Perturbation Analysis of Nonlinear Propagation in a Strongly Dispersive Optical Communication System

Pontus Johannisson and Magnus Karlsson

***Abstract*—We discuss an analytical model that predicts the impact of the Kerr nonlinearity in optical communication systems when the signal spectrum is wide and the accumulated dispersion during propagation is large. A detailed derivation of this model is given for a generalized system by means of a perturbation analysis of the Manakov equation with attenuation, gain, and third order dispersion included. As in the case with previous studies, three simplifying assumptions are necessary. These are that (i) the nonlinearity is weak, (ii) the input signal is of a given specific form, and (iii) the signal-noise interaction can be neglected. Under these assumptions, the result is found exactly. We also discuss the accuracy of the analytical result and show that third order dispersion has a small impact in practice.**

***Index Terms*—Communication system nonlinearities, nonlinear optics, optical fiber communication, wavelength division multiplexing.**

I. Introduction

In a fiber-optic transmission system operating using a wideband optical signal, the propagating signal is rapidly distorted due to chromatic dispersion (CD). For a polarization-multiplexed signal it has been found numerically that after a relatively short distance of propagation, the probability density function of all four real components of the electric field become Gaussian with zero mean and a variance related to the signal power [1]. This can be shown analytically by using the central limit theorem and is also intuitively understandable since the dispersed signal at every point in time can be viewed as a coherent superposition of many independent signal pulses. During the propagation, a nonlinear phase shift is induced in proportion to the local power by the Kerr nonlinearity. In modern systems, the CD is often compensated for in the receiver but after this step there will be residual signal distortion due to nonlinear effects. It has been shown both numerically and experimentally that this distortion has a statistical distribution that typically is very close to Gaussian [1], [2], also in the absence of any amplifier noise. This observation suggests that if no attempt is made to compensate for the nonlinear effects, then the nonlinear signal distortion should be modeled statistically. Such an approach is relevant in practice for systems that transmit a wide signal spectrum and perform CD compensation without any nonlinear compensation in the receiver. However, it should be noticed that the accuracy of this approach depends on the system parameters. As an example, it has been shown that the distribution for a 16 quadrature amplitude modulation signal deviates from the Gaussian distribution in a single-channel transmission [3], but we expect that by introducing interchannel nonlinear effects by using wavelength division multiplexing (WDM), the deviation from a Gaussian distribution would be smaller.

Four wave mixing (FWM) has been investigated thoroughly since it is an important limitation in optical communication systems. For example, WDM systems were analyzed already before the wide-spread use of erbium-doped fiber amplifiers (EDFAs) by calculating the FWM between individual frequency components [4]. Inoue has investigated FWM in systems with periodic amplification and varying amount of dispersion [5]–[7], and this work has also been generalized, e.g., to describe the FWM in quasi-distributed erbium-doped fiber [8]. While there are many publications covering the FWM process, there is also a need to describe the connection between the system geometry, the input signal, and the bit error rate (BER) analytically. Pioneering work in this respect was reported in [9]–[11] by calculating the power spectral density (PSD) of the nonlinear signal distortion, which in turn allowed the BER to be found. A statistical description of the FWM has also been given in, e.g., [12], but in this case it was not possible to completely describe the statistical distribution analytically. The FWM between the subcarriers of an OFDM system has been investigated in [13]–[15] and recently a similar analytical model was presented that predicts the noise-like nonlinear signal distortion for a system using polarization multiplexing and WDM [16], [17]. This distortion was called *nonlinear interference* (NLI) and recently an extensive paper was published, which discussed the theory and reported a comprehensive series of numerical simulations in order to test the analytical predictions [18]. Further comparison with numerical results were given in [19], and the theory also agrees well with experimental results [20].

In [16]–[18], a suggestion was given for how to model the signal, the FWM of the different signal spectral components was calculated, and the corresponding PSD was found. In this paper, we independently derive the corresponding result for a generalized system which allows, e.g., the inclusion of different types of fibers and simultaneous use of EDFAs and Raman amplification. The calculation is based on the Manakov equation with power gain, attenuation, and third order dispersion (TOD) included. The result [18, Eq. (18)] is then recovered as a special case. The fact that we provide a step-by-step derivation gives detailed insight into the model, see also [21]–[23].

Manuscript received September 06, 2012; revised January 01, 2013; accepted February 07, 2013. Date of publication February 11, 2013; date of current version February 27, 2013. This work was supported in part by the Swedish Research Council.

The authors are with the Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-41296 Göteborg, Sweden (e-mail: pontus.johannisson@chalmers.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2013.2246543

Previously, perturbation analysis has been used, e.g., to investigate intrachannel cross-phase modulation (XPM) and intrachannel FWM, which gives rise to "ghost pulses" in systems using on-off keying [24]–[26]. This approach, which coincides with the Volterra series approach [27], has later been used both to analytically study systems, [28]–[30], and to find a compensation scheme for the NLI, see for example [31]. We also mention that the signal variance due to nonlinear effects has been investigated in a different way by means of a discrete channel model [32] and that the fact that, under the assumption of weak nonlinear effects [28] and negligible signal-noise interaction, the NLI scales as the cube of the signal power allows some general statements to be made about system optimization [33].

The organization of this paper is as follows: In Section II, the perturbation analysis is introduced and a formal solution that is valid for any input signal and system is given. In Section III, we find the solution corresponding to a specific input signal, which is then used in Section IV to calculate the NLI PSD. We then obtain the result from [18, Eq. (18)] by selecting a specific system in Section V. The analysis is discussed in Section VI and finally we conclude.

II. Perturbation Analysis

In order to treat the Manakov equation analytically, we must make some simplifying assumptions. The first of these is that the nonlinear term only gives rise to a small signal distortion and this allows us to perform a perturbation analysis. This assumption is present also in [16] and is stated as "the pump is undepleted". We proceed by introducing the complex envelope of the electric field in the x and y polarizations according to $\mathbf{A} = (A^x, A^y)^{\mathrm{T}}$ and writing this as the sum of the linearly propagating signal, i.e., the signal in the absence of Kerr nonlinearity, and a small perturbation. This first assumption implies that the nonlinear effects may not become significant. The second assumption is that the input signal can be written on a specific form in the frequency domain, [16], which is stated in detail in Section III. The range of validity for this signal model is a separate question, which is outside the scope of this work. The third assumption is that the signal-noise interaction can be neglected. Also this assumption is present in [16] and noise corresponding to the total noise from the amplification is instead added just before the receiver. These three assumptions are sufficient to be able to carry out the analysis.

A. The Perturbation Equation From the Manakov Equation

To describe the signal transmission, we start from the Manakov equation [34] and include power gain, attenuation, and TOD [35]. If necessary, even higher order dispersion can be included and the way to do this should be clear from the derivation below. It should be noted that the Manakov equation is obtained by averaging over the polarization rotations which are assumed to be fast and this equation does not take polarization mode dispersion into account. Denoting the lowest order group-velocity dispersion by $\beta_2(z)$, the TOD parameter by $\beta_3(z)$, the power gain by $g(z)$, and the power attenuation by $\alpha(z)$ we have[1]

$$i\frac{\partial \mathbf{A}}{\partial z} = \frac{\beta_2(z)}{2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - i\frac{\beta_3(z)}{6}\frac{\partial^3 \mathbf{A}}{\partial t^3} - \gamma(z)(\mathbf{A}^{\mathrm{H}}\mathbf{A})\mathbf{A} + i\frac{g(z)-\alpha(z)}{2}\mathbf{A}, \quad (1)$$

where $\mathbf{A}^{\mathrm{H}}\mathbf{A} = |A^x|^2 + |A^y|^2$ is the sum of the power in the x and y polarizations and the nonlinear parameter $\gamma(z) = (8/9)(k_0 n_2/A_{\mathrm{eff}}(z))$. In the expression for γ, k_0 is the wavenumber corresponding to the center frequency, n_2 is the Kerr coefficient, and A_{eff} is the effective area of the optical fiber. The power gain $g(z)$, which can be set up using EDFAs and/or Raman amplification, is assumed to have no frequency dependence, i.e., the gain is flat over the bandwidth of the signal. Note that all parameters introduced in (1) have an arbitrary z-dependence, which is not specified at this stage. For notational compactness, this z-dependence is often implicit below. The perturbation is introduced according to

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1 = \begin{pmatrix} A_0^x \\ A_0^y \end{pmatrix} + \begin{pmatrix} A_1^x \\ A_1^y \end{pmatrix}, \quad (2)$$

where $\mathbf{A}_0$ solves the linear equation obtained from (1) by setting $\gamma = 0$. It should be noticed that we, by writing $\mathbf{A}$ in this way, only include the first-order perturbation. If necessary, the perturbation analysis can be generalized to any order [27]. Our intention is to study A^x and we will consider the two cases that either (i) $|A_0^x|$ and $|A_0^y|$ are of the same order of magnitude (transmission using polarization multiplexing) or (ii) $A_0^y = 0$ (single-polarization transmission). The first assumption above can then be strictly formulated as $|A_1^x| \ll |A_0^x|$ and $|A_1^y| \ll |A_0^x|$. This allows us to approximate the nonlinear term in (1) to leading order according to

$$(\mathbf{A}^{\mathrm{H}}\mathbf{A})\mathbf{A} \approx (|A_0^x|^2 + |A_0^y|^2)\begin{pmatrix} A_0^x \\ A_0^y \end{pmatrix}. \quad (3)$$

Substituting (2) into (1) and using that $\mathbf{A}_0$ by definition solves

$$i\frac{\partial \mathbf{A}_0}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}_0}{\partial t^2} - i\frac{\beta_3}{6}\frac{\partial^3 \mathbf{A}_0}{\partial t^3} + i\frac{g-\alpha}{2}\mathbf{A}_0. \quad (4)$$

gives

$$i\frac{\partial \mathbf{A}_1}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}_1}{\partial t^2} - i\frac{\beta_3}{6}\frac{\partial^3 \mathbf{A}_1}{\partial t^3} - \gamma(\mathbf{A}_0^{\mathrm{H}}\mathbf{A}_0)\mathbf{A}_0 + i\frac{g-\alpha}{2}\mathbf{A}_1. \quad (5)$$

We study the x-polarized perturbation, which is described by

$$\frac{\partial A_1^x}{\partial z} + i\frac{\beta_2}{2}\frac{\partial^2 A_1^x}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A_1^x}{\partial t^3} - \frac{g-\alpha}{2}A_1^x = S, \quad (6)$$

[1]To aid the comparison, we use the same Fourier transform definition as [18], see also, e.g., [36]. Thus, $\tilde{u}(f) = \int_{-\infty}^{\infty} u(t)e^{-i2\pi f t}dt$ and $u(t) = \int_{-\infty}^{\infty} \tilde{u}(f)e^{i2\pi f t}df$. Unfortunately, this is different from Agrawal, see [35, Eq. (3.2.6)]. The dispersion relation is defined in the frequency domain and this causes the TOD term in (1) to change sign compared to [35, Eq. (3.3.1)]. To see this we rewrite [35, Eq. (3.3.2)] with the here used Fourier transform definition and find the corresponding equation in the time domain. Thus, (1) is consistent with the conventional definition of the dispersion relation.

where, for convenience, we temporarily denote the source term by $S(z,t) = i\gamma(|A_0^x|^2 + |A_0^y|^2)A_0^x$.

B. Formal Solution

We proceed by deriving a general formal solution to (6) without making any assumptions about the input signal or the system parameters. To describe the power evolution, we introduce $P(z)$ as a function that satisfies the equation

$$\frac{dP}{dz} = [g(z) - \alpha(z)]P. \tag{7}$$

In the absence of any distributed amplification this function decreases as $e^{-\alpha z}$ between the amplifiers. The EDFA gain can be modeled by a δ-function in $g(z)$ to obtain the discontinuities in $P(z)$ at each z corresponding to the location of an amplifier. We define the accumulated dispersion up to the receiver, $B_2(z)$ and $B_3(z)$, by

$$B_2(z) = \int_0^z \beta_2(\zeta)\, d\zeta \quad \text{and} \quad B_3(z) = \int_0^z \beta_3(\zeta)\, d\zeta. \tag{8}$$

If there is lumped dispersion (such as a chirped fiber Bragg grating), which also can be modeled using δ-functions, then B_2 and B_3 will have discontinuities. The formal solution to (6), which is derived in Appendix A, is then

$$\tilde{A}_1^x(L,f) = \int_0^L \frac{\tilde{S}(z,f)}{\sqrt{p(z)}} e^{-i(2\pi f)^2(B_2(z)/2) - i(2\pi f)^3(B_3(z)/6)}\, dz, \tag{9}$$

where $\tilde{S}(z,f) = \mathcal{F}[i\gamma(|A_0^x|^2 + |A_0^y|^2)A_0^x]$ and Fourier transformation is denoted either by $\mathcal{F}$ or tilde. In this expression L is the total system length (possibly containing many fibers and amplifiers) and $p(z) = P(z)/P(0)$. It has been assumed that the amplification perfectly compensates for all losses and that perfect dispersion compensation is carried out in the receiver. However, the latter only affects the spectrum phase and not the PSD of the perturbation.

III. Signal Model

In order to use the formal solution (9), we need to select the input signal and the system parameters, find the linear solution, and perform the integration. When trying to do this with a detailed modeling of the signal, the calculations become cumbersome and the obtained expression must then be averaged over the data to obtain the expected value for the perturbation. We here instead use the second assumption, i.e., we write the initial field in the frequency domain as suggested in [16]. This model is

$$\tilde{A}_0^x(0,f) = \sqrt{f_0} \sum_{k=-\infty}^{\infty} \xi_k \sqrt{G_0^x(kf_0)}\, \delta(f - kf_0), \tag{10}$$

$$\tilde{A}_0^y(0,f) = \sqrt{f_0} \sum_{k=-\infty}^{\infty} \zeta_k \sqrt{G_0^y(kf_0)}\, \delta(f - kf_0), \tag{11}$$

where ξ_k and ζ_k are complex independent Gaussian random variables of unit variance and the input signal PSDs of the x and y polarizations are denoted by $G_0^x(f)$ and $G_0^y(f)$, respectively. It should be noticed that this signal model makes no attempt to describe the initial signal phase, but the expected value of the PSD agrees with the input signal. Furthermore, these signals are periodic in time since they consist of discrete frequency components spaced f_0 apart. However, the period will approach infinity as we later allow f_0 to approach zero. This signal model greatly simplifies the analytical calculations. For notational convenience, we will in the following suppress the infinite summation limits and the indication of the z-dependence. Furthermore, we will use the abbreviated notation $f_k \equiv kf_0, \omega_k \equiv 2\pi f_k$, and analogous expressions for l and m. The perturbation solution corresponding to this signal is derived in Appendix B as (46). The main steps are to first account for the linear effects during propagation and then insert the found expression into (9) and simplify the result. The solution is conveniently expressed in terms of the function

$$\mathcal{C}_{klm} \equiv \mathcal{C}(f_k, f_l, f_m) \equiv \int_0^L \gamma p e^{-i(\omega_k - \omega_m)(\omega_l - \omega_m)B_2} \times e^{-i(\omega_k + \omega_l)(\omega_k - \omega_m)(\omega_l - \omega_m)(B_3/2)}\, dz \tag{12}$$

and rearranging the summation and integration in (46), we find

$$A_1^x(L,t) = if_0^{3/2} \sum_{k,l,m} \mathcal{C}_{klm} e^{i(\omega_k + \omega_l - \omega_m)t} \xi_k \sqrt{G_0^x(f_k)} \times \left(\xi_l \xi_m^* \sqrt{G_0^x(f_l) G_0^x(f_m)} + \zeta_l \zeta_m^* \sqrt{G_0^y(f_l) G_0^y(f_m)} \right). \tag{13}$$

This is an infinite triple sum over k, l, and m. For later use we notice that

$$|\mathcal{C}_{klm}| \le \int_0^L \left| \gamma p e^{-i(\omega_k - \omega_m)(\omega_l - \omega_m)B_2} \times e^{-i(\omega_k + \omega_l)(\omega_k - \omega_m)(\omega_l - \omega_m)(B_3/2)} \right| dz \le \hat{\gamma}\hat{p}L, \tag{14}$$

where $\hat{p}$ and $\hat{\gamma}$ are the maximum values of $p(z)$ and $\gamma(z)$, respectively, in the interval $z \in [0, L]$. This means that $|\mathcal{C}_{klm}|$ is upper bounded by a constant as k, l, m, and f_0 are changed.

IV. Power Spectral Density

When the input signal is modeled by (10) and (11), the perturbation is given by (13). In order to find the NLI, we need to calculate the corresponding PSD. According to the Wiener-Khinchin theorem [36, p. 67], the PSD corresponding to $A_1^x(L,t)$ is the Fourier transform of the autocorrelation, i.e.,

$$G_1^x(f) = \mathcal{F}[R(\tau)] = \int_{-\infty}^{\infty} R(\tau) e^{-i2\pi f\tau}\, d\tau, \tag{15}$$

where

$$R(\tau) = \mathbb{E}\{A_1^x(L,t_1)(A_1^x(L,t_2))^*\} \tag{16}$$

and $\mathbb{E}\{\cdot\}$ denotes the expectation operator. As will be shown, the latter expression is a function of the time difference $\tau \equiv t_1 -$

t_2. We suppress the infinite limits for notational convenience. Using (13), we have

$$\begin{aligned}
&A_1^x(L,t_1)(A_1^x(L,t_2))^* \\
&= f_0^3 \sum_{\substack{k,l,m \\ k',l',m'}} \mathcal{C}_{klm}\mathcal{C}^*_{k'l'm'} \\
&\quad \times e^{i(\omega_k+\omega_l-\omega_m)t_1} e^{-i(\omega_{k'}+\omega_{l'}-\omega_{m'})t_2} \\
&\quad \times \xi_k \xi^*_{k'} \sqrt{G_0^x(f_k)}\sqrt{G_0^x(f_{k'})} \\
&\quad \times \left(\underbrace{\xi_l \xi^*_m \sqrt{G_0^x(f_l)G_0^x(f_m)}}_{1} + \underbrace{\zeta_l \zeta^*_m \sqrt{G_0^y(f_l)G_0^y(f_m)}}_{2} \right) \\
&\quad \times \left(\underbrace{\xi^*_{l'} \xi_{m'} \sqrt{G_0^x(f_{l'})G_0^x(f_{m'})}}_{3} + \underbrace{\zeta^*_{l'} \zeta_{m'} \sqrt{G_0^y(f_{l'})G_0^y(f_{m'})}}_{4} \right).
\end{aligned} \tag{17}$$

This is now a six-dimensional sum and we see that by expanding the brackets, the complete autocorrelation function will consist of four terms that each are products of the terms marked in (17). Thus, we write

$$R = R_{13} + R_{14} + R_{23} + R_{24}, \tag{18}$$

where the indices now indicate which of the terms in (17) that are being considered. Thus, as an example we have

$$\begin{aligned}
R_{13} = f_0^3 &\sum_{\substack{k,l,m \\ k',l',m'}} \mathcal{C}_{klm}\mathcal{C}^*_{k'l'm'} e^{i(\omega_k+\omega_l-\omega_m)t_1} e^{-i(\omega_{k'}+\omega_{l'}-\omega_{m'})t_2} \\
&\times \mathbb{E}\{\xi_k \xi_l \xi^*_m \xi^*_{k'} \xi^*_{l'} \xi_{m'}\} \\
&\times \sqrt{G_0^x(f_k)G_0^x(f_l)G_0^x(f_m)G_0^x(f_{k'})G_0^x(f_{l'})G_0^x(f_{m'})},
\end{aligned} \tag{19}$$

where we also used the fact the all terms except the product of stochastic variables are deterministic. The four terms in the autocorrelation will give rise to four different terms in the total PSD, which we denote by $G^x_{1,13}$, $G^x_{1,14}$, $G^x_{1,23}$, and $G^x_{1,24}$, respectively. We need to treat these four terms individually.

A. The Random Variables

As exemplified by (19), we need to carry out the expectation operation to obtain R. To do this, the theorem [37, Eq. (2.8–21)] can be directly used but we prefer to introduce this result by a discussion.

The ξ_k occurring in (10) have the properties $\mathbb{E}\{\xi_k\} = 0$, $\mathbb{E}\{\xi_k^2\} = 0$, $\mathbb{E}\{|\xi_k|^2\} = 1$, $\forall k$. Analogous expressions hold for ζ_k, $\forall k$ in (11). Using this, we can simplify the six-dimensional sums resulting from (17) in the following way. First assume that one of the summation variables, say k, has a value different from all other summation variables, i.e., k is unique. Due to the independence (which also means that the Gaussian random variables are uncorrelated), we then have

$$\mathbb{E}\{\xi_k \xi_l \xi^*_m \xi^*_{k'} \xi^*_{l'} \xi_{m'}\} = \mathbb{E}\{\xi_k\}\,\mathbb{E}\{\xi_l \xi^*_m \xi^*_{k'} \xi^*_{l'} \xi_{m'}\} = 0. \tag{20}$$

An identical argument holds if any of the ξ are replaced by ζ. This implies that the expected value is zero when one of the summation variables is unique. Thus, no summation variable can have a unique value. Second we assume that no summation variable is unique, but there are three pairwise equal values, where each pair has a unique value. Assume for example that $k = l$, $k' = l'$, $m = m'$. Then

$$\begin{aligned}
\mathbb{E}\{\xi_k \xi_l \xi^*_m \xi^*_{k'} \xi^*_{l'} \xi_{m'}\} &= \mathbb{E}\{\xi_k \xi_k \xi^*_m \xi^*_{k'} \xi^*_{k'} \xi_m\} \\
&= \mathbb{E}\{\xi_k \xi_k\}\,\mathbb{E}\{\xi^*_{k'} \xi^*_{k'}\}\,\mathbb{E}\{\xi^*_m \xi_m\} \\
&= \mathbb{E}\{\xi_k^2\}\mathbb{E}\{(\xi^*_{k'})^2\}\mathbb{E}\{|\xi_m|^2\} = 0.
\end{aligned} \tag{21}$$

From this we conclude that each random variable must be paired up with the complex conjugated version of the same random variable. This conclusion reduces the dimensionality of the summation to three or less. However, the dimensionality cannot be less than three, because then $G_1^x \to 0$ as $f_0 \to 0$. To see this we first notice that

$$|G_1^x| \le |G^x_{1,13}| + |G^x_{1,14}| + |G^x_{1,23}| + |G^x_{1,24}|, \tag{22}$$

and since all terms behave similarly in this respect, we can study, say, $G^x_{1,13}$. Let us select the one-dimensional case $k = l = m = k' = l' = m'$. We then have

$$R_{13} = f_0^3 \sum_k |\mathcal{C}_{kkk}|^2 e^{i\omega_k \tau}\, \mathbb{E}\{|\xi_k|^6\}[G_0^x(f_k)]^3, \tag{23}$$

which gives

$$G^x_{1,13} = f_0^3 \sum_k |\mathcal{C}_{kkk}|^2\, \mathbb{E}\{|\xi_k|^6\}[G_0^x(f_k)]^3 \delta(f - f_k). \tag{24}$$

We see that $G^x_{1,13} \to 0$ as $f_0 \to 0$ by identifying the Riemann sum and writing the expression as

$$\begin{aligned}
G^x_{1,13} &= f_0^2 \int |\mathcal{C}(f_1,f_1,f_1)|^2\, \mathbb{E}\{|\xi_k|^6\}[G_0^x(f_1)]^3 \delta(f - f_1)\, df_1 \\
&= f_0^2 |\mathcal{C}(f,f,f)|^2\, \mathbb{E}\{|\xi_k|^6\}[G_0^x(f)]^3.
\end{aligned} \tag{25}$$

Remembering that the value of $|\mathcal{C}|^2$ is upper bounded by a constant, the fact that $G^x_{1,13} \to 0$ for $f_0 \to 0$ is now obvious. An analogous argument can be made for the case of a two-dimensional sum. We conclude that the summation must have dimension exactly three, i.e., each random variable must be paired up with the complex conjugated version of the same random variable and all three pairs must have unique values.

B. The Perturbation PSD

The main result of this work is the general PSD of the perturbation, i.e., the sum of the terms in (18), which are calculated individually in Appendix C. The result is

$$\begin{aligned}
G_1^x &= G^x_{1,13} + G^x_{1,14} + G^x_{1,23} + G^x_{1,24} \\
&= 2 \int\int |\mathcal{C}(f_1,f_2,f_1+f_2-f)|^2 \\
&\quad \times G_0^x(f_1)G_0^x(f_2)G_0^x(f_1+f_2-f)\, df_1 df_2 \\
&\quad + \int\int |\mathcal{C}(f_1,f_2,f_1+f_2-f)|^2 \\
&\quad \times G_0^x(f_1)G_0^y(f_2)G_0^y(f_1+f_2-f)\, df_1 df_2 \\
&\quad + \mathcal{C}_0^2 (2P_x + P_y)^2 G_0^x(f),
\end{aligned} \tag{26}$$

where P_x and P_y are the average powers of the x and y polarizations, respectively, and $\mathcal{C}_0$ is defined by (54) in Appendix C-A2. We notice that the x polarization acting on itself is twice as effective as the y polarization acting on the x polarization. As is clear from the derivation, the reason for this is the level of degeneracy. The final term of (26) should be of no consequence for the transmission. As discussed in Appendix C, the origin of this term is the phase modulation from the entire propagating field acting on itself, which in the approximate perturbation analysis changes the PSD. In practice, we expect this to just cause a rotation of the received signal constellation. Evaluating (12) using the arguments in (26) we find

$$\mathcal{C}(f_1,f_2,f_1+f_2-f) = \int_0^L \gamma(z)p(z)e^{-i4\pi^2(f_1-f)(f_2-f)[B_2(z)+\pi(f_1+f_2)B_3(z)]}\,dz. \tag{27}$$

It should be noticed that $|\mathcal{C}|^2$ is a measure of the FWM efficiency and that $\mathcal{C}$ is determined when the physical parameters of the channel have been selected. Then, by choosing the input signal PSD we obtain the PSD of the NLI. The fact that the system and the input signal do not need to be jointly optimized is an interesting consequence of the model.

V. System Example

The result (26) is valid for a general system which, e.g., could contain different types of fiber and both EDFA and Raman amplification. However, in order to recover [18, Eq. (18)] we select the system to consist of N spans, each containing a standard single-mode fiber (SMF) followed by an EDFA. There are no dispersion-compensating fibers or any other optical dispersion compensation. Instead, the CD is compensated for by using digital signal processing in the receiver. The first span starts at $z=z_0=0$ and the last span ends at $z=z_N$. The fiber parameters α, β_2, β_3, and γ are assumed to have no z-dependence. Furthermore, we assume that the two polarization-multiplexed signals have the same PSD, i.e., we set $G_0^y=G_0^x$ and we remove the terms that are not due to FWM to rewrite (26) as

$$G_1^x = 3\int\int |\mathcal{C}(f_1,f_2,f_1+f_2-f)|^2 \times G_0^x(f_1)G_0^x(f_2)G_0^x(f_1+f_2-f)\,df_1df_2. \tag{28}$$

For this system, we have $B_2(z)=\beta_2 z$, $B_3(z)=\beta_3 z$, and temporarily using $\kappa \equiv 4\pi^2(f_1-f)(f_2-f)[\beta_2+\pi(f_1+f_2)\beta_3]$ we find

$$\begin{aligned}\mathcal{C}(f_1,f_2,f_1+f_2-f) &= \int_0^L \gamma p(z)e^{-i\kappa z}\,dz = \gamma\sum_{n=1}^{N}\int_{z_{n-1}}^{z_n} e^{-\alpha(z-z_{n-1})}e^{-i\kappa z}\,dz \\ &= \frac{\gamma}{\alpha+i\kappa}(1-e^{-\alpha\ell}e^{-i\kappa\ell})\frac{1-e^{-i\kappa\ell N}}{1-e^{-i\kappa\ell}},\end{aligned} \tag{29}$$

where we also used the assumption that all SMFs have the same length, denoted by $\ell = z_n - z_{n-1}$. We find

$$|\mathcal{C}(f_1,f_2,f_1+f_2-f)|^2 = \gamma^2\left|\frac{1-e^{-(\alpha+i\kappa)\ell}}{\alpha+i\kappa}\right|^2 \frac{\sin^2\left(\frac{\kappa\ell N}{2}\right)}{\sin^2\left(\frac{\kappa\ell}{2}\right)} \tag{30}$$

and

$$G_1^x = 3\gamma^2\int\int\left|\frac{1-e^{-(\alpha+i\kappa)\ell}}{\alpha+i\kappa}\right|^2 \frac{\sin^2\left[\frac{\kappa\ell N}{2}\right]}{\sin^2\left[\frac{\kappa\ell}{2}\right]} \times G_0^x(f_1)G_0^x(f_2)G_0^x(f_1+f_2-f)\,df_1df_2. \tag{31}$$

In order to see that this expression is equivalent to [18, Eq. (18)], we must set $\beta_3=0$ and account for the fact that a different convention for the Manakov equation is used in [18]. Thus, we need to replace γ with $8\gamma/9$. Furthermore, the NLI PSD in [18, Eq. (18)] is the summation over both polarizations, i.e., we get $G_{\mathrm{NLI}}=G_1^x+G_1^y=2G_1^x$. Similarly, we have $G_{\mathrm{Tx}}=G_0^x+G_0^y=2G_0^x$. Finally, the attenuation α is defined in [18] using the amplitude, not the power. For the definition used here, we refer to (1) and (7). When these differences in notation are taken into account, the results become identical. We notice that we obtain the "coherent" result, cf. [18, Section IV-B].

VI. Discussion

The range of validity of the model is an important but complex and difficult question. One challenge is the sheer size of the set of systems for which the model can be applied, which is exemplified by the simulation effort reported in [18]. Another issue is that the three assumptions above have different implications for the validity. A fundamental question that deserves further investigation is the accuracy of the signal model, (10) and (11). However, this question is outside the scope of this work and we will focus on the statement "the nonlinearity is weak".

In a full system simulation, the total signal distortion is due not only to nonlinear effects, but also, e.g., to amplifier noise, inter-symbol interference, and receiver imperfections. A further complication is that the perturbation PSD is not directly observable as the situation is analogous to the case when measuring the optical signal-to-noise ratio with an optical spectrum analyzer: The noise PSD can only be observed outside the signal band. However, a similar approach can be used to obtain a direct numerical check of the first assumption above.

In order to separate the assumption about the signal model and the signal-noise interaction from the assumption of a weak nonlinearity, we suggest that the solution (26) is compared to the full numerical solution of (1) without amplifier noise and using an input signal of the assumed type, i.e., (10) and (11). An example simulation for this is seen in Fig. 1. The red solid lines are the analytical perturbation PSDs and the blue lines are the numerical total output signal PSD, which has been obtained by averaging the power of the spectrum over 10000 simulations. The input signals consists of three 28 Gbaud polarization-multiplexed WDM channels with rectangular spectra and a frequency spacing of 50 GHz for a simplified system consisting of a single 80-km-long SMF, with $\alpha=0.2\,\mathrm{dB/km}$, $D=16\,\mathrm{ps/(nm\,km)}$, $\beta_3=0\,\mathrm{ps^3/km}$, $\gamma=1.3\,\mathrm{W^{-1}km^{-1}}$, and without amplifier

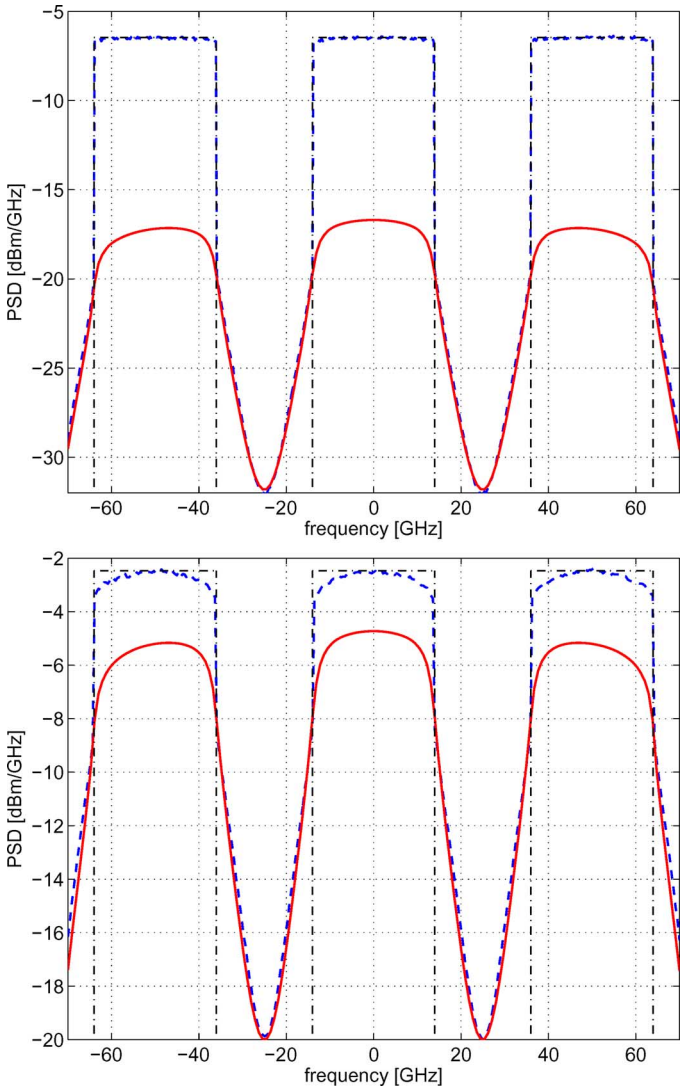

Fig. 1. A comparison of the total signal and the perturbation PSDs. The blue dashed line corresponds to the output signal and the red solid line is the analytical result. The black dash-dotted line (partially obscured) shows the input signal PSD. Top: Input power is 8 dBm per channel and polarization. Bottom: Input power is 12 dBm per channel and polarization. Notice the change of scale for the PSD.

noise. The input power is relatively high. In the top figure, it is 8 dBm per channel and polarization, giving a signal-to-noise ratio $\mathrm{SNR} \approx 10$ dB which is sufficient for quadrature phase-shift keying at a $\mathrm{BER} \approx 10^{-3}$. (Note that "SNR" here refers to the ratio of signal to NLI. Measuring in the middle of the center channel we find approximately 10.2 dB.) In the bottom figure, the input power is 12 dBm giving an SNR of approximately 2.3 dB. Increasing the input power beyond 13 dBm will lead to the SNR dropping to below 0 dB. The input signal PSD is indicated by the partially obscured black dash-dotted lines. It is seen that, although the theoretical results are somewhat too low, the agreement between the numerical and analytical results is very good with an error below 10% in the 8 dBm case. However, vary careful inspection of the bottom figure shows that the slopes of the blue and red curves are not identical at the edges of the three channels and this leads us to conclude that within the three channels, the analytical model instead gives a

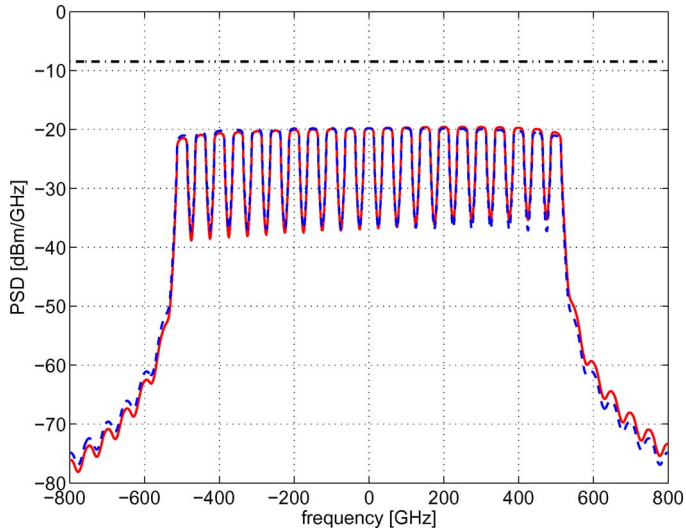

Fig. 2. The analytical PSD calculated for 21 WDM channels with (red solid line) and without (blue dashed line) TOD. Notice that β_3 has been increased ten times from the typical value in an SMF to get a noticeable difference. The black dash-dotted line shows the maximum value of the input signal PSD.

slightly too large value for the PSD. In the 12 dBm case there is also a clear rounding of the signal PSD due to the spectral broadening. At the edges of the center channel, the power loss is close to 1 dB. This illustrates that fact that the first assumption above is not valid for all input signals, but the SNR requirements should make this assumption sufficiently accurate in most cases of practical importance.

While it is true that TOD can, if left uncompensated, have an impact on a wide-band channel, it has a very weak impact on the perturbation PSD. To illustrate this we have extended the number of WDM channels in the system described above to 21 and compared the perturbation PSD with and without TOD. To keep the SNR above 10 dB, we have reduced the input power to 6 dBm per channel and polarization. We then find that using a typical value for the dispersion slope, $S = 0.07\ \mathrm{ps}/(\mathrm{nm}^2\ \mathrm{km})$, gives a very small effect. Increasing the value of β_3 by a factor of ten gives the result seen in Fig. 2. The results with (red solid line) and without (blue dashed line) TOD are still very similar. Most clear is the impact of TOD outside the signal band and gives rise to a clearly visible asymmetry in the spectrum. Within the signal band, the TOD changes the SNR by up to 0.6 dB with this high TOD value. The reason for this small impact is that with strong dispersive effects, which occur also with typical values for D, the TOD only changes the FWM efficiency slightly.

VII. Conclusion

We have derived the PSD of the NLI under the assumptions that (i) the nonlinear effects are weak, (ii) the signal is of the form suggested in [16], and (iii) the signal-noise interaction can be neglected. Using these three assumptions, we have obtained the NLI for a general system where all parameters in the modeling equation can have an arbitrary z-dependent value. This result allows the estimation of the impact of the Kerr nonlinearity in a very large set of practically relevant systems. Another benefit is that by using the model it is easy to separate the impact of the Kerr nonlinearity from other signal distortion due, e.g., to

amplifier noise and inter-symbol interference. While some questions about the model validity, in particular related to the signal model, still deserve more work, we believe that the discussed model [18] greatly contributes to other current efforts, such as [30], [32], to describe the impact of the Kerr nonlinearity in the highly dispersive regime.

Applying the general result to a system consisting of identical SMFs and EDFAs, we have recovered a result equivalent to that presented in [18, Eq. (18)]. We have also suggested a way to investigate the accuracy of the first assumption and found that the assumption is accurate within approximately 10% at an SNR of about 10 dB in the investigated example system. We have also investigated the impact of TOD and shown that it typically has a very small impact on the NLI PSD. The reason for this is that TOD changes the dispersion only by a small amount over the bandwidth of most WDM signals.

Appendix A
Derivation of the Formal Solution

Introducing $\psi(z,t) = A_1^x(z,t)/\sqrt{P(z)}$, we rewrite (6) to

$$\frac{\partial \psi}{\partial z} + i\frac{\beta_2}{2}\frac{\partial^2 \psi}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 \psi}{\partial t^3} = \frac{S}{\sqrt{P}}. \tag{32}$$

By Fourier transforming (32), and denoting this operation by tilde, we obtain an equation that can be written

$$\frac{\partial}{\partial z}\left(\tilde{\psi} e^{-i(2\pi f)^2(B_2(z)/2) - i(2\pi f)^3(B_3(z)/6)}\right) = e^{-i(2\pi f)^2(B_2(z)/2) - i(2\pi f)^3(B_3(z)/6)}\frac{\tilde{S}}{\sqrt{P}}. \tag{33}$$

Integrating $\int_0^z \cdot d\zeta$ and using that $\tilde{\psi}(0,f) = 0$ (since the perturbation is initially zero), we obtain

$$\tilde{\psi}(z,f) = e^{i(2\pi f)^2(B_2(z)/2) + i(2\pi f)^3(B_3(z)/6)} \times \int_0^z \frac{\tilde{S}(\zeta,f)}{\sqrt{P(\zeta)}} e^{-i(2\pi f)^2(B_2(\zeta)/2) - i(2\pi f)^3(B_3(\zeta)/6)}\, d\zeta. \tag{34}$$

The expression at the receiver is $\tilde{\psi}(L,f)$, where L is the total system length. Assuming that perfect dispersion compensation is carried out in the receiver, the exponential before the integral is canceled and we obtain

$$\tilde{\psi}(L,f) = \int_0^L \frac{\tilde{S}(z,f)}{\sqrt{P(z)}} e^{-i(2\pi f)^2(B_2(z)/2) - i(2\pi f)^3(B_3(z)/6)}\, dz. \tag{35}$$

Assuming that the amplification perfectly compensates for the losses, the power at $z = L$ is equal to the power at the transmitter and we have $A_1^x(L,t) = \sqrt{P(L)}\,\psi(L,t) = \sqrt{P(0)}\,\psi(L,t)$. We then directly obtain (9).

Appendix B
Solution Corresponding to the Signal Model

Accounting for the dispersion and the power variation during the propagation, (10) and (11) give

$$\tilde{A}_0^x(z,f) = \sqrt{f_0 p}\sum_k \xi_k \sqrt{G_0^x(f_k)} \times \delta(f - f_k) e^{i\omega_k^2(B_2/2) + i\omega_k^3(B_3/6)}, \tag{36}$$

$$\tilde{A}_0^y(z,f) = \sqrt{f_0 p}\sum_k \zeta_k \sqrt{G_0^y(f_k)} \times \delta(f - f_k) e^{i\omega_k^2(B_2/2) + i\omega_k^3(B_3/6)}. \tag{37}$$

To calculate the perturbation, we need the expression

$$\begin{aligned}\tilde{S}(z,f) &= \mathcal{F}[i\gamma(|A_0^x|^2 + |A_0^y|^2)A_0^x] \\ &= i\gamma\mathcal{F}[(A_0^x)^2(A_0^x)^*] + i\gamma\mathcal{F}[A_0^x A_0^y (A_0^y)^*],\end{aligned} \tag{38}$$

to write the solution (9) as

$$\begin{aligned}&\tilde{A}_1^x(L,f) \\ &= i\int_0^L \gamma e^{-i(2\pi f)^2(B_2/2) - i(2\pi f)^3(B_3/6)} \frac{\mathcal{F}[(A_0^x)^2(A_0^x)^*]}{\sqrt{p}}\, dz, \\ &\quad + i\int_0^L \gamma e^{-i(2\pi f)^2(B_2/2) - i(2\pi f)^3(B_3/6)} \frac{\mathcal{F}[A_0^x A_0^y (A_0^y)^*]}{\sqrt{p}}\, dz.\end{aligned} \tag{39}$$

The inverse Fourier transformation of (36) and (37) is trivial since the only frequency dependence occurs as δ-functions. The expressions $(A_0^x)^2(A_0^x)^*$ and $A_0^x A_0^y (A_0^y)^*$ are the products of three infinite sums and these expressions can now be written as a single triple sum and Fourier transformed to obtain

$$\begin{aligned}&\mathcal{F}[(A_0^x)^2(A_0^x)^*] \\ &= (f_0 p)^{3/2}\sum_{k,l,m} \xi_k \xi_l \xi_m^* \sqrt{G_0^x(f_k)G_0^x(f_l)G_0^x(f_m)} \\ &\quad \times \delta(f - (f_k + f_l - f_m)) \\ &\quad \times e^{i(\omega_k^2 + \omega_l^2 - \omega_m^2)(B_2/2) + i(\omega_k^3 + \omega_l^3 - \omega_m^3)(B_3/6)},\end{aligned} \tag{40}$$

$$\begin{aligned}&\mathcal{F}[A_0^x A_0^y (A_0^y)^*] \\ &= (f_0 p)^{3/2}\sum_{k,l,m} \xi_k \zeta_l \zeta_m^* \sqrt{G_0^x(f_k)G_0^y(f_l)G_0^y(f_m)} \\ &\quad \times \delta(f - (f_k + f_l - f_m)) \\ &\quad \times e^{i(\omega_k^2 + \omega_l^2 - \omega_m^2)(B_2/2) + i(\omega_k^3 + \omega_l^3 - \omega_m^3)(B_3/6)},\end{aligned} \tag{41}$$

where we used $f_k \equiv kf_0$, $\omega_k \equiv 2\pi f_k$, and analogous expressions for l and m. These expressions can now be inserted into the integrands of (39). To simplify the result, we move all exponential functions after the summation signs and use that

$$(\omega_k^2 + \omega_l^2 - \omega_m^2) - (\omega_k + \omega_l - \omega_m)^2 = -2(\omega_k - \omega_m)(\omega_l - \omega_m) \tag{42}$$

and

$$(\omega_k^3+\omega_l^3-\omega_m^3)-(\omega_k+\omega_l-\omega_m)^3 = -3(\omega_k+\omega_l)(\omega_k-\omega_m)(\omega_l-\omega_m) \quad (43)$$

to get

$$\begin{aligned}&\gamma e^{-i(2\pi f)^2(B_2/2)-i(2\pi f)^3(B_3/6)}\frac{\mathcal{F}[(A_0^x)^2(A_0^x)^*]}{\sqrt{p}}\\ &=f_0^{3/2}\gamma p\sum_{k,l,m}\xi_k\xi_l\xi_m^*\sqrt{G_0^x(f_k)G_0^x(f_l)G_0^x(f_m)}\\ &\quad\times\delta(f-(f_k+f_l-f_m))e^{-i(\omega_k-\omega_m)(\omega_l-\omega_m)B_2}\\ &\quad\times e^{-i(\omega_k+\omega_l)(\omega_k-\omega_m)(\omega_l-\omega_m)(B_3/2)}, \quad (44)\\ &\gamma e^{-i(2\pi f)^2(B_2/2)-i(2\pi f)^3(B_3/6)}\frac{\mathcal{F}[A_0^xA_0^y(A_0^y)^*]}{\sqrt{p}}\\ &=f_0^{3/2}\gamma p\sum_{k,l,m}\xi_k\zeta_l\zeta_m^*\sqrt{G_0^x(f_k)G_0^y(f_l)G_0^y(f_m)}\\ &\quad\times\delta(f-(f_k+f_l-f_m))e^{-i(\omega_k-\omega_m)(\omega_l-\omega_m)B_2}\\ &\quad\times e^{-i(\omega_k+\omega_l)(\omega_k-\omega_m)(\omega_l-\omega_m)(B_3/2)}. \quad (45)\end{aligned}$$

Inserting these expressions into (39) and performing the inverse Fourier transformation, we obtain

$$\begin{aligned}A_1^x(L,t)=&\,if_0^{3/2}\int_0^L\gamma p\sum_{k,l,m}\xi_k\xi_l\xi_m^*\sqrt{G_0^x(f_k)G_0^x(f_l)G_0^x(f_m)}\\ &\times e^{i(\omega_k+\omega_l-\omega_m)t}e^{-i(\omega_k-\omega_m)(\omega_l-\omega_m)B_2}\\ &\times e^{-i(\omega_k+\omega_l)(\omega_k-\omega_m)(\omega_l-\omega_m)(B_3/2)}\,dz\\ &+if_0^{3/2}\int_0^L\gamma p\sum_{k,l,m}\xi_k\zeta_l\zeta_m^*\sqrt{G_0^x(f_k)G_0^y(f_l)G_0^y(f_m)}\\ &\times e^{i(\omega_k+\omega_l-\omega_m)t}e^{-i(\omega_k-\omega_m)(\omega_l-\omega_m)B_2}\\ &\times e^{-i(\omega_k+\omega_l)(\omega_k-\omega_m)(\omega_l-\omega_m)(B_3/2)}\,dz. \quad (46)\end{aligned}$$

Appendix C
Derivation of the Perturbation PSD

The First Term in the PSD: Having discussed how to handle the expectation value of the stochastic variables, we can now go on to study the first expression, i.e., $G_{1,13}^x(f)=\mathcal{F}[R_{13}(\tau)]$. Following the rules in Section IV-A for how the indices can be chosen in (19), we have six possible combinations

$$\begin{aligned}&k=k' \quad l=l' \quad m=m',\\ &k=k' \quad l=m \quad m'=l',\\ &k=l' \quad l=k' \quad m=m',\\ &k=l' \quad l=m \quad m'=k',\\ &k=m \quad l=k' \quad m'=l',\\ &k=m \quad l=l' \quad m'=k'. \quad (47)\end{aligned}$$

We need to study these different possibilities individually but it should be noticed that there are only two qualitatively different types of terms. These can be called "FWM-like" and "XPM-like" and the former results when $m=m'$ and the latter results otherwise. (There are no "SPM-like" terms, as these cease to be meaningful as $f_0\to 0$.) This is also further commented on below.

The Case $k=k'$, $l=l'$, $m=m'$: We then have

$$R_{13}=f_0^3\sum_{k,l,m}|\mathcal{C}_{klm}|^2e^{i(\omega_k+\omega_l-\omega_m)\tau}G_0^x(f_k)G_0^x(f_l)G_0^x(f_m), \quad (48)$$

where we used that $\mathbb{E}\{|\xi_k|^2|\xi_l|^2|\xi_m|^2\}=1$. The corresponding PSD is

$$\begin{aligned}G_{1,13}^x(f)=&\,f_0^3\sum_{k,l,m}|\mathcal{C}_{klm}|^2G_0^x(f_k)G_0^x(f_l)G_0^x(f_m)\\ &\times\delta(f-(f_k+f_l-f_m)). \quad (49)\end{aligned}$$

Writing this as a three-dimensional Riemann sum and letting $f_0\to 0$, we get

$$\begin{aligned}G_{1,13}^x(f)=&\int\!\!\int\!\!\int|\mathcal{C}(f_1,f_2,f_3)|^2G_0^x(f_1)G_0^x(f_2)G_0^x(f_3)\\ &\times\delta(f-f_1-f_2+f_3)\,df_1df_2df_3\\ =&\int\!\!\int|\mathcal{C}(f_1,f_2,f_1+f_2-f)|^2\\ &\times G_0^x(f_1)G_0^x(f_2)G_0^x(f_1+f_2-f)\,df_1df_2. \quad (50)\end{aligned}$$

Although this is only part of the final result, this illustrates what we mean with "FWM-like". It is seen that the frequency components at f_1, f_2, and f_1+f_2-f interact to contribute to the NLI at f, and the efficiency of the process is determined by $|\mathcal{C}|^2$.

The Other Cases: Compared to the case above, the case $k=l'$, $l=k'$, $m=m'$ is obtained by swapping k' and l'. Since $\mathcal{C}_{klm}=\mathcal{C}_{lkm}$ and the rest of the expression is clearly invariant under this change, this yields an identical result as the case above. The other four cases are all, in a similar sense, equivalent to the case $k=k'$, $l=m$, $m'=l'$, for which we get

$$R_{13}=f_0^3\sum_{kll'}\mathcal{C}_{kll}\mathcal{C}_{kll'}^*e^{i\omega_k\tau}G_0^x(f_k)G_0^x(f_l)G_0^x(f_{l'}), \quad (51)$$

$$G_{1,13}^x(f)=f_0^3\sum_{kll'}\mathcal{C}_{kll}\mathcal{C}_{kll'}^*G_0^x(f_k)G_0^x(f_l)G_0^x(f_{l'})\delta(f-f_k). \quad (52)$$

Letting $f_0\to 0$, we get

$$\begin{aligned}G_{1,13}^x(f)=&\int\!\!\int\!\!\int\mathcal{C}(f_1,f_2,f_2)\mathcal{C}^*(f_1,f_3,f_3)\\ &\times G_0^x(f_1)G_0^x(f_2)G_0^x(f_3)\delta(f-f_1)\,df_1df_2df_3\\ =&\int\!\!\int\mathcal{C}(f,f_2,f_2)\mathcal{C}^*(f,f_3,f_3)\\ &\times G_0^x(f)G_0^x(f_2)G_0^x(f_3)\,df_2df_3. \quad (53)\end{aligned}$$

However

$$\begin{aligned}\mathcal{C}(f_1,f_2,f_2)=&\,\mathcal{C}(f_1,f_3,f_3)=\mathcal{C}(f_2,f_1,f_2)=\mathcal{C}(f_3,f_1,f_3)\\ =&\int_0^L\gamma(z)p(z)\,dz\equiv\mathcal{C}_0, \quad (54)\end{aligned}$$

which is a real number and this gives

$$\begin{aligned} G^x_{1,13}(f) &= \mathcal{C}_0^2 G_0^x(f) \int\int G_0^x(f_2) G_0^x(f_3)\, df_2 df_3 \\ &= \mathcal{C}_0^2 G_0^x(f) \int G_0^x(f_2)\, df_2 \int G_0^x(f_3)\, df_3 \\ &= \mathcal{C}_0^2 P_x^2 G_0^x(f). \end{aligned} \tag{55}$$

This is an "XPM-like" term and to understand the origin of this term, we study (13). As above, we set $l = m$ and since we are considering R_{13}, we set $G_0^y = 0$. The resulting expression can be written

$$\begin{aligned} &A_1^x(L,t) \\ &= i f_0^{3/2} \sum_{k,l} \mathcal{C}_0 e^{i\omega_k t} \xi_k \sqrt{G_0^x(f_k)} |\xi_l|^2 G_0^x(f_l) \\ &= \left[\sqrt{f_0} \sum_k \xi_k \sqrt{G_0^x(f_k)} e^{i\omega_k t} \right] \\ &\quad \times \left[i f_0 \mathcal{C}_0 \sum_l |\xi_l|^2 G_0^x(f_l) \right] \\ &= A_0^x(0,t) i \mathcal{C}_0 P_x. \end{aligned} \tag{56}$$

We interpret this term as the perturbation approximation, cf. (2), of the exact expression for an XPM[2] phase shift (without degeneracy), which would be

$$A_1^x(L,t) = A_0^x(0,t) e^{i\mathcal{C}_0 P_x} - A_0^x(0,t). \tag{57}$$

What happens is that the input field is acquiring a phase shift during the propagation due to the total power. This phase shift is not possible to describe exactly within the lowest-order perturbation analysis and the phase shift instead gives rise to a term that has non-zero PSD. However, a constant phase shift of the entire input field only amounts to a rotation of the received constellation and this should be of no consequence for the quality of the received signal. Therefore, the XPM-like terms are discarded from the NLI expression.

The Total PSD for the First Term: The six possible index selection cases split into two groups of degeneracy two and four, respectively, and we get the complete expression

$$\begin{aligned} G^x_{1,13}(f) &= 2 \int\int |\mathcal{C}(f_1, f_2, f_1 + f_2 - f)|^2 \\ &\times G_0^x(f_1) G_0^x(f_2) G_0^x(f_1 + f_2 - f)\, df_1 df_2 + 4\mathcal{C}_0^2 P_x^2 G_0^x(f). \end{aligned} \tag{58}$$

The Second Term in the PSD: We now study the second expression, i.e., $G^x_{1,14}(f) = \mathcal{F}[R_{14}(\tau)]$. Following the above rules for how the indices can be chosen, we have two combinations

$$\begin{aligned} &k = k' \quad l = m \quad m' = l' \\ &k = m \quad l = k' \quad m' = l'. \end{aligned} \tag{59}$$

[2]We emphasize that "XPM" here refers to phase shifts between frequency components, cf. (10)–(11), not between WDM channels. The input signal is treated as a single broadband signal without any division into channels and the individual channels are never seen in the analysis. A further consequence of this is that SPM is not part of the final result since the SPM of an individual frequency component becomes zero as f_0 approaches zero.

The reason that we have fewer possibilities is that the expression contains both ξ and ζ, which are independent.

The case $k = k'$, $l = m$, $m' = l'$ gives rise to a XPM-like term by a calculation in analogy with that described above. The result is

$$G^x_{1,14}(f) = \mathcal{C}_0^2 P_x P_y G_0^x(f). \tag{60}$$

The difference as compared with (55) is that the expression now involves both P_x and P_y. The case $k = m$, $l = k'$, $m' = l'$ is obtained from the above case by swapping k and l and will therefore give the same result. This gives us

$$G^x_{1,14}(f) = 2\mathcal{C}_0^2 P_x P_y G_0^x(f). \tag{61}$$

The Third Term in the PSD: The third expression, i.e., $G^x_{1,23}(f) = \mathcal{F}[R_{23}(\tau)]$ corresponds to the two combinations

$$\begin{aligned} &k = k' \quad l = m \quad m' = l' \\ &k = l' \quad l = m \quad m' = k'. \end{aligned} \tag{62}$$

The first case is identical to the first case for the second term in the PSD. The second case is obtained from the first case by swapping k' and l' and will therefore give the same result. We get

$$G^x_{1,23}(f) = 2\mathcal{C}_0^2 P_x P_y G_0^x(f). \tag{63}$$

The Fourth Term in the PSD: For the fourth expression, i.e., $G^x_{1,24}(f) = \mathcal{F}[R_{24}(\tau)]$, there are two combinations

$$\begin{aligned} &k = k' \quad l = l' \quad m = m' \\ &k = k' \quad l = m \quad m' = l'. \end{aligned} \tag{64}$$

The case $k = k'$, $l = l'$, $m = m'$ is a FWM-like term and in analogy with above, we get

$$R_{24} = f_0^3 \sum_{k,l,m} |\mathcal{C}_{klm}|^2 e^{i(\omega_k + \omega_l - \omega_m)\tau} G_0^x(f_k) G_0^y(f_l) G_0^y(f_m), \tag{65}$$

where we used that $\mathbb{E}\{|\xi_k|^2 |\zeta_l|^2 |\zeta_m|^2\} = 1$. Fourier transforming this expression and letting $f_0 \to 0$, we get

$$\begin{aligned} G^x_{1,24}(f) = \int\int &|\mathcal{C}(f_1, f_2, f_1 + f_2 - f)|^2 \\ &\times G_0^x(f_1) G_0^y(f_2) G_0^y(f_1 + f_2 - f)\, df_1 df_2. \end{aligned} \tag{66}$$

The case $k = k', l = m, m' = l'$ is an XPM-like term and we get

$$G^x_{1,24}(f) = \mathcal{C}_0^2 P_y^2 G_0^x(f). \tag{67}$$

The total PSD for the fourth term is the sum of (66) and (67).

ACKNOWLEDGMENT

The authors wish to thank A. Bononi and P. Serena for helpful discussions, G. Bosco and P. Poggiolini for valuable feedback, and coworkers at the Fibre Optic Communications Research Centre at Chalmers.

REFERENCES

[1] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. Tapia Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," in *Proc. ECOC*, 2010, Paper P4.07.

[2] F. Vacondio, C. Simonneau, L. Lorcy, J. C. Antona, A. Bononi, and S. Bigo, "Experimental characterization of Gaussian-distributed nonlinear distortions," in *Proc. ECOC*, 2011, Paper We.7.B.1.
[3] Y. Fan, L. Dou, Z. Tao, L. Lei, S. Oda, T. Hoshida, and J. C. Rasmussen, "Modulation format dependent phase noise caused by intrachannel nonlinearity," in *Proc. ECOC*, 2012, Paper We.2.C.3.
[4] M. W. Maeda, W. B. Sessa, W. I. Way, A. Yi-Yan, L. Curtis, R. Spicer, and R. I. Laming, "The effect of four-wave mixing in fibers on optical frequency-division multiplexed systems," *J. Lightw. Technol.*, vol. 8, no. 9, pp. 1402–1408, Sep. 1990.
[5] K. Inoue, "Four-wave mixing in an optical fiber in the zero-dispersion wavelength region," *J. Lightw. Technol.*, vol. 10, no. 11, pp. 1553–1561, Nov. 1992.
[6] K. Inoue, "Phase-mismatching characteristic of four-wave mixing in fiber lines with multistage optical amplifiers," *Opt. Lett.*, vol. 17, no. 11, pp. 801–803, Jun. 1992.
[7] K. Inoue and H. Toba, "Fiber four-wave mixing in multi-amplifier systems with nonuniform chromatic dispersion," *J. Lightw. Technol.*, vol. 13, no. 1, pp. 88–93, Jan. 1995.
[8] S. Radic, G. Pendock, A. Srivastava, P. Wysocki, and A. Chraplyvy, "Four-wave mixing in optical links using quasi-distributed optical amplifiers," *J. Lightw. Technol.*, vol. 19, no. 5, pp. 636–645, May 2001.
[9] A. Splett, C. Kurzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. ECOC*, 1993, Paper MoC2.4.
[10] H. Louchet, A. Hodžić, and K. Petermann, "Analytical model for the performance evaluation of DWDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 15, no. 9, pp. 1219–1221, Sep. 2003.
[11] H. Louchet, A. Hodžić, K. Petermann, A. Robinson, and R. Epworth, "Analytical model for the design of multispan DWDM transmission systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 1, pp. 247–249, Jan. 2005.
[12] M. Eiselt, "Limits on WDM systems due to four-wave mixing: A statistical approach," *J. Lightw. Technol.*, vol. 17, no. 11, pp. 2261–2267, Nov. 1999.
[13] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. Cho, R. Noé, I. Shpantzer, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Exp.*, vol. 16, no. 20, pp. 15777–15810, Sep. 2008.
[14] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Exp.*, vol. 18, no. 18, pp. 19 039–19 054, Aug. 2010.
[15] W. Shieh and X. Chen, "Information spectral efficiency and launch power density limits due to fiber nonlinearity for coherent optical OFDM systems," *IEEE Photon. J.*, vol. 3, no. 2, pp. 158–173, Apr. 2011.
[16] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 2011.
[17] P. Poggiolini, G. Bosco, A. Carena, V. Curri, and F. Forghieri, "A simple and accurate model for non-linear propagation effects in uncompensated coherent transmission links," in *Proc. ICTON*, 2011, Paper We.B1.3.
[18] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of non-linear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.
[19] P. Bayvel, S. Kilmurray, and R. I. Killey, "Nonlinear transmission performance of digital nyquist WDM and optical OFDM," in *Proc. ICTON*, 2012, Paper Mo.C1.6.
[20] E. Torrengo, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, and F. Forghieri, "Experimental validation of an analytical model for nonlinear propagation in uncompensated optical links," in *Proc. ECOC*, 2011, Paper We.7.B.2.
[21] P. Johannisson, "Analytical Modeling of Nonlinear Propagation in a Strongly Dispersive Optical Communication System," 2012 [Online]. Available: arXiv:1205.2193v2 [physics.optics]
[22] A. Bononi and P. Serena, "An Alternative Derivation of Johannisson's Regular Perturbation Model," 2012 [Online]. Available: arXiv:1207.4729v1 [physics.optics]
[23] P. Poggiolini, G. Bosco, A. Carena, V. Curri, and F. Forghieri, "A Detailed Analytical Derivation of the GN Model of Non-Linear Interference in Coherent Optical Transmission Systems," 2012 [Online]. Available: arXiv:1209.0394v2 [physics.optics]
[24] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.
[25] M. J. Ablowitz and T. Hirooka, "Resonant nonlinear intrachannel interactions in strongly dispersion-managed transmission systems," *Opt. Lett.*, vol. 25, no. 24, pp. 1750–1752, Dec. 2000.
[26] P. Johannisson, D. Anderson, A. Berntson, and J. Mårtensson, "Generation and dynamics of ghost pulses in strongly dispersion-managed fiber-optic communication systems," *Opt. Lett.*, vol. 26, no. 16, pp. 1227–1229, Aug. 2001.
[27] A. Vannucci, P. Serena, and A. Bononi, "The RP method: A new tool for the iterative solution of the nonlinear Schrödinger equation," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1102–1112, Jul. 2002.
[28] A. Bononi, P. Serena, and M. Bertolini, "Unified analysis of weakly-nonlinear dispersion-managed optical transmission systems using a perturbative approach," *Comptes Rendus Physique*, vol. 9, pp. 947–962, Nov. 2008.
[29] A. Bononi, P. Serena, N. Rossi, E. Grellier, and F. Vacondio, "Modeling nonlinearity in coherent transmissions with dominant intrachannel-four-wave-mixing," *Opt. Exp.*, vol. 20, no. 7, pp. 7777–7791, Mar. 2012.
[30] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.
[31] W. Yan, Z. Tao, L. Dou, L. Li, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Low complexity digital perturbation back-propagation," in *Proc. ECOC*, 2011, Paper Tu.3.A.2.
[32] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.
[33] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise," *Opt. Exp.*, vol. 19, no. 13, pp. 12 781–12 788, Jun. 2011.
[34] D. Wang and C. R. Menyuk, "Polarization evolution due to the Kerr nonlinearity and chromatic dispersion," *J. Lightw. Technol.*, vol. 17, no. 12, pp. 2520–2529, Dec. 1999.
[35] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. New York, NY, USA: Academic, 2001.
[36] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2008.
[37] J. W. Goodman, *Statistical Optics*. New York, NY, USA: Wiley-Interscience, 1985.

Pontus Johannisson received the Ph.D. degree from Chalmers University of Technology, Göteborg, Sweden, in 2006.

His thesis was focused on nonlinear intrachannel signal impairments in optical fiber communications systems. In 2006, he joined the research institute IMEGO in Göteborg, Sweden, where he worked with digital signal processing for inertial navigation with MEMS-based accelerometers and gyroscopes. In 2009, he joined the Photonics Laboratory, Chalmers University of Technology, where he currently holds a position as Assistant Professor. His research interests include, e.g., nonlinear effects in optical fibers and digital signal processing in coherent optical receivers.

Magnus Karlsson received the Ph.D. degree in 1994 from Chalmers University of Technology, Gothenburg, Sweden.

Since 1995, he has been with the Photonics Laboratory at Chalmers, first as Assistant Professor and since 2003 as Professor in photonics. He has authored or co-authored around 200 scientific journal and conference contributions. His research has been devoted to mainly nonlinear fiber optics and optical communications in a broad sense, from fundamental transmission effects such as fiber nonlinearities and polarization mode dispersion, to applied issues such as high-capacity data transmission and all-optical switching. Currently he is devoted to parametric amplification, multilevel modulation formats and coherent transmission in optical fibers.

Dr. Karlsson is currently Associate Editor of *Optics Express*. He has served in the technical committee for the Optical Fiber Communication Conference (OFC), and currently serves in the technical program committees for the European Conference of Optical Communication (ECOC), and the Asia Communications and Photonics Conference (ACP). He became Fellow of the OSA in 2012.