Check for updates

Nonlinear interference noise in space-division multiplexed transmission through optical fibers

CRISTIAN ANTONELLI,[1,*] ORI GOLANI,[2] MARK SHTAIF,[2] AND ANTONIO MECOZZI[1]

[1] *Department of Physical and Chemical Sciences, University of L'Aquila, 67100 L'Aquila, Italy*
[2] *School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel 69978*
*cristian.antonelli@univaq.it

Abstract: We study the nonlinear interference noise (NLIN) generated in SDM systems, and generalize the NLIN model introduced in the context of single-mode fibers to the multi-mode case. The generalized model accounts for the modulation-format dependence of the NLIN, and gives the scaling of the NLIN power with the number of transmitted modes. It also provides the tools for extending the results of the NLIN wizard to SDM. Unlike in the case of single-mode systems, the effect of MD cannot in general be ignored in the SDM case. We show that inclusion of MD erases the contribution of FWM effects, and significantly suppresses the effect of XPM.

References and links

1. H. Chen, R. Ryf, N. K. Fontaine, A. M. Velazquez-Benitez, J. Antonio-Lopez, C. Jin, B. Huang, M. Bigot-Astruc, D. Molin, F. Achten, and P. Sillard, "High Spectral Efficiency Mode-Multiplexed Transmission over 87-km 10-Mode Fiber," OFC 2016, paper Th4C.2 (2016).
2. N. K. Fontaine, R. Ryf, H. Chen, A. V. Benitez, B. Guan, R. Scott, B. Ercan, S. J. B. Yoo, L. E. Gruner-Nielsen, Y. Sun, R. Lingle, E. Antonio-Lopez, and R. Amezcua-Correa, "30×30 MIMO Transmission over 15 Spatial Modes," OFC2015, paper PDP Th5C.1 (2015).
3. P. Poggiolini, "The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems," J. Lightwave Technol. **30**, 3857–3879 (2012).
4. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," Opt. Express **21**, 25685–25699 (2013).
5. P. Serena and A. Bononi, "An Alternative Approach to the Gaussian Noise Model and its System Implications," J. Lightwave Technol. **31**, 3489–3499 (2013).
6. M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulations," IEEE Photon. Technol. Lett. **24**, 2016–2019 (2012).
7. P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," J. Lightwave Technol. **31**, 1273–1282 (2013).
8. O. Golani, R. Dar, M. Feder, A. Mecozzi, and M. Shtaif "Modeling the bit-error-rate performance of nonlinear fiber-optic systems," J. Lightwave Technol. **34**, 3482–3489 (2016).
9. A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," Opt. Express **22**, 16335–16362 (2014).
10. G. Rademacher and K. Petermann, "Nonlinear Gaussian Noise Model for Multimode Fibers With Space-Division Multiplexing," J. Lightwave Technol. **34**, 2280–2287 (2016).
11. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in fiber-optic systems," Opt. Express **22**, 14199–14211 (2014).
12. R. Ryf, R.-J. Essiambre, A. H. Gnauck, S. Randel, M. A. Mestre, C. Schmidt, P. J. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, T. Hayashi, T. Taru, and T. Sasaki, "SDM transmission over 4200-km 3-core microstructured fiber," OFC 2011, Paper PDP5C.2 (2011).
13. R. Ryf, N.K. Fontaine, B. Guan, R.-J. Essiambre, S. Randel, A. H. Gnauck, S. Chandrasekhar, A. Adamiecki, G. Raybon, B. Ercan, R. P. Scott, S.J. Ben Yoo, T. Hayashi, T. Nagashima, and T. Sasaki, "1705-km Transmission over Coupled-Core Fibre Supporting 6 Spatial Modes," ECOC 2014, PD.3.2 (2014).
14. C. Antonelli, A. Mecozzi, and M. Shtaif, "Nonlinear propagation in Space-Division Multiplexed fiber-optic transmission," IEEE Photonics Conference 2015, Paper MG2.1 (2015).
15. M. Brehler, D. Ronnenberg, and P. M. Krummrich, "Scaling of Nonlinear Effects in Multimode Fibers with the Number of Propagating Modes," OFC 2016, Paper W4I.3 (2016).

16. C. Antonelli, M. Shtaif and A. Mecozzi, "Modeling of Nonlinear Propagation in Space-Division Multiplexed Fiber-Optic Transmission," J. Lightwave Technol. **34**, 36–54 (2016).
17. In the regime of high MD, FWM is negligible. Other terms, that are related to the nonlinear spectral broadening of the nearest neighboring channels are also negligible in the high MD case.
18. NLIN wizard website at http://nlinwizard.eng.tau.ac.il/
19. 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).
20. D. Gloge, "Weakly guiding fibers," Appl. Opt. **10**, 2252–2258 (1971).
21. A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multi-mode fibers in the strong coupling regime," Opt. Express **20**, 11673–11678 (2012).
22. 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).
23. C. Antonelli, A. Mecozzi, and M. Shtaif, "The delay spread in fibers for SDM transmission: dependence on fiber parameters and perturbations," Opt. Express **23**, 2196–2202 (2015).
24. C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Stokes-space analysis of modal dispersion in fibers with multiple mode transmission," Opt. Express **20**, 11718–11733 (2012).
25. Here we are referring to spectral broadening of the nearest ICs that is caused by either self-phase modulation, cross-phase modulation with any of the other channels, or four-wave mixing contributions, whose central frequency coincides with that of the nearest ICs, and whose spectrum leaks into the bandwidth of the COI.
26. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-Channel Nonlinear Interference Noise in WDM Systems: Modeling and Mitigation," J. Lightwave Technol. **33**, 1044-1053 (2015).
27. We stress that since our goal is to validate the above presented analysis against numerical solutions, the particular choice of $\gamma \kappa_{aa}$ is of minor importance.
28. C. Antonelli, A. Mecozzi, and M. Shtaif, "Scaling of inter-channel nonlinear interference noise and capacity with the number of strongly coupled modes in SDM systems," OFC 2016, Paper W4I.2 (2016).
29. S. Mumtaz, R.-J. Essiambre, G. P. Agrawal, "Reduction of Nonlinear Penalties Due to Linear Coupling in Multicore Optical Fibers," Photon. Technol. Letters **24**, 1574-1576 (2012).
30. C. Antonelli, A. Mecozzi, O. Golani, and M. Shtaif, "Inter-modal nonlinear interference in SDM systems and its impact on information capacity," IEEE Summer Topicals 2016, Paper ME1.1 (2016).
31. G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. (Elsevier, 2006).
32. E. Temprana, E. Myslivets, B.P.-P. Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," Science **348**, 1445–1448 (2015).
33. The derivation of (45) makes use of the property $\sum_k X_{h,k,k} = 0$ for $h \neq 0$. It can be shown that the same property holds for $\sum_k Y_{h,k,k}$.

1. Introduction

Space-division multiplexed transmission on multi-mode and multi-core fibers is viewed as a promising approach for scaling the capacity of fiber communication systems. Numerous experimental demonstrations indicate the feasibility of the involved technology (record experiments have reported spectral efficiencies as high as 58 bits/second/Hertz [1] over fibers supporting up to 15 spatial modes [2]), thus making space-division multiplexing (SDM) a leading candidate for the future transport network. Similarly to the case of single-mode fiber based systems, the ultimate limit to the fiber information throughput is set by the fiber's nonlinearity. In the absence of full knowledge of the information transmitted through the fiber (in all channels), the nonlinear distortions must be treated as noise that penalizes transmission. In what follows we refer to this noise as nonlinear interference noise, or NLIN. In the context of single-mode systems it has been shown that a good estimate of the effects of nonlinearity on performance can be obtained based on the second-order statistics of the NLIN, i.e. the variances and correlations of the various noise components [3–7], and secondly their temporal correlation properties [8]. A significant amount of work on this topic has been carried out in the context of single-mode fiber (SMF) systems, cultivating into the well known Gaussian Noise (GN) model [3], and its extensions [4,9]. One attempt to characterize the NLIN in SDM fibers has been recently reported by Rademacher and Petermann [10], who essentially extended the formulation of the basic GN model [3] to the multi-mode-case. The limitations of this approach are in the fact that the GN model does not account for modulation format dependence and tends to be inaccurate in the case of short links [9, 11]. In addition, it does not account for the presence of modal dispersion (MD) within

quasi-degenerate mode-groups [12, 13], whose impact on the NLIN can be substantial, as has been pointed out in [14]. A purely numerical characterization of the NLIN power and its scaling with the number of modes was reported in [15], but in that study as well the effect of intra-group MD was not taken into account. In [16] the variance of the NLIN induced by inter-channel nonlinearities in SDM fibers was evaluated for the case of a single group of quasi-degenerate modes in the high MD regime, where cross-phase-modulation (XPM) is the dominant inter-channel nonlinearity [17].That work extended the analysis of [4], where the modulation format dependence is rigorously taken into account.

In this paper we conduct a comprehensive study of the NLIN power in SDM systems where the information is encoded into multiple groups of quasi-degenerate modes, and characterize all the relevant nonlinear processes. The major outcome of our study is a set of expressions for the scaling of the NLIN power with the number of modes. These expressions allow the extraction of the powers of the NLIN contributions from well established results that were derived and validated in the context of single-mode transmission [9, 18].We note that, given the assumptions on the SDM fiber characteristics that will be discussed in detail at a later stage, the derived expressions are exact within the same perturbation approach underpinning the treatment of the fiber Kerr nonlinearity introduced in the study of single-mode transmission [4].

The paper is organized as follows. In Section 3 we introduce the formalism that is necessary for the subsequent study of the NLIN, and review the coupled Manakov equations, which are the starting point of the analysis. In Section 4 we derive the expressions for the powers of the NLIN produced by intra-group nonlinear processes. These are self-phase modulation (SPM), XPM, and FWM involving either three, or four wavelength-division multiplexed (WDM) channels. In the same section we discuss the beneficial effect of intra-group MD, which is shown to notably reduce the contribution of XPM to the NLIN power, and to produce de-phasing that eliminates the contribution of FWM altogether. Section 5 is devoted to the study of NLIN produced by inter-group interactions. These are shown to be of two types, one is XPM, and another is non-degenerate FWM. We derive the expressions of the NLIN powers for the two interactions, and show that intra-group MD does not affect cross-group XPM, whereas it averages out cross-group FWM. The related phase-matching conditions are discussed. The paper ends with an appendix. In the first part of the appendix, we present a detailed derivation of the NLIN power expressions discussed in Sections 4 and 5. These expressions involve coefficients consisting of infinite sums of triple integrals. In the second part of the appendix, we express these coefficients in a form that is suitable for efficient numerical evaluation.

2. A note on terminology

The terms SPM, XPM, and FWM, which will be intensively used in what follows, need some clarification. We consider an SDM system where different mode groups do not couple strongly with one another, and hence they are received and processed separately. Conversely, all modes within the same mode group couple strongly, and hence they need to be detected and processed jointly, just like the two field polarizations in a single-mode system. Consistently with this notion, the signals transmitted in the various modes of the same group and at the same central wavelength are considered a single WDM channel, so that the total number of WDM channels transmitted in an SDM system equals the number of WDM wavelengths multiplied by the number of mode groups. We refer to one of the WDM channels as the channel of interest (COI), whose performance degradation due to NLIN is to be evaluated. Each of the other WDM channels is referred to as an interfering channel (IC). With these notions in mind, we use the terms SPM, XPM, and FWM as follows. SPM is a process through which the COI interferes nonlinearly with itself. XPM is a process where a single IC imposes a nonlinear interference on the COI. Finally, FWM is a process in which either two or three ICs impose a nonlinear interference on the COI.

3. Formalism and propagation equations

For the simplicity of notation and with no loss of generality, we consider propagation of two quasi-degenerate mode groups, which we denote as group a and group b. In this work we assume that the linear coupling between the mode groups is small, as should be the case in short-to-medium reach links, in which case its contribution to the nonlinear distortion is negligible. When this assumption is not satisfied, meaning that the inter-group coupling is large, then its inclusion in the analysis is per se a problem that requires a dedicated effort, which is left for future work. In any case, the inclusion of linear coupling is straightforward for the purpose of performing numerical studies [16, 19]. The number of modes in each group is $2N_a$ and $2N_b$, where the factor of two accounts for the existence of two degenerate orthogonal polarization modes in each spatial mode (as appropriate in the weakly guiding approximation [20]). The complex envelopes of the electric fields in the modes of groups a and b are stacked into two column vectors $\vec{E}_a$ and $\vec{E}_b$, containing $2N_a$ and $2N_b$ components, respectively. The evolution of the electric field vectors along the SDM fiber obeys the coupled multi-component Manakov equations [21, 22], supplemented with the terms accounting for MD within the groups (this type of MD should not be confused with the fact that the group velocities of two mode groups are different. On the contrary, it originates from the fact that what we refer to as *degenerate modes* consist of true modes characterized by somewhat different group velocities propagating in the regime of strong random coupling [23]),

$$\frac{\partial \vec{E}_a}{\partial z} = -\frac{\alpha}{2}\vec{E}_a + i\beta_a\vec{E}_a - \beta_a'\frac{\partial \vec{E}_a}{\partial t} - \mathbf{B}_a\frac{\partial \vec{E}_a}{\partial t} - i\frac{\beta_a''}{2}\frac{\partial^2 \vec{E}_a}{\partial t^2} + i\gamma\left(\kappa_{aa}|\vec{E}_a|^2 + \kappa_{ab}|\vec{E}_b|^2\right)\vec{E}_a \quad (1)$$

$$\frac{\partial \vec{E}_b}{\partial z} = -\frac{\alpha}{2}\vec{E}_b + i\beta_b\vec{E}_b - \beta_b'\frac{\partial \vec{E}_b}{\partial t} - \mathbf{B}_b\frac{\partial \vec{E}_b}{\partial t} - i\frac{\beta_b''}{2}\frac{\partial^2 \vec{E}_b}{\partial t^2} + i\gamma\left(\kappa_{ba}|\vec{E}_a|^2 + \kappa_{bb}|\vec{E}_b|^2\right)\vec{E}_b \quad (2)$$

A detailed derivation of Eqs. (1) and (2) from Maxwell's equations can be found in [16]. Here α is the fiber loss coefficient (in this work we neglect mode-dependent loss), β_a and β_b are the propagation constants of the two groups, β_a' and β_b' are the inverse group velocities, β_a'' and β_b'' are the chromatic dispersion coefficients. The terms $\mathbf{B}_a(z)$ and $\mathbf{B}_b(z)$ are $2N_a \times 2N_a$ and $2N_b \times 2N_b$ traceless matrices, respectively, which describe the local generalized birefringence within the two groups [24]. By γ we denote the usual nonlinearity coefficient appearing in the scalar nonlinear Schrödinger equation (NLSE), namely $\gamma = n_2\omega_0/cA_{\text{eff}}$, where n_2 is the glass nonlinearity coefficient, ω_0 is the center frequency of the optical signal, c is the speed of light in vacuum, and A_{eff} is the effective area of the fundamental mode. The coefficients κ_{aa}, κ_{ab}, κ_{ba}, and κ_{bb} are given by the following expression

$$\kappa_{uv} = \sum_{k,m}\sum_{j\in u}\sum_{h\in v} C_{jhkm}\frac{\delta_{hk}\delta_{jm} + \delta_{hm}\delta_{jk}}{(2N_u)(2N_v + \delta_{vu})}, \quad (3)$$

where δ_{uv} is the Kronecker delta function, and each of the indices u and v takes the values a and b depending on which of the coefficients is evaluated. For example, if $u = a$ and $v = b$, then in the summation the index j runs over all values corresponding to the modes in group a and the index h runs over all the indices corresponding to the modes in group b. The indices k and m in the summation span all the modes. The coefficients C_{jhkm} involve overlap integrals between the various mode profiles and their expressions can be found in [22]. Owing to the symmetry $C_{jhkm} = C_{hjkm}$ that characterizes them, the coefficients that nonlinearly couple the two groups are identical, $\kappa_{uv} = \kappa_{vu} = \mu$. Finally, it should be stressed that both γ and κ_{uv} depend on the number of modes, although this dependence is omitted in our notation for the sake of simplicity.

In order to describe the scaling, in what follows we also express the scalar NLSE,

$$\frac{\partial E}{\partial z} = -\frac{\alpha}{2}E - i\frac{\beta''}{2}\frac{\partial^2 E}{\partial t^2} + i\gamma|E|^2, \quad (4)$$

as well as the Manakov equation describing propagation in single-mode fibers,

$$\frac{\partial \vec{E}}{\partial z} = -\frac{\alpha}{2}\vec{E} - i\frac{\beta''}{2}\frac{\partial^2 \vec{E}}{\partial t^2} + i\frac{8}{9}\gamma|\vec{E}|^2. \tag{5}$$

so as to facilitate the comparison between the multi-mode and the scalar/single-mode cases in what follows. In the NLIN analysis, we consider a WDM signal, which we express as

$$\vec{E}_a = \sum_k \vec{E}_{a,k} e^{-ik\Omega t} \tag{6}$$

$$\vec{E}_b = \sum_k \vec{E}_{b,k} e^{-ik\Omega t} \tag{7}$$

where $\Omega/2\pi$ is the frequency spacing between adjacent channels. In what follows, we will assume that the vector $\vec{E}_{a,0}$ corresponds to the COI, whereas the vectors $\vec{E}_{a,k\neq 0}$ and $\vec{E}_{b,k}$ describe the ICs. By substituting Eqs. (6) and (7) into Eq. (1), the nonlinear terms that fall into group a and are within the bandwidth of the COI can be expressed as

$$i\gamma\kappa_{aa}\left(\vec{E}_{a,l}^{\dagger}\vec{E}_{a,m}\vec{E}_{a,l-m} + \eta_{l,m}\vec{E}_{a,l}^{\dagger}\vec{E}_{a,l-m}\vec{E}_{a,m}\right), \tag{8}$$

$$i\gamma\kappa_{ab}\vec{E}_{b,l}^{\dagger}\vec{E}_{b,m}\vec{E}_{a,l-m}, \tag{9}$$

where by the dagger we denote hermitian conjugation (that is, transposition and conjugation). The coefficient $\eta_{l,m}$ is 0 when $l = 2m$ (which includes the case $l = m = 0$), and it is equal to 1 in all other cases. Equation (8) represents intra-group nonlinear interference, whereas Eq. (9) represents the nonlinear interference between the two groups. Note that these terms do not account for the interference caused by the nonlinear spectral broadening of the nearest ICs [25].

   In the following section we start by characterizing intra-group interference, and proceed to inter-group interference in the subsequent section. Throughout this paper, we assume that the various space and polarization modes are modulated with statistically independent and identically distributed data streams. We note, however, that the assumption of identical distribution can easily be removed, at the expenses of the simplicity of the final results. Another important assumption concerns the effect of MD. In order to be able to analytically account for this phenomenon in the context of nonlinear interference, we assume that MD is sufficiently large to imply that different WDM channels undergo independent mode coupling processes, while at the same time the MD-induced distortion of the individual channels is negligible. For reference, we also consider the limit in which MD is negligible altogether, a situation of practical relevance for two-fold degenerate mode groups. Concerning the role of the modal dispersion terms proportional to the matrices $\mathbf{B}_a$ and $\mathbf{B}_b$ in Eqs. (1) and (2), we note that MD is assumed not to imply intra-channel distortions, which corresponds to setting $\mathbf{B}_a$ and $\mathbf{B}_b$ to zero in the analysis of the intra-channel dynamics. As for the inter-channel dynamics, we stress that the analysis focuses on the two limiting cases where MD is either absent or large, where the term *large* is used by assuming the spacing between adjacent WDM channels as a reference, so that MD is considered to be large when the MD bandwidth [24] is smaller than the channel spacing, but larger than the channel bandwidth. In the former case the terms $\mathbf{B}_a$ and $\mathbf{B}_b$ disappear, whereas in the second case their effect is to render the orientations of the hyper-polarizations of different WDM channels belonging to the same mode group mutually uncorrelated. The more general case where MD implies intra-channel distortions goes beyond the scope of this paper and is left for future work.

4. Intra-group nonlinear interference

The nonlinear interactions that are addressed in this section are sketched in Fig. 1. A detailed description of the various interactions is provided in the respective subsections below.
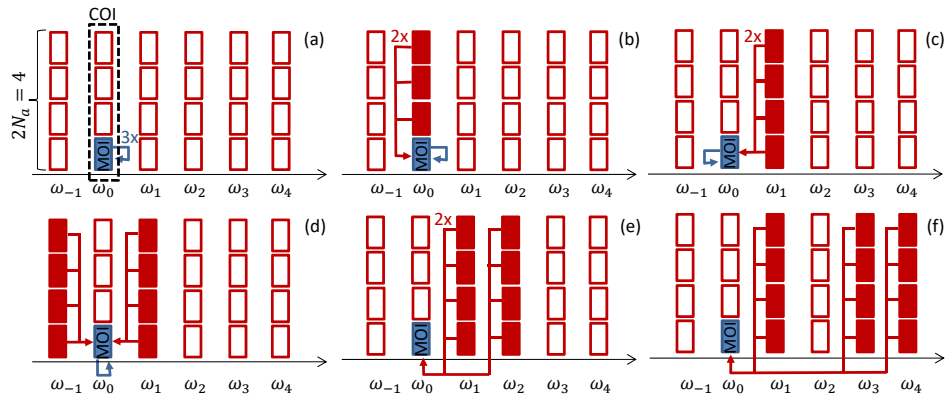
Fig. 1. Intra-group nonlinear interference processes analyzed in Section 4. (a) Contribution to self-phase modulation (SPM) from the mode of interest (MOI) itself (that is, the mode of the COI in which the nonlinear distortion is to be evaluated), and (b) from the other modes of the channel of interest (COI). (c) Cross-phase modulation (XPM). (d) Three-channel interaction of type I (3ch-I). (e) Three-channel interaction of type II (3ch-II). (f) Four-channel interaction (4ch). This illustration refers to the case of a groups consisting of four degenerate modes. In each of the interference processes represented here, the interfering modes are denoted by red filled rectangles. The red empty rectangles represent modes that are not involved. We denote by a blue filled rectangle the mode of interest. The symbols $2\times$ and $3\times$ are used to indicate that the nonlinear interaction is degenerate with respect to the channels to which they refer.

4.1. Intra-channel interaction

The first case that we consider is the one where $l = m = 0$. As will be seen in what follows, it is convenient to look at one component of the vector in Eq. (8), which can be expressed as follows,

$$i\gamma\kappa_{aa}\vec{E}_{a,0}^{\dagger}\vec{E}_{a,0}E_{a,0,j} = i\gamma\kappa_{aa}\left|E_{a,0,j}\right|^2 E_{a,0,j} + i\gamma\kappa_{aa}\sum_{k\neq j}^{2N_a}\left|E_{a,0,k}\right|^2 E_{a,0,j}. \qquad (10)$$

where by $E_{a,0,j}$ we denote the j-th component of vector $\vec{E}_{a,0}$. The first term at the right-hand side of the equation is identical to the self-phase modulation term that one would encounter in the analysis of the scalar NLSE (4), provided that γ is replaced by $\gamma\kappa_{aa}$. Each of the remaining $(2N_a - 1)$ terms is equivalent to the term that accounts for cross-polarization phase modulation encountered in the single-mode case, as follows from Eq. (5), where $8\gamma/9$ is to be replaced by $\gamma\kappa_{aa}$. In addition, the terms in the summation are statistically independent of each other, and hence their contributions to the NLIN power in the j-th mode are additive. For the sake of simplicity we refer to all the terms appearing in Eq. (10) as SPM, although this definition should be intended as a generalization of the standard definition of SPM. Based on the observations described above, the variance of the NLIN contribution due to intra-group SPM in each of the modes belonging to group a can be expressed as

$$\sigma_{\mathrm{SPM}}^2 = \sigma_{\mathrm{SPM,scalar}}^2 + (2N_a - 1)\left(\sigma_{\mathrm{SPM,1}}^2 - \sigma_{\mathrm{SPM,scalar}}^2\right), \qquad (11)$$

where $\sigma_{\mathrm{SPM,scalar}}^2$ is the NLIN contribution that would characterize SPM in the scalar case, and $\sigma_{\mathrm{SPM,1}}^2$ is the contribution of the generalized SPM to NLIN in a single-mode fiber governed by the standard Manakov Equation (5). The relevance of representing the result in this form follows from the fact that $\sigma_{\mathrm{SPM,scalar}}^2$ and $\sigma_{\mathrm{SPM,1}}^2$ are easily obtained from the literature relating

to NLIN in single-mode fibers, and in particular they can be readily extracted from the NLIN Wizard available in [18]. For completeness we provide their explicit expressions in this paper's appendix, since they do not seem to be available in previous related publications. We also remind the reader that $\sigma^2_{\text{SPM,scalar}}$ and $\sigma^2_{\text{SPM,1}}$ depend on the number of modes through the coefficient κ_{aa} and through A_{eff}, which determines γ.

4.2. Two-channel interaction

The second case that we consider is the one where $l = m$, with $l, m \neq 0$. This term describes XPM and hyper-polarization rotation (the latter coincides with the rotation of the standard polarization state vector in the single-mode case). In this case, the sum of the two contributions from Eq. (8) can be expressed as

$$i\gamma\kappa_{aa} \left(|\vec{E}_{a,l}|^2 \mathbf{I}_a + \vec{E}_{a,l}\vec{E}^\dagger_{a,l} \right) \vec{E}_{a,0}, \tag{12}$$

where $\mathbf{I}_a$ is the $2N_a \times 2N_a$ identity matrix.

In the limit of large MD, the nonlinear interference vector can be averaged with respect to the relative orientation of the two state vectors. In other words, the nonlinear interference vector is to be averaged with respect to the orientation of $\vec{E}_{a,l}$, in a reference frame rotating with the orientation of $\vec{E}_{a,0}$. As a result of this procedure (which has also been discussed in [16]), Eq. (12) simplifies to

$$i\gamma\kappa_{aa} \left(1 + \frac{1}{2N_a} \right) |\vec{E}_{a,l}|^2 \vec{E}_{a,0}. \tag{13}$$

This expression is obtained by averaging the matrix $\vec{E}_{a,l}\vec{E}^\dagger_{a,l}$ with respect to the isotropically distributed orientation of the state vector $\vec{E}_{a,l}$, which yields $\mathbf{I}_a |\vec{E}_{a,l}|^2/2N$. The form of Eq. (13) indicates that the terms considered here contribute to NLIN through XPM. It is now convenient to expand the expression of the j-th component of the vector in Eq. (13) as

$$i\gamma\kappa_{aa} \left(1 + \frac{1}{2N_a} \right) |\vec{E}_{a,l}|^2 E_{a,0,j} = i\gamma\kappa_{aa} \left(1 + \frac{1}{2N_a} \right) \sum_k \left| E_{a,l,k} \right|^2 E_{a,0,j}. \tag{14}$$

This form shows that the modes at the frequency $l\Omega$ provide $2N_a$ statistically independent contributions to NLIN, and hence the corresponding variances can be added to each other. The XPM-induced NLIN variance can in this case be expressed as

$$\sigma^2_{\text{XPM}} = 2N_a \left(1 + \frac{1}{2N_a} \right)^2 \frac{\sigma^2_{\text{XPM,scalar}}}{4} = \frac{(2N_a + 1)^2}{8N_a} \sigma^2_{\text{XPM,scalar}}. \tag{15}$$

where by $\sigma^2_{\text{XPM,scalar}}$ we denote the NLIN variance produced by a single WDM interferer in the scalar case described by Eq. (8) with $\gamma = \gamma\kappa_{aa}$, and where the factor of 1/4 accounts for the fact that the XPM term in the scalar NLSE has a factor of 2 in front. We note that Eq. (15) was first presented in [16].

In the opposite limit of small MD, the various WDM channels do not rotate relative to each other, and the contribution to the NLIN variance is calculated as follows. We first express explicitly the j-th component of the nonlinear interference vector as follows

$$2i\gamma\kappa_{aa}|E_{a,l,j}|^2 E_{a,0,j} + i\gamma\kappa_{aa} \sum_{k,k\neq j} |E_{a,l,k}|^2 E_{a,0,j} + i\gamma\kappa_{aa} \sum_{k,k\neq j} E_{a,l,j} E^*_{a,l,k} E_{a,0,k}. \tag{16}$$

The first term of Eq. (15) is identical to the XPM term given by the scalar NLSE (4), with $\gamma = \gamma\kappa_{aa}$. The second and third terms introduce $(2N - 1)$ statistically independent contributions, which are individually equivalent to the contribution of the second polarization in the single-mode

case. Based on this argument, the contribution of Eq. (15) to the NLIN variance can be expressed as

$$\sigma_{\text{XPM}}^2 = \sigma_{\text{XPM,scalar}}^2 + (2N_a - 1)\left(\sigma_{\text{XPM},1}^2 - \sigma_{\text{XPM,scalar}}^2\right), \tag{17}$$

where $\sigma_{\text{XPM},1}^2$ is the NLIN variance in each of the two polarizations of a single-mode fiber obeying the Manakov equation (5).

Using the results of [26], Eq. (15) and (17) can be expressed as

$$\frac{\sigma_{\text{XPM}}^2}{P^3} = \frac{(2N_a + 1)^2}{3N_a}\chi_1 + 2\frac{(2N_a + 1)^2}{5N_a}\chi_2\left(\frac{\langle|b|^4\rangle}{\langle|b|^2\rangle^2} - 2\right), \text{ large MD} \tag{18}$$

$$\frac{\sigma_{\text{XPM}}^2}{P^3} = 4\frac{2N_a + 1}{3}\chi_1 + 4\frac{2N_a + 3}{5}\chi_2\left(\frac{\langle|b|^4\rangle}{\langle|b|^2\rangle^2} - 2\right), \text{ negligible MD} \tag{19}$$

where P is the average signal power per WDM channel in each scalar mode, b stands for the constellation symbols used in each independent data stream (e.g. in the case of QPSK b receives the values $\pm 1 \pm i$ with equal probabilities), and angled brackets denote ensemble average. Using the terminology of [4, 26], the first term in the two expressions accounts for second order noise (SON), which coincides with the standard GN model [3, 10], whereas the second term is referred to as fourth-order noise (FON), and it accounts for the dependence on modulation format. The coefficients χ_1 and χ_2 are the SON and FON coefficients, respectively, and their expressions are given in the appendix of [26], except that the factor $8\gamma/9$ therein needs to be replaced by $\gamma\kappa_{aa}$. It is worth noting that the dependence of the NLIN power on the modulation format in the multi-mode caseis very similar to that seen in single-mode transmission. Indeed, the ratio between the FON and the SON terms is equal to 6/5 in the case of negligible MD, whereas it reduces from 1 for $N_a = 1$ to 3/5 for large mode counts, in the case of large MD. The situation is similar in the other nonlinear interference processes which are studied in what follows.

As will be discussed in more detail at a later stage, we note that MD yields a reduction of the SON coefficient. The ratio between the SON coefficients appearing in Eqs. (19) and (18) can be expressed as $2[1 - 1/(2N_a + 1)]$, which corresponds to a reduction by a factor of about 1.6 for $N_a = 2$, and approaches the value of 2, as the number of modes increases.

4.3. Three and four-channel interactions

Three and four-channel interactions involve either two or three different ICs whose nonlinear interaction has an effect on the COI. In the limit of large MD, the various channels undergo uncorrelated mode coupling processes and hence these interactions do not build up coherently, with the consequence that they can be safely neglected. It is only in the limit of low MD that the contribution of these terms needs to be considered. To the best of the authors' knowledge, this point has eluded previous work, where intra-group modal dispersion was neglected within the master model assumptions [10, 15].

We start by considering the case of three-channel interactions. There are two types of such interactions. One is degenerate with respect to the COI, and is of the form $i\gamma\kappa_{aa}\vec{E}_{a,0}^\dagger\vec{E}_{a,m}\vec{E}_{a,-m} + i\gamma\kappa_{aa}\vec{E}_{a,0}^\dagger\vec{E}_{a,-m}\vec{E}_{a,m}$, with $m \neq 0$. The other is degenerate with respect to one of the ICs, and is of the form $i\gamma\kappa_{aa}\vec{E}_{a,2m}^\dagger\vec{E}_{a,m}\vec{E}_{a,m}$, with $m \neq 0$. The j-th component of the nonlinear interference vector in the first case reads as

$$i\gamma\kappa_{aa}\left(2E_{a,0,j}^*E_{a,m,j}E_{a,-m,j} + \sum_{k,k\neq j}E_{a,0,k}^*E_{a,m,k}E_{a,-m,j} + E_{a,0,k}^*E_{a,-m,k}E_{a,m,j}\right), \tag{20}$$

and since all field triplets in the expansion are independent of each other and identically distrib-

uted, their contribution to the NLIN variance can be expressed as

$$\sigma^2_{\text{3ch-I}} = \frac{2N_a + 1}{2}\sigma^2_{\text{3ch-I,scalar}} \tag{21}$$

where $\sigma^2_{\text{3ch-I,scalar}}$ is the NLIN power due to the same type of three-channel interaction in the scalar case, as follows from Eq. (4) (in the scalar case this interaction is described by the first term inside the parentheses in Eq. (20)). In the second case, where $l = 2m$, the j-th component of the nonlinear interference vector reads as

$$i\gamma\kappa_{aa}E^*_{a,2m,j}E^2_{a,m,j} + i\gamma\kappa_{aa}\sum_{k,k\neq j}E^*_{a,2m,k}E_{a,m,k}E_{a,m,j}. \tag{22}$$

The first term in the expansion is the same as would follow from the scalar NLSE, whereas each of the other $(2N_a - 1)$ terms is equivalent to the second polarization in the single-mode fiber case described by the Manakov equation (5). Since they are all uncorrelated with each other, the overall NLIN power can be expressed as

$$\sigma^2_{\text{3ch-II}} = \sigma^2_{\text{3ch-II,scalar}} + (2N_a - 1)\left(\sigma^2_{\text{3ch-II,1}} - \sigma^2_{\text{3ch-II,scalar}}\right), \tag{23}$$

where $\sigma^2_{\text{3ch-II,scalar}}$ and $\sigma^2_{\text{3ch-II,1}}$ are the NLIN powers due to the same type of three-channel interaction in the scalar case and in the single-mode case, as follows from Eqs. (4) and (5), respectively. The explicit expressions for $\sigma^2_{\text{3ch-I,1}}$, $\sigma^2_{\text{3ch-II,1}}$, $\sigma^2_{\text{3ch-I,scalar}}$, and $\sigma^2_{\text{3ch-II,scalar}}$ have been first presented in [9], and they are also provided for convenience in the appendix.

Finally, we consider the contribution of four-channel interactions, which are accounted for by the nonlinear interference vector $i\gamma\kappa_{aa}\vec{E}^{\dagger}_{a,l}\vec{E}_{a,m}\vec{E}_{a,l-m} + i\gamma\kappa_{aa}\vec{E}^{\dagger}_{a,l}\vec{E}_{a,l-m}\vec{E}_{a,m}$, with $l \neq m$, $l \neq 2m$, and $l, m \neq 0$. Its j-th component is

$$i2\gamma\kappa_{aa}E^*_{a,l,j}E_{a,m,j}E_{a,l-m,j} + i\gamma\kappa_{aa}\sum_{k,k\neq j}E^*_{a,l,k}E_{a,m,k}E_{a,l-m,j} + E^*_{a,l,k}E_{a,l-m,k}E_{a,m,j}. \tag{24}$$

Here also all terms are identically distributed and uncorrelated with each other, and hence their contribution to the NLIN power reads as

$$\sigma^2_{\text{4ch}} = \frac{2N_a + 1}{2}\sigma^2_{\text{4ch,scalar}} \tag{25}$$

where by $\sigma^2_{\text{4ch,scalar}}$ we denote the corresponding contribution in the scalar case, as follows from the scalar NLSE (4). Its expression, first derived in [9], is also given in the appendix.

4.4. Numerical validation: intra-group nonlinear interference

We now proceed to validate the formulae derived in Section 4, which describe the scaling of the NLIN power with the number of strongly coupled modes. In this case the simulations are based on the numerical integration of Eq. (1), where the inter-group nonlinear coupling term $\gamma\kappa_{ab}|\vec{E}_b|^2\vec{E}_a$ is suppressed by setting $\kappa_{ab} = 0$. We considered a 5×100 km SDM system, with a fiber loss coefficient of 0.2 dB/km, and a nonlinearity coefficient equal to $\gamma\kappa_{aa} = 1.3/N_a$ W/km^{-1}, consistently [27]with the scaling discussed in [16, 28]. We transmitted a 16 QAM modulated signal using a square-root raised-cosine fundamental waveform with a roll-off factor of 0.01, and assumed coherent reception with a matched electrical filter. The signal power was set to -2 dBm per scalar (space and polarization) mode. Given the focus on nonlinear interference, amplification noise was not added to the propagating field. In Fig. 2 we plot the NLIN power as a function of the number of strongly coupled modes for various system settings, as discussed in what follows. All simulations were performed in the regime of negligible MD, which is the case
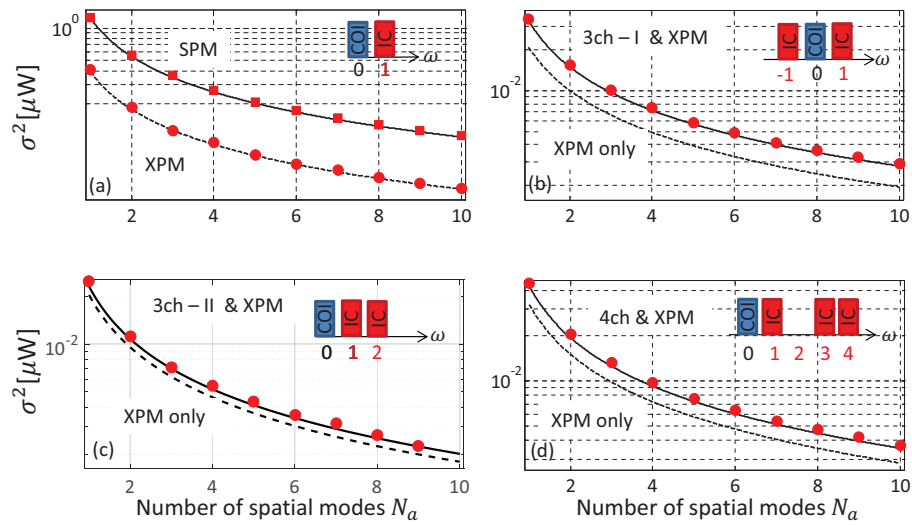
Fig. 2. The power of the intra-group NLIN versus the number of strongly coupled modes in a 5×100 km SDM system. Details on the simulation settings are provided in the main text. (a) Nonlinear interference noise power due to SPM (squares) and XPM (dots). Solid and dashed curves represent Eqs. (11) and (17), respectively. Panels (b)–(d) correspond to three- and four-channel interactions, and the NLIN power is plotted after removal of the SPM contribution. Panel (b) shows the NLIN due to three-channel interaction of type I, which was produced by transmitting the COI along with the two nearest channels. The solid curve was obtained by summing Eqs. (17) and (21), whereas the dashed curve shows only the XPM contribution (Eq. (17). Panel (c) shows the results obtained for the three-channel interaction of type II, produced by transmitting the COI along with the first and second neighboring channels. The solid curve was obtained by summing Eqs. (17) and (23), whereas the dashed curve shows the XPM contribution alone. Panel (d) refers to four-channel interaction, and was obtained by transmitting the first, third, and fourth neighboring channels along with the COI. The solid curve was obtained by summing Eqs. (17) and (25), whereas the dashed curve shows the XPM contribution alone. All plots were produced in the limit of negligible MD. In the presence of substantial MD, the contributions shown in (b)–(d) would vanish.

where theory needs to be validated numerically (in the opposite regime of large inter-channel MD, the scaling of XPM has already been tested numerically in [28] and [16], whereas FWM terms would simply vanish, as discussed in Section 4.3).

In Fig. 2(a) we plot the results obtained for SPM and XPM. We first extracted the SPM contribution by propagating the COI alone. Then, the XPM contribution was evaluated by propagating the COI along with the nearest IC ($l = m = 1$), and subtracting the SPM contribution from the overall NLIN power. In both runs the same set of 2^{15} pseudo-random symbols was used. The simulations of Fig. 2(a) were performed with a typical modulation rate of 32 Gbaud and a channel spacing of 50 GHz, and set the chromatic dispersion coefficient to the usual value $\beta_a'' = -21$ ps^2/nm/km. Squares and dots show the simulation results for SPM and XPM, respectively. Solid and dashed lines represent Eq. (11) and (17), respectively.

Figures 2(b) – 2(d) show the results for three- and four-channel interactions. In order to resolve the NLIN contribution of these interactions, it was necessary to reduce the baudrate and the dispersion coefficient to the extent that SPM and XPM are not dominant. Hence we assumed a baudrate of 3.2 GBaud, and set the chromatic dispersion coefficient to $\beta_a'' = -5$ ps^2/nm/km.
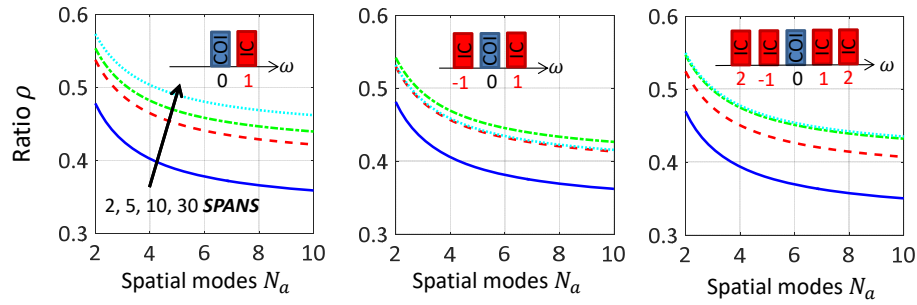
Fig. 3. Ratio ρ between the power values of the NLIN due two-channel interaction in the two cases of large and negligible MD, for the same system considered in Fig. 2(a). The four curves in each panel correspond to two, five, ten, and thirty spans. From left to right, the number of interfering channels is increased from one four, as shown in the insets.

In order to avoid the interference which is caused by the spectral broadening of the individual channels, we set the channel spacing to 6 GHz, so that the ratio between the channel separation and the baudrate is slightly larger than it is in Fig. 2(a). Figure 2(b) refers to three-channel interaction of type I, which was produced by transmitting the COI along with the two nearest ICs ($l = 0$, $m = 1$ in Eq. (8)). The NLIN power is plotted after removal of the SPM contribution. The solid curve was obtained by summing Eqs. (15) and (21), whereas the dashed curve shows only the XPM contribution (Eq. (17)). The appreciable difference between the two curves confirms that the three-channel interaction of type I was tested in a regime where its contribution to the NLIN power is substantial. Figure 2(c) shows the results obtained for the three-channel interaction of type II, produced by transmitting the COI along with the first and second neighboring channels ($l = 2$, $m = 1$ in Eq. (8)). The solid curve was obtained by summing Eqs. (15) and (23), whereas the dashed curve shows the XPM contribution alone. Finally, Fig. 2(d) refers to four-channel interaction, and was obtained by transmitting the first, third, and fourth neighboring channels along with the COI ($l = 4$, $m = 1$ in Eq. (8)). The agreement between simulation results and theory is excellent in all cases.

4.5. Effect of modal dispersion on intra-group NLIN

While the beneficial effect of mode coupling has been noted since early numerical studies [29], MD has been argued to reduce the strength of the nonlinear interference in SDM systems only recently [14, 16, 30]. This effect follows from the fact that MD imposes uncorrelated rotations of different WDM channels in the process of propagation. The immediate consequence of this reality is that three- and four-channel interactions are averaged out, as discussed in the previous section. In addition, as we explain below, MD also suppresses the NLIN caused by XPM. In order to quantify this effect, we define the parameter ρ as the ratio between the powers of the NLIN in the two cases of large and negligible MD, which are given in Eq. (15) and (17), respectively:

$$\rho = \frac{(2N_a + 1)^2}{8N_a} \frac{1}{1 + (2N_a - 1)\left(\sigma_{\text{XPM},1}^2 / \sigma_{\text{XPM,scalar}}^2 - 1\right)} \tag{26}$$

Interestingly, ρ is independent of the nonlinearity coefficient $\gamma\kappa_{aa}$, and the system parameters affect its value only through the ratio $\sigma_{\text{XPM},1}^2 / \sigma_{\text{XPM,scalar}}^2$, which can be conveniently computed by means of the NLIN Wizard [18]. In the limit of large mode count, $\rho \sim 0.25/(\sigma_{\text{XPM},1}^2 / \sigma_{\text{XPM,scalar}}^2 - 1)$.

Figure 3 shows a plot of ρ as a function of the number of modes N_a for the same system considered in Fig. 2(a). The four curves were obtained by increasing the number of spans from
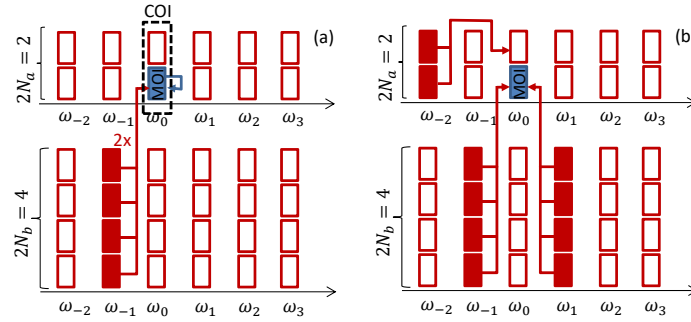
Fig. 4. Inter-group nonlinear interference processes. (a) Cross-group cross-phase modulation (XGXPM). (b) Cross-group four-wave mixing (XGFWM). This illustration refers to the case of one group consisting of two degenerate modes ($2N_a = 2$), and another consisting of four degenerate modes ($2N_b = 4$). The use of symbols is consistent with that adopted in Fig. 1.

two to thirty, whereas from the leftmost to the rightmost panel, the number of interfering channels was increased from one to four, as is sketched in the insets. The plots show that MD yields a substantial reduction in the NLIN power, implying that failure to account for it (e.g. as in [10, 15]) leads to exaggerated estimates of the nonlinear distortion due to inter-channel interactions.

5. Inter-group nonlinear interference

In this section we consider the nonlinear interference that channels propagating in mode group b impose on the COI, which is in mode group a, and whose central frequency is set to zero. We stress that in our assumptions the receiver processes separately the fields in group a and b, so that the COI receiver does not have access to the zero-frequency channel in group b. This assumption is consistent with the neglect of linear coupling between different mode groups.

Here we are concerned with nonlinear terms of the form $i\gamma\kappa_{ab}\vec{E}_{b,l}^\dagger \vec{E}_{b,m}\vec{E}_{a,l-m}$, and two cases need to be addressed separately. The first is $l = m$ and the second $l \neq m$. In the first case, which is illustrated in Fig. 4(a), the nonlinear interference vector can be expressed in the form

$$i\gamma\kappa_{ab} \sum_k \left|E_{b,l,k}\right|^2 \vec{E}_{a,0}, \tag{27}$$

where the terms in the summation are independent of each other when conditioning on the signal transmitted in the channel of interest. This conclusion is valid in both limits of small and large MD. The effect of these terms is identical to that of the the terms in Eq. (14), as they produce cross-group cross phase modulation (XGXPM), and their contribution to the NLIN variance is given by

$$\sigma^2_{\text{XGXPM}} = 2N_b \zeta^2_{\text{XPM,scalar}}. \tag{28}$$

The quantity $\zeta^2_{\text{XPM,scalar}}$ is very similar, yet not identical, to what we have denoted previously as $\sigma^2_{\text{XPM,scalar}}$ in Eq. (15). The difference is in how this quantity is evaluated from the scalar NLSE (4), while accounting for the difference in group velocities and chromatic dispersion between the mode groups. The expression for $\zeta^2_{\text{XPM,scalar}}$ is provided in the appendix. The dominant contribution to XGXPM comes from the WDM channel for which group velocity dispersion compensates for the group velocity mismatch between the two mode groups to the largest extent,

that is

$$\beta_b'' \, l\Omega = \beta_a' - \beta_b'. \tag{29}$$

We note that, when Eq. (29) is fulfilled, XGXPM could in principle prevail over intra-group XPM.

The second case that needs to be addressed is $l \neq m$, and it is illustrated in Fig. 4(b). The j-th component of the nonlinear interference vector in this case can be expressed in the form of the summation

$$i\gamma\kappa_{ab} \sum_k E_{b,l,k}^* E_{b,m,k} E_{a,l-m,j}, \tag{30}$$

which consists of four-channel FWM interaction terms. Similarly to the case of intra-group FWM, these contributions too vanish in the presence of substantial MD. Hence, in what follows we calculate what they would be in the absence of MD. To this end we note that all of the terms in (30) are uncorrelated with one another and with the COI, as well as with the terms in Eq. (27). Their contribution to the NLIN variance can hence be expressed as

$$\sigma_{\text{XGFWM}}^2 = 2N_b \zeta_{\text{4ch,scalar}}^2. \tag{31}$$

The quantity $\zeta_{\text{4ch,scalar}}^2$ is very similar to the quantity $\sigma_{\text{4ch,scalar}}^2$ in Eq. (25), except that its evaluation from the scalar NLSE also requires accounting for the difference in group velocities and chromatic dispersion between the mode groups, similarly to the calculation of $\zeta_{\text{XPM,scalar}}^2$. The expression for $\zeta_{\text{4ch,scalar}}^2$ is provided in the appendix. This type of nonlinear interference is most effective when the ICs fulfill the following phase matching condition,

$$\beta_a = \beta_a + \beta_a'(l-m)\Omega + \frac{\beta_a''}{2}(l-m)^2\,\Omega^2 - \left[\beta_b + \beta_b'\,l\Omega + \frac{\beta_b''}{2}l^2\,\Omega^2\right] + \beta_b + \beta_b'm\Omega + \frac{\beta_b''}{2}m^2\,\Omega^2, \tag{32}$$

whose physical meaning follows from the preservation of the total light momentum during the FWM process [31]. After some straightforward algebra, Eq. (32) can be simplified to the form

$$\beta_a' - \beta_b' = m\frac{\beta_a'' + \beta_b''}{2}\Omega - l\frac{\beta_a'' - \beta_b''}{2}\Omega, \tag{33}$$

which is a handy generalization of the phase matching condition given in Eq. (6) of [10] to the most customary case where the chromatic dispersion coefficients of the two mode groups are not identical. Equation (33) includes Eq. (29) for $m = l$.

5.1. Numerical validation: inter-group nonlinear interference

In this section we validate the formulae derived in Section 5. The transmission settings are identical to those used in Fig. 2(a), and in order to isolate the NLIN imposed by group b on group a, all undesired nonlinear interference processes were suppressed by setting $\kappa_{aa} = \kappa_{bb} = \kappa_{ba} = 0$. The only nonzero nonlinearity coefficient was $\gamma\kappa_{ab} = 1.3/N_b$ W/km^{-1} (here too, the inverse dependence on N_b is justified in [16]). The chromatic dispersion coefficients were set to $\beta_a'' = -10$ ps^2/km and $\beta_b'' = -25$ ps^2/km, while the inverse-group-velocity difference $\beta_a' - \beta_b'$ was set accordingly, so as to fulfill the phase-matching condition (33).

The contribution of XGXPM to the NLIN power was extracted by transmitting only the COI in group a, and two ICs in group b at $\pm\Omega$ ($l = -1$ and $m = 1$ in Eq. (9)). The results are shown in Fig. 5(a), where the power of the NLIN imposed by group b on group a is plotted versus the number of strongly coupled modes in group b. Markers refer to simulations, while the solid curve is the plot of Eq. (28).

The contribution of XGFWM to the NLIN power was computed for the same system configuration of Fig. 5(a), except that an additional IC at frequency -2Ω ($l - m = -2$ in Eq. (9)) was
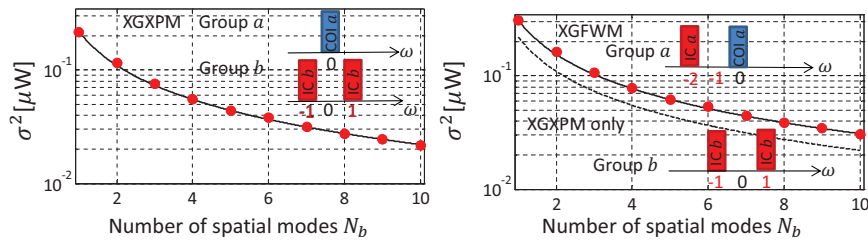
Fig. 5. The NLIN power versus the number of strongly coupled spatial modes in group b in a 5×100 km SDM system with two groups of quasi degenerate modes. Details on the simulation settings are provided in the main text. The left panel shows the NLIN power computed by transmitting the COI in mode group a and two interfering channels at $\pm\Omega$ in mode group b. The markers refer to the simulations, while the solid curve is a plot of Eq. (28). The right panel refers to the same system, which was modified by transmitting the additional WDM channel at -2Ω in mode group a, so as to produce XGFWM in the COI. Solid curve and markers in Fig. 5(b) show the total NLIN power, the dashed curve shows the XGXPM contribution alone. In both cases β'_a and β'_b were set so as to fulfill the phase matching condition of Eq. (33). All plots were produced in the limit of negligible MD. In the presence of substantial MD, the contributions shown in (b) would vanish.

also transmitted in group a along with the COI. The solid curve and markers in Fig. 5(b) show the total NLIN power, whereas the dashed curve shows the contribution of XGXPM only.

In both numerical examples considered in this section we assumed that mode group a consisted of a single spatial mode ($N_a = 1$).

6. Mitigation of the nonlinear distortions in SDM links

The nonlinear interference in SDM systems is strongly influenced by the mechanisms that are responsible for linear mode coupling. As has been extensively argued in [16], strong mode mixing implies a considerable reduction of the NLIN power. This effect follows from the fact that because of the strong mode mixing, the power transmitted in each spatial mode divides equally across the individual modes, with the consequence that the overall NLIN results from independent sources, each distributed over a wider effective area.

Intra-group MD yields an additional reduction of the NLIN power, by preventing the coherent build-up of the nonlinear distortions. This reduction follows from the fact that different spectral components of the SDM signal undergo uncorrelated rotations of their hyper-polarization states. The downside of this mechanism is that use of backward propagation techniques for removing the nonlinear noise becomes highly impractical, as it would require detailed knowledge of the linear SDM fiber channel characteristics across the entire signal spectrum at every position along the fiber. Nonetheless, the MD-induced NLIN reduction can be even more effective than backward propagation techniques whose implementation is conceivable in the foreseeable future. Moreover, the benefit of MD-induced NLIN reduction comes at essentially no cost (given that MD is unavoidable anyway). This is in sharp contrast to backward propagation, which requires detailed knowledge of the received signal in the entire bandwidth which is to be back-propagated, and is only effective when the WDM channels are mutually coherent [32].

Similar considerations apply to the nonlinear interaction between non-degenerate mode groups.

7. Conclusions

We evaluated the power of the nonlinear interference noise (NLIN) generated in SDM systems, focusing on the relevant situation where the various fiber modes can be classified into groups such that modes belonging to the same group are degenerate and strongly coupled, whereas modes in different groups are non-degenerate and their coupling is negligible. This work generalizes the model introduced in [4, 9] to the multi-mode case, and gives the scaling of the NLIN power with the number of transmitted modes. Similarly to the single-mode case, the generalized model accounts for the modulation-format dependence of the NLIN, and provides tools for extending the results of the NLIN Wizard [18] to SDM. Unlike in the case of single-mode systems, the effect of MD cannot in general be ignored in the SDM case. We show that inclusion of MD erases the contribution of FWM effects, and significantly suppresses the effect of XPM.

8. Appendix

This appendix is devoted to providing the expressions for the variances used in the main body of the paper. These are $\sigma^2_{\text{SPM,scalar}}$, $\sigma^2_{\text{SPM},1}$, $\sigma^2_{\text{XPM,scalar}}$, $\sigma^2_{\text{3ch}-\text{I,scalar}}$, $\sigma^2_{\text{3ch}-\text{II,scalar}}$, $\sigma^2_{\text{3ch}-\text{II},1}$, $\sigma^2_{\text{4ch,scalar}}$, ζ^2_{XPM}, and ζ^2_{FWM}. All quantities, except for ζ^2_{XPM} and ζ^2_{FWM} are available in the existing literature, and many of then can be conveniently extracted from the single-mode NLIN Wizard [18], as will be explained in what follows. We proceed by evaluating the quantities ζ^2_{XPM} and ζ^2_{FWM}, and will rely on their derivation to extract the others.

8.1. Derivation of the NLIN power expressions

Inter-group, or cross-group XPM (XGXPM) involves only a single IC in group b, whereas cross-group FWM (XGFWM) involves one IC in group a and two ICs in groups b. Using the indexes l and m as in Eq. (9) (meaning that group a contains the COI and the $(l-m)$-th IC, whereas group b contains the l-th and the m-th ICs) we can express the linearly propagating fields in groups a and b as

$$\vec{E}_a(z,t) = \sum_n g_{a,0}(z,t-nT)\vec{a}_n + (1-\delta_{l,m})e^{-i(l-m)\Omega t}\sum_n g_{a,l-m}(z,t-nT)\vec{b}_n \quad (34)$$

$$\vec{E}_b(z,t) = e^{-il\Omega t}\sum_n g_{b,l}(z,t-nT)\vec{c}_n + (1-\delta_{l,m})e^{-im\Omega t}\sum_n g_{b,m}(z,t-nT)\vec{d}_n, \quad (35)$$

where the components of the vector $\vec{a}_n$ are the data symbols transmitted in the n-th symbol slot in the channel of interest, whereas $\vec{b}_n$, $\vec{c}_n$, $\vec{d}_n$, represent the data transmitted in the various ICs. The vectors $\vec{a}_n$ and $\vec{b}_n$ are of length $2N_a$, whereas the vectors $\vec{c}_n$ and $\vec{d}_n$ are of lengths $2N_b$. The case $l=m$ corresponds to XGXPM, and the case $l \neq m$ corresponds to XGFWM. The functions $g_{a,k}(z,t)$ and $g_{b,k}(z,t)$ are the propagated versions of the fundamental waveform $g(t)$ in the k-th channel of groups a and b, respectively. The fundamental waveform is chosen so as to ensure ISI-free reception in the back-to-back configuration, that is $\int dt g^*(t-mT)g(t-nT) = \delta_{n,m}$, where the energy of $g(t)$ is normalized to 1 for convenience. Thus

$$g_{a,k}(z,t) = \exp\left[i\left(\beta_a + \beta'_a k\Omega + \frac{\beta''_a}{2}k^2\Omega^2\right)z\right]\exp\left(-\frac{i}{2}\beta''_a z\partial_t^2\right)g(t-\beta'_a z - k\Omega\beta''_a z) \quad (36)$$

$$g_{b,k}(z,t) = \exp\left[i\left(\beta_b + \beta'_b k\Omega + \frac{\beta''_b}{2}k^2\Omega^2\right)z\right]\exp\left(-\frac{i}{2}\beta''_b z\partial_t^2\right)g(t-\beta'_b z - k\Omega\beta''_b z) \quad (37)$$

The NLIN induced error in the received symbol vector of the COI in the 0-th time slot after chromatic dispersion compensation and matched filtering, is given by

$$\Delta\vec{a}_0 = i\gamma\kappa_{ab}\sum_{h,k,p}Y_{h,k,p}\vec{c}_k^\dagger\vec{c}_p\vec{a}_h, \quad \text{XGXPM}, l=m \quad (38)$$

$$\Delta \vec{d}_0 \;\; = \;\; i\gamma \kappa_{ab} \sum_{h,k,p} Y_{h,k,p} \vec{c}_k^\dagger \vec{d}_p \vec{b}_h, \quad \text{XGFWM}, \; l \neq m \tag{39}$$

where

$$Y_{h,k,p} = \int_0^L dz f(z) \int dt g_{a,0}^*(z,t) g_{a,l-m}(z, t - hT) g_{b,l}^*(z, t - kT) g_{b,m}(z, t - pT) \tag{40}$$

with $f(z)$ describing the z-dependent loss/gain profile [4]. By expanding Eqs. (38) and (39) with respect to the scalar products therein, the j-th component of $\Delta \vec{d}_0$ is seen to consist of $2N_b$ independent and identically distributed terms of the kind

$$i\gamma \kappa_{ab} \sum_q \sum_{h,k,p} Y_{h,k,p} c_{k,q}^* c_{p,q} a_{h,j}, \quad \text{XGXPM}, \; l = m \tag{41}$$

$$i\gamma \kappa_{ab} \sum_q \sum_{h,k,p} Y_{h,k,p} c_{k,q}^* d_{p,q} b_{h,j}, \quad \text{XGFWM}, \; l \neq m \tag{42}$$

where by $c_{k,q}$ we denote the q-th component of $\vec{c}_k$. The variance of each of these terms can be extracted by analogy with XPM and four-channel FWM contributions to NLIN in the scalar case. To do so, we express the error in the reception of the symbol a_0 in the scalar case

$$\Delta a_0 \;\; = \;\; i2\gamma \sum_{h,k,p} X_{h,k,p} c_k^* c_p a_h, \quad \text{XPM} \tag{43}$$

$$\Delta a_0 \;\; = \;\; i2\gamma \sum_{h,k,p} X_{h,k,p} c_k^* d_p b_h, \quad \text{4ch-FWM}, \tag{44}$$

where the coefficient $X_{h,k,p}$ is obtained from Eq. (40) by replacing the subscript b with a. The XPM variance of Δa_0 has been calculated in [4], and it can be expressed in the time-domain formalism as,

$$\sigma^2_{\text{XPM,scalar}} \;\; = \;\; P^3 \frac{8}{3} \chi_1 + P^3 \chi_2 \frac{16}{5} \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle^2} - 2 \right), \tag{45}$$

$$\chi_1 \;\; = \;\; \frac{3}{2} \gamma^2 \sum_{h,k,p} |X_{h,k,p}|^2 \tag{46}$$

$$\chi_2 \;\; = \;\; \frac{5}{4} \gamma^2 \sum_{h,k} |X_{h,k,k}|^2, \tag{47}$$

where by b we denote a symbol of the signal constellation used to modulate the various modes. Note that Eq. (45) contains coefficients that differ from those contained in Eq. (25) of [4], where the χ_1 and χ_2 were defined for the scalar case. Here we adopt the definition used in subsequent literature, where χ_1 and χ_2 correspond to the case of polarization-multiplexed transmission [26]. In addition, unlike in [26], the NLIN power and the signal power are intended to be both *per polarization* (as opposed to be per WDM channel), and the factor of 8/9 is absorbed into the definition of γ. Namely, we use the following expression for $\sigma^2_{\text{XPM,1}}$,

$$\sigma^2_{\text{XPM,1}} = 4P^3 \chi_1 + 4P^3 \chi_2 \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle^2} - 2 \right). \tag{48}$$

The quantity $\zeta^2_{\text{XPM,scalar}}$ is given by Eqs. (45)–(47), provided that the coefficient $X_{h,k,p}$ are replaced with $Y_{h,k,p}$ and $4\gamma^2$ is replaced with $\gamma^2 \kappa_{ab}^2$, namely [33]

$$\zeta^2_{\text{XPM,scalar}} \;\; = \;\; P^3 \frac{8}{3} \xi_1 + P^3 \frac{16}{5} \xi_2 \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle^2} - 2 \right), \tag{49}$$

$$\xi_1 = \frac{3}{8}\gamma^2\kappa_{ab}^2 \sum_{h,k,p} |Y_{h,k,p}|^2 \tag{50}$$

$$\xi_2 = \frac{5}{16}\gamma^2\kappa_{ab}^2 \sum_{h,k} |Y_{h,k,k}|^2, \tag{51}$$

Expressions for ξ_1 and ξ_2 which are suitable for efficient numerical evaluation are provided in the second part of this appendix.

We now switch to the extraction of the NLIN variance due to inter-group four-channel FWM, which we denoted as $\zeta_{4\text{ch,scalar}}^2$. To this end we evaluate the four-channel FWM variance of Δa_0 in the (single-mode) scalar case. The various FWM contributions have been evaluated previously (e.g. in [9]), but those derivations were performed in the frequency domain, and their relation with the coefficients $X_{h,k,p}$ is not obvious. The contribution of the four-channel FWM in the scalar case, which is used in Eq. (25), is

$$\sigma_{4\text{ch,scalar}}^2 = 4\gamma_0^2 \sum_{h,k,p}\sum_{h',k',p'} X_{h,k,p}X_{h',k',p'}^* \langle c_k^* c_{k'} d_p d_{p'}^* b_h b_{h'}^* \rangle = \frac{8}{3}\chi_1 P^3, \tag{52}$$

where the last equality follows from the statical independence of the symbols transmitted in the various channels, which yields $\langle c_k^* c_{k'}\rangle\langle d_p d_{p'}^*\rangle\langle b_h b_{h'}^*\rangle = \delta_{k,k'}\delta_{k,k'}\delta_{k,k'}P^3$. The quantity $\zeta_{4\text{ch,scalar}}^2$ is given by the same expression, using the same substitutions discussed earlier, namely

$$\zeta_{4\text{ch,scalar}}^2 = \frac{8}{3}\xi_1 P^3. \tag{53}$$

Inter-group XPM and FWM are most effective when the phase-matching condition is fulfilled. Inspection of (40) shows that inter-group XPM is strongest when $\Delta\beta_g = 0$, that is for

$$\beta_a' - \beta_b' = \beta_b'' l\Omega \tag{54}$$

On the other hand, inspection of Eqs. (34) and (35) shows that inter-group FWM is most effective when the following phase-matching condition is met

$$(\beta_a' - \beta_b')(l - m)\Omega + \left[\beta_a''(l-m)^2 - \beta_b''(l^2 - m^2)\right]\frac{\Omega^2}{2} = 0, \tag{55}$$

which simplifies to

$$\beta_a' - \beta_b' = m\frac{\beta_a'' + \beta_b''}{2}\Omega - l\frac{\beta_a'' - \beta_b''}{2}\Omega. \tag{56}$$

We now proceed to evaluating the variance $\sigma_{3\text{ch-I,scalar}}^2$, which results from the three-channel FWM interaction which is degenerate with respect to the COI. In this case we consider the scalar field

$$E(z,t) = \sum_n g(z,t-nT)a_n + e^{-im\Omega t}\sum_n g(z,t-nT)b_n + e^{im\Omega t}\sum_n g(z,t-nT)c_n \tag{57}$$

which yields the following error in the received symbol a_0

$$\Delta a_0 = i2\gamma_0 \sum_{h,k,p} W_{h,k,p}b_k c_p a_h^*, \tag{58}$$

where

$$W_{h,k,p} = \int_0^L \mathrm{d}z f(z)\int \mathrm{d}t g_{a,0}^*(z,t)g_{a,0}^*(z,t-hT)g_{a,m}(z,t-kT)g_{a,-m}(z,t-pT). \tag{59}$$

The variance of Δa_0 is

$$\sigma^2_{\text{3ch-I,scalar}} = 4\gamma_0^2 \sum_{h,k,p} \sum_{h',k',p'} W_{h,k,p} W^*_{h',k',p'} \langle b_k b^*_{k'} c_p c^*_{p'} a^*_h a_{h'} \rangle = 4\gamma_0^2 P^3 \sum_{h,k,p} \left| W_{h,k,p} \right|^2, \tag{60}$$

where in the last equality we used the independence of the symbols in the various channels. An expression for $\sigma^2_{\text{3ch-I,scalar}}$ which is suitable for numerical evaluation is provided in the second part of this appendix.

In order to compute the NLIN variance in the scalar case due to the three-channel FWM interaction that is degenerate with respect to one of the ICs, we consider the field

$$E(z,t) = e^{-2im\Omega t} \sum_n g(z, t - nT) b_n + e^{-im\Omega t} \sum_n g(z, t - nT) c_n, \tag{61}$$

which yields the following error in the detection of a_0,

$$\Delta a_0 = i\gamma_0 \sum_{h,k,p} Z_{h,k,p} b^*_h c_k c_p, \tag{62}$$

where we used the definition

$$Z_{h,k,p} = \int_0^L \mathrm{d}z f(z) \int \mathrm{d}t\, g^*_{a,0}(z,t) g^*_{a,2m}(z, t - hT) g_{a,m}(z, t - kT) g_{a,m}(z, t - pT) \tag{63}$$

The variance of Δa_0 is

$$
\begin{aligned}
\sigma^2_{\text{3ch-II,scalar}} &= \gamma_0^2 \langle |b|^2 \rangle \sum_{h,k,p} \sum_{h,k',p'} Z_{h,k,p} Z^*_{h,k',p'} \langle c_k c^*_{k'} c^*_p c_{p'} \rangle, \\
&= P^3 (\mu_1 + \mu_3) + P^3 \mu_2 \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle^2} - 2 \right),
\end{aligned}
\tag{64}
$$

where

$$\mu_1 = \gamma_0^2 \sum_{h,k,p} |Z_{h,k,p}|^2, \tag{65}$$

$$\mu_2 = \gamma_0^2 \sum_{h,k} |Z_{h,k,k}|^2, \tag{66}$$

$$\mu_3 = \gamma_0^2 \sum_h \left| \sum_k Z_{h,k,k} \right|^2, \tag{67}$$

and where $\mu_1 + \mu_3$, and μ_2 are the SON and FON coefficients, respectively, in this type of three-channel interaction. In order to obtain $\sigma^2_{\text{3ch-II,1}}$, the same procedure needs to be repeated for the single-mode case with two-polarizations. The field expression is

$$\vec{E}(z,t) = e^{-2im\Omega t} \sum_n g(z, t - nT) \vec{b}_n + e^{-im\Omega t} \sum_n g(z, t - nT) \vec{c}_n, \tag{68}$$

which yields the following error vector in the detection of a_0,

$$\Delta \vec{a}_0 = i\gamma_0 \sum_{h,k,p} Z_{h,k,p} \vec{b}^\dagger_h \vec{c}_k \vec{c}_p. \tag{69}$$

Its x component is given by

$$\Delta a_{0_x} = i\gamma_0 \left(\sum_{h,k,p} Z_{h,k,p} b^\dagger_{h,y} c_{k,y} c_{p,x} + \sum_{h,k,p} Z_{h,k,p} b^\dagger_{h,x} c_{k,x} c_{p,x} \right), \tag{70}$$

and the variance can straightforwardly be found to be

$$\sigma^2_{3\text{ch}-\text{II},1} = P^3 \mu_1 + \sigma^2_{3\text{ch}-\text{II,scalar}} = P^3(2\mu_1 + \mu_3) + P^3 \mu_2 \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle^2} - 2 \right). \tag{71}$$

We now conclude this part by evaluating the single-channel NLIN contributions $\sigma^2_{\text{SPM,scalar}}$ and $\sigma^2_{\text{SPM},1}$. We start with the scalar case, where we may write

$$\Delta a_0 = i\gamma_0 \sum_{h,k,p} S_{h,k,p} a_h a_k^* a_p \tag{72}$$

where [4]

$$S_{h,k,p} = \int_0^L \mathrm{d}z f(z) \int \mathrm{d}t g_{a,0}^*(z,t) g_{a,0}(z,t-hT) g_{a,0}^*(z,t-kT) g_{a,0}(z,t-pT). \tag{73}$$

Like in the case of XPM [4], the NLIN power is obtained as the mean square value of $\Delta a_0 - i\theta_{\text{RX}} a_0$, where θ_{RX} is the rotation of the symbol constellation removed by the receiver, hence $\sigma^2_{\text{scalar}} = \langle |\Delta a_0|^2 \rangle + \langle |b|^2 \rangle \theta^2_{\text{RX}} - 2\theta_{\text{RX}} \text{Im} \left\{ \langle a_0^* \Delta a_0 \rangle \right\}$. We start by computing $\langle |\Delta a_0|^2 \rangle$, which can be expanded as

$$\langle |\Delta a_0|^2 \rangle = \gamma_0^2 \sum_{h,k,p} \sum_{h',k',p'} S_{h,k,p} S^*_{h',k',p'} \langle a_h a_k^* a_p a_{h'}^* a_{k'} a_{p'}^* \rangle. \tag{74}$$

There are three cases that need to be distinguished. In the first case all the indexes are identical, with the result

$$\gamma_0^2 \langle |b|^6 \rangle \sum_h |S_{h,h,h}|^2. \tag{75}$$

The second case is the one with a quadruplet and a pair of identical indexes. In this case the result is proportional to $\langle |b|^4 \rangle \langle |b|^2 \rangle$, and there are nine possible combinations, with the result

$$\gamma_0^2 \langle |b|^4 \rangle \langle |b|^2 \rangle \left[\sum_{h,k} \left(4|S_{h,k,k}|^2 + |S_{k,h,k}|^2 \right) - 9 \sum_k |S_{k,k,k}|^2 + 4 S_{0,0,0} \sum_k S_{0,k,k} \right], \tag{76}$$

in whose derivation we used the property $\sum_k S_{h,k,k} = \delta_{h,0} \sum_k S_{0,k,k}$. The third case is the one with three different pairs of indexes. In this case there are six possible combinations (all possible triplets of pairs, where each pair contains one conjugated symbol and one non-conjugated symbol). The total contribution of these terms is

$$\gamma_0^2 \langle |b|^2 \rangle^3 \sum_{h \neq k \neq p} \left[2|S_{h,k,p}|^2 + 4 S_{h,k,k} S^*_{h,p,p} \right], \tag{77}$$

which, after some algebra, can be expressed as

$$\gamma_0^2 \langle |b|^2 \rangle^3 \left[2 \sum_{h,k,p} |S_{h,k,p}|^2 - 2 \sum_{h,k} \left(4|S_{h,k,k}|^2 + |S_{k,h,k}|^2 \right) + 12 \sum_h |S_{h,h,h}|^2 + 8 S_{0,0,0} \sum_k S_{0,k,k} \right], \tag{78}$$

In a similar way one can compute

$$2\text{Im} \left\{ \langle a_0^* \Delta a_0 \rangle \right\} = 2\gamma_0^2 \left(S_{0,0,0} \langle |b|^4 \rangle + 2\langle |b|^2 \rangle^2 \sum_{k \neq 0} S_{0,k,k} \right) \tag{79}$$

In order to finalize this calculation we need to specify the expression of θ_{RX}, based on what the receiver does.

The simplest assumption is that the receiver removes the average rotation produced by SPM. This can be computed by averaging all the triplets of the form $|a_k|^2 a_0$ with respect to $|a_k|^2$, with the result

$$\theta_{RX} = 2\gamma_0 \langle |b|^2 \rangle \sum_k S_{0,k,k}. \tag{80}$$

This type of phase recovery is rather simple (it does not imply any adaptive procedure), but at the same time it is clearly sub-optimal. More accurate approaches to removing the nonlinear phase shift imply adaptive capabilities of the receiver [8], or the use of pilots. In the ideal case, the rotation removed by the receiver follows from minimizing the mean-square distance between the received symbols and the transmitted ones. By straightforward algebra one can find the following expression for the rotation,

$$e^{i\theta_{RX}} = \frac{\langle a_0^*(a_0 + \Delta a_0) \rangle}{|\langle a_0^*(a_0 + \Delta a_0) \rangle|}, \tag{81}$$

which to first order in θ_{RX} becomes $\theta_{RX} = \text{Im}\left\{\langle a_0^* \Delta a_0 \rangle\right\} / \langle |a_0|^2 \rangle$, with the result

$$
\begin{aligned}
\theta_{RX} &= \gamma_0 \left(S_{0,0,0} \frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle} + 2\langle |b|^2 \rangle \sum_{k \neq 0} S_{0,k,k} \right) \\
&= \gamma_0 S_{0,0,0} \left(\frac{\langle |b|^4 \rangle}{\langle |b|^2 \rangle} - 2\langle |b|^2 \rangle \right) + 2\gamma_0 \langle |b|^2 \rangle \sum_k S_{0,k,k}
\end{aligned}
\tag{82}
$$

The difference between the two expressions is in the first term of the above. This term depends on the modulation format, and it vanishes for Gaussian modulation.

In the case of two polarizations the error vector can be expressed as

$$\Delta \vec{a}_0 = i\gamma_0 \sum_{h,k,p} S_{h,k,p} \vec{a}_h (\vec{a}_k^\dagger \vec{a}_p). \tag{83}$$

From this point it is convenient to look at the two components of $\Delta \vec{a}_0$ separately. In fact, the variance of the vector is the sum of the variances of its two components, which by symmetry arguments are identical. We hence express the x-component of the vector $\Delta \vec{a}_0$ as

$$\Delta a_{0,x} = i\gamma_0 \sum_{h,k,p} S_{h,k,p} a_{h,x} a_{k,x}^* a_{p,x} + i\gamma_0 \sum_{h,k,p} S_{h,k,p} a_{h,x} a_{k,y}^* a_{p,y} = \Delta x_0 + \Delta y_0, \tag{84}$$

where Δx_0 represents the interactions within the x-polarized components of the data symbols, whereas Δy_0 accounts for cross-polarization effects. We denote again by θ_{RX} the phase shift removed by receiver. The NLIN power can hence be expressed as

$$
\begin{aligned}
\sigma_{SPM,1}^2 &= \langle |\Delta x_0 + \Delta y_0|^2 \rangle - 2\theta_{RX} \text{Im}\left\{\langle a_{0,x}^*(\Delta x_0 + \Delta y_0) \rangle\right\} + \theta_{RX}^2 \langle |b|^2 \rangle \\
&= \langle |\Delta x_0|^2 \rangle + \langle |\Delta y_0|^2 \rangle + 2\text{Re}\left\{\langle \Delta x_0^* \Delta y_0 \rangle\right\} \\
&\quad - 2\theta_{RX} \text{Im}\left\{\langle a_{0,x}^*(\Delta x_0 + \Delta y_0) \rangle\right\} + \theta_{RX}^2 \langle |b|^2 \rangle.
\end{aligned}
\tag{85}
$$

The terms $\langle |\Delta x_0|^2 \rangle$ and $\langle |\Delta y_0|^2 \rangle$ are identical to those evaluated for SPM and XPM in the scalar case, respectively. The term $\langle \Delta x_0^* \Delta y_0 \rangle$ is new. After some straightforward algebra, it can be expressed in the following form,

$$\langle \Delta x_0^* \Delta y_0 \rangle = \gamma_0^2 \left(\sum_k S_{0,k,k} \right) \left(S_{0,0,0} \langle |b|^4 \rangle + 2\langle |b|^2 \rangle^2 \sum_{k \neq 0} S_{0,k,k} \right). \tag{86}$$

The remaining term is

$$\langle a_{0,x}^*(\Delta x_0 + \Delta y_0)\rangle = \gamma_0\left[S_{0,0,0}\langle|b|^4\rangle + 3\langle|b|^2\rangle^2\sum_{k\neq 0}S_{0,k,k}\right]. \tag{87}$$

The calculation ends by specifying the expression of θ_{RX}. The first definition of θ_{RX} discussed in the scalar case yields

$$\theta_{\mathrm{RX}} = 3\gamma_0\langle|b|^2\rangle\sum_k S_{0,k,k}, \tag{88}$$

whereas the second yields

$$\theta_{\mathrm{RX}} = \gamma_0 S_{0,0,0}\left(\frac{\langle|b|^4\rangle}{\langle|b|^2\rangle} - 2\langle|b|^2\rangle\right) + 3\gamma_0\langle|b|^2\rangle\sum_k S_{0,k,k}. \tag{89}$$

As in the scalar case, the two rotation angles are identical for Gaussian modulation.

8.2. Efficient calculation of the coefficients involved in the NLIN power expressions

The NLIN power expressions derived in the previous section involve the following coefficients,

$$\mathcal{S}_1 = \sum_{h,k,p}|S_{h,k,p}|^2, \quad \mathcal{S}_2 = \sum_{h,k}|S_{h,k,k}|^2, \quad \mathcal{S}_3 = \sum_{h,k}|S_{h,k,h}|^2, \tag{90}$$

$$\mathcal{S}_4 = \sum_h|S_{h,h,h}|^2, \quad \mathcal{S}_5 = \sum_k S_{0,k,k} \tag{91}$$

$$\mathcal{W}_1 = \sum_{h,k,p}|W_{h,k,p}|^2, \tag{92}$$

$$\mathcal{X}_1 = \sum_{h,k,p}|X_{h,k,p}|^2, \quad \mathcal{X}_2 = \sum_{h,k}|X_{h,k,k}|^2 \tag{93}$$

$$\mathcal{Y}_1 = \sum_{h,k,p}|Y_{h,k,p}|^2, \quad \mathcal{Y}_2 = \sum_{h,k}|Y_{h,k,k}|^2 \tag{94}$$

$$\mathcal{Z}_1 = \sum_{h,k,p}|Z_{h,k,p}|^2, \quad \mathcal{Z}_2 = \sum_{h,k}|Z_{h,k,k}|^2, \quad \mathcal{Z}_3 = \sum_h\left|\sum_k Z_{h,k,k}\right|^2. \tag{95}$$

By using the following definition

$$C_{h,k,p}(a,x,l,m) = \int_0^L \mathrm{d}z f(z)\int \mathrm{d}t g_{a,0}^*(z,t)g_{a,l-m}(z,t-hT)g_{x,l}^*(z,t-kT)g_{x,m}(z,t-pT), \tag{96}$$

where $x \in \{a,b\}$, the pulse-collision coefficients can be expressed in terms of $C_{h,k,p}(a,x,l,m)$ as follows: $S_{h,k,p} = C_{h,k,p}(a,a,0,0)$, $W_{h,k,p} = C_{p,h,k}(a,a,0,m)$, $X_{h,k,p} = C_{h,k,p}(a,a,l,m)$, $Y_{h,k,p} = C_{h,k,p}(a,b,l,m)$, $Z_{h,k,p} = C_{k,h,p}(a,b,2m,m)$. Given this, the different types of integrals that need to be computed reduce to

$$C_1 = \sum_{h,k,p}|C_{h,k,p}(a,x,l,m)|^2, \tag{97}$$

$$C_2 = \sum_{h,k}|C_{h,k,k}(a,x,l,m)|^2, \tag{98}$$

$$C_3 = \sum_{h,k}|C_{h,k,h}(a,x,l,m)|^2, \tag{99}$$

$$C_4 = \sum_k \left| \sum_h C_{h,k,h}(a,x,l,m) \right|^2, \tag{100}$$

$$C_5 = \sum_h |C_{h,h,h}(a,x,l,m)|^2, \tag{101}$$

$$C_6 = \sum_h C_{0,h,h}(a,x,l,m). \tag{102}$$

The numerical extraction of Eqs. (97)–(102) is highly nontrivial. A great simplification can be achieved by following the procedure introduced in [8]. Using the formalism of [8], the interaction coefficients $C_{h,k,p}(a,x,l,m)$ may be expressed as frequency-domain integrals,

$$C_{h,k,p}(a,x,l,m) = \frac{1}{(2\pi)^3} \int d\omega d\omega' d\omega'' \tilde{g}^*_{a,0}(\omega') \tilde{g}_{a,l-m}(\omega+\omega') \tilde{g}^*_{x,l}(\omega'') \tilde{g}_{x,m}(\omega''-\omega)$$
$$\phi(\omega,\omega',\omega'') e^{-ihT(\omega+\omega')} e^{ikT\omega''} e^{-ipT(\omega''-\omega)}, \tag{103}$$

where $\tilde{g}(w)$ is the Fourier transform of $g(t, z=0)$, and the *link function* $\phi(\omega,\omega',\omega'')$ is given by

$$\phi(\omega,\omega',\omega'') = \int_0^L dz f(z) e^{iz\eta(\omega,\omega',\omega'')}, \tag{104}$$

where we defined

$$\eta = \beta_a'' \left[\frac{1}{2}\omega^2 + \omega\omega' - (\omega'+\omega)(l-m)\Omega + \frac{1}{2}(l-m)^2\Omega^2 \right] \tag{105}$$
$$+ \beta_x'' \left[\frac{1}{2}\omega^2 - \omega\omega'' + \omega'' l\Omega - (\omega''-\omega)m\Omega - \frac{1}{2}(l^2-m^2)\Omega^2 \right] + \Delta\beta' \left[(l-m)\Omega - \omega \right].$$

Using these definitions, the coefficients C_n ($n = 1 \ldots 6$) may be expressed as 3rd order integrals. Their manipulation involves a number of straightforward, yet cumbersome steps, which are either identical or very similar to those detailed in [8]. As a result, the coefficient C_1 is found to be given by

$$C_1 = \frac{1}{(2\pi)^3 T^3} \int d\omega d\omega' d\omega'' I^{(1)}_{0,0,0} \sum_{h,k,p} I^{(1)*}_{h,k,p} \tag{106}$$

where

$$I^{(1)}_{h,k,p}(\omega,\omega',\omega'') = \tilde{g}^*_{a,0}\left(\omega' + \frac{2\pi}{T}(k-p-h)\right) \tilde{g}_{a,l-m}\left(\omega+\omega' - \frac{2\pi}{T}h\right) \tilde{g}^*_{b,l}\left(\omega'' - \frac{2\pi}{T}k\right) \tag{107}$$
$$\tilde{g}_{b,m}\left(\omega'' - \omega - \frac{2\pi}{T}p\right) \phi\left(\omega + \frac{2\pi}{T}(p-k), \omega' + \frac{2\pi}{T}(k-p-h), \omega'' - \frac{2\pi}{T}k\right).$$

The coefficient C_2 is given by

$$C_2 = \frac{1}{(2\pi)^4 T^2} \int d\omega d\omega' I^{(2)}_{0,0} \sum_{h,k} I^{(2)*}_{h,k}, \tag{108}$$

where

$$I^{(2)}_{h,k}(\omega,\omega') = \tilde{g}^*_a\left(\omega' - \frac{2\pi}{T}(k+h)\right) \tilde{g}_a\left(\omega+\omega' - \frac{2\pi}{T}h\right) \tag{109}$$
$$\int d\omega'' \tilde{g}^*_b(\omega'') \tilde{g}_b\left(\omega'' - \omega - \frac{2\pi}{T}k\right) \phi\left(\omega + \frac{2\pi}{T}k, \omega' - \frac{2\pi}{T}(k+h), \omega''\right).$$

The coefficient C_3 is given by

$$C_3 = \frac{1}{(2\pi)^4 T^2} \int d\omega' d\omega'' I_{0,0}^{(3)} \sum_{h,k} I_{h,k}^{(3)*}, \tag{110}$$

where

$$I_{h,k}^{(3)} = g_{a,0}^*(\omega') g_{x,l}^* \left(\omega'' - \frac{2\pi}{T}k\right) \int d\omega g_{a,l-m} \left(\omega + \omega' + \frac{2\pi}{T}(h+k)\right) \tag{111}$$

$$g_{x,m} \left(\omega'' - \frac{2\pi}{T}k\right) \phi \left(\omega, \omega' + \frac{2\pi}{T}(k+h), \omega'' - \frac{2\pi}{T}k\right).$$

The coefficient C_4 is given by

$$C_4 = \frac{1}{(2\pi)^3 T^3} \int d\omega' \sum_h I_{h,0}^{(4)} \sum_k I_{h,k}^{(4)*}, \tag{112}$$

where

$$I_{h,k}^{(4)} = g_{a,0}^* \left(\omega' - \frac{2\pi}{T}k\right) g_{x,l}^* \left(-\omega' - \frac{2\pi}{T}(h-k)\right) \int d\omega g_{a,l-m} \left(\omega + \omega' - \frac{2\pi}{T}k\right) \tag{113}$$

$$g_{x,m} \left(-\omega - \omega' - \frac{2\pi}{T}(h-k)\right) \phi \left(\omega, \omega' - \frac{2\pi}{T}k, -\omega' - \frac{2\pi}{T}(h-k)\right).$$

The coefficient C_5 is given by

$$C_5 = \frac{1}{(2\pi)^5 T} \int d\omega' I_{0,0}^{(5)} \sum_k I_k^{(5)*}, \tag{114}$$

where

$$I_k^{(5)}(\omega') = \tilde{g}_{a,0}^* \left(\omega' - \frac{2\pi}{T}k\right) \tag{115}$$

$$\int d\omega d\omega'' \tilde{g}_{a,l-m} \left(\omega + \omega' - \frac{2\pi}{T}k\right) \tilde{g}_{x,l}^* (\omega'') \tilde{g}_{x,m} (\omega'' - \omega) \phi \left(\omega, \omega' - \frac{2\pi}{T}k, \omega''\right).$$

Lastly, the coefficient C_6 is given by

$$C_6 = \frac{1}{(2\pi)^2 T} \sum_h \int d\omega' d\omega'' I_h^{(6)}, \tag{116}$$

where

$$I_h^{(6)} = g_{a,0}^* (\omega') g_{a,l-m}^* \left(\omega' + \frac{2\pi}{T}h\right) g_{x,l}^* (\omega'') g_{x,m}^* \left(\omega'' - \frac{2\pi}{T}h\right) \phi \left(\frac{2\pi}{T}h, \omega', \omega''\right). \tag{117}$$

The most important feature of this representation is that the summations are no longer infinite. Assuming that the pulse spectrum vanishes outside the range $[-1/T, 1/T]$, each summation is restricted to the only terms for which $h, k, p \in [-1, 0, 1]$. For the specific case of Nyquist pulses, the only non zero term is $h = k = p = 0$.

Funding

Acknowledgement