CHALMERS
UNIVERSITY OF TECHNOLOGY

Analytical Modeling of Nonlinear Fiber Propagation for Four Dimensional Symmetric Constellations

Citation for the original published paper (version of record):
Rabbani, H., Ayaz, M., Beygi, L. et al (2021). Analytical Modeling of Nonlinear Fiber Propagation for Four Dimensional Symmetric Constellations. Journal of Lightwave Technology, 39(9): 2704-2713. http://dx.doi.org/10.1109/JLT.2021.3055966

Analytical Modeling of Nonlinear Fiber Propagation for Four Dimensional Symmetric Constellations

Hami Rabbani, Mostafa Ayaz, Lotfollah Beygi, Gabriele Liga, *Member*, *IEEE*, Alex Alvarado, *Senior Member*, *IEEE*, Erik Agrell, *Fellow*, *IEEE*, Magnus Karlsson, *Senior Member*, *IEEE*, *Fellow*, *OSA*

***Abstract*—Coherent optical transmission systems naturally lead to a four dimensional (4D) signal space, i.e., two polarizations each with two quadratures. In this paper we derive an analytical model to quantify the impact of Kerr nonlinearity on such 4D spaces, taking the interpolarization dependency into account. This is in contrast to previous models such as the GN and EGN models, which are valid for polarization multiplexed (PM) formats, where the two polarizations are seen as independent channels on which data is multiplexed. The proposed model agrees with the EGN model in the special case of independent two-dimensional modulation in each polarization. The model accounts for the predominant nonlinear terms in a WDM system, namely self-phase modulation and and cross-phase modulation. Numerical results show that the EGN model may inaccurately estimate the nonlinear interference of 4D formats. This nonlinear interference discrepancy between the results of the proposed model and the EGN model could be up to 2.8 dB for a system with 80 WDM channels. The derived model is validated by split-step Fourier simulations, and it is shown to follow simulations very closely.**

***Index Terms*—Coherent transmission, Enhanced Gaussian noise model, Four dimensional signals, Gaussian noise model, Kerr nonlinearity, Optical fiber communications.**

I. Introduction

THE amount of traffic carried on optical backbone networks continues to grow at a rapid pace, and makes efficient use of available resources indispensable. The Kerr nonlinearity is the overriding factor that leads to signal distortion and limits the capacity of optical fiber transmission systems [1]. Studying the ultimate limits of such systems is key to avoid the capacity crunch. To circumvent the capacity crunch problem, spectrally-efficient modulation formats have attracted substantial attention.

H. Rabbani, M. Ayaz and L. Beygi are with the the EE Dept. of K. N. Toosi University of Technology. E-mails: hami.rabbani@email.kntu.ac.ir, m.ayaz@email.kntu.ac.ir and beygi@kntu.ac.ir

G. Liga and A. Alvarado are with the Information and Communication Theory Lab, Signal Processing Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands. E-mails: {g.liga,a.alvarado}@tue.nl

E. Agrell is with the Dept. of Electrical Engineering, Chalmers University of Technology, Sweden. E-mail: agrell@chalmers.se

M. Karlsson is with the Dept. of Microtechnology and Nanoscience, Photonics Laboratory, Chalmers University of Technology, Sweden. E-mail: magnus.karlsson@chalmers.se

G. Liga is funded by the EuroTechPostdoc programme under the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement No. 754462). This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 757791).

Optimized 2D modulation formats have become increasingly popular in optical communications. However, further optimization is possible if the full 4D signal space (which is inherent in optical coherent detection) is exploited. The idea of 4D modulation formats was introduced to optical communications as far back in time as the coherent receiver was explored [2]–[5]. Agrell and Karlsson [6], [7] began optimizing modulation formats in a 4D space for coherent optical communication systems in 2009. A number of 4D modulation formats have recently been proposed for purposes of maximizing generalized mutual information, optimizing power efficiency, and other equally compelling motivations [8]–[11]. 4D coded modulation with bit-wise decoders was studied in [12]. Recently, other 4D coded modulation schemes have been proposed in [13], [14].

Many 4D modulation formats have been experimentally demonstrated. The optical hardware for 4D transmission is the same as for polarization-multiplexed (PM) 2D modulation, namely a separate Mach–Zehnder modulator for each polarization, which in the 4D case are controlled by the same electronics (see, e.g., [15, Fig. 3]).

Although quite a few approximate analytical models for nonlinear fibre propagation are currently available in the literature [16]–[21], all of these models aim to predict the nonlinear interference (NLI) in polarization multiplexed (PM) systems. What follows is a short description of analytical models proposed for such PM optical systems.

To analytically evaluate the quality of transmissions of fiber-optic links, many research works have been devoted to extracting channel models both in the time and frequency domains [17], [21], [22]. The Gaussian noise (GN) models in highly dispersive optical communications systems were presented in [18], [22]–[24]. The 4D GN-type channel model was first proposed in [25]. The finite-memory GN model was introduced in [26]. Due to the Gaussianity assumption of the signal, GN model is not able to predict the modulation format dependence property of NLI.

The authors of [17] for the first time addressed a modulation-format-dependent time-domain model, assuming only the dominant nonlinear terms of cross-channel interference (XCI), known as cross-phase modulation (XPM) terms. Later, this time-domain model was studied comprehensively in [19] and compared with the GN model to address the discrepancy between these two models. In much the same way as in [19], the authors of [21] derived a new perturbation model (in the frequency domain) dropping the assumption of Gaussianity of the transmitted signal. This model was labelled

enhanced Gaussian noise (EGN) model. As its name suggests, the EGN model added a number of correction terms to the GN model formulation, which fully captured the modulation format dependency of the NLI. Moreover, the frequency-domain approach in [21] allows the model to fully account for all the different contributions of the NLI in a WDM spectrum, including: the self-channel interference (SCI), and unlike [19], all XCI and multi-channel interference (MCI) terms. It was shown in [27] that the GN and time-domain model [17], [19] failed to accurately predict the NLI, whilst the EGN model was able to capture both the modulation format and the symbol rate dependency of the NLI. The achievable rate in nonlinear WDM systems was evaluated in [28].

Recently, [29] proposed a modulation-format-dependent model in the presence of stimulated Raman scattering. The authors of [29] added a modulation format correction term to the XPM, while the SCI was computed under a Gaussian assumption. A general nonlinear model in the presence of Kerr nonlinearity and stimulated Raman scattering was proposed in [30], which accounts for the modulation-format-dependent SCI, XCI, and MCI terms. A survey of channel models proposed in the literature up to 2015 was presented in [31].

All of the aforementioned works are valid for PM modulation formats in which polarizations act as two independent channels. In this paper, we concentrate on symmetric[1] constellations and derive an accurate analytical model that is able to predict the impact of NLI on 4D optical transmission systems where data is jointly transmitted on both polarizations. Unlike the previous models [19], [21], [32], the derived model is built on the fact that the x- and y-polarization are dependent of one another, making it possible to predict the performance of optimized 4D modulation formats in the presence of fiber nonlinearities. A comprehensive approach to deriving the SCI term in the frequency domain was presented in [33], thus enabling the computation of the NLI power of arbitrary zero-mean 4D constellations.

The paper computes the SCI and XPM nonlinear terms. Our model is derived following a time-domain approach, as in [17], [19], and does not include other XCI terms apart from XPM, nor does it contain MCI [21, Fig. 7]. Although the derivation of a comprehensive analytical model that can take into account all terms of NLI (SCI, XCI, and MCI) goes beyond the scope of this paper, the model in this paper computes the lion's share of the NLI in the high dispersion multi-channel WDM systems, i.e., the SCI and XPM terms [27, Fig. 2]. The proposed model may therefore be unable to accurately predict the NLI in the low dispersion systems where the MCI contributions have an important role to play.

The rest of this paper is organized as follows. In Sec. II, we describe the electrical field in a 4D space and also review the first order solution to Manakov equation. The main result of this work is presented in Sec. III. In Sec. IV, we validate the proposed model by split-step Fourier simulations, and compare a wide variety of 4D formats in terms of the experienced NLI. Sec. V concludes the paper. The detailed derivations of the main result of this paper are included in the Appendix.

[1]Constellations which are symmetric with respect to the origin, and have the same power in both polarizations.

II. PRELIMINARIES

The electric field of the optical wave intrinsically comprises two polarizations, each with two quadratures, thus in total four degrees of freedom, any one of which can be considered as a dimension. The electrical field can therefore be written as[2]

$$\boldsymbol{E} = \begin{bmatrix} E_{\mathrm{x}} \\ E_{\mathrm{y}} \end{bmatrix} = \begin{bmatrix} E_{\mathrm{x,r}} + iE_{\mathrm{x,i}} \\ E_{\mathrm{y,r}} + iE_{\mathrm{y,i}} \end{bmatrix}, \tag{1}$$

where indices x and y stand for polarization states, and r and i the real and imaginary parts, resp., of the electrical field.

The propagation of dual-polarized signals in a dispersive and nonlinear optical fiber is governed by the Manakov equation [32, Eq. (2)]

$$\frac{\partial}{\partial z}\boldsymbol{u}(t,z) = -\frac{i\beta_2}{2}\frac{\partial^2}{\partial t^2}\boldsymbol{u}(t,z) + i\frac{8}{9}\gamma f(z)\boldsymbol{u}^{\dagger}(t,z)\boldsymbol{u}(t,z)\boldsymbol{u}(t,z), \tag{2}$$

where $\boldsymbol{u}(t,z)$ is related to $\boldsymbol{E}(t,z)$ via a distance-dependent scaling function to compensate for gain/loss as in [17, Sec. II]. In (2), γ is the nonlinearity coefficient, β_2 is the group velocity dispersion, and $f(z)$ accounts for the link's loss/gain profile. In the case of perfectly distributed amplification $f(z) = 1$, while in the case of lumped amplification $f(z) = \exp\{-\alpha \mathrm{mod}(z, L)\}$ where α is the loss coefficient, L is the span length and $\mathrm{mod}(z, L)$ is the modulo operation and shows the distance between the point z and the nearest preceding amplifier.

We wish to evaluate the variance of SCI (intra-channel interference) and XPM (inter-channel interference) terms based on the first order perturbation approach, as these terms contribute to the NLI as predominant factors. We consider a channel of interest (COI) whose central frequency is set to zero, and an interfering channel with central frequency Ω. The XPM contributions of multiple WDM channels sum up incoherently, so there is no need to consider more than one channel pair [19, Sec. 2]. The linear solution of the Manakov equation at distance z for two channels is [19, Eq. (1)]

$$\boldsymbol{u}(z,t) = \sum_k \boldsymbol{a}_k g_a(t - kT_a, z) + \mathrm{e}^{-i\Omega t + \frac{i\beta_2\Omega^2}{2}z} \sum_k \boldsymbol{b}_k g_b(t - kT_b - \beta_2\Omega z, z), \tag{3}$$

where $\boldsymbol{a}_k = [a_{k,\mathrm{x}}\ a_{k,\mathrm{y}}]^{\mathrm{T}}$ and $\boldsymbol{b}_k = [b_{k,\mathrm{x}}\ b_{k,\mathrm{y}}]^{\mathrm{T}}$ are column vectors containing two elements, which represent the k-th symbol transmitted by the COI (a) and interfering channel (b), resp. The dispersed pulse is represented by $g_{a,b}(t,z) = \exp(-iz\beta_2\partial_t^2/2)g_{a,b}(t,0)$ [34], where $g_{a,b}(t,0)$ is the input pulse used by the COI and interfering channel, resp., and ∂_t^2 is the time derivative operator. The symbol rate of the COI and

[2]Throughout this paper we use $(\cdot)_{\mathrm{x}}$ and $(\cdot)_{\mathrm{y}}$ to represent variables associated to polarizations x and y, resp. Expectations are denoted by $\mathbb{E}\{\cdot\}$, and two dimensional complex functions are denoted using boldface (e.g., $\boldsymbol{E}$) symbols whose Hermitian conjugate is shown by $(\cdot)^{\dagger}$.

interfering channel is denoted by T_a^{-1} and T_b^{-1}, resp., which can differ from each other.

Without loss of generality, we concentrate on detecting the zeroth symbol in the COI, i.e., $\boldsymbol{a}_0$. The receiver for the COI is assumed to fully compensate for the linear link's impairments. The received symbol at the end of the link is therefore expressed as $\boldsymbol{a}_0 + \Delta\boldsymbol{a}_0$, where $\Delta\boldsymbol{a}_0$ is the NLI contribution. The first order solution to Manakov equation is obtained based on the perturbation approach [34, Eq. (3)], which gives

$$\begin{aligned}\Delta\boldsymbol{a}_0(\Omega) =& i\frac{8}{9}\gamma\sum_{h,k,l} S_{h,k,l}\boldsymbol{a}_k^\dagger\boldsymbol{a}_h\boldsymbol{a}_l \\ &+ i\frac{8}{9}\gamma\sum_{h,k,l} X_{h,k,l}\left(\boldsymbol{b}_k^\dagger\boldsymbol{b}_h\mathbb{I} + \boldsymbol{b}_h\boldsymbol{b}_k^\dagger\right)\boldsymbol{a}_l. \end{aligned} \quad (4)$$

In (4) $\mathbb{I}$ is the 2×2 identity matrix, and $S_{h,k,l}$ and $X_{h,k,l}$ are [34, Eqs. (4) and (5)]

$$\begin{aligned}S_{h,k,l} = &\int_0^L \mathrm{d}z \int_{-\infty}^{\infty} \mathrm{d}t f(z) g_R^*(t,z) g_a(t-lT,z) \\ &\cdot g_a^*(t-kT,z) g_a(t-hT,z), \end{aligned} \quad (5)$$

and

$$\begin{aligned}X_{h,k,l} = &\int_0^L \mathrm{d}z \int_{-\infty}^{\infty} \mathrm{d}t f(z) g_R^*(t,z) g_a(t-lT,z) \\ &\cdot g_b^*(t-kT-\beta_2\Omega z,z) g_b(t-hT-\beta_2\Omega z,z), \end{aligned} \quad (6)$$

resp., where $g_R(z,t)$ is the matched filter's impulse response [32, Eq. (4)]. The first and second terms on the right-hand side of (4) are responsible for estimating the SCI and XPM terms, resp. Using the fact that $g(t,z) = \int \mathrm{d}w \tilde{g}(w)\exp(-iwt + iw^2\beta_2 z/2)/(2\pi)$, where $\tilde{g}(w)$ is the Fourier transform of $g(t,0)$ (see [34, Appendix] and [34, Eqs. (11) and (12)]), (5) and (6) are expressed in the frequency domain as

$$S_{h,k,l} = \int \frac{\mathrm{d}^3 w}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3)\mathrm{e}^{i(w_1 h - w_2 k + w_3 l)T}, \quad (7)$$

and

$$X_{h,k,l} = \int \frac{\mathrm{d}^3 w}{(2\pi)^3}\rho_{\mathrm{xp}}(w_1,w_2,w_3)\mathrm{e}^{i(w_1 h - w_2 k + w_3 l)T}, \quad (8)$$

resp., where $\mathrm{d}^3 w$ stands for $\mathrm{d}w_1\mathrm{d}w_2\mathrm{d}w_3$, and

$$\begin{aligned}\rho_{\mathrm{s}}(w_1,w_2,w_3) = &\tilde{g}_R^*(w_1 - w_2 + w_3) \\ &\cdot \tilde{g}_a(w_1)\tilde{g}_a^*(w_2)\tilde{g}_a(w_3)\Upsilon(w_1,w_2,w_3), \end{aligned} \quad (9)$$

and

$$\begin{aligned}\rho_{\mathrm{xp}}(w_1,w_2,w_3) = &\tilde{g}_R^*(w_1 - w_2 + w_3) \\ &\cdot \tilde{g}_b(w_1-\Omega)\tilde{g}_b^*(w_2-\Omega)\tilde{g}_a(w_3)\Upsilon(w_1,w_2,w_3), \end{aligned} \quad (10)$$

where

$$\Upsilon(w_1,w_2,w_3) = \int_0^L \mathrm{d}z f(z)\mathrm{e}^{i\beta_2(w_2-w_3)(w_2-w_1)z} \quad (11)$$

is the link function. For a more general model that also comprises heterogeneous fiber spans and third-order dispersion, $\Upsilon(f_1,f_2,f)$ in [30, Table III] can, after substituting $f_1 = w_1/2\pi$, $f_2 = w_3/2\pi$, and $f = w_2/2\pi$, be used instead of (11) in (10). One may want to take all the NLI terms such as SCI, XCI and MCI into account. In this regard, (3) should be extended to a general equation, which accounts for N terms, where N is the number of WDM channels occupying the full C-band spectrum, and as a result, (4) will contain N^3 terms, which stem from $\boldsymbol{u}^\dagger(t,z)\boldsymbol{u}(t,z)\boldsymbol{u}(t,z)$ in (2). Nonlinear analysis of all the NLI terms however falls outside the scope of the paper and is left for future work.

III. The key result: NLI variance

This section is devoted to providing the key result of this work, which is the variance of (4). Not only is the key result able to predict the NLI of most 4D constellations used in practice, it is straightforward enough to be easily calculated with even the simplest of computers. The detailed derivation of the key result will be given in the Appendix. The key result is obtained under some simplifying assumptions, which are discussed below.

The first assumption is that the data symbols in the x- and y-polarization are correlated with each other. The second assumption is that the data symbols in different time slots are independent of one another. Here, we consider a multi-channel WDM system where channels across the spectrum can have different bandwidths, different spacings, different launch powers and different 4D modulation formats. The probability distribution in each WDM channel is assumed uniform over all constellation points. We further assume that the launch power in the x- and y-polarization are the same, meaning that

$$\frac{P_a}{2} = \mathbb{E}\{|a_{\mathrm{x}}|^2\} = \mathbb{E}\{|a_{\mathrm{y}}|^2\}, \quad \frac{P_b}{2} = \mathbb{E}\{|b_{\mathrm{x}}|^2\} = \mathbb{E}\{|b_{\mathrm{y}}|^2\}, \quad (12)$$

where P_a and P_b are the total launch power transmitted in the COI and interfering channel, resp. It is also assumed that

$$\mathbb{E}\{|a_{\mathrm{x}}|^4\} = \mathbb{E}\{|a_{\mathrm{y}}|^4\}, \quad \mathbb{E}\{|b_{\mathrm{x}}|^4\} = \mathbb{E}\{|b_{\mathrm{y}}|^4\}. \quad (13)$$

The last key assumption is that $\mathbb{E}\{a_{\mathrm{x}}\} = \mathbb{E}\{a_{\mathrm{y}}\} = \mathbb{E}\{a_{\mathrm{x}}^2\} = \mathbb{E}\{a_{\mathrm{x}}a_{\mathrm{y}}^*\} = \mathbb{E}\{|a_{\mathrm{x}}|^2 a_{\mathrm{x}}\} = \mathbb{E}\{|a_{\mathrm{y}}|^2 a_{\mathrm{x}}\} = 0$. This assumption holds for most zero-mean symmetric constellations with respect to the origin that have the same power in both polarizations.

The NLI variance on the n-th channel (COI) caused by (4) is given by

$$\sigma^2_{\mathrm{NLI},n} = \mathrm{Var}\left\{\sum_{\Omega}\Delta\boldsymbol{a}_0(\Omega)\right\}, \quad (14)$$

Since the data symbols in different WDM channels are uncorrelated, we can write (14) as

$$\sigma^2_{\mathrm{NLI},n} = \sigma^2_{\mathrm{SCI}} + \sum_{j=1,j\neq n}^{N}\sigma^2_{\mathrm{XPM}}(\Omega), \quad \Omega = |\nu_j - \nu_n|, \quad (15)$$

where ν_j is the central frequency of channel j. The SCI and XPM variances given in (15) are expressed as

$$\sigma^2_{\mathrm{SCI}} = \sigma^2_{\mathrm{SCI,x}} + \sigma^2_{\mathrm{SCI,y}}, \quad (16)$$

and

$$\sigma^2_{\text{XPM}}(\Omega) = \sigma^2_{\text{XPM,x}}(\Omega) + \sigma^2_{\text{XPM,y}}(\Omega), \tag{17}$$

resp., in which $\sigma^2_{\text{SCI,x}}$ and $\sigma^2_{\text{XPM,x}}$ are the SCI and XPM variances in the x-polarization, resp. The same is true for the y-polarized terms given in (16) and (17). The terms $\sigma^2_{\text{SCI,x}}$ and $\sigma^2_{\text{XPM,x}}(\Omega)$, given in (16) and (17), resp., are equal to

$$\sigma^2_{\text{SCI,x}} = \frac{8}{81}\gamma^2 P_a^3 \left(\Psi_1 S_1 + \Psi_2 X_1 + \Psi_3 X_2 + 3Z_1\right), \tag{18}$$

and

$$\sigma^2_{\text{XPM,x}}(\Omega) = \frac{8}{81}\gamma^2 P_a P_b^2(\Omega) \left(\Phi_1(\Omega) X(\Omega) + 6Z(\Omega)\right). \tag{19}$$

The terms S_1, X_1, X_2, Z_1, $X(\Omega)$, and $Z(\Omega)$ in Table I depend on the spectral properties of the signal, in contrast with Ψ_1, Ψ_2, Ψ_3 and Φ_1, given in Table II, which depend on the modulation format. The SCI and XPM variances in the y-polarization can be obtained from (18) and (19), resp., by swapping x and y in (18), (19) and Table III.

The computational complexity of the proposed model is the same as the EGN model [21]. Just as [21, Appendix C], all the integral terms in Table I can be computed using a triple integral. The terms Ψ_1, Ψ_2, Ψ_3 and Φ_1 in Table II can be readily computed for any 4D modulation format from a list of its coordinates.

In the special case of independent polarizations that the same format is used in both polarizations, Table III yields $\varphi_4 = \varphi_3 = \varphi_2$ and $\varphi_7 = \varphi_5 = 1$. These values used in Table II give $\Psi_1 = \varphi_1 - 9\varphi_2 + 12$, $\Psi_2 = 5\varphi_2 - 10$, $\Psi_3 = \varphi_2 - 2$, and $\Phi_1 = 5\varphi_6 - 10$. These values used in combination with the integral expressions in Table I can be shown to coincide with the EGN model.

IV. Numerical Results

This section investigates the NLI of 4D modulation formats from the database [35], where the first two coordinates map onto the x polarization and the last two onto the y polarization. Two coherent transmission scenarios were simulated. The first scenario uses a standard single-mode fiber (SMF), whose spectrum is populated with 32 Gbaud channels with spacing 50 GHz. The second scenario concentrates on non-zero dispersion shifted fiber (NZDSF), where 12.5 Gbaud channels with spacing 25 GHz occupy the spectrum. We used the same fiber type in every span and uniform span length of 100 km. The following parameters were used for SMF and NZDSF respectively: Dispersion coefficient $D = 16.5$ and 3.8 ps/nm/km, nonlinear coefficient $\gamma = 1.3$ and 1.5 1/W/km, and attenuation $\alpha = 0.2$ and 0.2 dB/km. The EDFA noise figure and optical center wavelength were 5 dB and 1550 nm, resp. To compute the experienced NLI of 4D formats, we employ the 4D model defined in (15)–(19) together with Tables I and II. To relate our work to previously works, we compare our model with the EGN model.

In this section, we compare 4D constellations in terms of

$$\eta_n = \frac{\sigma^2_{\text{NLI},n}}{P^3}, \tag{20}$$

Table I
Integral expressions for the terms used in (18) and (19). The functions $\rho_{\text{S}}(\cdot)$ and $\rho_{\text{XP}}(\cdot)$ are given in (9) and (10), resp.

Term	Integral Expression
S_1	$\frac{1}{T_a}\int \frac{\mathrm{d}^3 w}{(2\pi)^3} \frac{\mathrm{d}^2 w'}{(2\pi)^2} \rho_{\text{s}}(w_1, w_2, w_3) \rho_{\text{s}}^*(w_1', w_2', w_1 + w_3 + w_2' - w_2 - w_1')$
X_1	$\frac{1}{T_a^2} \int \frac{\mathrm{d}^3 w}{(2\pi)^3} \frac{\mathrm{d}w_2'}{2\pi} \rho_{\text{s}}(w_1, w_2, w_3) \rho_{\text{s}}^*(w_1, w_2', w_2' - w_2 + w_3)$
X_2	$\frac{1}{T_a^2} \int \frac{\mathrm{d}^3 w}{(2\pi)^3} \frac{\mathrm{d}w_1'}{2\pi} \rho_{\text{s}}(w_1, w_2, w_3) \rho_{\text{s}}^*(w_1', w_2, w_1 + w_3 - w_1')$
Z_1	$\frac{1}{T_a^3} \int \frac{\mathrm{d}^3 w}{(2\pi)^3} \lvert\rho_{\text{s}}(w_1, w_2, w_3)\rvert^2$
X	$\frac{1}{T_a T_b} \int \frac{\mathrm{d}^3 w}{(2\pi)^3} \frac{\mathrm{d}w_2'}{2\pi} \rho_{\text{xp}}(w_1, w_2, w_3) \rho_{\text{xp}}^*(w_1 - w_2 + w_2', w_2', w_3)$
Z	$\frac{1}{T_a T_b^2} \int \frac{\mathrm{d}^3 w}{(2\pi)^3} \lvert\rho_{\text{xp}}(w_1, w_2, w_3)\rvert^2$

Table II
The terms used in (18) and (19). The values of $\varphi_1, \cdots, \varphi_7$ are given in Table III.

Term	Expression
Ψ_1	$\varphi_1 - 12\varphi_2 + 24 + 2\varphi_3 + \varphi_4 - 12\varphi_5$
Ψ_2	$5\varphi_2 - 15 + 5\varphi_5$
Ψ_3	$\varphi_2 - 3 + \varphi_5$
Φ_1	$5\varphi_6 - 15 + 5\varphi_7$

Table III
Expressions for the terms $\varphi_1, \cdots, \varphi_7$ used in Table II.

Term	Expression	Term	Expression	Term	Expression
φ_1	$\frac{\mathbb{E}\{\lvert a_{\text{x}}\rvert^6\}}{\mathbb{E}^3\{\lvert a_{\text{x}}\rvert^2\}}$	φ_2	$\frac{\mathbb{E}\{\lvert a_{\text{x}}\rvert^4\}}{\mathbb{E}^2\{\lvert a_{\text{x}}\rvert^2\}}$	φ_3	$\frac{\mathbb{E}\{\lvert a_{\text{x}}\rvert^4 \lvert a_{\text{y}}\rvert^2\}}{\mathbb{E}^3\{\lvert a_{\text{x}}\rvert^2\}}$
φ_4	$\frac{\mathbb{E}\{\lvert a_{\text{y}}\rvert^4 \lvert a_{\text{x}}\rvert^2\}}{\mathbb{E}^3\{\lvert a_{\text{x}}\rvert^2\}}$	φ_5	$\frac{\mathbb{E}\{\lvert a_{\text{x}}\rvert^2 \lvert a_{\text{y}}\rvert^2\}}{\mathbb{E}^2\{\lvert a_{\text{x}}\rvert^2\}}$	φ_6	$\frac{\mathbb{E}\{\lvert b_{\text{x}}\rvert^4\}}{\mathbb{E}^2\{\lvert b_{\text{x}}\rvert^2\}}$
φ_7	$\frac{\mathbb{E}\{\lvert b_{\text{x}}\rvert^2 \lvert b_{\text{y}}\rvert^2\}}{\mathbb{E}^2\{\lvert b_{\text{x}}\rvert^2\}}$				

assuming $P_a = P_b = P$, where $\sigma^2_{\text{NLI},n}$ is defined in (15).[3] The SNR of the COI n is $\text{SNR}_n = P/(\sigma^2_{\text{ASE}} + \sigma^2_{\text{NLI},n})$, where σ^2_{ASE} is the variance of the amplified spontaneous emission noise (ASE). The modeled NLI, namely $\sigma^2_{\text{NLI},n}$, is additive and Gaussian, as with the GN and EGN models. We first validate the 4D model using the split-step Fourier method (SSFM), and then compare a wide range of 4D constellations.

A. SSFM Simulations

Numerically solving the Manakov equation (2) for the entire C band is a big challenge. Part of the problem is, of course, high memory requirements, in addition to the excessive use of very large fast Fourier transforms. For this reason, the SSFM study was restricted to a bandwidth of 0.5 THz. To validate η_n, given in (20), ASE-noise-free SSFM numerical simulations were performed. In the absence of other noise sources, η_n can

[3]Note that the results in Sec. III can also be used for a system with mixed (heterogeneous) channel powers (i.e., $P_a \neq P_b$).

Table IV
THE VALUE Φ_1 FOR 4D CONSTELLATIONS CHOSEN FROM [35] ALONG WITH THREE NEW CONSTELLATIONS PROPOSED IN [8], [11], [36].

Modulation	Φ_1	**Modulation**	Φ_1	**Modulation**	Φ_1	**Modulation**	Φ_1	**Modulation**	Φ_1
biortho4_8	-5	tetra4_9	-3.75	PM-QPSK	-5	SO-PM-QPSK	-3	dicyclic4_24	-5
24cell4_24	-5	l4_25	-4.58	b4_32	-4.38	w4_40	-4.05	w4_49	-3.65
b4_64	-4.14	4D-2A8PSK [8]	-5	4D-64PRS [11]	-5	w4_88	-3.87	4D-OS128 [36]	-3.02
SP-QAM4_128	-3.4	w4_145	-3.99	w4_152	-3.77	w4_169	-3.88	PM-16QAM	-3.4
w4_256	-3.8	w4_313	-3.75	w4_409	-3.77	w4_464	-3.74	cross4_512	-3.57
sphere4_512	-3.8	SP-cross4_512	-3.45	120cell4_600	-5	w4_601	-3.81	w4_656	-3.76
w4_800	-3.77	cross4_2048	-3.51	SP-QAM4_2048	-3.09	PM-64QAM	-3.09		

be estimated via the received SNR for each channel n via the relationship

$$\eta_n \approx \frac{1}{\mathrm{SNR}_n^{\mathrm{est}} P^2}. \tag{21}$$

The approximate equality in (21) is due to the fact that the SSFM-based $\mathrm{SNR}^{\mathrm{est}}$ estimates also contain higher order perturbation terms. The SNR for a constellation with M symbols was estimated via

$$\mathrm{SNR}_n^{\mathrm{est}} = \frac{\sum_{i=1}^{M} |\bar{y}_i|^2}{\sum_{i=1}^{M} \mathbb{E}\{|Y - \bar{y}_i|^2 | X = x_i\}}, \tag{22}$$

where X and Y are the random variables representing the transmitted and received symbols, resp., x_i is the i-th constellation point, and $\bar{y}_i = \mathbb{E}\{Y|X = x_i\}$. A total number of 30000 symbols were used, of which the first 1500 and the last 1500 symbols were removed from the transmitted and received sequences. All channels used a flat launch power of $P = 0$ dBm.

A WDM system with $N = 10$ channels and four modulation formats, namely PM-QPSK, subset optimized PM-QPSK (SO-PM-QPSK) [35], [37], PM-16QAM and a4_256 [15], [35] was simulated. Fig. 1 shows the simulation results for η_n in $\mathrm{dB}(W^{-2}) = 10\log_{10}(\eta_n \cdot 1W^2)$ using markers for a transmission distance of 500 km. Fig. 1 (a) indicates that the 4D model results for SO-PM-QPSK perfectly follows the simulations, whereas the EGN model fails to estimate the NLI of this format. Fig. 1 (b) also illustrates that the results obtained from the 4D model for a4_256 are in good agreement with simulations, while the EGN model results depart from simulations. The results, shown in Figs. 1 (a) and (b), imply that the EGN model is inaccurate for the study of arbitrary 4D constellations, and that the NLI can be underestimated (SO-PM-QPSK) or overestimated (a4_256). The proposed 4D model instead has the capacity to predict the NLI of 4D formats with a good level of accuracy. The discrepancy between simulations and the results obtained from the 4D model is on average about 0.2 dB. For PM-QPSK and PM-16QAM, both the EGN model and 4D model give the same results that match the simulation results. In the following section, we attempt to identify the reasons behind an increase or decrease in the NLI estimated from the EGN model and 4D model.

Figs. 1 (c) and (d) show the results in the low dispersion regime, where low symbol rate channels, i.e., 12.5 Gbaud,

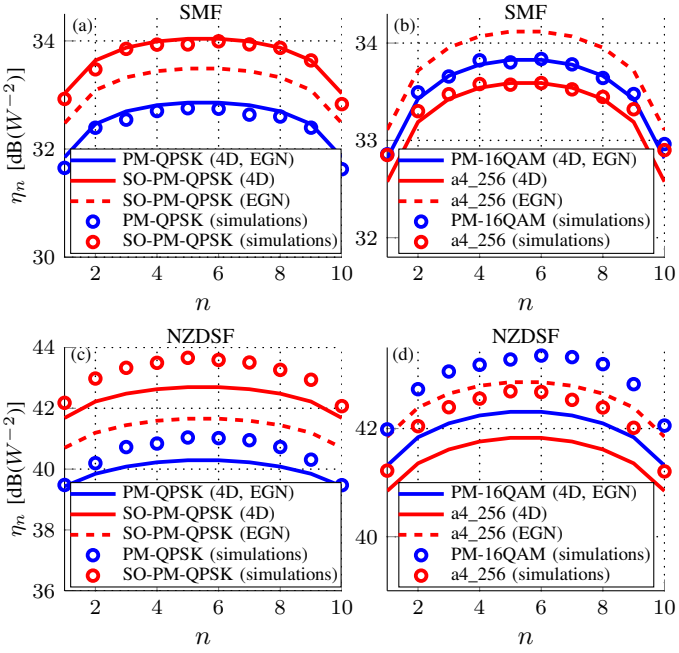

Figure 1. η_n as a function of channel number n after 5 spans. The link consisting of 5 spans. Two types of fiber (SMF and NZDSF) were simulated. The SMF supports $N = 10$ WDM channels with bandwidth 32 Gbaud and spacing 50 GHz, whereas NZDSF supports the same number of channels with bandwidth 12.5 Gbaud and spacing 25 GHz.

occupy the spectrum of an NZDSF. As shown in these figures, the proposed model underestimates the NLI by about 1 dB in comparison with the SSFM results. Neglecting the MCI terms explains this deviation. As the bandwidth and channel spacing decrease, this discrepancy will become greater. Comparing Figs. 1 (a) and (b) with (c) and (d) indicates that the model is accurate when the effect of MCI terms is small compared with SCI and XPM. This is often the case in the high dispersion regime [27]. On the other hand, in systems with strong MCI effects, such as NZDSF systems with low symbol rates, the model is less accurate, as shown in Figs. 1 (c) and (d). In such cases, full SSFM simulations may be preferred.

B. Comparing a wide range of constellations for an SMF

This section investigates the NLI of 4D constellations propagated over an SMF. We assume that the entire spectrum is populated with $N = 80$ WDM channels, and that the link comprises 10 spans. Figs. 2 (a) and (c) compare different

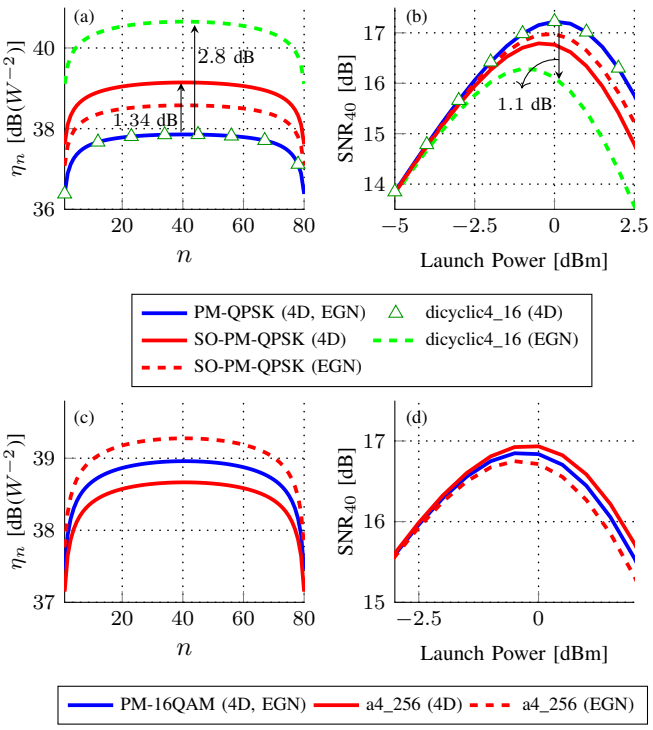

Figure 2. (a) and (c) illustrate η_n, defined in (20), as a function of channel number n after 10 spans, while (b) and (d) illustrate the SNR of COI, i.e., SNR$_{40}$, as a function of launch power after 10 spans of SMF. The full C-band spectrum can accommodate $N = 80$ WDM channels with bandwidth 32 Gbaud and spacing 50 GHz.

formats in terms of η_n, while Figs. 2 (b) and (d) compare them in terms of SNR$_{40}$. We interpret the first two coordinates in a coordinate list of [35] as the x polarization and the last two as the y polarization. Figs. 2 (a) and (b) give information about PM-QPSK, SO-PM-QPSK and dicyclic4_16 [35], [38]. We benchmark the 4D model against the EGN model in this figure. As can be seen in Fig. 2 (a), the curves are highest in the middle of spectrum. The 4D model indicates that over the entire spectrum shown in Fig. 2 (a), SO-PM-QPSK undergoes the most NLI, while PM-QPSK and dicyclic4_16 experience the least. It is also noticeable that PM-QPSK and dicyclic4_16 have the same NLI. The difference between the experienced NLI for SO-PM-QPSK and PM-QPSK is about 1.34 dB. This means that SO-PM-QPSK is more vulnerable to the Kerr nonlinearity than PM-QPSK. The SCI and XPM terms are responsible for this gap. The impact of SCI on the COI is high, but the XPM effects in multi-channel WDM systems are even higher, and therefore, the better part of this deviation stems form the XPM terms. To be more specific, the origin of this discrepancy comes form the fact that Φ_1, given in Table II, for SO-PM-QPSK ($\Phi_1 = -3$) is larger than for PM-QPSK ($\Phi_1 = -5$).

From the curve with triangles (4D model) to the green curve (EGN model), there is a 2.8 dB increase in the NLI for dicyclic4_16, with SNR falling by around 1.1 dB to approximately 16.1 dB (see Fig. 2 (b)). This implies that the EGN model significantly overestimates the NLI for dicyclic4_16. This is because $\varphi_7 = 0$ for this format according to Table III, whereas the EGN model corresponds to setting $\varphi_7 = 1$ for any format. On the other hand, we can see that the EGN model underestimates the NLI of SO-PM-QPSK in comparison with the 4D model. This is because the term φ_7 is lower for the EGN model ($\varphi_7 = 1$) than for the 4D model ($\varphi_7 = 1.2$).

Fig. 2 (c) and (d) compare the PM-16QAM and a4_256 formats in terms of η_n and SNR, resp. Fig. 2 (c) shows that PM-16QAM is at a disadvantage compared with a4_256. The deviation of the NLI between the PM-16QAM and a4_256 formats, as shown in Fig. 2 (c), is about 0.3 dB. This deviation may be rooted in the value of Φ_1 which is smaller for a4_256 ($\Phi_1 = -3.8$) than for PM-16QAM ($\Phi_1 = -3.4$). We can also see in Fig. 2 (c) that the EGN model overestimates the NLI of a4_256 by about 0.6 dB. It is clear from Fig. 2 (d) that the SNR for a4_256 falls from about 17 dB (4D model) to around 16.8 dB (EGN model) at 0 dBm launch power.

Given (15)–(19), as mentioned earlier, the experienced amount of NLI is dependent on Ψ_1, Ψ_2, Ψ_3, and Φ_1 given in Table II, none of which is more important than Φ_1, because the effect of XPM terms are higher than the SCI contribution for a fully loaded spectrum. For this reason, Table IV simply shows the value of Φ_1 for constellations selected from [35] as well as three constellations proposed in [8], [11], [36]. All the constellations shown in Table IV satisfy the assumptions made in Sec. III. It is clear that SO-PM-QPSK and four-dimensional orthant-symmetric 128-ary modulation (4D-OS128) [36] generate higher NLI than do other formats. For SO-PM-QPSK and 4D-OS128, Φ_1 is around -3, which is higher than the others. The lowest amount of NLI belongs to constellations whose Φ_1 equals to -5, meaning that these constellations undergo approximately the same NLI as PM-QPSK.

V. Conclusions

A nonlinear model which analytically models the impact of the Kerr nonlinearity on a 4D signal space was proposed and analyzed in detail. The model applies to zero-mean dual-polarization 4D formats which are symmetric with respect to the origin and have equal energy on the two polarization components. Unlike the GN and EGN models, we consider the interpolarization dependency so as to derive a 4D nonlinear model. The proposed model accounts for the SCI and XPM nonlinear terms. This is because the SCI and XPM are the predominant nonlinear terms in multi-channel WDM systems. We have compared different 4D modulation formats in terms of the experienced NLI, and showed that the derived model is a powerful tool to find 4D formats which are more resistant to the NLI.

Appendix

Our intention, in this section, is to study the variance of SCI (the first term on the right-hand side of (4)) and XPM (the second term on the right-hand side of (4)) terms. We proceed by computing the variance of SCI.

A. SCI variance

For the sake of brevity, we only focus on the x-polarized element of (4) because we can obtain the results for the y-

polarized component under the substitution x→y, y→x. Using (4), the x-polarized component of SCI term is given by

$$\Delta a_{0,\mathrm{SCI,x}} = i\frac{8}{9}\gamma \sum_{h,k,l} S_{h,k,l}\Big(a_{h,\mathrm{x}}a^*_{k,\mathrm{x}}a_{l,\mathrm{x}} + a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}\Big). \quad (23)$$

The variance of (23) is therefore equal to

$$\sigma^2_{\mathrm{SCI,x}} = \mathbb{E}\{\Delta a_{0,\mathrm{SCI,x}}\Delta a^*_{0,\mathrm{SCI,x}}\} - \mathbb{E}\{\Delta a_{0,\mathrm{SCI,x}}\}\mathbb{E}\{\Delta a^*_{0,\mathrm{SCI,x}}\}, \quad (24)$$

where $\mathbb{E}\{\Delta a_{0,\mathrm{SCI,x}}\} = 0$. This is because under the assumptions made in Sec. III, we have $\mathbb{E}\{a_{h,\mathrm{x}}\} = \mathbb{E}\{a^2_{h,\mathrm{x}}\} = \mathbb{E}\{a_{h,\mathrm{x}}a^*_{h,\mathrm{y}}\} = \mathbb{E}\{|a_{h,\mathrm{x}}|^2 a_{h,\mathrm{x}}\} = \mathbb{E}\{|a_{h,\mathrm{y}}|^2 a_{h,\mathrm{x}}\} = 0$ (see [32, Appendix A]), and as a result, $\mathbb{E}\{a_{h,\mathrm{x}}a^*_{k,\mathrm{x}}a_{l,\mathrm{x}}\} = \mathbb{E}\{a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}\} = 0$ for all h, k, and l. Substituting (23) into (24) gives

$$\sigma^2_{\mathrm{SCI,x}} = \frac{64}{81}\gamma^2 \sum_{h,k,l,h',k',l} S_{h,k,l}S^*_{h',k',l'}\Big(\mathbb{E}\{a_{h,\mathrm{x}}a^*_{k,\mathrm{x}}a_{l,\mathrm{x}}a^*_{h',\mathrm{x}}a_{k',\mathrm{x}}a^*_{l',\mathrm{x}}\} + \mathbb{E}\{a_{h,\mathrm{x}}a^*_{k,\mathrm{x}}a_{l,\mathrm{x}}a^*_{h',\mathrm{y}}a_{k',\mathrm{y}}a^*_{l',\mathrm{x}}\} + \mathbb{E}\{a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}a^*_{h',\mathrm{x}}a_{k',\mathrm{x}}a^*_{l',\mathrm{x}}\} + \mathbb{E}\{a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}a^*_{h',\mathrm{y}}a_{k',\mathrm{y}}a^*_{l',\mathrm{x}}\}\Big). \quad (25)$$

We can rewrite (25) as

$$\sigma^2_{\mathrm{SCI,x}} = \sum_{i=1}^{4} \sigma^2_{\mathrm{SCI,x},i}, \quad (26)$$

where $\sigma^2_{\mathrm{SCI,x},i}$ represents the i-th term in (25). We only give the procedure of calculating $\sigma^2_{\mathrm{SCI,x},4}$ in detail, and we can follow the same approach for the others. The contribution of $\sigma^2_{\mathrm{SCI,x},1}$ was calculated in the first term of Eq. (36) and Eq. (105) of [39], and $\sigma^2_{\mathrm{SCI,x},4}$ is more challenging to compute than the second and third terms, which is why we focus on calculating this term. The term $\sigma^2_{\mathrm{SCI,x},4}$ is given by

$$\sigma^2_{\mathrm{SCI,x},4} = \frac{64}{81}\gamma^2 \sum_{h,k,l,h',k',l} S_{h,k,l}S^*_{h',k',l'}\mathbb{E}\{a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}a^*_{h',\mathrm{y}}a_{k',\mathrm{y}}a^*_{l',\mathrm{x}}\}, \quad (27)$$

whose expectation term is equal to [32, Eqs. (26) and (27)]

$$\mathbb{E}\{a_{h,\mathrm{y}}a^*_{k,\mathrm{y}}a_{l,\mathrm{x}}a^*_{h',\mathrm{y}}a_{k',\mathrm{y}}a^*_{l',\mathrm{x}}\} = \begin{cases} \mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^4\}, & h=k=l=h'=k'=l', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^2\}, & h=h'\neq l=k=k'=l', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^2\}, & h=k\neq l=h'=k'=l', \\ \mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}\mathbb{E}\{|a_{\mathrm{y}}|^2a^*_{\mathrm{y}}a_{\mathrm{x}}\}, & h=l'\neq l=k=h'=k', \\ \mathbb{E}\{a^*_{\mathrm{y}}a_{\mathrm{x}}\}\mathbb{E}\{|a_{\mathrm{y}}|^2a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & k=l\neq h=h'=k'=l', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^2\}, & k=k'\neq h=l=h'=l', \\ \mathbb{E}\{a^*_{\mathrm{y}}a_{\mathrm{x}}\}\mathbb{E}\{|a_{\mathrm{y}}|^2a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & l=h'\neq h=k=k'=l', \\ \mathbb{E}\{|a_{\mathrm{x}}|^2\}\mathbb{E}\{|a_{\mathrm{y}}|^4\}, & l=l'\neq h=k=h'=k', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^2\}, & h'=k'\neq h=k=l=l', \\ \mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}\mathbb{E}\{|a_{\mathrm{y}}|^2a^*_{\mathrm{y}}a_{\mathrm{x}}\}, & k'=l'\neq h=k=l=h', \\ \mathbb{E}^2\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2\}, & h=h'\neq k=k'\neq l=l', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{a_{\mathrm{x}}a^*_{\mathrm{y}}\}\mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & h=k\neq l=h'\neq k'=l', \\ \mathbb{E}^2\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2\}, & h=k\neq l=l'\neq k'=h', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{a_{\mathrm{x}}a^*_{\mathrm{y}}\}\mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & h=h'\neq k=l\neq k'=l', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{a_{\mathrm{x}}a^*_{\mathrm{y}}\}\mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & h=l'\neq k=l\neq h'=k', \\ \mathbb{E}\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{a_{\mathrm{x}}a^*_{\mathrm{y}}\}\mathbb{E}\{a_{\mathrm{y}}a^*_{\mathrm{x}}\}, & h=l'\neq k=k'\neq l=h'. \end{cases} \quad (28)$$

We can hence write (27) as

$$\sigma^2_{\mathrm{SCI,x},4} = \sum_{j=1}^{16} \sigma^2_{\mathrm{SCI,x},4,j}, \quad (29)$$

where $\sigma^2_{\mathrm{SCI,x},4,j}$ stands for the contribution of the j-th case, given in (28), to (27). We first remove from (28) the terms which involve $\mathbb{E}\{a_{\mathrm{x}}a^*_{\mathrm{y}}\}$ or $\mathbb{E}\{a^*_{\mathrm{x}}a_{\mathrm{y}}\}$. By substituting (7) into (27), we can write $\sigma^2_{\mathrm{SCI,x},4,1}$, given in (29), as

$$\sigma^2_{\mathrm{SCI,x},4,1} = \frac{64}{81}\gamma^2\mathbb{E}\{|a_{\mathrm{x}}|^2|a_{\mathrm{y}}|^4\}\int \frac{\mathrm{d}^3w}{(2\pi)^3}\frac{\mathrm{d}^3w'}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3) \cdot \rho^*_{\mathrm{s}}(w'_1,w'_2,w'_3)\sum_h \mathrm{e}^{i(w_1-w_2+w_3-w'_1+w'_2-w'_3)hT_a}. \quad (30)$$

Assuming a rectangular or almost rectangular[4] channel spectrum $\tilde{g}(w)$ in (9), using the identity [34, Eq. (14)]

$$\sum_{k=-\infty}^{\infty} \mathrm{e}^{ikTw_1} = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty}\delta(w_1 - \frac{2\pi n}{T}), \quad (31)$$

and considering (12), we can write (30) as

$$\sigma^2_{\mathrm{SCI,x},4,1} = \frac{8}{81}\gamma^2 P^3_{\mathrm{coi}}\varphi_4 S_1, \quad (32)$$

where S_1 is given in Table I and φ_4 is given in Table III.

By using (7) once again in (27), and considering (12) the term $\sigma^2_{\mathrm{SCI,x},4,2}$, given in (29), is equal to

$$\sigma^2_{\mathrm{SCI,x},4,2} = \frac{8}{81}\gamma^2 P^3_{\mathrm{coi}}\varphi_5\int \frac{\mathrm{d}^3w}{(2\pi)^3}\frac{\mathrm{d}^3w'}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3) \cdot \rho^*_{\mathrm{s}}(w'_1,w'_2,w'_3)\sum_{h\neq l} \mathrm{e}^{i(w_1-w'_1)hT_a + i(w_3-w_2+w'_2-w'_3)lT_a}, \quad (33)$$

where φ_5 is given in Table III. Using the same approach given

[4]For large roll-off factors the excess terms must be taken into consideration [32, Appendix B, Eqs. (39)-(43)].

in [32, Eq. (29)], we have

$$\sum_{h\neq l}\mathrm{e}^{i(w_1-w_1')hT_a}\mathrm{e}^{i(w_3-w_2+w_2'-w_3')lT_a}=\sum_{h,l}\mathrm{e}^{i(w_1-w_1')hT_a}\cdot\mathrm{e}^{i(w_3-w_2+w_2'-w_3')lT_a}-\sum_{h}\mathrm{e}^{i(w_1-w_1'+w_3-w_2+w_2'-w_3')hT_a}. \tag{34}$$

Considering (31), we can rewrite (34) as

$$\sum_{h\neq l}\mathrm{e}^{i(w_1-w_1')hT_a}\mathrm{e}^{i(w_3-w_2+w_2'-w_3')lT_a}=\frac{4\pi^2}{T_a^2}\delta(w_3-w_2+w_2'-w_3')\cdot\delta(w_1-w_1')-\frac{2\pi}{T_a}\delta(w_1-w_1'+w_3-w_2+w_2'-w_3'). \tag{35}$$

By inserting (35) into (33), we get

$$\sigma^2_{\mathrm{SCI,x,4,2}}=\frac{8}{81}\gamma^2P_{\mathrm{coI}}^3\varphi_5\left(X_1-S_1\right), \tag{36}$$

where X_1 and S_1 are given in Table I. Considering (7), (12), (27), (29), (34) and (35), we can express $\sigma^2_{\mathrm{SCI,x,4,3}}$, given in (29), as

$$\sigma^2_{\mathrm{SCI,x,4,3}}=\frac{8}{81}\gamma^2P_{\mathrm{coI}}^3\varphi_5\int\frac{\mathrm{d}^3w}{(2\pi)^3}\frac{\mathrm{d}^3w'}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3)\cdot\rho_{\mathrm{s}}^*(w_1',w_2',w_3')\Big(\frac{4\pi^2}{T_a^2}\delta(w_3-w_1'+w_2'-w_3')\delta(w_1-w_2)-\frac{2\pi}{T_a}\delta(w_1-w_2+w_3-w_1'+w_2'-w_3')\Big). \tag{37}$$

The term $\delta(w_1-w_2)$ is a bias term and should be discarded. Bias terms are those for which $w_2=w_1$, $w_2=w_3$, $w_2'=w_1'$, or $w_2'=w_3'$. These terms create a constant phase shift, and thus, irrelevant for the noise variance we would like to compute (see [17, Sec. VIII, Eqs. (63)–(67)], [19, Sec. 3, Eq. (17)], [21, Appendix A], [40, Sec.IV-B and the text after (63)] and [18, Appendix C]). Eq. (37) is therefore reduced to

$$\sigma^2_{\mathrm{SCI,x,4,3}}=-\frac{8}{81}\gamma^2P_{\mathrm{coI}}^3\varphi_5S_1, \tag{38}$$

and we can express the same formula for $\sigma^2_{\mathrm{SCI,x,4,9}}$. Following the same approach, the term $\sigma^2_{\mathrm{SCI,x,4,6}}$, given in (29), contributes to (27) as

$$\sigma^2_{\mathrm{SCI,x,4,6}}=\frac{8}{81}\gamma^2P_{\mathrm{coI}}^3\varphi_5\left(X_2-S_1\right), \tag{39}$$

where X_2 is given in Table II.

The last step in calculating (29) is to investigate the impact of the last six situations on the NLI variance. We start with $\sigma^2_{\mathrm{SCI,x,4,11}}$, which contributes to (29) as

$$\sigma^2_{\mathrm{SCI,x,4,11}}=\frac{64}{81}\gamma^2\mathbb{E}^2\{|a_{\mathrm{y}}|^2\}\mathbb{E}\{|a_{\mathrm{x}}|^2\}\int\frac{\mathrm{d}^3w}{(2\pi)^3}\frac{\mathrm{d}^3w'}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3)\cdot\rho_{\mathrm{s}}^*(w_1',w_2',w_3')\sum_{h\neq k\neq l}\mathrm{e}^{i(w_1-w_1')hT_a}\mathrm{e}^{-i(w_2-w_2')kT}\mathrm{e}^{i(w_3-w_3')lT_a}, \tag{40}$$

where the triple summation is expressed as

$$\begin{aligned}\sum_{h\neq k\neq l}\mathrm{e}^{i(w_1-w_1')hT_a-i(w_2-w_2')kT+i(w_3-w_3')lT_a}=&\sum_{h,k,l}\mathrm{e}^{i(w_1-w_1')hT_a}\cdot\mathrm{e}^{-i(w_2-w_2')kT_a+i(w_3-w_3')lT_a}-\sum_{h=k\neq l}\mathrm{e}^{i(w_1-w_1'-w_2+w_2')hT_a+i(w_3-w_3')lT_a}\\&-\sum_{h=l\neq k}\mathrm{e}^{i(w_1-w_1'+w_3-w_3')hT_a}\mathrm{e}^{-i(w_2-w_2')kT_a}-\sum_{h\neq k=l}\mathrm{e}^{i(w_1-w_1')hT_a}\\&\cdot\mathrm{e}^{i(-w_2+w_2'+w_3-w_3')kT_a}+2\sum_{h}\mathrm{e}^{i(w_1-w_1'-w_2+w_2'+w_3-w_3')hT_a}.\end{aligned} \tag{41}$$

Considering (12), (31), (41), and (40), we have

$$\begin{aligned}\sigma^2_{\mathrm{SCI,x,4,11}}=&\frac{8}{81}\gamma^2P_a^3\int\frac{\mathrm{d}^3w}{(2\pi)^3}\frac{\mathrm{d}^3w'}{(2\pi)^3}\rho_{\mathrm{s}}(w_1,w_2,w_3)\\&\cdot\rho_{\mathrm{s}}^*(w_1',w_2',w_3')\Big(\frac{8\pi^3}{T_a^3}\delta(w_1-w_1')\delta(w_2'-w_2)\delta(w_3-w_3')\\&-\frac{4\pi^2}{T_a^2}\delta(w_1-w_1'+w_2'-w_2)\delta(w_3-w_3')-\frac{4\pi^2}{T_a^2}\delta(w_2'-w_2)\\&\cdot\delta(w_1-w_1'+w_3-w_3')-\frac{4\pi^2}{T_a^2}\delta(w_2'-w_2+w_3-w_3')\\&\cdot\delta(w_1-w_1')+\frac{8\pi}{T_a}\delta(w_1-w_1'+w_2'-w_2+w_3-w_3')\Big),\end{aligned} \tag{42}$$

which can be written as

$$\sigma^2_{\mathrm{SCI,x,4,11}}=\frac{8}{81}\gamma^2P_a^3\left(Z_1-2X_1-X_2+2S_1\right), \tag{43}$$

where Z_1, X_1, X_2 and S_1 are given in Table I. The same approach holds for $\sigma^2_{\mathrm{SCI,x,4,13}}$, but the bias terms should not be taken into account. Considering (32), (36), (38), (39) and (43), and using (12) and (13), we can express (27) as

$$\begin{aligned}\sigma^2_{\mathrm{SCI,x,4}}=&\frac{8}{81}\gamma^2P_a^3[(\varphi_4-4\varphi_5-\varphi_2+4)S_1+(\varphi_5+\varphi_2-3)X_1\\&+(\varphi_5-1)X_2+Z_1],\end{aligned} \tag{44}$$

which is called the interpolarization nonlinear effect, and the term φ_2 is given in Table III. This expression is not available in the literature.

The contributions of $\sigma^2_{\mathrm{SCI,x,1}}$, $\sigma^2_{\mathrm{SCI,x,2}}$, and $\sigma^2_{\mathrm{SCI,x,3}}$, given in (27), can be calculated through the same procedure, so their detailed derivations will not be repeated here, and we only give the final results for them as follows

$$\begin{aligned}\sigma^2_{\mathrm{SCI,x,2}}=&\sigma^2_{\mathrm{SCI,x,3}}=\frac{8}{81}\gamma^2P_a^3[(\varphi_3-4\varphi_5-\varphi_2+4)S_1\\&+(2\varphi_5-2)X_1],\end{aligned} \tag{45}$$

$$\begin{aligned}\sigma^2_{\mathrm{SCI,x,1}}=&\frac{8}{81}\gamma^2P_a^3[(\varphi_1-9\varphi_2+12)S_1+(4\varphi_2-8)X_1\\&+(\varphi_2-2)X_2+2Z_1],\end{aligned} \tag{46}$$

and we call (46) the intra-polarization nonlinear effect. By excluding the bias terms[5] from [39, Eq. (105)], and using it into the first term of [39, Eq. (36)], we can get (46). Putting

[5]Bias terms in [39, Eq. (105)] are those which involve δ_{m-n}, δ_{k-n}, $\delta_{m'-n'}$ and $\delta_{k'-n'}$.

(46), (45) and (44) together, we obtain the total variance of the SCI nonlinear term (25), which is expressed as (18) with coefficients from Tables I and III.

B. XPM variance

Here we calculate $\sigma^2_{\text{XPM,x}}(\Omega)$ in (19) for a single pair of channels with fixed separation Ω. For notational convenience, the dependence on Ω is dropped throughout the section. As mentioned in Sec. II, the second term of (4) gives rise to the XPM nonlinear term. The x-polarized component of this term is

$$i\frac{8}{9}\gamma\sum_{h,k,l}X_{h,k,l}\Big(2b_{h,\text{x}}b^*_{k,\text{x}}a_{l,\text{x}}+b_{h,\text{y}}b^*_{k,\text{y}}a_{l,\text{x}}+b_{h,\text{x}}b^*_{k,\text{y}}a_{l,\text{y}}\Big), \tag{47}$$

whose variance is equal to

$$\begin{aligned}\sigma^2_{\text{XPM,x}}=\frac{64}{81}\gamma^2\sum_{h,k,l,h',k',l'}X_{h,k,l}X^*_{h',k',l'}\Big(&4\mathbb{E}\{b_{h,\text{x}}b^*_{k,\text{x}}b^*_{h',\text{x}}b_{k',\text{x}}\}\\ &\cdot\mathbb{E}\{a_{l,\text{x}}a^*_{l',\text{x}}\}+2\mathbb{E}\{b_{h,\text{x}}b^*_{k,\text{x}}b^*_{h',\text{y}}b_{k',\text{y}}\}\mathbb{E}\{a_{l,\text{x}}a^*_{l',\text{x}}\}\\ &+2\mathbb{E}\{b_{h,\text{y}}b^*_{k,\text{y}}b^*_{h',\text{x}}b_{k',\text{x}}\}\mathbb{E}\{a_{l,\text{x}}a^*_{l',\text{x}}\}+\mathbb{E}\{b_{h,\text{y}}b^*_{k,\text{y}}b^*_{h',\text{y}}b_{k',\text{y}}\}\\ &\cdot\mathbb{E}\{a_{l,\text{x}}a^*_{l',\text{x}}\}+\mathbb{E}\{b_{h,\text{x}}b^*_{k,\text{y}}b^*_{h',\text{x}}b_{k',\text{y}}\}\mathbb{E}\{a_{l,\text{y}}a^*_{l',\text{y}}\}\Big).\end{aligned} \tag{48}$$

We now focus on the calculation of the first term of (48), and we can compute the others in a similar way. To evaluate the fourth order moment given in the first term of (48), the following cases should be taken into account.

$$\mathbb{E}\{b_{h,\text{x}}b^*_{k,\text{x}}b^*_{h',\text{x}}b_{k',\text{x}}\}=\begin{cases}\mathbb{E}\{|b_\text{x}|^4\}, & h=k=h'=k'\\ \mathbb{E}^2\{|b_\text{x}|^2\}, & h=k\neq h'=k'\\ \mathbb{E}^2\{|b_\text{x}|^2\}, & h=h'\neq k=k'.\end{cases} \tag{49}$$

Using (49) and (8), we can write the contribution of the first term of (48) to the NLI, as

$$\begin{aligned}\sigma^2_{\text{XPM,x,1st}}=&\frac{64}{81}\gamma^2\int\frac{\text{d}^3w}{(2\pi)^3}\frac{\text{d}^3w'}{(2\pi)^3}\rho_{\text{xp}}(w_1,w_2,w_3)\rho^*_{\text{xp}}(w'_1,w'_2,w'_3)\\ &\Big(4\mathbb{E}\{|b_\text{x}|^4\}\mathbb{E}\{|a_\text{x}|^2\}\sum_h\text{e}^{i(w_1-w_2-w'_1+w'_2)hT_b}\sum_l\text{e}^{i(w_3-w'_3)lT_a}\\ &+4\mathbb{E}^2\{|b_\text{x}|^2\}\mathbb{E}\{|a_\text{x}|^2\}\Big[\sum_{h\neq h'}\text{e}^{i(w_1-w_2)hT_b-(w'_1-w'_2)h'T_b}\sum_l\text{e}^{iw_3lT_a}\\ &\cdot\text{e}^{-w'_3lT_a}+\sum_{h\neq k}\text{e}^{i(w_1-w'_1)hT_b-(w_2-w'_2)kT_b}\sum_l\text{e}^{i(w_3-w'_3)lT_a}\Big]\Big).\end{aligned} \tag{50}$$

Considering (12), (31) and (35), we can express (50) as

$$\begin{aligned}\sigma^2_{\text{XPM,x,1st}}=&\frac{8}{81}\gamma^2P_aP_b^2\int\frac{\text{d}^3w}{(2\pi)^3}\frac{\text{d}^3w'}{(2\pi)^3}\rho_{\text{xp}}(w_1,w_2,w_3)\\ &\cdot\rho^*_{\text{xp}}(w'_1,w'_2,w'_3)\Big(\varphi_6\frac{16\pi^2}{T_aT_b}\delta(w_1-w_2-w'_1+w'_2)\\ &\cdot\delta(w_3-w'_3)+4\Big[(\frac{8\pi^3}{T_aT_b^2}\delta(w_1-w_2)\delta(w'_1-w'_2)-\frac{4\pi^2}{T_aT_b}\\ &\cdot\delta(w_1-w_2+w'_1-w'_2))\delta(w_3-w'_3)+(\frac{8\pi^3}{T_aT_b^2}\delta(w_1-w'_1)\\ &\cdot\delta(w_2-w'_2)-\frac{4\pi^2}{T_aT_b}\delta(w_1-w_2+w'_1-w'_2))\delta(w_3-w'_3)\Big]\Big).\end{aligned} \tag{51}$$

It should be noticed that the term $\delta(w_1-w_2)\delta(w'_1-w'_2)$ is a bias term and should be ignored. By excluding this term from (51), we have

$$\sigma^2_{\text{XPM,x,1st}}=\frac{8}{81}\gamma^2P_aP_b^2\left[(\varphi_6-2)\,4X+4Z\right], \tag{52}$$

where X and Z are given in Table I. Analogous expressions hold for other terms of (48). We can therefore express (48) as (19).

REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[2] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Exploiting fiber optics transmission capacity: 4-quadrature multilevel signalling," *Electron. Lett.*, vol. 26, no. 14, pp. 992–993, July 1990.

[3] ——, "A novel multilevel coherent optical system: 4-quadrature signaling," *J. Lightw. Technol.*, vol. 9, no. 4, pp. 514–523, Apr. 1991.

[4] S. Benedetto and P. Poggiolini, "Theory of polarization shift keying modulation," *IEEE Trans. Commun.*, vol. 40, no. 4, pp. 708–721, Apr. 1992.

[5] R. Cusani, E. Iannone, A. M. Salonico, and M. Todaro, "An efficient multilevel coherent optical system: M-4Q-QAM," *J. Lightw. Technol.*, vol. 10, no. 6, pp. 777–786, June 1992.

[6] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol. 27, no. 22, pp. 5115–5126, Nov. 2009.

[7] M. Karlsson and E. Agrell, "Which is the most power-efficient modulation format in optical links?" *Opt. Express*, vol. 17, no. 13, pp. 10 814–10 819, June 2009.

[8] K. Kojima, T. Yoshida, T. Koike-Akino, D. S. Millar, K. Parsons, M. Pajovic, and V. Arlunno, "Nonlinearity-tolerant four-dimensional 2A8PSK family for 5–7 bits/symbol spectral efficiency," *J. Lightw. Technol.*, vol. 35, no. 8, pp. 1383–1391, Feb. 2017.

[9] M. Reimer, S. O. Gharan, A. D. Shiner, and M. O'Sullivan, "Optimized 4 and 8 dimensional modulation formats for variable capacity in optical networks," in *Proc. Optical Fiber Communication Conf.*, Anaheim, CA, USA, Mar. 2016.

[10] T. Nakamura, E. L. T. de Gabory, H. Noguchi, W. Maeda, J. Abe, and K. Fukuchi, "Long haul transmission of four-dimensional 64SP-12QAM signal based on 16QAM constellation for longer distance at same spectral efficiency as PM-8QAM," in *Proc. European Conf. Optical Communication*, Valencia, Spain, Sep. 2015.

[11] B. Chen, C. Okonkwo, H. Hafermann, and A. Alvarado, "Polarization-ring-switching for nonlinearity-tolerant geometrically shaped four-dimensional formats maximizing generalized mutual information," *J. Lightw. Technol.*, vol. 37, no. 14, pp. 3579–3591, July 2019.

[12] A. Alvarado and E. Agrell, "Four-dimensional coded modulation with bit-wise decoders for future optical communications," *J. Lightw. Technol.*, vol. 33, no. 10, pp. 1993–2003, May 2015.

[13] J. Cai, M. V. Mazurczyk, H. G. Batshon, M. Paskov, C. R. Davidson, Y. Hu, O. V. Sinkin, M. A. Bolshtyansky, D. G. Foursa, and A. N. Pilipetskii, "Performance comparison of probabilistically shaped QAM formats and hybrid shaped APSK formats with coded modulation," *J. Lightw. Technol.*, vol. 38, no. 12, pp. 3280–3288, June 2020.

[14] F. Frey, S. Stern, J. K. Fischer, and R. F. H. Fischer, "Two-stage coded modulation for Hurwitz constellations in fiber-optical communications," *J. Lightw. Technol.*, vol. 38, no. 12, pp. 3135–3146, June 2020.

[15] T. A. Eriksson, S. Alreesh, C. Schmidt-Langhorst, F. Frey, P. W. Berenguer, C. Schubert, J. K. Fischer, P. A. Andrekson, M. Karlsson, and E. Agrell, "Experimental investigation of a four-dimensional 256-ary lattice-based modulation format," in *Proc. Optical Fiber Communication Conf.*, Los Angeles, CA, USA, Mar. 2015.

[16] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[17] A. Mecozzi and R. J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, June 2012.

[18] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 31, no. 8, pp. 1273–1282, Apr. 2013.

[19] 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. 25 685–25 699, Nov. 2013.
[20] 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. 3308–17, Feb. 2013.
[21] 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. 16 335–16 362, June 2014.
[22] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.
[23] 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.
[24] P. Serena and A. Bononi, "An alternative approach to the Gaussian noise model and its system implications," *J. Lightw. Technol.*, vol. 31, no. 22, pp. 3489–3499, Nov. 2013.
[25] 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.
[26] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.
[27] P. Poggiolini, A. Nespola, Y. Jiang, G. Bosco, A. Carena, L. Bertignono, S. M. Bilal, S. Abrate, and F. Forghieri, "Analytical and experimental results on system maximum reach increase through symbol rate optimization," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1872–1885, Apr. 2016.
[28] 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.
[29] D. Semrau, E. Sillekens, R. I. Killey, and P. Bayvel, "A modulation format correction formula for the Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5122–5131, Oct. 2019.
[30] H. Rabbani, G. Liga, V. Oliari, L. Beygi, E. Agrell, M. Karlsson, and A. Alvarado, "An improved model of nonlinear fiber propagation in the presence of Kerr nonlinearity and stimulated Raman scattering," *arXiv*, 2020. [Online]. Available: http://arxiv.org/abs/1909.08714v3.
[31] E. Agrell, G. Durisi, and P. Johannisson, "Information-theory-friendly models for fiber-optic channels: A primer," in *IEEE Information Theory Workshop (ITW)*, Jerusalem, Israel, Apr.-May 2015.
[32] O. Golani, R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Modeling the bit-error-rate performance of nonlinear fiber-optic systems," *J. Lightw. Technol.*, vol. 34, no. 15, pp. 3482–3489, Aug. 2016.
[33] G. Liga, A. Barreiro, H. Rabbani, and A. Alvarado, "Extending fibre nonlinear interference power modelling to account for general dual-polarisation 4D modulation formats," *Entropy*, vol. 22, no. 11, p. 1324, Nov. 2020.
[34] 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.
[35] E. Agrell, "Database of sphere packings," 2014–2020. [Online]. Available: http://codes.se/packings/.
[36] B. Chen, A. Alvarado, S. van der Heide, M. van den Hout, H. Hafermann, and C. Okonkwo, "Analysis and experimental demonstration of orthant-symmetric four-dimensional 7 bit/4D-sym modulation for optical fiber communication," *arXiv*, 2020. [Online]. Available: http://arxiv.org/abs/2003.12712.
[37] M. Sjödin, E. Agrell, and M. Karlsson, "Subset-optimized polarization-multiplexed PSK for fiber-optic communications," *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 838–840, May 2013.
[38] L. Zetterberg and H. Brändström, "Codes for combined phase and amplitude modulated signals in a four-dimensional space," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 943–950, Sep. 1977.
[39] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "On the accuracy of the GN-model and on analytical correction terms to improve it," *arXiv*, 2014. [Online]. Available: http://arxiv.org/abs/1401.6946.
[40] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A detailed analytical derivation of the GN model of non-linear interference in coherent optical transmission systems," *arXiv*, 2014. [Online]. Available: http://arxiv.org/abs/1209.0394.