.7-km-long fiber is 1.06 dB. A small part of the pump powers is transferred to the probe and idler waves through IM-FWM. Indeed, the 0.51 dB reduction in the probe power is less than the fiber loss of 1.06 dB. The difference of 0.55 dB corresponds to parametric amplification of the probe from the IM-FWM process.

We can estimate the idler power using a simple model. For the configurations shown in Fig. 1 and in the absence of linear coupling ($\mathbf{Q} = 0$), Eq. (7) is reduced to the following two equations for the fields in LP01x and LP11ax modes:

$$\frac{\partial A_1}{\partial z} = \hat{D}_1 A_1 + i\gamma f_{1111} |A_1|^2 A_1 + i\gamma f_{1331} A_3^2 A_1^* + i\gamma (f_{3131} + f_{3311}) |A_3|^2 A_1,$$
(14)

$$\frac{\partial A_3}{\partial z} = \hat{D}_3 A_3 + i\gamma f_{3333} |A_3|^2 A_3 + i\gamma f_{3113} A_1^2 A_3^* + i\gamma (f_{1313} + f_{1133}) |A_1|^2 A_3,$$
(15)

where \hat{D}_r is the dispersion operator for the *r*th mode in the form

$$\hat{D}_{r} = i\beta_{0}^{(r)} - \beta_{1}^{(r)}\frac{\partial}{\partial t} - i\frac{\beta_{2}^{(r)}}{2}\frac{\partial^{2}}{\partial t^{2}} + \frac{\beta_{3}^{(r)}}{6}\frac{\partial^{3}}{\partial t^{3}} - \frac{\alpha}{2}.$$
(16)

The nonlinear term that generates the idler wave for both PROC1 and PROC2 is the last term governing intermodal cross-phase modulation (IM-XPM). Since power reduction of the two pumps is mainly due to linear fiber losses, we can employ the undepleted-pump approximation [15]. Also, with the assumption of perfect phase matching and CW waves, the idler field A_I within the LP11ax mode satisfies

$$\frac{dA_I}{dz} \approx i\gamma (f_{3131} + f_{3311}) A_{\rm p1}(0) A_{\rm p2}(0) A_{\rm B}^*(0) e^{-3\alpha z/2} - (\alpha/2) A_I , \qquad (17)$$



Fig. 3. Same as in Fig. 2 but for PROC2. The wavelength of the idler is $\lambda_{idler} = 1536$ nm. A second idler is generated in the LP01 mode at $\lambda = 1540$ nm through cascaded IM-FWM (see main text).

where A_{p1} , A_{p2} , and A_B are the amplitudes for pump1, pump2, and probe, respectively. As the preceding equation is linear, it can be easily solved to obtain the idler power $P_I(z) = |A_I(z)|^2$. For the parameter values used in our numerical simulations, we obtain $P_I(L) = -4$ dBm, which is quite close to the numerical value of -3.78 dB obtained by solving Eq. (7). The reason our estimated value is slightly lower than the numerical value is that Eq. (17) does not consider parametric amplification of the probe through IM-FWM.

We also performed simulations for the PROC2 IM-FWM process by switching the wavelengths of the probe and the second pump. More specifically, we chose $\lambda_{\text{probe}} = 1552 \text{ nm}$ and $\lambda_{\text{pump2}} = 1546 \text{ nm}$, following the experimental work [8]. We adjusted the value of $\Delta\beta_1$ slightly (about 0.1%, $\beta_1^{11} - \beta_1^{01} = 299.7 \text{ ps/km}$) to ensure that PROC2 is perfectly phase-matched (phase matching in PROC2 is very sensitive to the wavelength and fiber parameters as will be seen in Sec. 5). The results are shown in Fig. 3. As expected, an idler wave is now created in the LP11a mode at a wavelength of 1536 nm through PROC2. The power of this idler wave is quite close to the idler in Fig. 2. This is expected since both processes are perfectly phase-matched in our simulations. Notice that another much weaker idler is also generated in the LP01 mode at 1540 nm. Its origin lies in the cascaded IM-FWM process (of type PROC1) in which the idler wave generated from PROC2 plays the role of the probe. The main reason for the low intensity is that this process is not fully phase-matched. As discussed in Sec. 2, in order to have efficient FWM, the averaged wavelength for the LP01 mode should be about 16 nm longer than that for the LP11 mode. For the cascaded IM-FWM process, the difference in the averaged wavelength is about 10 nm, and this condition is not satisfied.

5. Bandwidth analysis for PROC1 and PROC2

Since a FWM process requires phase matching, any detuning of the two pumps or the probe from the exact phase-matching conditions will reduce IM-FWM efficiency. This section focuses on the tuning bandwidth of the two IM-FWM processes discussed in Sec. 4. We assume for simplicity that the power levels are sufficiently low that the SPM and XPM do not affect phase

matching significantly. Then, the phase-matching bandwidth can be estimated from the linear phase mismatch given in Eq. (2). Its use predicts that the idler power in the undepleted-pump approximation is proportional to [15, Eq. (2.3.2)]

$$\eta = \left[\frac{\sin(\Delta\beta L/2)}{(\Delta\beta L/2)}\right]^2,\tag{18}$$

where $\Delta\beta$ is the linear phase mismatch and *L* is the fiber length. We stress that η is the phasematching efficiency relative to the ideal phase-matching conditions and should not be confused with the overall efficiency or gain of the underlying FWM process, which relates to the ratio of the generated idler power to the pumps or probe [24].

Before we can use Eq. (18), we need to calculate $\Delta\beta$ for the two IM-FWM processes considered. Using the numbering of the waves as shown in Fig. 1 and substituting Eq. (3) into Eq. (2), we obtain an identical expression for both PROC1 and PROC2:

$$\Delta \beta = (\Delta \omega_{1} - \Delta \omega_{4}) \Big\{ \beta_{1}^{11} - \beta_{1}^{01} + \frac{1}{2} \beta_{2}^{11} (\Delta \omega_{1} + \Delta \omega_{4}) - \frac{1}{2} \beta_{2}^{01} (\Delta \omega_{2} + \Delta \omega_{3}) \\ + \frac{1}{6} \beta_{3}^{11} \big[\Delta \omega_{1}^{2} + \Delta \omega_{1} \Delta \omega_{4} + \Delta \omega_{4}^{2} \big] - \frac{1}{6} \beta_{3}^{01} \big[\Delta \omega_{2}^{2} + \Delta \omega_{2} \Delta \omega_{3} + \Delta \omega_{3}^{2} \big] \Big\},$$
(19)

where $\Delta \omega_i = \omega_i - \omega_0$ for i = 1...4. Setting $\beta_3^{11} = \beta_3^{01} = 0$ leads to the phase-matching condition given in Eq. (6).

Let us assume that PROC1 is phase-matched for a specific probe frequency ω_2 so that $\Delta\beta = 0$ at that frequency. If we now detune the probe by an amount $\delta\omega$, the new probe and idler frequencies become $\omega'_2 = \omega_2 + \delta\omega$, and $\omega'_4 = \omega_4 - \delta\omega$. Using them, we obtain the phase mismatch due to detuning of the probe frequency by $\delta\omega$ for PROC1:

$$(\Delta\beta)_{1} = \frac{1}{2}\delta\omega(\Delta\omega_{2} - \Delta\omega_{3} + \delta\omega) \left[\beta_{2}^{01} + \beta_{2}^{11} + \frac{1}{3}\beta_{3}^{11}(3\Delta\omega_{1} - 2\Delta\omega_{2} + 2\Delta\omega_{3} - \delta\omega) + \frac{1}{3}\beta_{3}^{01}(\Delta\omega_{3} + 2\Delta\omega_{2} + \delta\omega)\right].$$
(20)

Although PROC2 shares the same expression for $\Delta\beta$ with PROC1, the impact of probe detuning is quite different for PROC2 because relative locations for pump2 and probe are switched (see Fig. 1). For a detuning $\delta\omega$ in the probe frequency, the idler frequency now changes to $\omega'_4 = \omega_4 + \delta\omega$. As a result, the phase mismatch for PROC2 is given by

$$(\Delta\beta)_{2} = \frac{1}{2}\delta\omega(\Delta\omega_{2} - \Delta\omega_{3} - \delta\omega) \left[\beta_{2}^{11} - \beta_{2}^{01} + \frac{1}{3}\beta_{3}^{11}(3\Delta\omega_{1} + 2\Delta\omega_{2} - 2\Delta\omega_{3} + \delta\omega) - \frac{1}{3}\beta_{3}^{01}(\Delta\omega_{3} + 2\Delta\omega_{2} + \delta\omega)\right].$$
(21)

The two expressions for the phase mismatch can be simplified further by noting that $\delta \omega$ is a small quantity in practice compared to any $\Delta \omega_i$ (i = 1-3). With this simplification, the phase mismatch for both processes is proportional to the probe detuning $\delta \omega$ as

$$(\Delta\beta)_1 \approx \frac{1}{2} \delta\omega (\Delta\omega_2 - \Delta\omega_3) M_1,$$
 (22)

$$(\Delta\beta)_2 \approx \frac{1}{2}\delta\omega(\Delta\omega_2 - \Delta\omega_3)M_2$$
, (23)

where the two constants M_1 and M_2 are given by

$$M_{1} = \beta_{2}^{11} + \frac{1}{3}\beta_{3}^{11}(3\Delta\omega_{1} - 2\Delta\omega_{2} + 2\Delta\omega_{3}) + \beta_{2}^{01} + \frac{1}{3}\beta_{3}^{01}(\Delta\omega_{3} + 2\Delta\omega_{2}),$$
(24)

$$M_2 = \beta_2^{11} + \frac{1}{3}\beta_3^{11}(3\Delta\omega_1 + 2\Delta\omega_2 - 2\Delta\omega_3) - \beta_2^{01} - \frac{1}{3}\beta_3^{01}(\Delta\omega_3 + 2\Delta\omega_2).$$
(25)



Fig. 4. (a) Efficiency as a function of λ_{probe} and λ_{pump2} for PROC1 when $\lambda_{\text{pump1}} = 1530$ nm. (b) Efficiency curve for $\lambda_{\text{pump2}} = 1554$ nm (full green curve) from Eq. (22). Numerical result (input power $P_{p1} = 4$, $P_{p2} = 6$, $P_B = -6$ dBm, dotted red curve) agrees well with the analytic approach.

The main difference between the two constant M_1 and M_2 is the sign change of the last term, indicating that phase mismatch depends on the difference in β_2 and β_3 values for the LP01 and LP11 modes for PROC2, while it depends on their sum for PROC1. The differences in β_2 and β_3 values for the two modes are quite small in practice (typically less than 10%). As a result, PROC2 has generally a much larger bandwidth compared to PROC1 because M_1 is much larger than M_2 .

To verify the accuracy of the analytical approximation, we use Eq. (18) to calculate the phase-matching efficiency η as a function of λ_{probe} and λ_{pump2} . The results for PROC1 using Eq. (22) are shown in Fig. 4(a). As can be seen, the ridge follows a straight line in the two-dimensional plane formed by λ_{probe} and λ_{pump2} . The slope of the ridge depends on the values of β_2 and β_3 for the two spatial modes, LP01 and LP11a, in which the four waves are propagating. The bandwidth with respect to the probe detuning is quite small (~0.01 nm) for PROC1. On the other hand, λ_{pump2} can vary over more than 2 nm. Figure 4(b) shows the efficiency curve for the specific value $\lambda_{\text{pump2}} = 1554$ nm (full green curve). We also show the numerical results with a dotted red curve obtained using Eq. (7) using low power levels for all waves to rule out the impact of nonlinearity on phase matching.

We performed similar efficiency calculations for PROC2, and the results using Eq. (23) are shown in Fig. 5(a). Notice that the x and y axes have been switched in this figure compared to Fig. 4(a). The efficiency map for PROC2 is similar to that of PROC1 in the sense that the peak position follows a straight line in the two-dimensional plane formed by λ_{probe} and λ_{pump2} . However there is a practical difference related to the switching of the axes. Now, for PROC2,



Fig. 5. Same as Fig. 4 but for PROC2 and with $\lambda_{pump2} = 1546$ nm from Eq. (23). Notice the difference in scale of the λ_{probe} axis and difference in bandwidth of the process. The analytical formula agrees well with the numerical results.

the bandwidth with respect to pump detuning is ~0.01 nm but the probe wavelength can be detuned more than 2 nm. This feature is evident in Fig. 4(b) where we show the efficiency curve at a specific value $\lambda_{pump2} = 1546$ nm (full green curve). We also show the numerical results with a dotted red curve obtained using Eq. (7), using once again low power levels for all waves. The excellent agreement between the two curves for both PROC1 and PROC2 supports well the analytic approximations using Eqs. (18), (22) and (23).

One may ask why PROC2 has a much larger bandwidth (~2 nm) than that of PROC1 (~0.01 nm) with respect to probe detuning. A physical reason behind this difference is related to the energy conservation conditions that determine the probe frequency. For PROC1, the idler and the probe frequencies move in the opposite directions, as evident from Eq. (1). In contrast, the two waves move in the same direction in the case of PROC2. Clearly the frequency difference $\omega_2 - \omega_4$ is preserved in the case of PROC2. It is this property that enhances the bandwidth of PROC2. As the probe wavelength is detuned, the idler wavelength changes in such a way that the phase-matching condition in Eq. (6) is maintained over a much wider range compared to PROC1. Using the parameter values for our FMF, we find from Eqs. (22) and (23) that $(\Delta\beta)_1 \approx 1.4963 \times 10^{-10} \delta \omega$ and $(\Delta\beta)_2 \approx -8.0439 \times 10^{-13} \delta \omega$. We can estimate the bandwidth from η in Eq. (18) by noting that $\eta = 0$ for $\Delta\beta L/2 = \pi$. Taking L = 4.7 km, it follows that $\delta v = \delta \omega/(2\pi) \approx 1.5$ GHz for PROC1 and 280 GHz for PROC2. In wavelength units, these numbers correspond to 12 pm and 2.2 nm, respectively. These values agree with the numerical curves in Figs. 4 and 5.

We now discuss briefly the impact of nonlinearity on the phase-matching conditions that

increases in importance as the power of the input waves increases. For this purpose, we need to find how the phase of each wave is affected by self-phase modulation (SPM) and cross-phase modulation (XPM). In our case, the powers of the two pumps (16 and 18 dBm) are much higher than those of the probe and idler waves (6 and -4 dBm). If we only retain the nonlinear contributions coming from the two pumps, the nonlinear contribution to the phase mismatch for PROC1 and PROC2 is found to be

$$(\Delta\beta)_1^{NL} = -\gamma L(f_{3333}P_{p1} + f_{1111}P_{p2}), \qquad (26)$$

$$(\Delta\beta)_2^{NL} = -\gamma L(f_{3333}P_{\rm p1} - f_{1111}P_{\rm p2}).$$
⁽²⁷⁾

It is interesting to note that nonlinearity affects the two IM-FWM processes differently. The nonlinear contributions to the phase matching increase with increasing pump powers for both pumps for PROC1 but go in opposite directions for pump2 in PROC2. Clearly, PROC2 is generally less sensitive to pump-power-dependent phase matching. The phase-matching condition become insensitive to the pump powers in PROC2 when $P_{p1}/P_{p2} = f_{1111}/f_{3333}$.

To help assess how nonlinearity impacts phase matching in the experiments, we performed a series of simulations for both IM-FWM processes where we vary the pump powers while keeping other parameters unchanged. We plotted the efficiency curves for PROC1 and PROC2 in Figs. 6(a) and (b), respectively. As shown in these two figures, the impact of nonlinearity is mainly to shift slightly the efficiency curves with minimum impact on the bandwidth. As seen from Fig. 6(a), the efficiency curve shifts towards the same direction (shorter wavelength) when either P_{p1} or P_{p2} increases. On the other hand, the efficiency curve shifts in the opposite directions when P_{p1} or P_{p2} increases for PROC2 (b). These observed behaviors agree with Eqs. (26) and (27). By carefully looking at the shift of the efficiency curves in Fig. 6, one can observe that increasing P_{p2} produces a larger shift than that from increasing P_{p1}. This behavior can be understood directly from Eqs. (26) and (27), since the nonlinear overlapping factor f_{1111} is greater than f_{3333} (For our FMF, $f_{1111} = 1, f_{3333} \approx 0.75$).

Looking from the efficiency curve for PROC2, the amount of wavelength shift between the dashed–green curve [P = (10, 10, 6) dBm] and solid–purple curve with "+" marker [P = (10, 16, 6) dBm] is about 0.09 nm. This wavelength shift can be estimated using Eq. (27). The nonlinear phase difference between these two curves is γf_{1111} (16 dBm - 10dBm) = 0.053 km⁻¹, which is compensated by an amount of frequency detuning about 65.9 GHz from Eq. (23). In wavelength units, this number corresponds to 0.085 nm, quite close to the value 0.09 nm from numerical simulation of Eq. (7). Similar estimation can also be done for PROC1.

6. Impact of linear mode coupling on IM-FWM

In this section, we consider the impact of random linear mode coupling on the two IM-FWM processes. For SMFs, linear coupling between the two polarization modes originates from the presence of random birefringence fluctuations that lead to random rotations of the polarization state of the field along the fiber length [11,12]. The length of fiber over which the state of polarization at the output becomes uncorrelated with its input value (i.e., the correlation length) depends on details of fiber design and fabrication. Experimental measurements on specific fibers have shown values ranging from a few meters to few tens of meters [25, 26]. In contrast, for fibers supporting multiple spatial modes, linear coupling between two different spatial modes varies greatly depending on the difference in their propagation constants [12, 27, 28]. Many physical effects lead to random linear coupling between modes. Examples of such physical mechanisms are bending, twisting, tension, kinks, pressure, and fluctuations in the fiber core shape and refractive index [11, 12]. Fluctuations in the refractive index can be modeled by a random quantity $\Delta \varepsilon(x, y, z)$ that is used to find the elements of the matrix $\mathbf{Q}(z)$ in Eq. (13). The



Fig. 6. Efficiency curves all shift to shorter wavelengths as P_{p1} or P_{p2} increase for (a) PROC1 while they shift in opposite directions for (b) PROC2 depending on whether P_{p1} or P_{p2} is increased. Powers vector is defined as $P = (P_{p1}, P_{p2}, P_B)$.

way polarization and spatial modes couple for a particular fiber depends on the strength and spatial frequency of each coupling mechanism for that fiber.

To evaluate the impact of random linear coupling on nonlinear propagation, we first assume that the linear coupling and phase evolution from difference in refractive indices occur on a length scale shorter than that associated with the nonlinear effects. Under this condition, we can treat the terms \mathbf{B}_0 and $\mathbf{Q}(z)$ separately from the other terms in Eq. (7) and write,

$$\frac{\partial \mathbf{A}(z)}{\partial z} = i[\mathbf{B}_0 + \mathbf{Q}(z)]\mathbf{A}(z) .$$
(28)

The exact way various fiber imperfections impact the coupling matrix $\mathbf{Q}(z)$ of multimode fibers still remains an active topic of investigation [28]. For our numerical simulations, we assume that the elements of the coupling matrix $\mathbf{Q}(z)$ follow Eq. (13). Because $\Delta \varepsilon(x, y, z)$ is a scalar field in Eq. (13), the matrix $\mathbf{Q}(z)$ does not introduce random coupling between the two polarization states of the same spatial mode. However, as mentioned earlier, we expect several physical mechanisms to randomly couple the two polarization states of the same spatial modes. To take this into account in a model that resembles what occurs in SMFs, we introduce the following matrix $\mathbf{R}(z)$ that randomly couples the two polarization states of each spatial

mode [29]:

$$\mathbf{R}(z) = \begin{bmatrix} r_{11p} & r_{12p} & 0 & 0 & 0 & 0 \\ r_{21p} & r_{22p} & 0 & 0 & 0 & 0 \\ 0 & 0 & r_{11l} & r_{12l} & 0 & 0 \\ 0 & 0 & r_{21l} & r_{22l} & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{11m} & r_{12m} \\ 0 & 0 & 0 & 0 & r_{21m} & r_{22m} \end{bmatrix},$$
(29)

where the subscripts p, l, m indicate that all 2×2 unitary matrices are independently random. $\mathbf{R}(z)$ is also unitary. Note that we omitted writing explicitly the z-dependence of each matrix element for clarity.

In the following numerical simulations, we solve Eq. (7) with the split-step Fourier method and incorporate the combined effects of B_0 and Q by multiplying the field after each numerical step Δz by the matrix

$$\mathbf{T}_{\Delta}(z) = \mathbf{R}(z) \exp[i(\mathbf{B}_0 + \mathbf{Q}(z))\Delta z].$$
(30)

Note that in this model, we assume full scrambling at each integration step, and therefore the rotation matrix **R** is not "proportional" to Δz . Even though not explicitly studied in this paper, this formulation of coupling allows to separately consider the impact of random linear coupling within each spatial mode from linear coupling between spatial modes. For instance, one can find an analytic derivation of the impact of random linear coupling between the two polarization modes of the same spatial mode alone on the nonlinear propagation equation in [5, 14].

The extent of linear coupling between modes depends on many fiber parameters and length of the system considered. In the Appendix, we present a simplified analysis that presents evidence of the existence of three different regimes of coupling. The analysis is based on Eq. (28) and considers a 6×6 coupling matrix. The three different regimes of propagation are found based on differences in propagation constants between spatial modes, the magnitude of the variances of the elements of $\mathbf{Q}(z)$ (all elements are assumed to have identical variances), and the factor by which the fiber length exceeds the correlation length L_c of random mode coupling. In the following, we therefore consider three distinct regimes of random linear coupling for $\mathbf{T}(z)$.

Model 1: No linear coupling among spatial modes In this case, none of the spatial modes linearly couple, but random linear coupling between the two polarization components of each spatial mode occurs. This simplified model is intended to address the case where all spatial modes have sufficiently distinct propagation constants. In this model, the 6×6 transfer matrix $\mathbf{T}(z)$ for a 6-mode fiber is identical to $\mathbf{R}(z)$ of Eq. (29) and assumes the following block-diagonal form

$$\mathbf{T_1}(z) = \begin{bmatrix} T_{11p} & T_{12p} & 0 & 0 & 0 & 0 \\ T_{21p} & T_{22p} & 0 & 0 & 0 & 0 \\ 0 & 0 & T_{11l} & T_{12l} & 0 & 0 \\ 0 & 0 & T_{21l} & T_{22l} & 0 & 0 \\ 0 & 0 & 0 & 0 & T_{11m} & T_{12m} \\ 0 & 0 & 0 & 0 & T_{21m} & T_{22m} \end{bmatrix} .$$
(31)

In Eq. (31) and equations below, we omit the *z*-dependence of the elements of the transfer matrices to simplify the notation.

Model 2: Strong random linear coupling between LP11a and LP11b In this case, the degenerate spatial modes LP11a and LP11b experience a strong random linear coupling, while their coupling to the LP01 mode is neglected by assuming a sufficiently large difference in their

Realization	Model 1	Model 2	Model 3
1	-7.41	-7.38	-49.28
2	-7.34	-7.40	-52.61
3	-7.37	-7.37	-50.87
4	-7.35	-7.38	-52.14
5	-7.40	-7.41	-51.25

Table 2. Power of the idler (in dBm) generated by the IM-FWM PROC2 for five different realizations of the random linear mode coupling for the three models considered.

propagation constants. As a result, the 6×6 transfer matrix for this case takes the following block-diagonal form

$$\mathbf{T_2}(z) = \begin{bmatrix} T_{11p} & T_{12p} & 0 & 0 & 0 & 0 \\ T_{21p} & T_{22p} & 0 & 0 & 0 & 0 \\ 0 & 0 & T_{11l} & T_{12l} & T_{13l} & T_{14l} \\ 0 & 0 & T_{21l} & T_{22l} & T_{23l} & T_{24l} \\ 0 & 0 & T_{31l} & T_{32l} & T_{33l} & T_{34l} \\ 0 & 0 & T_{41l} & T_{42l} & T_{43j} & T_{44j} \end{bmatrix},$$
(32)

where the 4×4 random unitary matrix takes into account random linear coupling among the LP11ax, LP11ay, LP11bx, and LP11by modes.

Model 3: Strong random linear coupling among all modes In this case, all modes are strongly linearly coupled, and the transfer matrix $T_3(z)$ is a full 6×6 random unitary matrix:

$$\mathbf{T_3}(z) = \begin{pmatrix} T_{11} & T_{12} & T_{13} & T_{14} & T_{15} & T_{16} \\ T_{21} & T_{22} & T_{23} & T_{24} & T_{25} & T_{26} \\ T_{31} & T_{32} & T_{33} & T_{34} & T_{35} & T_{36} \\ T_{41} & T_{42} & T_{43} & T_{44} & T_{45} & T_{46} \\ T_{51} & T_{52} & T_{53} & T_{54} & T_{55} & T_{56} \\ T_{61} & T_{62} & T_{63} & T_{64} & T_{65} & T_{66} \\ \end{pmatrix}$$
(33)

We performed numerical simulations for PROC2-type IM-FWM using these three random linear coupling models. The transfer matrix $T_i(z)$ (i = 1-3) was applied to the fields at each step of the split-step Fourier method. We propagated the six field components over 4.7 km of the FMF described in Tab. 1. A step size of 7.8 m was used as smaller step sizes generates similar results within the statistical uncertainty of the three models. Numerical simulations performed using Eq. (7) are repeated with different random seeds for the generation of the coupling matrices $T_i(z)$, with *i* being the model number. Table 2 shows how the idler power changes over five realizations of the random linear coupling process for the three linear coupling models. As in Sec. 4, the launched pump powers are 16 and 18 dBm, while the probe power is 6 dBm. As seen in Table 2, variations of the idler power for different random matrices realizations are relatively small (< 0.1 dBm) for Models 1 and 2, with an average value close to -7.4 dBm. These results suggest that strong linear coupling between the LP11a and LP11b modes does not affect much the IM-FWM process.

In contrast, in the case of Model 3 where the LP01 and LP11 modes are also linearly coupled, the idler power is drastically reduced and is in the range of -50 ± 2 dBm over five realizations of the random coupling process. Such larger changes indicate that fluctuations in phase

Realization	Model 1	Model 2	Model 3
1	-16.36	-16.27	-61.51
2	-16.19	-16.34	-59.58
3	-16.28	-16.29	-60.45
4	-16.30	-16.31	-61.75
5	-16.33	-16.32	-60.31

Table 3. Same as Table 2 but for 3 dB lower powers for the two pumps and the probe.

matching conditions resulting from linear coupling between the LP01 and LP11 modes are so important that IM-FWM is highly suppressed. These results indicate that phase matching cannot be achieves when three nondegenerate waves are launched in a FMF with nearly identical propagation constants.

Linear coupling between modes can be reduced in practice by designing fibers with large differences in their modal propagation constants [12], such as between the LP01 and LP11 modes for our FMF. Model 2 would be most realistic in this case. In Sec. 4, the power of the idler was about -3.93 dBm for PROC2 in the absence of any linear mode coupling. The average value of the idler power for Model 2 is about -7.4 dBm. Since random linear mode coupling decreases the idler power by 3.47 dB, it will help in reducing the penalty associated to the interference of the idler with other channels in a WDM system. To estimate the validity of using Model 1 in this section, we used the generalized Manakov equation [Eq. (26)] of Ref. [14] that includes the coupling model described by the matrix T_1 . The numerical value obtained for the idler power was -7.36 dBm, close to the values obtained in Table 2.

Analytic considerations can also be used to estimate the linear mode-coupling penalty. As discussed in Sec. 4, the nonlinear term responsible for generating the idler wave is proportional to $|A_3|^2A_1$, which is a "XPM-type" term as described in Refs. [14] and [20]. It was shown in Ref. [14] that the factor of 2 associated with the XPM terms in the absence of linear coupling is reduced to a factor 4/3 in the presence of random polarization rotations. As a result, the idler field is reduced by a factor of 2/3, which results in reduction by " $(3/2)^2$ " or 3.5 dB of the idler power.

In contrast, any linear coupling between the LP01 and LP11 (a or b) reduces the FWM efficiency drastically. The reason for this decrease in idler power is that the phase matching condition can no longer be satisfied. A generalized Manakov equation has also been derived for the strong coupling case (Model 3) [5, 14]. In fact, as can be seen from this Manakov equation, only "SPM" terms survive after the random averaging. Therefore, it is expected that no efficient IM-FWM happens in such a strong coupling region. Indeed, one can confirm this by performing simulations of Eq. (7) with increased number of steps. We found that the idler power did not change for Model 1 and Model 2, but it kept decreasing with the increasing number of steps in the case of Model 3. For a sufficient large step number (i.e., a sufficient small coupling length), the idler power became vanishingly small. Clearly, the results for Model 3 depend on the coupling length of the fiber. For a fixed coupling length, the idler power converges to a stable fixed value that depends on the fiber length. Note also that, in the strong coupling regime, a FMF approaches the behavior of a SMF with the noticeable difference that many modes can occupy the same frequencies.

To verify the power dependence of the IM-FWM processes studied, we reduced the power of all input waves by 3 dB, i.e., $P_{p1} = 13$ dBm, $P_{p2} = 15$ dBm, and $P_{probe} = 3$ dBm. All other parameters remain unchanged. The results for PROC2 are summarized in Table 3. Calculations based on the generalized Manakov equation [14, Eq. (26)] give -16.31 dBm for the idler power.

As can be seen by comparing Tables 2 and 3, the power of the idler generated for all cases is reduced by about 9 dB, a scaling in agreement with Eq. (17).

7. Conclusions

In this paper, we have studied a simple analytical model and performed numerical studies of intermodal four-wave mixing (IM-FWM) in few-mode fibers, including birefringence fluctuations and random linear mode coupling. Two different intermodal IM-FWM processes have been investigated. We found that one of these IM-FWM processes has a phase-matching bandwidth about 100 times larger than the other. The impact of nonlinearity on the phase-matching condition and the bandwidth was also studied. We included random linear mode coupling between fiber modes using three different models and found that random linear coupling always reduces the IM-FWM efficiency relative to the absence of any linear mode coupling. The reduction in efficiency is relatively small (on the order of 3 dB) for weakly coupled spatial modes, but becomes very large when all modes are strongly coupled. These results may prove useful in the context of SDM systems designed using wavelength-division multiplexing in multimode and multicore fibers.

8. Appendix: Asymptotic Behavior of Concatenated Random Mode Coupling Matrices

We are interested in the asymptotic behavior of the solution of Eq. (28), where \mathbf{B}_0 is a diagonal matrix of the form $\mathbf{B}_0 = \text{diag}(\beta_0^{01}, \beta_0^{01}, \beta_0^{11}, \beta_0^{11}, \beta_0^{11}, \beta_0^{11})$ and $\mathbf{Q}(z)$ is a Hermitian random matrix with all entries assumed to have the same standard deviation σ . Note that, in contrast to Eq. (13) of the main text, the model considered for $\mathbf{Q}(z)$ includes more general fiber deformations that randomly couple all modes, including those with orthogonal polarizations.

First, we introduce the following changes of variables,

$$\mathbf{B}_0 = \Delta \hat{\mathbf{B}} - \Sigma I, \qquad \mathbf{A} = \hat{\mathbf{A}} \exp\left(-i\Sigma z\right), \tag{34}$$

$$z = \hat{z}/\Delta, \qquad \mathbf{Q} = \Delta \sigma \hat{\mathbf{Q}},$$
 (35)

where $\Delta = (\beta_0^{01} - \beta_0^{11})/2$ and $\Sigma = (\beta_0^{01} + \beta_0^{11})/2$. In terms of the new normalized "hat" variables, Eq. (28) becomes

$$\frac{\partial \hat{\mathbf{A}}(\hat{z})}{\partial \hat{z}} = i[\hat{\mathbf{B}} + \sigma \hat{\mathbf{Q}}(\hat{z})] \hat{\mathbf{A}}(\hat{z}), \qquad (36)$$

where $\hat{\mathbf{B}} = \text{diag}(1, 1, -1, -1, -1, -1)$ is a normalized diagonal matrix and $\hat{\mathbf{Q}}$ is a normalized Hermitian random matrix with all elements having a standard deviation of 1. Mathematically, this model of $\hat{\mathbf{Q}}$ corresponds to the well known Gaussian Unitary Ensemble (GUE).

We define the evolution matrix in the normalized space $\hat{\mathbf{T}}(\hat{z})$ as

$$\hat{\mathbf{A}}(\hat{z}) = \hat{\mathbf{T}}(\hat{z})\,\hat{\mathbf{A}}'\,,\tag{37}$$

where $\hat{\mathbf{A}}'$ is stationary in \hat{z} . Using Eqs. (36) and (37), the solution for $\hat{\mathbf{T}}(\hat{z})$ from \hat{z}_1 to \hat{z}_2 is given by,

$$\hat{\mathbf{T}}(\hat{z}_2) = \exp\left[i\left(\hat{\mathbf{B}} + \sigma \langle \hat{\mathbf{Q}}(\hat{z}) \rangle\right)(\hat{z}_2 - \hat{z}_1)\right]\hat{\mathbf{T}}(\hat{z}_1) .$$
(38)

where $\langle \hat{\mathbf{Q}}(\hat{z}) \rangle$ is the z-independent value of $\hat{\mathbf{Q}}(\hat{z})$, assumed to be nearly constant over the interval $\hat{z}_2 - \hat{z}_1 \equiv L_c$. Let's separate the fiber of length L into $N = L/L_c$ segments of equal lengths L_c , where L_c is the correlation length over which the entries of $\hat{\mathbf{Q}}(\hat{z})$ can be considered as constant. Assuming a fiber that is statistically uniform along the length, the entries of $\langle \hat{\mathbf{Q}}(\hat{z}) \rangle$ can be considered as independent and identically distributed between distinct segments. Denoting

by $\hat{\mathbf{Q}}_j$ the random coupling matrix associated to the *i*th fiber segment of length L_c , the total evolution matrix over the fiber length **T** can then be written as,

$$\mathbf{T} = \prod_{j=1}^{N} \exp[(i/p) \left(\hat{\mathbf{B}} + \sigma \, \hat{\mathbf{Q}}_{j}\right)], \qquad (39)$$

where we defined $p = 1/(\Delta L_c)$ and made use of $z = \hat{z}/\Delta$ to return to physical units of distance. The ranges of the three variables that impact **T** in Eq. (39) are chosen to be $\sigma \in [10^{-5}, 10^2]$, $N \in [10^1, 10^4]$ and $p \in [10^{-5}, 10^6]$.

The asymptotic behavior in N of **T** in Eq. (39) for different values of p and σ is studied below. The parameter p can be interpreted as a ratio of a dephasing fiber length $(L_{\beta} = 1/\Delta = 2/(\beta_0^{01} - \beta_0^{11}))$ to the correlation length of random fluctuations (L_c) . The parameter σ is the amplitude of the random coupling coefficient. The ratio σ/p can therefore be loosely interpreted as the ratio of the amplitude of the random coupling to the dephasing length measured in units of correlation length. In the following we prove that, based on the relative values of σ and p, the product in Eq. (39) converges for $N \to \infty$ to a proper subgroup of the group of unitary matrices. To this end we recall the following well-known theorem:

Theorem 8.1 ([30,31]) Consider the group of unitary matrices U(d) of dimension d and a probability measure μ whose support is not contained in a proper subgroup of U(d). Then the measure μ^n , obtained by doing the convolution of μ with itself n times, converges in distribution as $n \to \infty$ to the uniform probability measure on U(d).

Specifically we study the following nine different regimes:

- *p* is fixed and
 - $\frac{\sigma}{p}$ is fixed: In this case, it is straightforward to prove, using 8.1, that the distribution of the product will converge to the uniform measure on the unitary group as $N \to \infty$.
 - $\frac{\sigma}{p}$ diverges: In this case also, one can prove using Theorem 8.1 that the distribution of the product will converge to the uniform measure on the unitary group as $N \to \infty$
 - $\frac{\sigma}{p}$ converges to zero: It is straightforward to prove that the matrix $\mathbf{B} + \sigma \mathbf{Q}_j$ converges to a diagonal matrix. Hence $\exp[(i/p)(\mathbf{B} + \sigma \mathbf{Q}_j)]$ will also be diagonal.
- *p* diverges
 - $\frac{\sigma}{p}$ is fixed: In this case, it is straightforward to prove using Theorem 8.1 that the distribution of the product will converge to the uniform measure on the unitary group as $N \to \infty$
 - $\frac{\sigma}{p}$ diverges: In this case the matrix $\mathbf{B} + \sigma \mathbf{Q}_j$ will behave more and more as a GUE matrix, and we can use Theorem 8.1 to conclude convergence to a full Haar matrix as $N \to \infty$.
 - $\frac{\sigma}{p}$ converges to zero: In this case the argument of the exponential goes to 0, hence the solution converges to the identity matrix.
- p converges to zero
 - $\frac{\sigma}{p}$ is fixed: This case will be analyzed in the rest of this appendix. We prove that $\exp[\frac{i}{n}(\mathbf{B} + \sigma \mathbf{Q}_i)]$ converges to a block diagonal Haar matrix.

- $\frac{\sigma}{p}$ diverges: In this case the matrix $\mathbf{B} + \sigma \mathbf{Q}_j$ will behave increasingly as a GUE matrix as $\frac{\sigma}{p}$ increases, and we can use Theorem 8.1 to conclude convergence to a full Haar matrix as $N \to \infty$.
- $\frac{\sigma}{p}$ converges to zero: We expect several different behaviors to be exhibited depending on the rate of convergence of σ to 0.

We now prove that when $p = \sigma$ and $\sigma \to 0$, the matrix $\exp[\frac{i}{p}(\mathbf{B} + \sigma \mathbf{Q}_j)]$ converges to a block diagonal matrix. To this end, we divide the individual entries of this matrix into three groups: *diagonal entries*, *off-diagonal entries* within a block, and *off-block entries*. The size of the blocks is dictated by the structure of the matrix **B**.

Off-block entries Let us suppose $\sigma \ll 1$. We can divide the eigenvalues of $\frac{1}{p}(\mathbf{B} + \sigma \mathbf{Q}_j)$ into two subsets: the eigenvalues S_{Δ} close to Δ and the eigenvalues $S_{-\Delta}$ close to $-\Delta$. Let $\mu \in S_{\Delta}$ (respectively $S_{-\Delta}$) and ν_{μ} the associated eigenvector. By writing down the associated eigenvalue eigenvector decomposition, we see that the last four (the first two) coordinates are of the order of σ . Since the off block entries of $\exp[\frac{i}{p}(\mathbf{B} + \sigma \mathbf{Q}_j)]$ are of the form $\sum_j e^{i\lambda_j} \nu_{\mu} \overline{\nu_{\mu'}}$ with μ and μ' in subsets attached to different eigenvalues, we conclude that *the magnitude of an off block entry is upper bounded by nr where n is the dimension of the matrix.* Hence, what remains to be shown is that the magnitude of the block entries does not vanish when $\sigma = p \rightarrow 0$.

Block entries We will show that the eigenvector matrix of $A = \frac{1}{p}((\mathbf{B} + \sigma \mathbf{Q}_j))$ is block diagonal in the limit $\sigma = p \rightarrow 0$, with non vanishing block entries. We prove this for $C = (1/p)^2[\sigma(\mathbf{B}\mathbf{Q}_j + \mathbf{Q}_j\mathbf{B}) + \sigma^2\mathbf{Q}_j^2]$ instead of *A*, as the two statements are equivalent. Indeed, *C* is A^2 minus a scalar matrix, which does not alter the eigenvectors, i.e., *C*, *A* and e^{iA} all have the same eigenvectors. Here, we use the fact the matrix *A* is Hermitian to be able to apply functional calculus to *A*.

Observe that the matrix $\mathbf{BQ}_j + \mathbf{Q}_j\mathbf{B}$ is block diagonal with 2 blocks of the desired dimensions. This is because of the particular structure of **B**. We will show that in the limit as $\sigma = p \rightarrow 0$, the eigenvectors of *C* converge to those of $\mathbf{BQ}_j + \mathbf{Q}_j\mathbf{B}$. This is the content of the following theorem, which bounds the Frobenius norm of the eigenvector matrix.

Theorem 8.2 ([32,33]) Let $M \in \mathbb{C}^{n \times n}$ be a Hermitian matrix with distinct eigenvalues and have spectral decomposition $M = U \wedge U^*$. Let $\tilde{M} = M + \delta M$ be a Hermitian matrix and

$$\varepsilon = rac{\|\delta M\|_F}{\min_{i\neq j}|\lambda_i - \lambda_j|}, \qquad lpha = rac{(1 + \sqrt{1 - 1/n})\|M\|_2}{\min_{i\neq j}|\lambda_i - \lambda_j|}.$$

If $\varepsilon(2\alpha + \sqrt{1+4\alpha^2}) \leq 1$, then \tilde{M} has a spectral decomposition such that

$$\|\tilde{U} - U\|_F \le \frac{\sqrt{2\varepsilon}}{[1 - 2\alpha\varepsilon + \sqrt{1 - 4\alpha\varepsilon - \varepsilon^2}]^{1/2}}.$$
(40)

We apply this theorem with $M = \frac{1}{p}(\mathbf{B}\mathbf{Q}_j + \mathbf{Q}_j\mathbf{B})$ and $\delta M = \mathbf{Q}_j^2$. Hence U is the eigenvector matrix of $\frac{1}{p}(\mathbf{B}\mathbf{Q}_j + \mathbf{Q}_j\mathbf{B})$ and \tilde{U} the eigenvector matrix of C. Fix p and observe that the eigenvalues of $\frac{1}{p}(\mathbf{B}\mathbf{Q}_j + \mathbf{B}\mathbf{Q}_j)$ are distinct with probability 1 since the density of the eigenvalue distribution is continuous. As $p \to 0$, $\min_{i \neq j} |\lambda_i - \lambda_j|$ and $||A||_2$ grow as 1/p. Using this, a straightforward computation shows that $\varepsilon \to 0$ and $\alpha \varepsilon \leq 1$ for p small enough, and the theorem is satisfied. Since the right side of Eq. (40) decreases to 0, the Frobenius norm also converges to 0. This implies that the entries of \tilde{U} converge to those of U. Since \mathbf{Q}_j is a GUE matrix, we

can say that the blocks in the eigenvector matrix of $\mathbf{BQ}_j + \mathbf{Q}_j \mathbf{B}$ (and therefore also those of *C*) will be Haar distributed.

Having this result in hand, we can show that e^{iA} is block diagonal, which is the result we seek. All off-block entries converge to 0 as stated earlier. Consider a block entry $(e^{iA})_{lm}$. The eigenvalue/eigenvector decomposition of A gives that this entry is of the form $\sum_k e^{i\lambda_k} v_k w_k$ where v and w are respectively the l^{th} and m^{th} eigenvectors of A. Observe that the only nonzero entries of v_k and w_k are those corresponding to the block under consideration. By a standard result in perturbation theory (**B** and **Q**_j being hermitian, see [34]), the eigenvalues of $\frac{1}{p}(\mathbf{B} + \sigma \mathbf{Q}_j)$ are close to those of \mathbf{B}/p , and the distance between the corresponding eigenvalues is of the order of σ/p . Since $\sigma/p = 1$, we conclude that the eigenvalues λ_k are distinct modulo 2π with probability 1. Hence, the sum $(e^{iA})_{lm} = \sum_k e^{i\lambda_k} v_k w_k$ is non zero. Putting all these results together, we conclude that in the limit $p = \sigma \rightarrow 0$, e^{iA} is block diagonal. Since multiplication respects the block structure, we use Theorem 8.1 to conclude that the blocks of the matrix will be Haar distributed in the limit as $N \rightarrow \infty$.

We show examples of the shape of the matrix **T** in Eq. (39) by varying parameters p, σ and N. In all numerical experiments, we consider an ensemble with 1000 members, each representing a random draw of the matrix T, and compute the empirical expected value of the absolute value of each entry. Figure 7 shows the gray-coded values of the matrix elements of **T** using N = 1and $\sigma = p$ for values of σ ranging from 10^{-1} to 10^{-4} . We observe two matrix blocks of sizes 2×2 and a 4×4 , and a strong diagonal over the entire 6×6 matrix. The bottom row shows the situation for N = 10. We clearly see two blocks; the distribution of the elements in each block become much more uniform for N = 10 compared to the N = 1 case. These numerical results show behavior predicted by the asymptotic analysis in this appendix.

Finally, Figure 8 shows an example where σ is fixed and p decreases using N = 10. In this



Fig. 7. Gray-coded values of the elements of the matrix **T** with white and black representing 0 and 1, respectively. The values of *p* with $p = \sigma$ are shown above each plot. Top row N = 1; bottom row N = 10.



Fig. 8. Gray-coded values of the elements of the matrix **T** for N = 10 when $\sigma = 1$ is fixed and *p* decreases from 0.1 to 10^{-4} .

situation, we do not obtain a block structure, and the matrix has its full dimension. As $N \rightarrow \infty$, the product in this case converged to a uniform distribution over the whole matrix.

Acknowledgments

This work was supported in part by Alcatel–Lucent in the framework of GreenTouch. We acknowledge an anonymous reviewer for the insightful comments.