A Closed-Form Expression to Evaluate Nonlinear Interference in Raman-Amplified Links

Daniel Semrau, *Student Member, IEEE*, Gabriel Saavedra, *Student Member, IEEE*, Domaniç Lavery, *Member, IEEE*, Robert I. Killey, *Senior Member, IEEE*, and Polina Bayvel, *Fellow, IEEE, Fellow, OSA*

Abstract—An accurate,closed-form expression to evaluate the nonlinear interference (NLI) noise power in Nyquist-spaced, coherent optical communication systems using backward-pumped Raman amplification is presented. This enables rapid estimation of the signal-to-noise ratio (SNR) and avoids the need of integral evaluations and split-step simulations.The accuracy of the proposed formula is compared to numerical integration of the Gaussian noise (GN) model and split-step simulations over a wide range of parameters, including three different fiber types. Additionally, the impact of pump depletion on the NLI noise power is studied and the formula is applied to a second-order Raman-amplified system. In the case of first-order amplification and negligible pump depletion, a maximum deviation of 0.34 dB in NLI coefficient between the GN model and the closed-form formula is found, which corresponds to a maximum deviation of ∼0.1 dB in optimal SNR or similar figures of merit (e.g., maximum reach). When pump depletion is considered,it is shown that the NLI coefficient becomes a function of launch power and as a result the cubic power dependence of the NLI noise power is no longer valid in such regimes. Finally, for the second-order Raman-amplified system, a maximum deviation of 0.39 dB in NLI coefficient is reported.

Index Terms—Channel model, Gaussian noise model, nonlinear distortion, nonlinear interference, optical fiber communications, Raman amplification.

I. Introduction

RAMAN amplification has become a key technology to continue the development of long-haul, high-capacity systems due to the possibility of expanding the amplifier bandwidth beyond the conventional erbium-doped fiber amplifier (EDFA) gain window [1], and the potential to reduce amplified spontaneous emission (ASE) noise accumulation by using the distributed amplification configuration [2]. However, nonlinear interference (NLI), induced by the Kerr effect, remains a major limitation of system performance.

Analytical models to estimate this nonlinear interference are key for rapid and efficient system design [3], achievable rate

estimations of point-to-point links [4]–[6] and physical layer aware network optimization. The latter is essential for optical network abstraction and virtualization leading to optimal and intelligent techniques to maximize optical network capacity [7]. In the literature nonlinear interference is also referred to as nonlinear distortions [8] or nonlinear interference noise [9]. A variety of models have been proposed in the past with a trade-off between complexity and accuracy [8]–[12]. Such analytical models enable a significant reduction in computational complexity with relatively small inaccuracies compared to split-step simulations and experiments [13]–[17], and enable actual-time network capacity optimization. The Gaussian noise (GN) model has been used as a basis for this approach as it provides low computational complexity as well as good accuracy [18], [19]. However, other than for lumped amplification schemes, the GN model requires numerical integration which is increasingly time-consuming with increasing bandwidth.

To date, approximate closed-form solutions only exist for lumped amplification or ideal distributed Raman amplification [17], [18]. These closed-form formulas are the basis of a variety of studies and are extensively used in the literature [3], [6], [8], [20]–[26]. In the case of (non-ideal) distributed Raman amplification, closed-form solutions are difficult to obtain. This is because the power profile is described by a system of coupled differential equations which generally does not admit simple, closed-form solutions.

In this paper, for the first time, a closed-form expression is presented that predicts the NLI for distributed, backward-pumped Raman-amplified optical fiber communication systems. A formula for the NLI coefficient and for the coherence factor are derived for an approximated signal power profile, obtained by a standard least-mean-squares method from the actual profile, following the approach in [27], [28]. This enables the solution to be applied in regimes where no analytical description of the power profile is known, e.g., cases that include pump depletion and second-order Raman amplification. The result was compared to numerical integration and split-step simulations over a wide range of parameters and excellent agreement is found for standard single mode fiber (SSMF), non-zero dispersion-shifted fiber (NZDSF) and ultra low-loss (ULL) fiber based spans. Additionally, the effect of pump depletion on the NLI was studied. It was found that the NLI coefficient becomes a function of launch power and the cubic power dependence of the NLI noise power is no longer valid. It is further shown that the closed-form expression is able to account for pump depletion. The formula

is applied to systems using second-order Raman amplification and good agreement is reported. Finally, the closed-form predictions are compared to quadrature amplitude modulated (QAM) signals.

II. APPROXIMATION OF THE POWER PROFILE

The aim is to obtain a closed-form formula for the NLI coefficient, which requires an analytical description of the power profile along a span in backward-pumped Raman-amplified links. The nonlinear interference induced by the Kerr effect depends on the instantaneous power along the fiber span. This nonlinear interference imposes a limitation on the maximum achievable signal-to-noise ratio (SNR). The SNR after electronic dispersion compensation is given by

$$\text{SNR} \approx \frac{P}{nP_{\text{ASE}} + \eta_n P^3}, \tag{1}$$

where P is the channel launch power, P_{ASE} is the ASE noise within the channel bandwidth per span and η_n is the nonlinear interference coefficient at span n. It is assumed that the signal is Nyquist-spaced and described by a Gaussian process. The assumption of signal Gaussianity, being the main assumption of the GN model [18], and its connection to quadrature amplitude modulation (QAM) formats is well studied in the literature [9], [12], [16], [17], [29]. For these assumptions, the NLI coefficient for the central channel can be obtained by [19, Eq. (6)]

$$\eta_n = \frac{256}{27} \frac{\gamma^2}{R_{\text{b}}^2} \int_0^{\frac{B}{2}} \rho\left(f\right) \frac{\sin^2\left(2n\pi^2 f^2 \beta_2 L\right)}{\sin^2\left(2\pi^2 f^2 \beta_2 L\right)} f \ln\left(\frac{B}{2f}\right) df, \tag{2}$$

where B is the total optical bandwidth, R_{b} is the symbol rate, L is the span length, γ is the nonlinear coefficient and β_2 is the group velocity dispersion (GVD) parameter. $\rho\left(f\right)$ is the four-wave mixing (FWM) efficiency factor which is given by

$$\rho\left(f\right) = \left| \int_0^L P(z) e^{j4\pi^2 \beta_2 f^2 z} dz \right|^2. \tag{3}$$

However, closed-form solutions are not reported in the literature for the case of (non-ideal) distributed Raman amplification. In general, a system of coupled differential equations must be solved in order to obtain the signal power profile $P(z)$ for a generic Raman pumping scheme which accounts for pump depletion and higher-order Raman pumping [30, Eqs. (1) and (2)]. Only in the simplest case of first-order Raman amplification and negligible pump depletion, these coupled differential equations can be solved analytically with the result [19, Eq. (11)]

$$P(z) = e^{-az} \exp\left[\frac{C_{\text{R}} P_{\text{p}}\left(e^{a_{\text{p}} z} - 1\right)}{a_{\text{p}}}\right], \tag{4}$$

where P_{p} is the pump power, a_{p} is the attenuation coefficient at the pump wavelength λ_{p}, a is the fiber attenuation at the signal wavelength and C_{R} is the Raman gain coefficient. In this paper all attenuation coefficients describe the power attenuation and are considered in SI units (i.e. $\alpha \left[\text{Np/m}\right]$). Although, closed-form and approximate closed-form expressions exist for the FWM efficiency $\rho(f)$ in [31, Eq. (3)] and [28, Eq. (1)], a closed-form solution for the NLI coefficient (2) is yet to be reported.
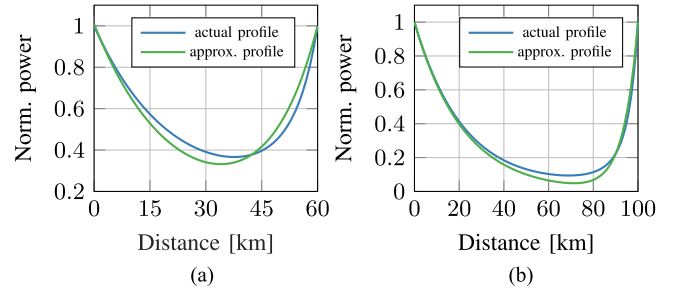

Fig. 1. Actual power profile and approximation for first-order Raman amplification and negligible pump depletion with (a) 60 km span length yielding $a_2 = 7.811 \cdot 10^{-5}$ Np/m, $b_2 = 0.937$ and (b) 100 km span length yielding $a_2 = 1.568 \cdot 10^{-4}$ Np/m and $b_2 = 0.99$.

In this work, the actual power profile $P(z)$ is approximated by two exponential functions, as originally proposed in [27, Sec. V], written as

$$P_{\text{a}}(z) = e^{-az} + b_2 e^{-a_2 L} e^{a_2 z}, \tag{5}$$

where b_2 is connected to the lumped gain at the end of the span to compensate for its loss ($b_2 = 1 - \exp\left(-aL\right)$ for transparent spans) and a parameter a_2 that is obtained via a standard least-mean-squares method. The authors in [28] use a similar approach to give a closed-form approximation of the four-wave mixing efficiency factor $\rho(f)$. In this work, we significantly extend this result by deriving a closed-form for the nonlinear interference factor η_n in (2).

As an example, the actual power profile $P(z)$ and its approximation $P_{\text{a}}(z)$ for first-order Raman amplification, negligible pump depletion and attenuation coefficients $a = 0.2$ dB/km at the signal wavelength and $a_{\text{p}} = 0.24$ dB/km at the pump wavelength are shown in Fig. 1. The attenuation coefficients can be converted to a linear scale using $\alpha \left[\text{Np/m}\right] = \alpha \left[\text{dB/km}\right] / \log_{10}\left(e\right) / 10000$. The pump powers are 27.2 dBm and 29.3 dBm for 60 km and 100 km span lengths, respectively. We define a root relative squared error (RRSE) as

$$\text{RRSE} = \sqrt{\frac{\int_0^L \left[P(z) - P_a(z)\right]^2 dz}{\int_0^L P^2(z) dz}}. \tag{6}$$

The RRSE amounts to 7.8% for the 60 km span in Fig. 1(a) and 8.2% for the 100 km span in Fig. 1(b). The deviation of the corresponding η_n between the actual profile and its approximation is typically smaller due to two reasons. First, the approximation might slightly overestimate the nonlinear process over a particular distance within the span length and underestimate the nonlinear process at a different location within the same span (cf. Fig. 1(a)), which leads to an error cancellation. Second, the deviation is often in the region of relatively low power (cf. Fig. 1(b)) where the nonlinear process is weak. Comparing the area under the curve for the actual profile and the approximation takes cancellation partly into account. For this measure, the difference is 1.8% for the 60 km span and 6.7% for the 100 km span. However, both error estimates are only rough indicators on the accuracy of the NLI coefficient, as the exact location of the power deviation and the corresponding frequency walk-off

is crucial for the nonlinear process. Overall, (5) appears as a viable candidate for approximating the signal power profile and its effectiveness with respect to the NLI coefficient will be studied in the following sections.

III. CLOSED-FORM FORMULA

In this section the closed-form expression to calculate the nonlinear interference coefficient η_n are presented. The nonlinear interference coefficient is first rewritten as [18, Sec. IX]

$$\eta_n = \eta_1 \cdot n^{1+\epsilon}, \tag{7}$$

where η_1 is the nonlinear interference factor for one span and ϵ is the so-called coherence factor. The integral expression (2) is then approximately solved for both quantities assuming a profile as in (5). For a detailed derivation, the reader is referred to appendices A and B. The result for the NLI coefficient at one span is

$$\eta_1 = \frac{8}{27} \frac{\gamma^2 \eta'}{\pi \, |\beta_2| \, R_b^2}, \tag{8}$$

with

$$\eta' = \frac{1}{a} \ln \left(\frac{\pi \phi}{a} \right) + \frac{b_2^2}{a_2} \ln \left(\frac{\pi \phi}{a_2} \right) + \frac{6}{5} \frac{b_2}{a a_2 L}$$
$$+ \frac{4 b_2 \ln (2L\phi) (e_1 - e_2)}{a_2 - a}, \tag{9}$$

and the resulting coherence factor is

$$\epsilon = \frac{1}{3} \ln \Bigg\{ 1 + \frac{26}{5} \frac{L_{\text{eff}}^2 a_2^2 + b_2^2}{a_2^2 L \eta'} + \frac{171}{40} \frac{b_2}{L a a_2 \eta'}$$
$$+ \frac{19}{5} \frac{b_2}{\eta'} \left[\frac{\ln (4L\phi) (e_1^2 - e_2^2)}{(a - a_2) \tilde{e}_1 \tilde{e}_2} + \frac{7}{5} \frac{a \tilde{e}_1 - a_2 \tilde{e}_2}{(a^2 - a_2^2) \tilde{e}_1 \tilde{e}_2} \right] \Bigg\}, \tag{10}$$

with $e_1 = e^{-aL}$, $e_2 = e^{-a_2 L}$, $\tilde{e}_1 = 1 - e^{-aL}$, $\tilde{e}_2 = 1 - e^{-a_2 L}$, $L_{\text{eff}} = \left(1 - e^{-aL} \right)/a$ and $\phi = B^2 \pi \, |\beta_2|$. All quantities must be inserted in SI units and the power attenuation coefficients are considered in Neper per meter.

As the power profile approximation is the main approximation, the proposed closed-form expression was first validated as a function of span length for three different fiber types, with parameters shown in Table I. The validation is performed for a standard single mode fiber (SSMF), as well as, a non-zero dispersion-shifted fiber (NZDSF) and an ultra low-loss fiber (ULL). The latter two were chosen to explore the impact of low dispersion and low loss on the accuracy of the closed-form expression. The dispersion parameter can be converted to the group velocity dispersion parameter using $\beta_2 = -D\lambda^2 / (2\pi c)$ with $c = 3 \cdot 10^8$.

The nonlinear interference coefficient as a function of span length at one span and at 20 spans is plotted in Fig. 2. The NLI coefficient of the actual power profile was obtained by numerically integrating (2). In order to validate the integral evaluations, the transmission systems were simulated by numerically solving the Manakov equation with the split-step algorithm using Gaussian modulated signals. The Gaussian modulation was used to

TABLE I
SIMULATION PARAMETERS

Parameters	SSMF	NZDSF	ULL
Loss @ λ (α) [dB/km]	0.2	0.2	0.165
Loss @ λ_{p} (α_p) [dB/km]	0.24	0.24	0.2
Loss @ λ_{p2} (α_{p2}) [dB/km]	0.26	0.26	0.22
Dispersion (D) [ps/nm/km]	17	3.8	16.3
GVD ($|\beta_2|$) [ps^2/km]	21.7	4.84	20.8
NL coefficient (γ) [1/W/km]	1.2	1.2	1.2
Raman gain (C_{R}) [1/W/km]		0.3	
Number of channels		31	
Optical bandwidth (B) [THz]		1.022	
Symbol rate (R_b) [GBd]		32	
Channel spacing [GHz]		33	
Roll-off factor [%]		0.1	

emulate signal Gaussianity. The Manakov equation is well established in the literature for simulating dual polarization signals. At the receiver, electronic dispersion compensation and matched filtering was performed. The SNR was ideally estimated as the ratio between the variance of the transmitted symbols $E[|X|^2]$ and the variance of the noise σ^2, where $\sigma^2 = E[|X - Y|^2]$ and Y represents the received symbols after digital signal processing as detailed in [32]. The NLI coefficient was extracted from the SNR using (1). It can be seen in Fig. 2 that the simulation results are in excellent agreement with numerical integration of the actual profile. The NLI coefficient of the approximated profile, obtained by numerically integrating (2), and the proposed closed-form expression are also plotted in Fig. 2, with an excellent agreement between the two. The NZDSF exhibits the largest nonlinearity and coherence due to its low dispersion. The ULL fiber exhibits a higher nonlinearity than the SSMF due to its higher average (normalized) power along the span. The discrepancy between numerically integrating the actual profile and the closed-form arises from the power profile approximation. The maximum discrepancy for all three fibers stays below 0.33 dB for the NLI coefficient at one span and below 0.34 dB for the NLI coefficient at 20 spans. Overall it can be seen that the proposed formula provides an excellent alternative to time consuming numerical integration and split-step simulations.

Although the NLI coefficient η_{20} already includes the coherence factor via (7), ϵ is validated separately as a function of optical bandwidth. The coherence factor can be calculated, similar to [18, Eq. (21)], as

$$\epsilon = \frac{\ln \left(\frac{\eta_{20}}{\eta_1} \right)}{\ln(20)} - 1. \tag{11}$$

Equation (2) is used in order to compute η_1 and η_{20} which are then inserted in (11) to get ϵ. The result as a function of optical bandwidth is shown in Fig. 3 for 80 km spans. The coherence factor rapidly decays with increasing bandwidth due to the faster walk-off between widely spaced channels. The closed-form formula exhibits remarkable accuracy compared to numerically integrating the actual power profile. For instance, at a bandwidth of 3 THz the deviation is 1.4% for the NZDSF,
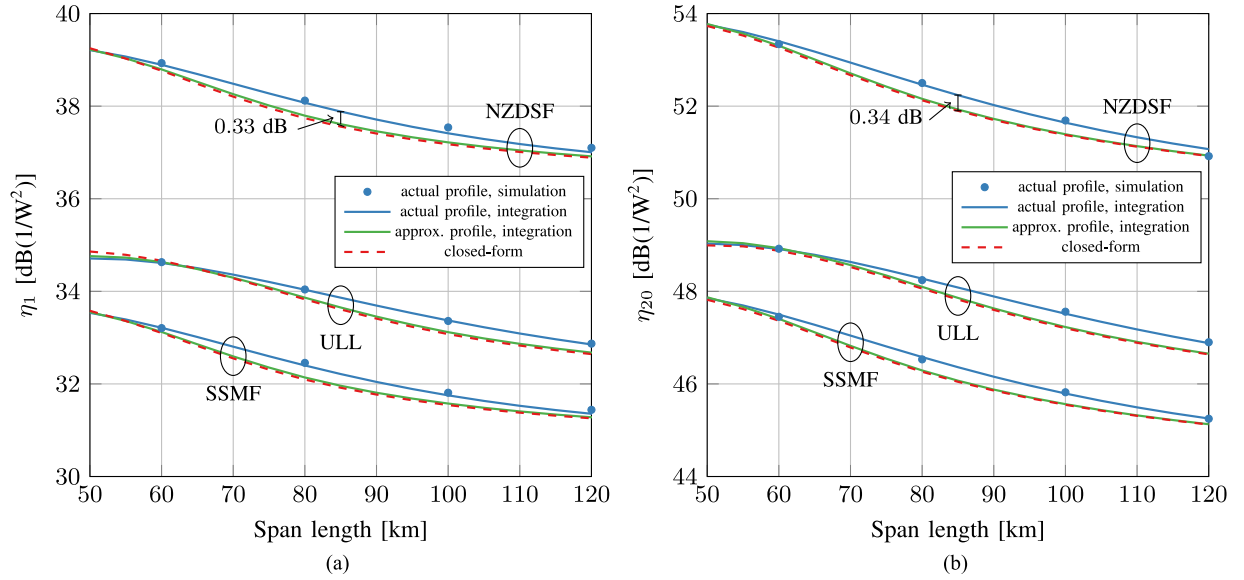
Fig. 2. The nonlinear interference coefficient as a function of span length at (a) 1 span and (b) 20 spans for an optical bandwidth of ~ 1 THz. Blue solid lines represent numerical integration of the actual profile, blue markers split step simulations, green solid lines the approximated profile with numerical integration and dashed red lines represent the closed-form solution.
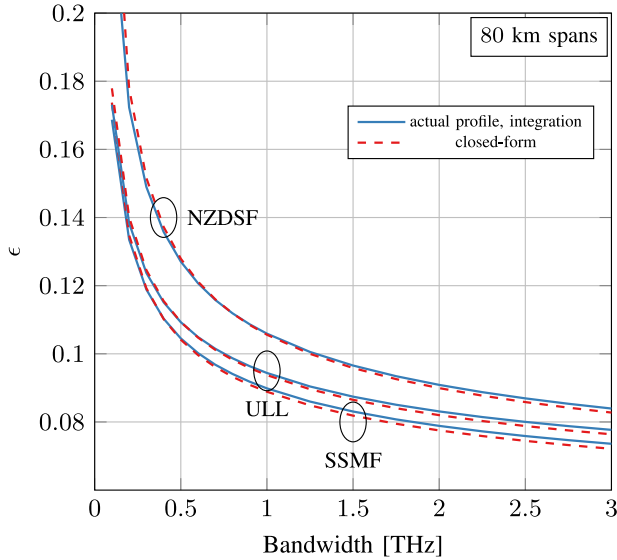

Fig. 3. The coherence factor as a function of optical bandwidth for 80 km spans. Blue solid lines represent numerical integration of the actual profile and dashed red lines represent the closed-form solution.

1.7% for the ULL fiber and 2% for the SSMF. The accumulation of nonlinear interference can be separated into an incoherent and a coherent contribution by rewriting (7) in decibel scale

$$\eta_1[\text{dB}] + n^{1+\epsilon}[\text{dB}] = \eta_1[\text{dB}] + \underbrace{n[\text{dB}]}_{\text{inc.}} + \underbrace{\epsilon \cdot n[\text{dB}]}_{\text{coh.}}. \quad (12)$$

For a bandwidth of 1 THz and $\epsilon = 0.093$ the incoherent accumulation amounts to an increase of $n[\text{dB}] = 13.01$ dB and the coherent accumulation to an increase of $\epsilon \cdot n[\text{dB}] = 1.21$ dB in NLI coefficient at $n = 20$ spans. The coherence factor for Raman amplified systems is generally higher than for their

lumped-amplified counterparts, due to the lower effective loss along the span.

The NLI coefficient as a function of the span number using 60 km spans is shown in Fig. 4(a). The proposed closed-form formula exhibits very good accuracy compared to numerically integrating the actual profile and split-step simulations. The coherence factors, obtained in closed-form, are 0.11 for the NZDSF, 0.096 for the ULL fiber and 0.093 for the SSMF. Fig. 4(a) also shows that the NLI coefficient grows according to (7) and that the coherence factor is virtually independent of the span number.

The SNR as a function of channel launch power at 20 spans, corresponding to a total distance of 1200 km, is shown in Fig. 4(b) with a P_{ASE} of -42.68 dBm. At optimum launch power the deviation between the numerical integration of the actual profile and the closed-form is 0.04 dB at -6.83 dBm/ch launch power and 0.13 dB in the highly nonlinear regime at 0 dBm/ch launch power. As expected, this shows that the difference in optimal SNR (or other figures of merit e.g., maximal reach, mutual information) is even smaller than the difference in NLI coefficient. Similar to [18, Eq. (32)], the inaccuracy in NLI coefficient $\Delta\eta$ translates to the inaccuracy in optimal SNR ΔSNR according to the rule

$$\Delta\text{SNR}[\text{dB}] \approx -\frac{1}{3}\Delta\eta[\text{dB}]. \quad (13)$$

To aid comparison with systems that use QAM modulation, a 16-QAM signal was simulated and its performance shown in Fig. 4(b). Gaussian modulation always underestimates the performance (e.g., SNR, maximal reach) compared to QAM constellations. The optimum SNR for 16-QAM is 21.84 dB at -6 dBm/ch launch power while the optimum SNR for Gaussian modulation is 21.35 dB at -6.7 dBm/ch launch power. This
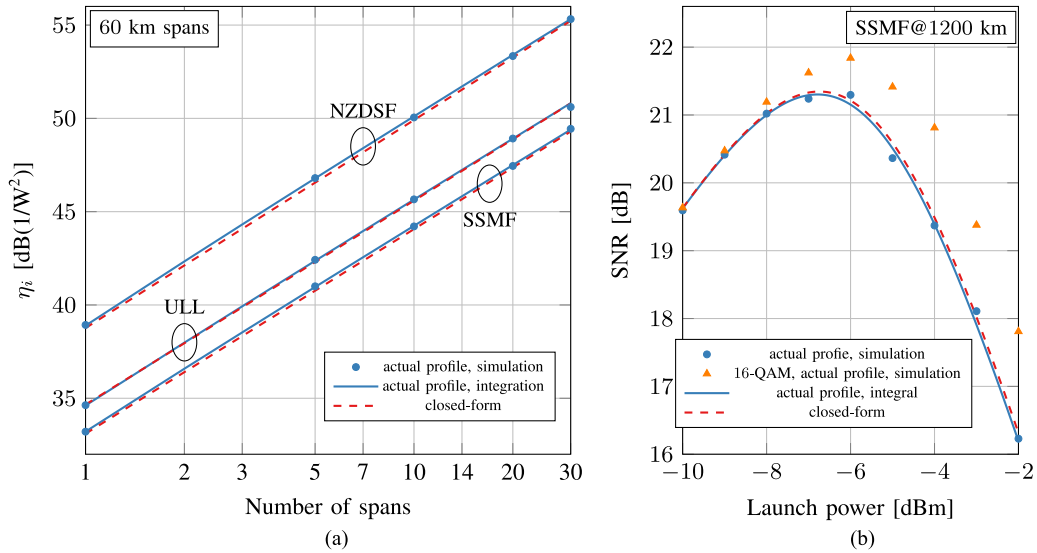
Fig. 4. (a) The NLI coefficient as a function of span number and (b) the SNR as a function of launch power. Blue solid lines represent numerical integration of the actual profile, blue markers split step simulations, yellow markers split step simulations with 16-QAM modulation format and dashed red lines represent the closed-form solution.

results in a difference in SNR of 0.51 dB between Gaussian modulation and 16-QAM modulation.

IV. PUMP DEPLETION

In the previous sections the results were obtained assuming that pump depletion is negligible. In this section, this assumption is dropped and the effect of pump depletion on the power profile and the nonlinear interference coefficient is investigated. If the total signal power is comparable to the pump power, then the effect of pump to signal power depletion must be considered. This effect becomes more significant as optical bandwidth is increased since the overall signal power, at optimum launch power, is monotonically increasing with bandwidth. Additionally, the possibility of using nonlinear compensation schemes will lead to the use of higher signal powers. When pump depletion is non-negligible, a system of nonlinear differential equations must be solved numerically, as there is no analytical solution. The system of differential equations is given by [27, Eqs. (1) and (2)]

$$\frac{\partial P}{\partial z} = aP - C_{\mathrm{R}} P P_{\mathrm{p}},$$

$$\frac{\partial P_{\mathrm{p}}}{\partial z} = -a_{\mathrm{p}} P_{\mathrm{p}} - \frac{\lambda}{\lambda_p} C_{\mathrm{R}} P P_{\mathrm{p}}, \qquad (14)$$

where λ is the signal wavelength and λ_{p} is the pump wavelength. The power profiles, obtained by numerically integrating (14) for no pump depletion, -8 dBm per channel, -2 dBm per channel and 8 dBm per channel are shown in Fig. 5(a). The pump power required to compensate for the span loss was 27.2 dBm for the case of no pump depletion and 28.96 dBm for the case of pump depletion (8 dBm/ch). No significant change in the power profile, as a result of pump depletion, occurs in the range of -8 dBm/ch to -2 dBm/ch, which shows that pump depletion is negligible for the systems studied in the previous section. However, the situation is different for higher launch powers or larger

optical bandwidths which both increase the total signal power. For instance, a launch power of 8 dBm/ch for a 1 THz (31 channels) signal has the same total signal power as a 5 THz (C-band) signal with 1 dBm per channel. At these powers substantial pump depletion is observed. Fig. 5(b) shows the NLI coefficient as a function of the channel launch power when pump depletion is included. The actual power profiles are again approximated by two exponential functions and then the closed-form formula is applied. The NLI coefficient becomes a function of launch power when pump depletion is no longer negligible; starting at approximately -4 dBm/ch. As pump depletion reduces the average power along a span the NLI coefficient reduces accordingly. The reduction in NLI coefficient at 8 dBm/ch is 0.98 dB for the NZDSF, 1.2 dB for the ULL fiber and 0.97 dB for the SSMF. The ULL shows the highest decrease in NLI due to its higher average (normalized) signal power along a span. The result of pump depletion is that the cubic law of NLI noise power in (1) is no longer valid for Raman-amplified systems where pump depletion is significant and the NLI coefficient must be evaluated at each launch power individually. Nonlinear compensation techniques (e.g., optical phase conjugation or digital back-propagation) increase the channel launch power [33]. In systems utilizing NLC, special attention should be given to the effect of pump depletion and the resulting changes in nonlinear interference. The closed-form expression correctly follows the dependency of the launch power with error of 0.21 dB for the NZDSF, 0.13 dB for the ULL fiber and 0.2 dB for the SSMF. The proposed formula eliminates the necessity to perform time consuming numerical integrations over each channel launch power.

V. SECOND-ORDER RAMAN AMPLIFICATION

Second-order amplification involves adding a second pump which alters the power profile of the first pump and hence the signal power profile. This scheme is of relevance due to the
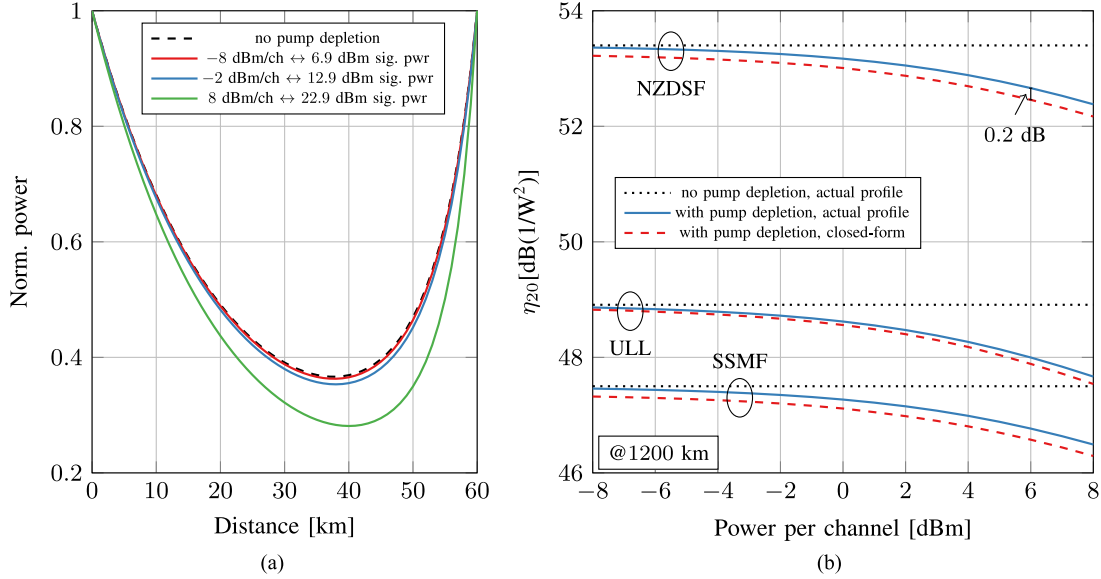
Fig. 5. (a) The power profile for different channel launch powers including pump depletion for SSMF and NZDSF. (b) The NLI coefficient as a function of launch power including pump depletion. Black dashed lines represent numerical integration of the actual profile without pump depletion, blue solid lines represent numerical integration of the actual profile including pump depletion and dashed red lines represents the closed-form solution including pump depletion.

improved ASE noise performance they offer compared to first-order Raman amplification. Recently, it has been shown that second-order Raman amplification allows for the extension of the optical system reach [34]. Additionally, it offers the possibility of a symmetrical power profile which is key to optical phase conjugation techniques [35]. In this section, it is shown that the proposed closed-form expression can be applied to second-order Raman amplification. In order to obtain the signal power profile, an additional equation must be added to the system of differential equations. The system is given by [36, Eqs. (1), (2), and (3)]

$$\frac{\partial P}{\partial z} = aP - C_R P P_p,$$

$$\frac{\partial P_p}{\partial z} = -a_p P_p - \frac{\lambda}{\lambda_p} C_R P P_p + C_R P_p P_{p2},$$

$$\frac{\partial P_{p2}}{\partial z} = -a_{p2} P_{p2} - \frac{\lambda_p}{\lambda_{p2}} C_R P_p P_{p2}, \quad (15)$$

where P_{p2} is the power profile of the second-order pump at wavelength λ_{p2}. Second-order Raman amplification generally exhibits a higher average (normalized) power along the span than first-order Raman amplification, making it more susceptible to pump depletion (at same launch power). Fig. 6(a) shows the power profile of the signal for $a = 0.165$ dB/km, $a_p = 0.2$ dB/km, $a_{p2} = 0.22$ dB/km and $P_{p2}(0) = 28.6$ dBm. The actual power profile is obtained by numerically solving (15) with a fixed first-order pump power of 20 dBm at 1455 nm and the second-order pump at 1366 nm. This power was chosen such that the fiber loss is compensated for. The second-order pump results in a peak at the end of the fiber span leading to a stronger deviation between the actual and approximated power profile.

The variation of the NLI coefficient as a function of span length is shown in Fig. 6(b). The coefficients were obtained from split-step simulations and numerical integration of (2) with the actual profile obtained from (15) shown in blue and the approximation in green. The NLI coefficients are generally larger compared to first-order Raman amplification due to the higher average (normalized) power along the span. The main difference of second-order amplification compared to first-order amplification is that at long span lengths a discrepancy remains due to the peak in the actual power profile that cannot be approximated by the two exponentials. At a 120 km span length, a maximum deviation of 0.37 dB, 0.38 dB and 0.39 dB is observed for the NZDSF, for the ULL fiber and SSMF, respectively. Overall, Fig. 6(b) shows that the proposed closed-form can predict the nonlinear interference in the case of second-order Raman amplification, when the second-order pump does not dominate the signal power profile.

VI. PERFORMANCE PREDICTION OF REALISTIC SYSTEMS

In this section the proposed closed-form expression is used to compute signal-to-noise ratios for practical systems and is compared to QAM-modulated signals obtained by split-step simulations. Two scenarios are studied. One hybrid Raman/EDFA system where the first-order Raman gain is held well under the span loss. This is typically done to ensure operation without pump-depletion and yields better noise performance than an EDFA only system. In the second case, an all-Raman configuration is analyzed where the Raman gain is several dB higher than the span loss to offset passive devices such as pump coupling devices, power monitoring equipment and gain equalizing filter. The factor b_2 can be calculated for arbitrary configurations as

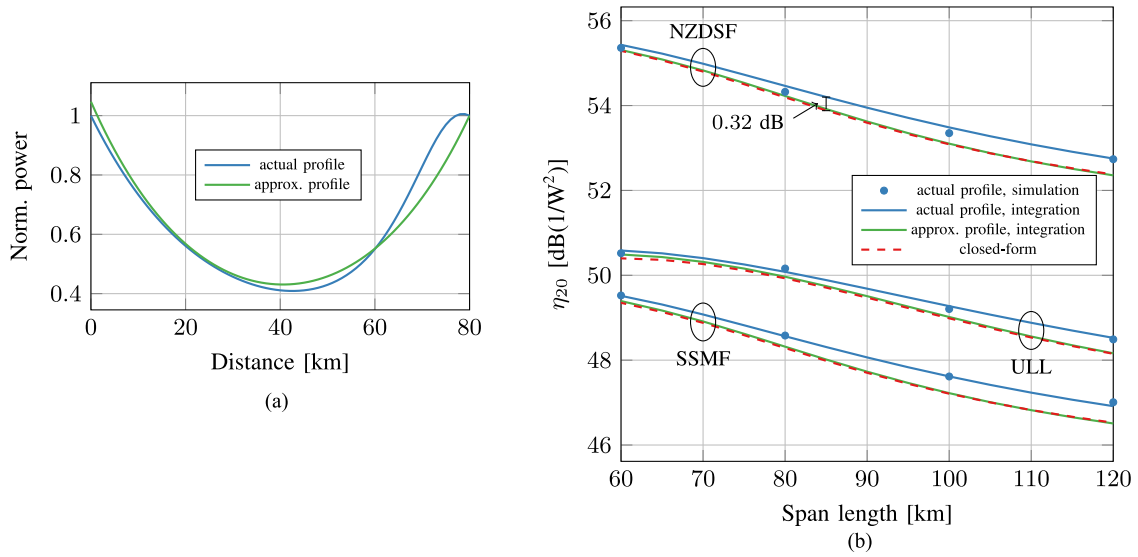$$b_2 = 10^{\frac{G}{10}} - e^{-aL} \quad (16)$$

Fig. 6.　(a) The power profile for second-order Raman amplification and approximation for ULL yielding $a_2 = 3.754 \cdot 10^{-5}$ Np/m and $b_2 = 0.952$. The blue line represent the actual profile and the green line represents the approximate profile. (b) NLI coefficient as a function of span length for second order Raman pumping. Blue solid lines represent numerical integration of the actual profile, blue markers split step simulations, green solid lines the approximated profile with numerical integration and dashed red lines represent the closed-form solution.
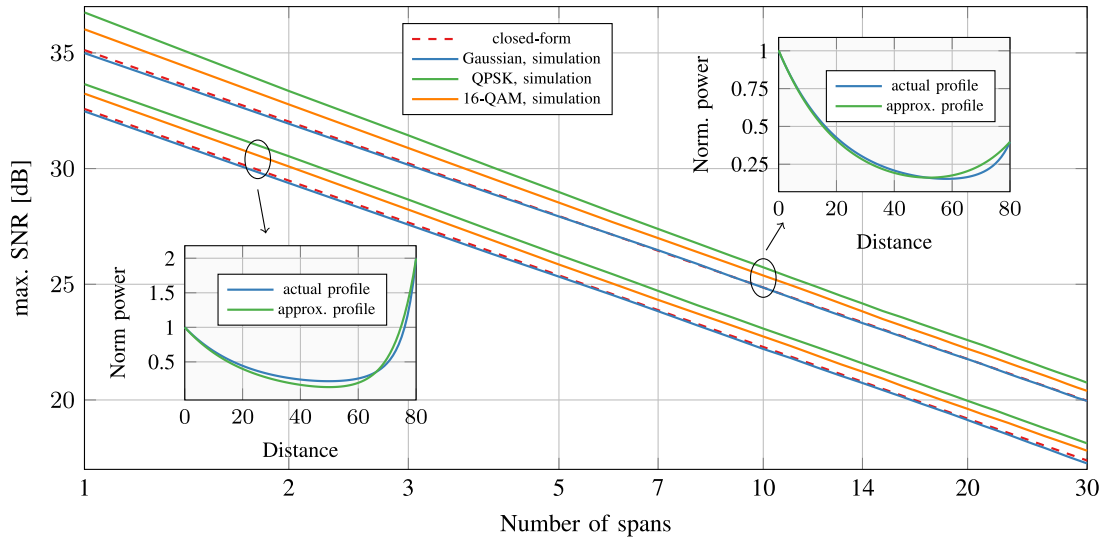

Fig. 7.　The SNR at optimum launch power as a function of span number for 2 different power profiles. The closed-form expression as well as QPSK, 16-QAM, and Gaussian modulation obtained from simulations are shown. The all-Raman system yields $a_2 = 1.346 \cdot 10^{-4}$ Np/m, $b_2 = 1.970$ and the hybrid Raman/EDFA system yields $a_2 = 5.964 \cdot 10^{-5}$ Np/m and $b_2 = 0.373$.

where G is the excess Raman gain in decibels with respect to the span loss (i.e., the additional gain which is beyond the gain required to exactly overcome span loss). 80 km spans of SSMF are chosen for both configurations yielding a span loss of 16 dB. In the first case, the Raman gain is chosen to be 12 dB ($G = -4$ dB) and in the second case 19 dB ($G = 3$ dB). The hybrid amplified transmission system exhibits a pump power of 27.47 dBm and a P_{ASE} of -42.4 dBm and the all-Raman system exhibits a pump power of 29.12 dBm and a P_{ASE} of -39.93 dBm.

The result is shown in Fig. 7, where the red dashed line represents the SNR obtained in closed-form using (1) and the blue line represents Gaussian modulation obtained from simulations. The green and yellow line represent simulations using QPSK and 16-QAM modulated signals, respectively. The closed-form predictions are in excellent agreement with the Gaussian modulated signal with a maximum deviation of 0.13 dB in SNR throughout Fig. 7. After one span, where the Gaussianity assumption is known to be less accurate, the difference between the closed-form and QPSK is 1.63 dB for the hybrid Raman/EDFA system and 1.07 dB for the all-Raman system. These differences decrease with increasing distance to 0.91 dB and 0.77 dB at 20 spans for the hybrid Raman/EDFA and the all-Raman case, respectively. With higher QAM-order the 4th and 6th

order moments of the signal constellation increase, resulting in more Gaussian-like statistics. Therefore, QAM modulated signals higher than QPSK are expected to be closer to the predictions of the presented closed-form. The difference between the closed-form and 16-QAM at 20 spans are 0.53 dB and 0.43 dB for the hybrid and the all-Raman system, respectively. The gap between QAM signals and Gaussian modulation in Raman-amplified systems is typically larger compared to its lumped counterpart [37]. Its physical sources and implications has been studied in [38]. The modulation format dependence in closed-form for Raman-amplified links is left for future work. However, (8) is a sum of three contributions, namely a "lumped" contribution, a contribution arising from the gain at the end of the span and its cross term. The lumped contribution can be corrected in closed-form with the result in [17]. As an example, correcting only the lumped contribution, this reduces the difference between QPSK and the closed-form from 0.77 dB to 0.58 dB and from 0.43 dB to 0.3 dB for 16-QAM at 20 spans for the all-Raman amplified case.

It has been shown that a fraction of the NLI, particularly in the case of Raman amplification, is phase-noise characterized by long temporal correlations [37]–[39]. It has been suggested that those correlations can be partially mitigated by adaptive equalizers or carrier phase recovery algorithms. However, the work in [40] experimentally compares a variety of NLI mitigation schemes for an Raman-amplified system and concludes that the NLI mitigation by simple receivers are rather modest for long distances, even in the context of strong temporal/spectral correlations. The closed-form expression in this paper therefore remains of practical interests for current receivers that almost always assume Gaussian channel statistics.

VII. CONCLUSION

A closed-form formula was presented to compute the nonlinear interference for multi-span, Nyquist-spaced, coherent optical communication systems that use first or second-order distributed Raman amplification including pump depletion. The presented formula significantly reduces the computational complexity required for the design and modeling of Raman-amplified WDM links. It was verified using split-step simulations and numerical integration with excellent agreement throughout. The deviation between the closed-form and simulations were within 0.34 dB in terms of NLI coefficient for first-order Raman amplification and 0.39 dB for second-order Raman amplification. This deviation was found to be linked to the deviation between the actual power profile and approximation by two exponential functions. It has been shown that pump depletion imposes a power dependence and the cubic power dependence of the NLI noise power becomes invalid in such regimes. The proposed formula predicts the nonlinear interference with minor inaccuracies, making it an excellent alternative to time-consuming numerical integrations or split-step simulations. This approach allows for rapid evaluation of performance (e.g., SNR, maximum reach) in Raman-amplified systems, an essential step towards dynamic optical network capacity optimization and intelligent information infrastructure design.

APPENDIX A
DERIVATION OF THE NONLINEAR INTERFERENCE COEFFICIENT η_1

In this section the closed-form expression for the nonlinear interference coefficient η_1 is derived. A prefactor of $\frac{16\gamma^2}{27R_b^2}$ is not shown throughout the derivation for notational convenience. With a change of variables $f^2 = \nu$, setting $n = 1$ and extending the integration limit to infinity in (2) we obtain

$$\eta_1 \propto 8 \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)$$
$$\times \left| \underbrace{\frac{e^{-aL+j\Phi\nu L}-1}{-a+j\Phi\nu}}_{X} + b_2 \underbrace{\frac{e^{j\Phi\nu L}-e^{-a_2 L}}{a_2+j\Phi\nu}}_{Y} \right|^2 d\nu, \quad (17)$$

with $\Phi = 4\pi^2 |\beta_2|$. Extending the integration limit to infinity yields a negligible error for typical bandwidths in optical communications ($B > 100$ GHz), due to the decay of the four-wave mixing efficiency. For the SSMF, the approximation error by setting B to infinity is 0.02 dB in NLI coefficient for $B = 100$ GHz and < 0.01 dB for $B = 300$ GHz. We apply the identity

$$|X + Y|^2 = |X|^2 + |Y|^2 + 2\Re(X)\Re(Y) + 2\Im(X)\Im(Y), \quad (18)$$

and solve the resulting terms individually. The integral corresponding to the first term $|X|^2$ can be solved by assuming $e^{-aL} \ll 1$ and using (42) and (43)

$$8\int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)|X|^2 d\nu$$
$$= 8\int_0^\infty \left[\ln\left(\frac{B}{2}\right) - \frac{1}{2}\ln(\nu)\right] \left|\frac{e^{-aL+j\Phi\nu L}-1}{-a+j\Phi\nu}\right|^2 d\nu$$
$$\approx 8\int_0^\infty \left[\ln\left(\frac{B}{2}\right) - \frac{1}{2}\ln(\nu)\right] \frac{1}{a^2+\Phi^2\nu^2} d\nu$$
$$= \frac{2\pi}{a\Phi} \ln\left(\frac{B^2\Phi}{4a}\right). \quad (19)$$

The integral corresponding to the second term $|Y|^2$ can be solved similarly by assuming $e^{a_2 L} \gg 1$ yielding

$$8\int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)|Y|^2 d\nu$$
$$\approx 8b_2^2 \int_0^\infty \left[\ln\left(\frac{B}{2}\right) - \frac{1}{2}\ln(\nu)\right] \frac{1}{a_2^2+\Phi^2\nu^2} d\nu$$
$$= \frac{2b_2^2\pi}{a_2\Phi} \ln\left(\frac{B^2\Phi}{4a_2}\right). \quad (20)$$
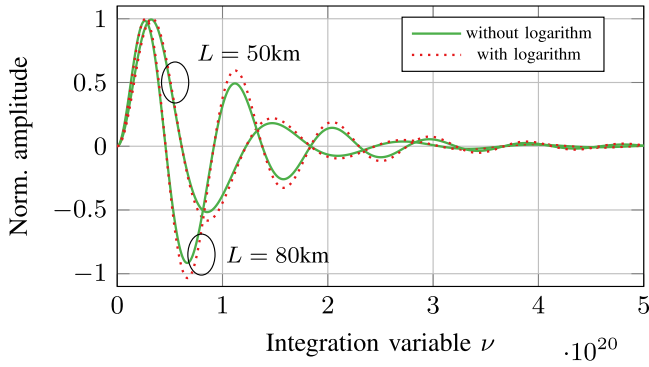
Fig. 8. Illustration of the approximation in (22) approximation using 2 different span length, SSMF and 1 THz optical bandwidth.

The integral corresponding to the cross-term can be rewritten as

$$16 \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right) \left[\Re(X)\Re(Y) + \Im(X)\Im(Y)\right] d\nu$$

$$= 16b_2 \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right) \left[\frac{(a_2 + a)\Phi\nu \sin(\Phi\nu L)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)}\right.$$

$$+ \frac{+\Phi^2\nu^2(e^{-aL} + e^{-a_2 L}) - \Phi^2\nu^2 \cos(\Phi\nu L)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)}$$

$$\left. + \frac{-aa_2(e^{-aL} + e^{-a_2 L}) + aa_2 \cos(\Phi\nu L)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)}\right] d\nu. \qquad (21)$$

In the following, we individually solve each of the five integrals resulting from (21). As each of these integrals does not seem to have an analytical solution, we propose to solve it approximately by replacing $\ln\left(\frac{B}{2\sqrt{\nu}}\right)$ by a suitable constant $\ln\left(\frac{B}{2\sqrt{\nu_0}}\right)$ with $\nu_0 = \frac{\pi}{2\Phi L}$. The constant is chosen such that the sine in (21) reaches its first maximum. The first term in (21) can then be solved using (44)

$$16b_2(a_2 + a) \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \int_0^\infty \frac{\Phi\nu \sin(\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$= \frac{8b_2\pi(a_2 + a)}{\Phi} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{e^{-aL} - e^{-a_2 L}}{a_2^2 - a^2}. \qquad (22)$$

To illustrate the approximation Fig. 8 shows the integrand of (22) with and without the logarithm for two different span lengths and the parameters of the SSMF. It can be seen that this approach is fairly accurate as the logarithm varies little over the domain where the contribution to the integral is substantial for typical parameters in optical communications.

The second term in expression (21) is solved using (45)

$$16b_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \int_0^\infty \frac{\left(e^{-aL} + e^{-a_2 L}\right) \Phi^2\nu^2}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$= \frac{8b_2\pi}{\Phi} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{e^{-aL} + e^{-a_2 L}}{a + a_2}. \qquad (23)$$

The third term in (21) is solved using (46)

$$16b_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \int_0^\infty \frac{-\Phi^2\nu^2 \cos(\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$= \frac{8b_2\pi}{\Phi} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{ae^{-aL} - a_2 e^{-a_2 L}}{a_2^2 - a^2}, \qquad (24)$$

and the fourth term in (21) is solved using (47)

$$-16b_2 aa_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \int_0^\infty \frac{e^{-aL} + e^{-a_2 L}}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$= -\frac{8b_2\pi}{\Phi} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{e^{-aL} + e^{-a_2 L}}{a + a_2}. \qquad (25)$$

In order to solve the fifth term in (21) the integration domain is split into two domains to improve the accuracy of the approximation as the cosine does not vanish for small ν, where the logarithm varies more rapidly. In the integration domain $\nu \in [0, B_0]$ the integral is approximated by

$$16b_2 aa_2 \int_0^{B_0} \ln\left(\frac{B}{2\sqrt{\nu}}\right) \frac{\cos(\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$\approx \frac{16b_2}{aa_2} \int_0^{B_0} \ln\left(\frac{B}{2\sqrt{\nu}}\right) d\nu, \qquad (26)$$

where B_0 is chosen such that $\cos(\Phi L\nu) \approx 1$, $a^2 \gg \Phi^2\nu^2$ and $a_2^2 \gg \Phi^2\nu^2$ for $\nu \in [0, B_0]$. $B_0 = \frac{\pi}{8\Phi L}$ yields $\cos(\Phi L B_0) \approx 0.92$ and $a^2 \gg \frac{\pi^2}{64 L^2}$. Even with pessimistic values the inequality gives $212 \gg 6.2$ for $a = 0.2$ dB/km and $L = 50$ km. For $\nu > B_0$ the logarithm is again replaced by a constant and the whole term is solved as

$$16b_2 aa_2 \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right) \frac{\cos(\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$\approx 16b_2 aa_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \int_0^\infty \frac{\cos(\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu$$

$$+ \frac{8b_2}{aa_2} \int_0^{B_0} \ln\left(\frac{\nu_0}{\nu}\right) d\nu. \qquad (27)$$

Finally, using (48) and (49) yields

$$8b_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{\pi}{\Phi} \frac{ae^{-La_2} - a_2 e^{-La}}{a^2 - a_2^2}$$

$$+ \frac{8b_2 B_0}{aa_2} \left[\ln\left(\frac{\nu_0}{B_0}\right) + 1\right]$$

$$= 8b_2 \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{\pi}{\Phi} \frac{ae^{-La_2} - a_2 e^{-La}}{a^2 - a_2^2} + \frac{8b_2 B_0}{aa_2} [\ln(4) + 1]$$

$$\approx \frac{8b_2\pi}{\Phi} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{a_2 e^{-La} - ae^{-La_2}}{a_2^2 - a^2} + \frac{96b_2 B_0}{5aa_2}. \qquad (28)$$

All 5 terms are summed up to obtain the cross-term as

$$
16 \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right) \left[\Re(X)\Re(Y) + \Im(X)\Im(Y)\right] d\nu
$$

$$
\approx \frac{96 b_2 B_0}{5 a a_2} + \frac{4 b_2}{\pi |\beta_2|} \ln\left(\frac{B}{2\sqrt{\nu_0}}\right) \frac{e^{-aL} - e^{-a_2 L}}{a_2 - a}
$$

$$
= \frac{3 b_2}{5 a a_2 \pi |\beta_2| L} + \frac{2 b_2 \ln\left(2 B^2 \pi |\beta_2| L\right)}{\pi |\beta_2|} \frac{e^{-aL} - e^{-a_2 L}}{a_2 - a}. \tag{29}
$$

The nonlinear interference coefficient η_1 is then obtained by summing up the solution of (19), (20) and the solution of the cross-term (29).

APPENDIX B
DERIVATION OF THE COHERENCE FACTOR ϵ

In this section the closed-form expression for the coherence factor ϵ is derived. A general formula for the coherence factor is given as [18, Eq. (21)]

$$
\epsilon = \frac{1}{\ln(20)} \ln\left[1 + \frac{1}{20} \frac{\eta_{\mathrm{coh},20}}{\eta_1}\right], \tag{30}
$$

where the enumerator η_{coh} in the logarithm stands for the coherent accumulation of nonlinear interference and the denominator $20\eta_1$ for the incoherent accumulation of the nonlinear interference coefficient at 20 spans. The coherence factor is approximately independent of the number of spans for typical values in optical communications. As in [18], coherence factor is calculated for 20 spans. The coherent part of the nonlinear interference coefficient at span n is given by [18, Appendix H]

$$
\eta_{\mathrm{coh}} \propto 16 \sum_{k=1}^{n-1} \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)(n-k)\cos\left(k\Phi L\nu\right)
$$

$$
\left|\underbrace{\frac{e^{-aL+j\Phi\nu L}-1}{-a+j\Phi\nu}}_{X} + \underbrace{b_2 \frac{e^{j\Phi\nu L}-e^{-a_2 L}}{a_2+j\Phi\nu}}_{Y}\right|^2 d\nu, \tag{31}
$$

where a prefactor of $\frac{16\gamma^2}{27 R_b^2}$ is not shown throughout the derivation for notational convenience. Similar to the derivation of η_1, (31) is rewritten using (18) and the resulting terms are solved individually. For the first two terms we make use of the result in [18, Appendix H] and obtain

$$
16 \sum_{k=1}^{n-1} \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)(n-k)\cos\left(k\Phi L\nu\right)|X|^2 \, d\nu
$$

$$
\approx \sum_{k=1}^{n-1} \frac{n-k}{k} \frac{L_{\mathrm{eff}}^2}{\beta_2 \pi L}, \tag{32}
$$

and

$$
16 \sum_{k=1}^{n-1} \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)(n-k)\cos\left(k\Phi L\nu\right)|Y|^2 \, d\nu
$$

$$
\approx \sum_{k=1}^{n-1} \frac{n-k}{k} \frac{b_2^2}{a_2^2 \beta_2 \pi L}. \tag{33}
$$

The cross-term is rewritten as

$$
32 \sum_{k=1}^{n-1} \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)(n-k)\cos\left(k\Phi L\nu\right)
$$

$$
\times \left[\Re(X)\Re(Y) + \Im(X)\Im(Y)\right] d\nu
$$

$$
= 32 b_2 \sum_{k=1}^{n-1} \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)(n-k)\cos\left(k\Phi L\nu\right)
$$

$$
\times \left[\frac{(a_2+a)\Phi\nu \sin\left(\Phi\nu L\right) + \Phi^2\nu^2(e^{-aL} + e^{-a_2 L})}{(a^2+\Phi^2\nu^2)(a_2^2+\Phi^2\nu^2)}\right.
$$

$$
+ \frac{-\Phi^2\nu^2\cos\left(\Phi\nu L\right) - a a_2(e^{-aL}+e^{-a_2 L})}{(a^2+\Phi^2\nu^2)(a_2^2+\Phi^2\nu^2)}
$$

$$
\left. + \frac{a a_2 \cos\left(\Phi\nu L\right)}{(a^2+\Phi^2\nu^2)(a_2^2+\Phi^2\nu^2)}\right] d\nu, \tag{34}
$$

and solved term by term with replacing the logarithm by a constant. The first term of (34) can be approximately solved using $\sin(x)\cos(kx) = \frac{1}{2}\sin[(k-1)x] + \frac{1}{2}\sin[(k+1)x]$ and (44)

$$
32 b_2 (a_2+a) \ln\left(\frac{B}{2\sqrt{\nu_3}}\right) \sum_{k=1}^{n-1}(n-k)
$$

$$
\times \int_0^\infty \frac{\Phi\nu \sin(\Phi L\nu)\cos(k\Phi L\nu)}{(a^2+\Phi^2\nu^2)(a_2^2+\Phi^2\nu^2)} d\nu
$$

$$
= 16 b_2 (a_2+a) \ln\left(\frac{B}{2\sqrt{\nu_3}}\right) \sum_{k=1}^{n-1}(n-k)
$$

$$
\times \int_0^\infty \Phi\nu \frac{\sin[(k-1)\nu] + \sin[(k+1)\nu]}{(a^2+\Phi^2\nu^2)(a_2^2+\Phi^2\nu^2)} d\nu
$$

$$
= \frac{16 b_2 \pi}{\Phi(a-a_2)} \ln\left(\frac{B}{2\sqrt{\nu_3}}\right) \sum_{k=1}^{n-1}(n-k)
$$

$$
\times \left[e^{-aLk}\cosh(aL) - e^{-a_2 Lk}\cosh(a_2 L)\right], \tag{35}
$$

with $\nu_3 = \frac{5\pi}{\Phi L}$. We observe that the sums in (35) converge quickly for typical values of the argument in the exponent. A negligible error is introduced when the sum is extended to infinity and the sum can be written in closed-form as $\sum_{k=1}^n e^{-ak} \approx \sum_{k=1}^\infty e^{-ak} = \frac{1}{e^a - 1}$ and $\sum_{k=1}^n k e^{-ak} \approx \sum_{k=1}^\infty k e^{-ak} = \frac{0.5}{\cosh(a)-1}$. To further simplify the result, it is assumed that $2\cosh(aL) \approx e^{aL}$ and $\exp(aL) \gg 1$. Expression

(35) then yields

$$
\frac{16 b_2 \pi}{\Phi(a - a_2)} \ln\left(\frac{B}{2\sqrt{\nu_3}}\right) \sum_{k=1}^{n-1} (n-k)
$$
$$
\times \left[e^{-aLk}\cosh(aL) - e^{-a_2 Lk}\cosh(a_2 L)\right]
$$
$$
\approx \frac{\pi 8 b_2 (n-1)}{\Phi(a - a_2)} \ln\left(\frac{B}{2\sqrt{\nu_3}}\right)\left[\frac{\exp(aL)}{e^{aL}-1} - \frac{\exp(a_2 L)}{e^{a_2 L}-1}\right].
$$
(36)

The second term of expression (21) is approximately solved using (46) as

$$
32 b_2 \ln\left(\frac{B}{2\sqrt{\nu_2}}\right) \sum_{k=1}^{n-1}(n-k)
$$
$$
\times \int_0^\infty \frac{\left(e^{-aL} + e^{-a_2 L}\right)\Phi^2\nu^2 \cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
= \frac{16 b_2 \pi}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\left(e^{-aL} + e^{-a_2 L}\right)
$$
$$
\times \sum_{k=1}^{n-1}(n-k)\left(a e^{-aLk} - a_2 e^{-a_2 Lk}\right)
$$
$$
\approx \frac{16 b_2 (n-1)\pi}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)
$$
$$
\times \left[\frac{a\left(e^{-aL} + e^{-a_2 L}\right)}{e^{aL}-1} - \frac{a_2\left(e^{-aL} + e^{-a_2 L}\right)}{e^{a_2 L}-1}\right],
$$
(37)

with $\nu_2 = \frac{\pi}{4\Phi L}$. The third term of expression (21) is approximately solved using (46)

$$
-32 b_2 \ln\left(\frac{B}{2\sqrt{\nu_3}}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \int_0^\infty \frac{\Phi^2\nu^2 \cos(\Phi L\nu)\cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
= -16 b_2 \ln\left(\frac{B}{2\sqrt{\nu_3}}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \int_0^\infty \Phi^2\nu^2 \frac{\cos[(k-1)\nu] + \cos[(k+1)\nu]}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
= -\frac{8 b_2 \pi}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_3}}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \left[a e^{-aLk}\cosh(aL) - a_2 e^{-a_2 Lk}\cosh(a_2 L)\right]
$$
$$
\approx -\frac{8 b_2 \pi(n-1)}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_3}}\right)
$$
$$
\times \left[\frac{a\exp(aL)}{e^{aL}-1} - \frac{a_2\exp(a_2 L)}{e^{a_2 L}-1}\right].
$$
(38)

The fourth term of expression (21) is approximately solved using (48)

$$
32 b_2 a a_2 \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\left(-e^{-aL} - e^{-a_2 L}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \int_0^\infty \frac{\cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
= -\frac{16 b_2 \pi}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\left(e^{-aL} + e^{-a_2 L}\right)
$$
$$
\times \sum_{k=1}^{n-1}(n-k)\left(a e^{-a_2 Lk} - a_2 e^{-aLk}\right)
$$
$$
\approx -\ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\frac{16 b_2 \pi(n-1)}{\Phi(a^2 - a_2^2)}
$$
$$
\times \left[\frac{a\left(e^{-aL} + e^{-a_2 L}\right)}{e^{a_2 L}-1} - \frac{a_2\left(e^{-aL} + e^{-a_2 L}\right)}{e^{aL}-1}\right].
$$
(39)

The fifth term of expression (21) is approximately solved using (48) and $\cos(x)\cos(kx) = \frac{1}{2}\cos[(k-1)x] + \frac{1}{2}\cos[(k+1)x]$ and (48). Similar to the derivation of η_1, the integration domain is split into two domains in order to improve the accuracy of the approximation (replacing the logarithm with a constant) as this term does not vanish for small ν.

$$
32 b_2 a a_2 \sum_{k=1}^{N_s - 1}(n-k)
$$
$$
\times \int_0^\infty \ln\left(\frac{B}{2\sqrt{\nu}}\right)\frac{\cos(\Phi L\nu)\cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
\approx 32 b_2 a a_2 \sum_{k=1}^{n-1}(n-k)\left[\int_{B_k}^\infty \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\right.
$$
$$
\times \frac{\cos(\Phi L\nu)\cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
+ \left. \int_0^{B_k} \ln\left(\frac{B}{2\sqrt{\nu}}\right)\frac{\cos(\Phi L\nu)\cos(k\Phi L\nu)}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu\right]
$$
$$
\approx 16 b_2 a a_2 \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \int_0^\infty \frac{\cos[(k-1)\Phi L\nu] + \cos[(k+1)\Phi L\nu]}{(a^2 + \Phi^2\nu^2)(a_2^2 + \Phi^2\nu^2)} d\nu
$$
$$
+ \frac{16 b_2}{a a_2}\sum_{n=1}^{n-1}(n-k)\int_0^{B_k}\ln\left(\frac{\nu_2}{\nu}\right) d\nu
$$
$$
= \frac{16 b_2 \pi}{\Phi(a^2 - a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\sum_{k=1}^{n-1}(n-k)
$$
$$
\times \left[a e^{-a_2 Lk}\cosh(a_2 L) - a_2 e^{-aLk}\cosh(aL)\right]
$$
$$
+ \frac{16 b_2}{a a_2}\sum_{k=1}^{n-1}(n-k) B_k\left[\ln\left(\frac{\nu_2}{B_k}\right) + 1\right]
$$

$$\approx \frac{8b_2\pi\,(n-1)}{\Phi\,(a^2-a_2^2)} \ln\left(\frac{B}{2\sqrt{\nu_2}}\right)\left[\frac{a\exp(a_2L)}{e^{a_2L}-1} - \frac{a_2\exp(aL)}{e^{aL}-1}\right]$$

$$+ \frac{8b_2B_0}{aa_2}\sum_{k=1}^{n-1}\frac{n-k}{k}\left[\ln(4k)+1\right], \tag{40}$$

with $B_k = \frac{B_0}{2k} = \frac{\pi}{16k\Phi L}$, where we divide by 2 due to the multiplication of two cosines and by k due to presence of k within the argument of the cosine.

Summing up the terms in (32), (33), all cross-terms, multiplying by $\frac{16\gamma^2}{27R_b^2}$, and setting $n=20$ yields

$$\eta_{\text{coh},20} = \frac{160\gamma^2}{27\pi\,|\beta_2|\,R_b^2}\left\{\frac{26}{5}\frac{L_{\text{eff}}^2 a_2^2 + b_2^2}{a_2^2 L} + \frac{171}{40}\frac{b_2}{Laa_2}\right.$$

$$\left.+ \frac{19b_2}{5}\left[\frac{\ln(4L\phi)\left(e_1^2 - e_2^2\right)}{(a-a_2)\hat{e}_1\hat{e}_2} + \frac{7}{5}\frac{a\hat{e}_1 - a_2\hat{e}_2}{(a^2-a_2^2)\hat{e}_1\hat{e}_2}\right]\right\}.$$
$$\tag{41}$$

Applying (30) yields the coherence factor ϵ in (10).

APPENDIX C
INTEGRALS

This section contains the integrals that were used during the derivation of the closed-form expression.

$$\int_0^\infty \frac{1}{a^2+x^2}dx = \frac{\pi}{2}\frac{1}{a}, \tag{42}$$

$$\int_0^\infty \frac{\ln(x)}{a^2+x^2}dx = \frac{\pi}{2}\frac{\ln(a)}{a}, \tag{43}$$

$$\int_0^\infty \frac{x\sin(ax)}{(b^2+x^2)(c^2+x^2)}dx = \frac{\pi}{2}\frac{e^{-ab}-e^{-ac}}{c^2-b^2}, \tag{44}$$

$$\int_0^\infty \frac{x^2}{(b^2+x^2)(c^2+x^2)}dx = \frac{\pi}{2}\frac{1}{b+c}, \tag{45}$$

$$\int_0^\infty \frac{x^2\cos(ax)}{(b^2+x^2)(c^2+x^2)}dx = \frac{\pi}{2}\frac{be^{-ab}-ce^{-ac}}{b^2-c^2}, \tag{46}$$

$$\int_0^\infty \frac{1}{(b^2+x^2)(c^2+x^2)}dx = \frac{\pi}{2}\frac{1}{b^2c+bc^2}, \tag{47}$$

$$\int_0^\infty \frac{\cos(ax)}{(b^2+x^2)(c^2+x^2)}dx = \frac{\pi}{2bc}\frac{be^{-ac}-ce^{-ab}}{b^2-c^2}, \tag{48}$$

$$\int_0^{x_0}\ln\left(\frac{a}{x}\right)dx = x_0\ln\left(\frac{a}{x_0}\right)+x_0, \tag{49}$$

ACKNOWLEDGMENT

REFERENCES

[1] T. Naito, T. Tanaka, K. Torii, N. Shimojoh, H. Nakamoto, and M. Suyama, "A broadband distributed Raman amplifier for bandwidths beyond 100 nm," *in Proc. Opt. Fiber Commun. Conf.*, 2002, pp. 1–3.

[2] V. Curri and A. Carena, "Merit of Raman pumping in uniform and uncompensated links supporting NyWDM transmission," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 554–565, Jan. 2016.

[3] T. Hasegawa, Y. Yamamoto, and M. Hirano, "Optimal fiber design for large capacity long haul coherent transmission [Invited]," *Opt. Express*, vol. 25, no. 2, pp. 706–712, 2017.

[4] D. Semrau *et al.*, "Achievable information rates estimates in optically amplified transmission systems using nonlinearity compensation and probabilistic shaping," *Opt. Lett*, vol. 42, no. 1, pp. 121–124, 2017.

[5] N. A. Shevchenko *et al.*, "Achievable information rates estimation for 100-nm Raman-amplified optical transmission system," in *Proc. Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.

[6] G. Bosco, P. Poggiolini, A. Carena, V. Curri, and F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection," *Opt. Express*, vol. 19, no. 26, pp. B440–B451, 2011.

[7] V. Anagnostopoulos, C. Politi, C. Matrakidis, and A. Stavdas, "Physical layer impairment aware wavelength routing algorithms based on analytically calculated constraints," *Opt. Commun.* vol. 270, no. 2, pp. 247–254, 2006.

[8] A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. Eur. Conf. Opt. Commun.*, 1993, pp. 1–3.

[9] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25685–25699, 2013.

[10] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Express*, vol. 18, no. 18, pp. 19039–19054, 2010.

[11] M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[12] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16335–16362, 2014.

[13] G. Saavedra *et al.*, "Experimental investigation of nonlinear signal distortions in ultra-wideband transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, 2017, pp. 1–3.

[14] A. Nespola *et al.*, "GN-model validation over seven fiber types in uncompensated PM-16QAM Nyquist-WDM links," *Photon. Technol. Lett.*, vol. 26, no. 2, pp. 206–209, 2014.

[15] A. Nespola *et al.*, "Experimental validation of the EGN-model in uncompensated optical links," in *Proc. Opt. Fiber Commun. Conf.*, 2015, pp. 1–3.

[16] L. Galdino *et al.*, "Experimental demonstration of modulation-dependent nonlinear interference in optical fibre communication," in *Proc. Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.

[17] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A Simple and effective closed-form GN model correction formula accounting for signal non-Gaussian distribution," *J. Lightw. Technol.*, vol. 33, no. 2, pp. 459–473, Jan. 2015.

[18] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[19] V. Curri, A. Carena, P. Poggiolini, G. Bosco, and F. Forghieri, "Extension and validation of the GN model for non-linear interference to uncompensated links using Raman amplification," *Opt. Express*, vol. 21, no. 3, pp. 3308–3317, 2013.

[20] D. Lavery, R. Maher, G. Liga, D. Semrau, L. Galdino, and P. Bayvel, "On the bandwidth dependent performance of split transmitter-receiver optical fiber nonlinearity compensation," *Opt. Express*, vol. 25, no. 4, pp. 4554–4563, 2017.

[21] L. Galdino *et al.*, "On the limits of digital back-propagation in the presence of transceiver noise," *Opt. Express*, vol. 25, no. 4, pp. 4564–4578, 2017.

[22] B. Li *et al.*, "The role of effective area in the design of weakly coupled MCF: Optimization guidance and OSNR improvement," *IEEE J. Sel. Topics Quantum Electron.*, vol. 22, no. 2, pp. 81–87, Mar./Apr. 2016.

[23] I. F. de Jauregui Ruiz, A. Ghazisaeidi, and G. Charlet, "Optimization rules and performance analysis of filtered digital backpropagation," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.

[24] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities [Invited]," *Opt. Express*, vol. 25, no. 3, pp. 1916–1936, 2017.

[25] A. Ghazisaeidi *et al.*, "Submarine transmission systems using digital nonlinear compensation and adaptive rate forward error correction," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1886–1895, Apr. 2016.

[26] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a non-linear optical channel with finite memory," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

[27] W. S. Pelouch, "Raman amplification: An enabling technology for long-haul coherent transmission systems," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 6–19, Jan. 2016.

[28] M. A. Z. Al-Khateeb, M. Tan, M. A. Iqbal, M. McCarthy, P. Harper, and A. D. Ellis, "Four wave mixing in distributed Raman amplified optical transmission systems," in *Proc. 2016 IEEE Photon. Conf.*, Waikoloa, HI, USA, 2016, pp. 795–796.

[29] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, "Impact of the transmitted signal initial dispersion transient on the accuracy of the GN model of non-linear propagation," in *Proc. Eur. Conf. Opt. Commun.*, 2013, pp. 1–3.

[30] J. Bromage, "Raman amplification for fiber communications systems," *J. Lightw. Technol.*, vol. 22, no. 1, pp. 79–93, Jan. 2004.

[31] V. Curri, A. Carena, P. Poggiolini, G. Bosco, and F. Forghieri, "Evaluation of non-linear interference in uncompensated links using Raman amplification," in *Proc. Eur. Conf. Opt. Commun.*, 2012, pp. 1–3.

[32] R. Maher, A. Alvarado, D. Lavery, and P. Bayvel, "Increasing the information rates of optical communications via coded modulation: A study of transceiver performance," *Sci. Rep.*, vol. 6, 2016, Art. no. 21278.

[33] G. Saavedra, D. Semrau, L. Galdino, R. I. Killey, and P. Bayvel, "Digital back-propagation for nonlinearity mitigation in distributed Raman amplified links," *Opt. Express*, vol. 25, no. 5, pp. 5431–5439, 2017.

[34] M. Tan, P. Rosa, S. T. Le, Md. A. Iqbal, I. D. Phillips, and P. Harper, "Transmission performance improvement using random DFB laser based Raman amplification and bidirectional second-order pumping," *Opt. Express*, vol. 24, no. 3, pp. 2215–2221, 2016.

[35] A. D. Ellis *et al.*, "4 Tb/s transmission reach enhancement using 10×400 Gb/s super-channels and polarization insensitive dual band optical phase conjugation," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1717–1723, Apr. 2016.

[36] J. D. Ania-Castañón, "Quasi-lossless transmission using second-order Raman amplification and fibre Bragg gratings," *Opt. Express*, vol. 12, no. 19, pp. 4372–4377, 2004.

[37] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation," *J. Lightw. Technol.*, vol. 33, no 5, pp. 1044–1053, Mar. 2015.

[38] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *J. Lightw. Technol.*, vol. 34, no 2, pp. 593–607, Jan. 2016.

[39] P. Poggiolini and Y. Jiang, "Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems," *J. Lightw. Technol.*, vol. 35, no 3, pp. 458–480, Feb. 2017.

[40] M. P. Yankov *et al.*, "Nonlinear phase noise compensation in experimental WDM systems with 256QAM," *J. Lightw. Technol.*, vol. 35, no 8, pp. 1438–1443, Apr. 2017.

Daniel Semrau received the M.Sc. degree in photonic networks engineering from Scuola Superiore Sant'Anna, Pisa, Italy and Aston University, Birmingham, U.K., in 2015. In 2015, he joined the Optical Networks Group, University College London, London, U.K., where he is currently working toward the Ph.D. degree. His research interests include channel modeling and ultrawideband transmission for long-haul coherent optical communications.

Gabriel Saavedra, biography not available at the time of publication.

Domaniç Lavery, biography not available at the time of publication.

Robert I. Killey, biography not available at the time of publication.

Polina Bayvel, biography not available at the time of publication.