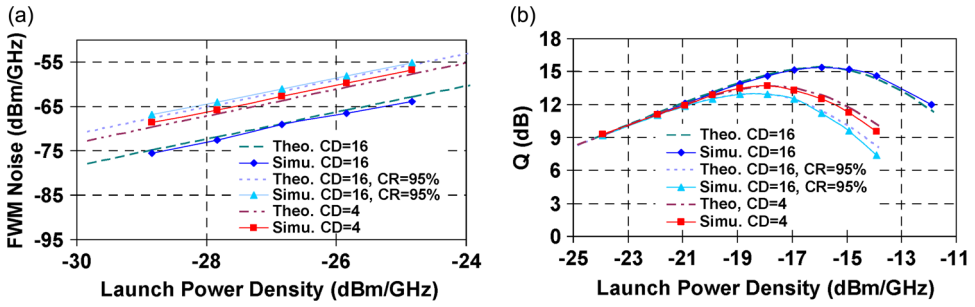
IEEE Photonics Journal
An IEEE Photonics Society Publication

Invited Paper

Information Spectral Efficiency and Launch Power Density Limits Due to Fiber Nonlinearity for Coherent Optical OFDM Systems

Volume 3, Number 2, April 2011

William Shieh
Xi Chen



DOI: 10.1109/JPHOT.2011.2112342
1943-0655/$26.00 ©2011 IEEE

Information Spectral Efficiency and Launch Power Density Limits Due to Fiber Nonlinearity for Coherent Optical OFDM Systems

William Shieh and Xi Chen

(Invited Paper)

National ICT Australia and Centre for Energy Efficient Telecommunications, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne Vic. 3010, Australia

DOI: 10.1109/JPHOT.2011.2112342
1943-0655/$26.00 ©2011 IEEE

Manuscript received December 17, 2010; accepted January 28, 2011. Date of publication March 2, 2011; date of current version March 8, 2011. Corresponding author: W. Shieh (e-mail: shiehw@unimelb.edu.au).

Abstract: We derive closed-form expressions for nonlinear transmission performance of dual-polarization closely spaced coherent optical orthogonal frequency-division multiplexing (CO-OFDM) systems. We find that the fiber nonlinear noise exhibits a signature of the flicker noise or $1/f$ noise beyond the corner frequency that is inversely proportional to the total participating bandwidth. We derive a noise enhancement factor that captures the interference effect of nonlinear noises among different spans. For a 10×100-km standard single-mode-fiber (SSMF) link with no dispersion compensation at an optimal launch power density of -15.9 dBm/GHz, the spectral efficiency of 9.90 b/s/Hz can be achieved for dual-polarization transmission, which is about 93% increase over single-polarization transmission. The closed-form expressions are also applicable to the closely spaced coherent single-carrier systems where the symbol rate is much larger than the dispersion walk-off bandwidth, and the optical dispersion is uncompensated.

Index Terms: Fiber nonlinearity, polarization effects, coherent communications, orthogonal frequency-division multiplexing (OFDM).

1. Introduction

The maximum spectral efficiency of optical fiber transmission can be achieved by removing the frequency guard bands between wavelength channels. The potential crosstalk due to such dense wavelength packing can be resolved with the concept of coherent optical orthogonal frequency-division multiplexing (CO-OFDM) [1]. In such systems, the CO-OFDM wavelength channels can be either continuously spaced without frequency guard band [2]–[5] or densely spaced with extremely small frequency guard band [6], [7]. Most recently, the nonlinear transmission performance of CO-OFDM systems has been reported, including analytical results for single-channel transmission without consideration of chromatic dispersion in [8], complete analytical expressions involving summation of a large number of nonlinearity products in [9], and computer numerical simulation of CO-OFDM system performance in [10]–[12]. It would be of great interest to derive concise closed-form solutions that capture the dependence of nonlinear performance on some major system parameters such as chromatic dispersion and dispersion compensation ratio. This body of analytical

work on nonlinear system performance was pioneered in [13] and [14], where nonlinear launch power and information capacity are derived in closed form. However, there are two limitations for the reports in [13] and [14]: i) They assume that nonlinear phase noise is generated independently in different spans, ignoring an important phase array effect of the four-wave-mixing (FWM) products that accounts for the interference among multiple spans [9], and ii) they only present the results for single polarization and thereby cross-polarization nonlinear interaction is not included. In this paper, we extend our previous work on single-polarization transmission [15] and derive closed-form analytical expressions for nonlinear system performance of dual-polarization CO-OFDM systems. The closed-form solution entails the results for achievable Q factor, optimum launch power density, nonlinear threshold of launch power density, and information spectral efficiency limit. These analytical results clearly identify the nonlinear performance dependence on system parameters including fiber dispersion, number of spans, dispersion compensation ratio, and overall bandwidth.

2. Analytical Derivation of Polarization Dependent FWM Noise

Because of random fluctuation of the birefringence eigen axis in the fiber, the optical OFDM signal will go through many small segments of birefringence fiber in a commonly used transmission fiber. Wai *et al.* have developed a nonlinear Schrödinger equation that includes both linear and nonlinear polarization effects given by [16], [17]

$$i\frac{\partial \boldsymbol{A}}{\partial z} + i\frac{1}{2}[\alpha + \vec{\alpha}'(z)\cdot\vec{\sigma}]\boldsymbol{A} - i\frac{1}{2}\vec{\gamma}'(z)\cdot\vec{\sigma}\frac{\partial \boldsymbol{A}}{\partial t} - \frac{1}{2}\beta_2\frac{\partial^2 \boldsymbol{A}}{\partial t^2} + \gamma_0\frac{8}{9}|\boldsymbol{A}|^2\boldsymbol{A} = 0 \tag{1}$$

where $\vec{\gamma}'(z)$ and $\vec{\alpha}'(z)$ are the Polarization-Mode Dispersion (PMD) and Polarization-Dependent Loss (PDL) vector in the direction of the eigen axis of a local large segment of the fiber [18], [19], α and β_2, respectively, represent the fiber loss and chromatic dispersion, $\vec{\sigma}$ is the Pauli matrix vector [16]–[19], γ_0 is the third-order nonlinear coefficient for a fiber with a fixed birefringence eigen axis, superscript "T" stands for matrix transpose, the bold font stands for Jones vector or Jones Matrix, the overhead arrow stands for 3-D vector, and $\boldsymbol{A}(z,t)$ is the time-domain CO-OFDM signal expressed as

$$\boldsymbol{A}(t) = \sum_{k=-N/2+1}^{N/2} \boldsymbol{c}'_k \exp(j2\pi f_k t), \qquad f_k = k\Delta f \tag{2}$$

$$\boldsymbol{A}(t) = [A_x, A_y]^T, \quad \boldsymbol{c}'_k = \left[c_x'^k, c_y'^k\right]^T \tag{3}$$

where f_k is the frequency for the kth subcarrier, Δf is the subcarrier spacing, N is the number of subcarriers, $\boldsymbol{c}'_k$ is the OFDM information symbol for the kth subcarrier as a Jones vector. $\boldsymbol{A}(z,t)$ is the time-domain signal using the rotating local polarization co-ordinate and is subsequently different from the signal using the fixed coordinate [16], [17]. Nevertheless, these two representations are only different by a linear transform in the frequency domain, and therefore, which coordinate to use does not matter as far as the system performance is concerned. In this paper, we investigate the densely spaced OFDM (DS-OFDM) systems where the frequency guard band is much narrower than the bandwidth of each wavelength [15]. As such, the frequency guard band can be omitted in the remainder of the investigation. In such DS-OFDM systems, all the nonlinear effects such as XPM, FWM, and SPM can be considered as FWM between all the subcarriers if we treat multiple densely spaced wavelength channels as an effective big "single-band" OFDM channel that encompasses all the subcarriers [15]. Additionally, we omit the effects of PDL and PMD, the study of which warrantees a separate treatment. With these assumptions, setting both $\vec{\gamma}'(z)$ and $\vec{\alpha}'(z)$ to zero, we are left with much simplified Manakov equation

$$i\frac{\partial \boldsymbol{A}}{\partial z} + i\frac{\alpha}{2}\boldsymbol{A} - \frac{1}{2}\beta_2\frac{\partial^2 \boldsymbol{A}}{\partial t^2} + \gamma|\boldsymbol{A}|^2\boldsymbol{A} = 0 \tag{4}$$

where $\gamma = 8\gamma_0/9$ is the third-order nonlinear coefficient for a fiber with randomly varying birefringence eigen axis, which is applicable to any commonly deployed transmission fiber.

Fig. 1. Schematic of one span of fiber link consisting of a transmission fiber and a dispersion compensation fiber (DCF), which is sandwiched between two Erbium-doped-fiber-amplifiers (EDFA), i.e., EDFAs I and II.

FWM is the third-order nonlinearity effect and its impact on optical fiber communications has been extensively studied [20], [21]. Due to the FWM, the interaction of subcarriers at the frequencies of f_i, f_j, and f_k produces a mixing product at the frequency of $f_g = f_i + f_j - f_k$. For explanatory simplicity, we assume for now the existence of only three initial subcarriers at f_i, f_j, and f_k, which produces the fourth frequency at f_g. The overall FWM effects will be a summation over the arbitrary f_j, and f_k, which will be discussed later in this section. With this assumption, optical signal in the time-domain in (2) becomes

$$\boldsymbol{A}(t) = \boldsymbol{c}'_i \exp(j\omega_i t) + \boldsymbol{c}'_j \exp(j\omega_j t) + \boldsymbol{c}'_k \exp(j\omega_k t) + \boldsymbol{c}'_g \exp(j\omega_g t) \tag{5}$$

where $\omega_{i,j,k,g} = 2\pi f_{i,j,k,g}$. Substituting (5) into the Manakov Equation (4) and after some simple reordering according to $\exp(j\omega_{i,j,k,g}t)$, the propagation equation for the subcarrier f_g is given by [20]

$$i\frac{\partial \boldsymbol{c}'_g}{\partial z} + i\frac{\alpha}{2}\boldsymbol{c}'_g + \frac{1}{2}\beta_2\omega_g^2\boldsymbol{c}'_g + \gamma\left[\left(\boldsymbol{c}'^{+}_k\boldsymbol{c}'_i\right)\boldsymbol{c}'_j + \left(\boldsymbol{c}'^{+}_k\boldsymbol{c}'_j\right)\boldsymbol{c}'_i\right] = 0 \tag{6}$$

where the superscript "+" stands for transpose conjugate. We have assumed that the FWM product $\boldsymbol{c}'_g$ is much smaller than the inducing field $\boldsymbol{c}'_{i,j,k}$. We now adopt the commonly used undepleted subcarrier approach where each subcarrier field $\boldsymbol{c}'_{i,j,k}$ can be expressed as [20]

$$\boldsymbol{c}'_{i,j,k}(z) = \boldsymbol{c}_{i,j,k} \exp\left(-\frac{1}{2}\alpha z - i\beta_{i,j,k}z\right) \tag{7}$$

$$\beta_{i,j,k} = \frac{1}{2}\beta_2\omega_{i,j,k}^2 \tag{8}$$

where $\boldsymbol{c}_{i,j,k}$ is the OFDM symbol for i, j, or kth subcarrier at the input of the fiber. This distinguishes from the OFDM symbols during the transmission $\boldsymbol{c}'_{i,j,k}$, represented by the prime in the superscript.

Substituting (7) into (6), the magnitude of the FWM product $\boldsymbol{c}'_g$ after transmission of one-span of the transmission fiber is given by

$$\boldsymbol{c}'_g = \gamma\left[\left(\boldsymbol{c}_k^{+}\boldsymbol{c}_i\right)\boldsymbol{c}_j + \left(\boldsymbol{c}_k^{+}\boldsymbol{c}_j\right)\boldsymbol{c}_i\right]e^{-\alpha L/2 - i\beta_g L}\frac{1 - e^{-\alpha L}e^{-j\Delta\beta_{ijk}L}}{j\Delta\beta_{ijk} + \alpha}. \tag{9}$$

We now extend to N_s spans of fiber link with a periodic dispersion map. Each span consisting of a transmission fiber, and a dispersion compensation fiber (DCF) sandwiched between two EDFAs, as shown in Fig. 1. For simplicity, we assume that the FWM products are dominated by those from the transmission fiber, and the complete analysis including the nonlinearity and loss effects of DCF will be performed in the Appendix B. Summing the FWM mixing over N_s spans, the accumulated FWM products at the N_sth span c'_{g,N_s} becomes

$$\boldsymbol{c}'_{g,N_s} = \boldsymbol{c}'_g \sum_{M=0}^{N_s-1} \exp(jM \cdot \Delta\tilde{\beta}_{ijk}) = \boldsymbol{c}'_g \frac{1 - \exp(jN_s \cdot \Delta\tilde{\beta}_{ijk})}{1 - \exp(j\Delta\tilde{\beta}_{ijk})} \tag{10}$$

$$\Delta\beta_{ijk} \equiv \beta_i + \beta_j - \beta_k - \beta_g \tag{11}$$

$$\Delta\tilde{\beta}_{ijk} = \Delta\beta_{ijk}L + \Delta\beta_{ijk,1}L_1 \tag{12}$$

where $\Delta\beta_{ijk}$ is the phase mismatch in the transmission fiber, and L is the transmission fiber length per span. In (12), the subscript "1" stands for the parameters associated with the dispersion compensation fiber (DCF). $\boldsymbol{c}'_g$ is the FWM product for one-span as shown by (9). The expression of (10) derived here from the Manakov equation extends the result in [22] by including the multispan chromatic dispersion effects. It is noted that we have made an important assumption that the phase mismatch and phase walk-off per span are dominated by the chromatic dispersion instead of PMD effects as we omit the PMD and PDL impact in this analysis. From (9), the power of the FWM component at subcarrier f_g is

$$P'_g = \left|\boldsymbol{c}'_{g,N}\right|^2 = \left|(\boldsymbol{c}_k^+\boldsymbol{c}_i)\boldsymbol{c}_j + (\boldsymbol{c}_k^+\boldsymbol{c}_j)\boldsymbol{c}_i\right|^2\gamma^2 e^{-\alpha L}\eta' \tag{13}$$

where η' is the FWM coefficient which has a strong dependence on the relative frequency spacing between the FWM components given by

$$\eta' = \eta'_1\eta'_2 \tag{14}$$

$$\eta'_1 = \left|\frac{1 - e^{-\alpha L}e^{-j\Delta\beta_{ijk}L}}{j\Delta\beta_{ijk} + \alpha}\right|^2 \approx \frac{1}{(\Delta\beta_{ijk})^2 + \alpha^2} \tag{15}$$

$$\eta'_2 = \left|\frac{1 - \exp(jN_s\cdot\Delta\tilde{\beta}_{ijk})}{1 - \exp(j\Delta\tilde{\beta}_{ijk})}\right|^2 = \frac{\sin^2\{N_s\Delta\tilde{\beta}_{ijk}/2\}}{\sin^2\Delta\tilde{\beta}_{ijk}/2}. \tag{16}$$

In (14), the overall FWM efficiency is decomposed into two separate contributions: i) η'_1, which is the FWM efficiency coefficient for single span, and ii) η'_2, which is the interference effect among N_s spans of FWM products, which is also known as phase array effect [9]. In (15), we have assumed that the span loss $e^{\alpha L}$ is much larger than 1, and therefore, $e^{-\alpha L}e^{-j\Delta\beta_{ijk}L}$ is removed from the nominator. Substituting mth subcarrier frequency having the form of $f_m = m\cdot\Delta f$ into (15) and (16), the phase mismatch terms $\Delta\beta_{ijk}$ and $\Delta\tilde{\beta}_{ijk}$ can be rewritten as

$$\Delta\beta_{ijk} = -4\pi^2\beta_2\Delta f^2(i-k)(j-k) \tag{17}$$

$$\Delta\tilde{\beta}_{ijk} = -4\pi^2\beta_2\Delta f^2 L(1-\rho)(i-k)(j-k) \tag{18}$$

where ρ is the dispersion compensation (or residual dispersion) ratio. We assume that the launched OFDM information symbol $\boldsymbol{c}_{i,j,k}$ for different subcarriers are statistically independent. It is shown in Appendix A that

$$\left\langle\left|(\boldsymbol{c}_k^+\boldsymbol{c}_i)\boldsymbol{c}_j + (\boldsymbol{c}_k^+\boldsymbol{c}_j)\boldsymbol{c}_i\right|^2\right\rangle = \frac{3}{2}P_iP_jP_k \tag{19}$$

where $\langle\ \rangle$ stands for ensemble average, and $P_{i,j,k}$ is the launch power at the frequency of $f_{i,j,k}$. Consequently, the ensemble average of the FWM power becomes

$$\left\langle P'_g\right\rangle = \left\langle\left|(\boldsymbol{c}_k^+\boldsymbol{c}_i)\boldsymbol{c}_j + (\boldsymbol{c}_k^+\boldsymbol{c}_j)\boldsymbol{c}_i\right|^2\gamma^2 e^{-\alpha L}\eta'\right\rangle = \frac{3}{2}P_iP_jP_k\gamma^2 e^{-\alpha L}\eta'. \tag{20}$$

At the end of each span, the FWM product P'_g along with the signal will be amplified by a gain of G equal to the loss of each span $e^{\alpha L}$ and the FWM product becomes

$$\langle P_g\rangle = \frac{3}{2}\gamma^2 P_iP_jP_k\eta'. \tag{21}$$

We adopt the approach used in [13], where the nonlinear effect is considered as the multiplicative noise to the signal. In essence, we consider ith subcarrier as the reference frequency, and j and k frequencies as the interferers, namely, frequency j and frequency k generate a beating frequency

component at $(f_j - f_k)$, which in turn modulates the subcarrier i, creating fourth components of f_g. Consequently, the nonlinearity impinging on subcarrier i, P_{NL}^i is given by

$$P_{NL}^i = \frac{3}{4}\gamma^2 P_i \sum_{k=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} P_j P_k \eta'. \tag{22}$$

A factor of one half is added in (22) because of the double counting in the dual summation. The lower summation boundary is reduced by 1 for simplicity. Equation (22) can be understood as the number of photons or amount of energy scattered off the subcarrier i and should be equivalent to the photons scattered into this subcarrier i with large bandwidth assumption, which we will clarify later. From now on, we drop index i and set it to zero, or equivalently, we are investigating the performance of center wavelength channels in broad bandwidth DS-OFDM systems. We also assume all the subcarriers have the same power of P for the sake of simplicity. The FWM power at the center subcarriers becomes

$$\begin{aligned} P_{NL} &= \frac{3}{4}\gamma^2 P^3 \sum_{k=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} \eta_1' \eta_2' \\ \eta_1' &= \left| \frac{\sin\left(N_s j(k-j)\Delta f^2/(2f_{PA}^2)\right)}{\sin\left(j(k-j)\Delta f^2/(2f_{PA}^2)\right)} \right|^2, \quad \eta_2' = \frac{1}{\beta_2^2(2\pi)^4} \frac{1}{\Delta f^4 j^2 (k-j)^2 + f_W^4} \\ f_{PA} &\equiv \frac{1}{2\pi}\sqrt{\frac{1}{|\beta_2| L \zeta}}, \qquad f_W \equiv \frac{1}{2\pi}\sqrt{\frac{\alpha}{|\beta_2|}} \end{aligned} \tag{23}$$

where f_{PA} is defined as the phase array bandwidth indicating frequency range of the effectiveness of phase array effects, and f_W is the defined as the dispersion walk-off bandwidth indicating the frequency range of the effectiveness of FWM nonlinearity in the presence of the dispersion. Substituting a new variable $m = k - j$, (23) becomes

$$P_{NL} = \frac{3\gamma^2 P^3}{4\beta_2^2(2\pi)^4} \sum_{m=-N/2-j}^{N/2-j} \sum_{j=-N/2}^{N/2} \left| \frac{\sin\left(N_s jm \cdot \Delta f^2/(2f_{PA}^2)\right)}{\sin\left(jm \cdot \Delta f^2/(2f_{PA}^2)\right)} \right|^2 \frac{1}{\Delta f^4 j^2 m^2 + f_W^4}. \tag{24}$$

Similar to the single-polarization scenario [15], it can be shown that under the condition of

$$f_{PA} \gg \Delta f \tag{25}$$

$$f_W \gg \Delta f. \tag{26}$$

FMW expression of (24) can be converted from discrete summation to integration. We call the conditions of (25) and (26) "dense subcarrier" assumptions. Under the assumptions of (25) and (26), substituting the continuous integral variable f for $m \cdot \Delta f$, f_1 for $j \cdot \Delta f$, the FWM power is transformed into

$$\begin{aligned} P_{NL} &= \frac{3\gamma^2}{4\beta_2^2(2\pi)^4} \frac{P^3}{\Delta f^2} \int_{-B/2-f_1}^{B/2-f_1} \int_{-B/2}^{B/2} \eta_1(f, f_1)\eta_2(f, f_1)\, df_1\, df \\ \eta_1(f, f_1) &= \left| \frac{\sin\left(N_s f_1 f/(2f_{PA}^2)\right)}{\sin\left(f_1 f/(2f_{PA}^2)\right)} \right|^2, \qquad \eta_2(f, f_1) = \frac{1}{(f_1 f)^2 + f_W^4} \end{aligned} \tag{27}$$

where $B = N\Delta f$ is the total bandwidth of the DS-OFDM systems. According to the definition of m in (24), the variable f represents the frequency of the multiplicative noise impairing the channel. We

now introduce more convenient and fundamentally more important terms: power (spectral) densities given by

$$I_{NL} \equiv \frac{P_{NL}}{\Delta f}, \quad I \equiv \frac{P}{\Delta f} \tag{28}$$

where I_{NL} and I are, respectively FWM noise (spectral) density and launch power (spectral) density. Substituting (28) into (27), we arrive at the FWM noise density

$$I_{NL} = \frac{3\gamma^2}{4\beta_2^2(2\pi)^4} I^3 \int_{-B/2-f_1}^{B/2-f_1} \int_{-B/2}^{B/2} \eta_1(f, f_1)\eta_2(f, f_1)\, df_1\, df. \tag{29}$$

The important conclusion from (29) is that under "dense subcarrier" assumption, the result of the nonlinearity is independent of the subcarrier spacing.

Similar to the derivation in [15, App. A], the FWM power density can be rewritten as

$$I_{NL} = \frac{3\gamma^2}{\beta_2^2(2\pi)^4} I^3 \int_{B_0/2}^{B/2} \int_{0}^{\infty} \eta_1(f, f_1)\eta_2(f, f_1)\, df_1\, df \tag{30}$$

$$B_0 = 2f_W^2/B \tag{31}$$

$$B \gg f_W. \tag{32}$$

Equation (32) is another important assumption that is used for deriving (30), which states that the overall bandwidth is much larger than the walk-off bandwidth. We call this condition the "large bandwidth" assumption. Since f is the nonlinearity noise frequency, the integration over f_1 in (30) would produce the nonlinear noise spectral density. We rewrite (30) in terms of the one-sided nonlinear multiplicative noise spectral density $i_{NL}(f)$ given by

$$I_{NL} = I \int_{B_0/2}^{B/2} i_{NL}(f)\, df, \qquad i_{NL}(f) = \frac{3\gamma^2}{\beta_2^2(2\pi)^4} I^2 \int_{0}^{\infty} \eta_1(f, f_1)\eta_2(f, f_1)\, df_1. \tag{33}$$

The nonlinear noise spectral density $i_{NL}(f)$ has the unit of dBc/Hz, which is similar to phase noise or relative intensity noise (RIN). $i_{NL}(f)$ can be integrated in closed form, the derivation of which is shown in [15, App. B]. The result of the integration gives

$$i_{NL}(f) = \gamma^2 I^2 \frac{3}{4\pi\alpha|\beta_2|} \left(\frac{(N_s - 1 + e^{-\alpha\zeta L N_s} - N_s e^{-\alpha\zeta L})e^{-\alpha\zeta L}}{(e^{-\alpha\zeta L} - 1)^2} + \frac{N_s}{2} \right) \frac{1}{f}. \tag{34}$$

Equation (34) is the first important result of the paper. The significance of (34) shows that the multiplicative nonlinear noise spectral density is essentially a well-known flicker noise or $1/f$ noise. This finding makes the authors deduce that our derivation may help explain one class of the flicker noise, namely, third-order nonlinearity and dispersion may be one type of mechanisms to produce of $1/f$ noise. B_0 is the corner frequency of the $1/f$ noise below which the nonlinear noise starts to roll off. It follows from (31) that the corner frequency B_0 is inversely proportional to the bandwidth of the participating noise B. Substituting (34) into (33), we finally arrive at the closed-form expression for the nonlinear noise power density I_{NL}

$$I_{NL} = \gamma^2 I^3 \frac{4}{\pi\alpha|\beta_2|} \left(\frac{2(N_s - 1 + e^{-\alpha\zeta L} - N_s e^{-\alpha\zeta L})e^{-\alpha\zeta L}}{(e^{-\alpha\zeta L} - 1)^2} + N_s \right) = \frac{3\gamma^2 N_s \ln(B/B_0) \cdot h_e}{8\pi\alpha|\beta_2|} I^3 \tag{35}$$

$$h_e \equiv \frac{2(N_s - 1 + e^{-\alpha\zeta L N_s} - N_s e^{-\alpha\zeta L})e^{-\alpha\zeta L}}{N_s(e^{-\alpha\zeta L} - 1)^2} + 1 \tag{36}$$

where h_e is the (noise) enhancement factor accounting for the FWM noise interference among different spans. We will discuss this interesting nonlinear enhancement factor h_e in more detail in the next section. We further express the nonlinear noise power density I_{NL} of (35) in a more concise form with the definition of nonlinear characteristic power density I_0 as follows:

$$I_{NL} = \left(\frac{I}{I_0}\right)^2 I, \qquad I_0 \equiv \frac{1}{\gamma}\sqrt{\frac{8\pi\alpha|\beta_2|}{3N_s h_e \ln(B/B_0)}}. \tag{37}$$

2.1. Signal-to-Noise Ratio and Spectral Efficiency Limit in the Presence of Nonlinearity

The signal power in presence of the nonlinear interference can be expressed as [13]

$$I = I\exp\left(-(I/I_0)^2\right) \cong I. \tag{38}$$

The noise can be considered as the summation of the white optical amplified-spontaneous-noise (ASE), i.e., n_0, and the FWM noise and is shown given by [13]

$$n = 2n_0 + I\left(1 - \exp\left(-(I/I_0)^2\right)\right) \tag{39}$$

$$n_0 = N_s(G-1)n_{sp}h\upsilon \cong 0.5N_s e^{\alpha L}h\upsilon \cdot NF \tag{40}$$

where n_{sp} is the spontaneous noise factor equal to half of the noise figure of the optical amplifier NF, h is the Planck constant, and υ is the light frequency. The factor of 2 in (39) accounts for the ASE noise for both polarizations, namely, the signal power and noise power density in (38) and (39) include the contribution from both polarizations. The signal-to-noise ratio (SNR) is thus given by

$$\mathrm{SNR} = \frac{I\exp\left(-(I/I_0)^2\right)}{2n_0 + I\left(1 - \exp\left(-(I/I_0)^2\right)\right)}. \tag{41}$$

For the SNR larger than 10, (41) can be approximated as

$$\mathrm{SNR} \cong \frac{I}{2n_0 + I(I/I_0)^2}. \tag{42}$$

The simplification is generally valid for the case of interests where the signal power density is much smaller than I_0.

We have verified through our simulation under "dense subcarrier" and "large bandwidth" assumptions of (25), (26), and (32) that the FWM noise is of Gaussian distribution. This is also verified previously in [23]. Under the assumption of Gaussian noise distribution, the information spectral efficiency (defined as the maximum information capacity C normalized to bandwidth B) for dual-polarization is readily given by [24]

$$S = 2\log_2(1+\mathrm{SNR}) = 2\log_2\left(1 + \frac{I\exp\left(-(I/I_0)^2\right)}{2n_0 + I\left(1 - \exp\left(-(I/I_0)^2\right)\right)}\right)$$

$$\cong 2\log_2\left(1 + \frac{I}{2n_0 + I(I/I_0)^2}\right). \tag{43}$$

From (43), the maximum spectral efficiency S_{opt} in the presence of fiber nonlinearity can be easily shown as

$$S_{opt} = 2\log_2\left(1 + \frac{1}{3}(I_0/n_0)^{2/3}\right). \tag{44}$$

2.2. Optimal Launch Power Density, Maximum Q, and Nonlinear Threshold of Launch Power Density

In (43), the ultimate spectral efficiency is obtained. However, in practice, the performance is always lower because of the practical implementation of modulation and coding. We therefore derive a few important parameters that are commonly used in the optical communications community. The first one is the maximum achievable Q factor. Under the Gaussian noise assumption and quadrature phase-shift keying (QPSK) modulation, the Q factor is equal to the SNR given by

$$Q = \mathrm{SNR} = \frac{I\exp\left(-(I/I_0)^2\right)}{2n_0 + I\left(1-\exp\left(-(I/I_0)^2\right)\right)} \cong \frac{I}{2n_0 + I(I/I_0)^2}. \tag{45}$$

The optimum launch power density is another important parameter and is defined as the launch power density where the maximum Q takes place. By simply differentiating Q of (45) over I and setting it to zero, we obtain that the optimum launch power density and the optimal Q is given by

$$I_{opt} = \left(n_0 I_0^2\right)^{1/3} = \left(\frac{8n_0\pi\alpha|\beta_2|}{3\gamma^2 N_s h_e \ln(B/B_0)}\right)^{1/3} \tag{46}$$

$$Q_{max} = \frac{1}{3}\left(\frac{I_0}{n_0}\right)^{2/3} = \frac{(8\pi\alpha|\beta_2|)^{1/3}}{3\left(3n_0^2\gamma^2 N_s h_e \ln(B/B_0)\right)^{1/3}}. \tag{47}$$

One of the inconveniences of using the optimum launch power expression in (46) is that it is dependent on the amplifier noise figure. The other commonly used term is nonlinear threshold launch power density that is defined as the maximum launch power density at which the BER due to the nonlinear noise can no longer be corrected by a certain type of forward-error-correction (FEC). For standard Reed–Solomon code RS(255, 239), the threshold Q is 9.8 (dB), or linear q_0 of 3.09. In (45), setting n_0 to zero and Q to q_0^2, we arrive at the nonlinear threshold of power density

$$I_{th} = \frac{I_0}{q_0} = \frac{1}{q_0\gamma}\sqrt{\frac{8\pi\alpha|\beta_2|}{3N_s h_e \ln(B/B_0)}} \tag{48}$$

where q_0 is the correctable linear Q for a specific FEC.

The closed-form expressions for nonlinear noise spectral density $i_{NL}(f)$ in (34), nonlinear noise power density I_{NL} in (35), nonlinear multispan noise enhancement factor h_e in (36), nonlinear characteristic power density I_0 in (37), information spectral efficiency S in (43) and its optimal value in (44), system Q factor in (45) and its optimal value in (47), optimal launch power density in (46), and nonlinear threshold of launch power density in (48) comprise the major findings in this work.

3. Application of the Closed-Form Expressions

3.1. Corroboration of the Theories With Numerical Simulation

We first conduct simulation of some commonly used systems to substantiate the theories developed in Section 2. The parameters for the simulated dual-polarization transmission systems are as follows: 16 wavelength channels, each covering 31-GHz bandwidth, giving total bandwidth B of 496 GHz; OFDM subcarrier frequency spacing of 85 MHz; QPSK modulation for each subcarrier; no frequency guard band between wavelength channels; 10-span of 100 km fiber link; fiber loss α of 0.2 dB/km; nonlinear coefficient $\gamma = 1.22$ w^{-1}km^{-1}; noise figure of the amplifier of 6 dB. The FWM noise density is simulated by using a perfect optical notch filter to notch out a 100-MHz gap at the center of the input signal spectrum, and the power density is measured at the output after 1000-km transmission. Fig. 2(a) shows the simulated nonlinear noise density compared with the computed nonlinear density using the closed-form expression of (37). Three transmission systems are investigated: i) single-mode-fiber (SSMF) type system with CD of 16 ps/nm/km with no dispersion

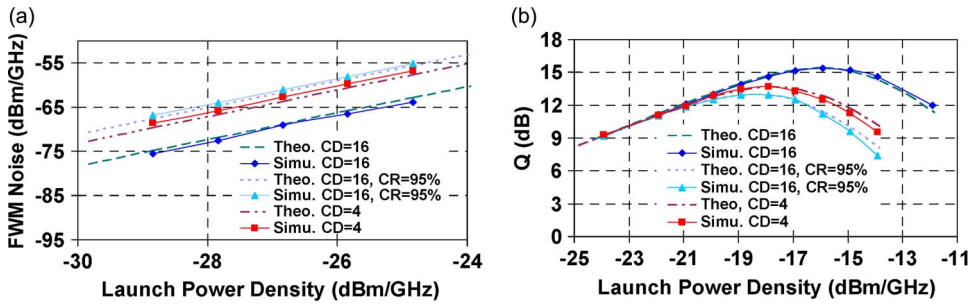


Fig. 2. Comparison of closed-form theory and simulation results for (a) FWM power density and (b) Q factor as a function of the launch power density. Theo.: Theory; Simu.: Simulation; CD: Chromatic dispersion with a unit of ps/nm/km; CR: (CD) Compensation ratio. Both (a) and (b) assume 10×100 km dual-polarization transmission systems.

compensation, abbreviated as "system I," ii) CD of 16 ps/nm/km but with dispersion 95% compensated per span, abbreviated as system II, and iii) nonzero dispersion-shifted type fiber with CD of 4 ps/nm/km, abbreviated as "system III." For systems I, II, and III, the average difference of FWM density is 25%, 20%, and 20%. This shows excellent match between the closed-form formula and simulation, considering the extreme sensitivity of the FWM density as a function of launch power density (cubic dependence). We also perform the simulation of the system Q factors with the above-described three systems, the results of which are shown in Fig. 2(b). We can see a good match between theoretical expressions based on (45) and simulation results. For instance, the difference between the optimal Q from theory and simulation is within 0.2 dB for all simulated dispersion maps. The difference of launch power between the simulation and closed-form theory for the same Q factor is less than 0.3 dB for wide range of launch power density of -25 to -12 dBm/GHz. All these confirm the excellent match between the simulation and the closed-form expression of Q factors in (45).

3.2. System Q Factor and Optimum Launch Power

Because the concise closed-form expressions are available, we are ready to quickly identify their dependence on system parameters including fiber dispersion, number of spans, dispersion compensation ratio, and overall bandwidth. In this section, we will discuss in detail the achieved system Q factor, optimum launch power density, information spectral efficiency, and multispan noise enhancement factor.

The immediate benefits of having closed-form formulas of (46) and (47) for system Q factor and optimum launch power density are their scaling over the underlying parameters. From (46) and (47), it follows that for every 3-dB increase in fiber dispersion, there is a 1-dB increase in the optimal launch power density and the achievable Q; for every 3-dB increase in fiber nonlinear coefficient γ, there is a 2-dB decrease in the optimal launch power density and achievable Q.

We can quickly generate the optimum launch power density and achievable Q for variety of dispersion maps. In particular, we investigate the three systems: systems I, II, and III that are described in Section 3.1. As shown in Fig. 3(a), system I has the best performance due to large local dispersion and no per-span dispersion compensation. The advantage of system I over system II increases with the increase of the number of spans, for instance, from 0 dB to 2.4 dB when the reach increases from single-span to 10 spans. The advantage of system I over system III is maintained at 1.7 dB, independent of the number of spans. Fig. 3(b) shows the optimal launch power versus number of spans. The optimum launch powers for noncompensated systems, i.e., systems I and III, are constant. This is because both the linear and nonlinear noises increase linearly with the number of the spans that leads to the optimum power independent of the number of spans. However, for the dispersion compensated system II, the optimum launch power density decreases with number of spans due to the multispan noise enhancement effect. Another interesting observation from (46) and (47) is that both the optimal Q factor and launch power has

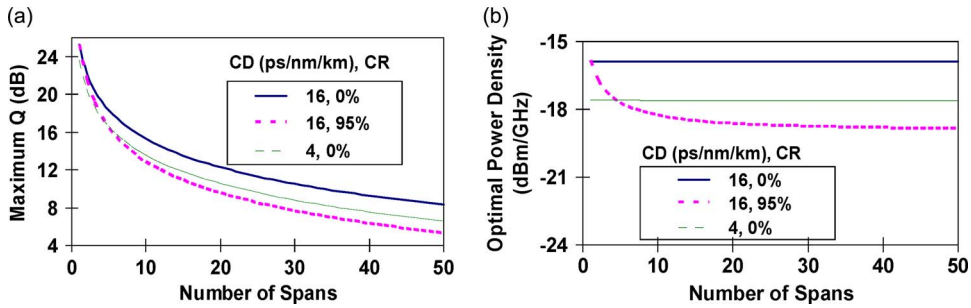


Fig. 3. (a) Maximum Q factor and (b) the optimal launch power density versus number of spans with various dispersion maps. CD: Chromatic dispersion. CR: (CD) Compensation ratio.

very weak dependence on the overall system bandwidth: proportional to 1/3 power of logarithm of the overall bandwidth. It can be easily shown that for both systems I and III, the Q is decreased by only about 0.7 dB with the tenfold increase of the overall system bandwidth from 400 to 4000 GHz, whereas system II incurs a larger decrease of the Q factor of 0.84 dB with the same bandwidth increase.

3.3. Information Spectral Efficiency

The information spectral efficiency is important as it represents the ultimate bound of what we can achieve by employing all possible modulations (of course not limited to QPSK) and codes. For large SNR, we simplify (43) into

$$S = 2\log_2\left(1+\frac{1}{3}(I_0/n_0)^{2/3}\right) \cong 2\log_2\left(\frac{1}{3}(8\pi\alpha|\beta_2|)^{1/3}\left(3\gamma^2 n_0^2 N_s h_e \ln(B/B_0)\right)^{-1/3}\right). \quad (49)$$

Equation (49) clearly shows the challenges of improving spectral efficiency by redesigning the fiber system parameters: To increase spectral efficiency by 2 bit/s/Hz, the dispersion needs to be increased by a factor of 8, the nonlinear coefficient γ needs to be decreased by a factor of 2.8, or the number of spans needs to be reduced by a factor of 2, all of which are difficult to achieve. In a nutshell, it is of diminishing return to improve the spectral efficiency by modifying the optical fiber system parameters. The only effective method to substantially improve the spectral efficiency is to add more dimensions such as resorting to polarization multiplexing that leads almost a factor of 2 improvement, as discussed in the paper, or fiber mode multiplexing by at least a factor of two or more dependent on the capability of achievable digital signal processing (DSP). Fig. 4 shows the achievable spectral efficiency for the three systems studied in Section 3.1. The only modification is that we assume 40 nm or 5 THz for the total bandwidth. The spectral efficiency for the systems I, II, and III are, respectively 9.90, 8.38, and 8.63 b/s/Hz over 10 spans of transmission. This shows a total capacity of 49.5 Tb/s can be achieved for 10 $\times$ 100 km SSMF uncompensated EDFA-only dual-polarization systems within C-band.

3.4. Multispan Noise Enhancement Factor

The multispan noise enhancement factor h_e of (36) is one of the most important findings in this report. This noise enhancement effect is ignored in the prior analytical results [13], [14]. This multispan interference effect can be understood as the phase array effect that has been discussed in [9]. The noise enhancement is referred to the important fact that the overall nonlinear FWM noise of multispan systems is enhanced by a factor of h_e over the scenario for which the nonlinear FWM noise originated in each span is assumed independent without interference with each other. We note the expression of h_e in (36) is the first concise closed-form result where the multispan interference effects of all the FWM products are accounted for. From (36), we conclude that as long as the factor $\alpha \boldsymbol{L}\zeta$ is much larger than 1, h_e approaches 1, namely, the nonlinear noise generated in each span can be treated independently in this regime. However, even when the fiber loss $\alpha \boldsymbol{L}$ is

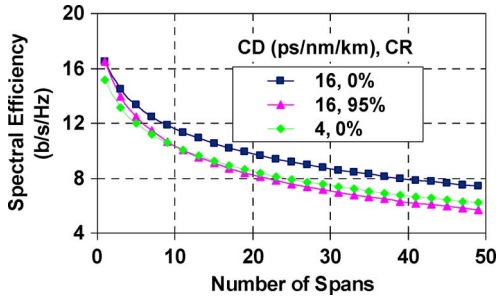


Fig. 4. Information spectral efficiency as a function of the number of spans for various dispersion maps. The total bandwidth B is assumed to be 40 nm. The other OFDM and link parameters are the same as those described in Section 3.1.

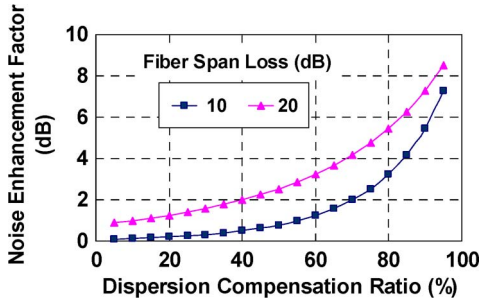


Fig. 5. Multi-span noise enhancement factor as a function of the dispersion compensation ratio with fiber span losses of 10 and 20 dB. The number of spans is maintained at 10. The link losses of 10 and 20 dB are obtained by setting the span length to 50 and 100 km, respectively.

large but ζ is small, or dispersion compensation ratio is large, h_e can be significantly high. Fig. 5 shows the noise enhancement factor h_e as a function of dispersion compensation ratio $\rho = 1 - \zeta$ for span losses of 10 and 20 dB. It can be seen that for a dispersion compensation ratio of 95%, the nonlinear noise is enhanced by 8.5 and 7.3 dB for span losses of 10 and 20 dB, respectively. It shows that multispan noise enhancement cannot be ignored, even when the span loss is as large as 20 dB if the compensation ratio is higher than 50%.

3.5. Comparison Between Single-Polarization and Dual-Polarization Transmission

We compare the performance difference between dual-polarization transmission studied in this paper and single-polarization transmission in [15], assuming the same link configuration, such as of the same transmission fiber, dispersion map, and amplifier noise figure. As shown in Table 1, the FWM noise power density of dual-polarization is 4.26 dB less than that of single-polarization for the same launch power density; the maximum Q of dual-polarization is 0.59 dB less than that of single-polarization; the optimal launch power density is 2.42 dB higher than that of single-polarization; the nonlinear threshold of launch power density is 2.13 dB higher than that of single-polarization. The comparison results of $\boldsymbol{I}_{NL}$, Q_{max}, $\boldsymbol{I}_{Opt}$, and $\boldsymbol{I}_{th}$ are independent of the fiber link configuration. However, the comparison of the spectral efficiency between single- and dual-polarization is link dependent. For the 10 $\times$ 100 SSMF fiber link, the spectral efficiency difference between dual-polarization and single-polarization is about 4.76 b/s/Hz, or the spectral efficiency of dual-polarization is about 7% away from doubling that of single-polarization.

3.6. Applicability to the Densely Spaced Coherent Single-Carrier Systems

Although our derivation is based on densely spaced CO-OFDM systems, we argue that our closed-form results are applicable to some of the densely spaced coherent single-carrier systems.

TABLE 1

Dual-Polarization Transmission Performance Normalized to That of Single-Polarization Transmission. The Information Spectral Efficiency S Is Evaluated for 10 × 100 km Uncompensated SSMF Link With 40 nm OFDM Signal Continuous Bandwidth

FWM Noise Power Density: ΔI_{NL} (dB)	Maximum Q: $\Delta\, Q_{max}$ (dB)	Optimal Power Density: ΔI_{Opt} (dB)	Nonlinear Threshold: ΔI_{th} (dB)	Spectral Efficiency: ΔS (b/s/Hz) or $(\Delta S)/S$
-4.26	-0.59	2.42	2.13	4.76 or 93 %

As explained in [25], in the single-carrier systems, when the symbol rate is much higher than the dispersion walk-off bandwidth f_W, due to rapid walk-off among different frequency components, the performance of the single-carrier systems with large symbol rate approaches that of the multicarrier systems with very small symbol rate. Consequently, we conclude that the closed-form expressions can be extended to the coherent single-carrier systems with high symbol rate. The closed-form expressions cannot be applied to those densely spaced multicarrier systems where the symbol rate is in the neighborhood of the dispersion walk-off bandwidth. As elucidated in [25], these systems have the optimal performance with their Q factor about 1 dB better than the ones with either very low symbol rate or very high symbol rate.

4. Conclusion

In this paper, we have derived closed-form expressions for nonlinear transmission performance of dual-polarization closely spaced CO-OFDM systems. We find that the FWM nonlinear noise exhibits a signature of the flicker noise or $1/f$ noise beyond the corner frequency that is inversely proportional to the total participating bandwidth. We derive a noise enhancement factor that captures the interference effect of nonlinear noises among different spans. The closed-form expressions are also applicable to the closely spaced coherent single-carrier systems where the symbol rate is much larger than the dispersion walk-off bandwidth, and the optical dispersion is uncompensated.

Appendix A
Derivation of the Average FWM Power

Let us denote $\boldsymbol{U} = (U_x, U_y)^T = (c_k^+ c_i)c_j + (c_k^+ c_j)c_i$. Expanding both sides of the denotation into the two polarization components, we have U_x given by

$$U_x = 2\boldsymbol{c}_x^{k*}\boldsymbol{c}_x^i\boldsymbol{c}_x^j + \boldsymbol{c}_y^{k*}\boldsymbol{c}_y^i\boldsymbol{c}_x^j + \boldsymbol{c}_y^{k*}\boldsymbol{c}_y^j\boldsymbol{c}_x^i. \tag{50}$$

The ensemble average of power of $\boldsymbol{U}_x$ is expressed as

$$\left\langle |U_x|^2 \right\rangle = \left\langle \left| 2\boldsymbol{c}_x^{k*}\boldsymbol{c}_x^i\boldsymbol{c}_x^j + \boldsymbol{c}_y^{k*}\boldsymbol{c}_y^i\boldsymbol{c}_x^j + \boldsymbol{c}_y^{k*}\boldsymbol{c}_y^j\boldsymbol{c}_x^i \right|^2 \right\rangle. \tag{51}$$

We assume that the OFDM information symbols on different polarizations or different subcarriers are uncorrelated, namely

$$\left\langle \boldsymbol{c}_{x_1}^{i*}\boldsymbol{c}_{x_2}^k \right\rangle = \frac{\boldsymbol{P}_j}{2}\delta_{x_1x_2}\delta_{jk} \tag{52}$$

where $\boldsymbol{x}_{1,2}$ stands for x or y polarization component, and $\boldsymbol{P}_j$ is the power of the jth subcarrier. Expanding the right side of (51) and substituting (52), we have

$$\left\langle |U_x|^2 \right\rangle = 4\left\langle |\boldsymbol{c}_x^k|^2|\boldsymbol{c}_x^i|^2|\boldsymbol{c}_x^j|^2 \right\rangle + \left\langle \left|\boldsymbol{c}_y^k\right|^2\left|\boldsymbol{c}_y^i\right|^2|\boldsymbol{c}_x^j|^2 \right\rangle + \left\langle \left|\boldsymbol{c}_y^k\right|^2\left|\boldsymbol{c}_y^j\right|^2|\boldsymbol{c}_x^j|^2 \right\rangle = \frac{3}{4}P_iP_jP_k. \tag{53}$$

Similarly, we arrive at average power of U_y given by

$$\left\langle |U_y|^2 \right\rangle = \frac{3}{4} P_i P_j P_k. \tag{54}$$

Subsequently, we obtain the average power of U given by

$$\left\langle |\boldsymbol{U}|^2 \right\rangle = \left\langle |U_x|^2 + |U_y|^2 \right\rangle = \frac{3}{2} P_i P_j P_k. \tag{55}$$

Appendix B
Nonlinearity Performance Including the Loss and Nonlinearity of the DCF

As shown in Fig. 1, each span of the fiber link is consisted of a transmission fiber and a DCF which is sandwiched between two DEFAs. To include the nonlinear contribution from DCF, the FWM product from each span equivalent to (9) is modified as

$$\begin{aligned} \boldsymbol{c}'_{g,t} &= \boldsymbol{c}'_{g,0} G_1^{\frac{1}{2}} e^{-\frac{1}{2}\alpha_1 L_1 - i\beta_{g,1} L_1} + \boldsymbol{c}'_{g,1} \\ \boldsymbol{c}'_{g,m} &= \gamma_m \left[\left(\boldsymbol{c}^+_{k,m} \boldsymbol{c}_{i,m} \right) \boldsymbol{c}_{j,m} + \left(\boldsymbol{c}^+_{k,m} \boldsymbol{c}_{j,m} \right) \boldsymbol{c}_{i,m} \right] e^{-\frac{1}{2}\alpha_m L_m - i\beta_{g,m} L_m} \frac{1 - e^{-\alpha_m L_m} e^{-j\Delta\beta_{ijk,m} L_m}}{j\Delta\beta_{ijk,m} + \alpha_m} \end{aligned} \tag{56}$$

where $\boldsymbol{c}'_{g,t}$ is the total FWM product measured at the output of the DCF of the first span, which consists of the contribution from both transmission fiber and DCF. The subscript "m" of "0" and "1" stands for the parameters associated with the transmission fiber and DCF, respectively. When m is equal to zero, the subscript can be dropped for brevity. For instance, $\boldsymbol{L}_0$ and $\boldsymbol{L}$ are synonymous. $\boldsymbol{c}_{i,m}$ is the field at the input of the transmission fiber $(m = 0)$ or DCF $(m = 1)$, and G_1 is the gain of the EDFA I. From Fig. 1, the field at the input of the DCF is related to that at the input of the transmission fiber by

$$\boldsymbol{c}_{i,j,k,1} = e^{-\frac{1}{2}\alpha L - i\beta_{i,j,k} L} G_1^{\frac{1}{2}} \boldsymbol{c}_{i,j,k}. \tag{57}$$

Substituting (57) into (56), we obtain

$$\begin{aligned} \boldsymbol{c}_{g,t} &= \left[\left(\boldsymbol{c}^+_k \boldsymbol{c}_i \right) \boldsymbol{c}_j + \left(\boldsymbol{c}^+_k \boldsymbol{c}_j \right) \boldsymbol{c}_i \right] e^{-\frac{1}{2}\alpha L - i\beta_g L} e^{-\frac{1}{2}\alpha_1 L_1 - i\beta_{g,1} L_1} G_1^{\frac{1}{2}} \\ &\quad \times \left\{ \gamma \frac{1 - e^{-\alpha L} e^{-j\Delta\beta_{ijk} L}}{j\Delta\beta_{ijk} + \alpha} + \gamma_1 \left(G_1^{\frac{1}{2}} e^{-\frac{1}{2}\alpha L} \right)^2 e^{-i\Delta\beta_{ijk} L} \frac{1 - e^{-\alpha_1 L_1} e^{-j\Delta\beta_{ijk,1} L_1}}{j\Delta\beta_{ijk,1} + \alpha_1} \right\} \\ &= \left[\left(\boldsymbol{c}^+_k \boldsymbol{c}_i \right) \boldsymbol{c}_j + \left(\boldsymbol{c}^+_k \boldsymbol{c}_j \right) \boldsymbol{c}_i \right] e^{-\frac{1}{2}\alpha L - i\beta_g L} e^{-\frac{1}{2}\alpha_1 L_1 - i\beta_{g,1} L_1} \left\{ \frac{\gamma}{j\Delta\beta_{ijk} + \alpha} + \frac{\gamma_1 \left(G_1^{\frac{1}{2}} e^{-\frac{1}{2}\alpha L} \right)^2 e^{-i\Delta\beta_{ijk} L}}{j\Delta\beta_{ijk,1} + \alpha_1} \right\}. \end{aligned} \tag{58}$$

We have assumed that $e^{-\alpha L} \gg 1$ and $e^{-\alpha_1 L_1} \gg 1$. The former is generally true which indicates the loss of the transmission fiber is much larger than 1. The later also requires the large loss of the DCF, which may not be true. However, this can be justified because when the DCF is short and its loss is low, the nonlinearity contribution becomes insignificant, and subsequently, the inaccuracy due to this approximation is also insignificant.

Similar to (10), the FWM product after the N_s-span transmission can be considered as the superposition of the contribution from N_s individual spans given by

$$\boldsymbol{c}'_{g,N} = \boldsymbol{c}'_{g,t} \sum_{M=0}^{N_s-1} \exp(jM \cdot \Delta\tilde{\beta}_{ijk}) = c'_{g,t} \frac{1-\exp(jN_s \cdot \Delta\tilde{\beta}_{ijk})}{1-\exp(j\Delta\tilde{\beta}_{ijk})}. \tag{59}$$

The power of the FWM after N_s spans of transmission is therefore

$$\begin{aligned} \left\langle P'_g \right\rangle &= \left\langle \left| \boldsymbol{c}'_{g,t} \frac{1-\exp(jN_s \cdot \Delta\tilde{\beta}_{ijk})}{1-\exp(j\Delta\tilde{\beta}_{ijk})} \right|^2 \right\rangle \\ &= \left\langle \left| (\boldsymbol{c}_k^+ \boldsymbol{c}_i) c_j + (\boldsymbol{c}_k^+ \boldsymbol{c}_j) c_i \right|^2 \right\rangle G_1 e^{-\alpha L - \alpha_1 L_1} \left\{ \frac{\gamma^2}{|\Delta\beta_{ijk}|^2 + \alpha^2} + \frac{\gamma_1^2 (G_1 e^{-\alpha L})^2}{|\Delta\beta_{ijk,1}|^2 + \alpha_1^2} \right\}. \end{aligned} \tag{60}$$

We have dropped a term proportional to $\cos(\Delta\beta_{ijk} L)$, which is rapidly varying as a function of the subcarrier frequency for practical systems. At end of each span as shown in Fig. 1, the FWM product P'_g along with the signal will be amplified further by a gain of G_2 equal to $e^{\alpha L + \alpha_1 L_1} G_1^{-1}$, and the FWM product becomes

$$\langle P_g \rangle = G_2 \left\langle P'_g \right\rangle = \frac{3}{2} \gamma^2 P_i P_j P_k \eta_2 \left(\frac{\gamma^2}{|\Delta\beta_{ijk}|^2 + \alpha^2} + \frac{\gamma_1^2 (G_1 e^{-\alpha L})^2}{|\Delta\beta_{ijk,1}|^2 + \alpha_1^2} \right). \tag{61}$$

Following the same procedures from (17)–(35), the closed-form expression for the nonlinear noise power density I_{NL} becomes

$$\begin{aligned} I_{NL} &= \left(\rho(\gamma, \alpha, \beta_2) + \rho(\gamma_1, \alpha_1, \beta_{2,1}) (G_1 e^{-\alpha L})^2 \right) N_s h_e I^3 \\ \rho(\gamma, \alpha, \beta_2) &= \frac{3\gamma^2 \ln(B/B_0)}{8\pi\alpha|\beta_2|}, \qquad B_0 = \frac{\alpha}{\pi|\beta_2|B}. \end{aligned} \tag{62}$$

It can be seen that the difference between nonlinear noise power density $\boldsymbol{I}_{NL}$ in (35) and (62) is the additional term of $\rho(\gamma_1, \alpha_1, \beta_{2,1})(G_1 e^{-\alpha L})^2$, which accounts for FWM contribution from the DCF. The nonlinear noise power density $\boldsymbol{I}_{NL}$ can be expressed in terms of characteristic power density $\boldsymbol{I}_0$, namely

$$I_{NL} = \left(\frac{I}{I_0} \right)^2 I, \qquad I_0 \equiv \sqrt{\frac{1}{\left(\rho(\gamma, \alpha, \beta_2) + \rho(\gamma_1, \alpha_1, \beta_{2,1}) (G_1 e^{-\alpha L})^2 \right) N_s h_e}}. \tag{63}$$

In addition to nonlinearity, the second effect of the DCF is the additional optical ASE noise due to using two EDFAs to compensate its loss. The two EDFAs with a DCF in between can be considered as a composite amplifier with a gain of $G_1 G_2 e^{-\alpha_1 L_1}$ and noise figure NF_c of

$$NF_c = NF + \frac{NF}{G_1 e^{-\alpha_1 L_1}} \tag{64}$$

where NF is the noise figure of the EDFAs I and II, which are assumed to be equal. Similar to (40), the ASE noise after N_s spans of transmission is given by

$$n_0 = 0.5 N_s e^{\alpha L} h\upsilon \cdot NF_c. \tag{65}$$

Using the new expressions for I_0 of (63) and n_0 of (65), the information spectral efficiency S of (43), system Q factor, and its optimal value of (45) and (47), optimal launch power density in (46) remains valid. For instance, the optimal Q factor can be expressed as

$$Q_{max} = \frac{1}{3}\left(\frac{I_0}{n_0}\right)^{2/3} = \frac{1}{3}\left(\frac{1}{\left[0.5N_s e^{\alpha L} h\upsilon \cdot \left(NF + \frac{NF}{G_1 e^{-\alpha_1 L_1}}\right)\right]^2 \left(\rho(\gamma,\alpha,\beta_2) + \rho(\gamma_1,\alpha_1,\beta_{2,1})(G_1 e^{-\alpha L})^2\right) N_s h_e}\right)^{1/3}. \quad (66)$$

References

[1] W. Shieh and C. Athaudage, "Coherent optical orthogonal frequency division multiplexing," *Electron. Lett.*, vol. 42, no. 10, pp. 587–589, May 2006.

[2] E. Yamada, A. Sano, H. Masuda, E. Yamazaki, T. Kobayashi, E. Yoshida, K. Yonenaga, Y. Miyamoto, K. Ishihara, Y. Takatori, T. Yamada, and H. Yamazaki, "1 Tb/s (111 Gb/s/ch × 10 ch) no-guard-interval CO-OFDM transmission over 2100 km DSF," presented at the Opto-Electronics Commun. Conf./Australian Conf. Optical Fiber Technol., Sydney, Australia, 2008, Paper PDP6.

[3] Y. Ma, Q. Yang, Y. Tang, S. Chen, and W. Shieh, "1-Tb/s per channel coherent optical OFDM transmission with subwavelength bandwidth access," presented at the Optical Fiber Commun. Conf., San Diego, CA, 2009, Paper PDP C1.

[4] S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval coherent OFDM superchannel over 7200-km of ultra-large-area fiber," presented at the European Conf. Optical Commun., Vienna, Austria, 2009, Post-deadline Paper PD2.6.

[5] G. Goldfarb, G. F. Li, and M. G. Taylor, "Orthogonal wavelength-division multiplexing using coherent detection," *IEEE Photon. Technol. Lett.*, vol. 19, no. 24, pp. 2015–2017, Dec. 2007.

[6] R. Dischler and F. Buchali, "Transmission of 1.2 Tb/s continuous waveband PDM-OFDM-FDM signal with spectral efficiency of 3.3 bit/s/Hz over 400 km of SSMF," presented at the Optical Fiber Commun. Conf., San Diego, CA, 2009, paper PDP C2.

[7] H. Takahashi, A. Al Amin, S. L. Jansen, I. Morita, and H. Tanaka, "Highly spectrally efficient DWDM transmission at 7.0 b/s/Hz using 8 × 65.1-Gb/s coherent PDM-OFDM," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 406–414, Feb. 2010.

[8] A. J. Lowery, S. Wang, and M. Premaratne, "Calculation of power limit due to fiber nonlinearity in optical OFDM systems," *Opt. Express*, vol. 15, no. 20, pp. 13 282–13 287, Oct. 2007.

[9] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. Cho, R. Noe, I. Shpantzer, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Express*, vol. 16, no. 20, pp. 15 777–15 810, Sep. 2008.

[10] M. Mayrock and H. Haunstein, "Monitoring of linear and nonlinear signal distortion in coherent optical OFDM transmission," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3560–3566, Aug. 2009.

[11] X. Liu, F. Buchali, and R. W. Tkach, "Improving the nonlinear tolerance of polarization-division-multiplexed CO-OFDM in long-haul fiber transmission," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3632–3640, Aug. 2009.

[12] Y. Tang and W. Shieh, "Coherent optical OFDM transmission up to 1 Tb/s per channel," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3511–3517, Aug. 2009.

[13] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, no. 6841, pp. 1027–1030, Jun. 2001.

[14] J. Tang, "The channel capacity of a multispan DWDM system employing dispersive nonlinear optical fibers and an ideal coherent optical receiver," *J. Lightwave Technol.*, vol. 20, no. 7, pp. 1095–1101, Jul. 2002.

[15] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Express*, vol. 18, no. 18, pp. 19 039–19 054, Aug. 2010.

[16] P. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 14, no. 2, pp. 148–157, Feb. 1996.

[17] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[18] N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Opt. Commun.*, vol. 142, no. 1–3, pp. 119–125, Feb. 1997.

[19] W. Shieh and I. Djordjevic, *OFDM for Optical Communications*. New York: Academic, 2010.

[20] 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.

[21] R. W. Tkach, A. R. Chraplyvy, F. Forghieri, A. H. Gnauck, and R. M. Derosier, "Four-photon mixing and high-speed WDM systems," *J. Lightwave Technol.*, vol. 13, no. 5, pp. 841–849, May 1995.
[22] K. Inoue, "Polarization effect on four-wave mixing efficiency in a single-mode fiber," *IEEE J. Quantum Electron.*, vol. 28, no. 4, pp. 883–894, Apr. 1992.
[23] X. Liu and F. Buchali, "Intra-symbol frequency-domain averaging based channel estimation for coherent optical OFDM," *Opt. Express*, vol. 16, no. 26, pp. 21 944–21 957, Dec. 2008.
[24] C. E. Shannon, "A mathematical theory of communication," *Bell Syst.Tech. J.*, vol. 27, pp. 379–423, Jul.–Oct. 1948.
[25] Y. Tang, W. Shieh, and B. S. Krongold, "DFT-spread OFDM for fiber nonlinearity mitigation," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1250–1252, Aug. 2010.