Fast Estimation of Nonlinear Interference Noise in Distributed Raman-Amplified Few-Mode Fiber Links

Francesco Lorenzi(*), Gianluca Marcon, Luca Schenato *Member, IEEE*, Antonio Mecozzi *Fellow, IEEE, Fellow, OPTICA*, Cristian Antonelli *Senior Member, IEEE, Fellow, OPTICA*, Marco Santagiustina *Member, IEEE*

***Abstract*—We present a model, based on the calculation of collision integrals, for evaluating nonlinear interference noise (NLIN) in a distributed Raman-amplified, few-mode fiber (FMF), wavelength division multiplexed (WDM) link. It is first demonstrated, neglecting dispersion and power evolution along the fiber, that the collision integrals have asymptotic limits for low and high differential group delay (DGD) between the channels. A general approximated model is then derived for any DGD regime. Dispersion and power evolution are incorporated into the model via correction factors applied to the asymptotic limits, which require only a very limited number of integrals to be numerically evaluated. The accuracy of the analytical estimates is confirmed through comparison with full numerical evaluations of the collision integrals. The proposed method offers a substantial reduction in computational effort compared with direct numerical calculations, which become prohibitive in FMFs due to the large number of interfering channels and the large number of collisions caused by large DGDs between mode groups. We apply the model to a link based on a FMF supporting four mode groups and with counterpropagating distributed Raman amplification. We confirmed the significant computational efficiency and accuracy, since the largest number of channel pairs present high DGD where the model is fast and accurate. The comparison of the FMF-based system with one based on a standard single-mode fiber (SMF), loaded by the same WDM grid of each mode group of the FMF, and comparable Raman amplification scheme, shows, as expected, that the NLIN is lower in the FMF link. However, NLIN in FMF can vary significantly, with subsets of channels exhibiting notably stronger intermodal NLIN due to group-velocity (GV) matching conditions.**

I. Introduction

NONLINEAR interference noise (NLIN) represents one of the major capacity limitations to current optical fiber communication links [1, 2]. This effect arises from the nonlinear interaction between pulses that propagate together in the fiber across different channels of a wavelength-division-multiplexed (WDM) system [3]. A pulse on a given channel propagates with a group velocity (GV) determined by the chromatic and modal dispersion of the fiber, which may exhibit a nontrivial profile across a wide WDM bandwidth. Pulse self-interference, yielding NLIN due to the self-phase modulation (SPM), can be compensated through digital backpropagation; however, cross-phase modulation (XPM) and four wave mixing (FWM) processes are more complex to mitigate, and they contribute to set the current capacity crunch [4–6].

From a theoretical perspective, models have been developed to estimate NLIN in single-mode fibers (SMF), using two distinct theoretical approaches. In the frequency domain, the GN model [7–10] is based on a specific signal model and the properties of Gaussian noise. Its extension, known as enhanced GN (EGN), relaxes some of the model's hypotheses, such as the fixed modulation of the model signal, and has been widely applied [11]. The GN model was extended to FMFs in Refs. [12, 13], providing a good match to experimental results.

A complementary time domain approach has been shown to be effective for a concise description of the NLIN as the composition of a modulation-dependent and a modulation-independent part, where the modulation-independent part coincides with the one computed using the GN model [14, 15]. This theory uses a perturbative approach on the nonlinear Schrödinger equation for weak nonlinearity to derive time and space integrals modeling the magnitude of interaction of each pair of interacting pulses [16]. Since the central role of these integrals, this is also known as pulse collision model [16].

This time-domain model has also been proposed for systems based on few-mode fibers (FMF), extending the methodology to the coupled Manakov equation under the assumption of strong intra-group mixing [17]. The noise variance has been determined as a function of the pulse-collision integrals for all possible channel interactions, categorized as self-interactions, cross-group interactions, and four-wave mixing [18].

In fiber-optic communication systems, distributed Raman amplification is a well-known technique for achieving lower noise figures than lumped amplification with rare-earth-doped fiber amplifiers [19]. Evaluation of NLIN within the pulse collision model in the presence of Raman amplification poses challenges due to the need to weight pulse collisions by the Raman amplification profile along the fiber. Closed-form analytical results within the GN model can be obtained in the undepleted pump approximation, or within similar semi-analytical approaches [20–22]. In a previous study [23, 24], the problem of computing the NLIN within the time domain model in the presence of arbitrary Raman amplification over the WDM grid of a SMF was addressed, adopting a fully numerical approach to find optimal transmission parameters and comparing Raman pumping schemes. It was shown that,

Francesco Lorenzi, Luca Schenato and Marco Santagiustina are with the Department of Information Engineering, University of Padova, 35131 Padova, Italy (* e-mail: francesco.lorenzi.2@phd.unipd.it). Luca Schenato and Marco Santagiustina are also with CNIT – National Inter-University Consortium for Telecommunications. Cristian Antonelli and Antonio Mecozzi are with the Dipartimento di Scienze fisiche e chimiche, University of L'Aquila, L'Aquila, Italy, and with CNIT – National Inter-University Consortium for Telecommunications.

in a wideband system, the signal-to-signal Raman interaction is also present, but its effect on the noise is weak compared to NLIN [24].

As mentioned, NLIN evaluation in the time domain is based on the calculation of pulse collision integrals, which consist of double integrations over time and space, specific to a given couple of pulses from two different channels. These integrals are generally not amenable to closed-form solutions and depend on the dispersion properties of the interaction channels. The number of collision integrals to be evaluated depends on the number of WDM channels and the differential group delay (DGD) between the channels. For these reasons, the computation in the FMF link can become prohibitive, as the number of collisions can become extremely large due to inter-modal dispersion, which can significantly surpass intra-modal dispersion.

In this work, we develop and validate a model that enables accurate, fast estimation of the NLIN in FMF, including Raman amplification effects and fully accounting for fiber dispersion. The model is derived analytically for Nyquist (zero roll-off) and Gaussian pulses, but it can be extended to any pulse shape. Although Gaussian pulses are not used in real systems, they have been considered because they provide a fully analytical framework that highlights a general behaviour that is independent of the pulse format and yields physical insights. In particular, the model, presented in Sect. II, is based on the fact that analytical expressions of the collision integral exist in the case of high and low DGD under ideal conditions (negligible dispersion and no power evolution along the fiber). Starting from the asymptotic results, a fitting model, that enables evaluation on NLIN for any value of the DGD, is found. Finally, we showed how the power evolution of channels (e.g. due to losses and Raman amplification) and dispersion can be also accurately accounted within the analytical approximation. To demonstrate the effectiveness of the proposed method, in Sect. III we evaluate the NLIN of a FMF with four mode groups and compare it with that of a single-mode fiber (SMF) system. Finally, In Sect. IV, we draw our conclusions.

II. Model description

The starting point of the analysis is the model for NLIN calculation reported in Ref. [18], in which expressions are derived from all nonlinear channel interactions. In our notation, channels are denoted by both the mode group and their position in the WDM grid. The channel of interest is considered to belong to mode group a and has an index p in the WDM grid. In this notation, ap is the channel of interest. To logically distinguish interfering channels, we denote them, in the same notation, with the mode and frequency pair bq. Based on these definitions, nonlinear interactions can be identified and associated with the corresponding independent noise contribution, denoted by σ^2. Here, we focus on the XPM contribution to NLIN, and also work within the assumption that the modes inside the same group are strongly mixed, whereas different mode groups have negligible mixing [12, 13, 15, 18]. Within this assumption, FWM terms are neglected: some remarks on this approximation will be given at the end of the paper. The interactions of interest are classified as follows: if $a = b$ and $p = q$, the interaction is self-phase modulation (SPM), otherwise it is cross-phase modulation (XPM). Considering $p \neq q$, if $a = b$ the interaction is in the same modal group (SG), so it is called SGXPM, and if $a \neq b$ it is on crossed groups (XG), so it is called XGXPM. Within this notation, the total NLIN power due to XPM can be written as a sum of independent noise contributions

$$\sigma^2_{\mathrm{XPM},ap} = \sum_{q\neq p} \sigma^2_{\mathrm{SGXPM},apaq} + \sum_{b\neq a}\sum_{q} \sigma^2_{\mathrm{XGXPM},apbq}\,. \quad (1)$$

Each of these terms depends on the properties of the XPM-type collisions between pulses [18], namely,

$$\begin{aligned} X_{hkm;apbq} = &\int_0^L dz\ f_{bq}(z) \int_{-\infty}^{\infty} dt \\ &\times\ g^{(0)}_{ap}(z,t)^* g^{(0)}_{ap}(z,t+hT) g^{(0)}_{bq}(z,t+kT-\Delta\beta_{1;apbq}\,z)^* \\ &\times g^{(0)}_{bq}(z,t+mT-\Delta\beta_{1;apbq}\,z)\,, \end{aligned} \quad (2)$$

where L is the fiber length, T is the symbol period, the functions ${g_{ap}}^{(0)}(z,t)$, ${g_{bq}}^{(0)}(z,t)$ are the pulse shapes linearly propagated to position z, and normalized so the integral of their modulus square in time is equal to 1 [15], $\Delta\beta_{1;apbq}$ is the DGD per unit length among channels ap and bq, and h,k,m are the symbol indexes in the transmission sequence. The symbol indexes denote shifts in time, indicating different pulses within the same data stream of a given channel ap or bq. They span the full set of the integer numbers. The function f_{bq} is introduced to model losses and distributed amplification along the fiber

$$f_{bq}(z) = \frac{P_{bq}(z)}{P_{bq}(0)}\,, \quad (3)$$

where $P_{bq}(z)$ is the average power of the interfering channel along the fiber. This quantity is, in general, dependent on the choice of the channel bq. Since the expression (2) includes pulses which have vanishing amplitude for all choices of coordinates (z,t) except the ones related to the position of the peak, the dominant choice of indices h,k,m is given by $h=0, k=m$. It is sufficient to consider this term for most practical XPM noise computations [15], namely ,the task is to compute the collision integrals

$$X_{0mm;apbq} = \int_0^L dz\ f_{bq}(z)\ I_{m;apbq}(z)\,, \quad (4)$$

$$\begin{aligned} I_{m;apbq}(z) = &\int_{-\infty}^{\infty} dt\ |g^{(0)}_{ap}(z,t)|^2 \\ &\times |g^{(0)}_{bq}(z,t+mT-\Delta\beta_{1;apbq}\,z)|^2\,. \end{aligned} \quad (5)$$

This subset of collision integrals, with the index choices described above, is real, unlike other choices of indices in Eq. (2).

Considering a transmission of a sequence of i.i.d. symbols following the complexrandom variable b_0 for each channel, with pulse time T and the same average energy per symbol $\langle |b_0|^2\rangle = PT$, with P the average power per polarization, we obtain the following expression for the noise contributions [18]

(see also the Appendix VI-A). The SGXPM can be simplified as

$$\sigma^2_{\mathrm{SGXPM},apaq} \approx P^3T^2\gamma^2\kappa^2_{aa}\left[\frac{\langle|b_0|^4\rangle}{\langle|b_0|^2\rangle^2}(2N_a+3)-4\right]\times\mathcal{N}_{apaq}\,, \tag{6}$$

when $b=a$, and the XGXPM becomes

$$\sigma^2_{\mathrm{XGXPM},apbq} \approx 2N_b\ P^3T^2\gamma^2\kappa^2_{ab}\left(\frac{\langle|b_0|^4\rangle}{\langle|b_0|^2\rangle^2}-1\right)\times\mathcal{N}_{apbq}\,. \tag{7}$$

when $b\neq a$, where P is the average power per polarization, κ_{ab} are mode dependent overlap integrals (defined in Appendix VI-D) N_a and N_b are the number of spatial mode per mode groups a and b, respectively, $\gamma=\omega_0 n_2/(cA_{\mathrm{eff}})$, A_{eff} being the effective area of the fundamental mode, and ω_0 the angular frequency at the WDM band center. The term $\mathcal{N}_{abpq}$ is defined as

$$\mathcal{N}_{apbq}=\sum_m {X_{0mm;apbq}}^2\,, \tag{8}$$

and contains all the information about the pulse collisions. In the following, the subscripts $apbq$ of the expressions $\mathcal{N}_{apbq}$, $X_{0mm;apbq}$, $I_{m;apbq}$, and $\Delta\beta_{1;apbq}$ are dropped for brevity, as they can be easily inferred from the context.

The result of the integration of Eq. (4-5) depends on the chromatic dispersion of the fiber for the two channels and on the relative DGD. In fact, the local collision integral (Eq. (5)) can be rewritten into

$$I_m(z)=\int_{-\infty}^{+\infty} dt\,|c(z,t)|^2\,, \tag{9}$$

with

$$c(z,t):=g^{(0)}_{ap}(z,t)\,g^{(0)}_{bq}(z,t+mT-\Delta\beta_1 z) \tag{10}$$

and can be evaluated in the frequency domain using the Parseval identity. In fact, in the frequency domain, exact expressions of the Fourier transforms of the propagated pulses $G_{ap}(z,\omega)$ and $G_{bq}(z,\omega)$, can be readily calculated using the dispersion propagation operators (for Nyquist and Gaussian pulses – see also Appendix VI-B – there also exist analytical expressions). The resulting expression is

$$I_m(z)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega|C(z,\omega)|^2, \tag{11}$$

where $C(z,\omega)$ is the Fourier transform of $c(z,t)$. Expressing explicitly the dispersion propagation operators,

$$C(z,\omega')=\frac{1}{2\pi T}\exp\left(i\frac{z\omega'^2}{2L_{Dbq}}\right)\exp\left[-i\omega T\left(\frac{z}{L_W}-m\right)\right]\times\int_{-\infty}^{+\infty}d\Omega'\,G_{ap}(0,\Omega')\,G_{bq}(0,\omega'-\Omega')\exp\left(i\frac{z}{L_{Dm}}\Omega'^2\right)\times\exp\left(-i\frac{z}{L_{Dbq}}\omega'\Omega'\right)\exp\left[i\Omega'\left(\frac{z}{L_W}-m\right)\right] \tag{12}$$

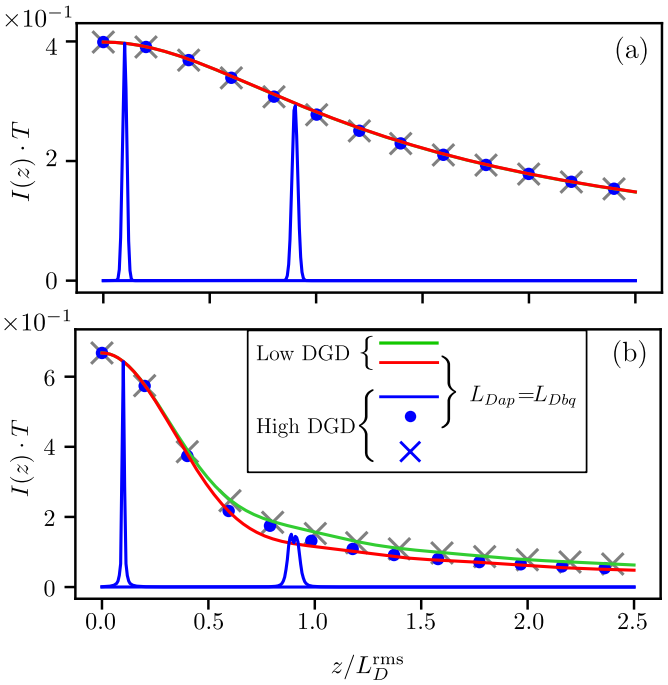

Fig. 1: Local collision integral I_m (from numerical integrations of Eq. (5)) as a function of the distance z (normalized to the rms dispersion length $L_D^{\mathrm{rms}}=\sqrt{2}L_{Dap}L_{Dbq}/\sqrt{{L_{Dap}}^2+{L_{Dbq}}^2}$), in panel (a) for Gaussian pulses and in panel (b) for Nyquist pulses. For a channel pair with high DGD ($L_W/L_D^{\mathrm{rms}}=10^{-2}$) and with equal dispersion lenghts $L_{Dap}=L_{Dbq}=L_D^{\mathrm{rms}}$, the results are represented by the blue curves, which refer to collisions with indices $m=10$ and $m=90$, and by the blue dots that are the peak values for other collisions (displayed once every 20 collisions). The blue crosses are the peak values for channels with different dispersion lengths still yielding the same rms dispersion length (in particular, $L'_{Dap}=2L_D^{\mathrm{rms}}$ and $L'_{Dbq}=4L_D^{\mathrm{rms}}/\sqrt{3}$). The red and green curves are obtained for vanishing DGD, respectively, for the conditions of equal or different dispersion lengths as specified above.

where $G_X(0,\omega)$, $X=ap,\ bq$ are the Fourier transform of the input pulses, $\omega'=\omega T$, $\Omega'=\Omega T$ are normalized frequencies, and $L_{Dm}=T^2/|\beta_{2m}|=2T^2/|\beta_{2,ap}+\beta_{2,bq}|$. We remark the presence of different length scales relative to the dispersion properties, namely the dispersion length associated to channel ap, $L_{Dap}=T^2/|\beta_{2,ap}|$, that associated to channel bq, $L_{Dbq}=T^2/|\beta_{2,bq}|$, and the walkoff length $L_W=T/|\Delta\beta_1|$.

A remarkably simple closed form expression is available for the case of Gaussian pulses. Even if they are not used in practice, they provide analytical insight as the integral (5) can be solved directly in the time domain by square completion. The result is

$$I_m(z)=\frac{1}{T\sqrt{2\pi}}\frac{1}{\sqrt{1+(z/L_D^{\mathrm{rms}})^2}}\times\exp\left[-\frac{(\Delta\beta_1 z-mT)^2}{2T^2(1+(z/L_D^{\mathrm{rms}})^2)}\right]. \tag{13}$$

where L_D^{rms} is defined as the dispersion length

corresponding to the root mean squared (RMS) value $\beta_2^{\rm rms} = \sqrt{({\beta_{2,ap}}^2 + {\beta_{2,bq}}^2)/2}$, namely $L_D^{\rm rms} = \sqrt{2} L_{Dap} L_{Dbq} / \sqrt{{L_{Dap}}^2 + {L_{Dbq}}^2}$.

The local collision integrals, for Gaussian and Nyquist pulses are represented, as a function of distance z (normalized to the rms dispersion length $L_D^{\rm rms}$), respectively in panels (a) and (b) of Fig. 1, for an ideal lossless fiber ($f_{bq}(z) = 1$). The lossless case is considered to isolate the effects of dispersion. The blue curves and markers denote the high-DGD regime. In particular, the blue curves present the local collision integrals for two specific collision indices ($m = 10, 90$). Each collision occurs over a short portion of the fiber, which scales with the walk-off length L_W, and is peaked around $z_m = mT/\Delta\beta_1 = mL_W$. The blue circles indicate the peak values of the other collision indices for two channels with identical dispersion lengths. In contrast, the blue crosses show the peak value for channels with different dispersion lengths. The Gaussian pulse (panel (a)) has the remarkable property to present a local collision integral only dependent on the rms dispersion length $L_D^{\rm rms}$ and not on the single channel dispersion lengths, as can be seen from Eq. (13). The collision peak position can be also interpreted by looking at zeros of the argument of the exponential function in the latter equation. For the Nyquist pulse (panel (b)), the sole dependence on $L_D^{\rm rms}$ holds for $z/L_D^{\rm rms} \lesssim 0.5$, but for larger values, a weak dependence on specific L_{Dap}, L_{Dbq} values exists. In practical systems, the weak dispersion case ($z/L_D^{\rm rms} \lesssim 0.5$) refers to low baud-rate (e.g. 10Gbaud) systems, while the larger baud-rates start to present such weak dependence on the difference between the GVD coefficients of the channels.

The red and green curves refer to the vanishing DGD regime. In this case, only one collision ($m = 0$) is presented, as it occurs along the entire fiber. Remarkably, as shown in the Appendix VI-B1, the same formulas used to calculate the peak values $I_m(z_m)$ in the high DGD regime can be used to calculate $I_m(z)$ in the vanishing DGD regime. For this reason, the blue circle and cross markers are overlapped to the red and green curves in Fig. 1.

The role of the difference in the dispersion properties of the two channels is further illustrated in Fig. 2, in the vanishing DGD regime. For the Gaussian pulse (panel (a)), $I_m(z)$ (Eq. (5)) depends only on the rms dispersion length $L_D^{\rm rms}$. For the Nyquist pulse (panel (b)), the calculation shows that such a property is no longer valid beyond $z/L_D^{\rm rms} \approx 0.5$. This fact will be taken into account when developing a general model to approximate the local collision integral in Sect. II-C.

We will now show that asymptotic approximations for the spatial integral X_{0mm} exist in the vanishing and high-DGD regimes.

A. High DGD regime

In the high DGD regime, $L_W \to 0$ and along the fiber there is a large number of collisions.

1) Gaussian pulse: Let us define a small dimensionless parameter $\varepsilon = L_W/L_D^{\rm rms}$; in Appendix VI-C2 it is shown that the integral (4) can be expanded as

$$X_{0mm} = \frac{1}{|\Delta\beta_1|}\left(a_1 + a_3\varepsilon^2 + ...\right), \tag{14}$$

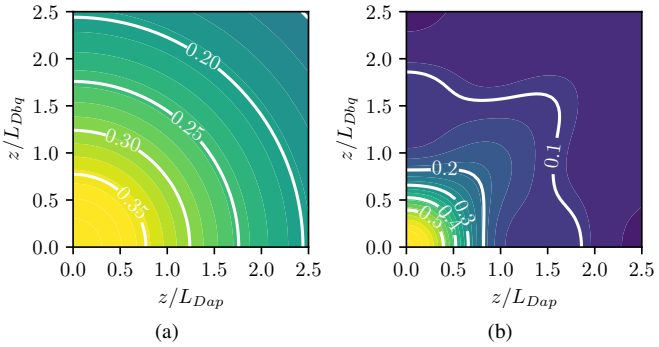

Fig. 2: The local collision integral $I_{m=0}\,T$ as a function of the distance normalized to the single channel dispersion lengths, for vanishing DGD: (a) Gaussian pulse. (b) Nyquist pulse.

with $a_1 = 1$, and $a_3 = 18$. In the following, we will keep only the zero-order term

$$X_{0mm} \approx \frac{1}{|\Delta\beta_1|}. \tag{15}$$

Note that this is the same expression as Ref. [15, Eq. (12)], under the substitution $\Delta\beta_1 = \beta_2\Omega$, where β_2 and Ω represent, within the assumptions of that work, the constant GVD coefficient and the channel frequency detuning. We remark that the result of Ref. [15] is derived under the hypothesis $z \gg L_D^{rms}$; however, our expression (15) is derived independently of the dispersion lengths. Moreover, the integral does not depend on the collision order m. This means that the integral of each blue curve of panel (a) of Fig. 1 is a constant depending on the DGD only.

Therefore the NLIN can be estimated by a sum of $M = L/L_W = L|\Delta\beta_1|/T$ equal collision integrals of value $X_{0mm}^2 \approx 1/|\Delta\beta_1|^2$, and so it scales inversely with respect to the DGD

$$\mathcal{N}^{\rm HI} = \frac{L}{T|\Delta\beta_1|}. \tag{16}$$

The latter expression is that obtained in Ref. [15], though we stress once more that the hypothesis of large channel dispersion is not required here. As an example, considering: $\Delta\beta_1 = 100$ ps/km, $L = 100$ km, $T = 100$ ps, we will have $M = \Delta\beta_1 L/T = 100$ collisions along the fiber link. If the RMS dispersion is $\beta_2 = -23$ ps^2/km, the expansion parameter is $\varepsilon = 0.002$, and the approximation holds.

2) Nyquist pulse: In the case of Nyquist pulses, the frequency integral (11) can be analytically computed for collisions occurring for $z_m/L_D^{\rm rms} \ll 1$ (see Appendix VI-C1):

$$I_m(z_m) \approx \frac{2}{3T}, \tag{17}$$

that correspond to the first blue cross marker on the left of Fig. 1 (b). As shown by the blue curve in the same panel, if $z_m/L_D^{\rm rms} \ll 1$, the spatial extension of the collision is also small, and the collision shape can be well approximated by a triangle. In this limit, the peak value is given by Eq. (17). At the same time, the base of the triangle can be approximated by considering the length over which the pulse of channel bp interacts significantly with that of channel ap (see (Eq. (5))).

In particular, the absolute value of the Nyquist pulse (see also Eq. (34) in the Appendix) presents secondary peaks for $t = \pm 3T/2$; so, approximating the pulsewidth by $3T$, the spatial extension of the collision is $3L_W$. Finally, the spatial integral (4) can be approximated by the triangle area, i.e. $X_{0mm} \approx 2/(3T) \cdot 3\,L_W/2 = 1/|\Delta\beta_1|$, i.e. the same value of the Gaussian pulse. For larger values of $z_m/L_D^{\rm rms}$, as can be noted in Fig. 1(b), the shape and peak value of the local collision integral of the Nyquist pulse change. Nonetheless numerical integrations of Eq. (11) show that $X_{0mm} \approx 1/|\Delta\beta_1|$ is valid like for the Gaussian case. The explanation of this behaviour can be given with reference to Eq. (12). In particular, the value of $|C(z,\omega)|$ is determined by the absolute value of the integral expression. As long as the walkoff lenght L_W is much smaller than L_{Dap} and L_{Dbq}, the application of the method of stationary phase implies that the values of the dispersion lenght do not play a relevant role in the calculation of the integral with respect to the walkoff length. So, summing for the number of collisions $M = L/L_W = L|\Delta\beta_1|/T$, we get the same noise value of the Gaussian pulse Eq. (16).

In panel (a) of Fig. 3, we demonstrate the validity of the approximations in the high DGD regime, by comparing formula (16) (orange line) to the full numerical integration (blue crosses for Gaussian pulses and green stars for Nyquist pulses).

To summarize the results of this subsection, in the high-DGD regime, channels' dispersion does not affect NLIN. An approximate formula (analytically derived for Gaussian and Nyquist pulses and therefore independent of the pulse shape) indicates that NLIN scales as $1/|\Delta\beta_1|$. In this regime, the direct evaluation of the integrals would require the calculation of a large number of time and space integrals; thus, our approximation enables a substantial simplification.

B. Low DGD regime

In the low DGD regime, $1/L_W \to 0$, there is a very low number of significant collisions and each collision occurs all along the fiber, as shown for example by the green and red curves in Fig. 1 that presents the local integral of the collision $m = 0$. Unlike the results presented in the previous section for the high-DGD case, NLIN depends on GVD in this regime. Asymptotic limits can still be found analytically for negligible dispersions $L \ll L_{Dap}$ and $L \ll L_{Dbq}$.

1) Gaussian pulse: This case can be analyzed in the limit $\varepsilon \to \infty$. For the most relevant collision, $m = 0$, we obtain (see Appendix VI-C2 for full calculation)

$$\mathcal{N}^{\rm LO,Gaussian} = \frac{1}{2\sqrt{\pi}} \left(\frac{L}{T}\right)^2 . \tag{18}$$

2) Nyquist pulse: Considering the collision $m = 0$, Eq. (11) reduces to the same form as Eq. (17) as soon as the substitution $z_m \to z$ is made (see also Eqs. (42 -43) in the Appendix). The local collision integral is approximated by $I_{m=0}(z) \approx 2/(3T)$ and so integrating over $z \in [0, L]$, and squaring, one gets:

$$\mathcal{N}^{\rm LO,Nyquist} = \frac{4}{9} \left(\frac{L}{T}\right)^2 . \tag{19}$$

The results show that when GV matching is achieved among channels, the NLIN tends to a constant that, for negligible dispersion, is proportional to the square of the ratio between the fiber length and the baudrate

$$\mathcal{N}^{\rm LO} = \xi \left(\frac{L}{T}\right)^2 . \tag{20}$$

through a pulse shape dependent coefficient ($\xi = 1/(2\sqrt{\pi})$ in the Gaussian case, and $\xi = 4/9$ in the Nyquist case). Note that the NLIN in the case of a low DGD is much larger than that for a high DGD. The low DGD approximations are shown in Fig. 3 (a), by the blue solid line for the Gaussian pulse, and the green dashed line for the Nyquist pulse. Blue crosses show the result of the full numerical integration for the Gaussian pulse, and the green stars that for the Nyquist pulse. The agreement between analytical and numerical results is very good.

When channel dispersion is no longer negligible, the local collision integral is no longer constant along z. In particular, as shown by the red and green curves in Fig. 1, the integral decreases with z, and so NLIN is expected to be smaller for any value of the DGD.

C. General model

In the previous subsections, we demonstrated that asymptotic limits exist for the NLIN as a function of the DGD. In this section, we present a general model to approximate NLIN for arbitrary values of the DGD, accounting also for channel dispersion and the power variation along the fiber (note that, so far, we have assumed $f_{bq}(z) = 1$). The behaviour presented in panel (a) of Fig. 3 can be, in fact, interpolated by the following function:

$$\mathcal{N}\left(\frac{L}{L_W}\right) = \mathcal{N}^{\circ}\left(1 + \left(\frac{1}{\Lambda}\frac{L}{L_W}\right)^{1/\eta}\right)^{-\eta} \tag{21}$$

with fitting parameters $\mathcal{N}^{\circ}$, Λ, and η. The fitting parameters have the following interpretation: $\mathcal{N}^{\circ}$ corresponds to the asymptotic constant value in the limit of low DGD; Λ sets the transition DGD between the asymptotic regimes, and η regulates the smoothness of the transition. The asymptotic power law behaviour for high DGD is fixed to $\mathcal{N} \sim \Delta\beta_1^{-1}$.

In the regime where $L \ll L_{Dap}$ and $L \ll L_{Dbq}$, $\mathcal{N}^{\circ}$ and Λ can be estimated starting from the asymptotic values given in Eq. (16), (18) and (19), or fitted to numerical data. The value of η can instead only be obtained by fitting to numerical data. For example, by fixing: $L = 70$ km, $T = 100$ ps, $\beta_{2,ap} = \beta_{2,bq} = -20$ ps^2/km, we have $L/L_D^{\rm rms} \approx 0.14$, showing that low dispersion approximation is valid. The parameters relative to the results of panel (a) of Fig. 3 are: $\mathcal{N}^{\circ} = 2.516$ km^2/ps^2, $\Lambda = 2.12$, and $\eta = 0.921$ for the Nyquist case, and $\mathcal{N}^{\circ} = 1.506$ km^2/ps^2, $\Lambda = 3.29$, and $\eta = 0.493$ for the Gaussian case.

We now show that corrections to the parameters of (21) can be introduced to account for the signal power profile evolution and the chromatic dispersion, which, remarkably, can be treated on a similar footing. In fact, both effects lead to a z dependent integrand function in Eq. 5 and so to a decrease of the NLIN with respect to the calculations carried out so far.

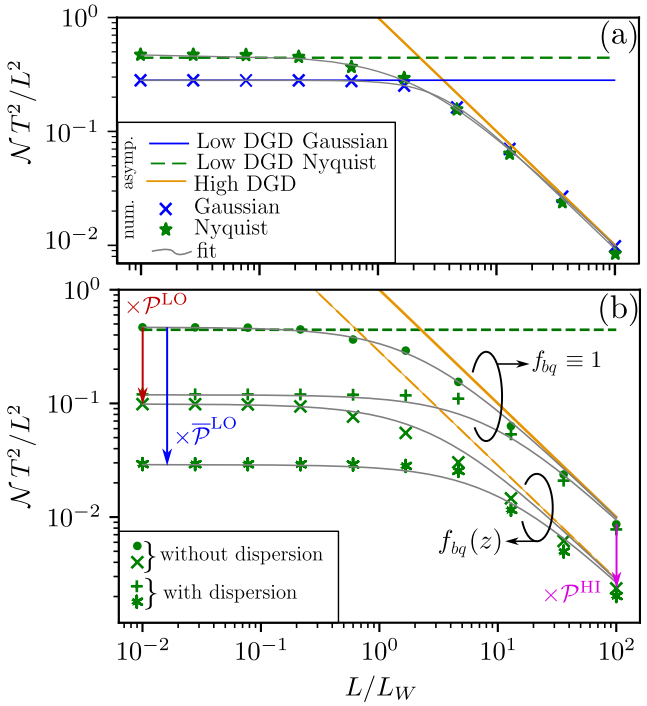

Fig. 3: Channel-pair NLIN, $\mathcal{N}$ as a function of the DGD. The results for the Nyquist and Gaussian pulses are respectively shown in green and blue. In panel (a), the ideal case with $f_{bq}(z)=1$ and negligible dispersion is presented. In particular, the green dashed and blue solid lines represent the low DGD approximations given by Eqs. (18) and (19). The orange line corresponds to the high DGD approximation for both Nyquist and Gaussian pulses (Eq. (15)). The green asterisk and blue cross markers are the result of the numerical solutions. Gray solid curves are the result of the general model, Eq. (21). In panel (b), only for the Nyquist pulse, the model corrections, due to losses/distributed amplification and the dispersion, are shown. The function $f_{bq}(z)$ used corresponds to the cyan curve of Fig. 7, taken as a representative curve of the counter-propagating amplification case. The dashed green and solid orange lines are the same of panel (a). The dashed orange line represents the high DGD limit for a case of a lossy and Raman amplified link. Green markers present the result of numerical calculations: dot (negligible dispersion and constant power); plus (dispersive channel and constant power); cross (negligible dispersion and varying power); asterisk (dispersive channel and varying power). Dispersive channel case has $L/L_{Dap}=2.3$ and $L/L_{Dbq}=0$. Gray curves are the fitting equation modified by the calculated correction factors $\mathcal{P}^{\mathrm{HI}},\mathcal{P}^{\mathrm{LO}},\overline{\mathcal{P}}^{\mathrm{LO}}$.

First, fiber losses and Raman distributed amplification are included in the model. A feature of the evolution function $f_{bq}(z)$ for fiber lengths of a few tens of km is that its amplitude can easily span about an order of magnitude, because of the exponential decay (due to losses) and exponential growth (due to the pumps) [19]. In a WDM system each channel will experience a different evolution, as it will be shown also in the case study of Sect. III (see Fig. 7).

In the high DGD regime, the local collision integral peaks around the values z_m (blue curves in Fig. 1). If we approximate $f_{bq}(z)\approx f_{bq}(z_m)$ in the vicinity of the pulse collision peak, the total noise will be a sum of the form:

$$\mathcal{N}^{\mathrm{HI}}\approx\sum_{m=0}^{M}\frac{f_{bq}(z_m)^2}{|\Delta\beta_1|^2}\,. \tag{22}$$

However, if we consider that the collision peaks are separated by L_W, which for high DGD is much smaller than L, we can map the discrete sum over m into the integral:

$$\mathcal{N}^{\mathrm{HI}}\approx\frac{L}{T|\Delta\beta_1|}\,\mathcal{P}^{\mathrm{HI}}\,, \tag{23}$$

where

$$\mathcal{P}^{\mathrm{HI}}=\frac{1}{L}\int_0^L dz\ {f_{bq}}^2(z) \tag{24}$$

is a correction factor with respect to the high DGD limit analytically found for $f_{bq}(z)=1$ in Eq. (16). This integral can be evaluated numerically efficiently, since $f_{bq}(z)$ is smooth across all channels. No further correction is required for accounting dispersion in the high DGD regime. We checked this for Nyquist pulses, and the results are presented in Fig. 3(b). The orange solid line represents Eq. (16) (for $f_{bq}(z)=1$) while the orange dashed line Eq. (23) for an evolving signal. In particular, we used the cyan curve in Fig. 7, which represents the evolution of $f_{bq}(z)$ for a counter-propagating pump Raman-amplified signal, including fiber losses. Note that the selected evolution is very typical of this amplification scheme. Noticeably, the dashed line is a translation of the solid line by the correction term $\mathcal{P}^{\mathrm{HI}}$ (magenta arrow in figure). Full numerical integrations with negligible (green crosses) and non-negligible dispersion (green asterisks) confirm that for high DGD, the NLIN is weakly dispersion dependent, and the approximation used is very good.

For the low-DGD regime, with negligible dispersion, the dominant contribution is given by $m=0$ collision, and the local collision integral $I_0(z)$, i.e. the integral (5) with $m=0$, is approximately constant, as derived in the previous subsection. Therefore, from Eq. (5), one gets:

$$\mathcal{N}^{\mathrm{LO}}\approx\xi\,\frac{L^2}{T^2}\,\mathcal{P}^{\mathrm{LO}}=\mathcal{N}^{\circ}\,\mathcal{P}^{\mathrm{LO}}\,, \tag{25}$$

where the correction factor is:

$$\mathcal{P}^{\mathrm{LO}}=\left(\frac{1}{L}\int_0^L dz\ f_{bq}(z)\right)^2\,. \tag{26}$$

Eqs. (23-26) provide some direct insight into the effect of including a distributed amplification and losses. Due to the Hölder inequality (see section VI-C3 of the Appendix) $1>\mathcal{P}^{\mathrm{HI}}\geq\mathcal{P}^{\mathrm{LO}}$, i.e., the decrease in NLIN due to different power evolution along the fiber is more significant in the case of low DGD compared to that of high DGD. Therefore, channels with similar GV are expected to be less sensitive to noise degradation induced by changes in pumping schemes. Numerical solutions (green crosses in Fig. 3(b)) confirm that in the low DGD regime, a translation of the curve by an amount equal to Eq. (26) (red arrow) is a very good approximation.

The correction factors (24) and (26) can be included in the estimation of the NLIN of Eq. (21) through the modification of the parameters. Given the pulse-shape dependent fitting parameters $(\mathcal{N}^{\circ}, \Lambda, \eta)$, obtained under the assumption of constant power evolution and negligible dispersion, the corrected fit parameters will be $(\mathcal{N}^{\circ}\mathcal{P}^{\mathrm{LO}}, \Lambda\mathcal{P}^{\mathrm{LO}}/\mathcal{P}^{\mathrm{HI}}, \eta)$.

In the low DGD regime, when dispersion is non negligible, both f_{bq} and I_m are z dependent. Moreover, though $m = 0$ remains the most relevant term, the tails of higher order (up to a value m_T to be evaluated) collisions are not always negligible. Following Eq. (25), the correction factor accounting for dispersion can be calculated as

$$\overline{\mathcal{P}}^{\mathrm{LO}} = \frac{\mathcal{N}^{\mathrm{LO}}}{\mathcal{N}^{\circ}} \approx \frac{1}{\mathcal{N}^{\circ}} \left(\int_0^L dz \; f_{bq}(z) \sum_{m=-m_T}^{m_T} I_m(z) \right)^2 . \tag{27}$$

Eq. (27) shows that dispersion effects can be evaluated first, by calculating the summation term. In Fig. 4, the local collision integrals $I_m, m = 1, 2, 3$ are shown as a function of the distance normalized to dispersion lengths $(z/L_{Dap}, z/L_{Dbq})$ for the Nyquist pulse ($m = 0$ is shown in Fig. 2). Note that: a) the $m = 0$ integral decreases monotonically with z; b) higher order integrals are of smaller and smaller amplitude as m increases (actually down to being negligible for $m > 3$); c) higher order integrals present a maximum that is achieved at higher and higher values of the normalized distance, as m increases. All these characteristics ultimately lead the sum to vary slowly and weakly with respect to $f_{bq}(z)$ as a function of z. For this reason, the dynamic range of values of Eq. (27) is dominated by the function $f_{bq}(z)$, the dispersion being a weak perturbation along z.

Therefore, an interpolation can be used to calculate the correction factor $\overline{\mathcal{P}}^{\mathrm{LO}}$ in the presence of dispersion from values calculated when dispersion is negligible. For negligible dispersion, through Eq. (25), the correction factors $(\mathcal{P}^{\mathrm{LO}}_{\max}, \mathcal{P}^{\mathrm{LO}}_{\min})$ for the WDM channels presenting the maximum and minimum average power evolution can be evaluated. Then, the same correction factors are numerically calculated using Eq. (27), so fully accounting for dispersion. Let us call the resulting values $\overline{\mathcal{P}}^{\mathrm{LO}}_{\max}, \overline{\mathcal{P}}^{\mathrm{LO}}_{\min}$ and note that the latter integrals must be calculated only once. For any power profile evolution, the correction factor accounting for dispersion $\overline{\mathcal{P}}^{\mathrm{LO}}$, can be simply linearly interpolated as follows

$$\frac{\overline{\mathcal{P}}^{\mathrm{LO}} - \overline{\mathcal{P}}^{\mathrm{LO}}_{\min}}{\overline{\mathcal{P}}^{\mathrm{LO}}_{\max} - \overline{\mathcal{P}}^{\mathrm{LO}}_{\min}} \approx \frac{\mathcal{P}^{\mathrm{LO}} - \mathcal{P}^{\mathrm{LO}}_{\min}}{\mathcal{P}^{\mathrm{LO}}_{\max} - \mathcal{P}^{\mathrm{LO}}_{\min}}, \tag{28}$$

The relative interpolation error will be calculated in the case study, presented in Sect. III and its maximum is in the order of 2%.

In panel (b) of Fig. 3, the low DGD correction for non-negligible dispersion is shown by the blue arrow, for the same power evolution $f_{bq}(z)$ used before. The agreement with numerical simulations (indicated by asterisk markers) is excellent.

To summarize, in this section we derived analytical asymptotic expressions of NLIN for low and high DGD, under the ideal hypotheses of negligible dispersion and negligible losses/amplification. A general equation (Eq. (21)) satisfying asymptotic expressions can be used to predict NLIN for any value of the DGD among WDM channels. Dispersion and losses/amplification can be accounted for in the modelling equation using corrective factors derived from the evaluation of a reduced set of integrals.

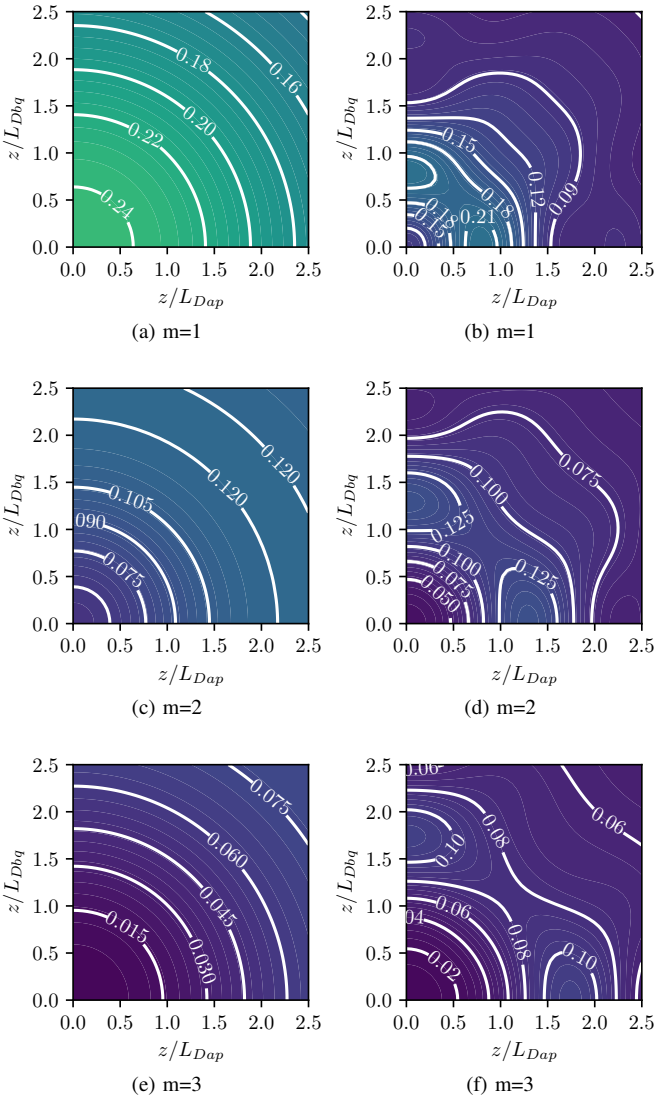

Fig. 4: Local collision integral $I_m \cdot T$, for collision orders $m = 1, 2, 3$, as a function of the position z normalized to single channel dispersion lengths, for vanishing DGD. The plots on the left are for Gaussian pulses; those on the right are for the Nyquist pulse. Parameters are the same as in Fig. 1.

The general model was developed for Gaussian and Nyquist pulses, however it can be applied to any pulse shape by proper calculation of the modelling parameters.

We summarize the method in pseudocode form. The computation of the total system NLIN is organized into three main steps (S1–S3), indicated on the pseudocode.

In Step S1, the ideal fitting parameters—i.e., those corresponding to a dispersionless system with a flat power profile—are determined using the numerical data and asymptotic ex-

pansions discussed in Sect. II.A and B. Once the pulse shape is fixed, these parameters are obtained by numerically evaluating $\mathcal{N}$ for ten different values of the DGD, spanning a range sufficient to capture all regimes. These parameters depend only on the pulse type and therefore serve as a baseline for all subsequent calculations. Step S2 performs a precomputation of channel-specific correction factors, which account for the actual signal power profiles and dispersion properties of the system. In particular, all signal power profiles are inserted in Eq. (26) and Eq. (24). For the maximum and minimum power profiles, Eq. (27) is evaluated over a set of 20 dispersion values (forming a 20×20 grid), up to a truncation index m_T chosen to ensure convergence for the maximum dispersion considered.

Finally, in Step S3, the correction factors are applied to the ideal fitting parameters. In particular, for the low DGD case $\overline{\mathcal{P}}^{\mathrm{LO}}_{\min}$ and $\overline{\mathcal{P}}^{\mathrm{LO}}_{\max}$ are obtained by interpolating the precomputed grid at the specific dispersion lengths of the channel pair under consideration. The proportional scaling of Eq. (28) is then performed based on the corresponding channel-specific signal power profile. Using the corrected fitting formula, Eq. (21), the noise coefficient $\mathcal{N}$ is evaluated, and the resulting NLIN contribution is added to the noise of the channel under analysis. This procedure is repeated for all channel pairs, and the contribution of each interfering channel is stored.

Finally, we note that the signal power profiles are treated as input data, since their calculation is not the primary focus of the present work.

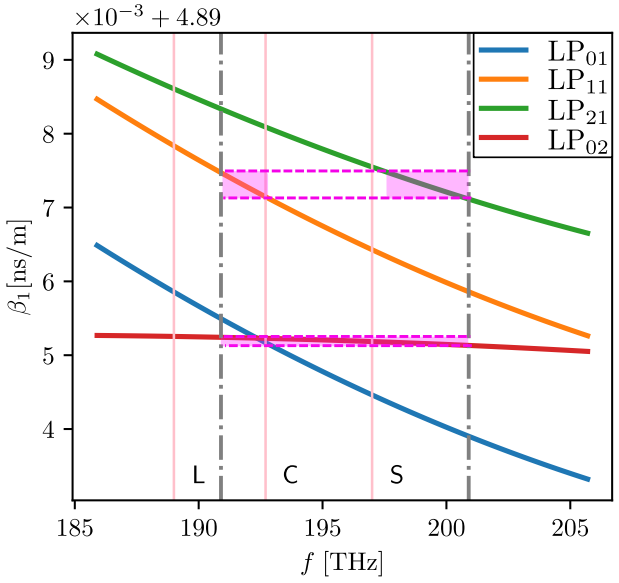

Fig. 5: Group delay per unit length for the FMF under study. The pink-shaded boxes highlight the bands over which WDM channels can exhibit GV matching. The gray vertical dashed-dotted lines represent the limits of the band considered in the case study.

III. Case Study

We applied the model to a FMF case study, to show its effectiveness and to clarify how dispersion properties affects

Pseudocode 1: Estimation of NLIN for every channel. Steps S1–S3 are described in the text.

Inputs : Channels of interest $\mathcal{A} = \{(a,p)\}$ for all mode groups a and WDM frequencies p
Link parameters: $L, T, \{\beta_2\}_{\mathcal{A}}, \{\Delta\beta_1\}_{\mathcal{A}\times\mathcal{A}}, \gamma, \kappa_{ab}, N_b$
$(L/L_D)_{\max} \equiv \max_{(a,p)} L\beta_{2,ap}/T^2$, heuristic m_T
Symbol statistics: $E_2 = \langle|b_0|^2\rangle$, $E_4 = \langle|b_0|^4\rangle$
Set of power profiles: $f_{bq}(z) = P_{bq}(z)/P_{bq}(0)$
Extremal power profiles: $f_{B,\min}(z) = \min_{bq} f_{bq}(z)$, $f_{B,\max}(z) = \max_{bq} f_{bq}(z)$
Pulse type
Output: List of $\{\sigma^2_{\mathrm{XPM}}\}_{\mathcal{A}}$
begin
S1 $(\mathcal{N}^\circ, \Lambda, \eta)_{\text{ideal}} \leftarrow$ fitting Eq. (21) on pulse-specific dispersionless data with flat power profile (analytic or numeric)
S2 **foreach** $(a,p) \in \mathcal{A}$ **do**
$\{\mathcal{P}^{\mathrm{LO}}\}_{\mathcal{A}} \leftarrow$ Eq. (26)
$\{\mathcal{P}^{\mathrm{HI}}\}_{\mathcal{A}} \leftarrow$ Eq. (24)
$\overline{\mathcal{P}}^{\mathrm{LO}}_{\min,\max} \leftarrow$ Eq. (27), with $I_m(z)$ computed up to $(L/L_D)_{\max}$ and $|m| \le m_T$
S3 **foreach** $(a,p) \in \mathcal{A}$ **do**
$\mathcal{B} \leftarrow \mathcal{A} \setminus \{(a,p)\}$
$\sigma^2_{\mathrm{XPM},ap} \leftarrow 0$
foreach $(b,q) \in \mathcal{B}$ **do**
$\mathcal{P}^{\mathrm{HI}} \leftarrow$ from precomputed $\{\mathcal{P}^{\mathrm{HI}}\}_{\mathcal{A}}$
$\overline{\mathcal{P}}^{\mathrm{LO}} \leftarrow$ from precomputed $\{\mathcal{P}^{\mathrm{LO}}\}_{\mathcal{A}}$, $\overline{\mathcal{P}}^{\mathrm{LO}}_{\min,\max}$ and Eq. (28)
$(\mathcal{N}^\circ, \Lambda, \eta)_{\text{corr}} \leftarrow (\mathcal{N}^\circ\, \overline{\mathcal{P}}^{\mathrm{LO}}, \Lambda\, \overline{\mathcal{P}}^{\mathrm{LO}}/\mathcal{P}^{\mathrm{HI}}, \eta)$
$\mathcal{N} \leftarrow$ Eq. (21) with $(\mathcal{N}^\circ, \Lambda, \eta)_{\text{corr}}$ and $\Delta\beta_{1,apbq}$
if $b = a$ **then**
$C \leftarrow P^3T^2\gamma^2\kappa^2_{aa}\left[(E_4/E_2^2)(2N_a+3)-4\right]\mathcal{N}$ (Eq. (6))
else
$C \leftarrow 2N_bP^3T^2\gamma^2\kappa^2_{ab}\left[(E_4/E_2^2)-1\right]\mathcal{N}$ (Eq. (7))
$\sigma^2_{\mathrm{XPM},ap} \leftarrow \sigma^2_{\mathrm{XPM},ap} + C$
return $\{\sigma^2_{\mathrm{XPM}}\}_{\mathcal{A}}$

NLIN in this type of fibers. The fiber under consideration supports 4 mode groups: LP_{01}, LP_{11}, LP_{21}, LP_{02}, with a number N of spatial modes of $1, 2, 2, 1$, respectively. The fiber and system parameters are specified in Tab. I.

The group delay per unit length of the FMF modes, calculated through numerical solutions (Finite-Element-Method, FEM), are plotted in Fig. 5; the FMF design parameters were selected so to present dispersion features that enable to observe all the significant dispersion regimes, and are very close to those found in the literature [26, 27]. We consider 200 WDM channels per group mode, over the S, C, and part of the L band (see dashed-dot lines in Fig. 5). WDM bands presenting low DGD are highlighted by the pink shadowed regions. There

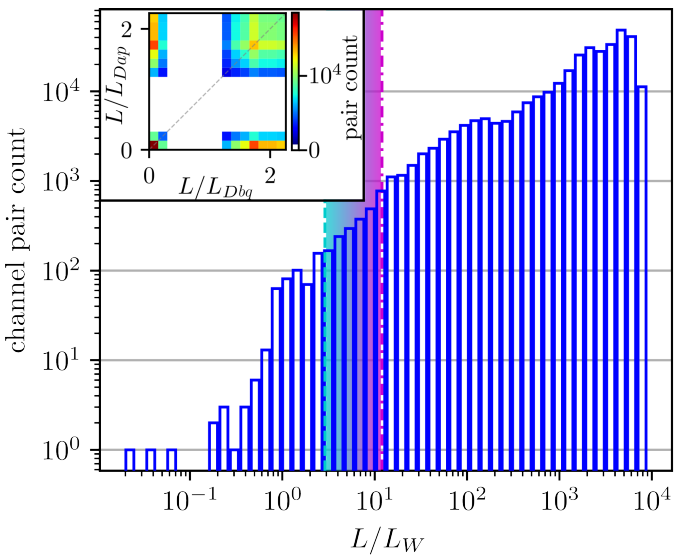

Fig. 6: Distribution of the channel pairs as a function of the DGD for the FMF under study. The shaded box highlights the estimated range of the fitting parameter Λ . It spans from the minimum value (cyan dashed line) to the maximum value (magenta dashed-dotted line) over all the channel pairs.It determines the transition between the high- and low-DGD regimes. In the inset, the distribution of the chromatic dispersion of the channel pairs is also shown.

TABLE I: Fiber and system parameters

Parameter	**Value**
Effective area (A_{eff})	220.7 µm^2 (FMF)
	80 µm^2 (SMF)
Kerr coefficient (n_2, see [25])	2.7×10^{-20} m^{2}W^{-1}
Raman coefficient (g_R, see [25])	7×10^{-14} m W^{-1}
Attenuation (at band center)	0.19 dB km^{-1}
Fiber Length (L)	70 km
Baud rate ($1/T$)	33 Gbaud
Modulation format	64-QAM
Channel spacing (Ω_0)	$2\pi\times$ 50 GHz
WDM carriers per group mode (N)	200
WDM band	190.94 THz to 200.94 THz
	1492.0 nm to 1570.1 nm (S+C+L)

are essentially three possible scenarios to get low DGD (i.e. GV matching). The first is that of channels on different modes separated by large frequency shifts; this is the case of large and low frequency channels on mode groups LP_{21}, LP_{11} in Fig. 5. The second is that of channels of about the same frequency on different mode groups; this is the case of LP_{01}, LP_{02} for which the group delay curves cross at a certain frequency. The third is the case of low dispersion mode groups, like LP_{02} for which β_2 spans from -2 to 0 ps^2/km, and so low DGD is found over many WDM channels propagating in the same mode.

In Fig. 6, the statistics of the DGD distribution for all the possible unsorted channel pairs (6.4×10^5) of the selected WDM grid are shown. The dispersion data are calculated from a direct interpolation of the dispersion data obtained from the FEM simulation with a order 3 polynomial, that is

TABLE II: Optimized pump launch powers at the end of the link.

FMF					
Pump #	**λ [nm]**	**Launch Power [dBm]**			
		LP_{01}	LP_{11}	LP_{02}	LP_{21}
1	1384.26	-1.9	-2.4	31.9	29.2
2	1397.61	0.7	3.4	28.0	24.8
3	1413.34	-9.0	-6.1	23.9	20.6
4	1429.42	-20.6	-16.7	21.8	18.7
5	1455.33	-29.6	-25.1	11.1	14.4
6	1457.97	-32.2	-29.0	17.0	8.3
SMF					
Pump #	**λ (nm)**	**Launch Power (dBm)**			
1	1380.98	26.9			
2	1394.51	27.6			
3	1409.75	24.8			
4	1426.05	21.4			
5	1454.22	12.3			
6	1456.08	19.2			

then differentiated and evaluated for the central frequencies of the WDM grid (see also Fig. 5 for the group delay). In the figure, the shaded box is delimited by the estimated maximum and minimum values (over all possible dispersion lengths and $f(z)$ evolution functions) of the fitting parameter Λ. The values, as discussed in Sect. II-C are calculated using the baseline Λ computed by numerical integration (step S1 in the pseudocode 1). Then, the value of Λ is corrected using the expressions Eq. (24), and Eq. (26) approximated using Eq. (28). The minimum value (cyan dashed line) is estimated by considering the maximum power profile and negligible dispersion, whereas to the minimum value (magenta dashed-dotted line), is estimated taking the minimum value of f and the most dispersive channels ($L/L_D = 2.3$). Channel pairs above the maximum are mostly in the high DGD regime, while those below the minimum are mostly in the low DGD regime. Channel pairs are therefore mostly in the high DGD regime, while a much more limited number of pairs show a low DGD. The statistics regarding the dispersion values are shown in the inset of Fig. 6. As can be seen, the system under consideration includes both high- and low-dispersion channel pairs. The proposed model is used to estimate the NLIN accurately (as demonstrated in Fig. 3 (b)).

We designed a counterpropagating Raman-amplified system with 6 pump wavelengths. The pump input powers and wavelengths are selected to equalize the gain at the end of the span, through the method defined in [28]. The resulting launch powers are reported in Tab. II, for both the SMF and the FMF case. However, let us remark that our calculations can be performed for any type of power distributions along the fiber. In the FMF case, the estimated power evolutions along the fiber $f(z)$, are presented (green/yellow curves) in Fig. 7. As it can be noticed the power evolution dynamics for all channels is bounded between a curve of maximum (dashed-dotted cyan curve) and a curve of minimum (dashed-dotted magenta curve), that can be used to estimate the corrections $\mathcal{P}^{\text{LO}}_{\text{max}}, \mathcal{P}^{\text{LO}}_{\text{min}}$ as indicated

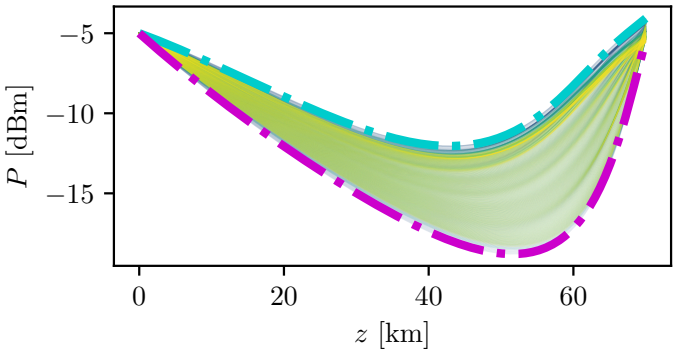

Fig. 7: Power evolution profiles $f(z)$ of all WDM channels, for the Raman counter-propagating pumping scheme of the case study. The profiles corresponding to the maximum and minimum interfering power levels are highlighted by the cyan and magenta dashed-dotted curves, respectively. Different shades of green and yellow denote channels pertaining to different mode groups.

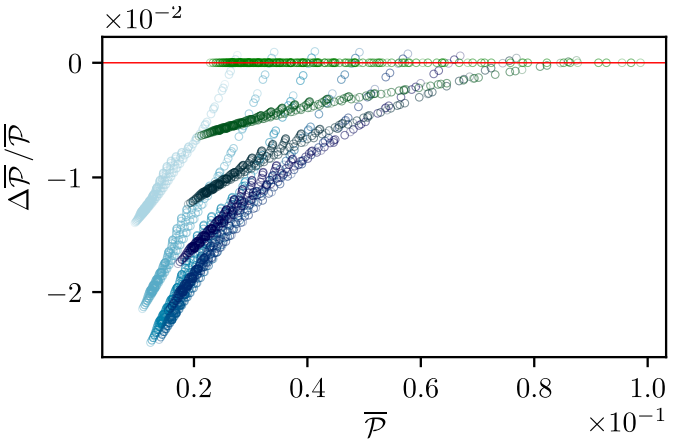

Fig. 8: Relative error $\Delta\overline{\mathcal{P}}/\overline{\mathcal{P}}$ between the Raman correction factors numerically computed (Eq. (27)) and approximate (Eq. (28)). Each marker corresponds to one computed value, plotted as a function of $\overline{\mathcal{P}}$. Shades of blue and green correspond to different curves f, the light blue being closer to the minimum power profile and the green being closer to the maximum. Markers with the same color are taken for different values of dispersion such that $L/L_D \in [0, 2.3]$ (see also the inset of Fig. 6.

in the previous section. In Fig. 8 we present the relative error of the estimated correction terms. The relative error $\Delta\overline{\mathcal{P}}/\overline{\mathcal{P}}$ (the label LO is dropped for compactness) is estimated by computing the Raman correction factor using the definition of $\overline{\mathcal{P}}$ in Eq. (27) and using the indirect approximated relation Eq. (28), and taking their difference $\Delta\overline{\mathcal{P}}$. The results confirm the validity of the approximation taken in Eq. (28). Crucially, this justify the usage of the fast approximated calculation of such correction on the sole basis of the precomputed data for the maximum and minimum profiles, along with the simpler integrals of Eq. (26).

The computation of the total NLIN for each WDM channel in the FMF is performed by calculating $\sum_{(bq)\neq(ap)} \mathcal{N}$, where the sum is extended over all the interfering channels. The results, presented in Fig. 9, enable insight into the various effects contributing to the NLIN.

By setting the overlap integrals $\kappa_{ab} = 0, a \neq b$ we can first quantify the strength of intramodal NLIN. This correponds to the use of Eq. (6) only, not considering the contribution of Eq. (7); the results are shown in Fig. 9 by the dashed-dotted curves for each mode group. The general trend is similar to that observed in SMFs [24]: channels at the grid center exhibit higher NLIN due to the presence of a larger number of significantly interfering channels. Moreover, the low dispersion mode (LP_{02}, dashed-dotted red curve) is the most affected by NLIN. This is due to the low DGD among the WDM channels, which is the condition yielding the highest interference noise, as shown by the result of Fig. 3.

The solid curves show the total NLIN resulting from both intermodal and intramodal interference. NLIN is slightly larger across all modes due to the increase in the number of interfering channels, and, more notably, the NLIN enhancement for low DGD (pink-shaded regions in Fig. 5) is clearly observable in two of the three previously mentioned scenarios. For channels on different modes separated by large frequency shifts (LP_{21}, LP_{11}, orange and green solid curves), the NLIN is clearly larger for all channel pairs close to GV matching. In the case of group delay curves crossing (Fig. 5), the NLIN is even more significantly enhanced for the mode with large GVD (LP_{01}, blue solid curve). Finally, the low-GVD mode (LP_{02}) channels experience only a slight increase due to the additional channels.

We finally compare the NLIN of a Raman-amplified system based on the FMF to that based on a SMF. The comparison is performed under the assumption of the same total bit rate per mode, i.e., assuming that the SMF is loaded with the same WDM setup as one of the group modes of the FMF. The dispersion of the fundamental mode of the SMF is equal to that of the LP_{01} (fundamental) group mode of the FMF. The Raman amplification scheme is optimized with the same method, with 6 counterpropagating pumps yielding comparable average ASE noise, as shown in Tab. III. The NLIN in both systems is computed for −5 dBm per-channel launch power, and the results are presented in Fig. 9 as the dashed gray curve. The WDM channels of the FMF exhibit better performance in terms of NLIN compared to the SMF, as summarized in Tab. III; however, NLIN can vary substantially among group modes and across wavelengths within the same group mode due to the different dispersion properties.

Finally, although not considered in this analysis, it is worth noting the potential contribution of FWM. As calculated in [18] (see Fig. 5 of that paper), the strength of NLIN FWM is comparable to that due to XPM when both group velocity and phase matching are achieved, the latter condition being hard to achieve in general. Therefore, for the largest majority of group modes and WDM channels, FWM will give a negligible contribution. The channels of the low dispersion mode group LP_{02} might indeed present an additional penalty due to FWM.

In order to assess the computational complexity of the proposed method, we compare the number of evaluations of the function $I_m(z)$ defined in Eq. (5), which must be computed

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2026.3680666

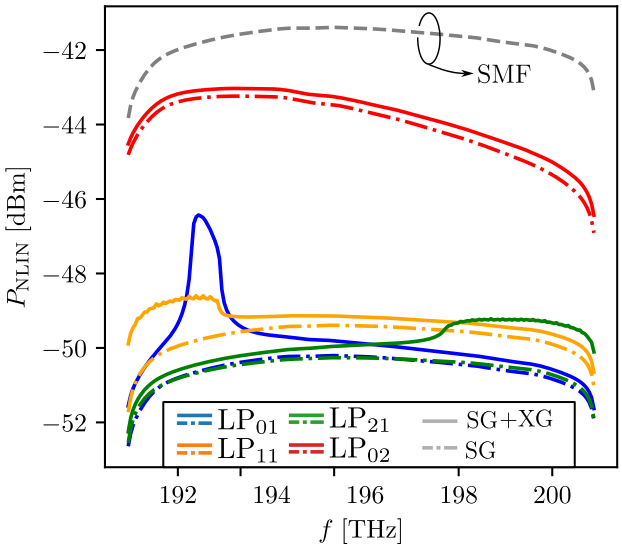

Fig. 9: NLIN power for the FMF case study with parameters in Tab. I. The dashed-dotted curves are the NLIN due to intragroup interactions only (SG); the solid curves are the NLIN due to both inter- and intragroup interactions (SG+XG). Color codes for mode groups are: blue LP_{01}, orange LP_{11}, green LP_{21}, red LP_{02}. The gray dashed curve represents NLIN for the SMF system.

TABLE III: Comparison between SMF and FMF systems. Launch power is −5 dBm.

Parameter	**FMF** [dBm]	**SMF** [dBm]
Average NLIN	−47.3	−41.7
Median NLIN	−49.3	−41.6
Maximum NLIN	−43.0	−41.4
Minimum NLIN	−52.3	−43.8
Average ASE	−29.9	−30.1

through numerical integration for a set of values of z, and can be regarded as the most computationally intensive operation required in the NLIN calculation. We consider a worst-case scenario in which the computational cost of evaluating $I_m(z)$ is independent of both the index m and the dispersion and DGD characteristics of the channels under consideration.

Breaking down the analysis according to the algorithmic steps (see also pseudocode 1), the proposed method requires the following operations. In Step S1, 8 sampling points of $\mathcal{N}$ are evaluated for different DGDs under $f(z) \equiv 1$ and no dispersion in order to estimate the coefficients of the ideal fit. This procedure requires the evaluation of about 200 functions $I_m(z)$. In Step S2, $I_m(z)$ are evaluated for $m = 0, 1, 2, 3$ (Eq. 5) in the low-DGD limit over a sampling grid with varying dispersion. This corresponds to $(20 \text{ dispersion sampling points})^2 \times 4$ values of m, but can be reduced to an equivalent of approximately 20×4 evaluations by using the scaling properties of I_m in the low DGD regime, for which the function depends only on the ratios z/L_{Dap} and z/L_{Dbq}. Another set of operations consists of the calculation of the integrals in Eqs. (24) and (26) for (200 channels $\times$ 4 mode groups), however these are single integrals and therefore substantially less demanding. Step S3 requires the evaluation of the fitting expressions with the appropriate corrections and the summation of the resulting contributions. By contrast, a direct-integration approach would require calculating $I_m(z)$, $M = L/L_W$ times for all channel pairs, 6.4×10^5. Taking into account the distribution of DGDs in the system under study (see Fig. 6), the number of relevant values of M introduces an additional multiplicative factor of about 10^3.

We report the total computation time measured on typical hardware, namely a laptop equipped with an Intel i7-12700H CPU and 16 GB of RAM. Multiprocessing is employed for Steps S1 and S2 of the algorithm, which are the most computationally demanding. The measured execution times (averaging over 5 calculations) are as follows: Step S1 ~ 50 s, Step S2 ~ 86 s, and Step S3 ~ 20 s. As discussed in Sect. II, Step S1 is largely system-independent, since neither modal nor chromatic dispersion nor the signal power profile are taken into account. Similarly, the results obtained in Step S2 can be reused for systems with lower chromatic dispersion, as well as for systems whose signal power profiles can be reasonably approximated by the ones considered here.

A direct timing comparison with the naïve integration approach is impractical because the total execution time would be prohibitively long. We estimate the speedup on the basis of the number of the computed collisions, that in our method is decreased by a factor of at least $\times 10^6$. The great advantage of our approach is based on the fact that we sample only a few values of NLIN in relevant regimes, and extrapolate for any other value of DGD, dispersion, and power evolution.

IV. Conclusions

In this work, we extended the time-domain model to evaluate NLIN in distributed Raman-amplified, spatial-division multiplexed systems based on few-mode fibers. In such systems, the increased number of WDM channels, together with the large DGD between them can make the number of pulse collisions to be evaluated significantly larger than in SMF systems, making the numerical calculation of the collision integrals highly computationally demanding.

We first show that asymptotic approximations of the integrals for low and high DGD can be analytically derived for Gaussian and Nyquist pulses under ideal conditions (negligible dispersion and no power evolution along the fiber). The low-DGD regime exhibits higher NLIN values than the high-DGD regime. These results enable the definition of a universal fitting model that matches numerical simulations accurately across all DGD values. The model can then be extended to account for channel power evolution (distributed Raman amplification and/or losses) and larger dispersion by calculating simple correction factors, whose evaluation requires a very reduced amount of numerical simulations. Although tested on Gaussian and Nyquist pulses, the methodology can be extended to other pulse formats by numerically determining their asymptotic limits. Moreover, the method can also be applied to single-

mode fiber systems, where, in particular, it would be advantageous for modeling multi-band transmission including the O and E bands.

As an application, we analyze a Raman-amplified system for a FMF supporting four mode groups. The fiber modes and dispersion properties were evaluated numerically. The comparison with a SMF-based system, at the same baud rate, shows that the NLIN is smaller for all channels of the FMF, confirming the main advantage of spatial division multiplexing. However, in FMFs, the NLIN distribution can vary substantially, and several channels can exhibit significantly larger NLIN values. Intermode NLIN can highly affect channels at wavelengths close to mode group velocity matching, while intramode NLIN can highly affect low-dispersion modes. Future work may exploit this fast time-domain model for the optimization of distributed Raman amplification with respect to NLIN. As discussed, NLIN is significantly enhanced in mode groups and wavelength regions where group-velocity matching occurs due to stronger inter-group interactions, indicating that Raman pump configurations optimized solely for gain equalization may be highly suboptimal from an NLIN perspective in FMF systems. The proposed model therefore offers a promising tool for Raman pump power and wavelength optimization that explicitly accounts for mode- and wavelength-dependent NLIN [21, 29]. Moreover, since the model is inherently modulation-format dependent, it may also be relevant for the optimization of probabilistic shaping aimed at maximizing generalized mutual information in nonlinear fiber channels [30].

V. Acknowledgements

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE0000001 - program "RESTART").

VI. Appendix

A. Derivation of the noise variance equations

The expressions for the noise variance given in Eqs. (6) and (7) are derived from Ref. [18, Eqs. (19) and (28)] as follows. We first consider the SGXPM case. The starting point is the fundamental relation given in Eq. (15) of Ref. [18],

$$\sigma^2_{\mathrm{SGXPM}} = \frac{(2N_a+1)^2}{8N_a}\,\sigma^2_{\mathrm{XPM,scalar}}, \tag{29}$$

where $\sigma^2_{\mathrm{XPM,scalar}}$ denotes the XPM variance obtained from a scalar perturbative treatment of the NLSE (see for example Ref. [15]). We recall that each mode group, labeled by a, contains N_a spatial modes.

Under the approximation that spatial modes within the same mode group experience negligible modal dispersion, the scalar XPM variance can be recast into the form of Eq. (19) of Ref. [18]. In that formulation, the second- and fourth-order noise coefficients appearing in Eqs. (46) and (48) encode the details of pulse collisions through integrals of the type shown in Eq. (2). Depending on the choice of indices, these integrals describe two-, three-, or four-pulse collision processes. As discussed in Ref. [16], the two-pulse collision term is the dominant contribution. Therefore, by neglecting higher-order collision terms, i.e., approximating

$$X_{lkm} = X_{0mm}\,\delta_{l0}\delta_{km}, \tag{30}$$

Eq. (19) of Ref. [18] can be rewritten as

$$\sigma^2_{\mathrm{XPM}} = 4P^3\kappa_{aa}^2\gamma^2 \sum_m X^2_{0mm} \times \left[\frac{2N_a+1}{2} + \frac{2N_a+3}{4}\left(\frac{\langle|b_0|^4\rangle}{\langle|b_0|^2\rangle^2} - 2\right)\right], \tag{31}$$

where the dependence on the factor $(\langle|b_0|^4\rangle/\langle|b_0|^2\rangle^2 - 2)$ is retained in the same form as in the original derivation. We note that, for $l = 0$ and $k = m$, the coefficients X_{lkm} are real. Upon simplification and regrouping of terms, this expression directly leads to Eq. (6).

The derivation of the XGXPM expression follows an analogous procedure, starting from Eq. (28) of Ref. [18]. In the notation of that work,

$$\sigma^2_{\mathrm{XGXPM}} = 2N_b\,\zeta^2_{\mathrm{XPM,scalar}}, \tag{32}$$

where $\zeta^2_{\mathrm{XPM,scalar}}$ differs from $\sigma^2_{\mathrm{XPM,scalar}}$ by accounting for modal dispersion between different mode groups (see Ref. [6, Eq. (49)]). This effect is incorporated through generalized pulse-collision coefficients Y_{hkp}, which reduce to X_{hkp} in the limit of XPM-only interactions, namely neglecting the three and four waves FWM contribution. Applying the same two-pulse-collision approximation and performing the corresponding simplifications yields Eq. (7).

B. Analytic expressions for the pulse models

Analytical results are available for Gaussian and Nyquist pulses. In the Gaussian case, with time normalized in units of the symbol time T, and space normalized in units of the dispersion length $L_D = T^2/|\beta_2|$, the pulse propagated linearly to the position z is given by:

$$g^{(0)}(z,t) = \frac{1}{T^{1/2}(\pi(1+(z/L_D)^2))^{1/4}} \times \exp\left[\frac{i}{2}\arctan\left(\frac{z}{L_D}\right)\right]\exp\left[-\frac{t^2}{2T^2}\frac{1+i(z/L_D)}{1+(z/L_D)^2}\right]. \tag{33}$$

The input ($z = 0$) Nyquist pulse with zero roll-off is defined in the time and frequency domains by

$$g^{(0)}(t) = \frac{1}{\sqrt{T}}\frac{\sin \pi t/T}{\pi t/T}, \quad G(\omega) = \begin{cases}\sqrt{T} & |\omega| \le \frac{\pi}{T} \\ 0 & |\omega| > \frac{\pi}{T}\end{cases} \tag{34}$$

and the propagated pulse is [31, Eq. (7)],

$$g^{(0)}(z,t) = \frac{1}{4}\sqrt{\frac{2}{i\pi T(z/L_D)}} \times \left[F\left(t,\frac{\pi}{T}\right) - F\left(t,-\frac{\pi}{T}\right)\right], \tag{35}$$

where

$$F(t,\beta) = \exp\left(-i\frac{t^2}{2\beta_2 z}\right) \times \mathrm{erfi}\left(\sqrt{\frac{i\,\beta_2 z}{2}}\left(-\frac{t}{\beta_2 z}+\beta\right)\right), \quad (36)$$

with $\mathrm{erfi}(z) = \mathrm{erf}(iz)/i$, $\mathrm{erf}(\cdot)$ being the error function [32].

1) Derivation of the general expression for $C(z,\omega)$: The time integral (9) can be evaluated in the frequency domain using the Parseval identity. By noting that $c(z,t)$ is the product of two functions we can derive its Fourier transform as a convolution integral of the Fourier transforms of $g_{ap}^{(0)}(z,t)$ and of $g_{bq}^{(0)}(z, t-mT-\Delta\beta_1 z)$, denoted respectively by $G_{ap}(z,\omega)$ and $G_{bq}(z,\omega)$. In the latter expression, $\Delta\beta_1 = \Delta\beta_{1,apbq}$ is the DGD, and m is the collision order (see also Sect. II for the definition of the subindices). The relation is

$$\int_{-\infty}^{+\infty} dt|c(z,t)|^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega|C(z,\omega)|^2, \quad (37)$$

where $C(z,\omega)$ is the Fourier transform of $c(z,t)$.

$$C(z,\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\; G_{ap}(z,\Omega)G_{bq}(z,\omega-\Omega)\,. \quad (38)$$

The Fourier transforms of the propagated pulses can be readily written as:

$$G_{ap}(z,\omega) = G_{ap}(0,\omega)\exp\left(i\frac{\beta_{2,ap}z\omega^2}{2}\right), \quad (39)$$

$$G_{bq}(z,\omega) = G_{bq}(0,\omega)\exp\left(i\frac{\beta_{2,bq}z\omega^2}{2}\right) \times \exp\left[-i\omega(-mT+\Delta\beta_1 z)\right] \quad (40)$$

where $\beta_{2,ap}, \beta_{2,pq}$ are the GVD coefficients respectively for channels ap and bq, and m the collision order. The transform (38) becomes:

$$\begin{aligned} C(z,\omega) = & \\ & \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\, G_{ap}(0,\Omega)\exp\left(i\frac{\beta_{2,ap}z\Omega^2}{2}\right) G_{bq}(0,\omega-\Omega) \\ & \times \exp\left[i\frac{\beta_{2,bq}z(\omega-\Omega)^2}{2}\right]\exp\left[-i(\omega-\Omega)(-mT+\Delta\beta_1 z)\right] = \\ & = \frac{1}{2\pi}\exp\left(i\frac{\beta_{2,bq}z\omega^2}{2}\right)\exp\left[-i\omega(-mT+\Delta\beta_1 z)\right] \\ & \times \int_{-\infty}^{+\infty} d\Omega\, G_{ap}(0,\Omega)\, G_{bq}(0,\omega-\Omega) \\ & \times \exp\left(i\beta_{2m}z\Omega^2\right)\exp(-i\beta_{2,bq}z\omega\Omega)\exp\left[i\Omega(-mT+\Delta\beta_1 z)\right] \end{aligned} \quad (41)$$

where $\beta_{2m} = (\beta_{2,ap}+\beta_{2,bq})/2$. By introducing dispersion length scales for a pulse of duration T, $L_{Dap} = T^2/|\beta_{2ap}|, L_{Dbq} = T^2/|\beta_{2bq}|, L_{Dm} = T^2/|\beta_{2m}|$ and the walk-off length for the pulses propagating in the two channels, $L_W = T/|\Delta\beta_1|$, the previous integral can be cast in a dimensionless form as shown in Eq. (12).

2) Nyquist pulse: For a Nyquist pulse with zero roll-off, Eq. (41) is non null only for values of Ω between $\Omega^- = \max\{-\pi/T, \omega-\pi/T\}$ and $\Omega^+ = \min\{\pi/T, \omega+\pi/T\}$. Therefore, noticing the quadratic dependence of the exponent in Ω, the integral, which is of the form $\int_{\Omega^-}^{\Omega^+} d\Omega\; e^{iA_1\Omega^2+iA_2\Omega}$ has an explicit expression in terms of error functions with a complex argument [32] as follows

$$|C(z,\omega')|^2 = \frac{\Theta(2\pi-|\omega'|)}{(2\pi)^2}\frac{\pi L_{Dm}}{4z} \times \left|\mathrm{erf}\left[\mathscr{C}_{ap}^{+}(z,\omega')\right] - \mathrm{erf}\left[\mathscr{C}_{bq}^{-}(z,\omega')\right]\right|^2, \quad (42)$$

where $\omega' = \omega T$ is the normalized frequency, Θ is the Heaviside step function denoting the band limit, and

$$\mathscr{C}_X^{\pm}(z,\omega') = i^{3/2}\frac{\sqrt{L_{Dm}z}}{2}\left(\frac{1}{L_W}-\frac{m}{z}\pm\frac{\omega'}{L_{DX}}\mp\mathrm{sgn}(\omega')\frac{2\pi}{L_{Dm}}\right). \quad (43)$$

for $X = ap, bq$.

C. Asymptotic expressions

1) High DGD asymptotics: For the Gaussian pulse, by considering the small parameter $\varepsilon = L_W/L_D$, where the rms superscript in L_D^{rms} is dropped for brevity, the noise coefficient (4) can be written as

$$X_{0mm} = \frac{L_D}{T\sqrt{2\pi}}\int_0^{L/(L_D\varepsilon)} d\zeta\; F(\zeta)\,, \quad (44)$$

where

$$F(\zeta) = \frac{\varepsilon}{\sqrt{1+(\varepsilon\zeta)^2}}\exp\left[-\frac{(\zeta-m)^2}{2(1+(\varepsilon\zeta)^2)}\right]. \quad (45)$$

By expanding the function $F(\zeta)$ in powers of ε, $F(\zeta) \sim \sum_{j=0}\varepsilon^j\; F_j(\zeta)$, the coefficients are obtained as

$$F_1(\zeta) = \exp\left[-\frac{(m+\zeta)^2}{2}\right], \quad (46)$$

$$F_3(\zeta) = \exp\left[-\frac{(m+\zeta)^2}{2}\right]3\zeta^2\left((m+\zeta)^2-1\right), \quad (47)$$

and $F_0(\zeta) = F_2(\zeta) = F_4(\zeta) = 0$, so the integral can be estimated by a truncation of the power series. The first two nonzero terms can be expressed analytically by taking the integration limit to infinity

$$\int_0^{\infty} d\zeta\; F_1(\zeta) = \sqrt{\frac{\pi}{2}}\;\mathrm{erfc}\left(-\frac{m}{\sqrt{2}}\right), \quad (48)$$

$$\int_0^{\infty} d\zeta\; F_3(\zeta) = 3\sqrt{2\pi}\;\mathrm{erfc}\left(-\frac{m}{\sqrt{2}}\right). \quad (49)$$

Recalling that $\mathrm{erfc}(x) \to 2$ when $x \to -\infty$, choosing a number $m \gg 1$ (i.e., a collision far from $\zeta = 0$), the relation (14) is retrieved. We remark that Eq. (14) is considered nonzero, and speficied by the above expression, only for the values of m for which $z_m < L$.

In the Nyquist case for $z_m < L \ll L_D^{\mathrm{rms}}$, the first-order expansion $\mathrm{erf}(\mathscr{C}_X^{\pm}) \approx 2\mathscr{C}_X^{\pm}/\sqrt{\pi}$ can be used in Eqs. (42,43). In particular, the sum of the contributions from the expanded

Eq. (43), which are linear in the argument of the erf function, Eq. (42) simplifies to

$$|C(z_m,\omega')|^2 \approx \frac{\Theta(2\pi-|\omega'|)}{(2\pi)^2}\frac{L_{Dm}^2}{4}\left|\frac{4\pi}{L_{Dm}}-|\omega'|\left(\frac{1}{L_{Dap}}+\frac{1}{L_{Dbq}}\right)\right|^2 = \frac{\Theta(2\pi-|\omega'|)}{(2\pi)^2}\left(2\pi-|\omega'|\right)^2 \quad (50)$$

where the relation $L_{Dm}^{-1}=2(L_{Dap}^{-1}+L_{Dbq}^{-1})$ has been used. The integration of this quantity in (9) is then performed using the Parseval relation Eq. (11). This relation becomes an elementary integral after the truncation of the integration region due to the band-limitedness. Within this approximation, Eq. (17) is obtained.

2) Low DGD asymptotics: In the Gaussian case for low DGD, we can take the opposite limit for the parameter, i.e. $\varepsilon\to\infty$. Therefore, it is sufficient to evaluate the integral (4) as

$$X_{0mm}\approx L_D\int_0^{L/L_D}dz'\,\frac{1}{T\sqrt{2\pi}\sqrt{1+z'^2}}\times\exp\left[-\frac{m^2}{2(1+z'^2)}\right] \quad (51)$$

with $z'=z/L_D$. The latter expression contains an exponential dependence on the index m, but setting $m=0$, the integral can be computed, resulting in

$$X_{000}=\frac{L_D}{\sqrt{2\pi}T}\,\mathrm{asinh}\left(\frac{L}{L_D}\right). \quad (52)$$

This expression can be used to get a lower bound to the channel pair noise generated by extremely low-DGD channel pairs. The lower bound is $\mathcal{N}^{(0)}:=X_{000}^2\le\sum_m X_{0mm}^2$. By using $z\le L$ in Eq. (51), one gets the bound on the noise coefficients

$$X_{0mm}<X_{000}\exp\left[-\frac{m^2}{2(1+(L/L_D)^2)}\right]. \quad (53)$$

After performing the (exponentially decaying) summation over all the $m\in\mathbb{Z}$, to the NLIN channel-pair upper bound

$$\mathcal{N}:=X_{000}^2\sum_m\exp\left[-\frac{m^2}{1+(L/L_D)^2}\right] = X_{000}^2\,\vartheta_3\left(0,\exp\left[-\frac{1}{1+(L/L_D)^2}\right]\right). \quad (54)$$

where the function ϑ_3 is the elliptic theta function 3, evaluated in this case with *nome* $\exp\left[-\left(1+(L/L_D)^2\right)^{-1}\right]$ [33, Eq. (8.180.4)]. We remark that the sum over all the m can be approximated using a Gaussian integral instead of a discrete sum. For negligible dispersion, $L/L_D\sim 10^{-1}$, this expression for $\mathcal{N}^{(0)}$ can be approximated by expanding $\mathrm{asinh}(L/L_D)\sim L/L_D$, so to have

$$\mathcal{N}^{(0)}\approx\frac{1}{2\pi}\left(\frac{L}{T}\right)^2. \quad (55)$$

Accordingly, the theta function becomes $\vartheta_3(0,e^{-1})$, and it can be approximated with the value of the standard Gaussian integral. Indeed, the error is small: $\vartheta_3(0,e^{-1})-\sqrt{\pi}\approx 0.0001833$. By substituting this approximation along with Eq. (55) in Eq. (54), we obtain Eq. (18).

3) Hölder inequality: As a consequence of the Hölder inequality in $L_1([0,L])$, if f is an integrable function between $z=0$ and $z=L$, it holds

$$\left(\int_0^L dz\; f_{bq}(z)\right)^2\le L\int_0^L dz\;(f_{bq}(z))^2. \quad (56)$$

D. Definition of the overlap integral coefficients

The overlap integrals used in this article are the same defined in Ref. [17, Eq. 4], and [34]. In the notation of Ref. [17], they depend on the coefficients C_{jhkp}, i.e., overlaps between vector mode functions. They are connected to the expressions originally derived in the context of the multimode generalized Schrödinger equation in Ref. [34], and are widely used in the case of the analysis of the mixing properties of modal groups in subsequent works (see Refs. [6, 17]). Linking the notation of Ref. [17], defining the tensor C, and of Ref. [34], defining the tensors $Q^{(1)}$ and $Q^{(2)}$, one gets

$$C_{jhkp}=A_{\mathrm{eff}}(2Q^{(1)}_{jphk}+Q^{(2)}_{jhkp}). \quad (57)$$

The resulting κ_{ab} matrix for the fiber under analysis are reported in Tab. IV.

	LP_{01}	LP_{11}	LP_{02}	LP_{21}
LP_{01}	0.8908	0.8610	0.5522	0.8169
LP_{11}	0.8610	0.7169	0.7364	0.4089
LP_{02}	0.5522	0.7364	0.5758	0.3679
LP_{21}	0.8169	0.4089	0.3679	0.6199

TABLE IV: Values of κ_{ab} for all the mode groups.

REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave technology*, vol. 28, no. 4, pp. 662–701, 2010. DOI: 10.1109/JLT.2009.2039464.

[2] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 157–179, 2015. DOI: 10.1109/JLT.2015.2463719.

[3] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, 2012. DOI: 10.1109/JLT.2012.2190582.

[4] M. Secondini and E. Forestieri, "Scope and limitations of the nonlinear Shannon limit," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 893–902, 2016. DOI: 10.1109/JLT.2016.2620721.

[5] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008. DOI: 10.1109/JLT.2008.927791.

[6] C. Antonelli, M. Shtaif, and A. Mecozzi, "Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 36–54, 2015. DOI: 10.1109/JLT.2015.2510511.

[7] G. Bosco *et al.*, "Experimental Investigation of Nonlinear Interference Accumulation in Uncompensated Links," *IEEE Photonics Technology Letters*, vol. 24, no. 14, pp. 1230–1232, 2012. DOI: 10.1109/LPT.2012.2200672.

[8] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical Modeling of Nonlinear Propagation in Uncompensated Optical Transmission Links," *IEEE Photonics Technology Letters*, vol. 23, no. 11, pp. 742–744, 2011. DOI: 10.1109/LPT.2011.2131125.

[9] P. Poggiolini, "The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857–3879, 2012. DOI: 10.1109/JLT.2012.2217729.

[10] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, *A Detailed Analytical Derivation of the GN Model of Non-Linear Interference in Coherent Optical Transmission Systems*, Issue: arXiv:1209.0394 arXiv:1209.0394 [physics], 2014. DOI: 10.48550/arXiv.1209.0394.

[11] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Optics Express*, vol. 22, no. 13, pp. 16 335–16 362, 2014. DOI: 10.1364/OE.22.016335.

[12] G. Rademacher *et al.*, "Investigation of intermodal nonlinear signal distortions in few-mode fiber transmission," *Journal of Lightwave Technology*, vol. 37, no. 4, pp. 1273–1279, 2019. DOI: 10.1109/JLT.2019.2892052.

[13] G. Rademacher *et al.*, "Intermodal nonlinear signal distortions in multi-span transmission with few-mode fibers," *IEEE Photonics Technology Letters*, vol. 32, no. 18, pp. 1175–1178, 2020. DOI: 10.1109/LPT.2020.3016348.

[14] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in fiber-optic systems," *Optics Express*, vol. 22, no. 12, pp. 14 199–14 211, 2014. DOI: 10.1364/OE.22.014199.

[15] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Optics Express*, vol. 21, no. 22, pp. 25 685–25 699, 2013.

[16] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 593–607, 2016. DOI: 10.1109/JLT.2015.2428283.

[17] A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," *Optics Express*, vol. 20, no. 21, pp. 23 436–23 441, 2012. DOI: 10.1364/OE.20.023436.

[18] C. Antonelli, O. Golani, M. Shtaif, and A. Mecozzi, "Nonlinear interference noise in space-division multiplexed transmission through optical fibers," *Optics Express*, vol. 25, no. 12, pp. 13 055–13 078, 2017. DOI: 10.1364/OE.25.013055.

[19] J. Bromage, "Raman amplification for fiber communications systems," *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 79–93, 2004. DOI: 10.1109/JLT.2003.822828.

[20] D. Semrau, G. Saavedra, D. Lavery, R. I. Killey, and P. Bayvel, "A Closed-Form Expression to Evaluate Nonlinear Interference in Raman-Amplified Links," *Journal of Lightwave Technology*, vol. 35, no. 19, pp. 4316–4328, 2017. DOI: 10.1109/JLT.2017.2741439.

[21] H. Buglia, M. Jarmolovičius, L. Galdino, R. I. Killey, and P. Bayvel, "A closed-form expression for the Gaussian noise model in the presence of Raman amplification," *Journal of Lightwave Technology*, vol. 42, no. 2, pp. 636–648, 2023. DOI: 10.1109/JLT.2023.3315127.

[22] J. Yang *et al.*, "Experimental validation of the closed-form GN model accounting for distributed Raman amplification in an S+ C+ L-band hybrid amplified long-haul transmission system," in *ECOC 2024; 50th European Conference on Optical Communication*, VDE, 2024, pp. 67–70.

[23] F. Lorenzi *et al.*, "Model for Nonlinear Interference Noise in Raman-amplified WDM Systems," in *2022 European Conference on Optical Communications (ECOC)*, 2022.

[24] F. Lorenzi *et al.*, "Nonlinear interference noise in Raman-Amplified WDM systems," *Journal of Lightwave Technology*, vol. 41, no. 20, pp. 6465–6473, 2023. DOI: 10.1109/JLT.2023.3287650.

[25] G. P. Agrawal, *Fiber-Optic Communication Systems*. John Wiley & Sons, 2012.

[26] P. Sillard, M. Bigot-Astruc, D. Boivin, H. Maerten, and L. Provost, "Few-mode fiber for uncoupled mode-division multiplexing transmissions," in *37th European Conference and Exposition on Optical Communications*, ser. ECOC, OSA, 2011, Tu.5.LeCervin.7. DOI: 10.1364/ecoc.2011.tu.5.lecervin.7.

[27] R. Gabet *et al.*, "Complete dispersion characterization of few mode fibers by OLCI technique," *Journal of Lightwave Technology*, vol. 33, no. 6, pp. 1155–1160, Mar. 2015. DOI: 10.1109/jlt.2014.2376702.

[28] G. Marcon, A. Galtarossa, L. Palmieri, and M. Santagiustina, "Model-Aware Deep Learning Method for Raman Amplification in Few-Mode Fibers," *Journal of Lightwave Technology*, vol. 39, no. 5, pp. 1371–1380, Mar. 2021. DOI: 10.1109/jlt.2020.3034692.

[29] H. Buglia, E. Sillekens, L. Galdino, R. Killey, and P. Bayvel, "Throughput maximisation in ultra-wideband hybrid-amplified links," in *Optical Fiber Communication Conference*, Optica Publishing Group, 2024, Tu3H–5.

[30] M. T. Askari and L. Lampe, "Probabilistic Shaping for Nonlinearity Tolerance," *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1565–1580, 2025. DOI: 10.1109/JLT.2024.3521642.

[31] T. Hirooka and M. Nakazawa, "Linear and nonlinear propagation of optical Nyquist pulses in fibers," *Optics Express*, vol. 20, no. 18, pp. 19 836–19 849, 2012. DOI: 10.1364/OE.20.019836.

[32] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Courier Corporation, Jan. 1965.

[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2014.

[34] F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," *Journal of the Optical Society of America B*, vol. 25, no. 10, pp. 1645–1654, 2008. DOI: 10.1364/JOSAB.25.001645.