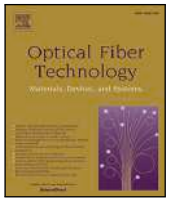
Optical Fiber Technology 75 (2023) 103187

Contents lists available at ScienceDirect

Optical Fiber Technology

journal homepage: www.elsevier.com/locate/yofte



Improving MDM–WDM optical network performance via optimized power allocation using Gaussian noise model

M.A. Amirabadi [b], S.A. Nezamalhosseini [b,*], M.H. Kahaei [b], Lawrence R. Chen [a,b]

[a] Electrical and Computer Engineering, McGill University, Canada
[b] School of Electrical Engineering, Iran University of Science and Technology (IUST), Tehran, 1684613114, Iran

ARTICLE INFO

Keywords:
Few-mode fiber
Gaussian noise model
Power allocation
Mode division multiplexing
Wavelength division multiplexing

ABSTRACT

The few-mode fiber nonlinearity has been well modeled by an additive Gaussian Noise (GN) source. The available GN models in Mode Division Multiplexing–Wavelength Division Multiplexing (MDM–WDM) systems are incoherent and neglect the dispersion slope. The incoherent GN model assumes that the Nonlinear Interference (NLI) noise created at each span is summed up in power incoherently at the receiver and leads to underestimating the NLI noise power. In the first part of this paper, a coherent GN model is derived by taking into account the dispersion slope and verified through Split-Step Fourier Method (SSFM) simulations. In the second part of this paper, the total capacity maximization, as well as the minimum Signal to Noise Ratio (SNR) margin maximization problems are presented and solved based on the obtained GN model. MDM–WDM multi-node nonlinear network under different scenarios including Equal/Non-Equal Required SNR (ERS/NERS), and Flat/Non-Flat Amplifier Gain (FG/NFG) are considered in the numerical simulations. Considering 200 different sets of randomized traffic demands (random channel-mode distributions) the cumulative distribution function of minimum SNR margin improvement in optimized power allocation compared with the best flat power allocation is in 99% more than 1 dB, 1.2 dB, 2.4 dB, and 2.8 dB, for ERS-FG, ERS-NFG, NERS-FG, and NERS-NFG scenarios, respectively.

1. Introduction

The rapid spread of internet services led to the increasing capacity demand in long-haul optical fiber communication systems. Several techniques including Wavelength Division Multiplexing (WDM), polarization division multiplexing, and coherent modulations have been proposed in Single-Mode Fiber (SMF) to increase the optical fiber communication data rate [1]. However, the transmission capacity of SMFs is now reaching the theoretical limits confined by nonlinear effects [2,3]. Recently, Space Division Multiplexing (SDM) techniques in Few-Mode Fiber (FMF) have been proposed to overcome this capacity limit [4]. In the aforementioned systems, the combination of Mode Division Multiplexing (MDM) with WDM which is respectively the multiplexing of signals with different spatial modes and different channels has been developed to dramatically increase the capacity of optical fiber communication systems [4].

The MDM–WDM system suffers from both linear and nonlinear effects of FMF. The FMF linear effects include attenuation, modal dispersion, chromatic dispersion, and linear coupling. For compensating FMF linear effects in the weak coupling, each mode is separately processed without Multiple-Input-Multiple-Output (MIMO) Digital Signal Processing (DSP) [5]. However, in the strong coupling, a MIMO DSP is required for compensating FMF linear effects [6]. The MIMO DSP complexity can be handled only in quiet small differential mode group delays between propagating modes [7,8]. The FMF nonlinear effects are the fundamental limitation factors of the MDM–WDM system which include Kerr nonlinearity and nonlinear coupling.

1.1. Related works

During the last decade, several investigations analyzed the FMF nonlinear effects by mainly focusing on numerical simulations [9] and analytical predictions [10] combined with experimental verification [11]. To analyze FMF nonlinear effects, the Manakov equation should be solved, the perturbation-based methods are the most well-known techniques for this aim [12]. The Gaussian Noise (GN) model is the most practical perturbation-based model which describes the nonlinear effects by an additive Gaussian noise [13]. The available GN model formulations for MDM–WDM system can be divided in two

* Corresponding author.
E-mail addresses: m_amirabadi@elec.iust.ac.ir (M.A. Amirabadi), nezam@iust.ac.ir (S.A. Nezamalhosseini), kahaei@iust.ac.ir (M.H. Kahaei), lawrence.chen@mcgill.ca (L.R. Chen).

https://doi.org/10.1016/j.yofte.2022.103187
Received 20 February 2022; Received in revised form 10 October 2022; Accepted 27 November 2022

categories including; integral-form [4,14–17] and closed-form [18–20]. Note that these GN model formulations are obtained based on incoherency assumption wherein the Nonlinear Interference (NLI) noise created at each span is incoherently accumulated at the receiver. The Incoherent GN (IGN) model leads to underestimating the NLI noise Power Spectral Density (PSD) [21]. Moreover, only first-order and second-order dispersion terms are considered in these models. The effects of third-order dispersion may be observed in MDM–WDM systems operating over a broad wavelength range. The dispersion slope has been included in the GN model in [21–25] for SMF–WDM systems to improve modeling accuracy, especially in cases involving large bandwidth. In a recent work, we compared the complexity-performance of the closed-form IGN model, integral-form IGN model, and SSFM for MDM–WDM systems [20]. Although the closed-form IGN model is very fast, it can only be applied to rectangular shaped Nyquist WDM signals with the channel spacing close to the symbol-rate [26–28].

The performance of the MDM–WDM system can be evaluated by using the minimum SNR margin and the total capacity which are dependent on SNR that includes NLI noise contribution. The NLI noise is related to system parameters such as launched power per channel and mode, therefore, power allocation [29–32] has an important role in optimizing performance of MDM–WDM systems and networks.

1.2. Novelties and contributions

Previous investigations on FMF nonlinearity resulted in the development of closed-form and integral-form IGN models. Despite the improvements provided by coherent GN model for simulating and analyzing SMF systems, no such formulation exists for FMF systems. Therefore, in the first part of this paper, we derive a coherent GN model formulation for the MDM–WDM system by taking into account the first, second, and third-order dispersion terms. The significance of this paper in first part include;

- Presenting a coherent GN model for MDM–WDM system, showing proposed coherent GN model provides better accuracy by removing the 1 dB and 2 dB SNR margin of IGN model presented by [4] respectively in weak and strong coupling at the optimal (best equal) launched power per channel-mode. This is very important in different applications such as resource allocation, optical performance monitoring, and quality of transmission estimation in unestablished lightpaths.
- Considering dispersion slope in GN model formulation for MDM–WDM system as an extension of previous work in [21] for SMF, describing the impact of dispersion slope induced inter-modal phase-matching on NLI noise in MDM–WDM systems, providing a normalized parameter beyond which the dispersion slope should be considered.
- Presenting a clear relation between NLI noise and launched power of different channels-modes in the proposed coherent GN model formulation, making it proper for power allocation in link, lightpath and network level.

The above mentioned researches apply convex optimization algorithms for power allocation in SMF links and networks. Recently, we solved a power allocation in FMF links using convex optimization [20]. However, the optimized power allocation in FMF networks is not investigated yet. Therefore, in the second part of this paper, based on the proposed GN model formulation, the optimized power allocation by minimum SNR margin maximization and total capacity maximization is investigated considering an MDM–WDM multi-node nonlinear network structure. Considering different practical scenarios, the obtained results show remarkable minimum SNR margin improvement, which is very important in the practical implementation of MDM–WDM networks. Therefore, the significance of the second part of this paper include;

- Deploying power allocation in MDM–WDM network by maximizing total capacity and minimum SNR margins by the proposed coherent GN model in DTG network.
- Providing comprehensive investigation over optimized power allocation for different MDM–WDM network scenarios, showing minimum SNR margin improvements are 99% more than 1 dB, 1.2 dB, 2.4 dB and 2.8 dB for flat/non-flat EDFA gain with equal/non-equal modulation format scenarios, respectively.

The rest of this paper is organized as follows. Section 2 describes the link topology and terminology, signal model, and Four-Wave Mixing (FWM) in MDM–WDM systems. Section 3 presents the proposed GN model formulation. The problem statement is given in Section 4. Section 5 presents the simulation results, and Section 6 concludes the paper.

2. Signal propagation in MDM-WDM system

2.1. Link topology and terminology

The considered MDM–WDM system is depicted in Fig. 1 [33]. In this structure, the input data carries N_{ch} WDM channels which each channel is the multiplexing of D spatial modes, and each mode is the multiplexing of 2 polarization modes. The generated signal is transmitted through an MDM–WDM link with N_s spans, each with length L_s. At the end of each span, the signal is amplified by an Erbium-Doped Fiber Amplifier (EDFA), and in turn, the Amplified Spontaneous Emission (ASE) noise is produced. The considered MDM–WDM link is assumed to be dispersion uncompensated to reduce FMF nonlinear effects [4]. The Kerr nonlinearity is generated by inter and intra channel interactions, as well as inter and intra mode interactions [4]. We consider both weak and strong linear coupling among spatial modes to be applicable in both short and long-range links [9,34]. Furthermore, a MIMO processor at the receiver is used to compensate FMF linear distortions [6]. However, no processing is used for nonlinearity compensation.

2.2. Signal model

The Linearly-Polarized (LP) spatial modal fields are considered in a transverse plane, each field contains two polarizations. The launched electrical field into the FMF is presented by a sum of spatial modes as [9]

$$\tilde{\boldsymbol{E}}(x,y,z,f)=\sum_{m=1}^{D} e^{j\beta_m(f)z}\tilde{\boldsymbol{A}}_m(z,f)F_m(x,y)/\sqrt{N_m}, \tag{1}$$

where $\tilde{(\cdot)}$ represents the frequency domain of parameters and for the mth mode $\tilde{\boldsymbol{A}}_m(z,f)=[\tilde{A}_{m,x}(z,f)\ \tilde{A}_{m,y}(z,f)]$ shows the slowly-varying amplitude of x and y polarizations, $F_m(x,y)$ is the spatial distribution, $\beta_m(f)$ is the propagation constant, $N_m=0.5\epsilon_0 n_{eff,m}cI_m$ is the mode power with $I_m=\iint F_m^2(x,y)dxdy$, ϵ_0 shows the vacuum permittivity, c is the light speed, and $n_{eff,m}$ denotes the fiber effective refractive index. The application of GN model analysis is along with some important assumptions including the Gaussian distribution for the transmitted signal, the statistical independence of the NLI noise from both the ASE noise and the transmitted signal [35]. Accordingly, a complex periodic process is chosen for the amplitude of transmitted signal which is spectrally shaped to satisfy the above assumptions [35]. The amplitude of transmitted signal has the following frequency domain descriptions in x and y polarizations of mth mode [18]

$$\begin{aligned}\tilde{A}_{(m,x)}(0,f)&=\sqrt{f_0}\sum_{n=-\infty}^{\infty}\xi_{n,m}\sqrt{G_{Tx}^{(m,x)}(nf_0)}\delta(f-nf_0)\\ \tilde{A}_{(m,y)}(0,f)&=\sqrt{f_0}\sum_{n=-\infty}^{\infty}\zeta_{n,m}\sqrt{G_{Tx}^{(m,y)}(nf_0)}\delta(f-nf_0),\end{aligned} \tag{2}$$

where f_0 is the frequency separation between any two successive frequency-components, $\xi_{n,m}$, and $\zeta_{n,m}$ are random variables of mth mode

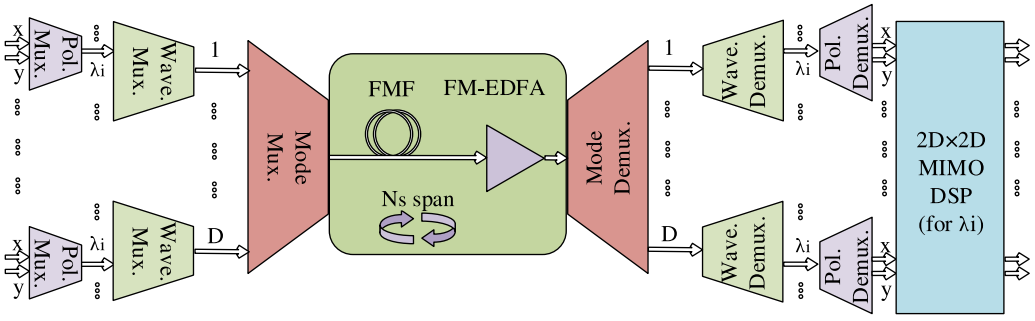


Fig. 1. Schematic diagram of MDM–WDM system.

and frequency nf_0, with zero mean and unit variance, $G_{Tx}^{(m,x)}(nf_0)$ and $G_{Tx}^{(m,y)}(nf_0)$ are the PSDs of the transmitted signal of x and y polarizations of mth mode, respectively. Note that the PSD of the transmitted signal is continuous in general. However, the GN model is an FWM-like model which is based on ideally splitting up the transmitted signal spectrum into spectral components with f_0 spacing whose nonlinear interactions during the propagation are then analytically explored in an approach similar to the FWM formulas [35].

The amplitude of transmitted signal has the following time domain descriptions in x and y polarizations of mth mode

$$A_{(m,x)}(0,t) = f_0^{3/2} \sum_{n=-\infty}^{\infty} \xi_{n,m} \sqrt{G_{Tx}^{(m,x)}(nf_0)} e^{j2\pi n f_0 t}$$
$$A_{(m,y)}(0,t) = f_0^{3/2} \sum_{n=-\infty}^{\infty} \zeta_{n,m} \sqrt{G_{Tx}^{(m,y)}(nf_0)} e^{j2\pi n f_0 t}. \tag{3}$$

2.3. FWM in MDM-WDM systems

The Kerr effect in MDM–WDM systems includes Self Phase Modulation (SPM), Cross Phase Modulation (XPM), and FWM which produce intra/inter channel as well as intra/inter modal nonlinear interactions in frequency and spatial domains, respectively. The FWM in MDM–WDM systems includes the nonlinear interactions between different frequencies-modes, and results in an energy transfer into an idler frequency-mode. The relation between the interacting frequencies-modes is defined through the so-called frequency condition as [2]

$$[nf_0]_m = [kf_0]_q - [lf_0]_s + [hf_0]_t, \tag{4}$$

where m, q, s, t and n, k, l, h are mode and channel indices, respectively. Moreover, the relation between the phase of the interacting frequencies-modes is defined through the so-called phase matching condition as [2]

$$\Delta\beta_{mqst}(nf_0, kf_0, lf_0, hf_0) = \beta_m(nf_0) - \beta_q(kf_0) - \beta_s(lf_0) + \beta_t(hf_0), \tag{5}$$

where $\beta_m(nf_0)$ is the propagation constant. For the further analysis, we expand the propagation constant into a Taylor series as [2]

$$\beta_m(nf_0) = \beta_{0m} + \beta_{1m}(2\pi n f_0) + \frac{\beta_{2m}}{2}(2\pi n f_0)^2 + \frac{\beta_{3m}}{6}(2\pi n f_0)^3, \tag{6}$$

where β_{im} is the ith order dispersion coefficient.

3. GN model formulation

To describe the signal propagation in the MDM–WDM system, the Manakov equation is used which includes attenuation, dispersion, Kerr nonlinearity, and nonlinear coupling, and can be expressed as [9,34]

$$\frac{\partial A_{(m,x)}(z,t)}{\partial z} = -\frac{\alpha_m}{2} A_{(m,x)}(z,t) + j\beta_{0m} A_{(m,x)}(z,t) - \beta_{1m} \frac{\partial A_{(m,x)}(z,t)}{\partial t}$$
$$- j\frac{\beta_{2m}}{2} \frac{\partial^2 A_{(m,x)}(z,t)}{\partial t^2} - \frac{\beta_{3m}}{6} \frac{\partial^3 A_{(m,x)}(z,t)}{\partial t^3} +$$
$$j \sum_{q=1}^{D} \tilde{\gamma}_{mq} |A_q(z,t)|^2 A_{(m,x)}(z,t), \tag{7}$$

where $|A_q(z,t)|^2 = A_{(q,x)}^2(z,t) + A_{(q,y)}^2(z,t)$, α_m is the attenuation of mth mode,

$$\tilde{\gamma}_{mq} = \begin{cases} \frac{4}{3}(\frac{2}{3})^{\delta_{mq}} f_{mq}\gamma & weak\ coupling \\ \kappa\gamma & strong\ coupling, \end{cases} \tag{8}$$

with

$$\kappa = \sum_{\substack{q \le m \\ m,q \in \{1,2,\ldots,D\}}} \frac{32}{2^{\delta_{mq}}} \frac{f_{mq}}{6D(2D+1)}, \tag{9}$$

where γ is the Kerr nonlinearity coefficient, $f_{mq} = \frac{A_{eff}}{I_m I_q} \iint F_m^2(x,y) F_q^2(x,y) dxdy$ is the nonlinear coupling coefficient between modes m and q, and A_{eff} is the effective area of the fundamental mode. The GN model is based on the first-order perturbation which expresses the solution of the Manakov equation as [36]

$$A_{(m,x)}(z,t) \simeq e^{\mathcal{L}z} A_{(m,x)}(0,t) + \int_0^z e^{\mathcal{L}(z-\xi)} \mathcal{N}\left(e^{\mathcal{L}z} A_{(m,x)}(0,t) \right) \partial\xi, \tag{10}$$

where $\mathcal{L}$ and $\mathcal{N}$ respectively account for linear and nonlinear propagation parts of the Manakov equation. $\mathcal{N}$ can be represented in the time domain as

$$\mathcal{N} \triangleq -j \sum_{q=1}^{D} \tilde{\gamma}_{mq} |A_q(z,t)|^2 A_m(z,t). \tag{11}$$

The Fourier transform of $\mathcal{L}$ is described as $\mathcal{F}(e^{\mathcal{L}z}) \triangleq e^{\vartheta(z,f)}$ that can be expressed as [36]

$$\vartheta(z,f) = -\int_0^z \left(\frac{\alpha_m(\xi)}{2} + j\beta_m(f,\xi) \right) \partial\xi. \tag{12}$$

By mitigating the dispersion effects at the receiver (10) can be expressed as

$$A_{(m,x)}(z,t) = A_{(m,x)}(0,t) + \int_0^z e^{-\mathcal{L}\xi} \mathcal{N}\left(e^{\mathcal{L}z} A_{(m,x)}(0,t) \right) \partial\xi, \tag{13}$$

where the first term is the transmitted signal and the second term is the NLI noise of x polarization of mth mode. Considering the second term of (13), and by using (3), (11), and (12), the NLI noise of x polarization of mth mode can be expressed as

$$n_{m,x}(z,t) = j f_0^{3/2} \sum_{q=1}^{D} \sum_{k,l,h} \tilde{\gamma}_{mq} \eta(kf_0, lf_0, hf_0) e^{2\pi(k+l-h)f_0 t} \xi_{k,q} \sqrt{G_{Tx}^{(m,x)}(kf_0)}$$
$$\times \left(\xi_{l,q}^* \xi_{h,q} \sqrt{G_{Tx}^{(q,x)}(lf_0) G_{Tx}^{(q,x)}(hf_0)} + \right.$$
$$\left. \zeta_{l,q}^* \zeta_{h,q} \sqrt{G_{Tx}^{(q,y)}(lf_0) G_{Tx}^{(q,y)}(hf_0)} \right) \tag{14}$$

where

$$\eta(kf_0, lf_0, hf_0) = \int_0^z e^{\vartheta(\xi,kf_0) + \vartheta(\xi,lf_0) + \vartheta^*(\xi,(k+l-h)f_0) - \vartheta(\xi,hf_0)} \partial\xi. \tag{15}$$

In the MDM–WDM link, the attenuation and dispersion parameters can be assumed to be constant in z [4]. Therefore, (15) can be expressed as

$$\eta(kf_0, lf_0, hf_0) = \frac{1-e^{-\alpha_m L_s} e^{j|\beta_q(kf_0+lf_0-hf_0)-\beta_m(kf_0)+\beta_m(hf_0)-\beta_q(lf_0)|L_s}}{\alpha_m - j|\beta_q(kf_0+lf_0-hf_0)-\beta_m(kf_0)+\beta_m(hf_0)-\beta_q(lf_0)|} \times \frac{1-e^{j|\beta_q(kf_0+lf_0-hf_0)-\beta_m(kf_0)+\beta_m(hf_0)-\beta_q(lf_0)|N_s L_s}}{1-e^{j|\beta_q(kf_0+lf_0-hf_0)-\beta_m(kf_0)+\beta_m(hf_0)-\beta_q(lf_0)|L_s}}. \quad (16)$$

Considering (14), after some mathematical manipulations as described by Appendix A, the variance of NLI noise of nth channel and mth mode takes the following form

$$\sigma^2_{GN_{m,n}} = \sum_{q=1}^{D} \sum_{n_1,n_2,n_3=1}^{N_{ch}} P_{n_1,m} P_{n_2,q} P_{n_3,q} D_{n,m}(n_1,n_2,n_3,q), \quad (17)$$

where

$$D_{n,m}(n_1,n_2,n_3,q) = 3/4 \int_{f_{c_n}-B_n/2}^{f_{c_n}+B_n/2} \int_{f_{c_1}-B_{n_1}/2}^{f_{c_1}+B_{n_1}/2} \int_{f_{c_2}-B_{n_2}/2}^{f_{c_2}+B_{n_2}/2} \tilde{\gamma}^2_{mq} |\eta(f_1,f_2,f)|^2 g_{n_1,m}(f_1) g_{n_2,q}(f_2) g_{n_3,q}(f_1+f_2-f) R_n g^*_{n,m}(f) df_1 df_2 df. \quad (18)$$

4. Problem statement

The MDM–WDM system performance can be evaluated using the minimum SNR margin and the total capacity criteria which are dependent on SNR that in turn includes both ASE and the NLI noise contributions. In this section, optimized power allocation based on the aforementioned criteria is obtained.

4.1. Minimum SNR margin maximization

The SNR of the nth channel and the mth mode can be expressed as [16]

$$SNR_{n,m} = \frac{P_{n,m}}{\sigma^2_{ASE} + \sigma^2_{GN,n,m}}, \quad (19)$$

where $\sigma^2_{ASE} = N_s F(G-1)h\nu B_n$ is the variance of ASE noise, F is the amplifier noise figure, G is the amplifier gain compensating loss in each fiber span, h is Planck constant, ν is the central frequency, and B_n is bandwidth of nth channel. By substituting (17) into (19), the $SNR_{n,m}$ can be rewritten as

$$SNR_{n,m} = \frac{P_{n,m}}{N_s F(G-1)h\nu B_n + \sum_{q=1}^{D} \sum_{n_1,n_2,n_3=1}^{N_{ch}} P_{n_1,m} P_{n_2,q} P_{n_3,q} D_{n,m}(n_1,n_2,n_3,q)}. \quad (20)$$

The SNR margin of the nth channel and the mth mode can be defined as

$$M_{n,m} = \frac{SNR_{n,m}}{SNR^{req}_{n,m}}, \quad (21)$$

where $SNR^{req}_{n,m}$ is the required SNR of the nth channel and the mth mode. Therefore, the minimum SNR margin maximization problem can be expressed as

$$\max_{P_{n,m}} \min_{n,m} \frac{SNR_{n,m}}{SNR^{req}_{n,m}}, \quad (22)$$

which is equivalent to

$$\min_{P_{n,m}} \max_{n,m} \frac{SNR^{req}_{n,m}}{SNR_{n,m}}. \quad (23)$$

Considering SNR formulation of (20), (23) is a non-convex optimization problem. Dealing with non-convex optimization problems has many barriers such as complexity and converging to local optimum. To tackle this problem, by using the variable $\hat{P}_{n,m} \triangleq log(P_{n,m})$ in (20) and noting that $log(x)$ is a monotonic function in x, the following formulation is obtained with the same minimum as (23),

$$\min_{\hat{P}_{n,m}} \max_{n,m} \left[log\left(SNR^{req}_{n,m}\right) + log\left(N_s F(G-1)h\nu B_n + \sum_{q=1}^{D} \sum_{n_1,n_2,n_3=1}^{N_{ch}} P_{n_1,m} P_{n_2,q} P_{n_3,q} D_{n,m}(n_1,n_2,n_3,q)\right) - \hat{P}_{n,m} \right]. \quad (24)$$

We define the slack variable β and rewrite (24) as

$$\min_{\beta,\hat{P}_{n,m}} \beta \quad s.t.\ log(SNR^{req}_{n,m}) + log\left(N_s F(G-1)h\nu B_n + \sum_{q=1}^{D} \sum_{n_1,n_2,n_3=1}^{N_{ch}} e^{\hat{P}_{n_1,m}+\hat{P}_{n_2,q}+\hat{P}_{n_3,q}} D_{n,m}(n_1,n_2,n_3,q)\right) - \hat{P}_{n,m} \le \beta. \quad (25)$$

To solve (25), we use the gradient descent algorithm in vector form. This is performed by introducing a vector $\mathbf{p}$ of dimension DN_{ch} whose elements $P_l; l = 1,2,\ldots,DN_{ch}$ are given by $P_{n,m}$, $n = 1,2,\ldots,N_{ch}$, $m = 1,2,\ldots,D$. In order to incorporate the values of B_n, we use a vector with the same dimension as $\mathbf{p}$ defined by $\boldsymbol{B} = [B_1, B_1, \ldots, B_1, B_2, B_2, \ldots, B_2, \ldots, B_{N_{ch}}, B_{N_{ch}}, \ldots, B_{N_{ch}}]$ in which each B_n has been repeated D times. In accordance with this change, the $D_{n,m}(n_1,n_2,n_3,q)$ changes to $D_l(l_1,l_2,l_3)$ in different subscripts. It should be noted that the values are equal and do not change. Then, by deploying the values of $\mathbf{p}$ and $\boldsymbol{B}$ in (25), we get the following formulation which is a convex optimization problem (see Appendix B),

$$\min_{\beta,\hat{P}_l} \beta \quad s.t.\ log(SNR^{req}_l) + log\left(N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3)\right) - \hat{P}_l \le \beta, \quad (26)$$

Many methods can be used for solving constrained convex optimization problems. The barrier method has been used for maximizing the minimum SNR margin in literature [37]. Another promising solution is the Bisection method [38,39]. This method is used in Appendix C for solving the proposed optimization problem (26) since it achieves convergence using fewer iterations without sacrificing numerical accuracy.

4.2. Total capacity maximization

The total capacity of an MDM–WDM system is the sum of the capacity of each channel and each mode which can be defined as [37]

$$C = 2 \sum_{m=1}^{D} \sum_{n=1}^{N_{ch}} R_n log_2\left(1 + SNR_{n,m}\right), \quad (27)$$

which by using (20), can be expressed as

$$C = 2 \sum_{m=1}^{D} \sum_{n=1}^{N_{ch}} R_n log_2 \times \left(1 + \frac{P_{n,m}}{N_s F(G-1)h\nu B_n + \sum_{q=1}^{D} \sum_{n_1,n_2,n_3=1}^{N_{ch}} P_{n,m_1} P_{n,m_2} P_{n,m_3} D_n(n_1,n_2,n_3)}\right). \quad (28)$$

To use the gradient ascent method for solving (28), n,m should be substituted by l, for the same reasons as discussed before. Therefore,

Table 1
Nonlinear coupling coefficient (γf_{pq}(1/W/km)) [41].

mq	LP01	LP11a	LP11b	LP02
LP01	0.73	0.36	0.36	0.36
LP11a	0.36	0.18	0.55	0.18
LP11b	0.36	0.18	0.18	0.55
LP02	0.36	0.36	0.18	0.18

Table 2
Attenuation (α_p (dB/km)), and dispersion terms (β_{1_p}(ps/km), β_{2_p} (ps^2/km), and β_{3_p} (ps^3/km)) [41].

	LP01	LP11a	LP11b	LP02
α_p	0.2	0.2	0.2	0.2
β_{1_p}	−0.29	−0.66	−0.66	−2.93
β_{2_p}	28.27	26.96	26.96	27.47
β_{3_p}	0.1452	0.1452	0.1452	0.1452

(28) can be rewritten as

$$C = 2\sum_{l=1}^{DN_{ch}} R_l log_2\left(1 + \frac{P_l}{N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} P_{l_1}P_{l_2}P_{l_3}D_l(l_1,l_2,l_3)}\right). \tag{29}$$

In this way, the non-convex power allocation problem for the MDM–WDM link is given by

$$\max_{P_l} C. \tag{30}$$

To deal with such problems, the Successive Convex Approximation (SCA) algorithm has been used to successfully approximate the objective function at each iteration by a concave function. Then, the respective parameters are used to update the approximated function in the next iteration. This procedure continues until converging to the optimum solution [40]. In this paper, SCA is used in Appendix D to solve the proposed optimization problem (30).

5. Simulation results

In the first part of this section, we investigate the accuracy of the proposed GN model in both weak and strong coupling regimes by comparing with SSFM simulations, integral-form IGN model [4] and closed-form IGN model [20]. Following, in the second part, we present the optimized power allocation results considering a multi-node nonlinear MDM–WDM network with strong coupling as the link ranges are long. The following MDM–WDM link parameters are used in simulations; the symbol rate is 32 $GBaud$, the channel spacing is 50 GHz, and the center frequency is 193.5 THz. The values for nonlinear coupling coefficient, dispersion, and attenuation are presented by Tables 1 and 2. At each span, an EDFA with 5 dB noise figure compensates for the FMF attenuation. It should be mentioned that despite SMF, there is no standard on FMF parameters and different values are being reported and used by literature which are of course dedicated from experimental FMF transmission setups. However, as shown in [21,22] the fiber type does not affect generality of GN model, since the main assumptions of GN model are about propagating signal and there is no assumption about link parameters. The provided GN model is a general model wherein number of modes and channels can be variable. The main assumptions of GN model are about signal propagation and there is no assumption or limitation on number of channels and number of modes [21,22]. Therefore, the presented GN model works with any number of modes, however, it should be noted that consideration of high number of modes is impractical due to the produced large differential mode delay which results in high receiver complexity.

5.1. Accuracy of proposed GN model

The SSFM simulation [42] is carried out in the Python/Tensorflow environment [43], considering an MDM–WDM system with $D = 3$, $N_{ch} = 9$ [4,13,16,17,21,44]. The GN model assumes Gaussian distribution for transmitted signal, therefore, we use Gaussian symbols taken from a circular-symmetric Gaussian distribution to verify the proposed GN model [45,46]. In addition, we use QPSK symbols for transmission to compare the performance with a standard modulation format [45,46]. In SSFM, we numerically simulate signal transmission by approximating solving the Manakov equation via successive numerical steps as explained in [20]. Note that the simulation parameters should be defined before deploying SSFM in Tensorflow, and thereby we cannot adjust step size adaptively based on maximum nonlinear phase rotation while computations [43]. Besides, fixed step size value results in slow and impractical SSFM computations. To have an SSFM with useful results, the step size should be small enough such that the nonlinear phase rotation does not exceed 0.05 radians, and the number of steps should be defined accordingly. Due to the proportion of nonlinear phase rotation and signal power (that is attenuated along each span), the step sizes at the beginning of the span should be smaller than the ones at the end, therefore, we use a logarithmic step size [42].

Figs. 2(a) and 2(b) plot the SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model versus launched power per channel-mode, respectively for weak coupling ($N_s = 1, L_s = 70$ km) and strong coupling ($N_s = 6, L_s = 80$ km), considering the central channel of LP01 and LP11a/b modes. The proposed GN model matches the SSFM simulation-Gaussian in all cases. As shown in Fig. 2(a), there is an agreement between the proposed GN model, SSFM simulation-Gaussian, and the IGN model in the low power region, since the linear effects are dominant in this region. However, the IGN model overestimates SNR by nearly 1 dB in the weak coupling, and 2 dB in the strong coupling, at the optimal (best equal) launched power per channel-mode, for both LP01 and LP11a/b modes. The error incurred by the IGN model is due to the incoherence assumption which leads to underestimation of NLI noise power [21]. A fixed gap appears between the SSFM simulation Gaussian and QPSK in nonlinear region, since NLI noise power is higher in Gaussian constellation. During the linear region, as opposed to the nonlinear part, a linear channel like AWGN is considered, and thus the SNR is independent of the modulation. A complete investigation over the effect of modulation format on the NLI noise PSD is presented in [45,47].

Figs. 3(a) and 3(b) show the SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model [4] versus number of channels, respectively for weak coupling ($N_s = 1, L_s = 70$ km) and strong coupling ($N_s = 6, L_s = 80$ km). Fig. 4 depicts the SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model [4] versus the number of spans for strong coupling with $L_s = 80$ km span length. In Figs. 3 and 4, we consider the optimum (best equal) launched power per channel-mode and plot the curves of the central channel of LP01 and LP11a/b modes. The proposed GN model matches the SSFM simulation-Gaussian in the different numbers of channels and spans. However, there is a 1 dB to 2 dB difference between the IGN model and the SSFM simulation-Gaussian in weak coupling and strong coupling regimes, respectively.

We present coherent GN model, and generally, GN model does not capture the modulation format effect, since it originally assumes the transmitted signal with Gaussian constellation. Although the modulation format correction is provided by enhanced GN (EGN) model, this formulation is very complicated takes too much time to estimate NLI noise variance. As seen in Figs. 3, 3 and 4, the NLI noise variance gap between GN and QPSK is fixed over different powers, number of channels, and span lengths. As an alternative to use EGN model in practical applications (e.g., network planning), one can use the GN model with directly applying the modulation format corrections. Since,

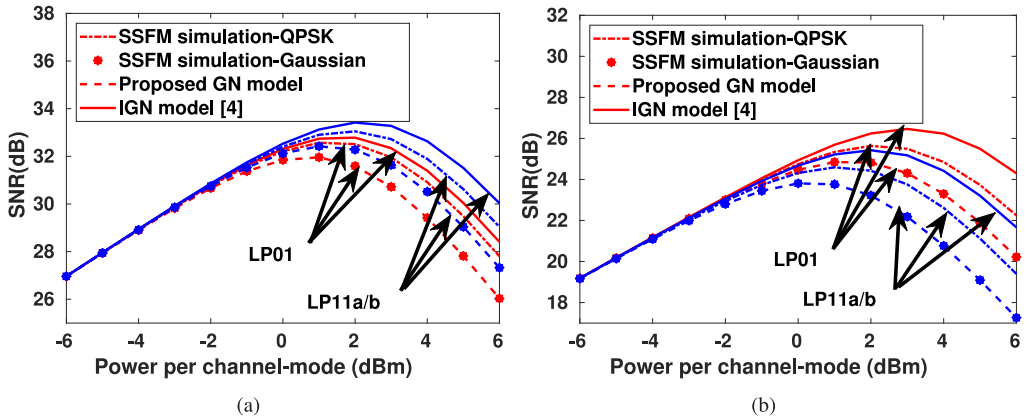


Fig. 2. SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model [4] versus launched power per channel-mode for (a) weak coupling ($N_s = 1, L_s = 70$ km) and (b) strong coupling ($N_s = 6, L_s = 80$ km) regimes.

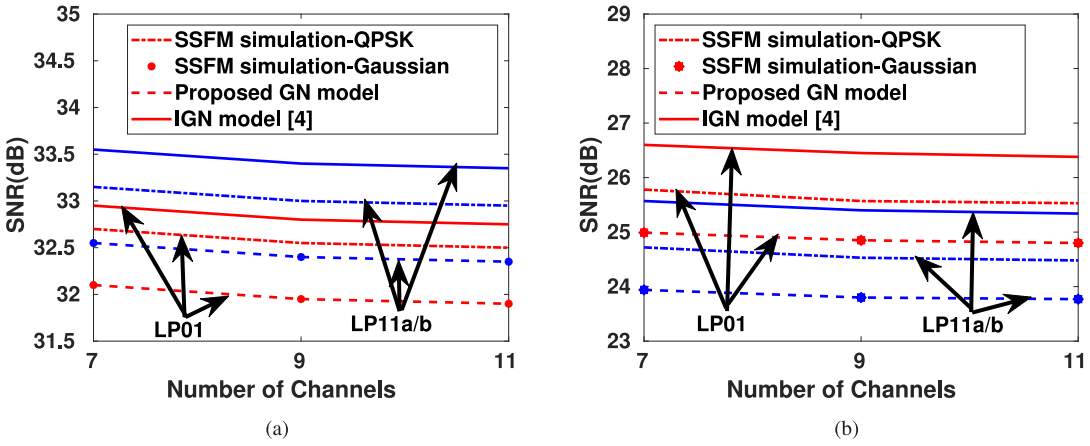


Fig. 3. SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model [4] versus number of channels for (a) weak coupling ($N_s = 1, L_s = 70$ km) and (b) strong coupling ($N_s = 6, L_s = 80$ km) regimes.

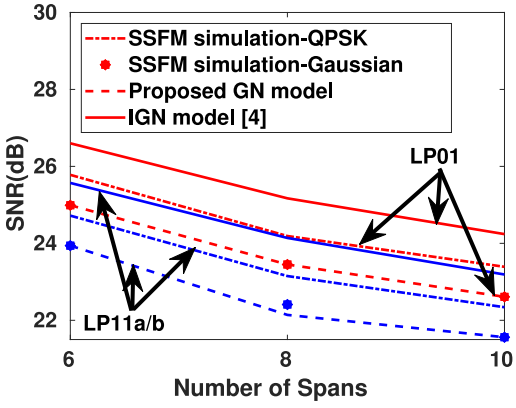


Fig. 4. SNR of SSFM simulation-QPSK, SSFM simulation-Gaussian, proposed GN model, and IGN model [4] versus number of spans for strong coupling regime with $L_s = 80$ km span length.

the NLI noise variance gap is related to the physical layer parameters, therefore, it is enough to calculate it once for a deployed link, lightpath or network.

Although the best way to verify GN model is by experiment, we could not access to an experimental FMF setup, therefore, we provided verifications using SSFM simulations. Note that the well-known SSFM is verified by experiments in some works reflecting the real transmission environment [21], and is the most reliable frequently used tool for verifying GN model in SMF [12,21,22,35,36] and FMF [4,20] literature. As a result, the proposed GN model is reliable and can be used for different applications such as power allocation, spectrum allocation, routing, quality of transmission estimation, network planning, etc [29,37,48].

Note that the presented GN model only works for FMF (and SMF) not for other SDM media such as multicore fiber. This is because the Manakov equation, which serves as the starting point in GN modeling, is different for each media (see [9]).

Table 3 shows the complexity analysis of the proposed GN model, SSFM, integral-form IGN model, and closed-form IGN model. The SSFM simulator is composed of several processing blocks including; MDM/WDM multiplexer and demultiplexer, EDFA at each span, dispersion compensation at receiver. Moreover, SSFM contains two blocks for modeling FMF linear and nonlinear effects. Note that each span has N_{step} steps. The dimension of the transmitted WDM multiplexed signal is $2DN_{sym}$ where N_{sym} is the number of QPSK/16QAM symbols generated for each channel of each mode, thereby the FFT/IFFT in accordance with dispersion has $2DN_{sym}log(2DN_{sym})$ multiplications/summations. Note that $exp(x) = \sum_{i=0}^{n_1} x_i/i!$ can be computed by $2n_1$ multiplications and n_1 summations where n_1 is an integer. As a result, we achieve better accuracy with a larger n_1. The complexity analysis of the proposed GN model is based on (17). At the first, second, and third dimensions of the 3D integration area, n_2, n_3 and n_4 points with identical distances should be considered for numerically solving a 3D integration. Therefore, the 3D integration area becomes $n_2 \times n_3 \times n_4$ small 3D areas over which taking a 3D summation

Table 3
Number of multiplications/summations of different methods.

Methods	Number of multiplications	Number of summations
SSFM simulation	$(N_{step}N_s)(8DN_{sym}log(DN_{sym})+6DN_{sym}+2n_1+1)+(4DN_{sym}log(DN_{sym})+4DN_{sym}+2n_1)$	$(N_{step}N_s)(8DN_{sym}log(DN_{sym})+2DN_{sym}+n_1+2)+(2DN_{sym}log(DN_{sym})+4DN_{sym}+n_1)$
Proposed GN model	$2Dn_2n_3n_4(2n_1+3)$	$2Dn_2n_3n_4(n_1+5)$
Integral-form IGN model	$Dn_2n_3(2n_1+3)$	$Dn_2n_3(n_1+5)$
Closed-form IGN model	$DN_{ch}(8n_1+5)+2$	$DN_{ch}(4n_1+3)$

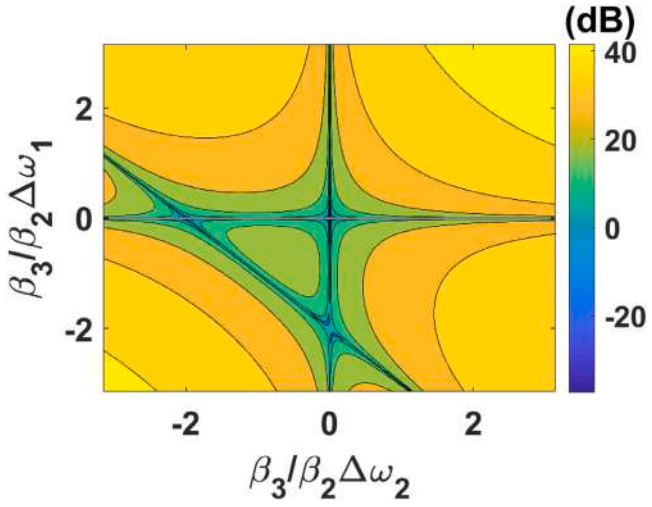


Fig. 5. Contour plot of inter-modal phase-matching condition (5) versus frequency separation ($\Delta\omega$) normalized to the ratio of dispersion to dispersion slope (β_2/β_3) (values are in dB).

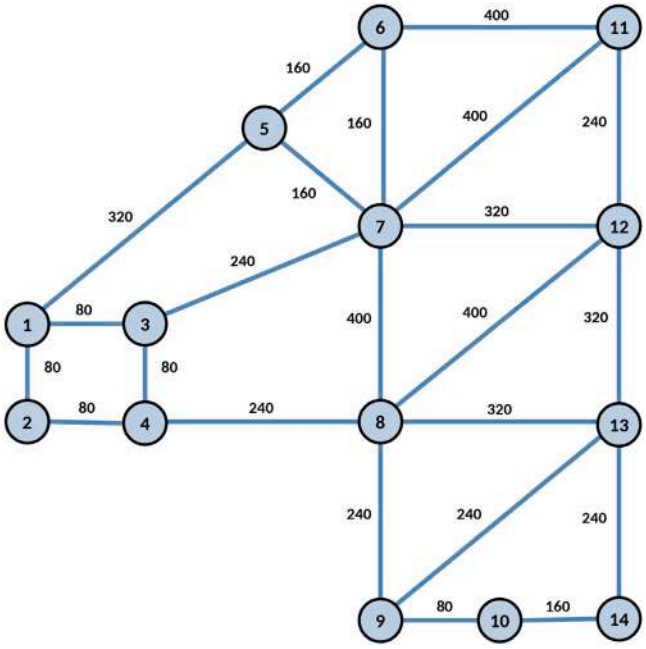


Fig. 6. DTG multi-node nonlinear network.

would be equivalent to a 3D integration over the whole 3D area. The larger values selected for n_2, n_3, and n_4, the higher accuracy can be obtained at the cost of more complexity. The complexity analysis of the integral-form IGN model is based on equation (15) of [4]. For numerically solving a 2D integration, the same procedure as the 3D integration should be considered despite that the summation over the third dimension would not appear as this dimension does not exist. The complexity analysis of the closed-form IGN model is as reported by Table. V of [20]. The proposed GN model provides much less complexity compared with SSFM simulation. Due to incoherency and NLI noise power flatness assumptions made by the closed/integral-form IGN model, it has less complexity compared with the proposed GN model at the cost of less accuracy in modeling the nonlinear effect. It should be noted that closed-form GN model provides the same accuracy with much less complexity compared with integral-form IGN model (see [20]).

The phase-matching criterion has been previously approximated by ignoring the dispersion slope and chromatic dispersion while evaluating at the central frequency. Fig. 5 shows the contour plot of inter-modal phase-matching condition (5) versus frequency separation ($\Delta\omega$) normalized to the ratio of dispersion to dispersion slope (β_2/β_3), for 80 km link length. Weak phase matching (poor FWM efficiency) is shown in green, whereas strong phase matching ($\Delta\beta \approx 0$) is shown in blue. Note that without considering β_3 only one strong phase matching occurs (see [4] Figs. 1b and 2b). The appearance of an additional strong phase matching area (diagonal yellow peak), the modification of phase matching in the region bounded by the three strong phase matching areas, and a slight change in the decay rate of the weak phase matching region are the three main effects of considering dispersion slope. All three phase matching peaks are expected to produce additional strong phase matching products within the area $2\beta_2/\beta_3 < \Delta\omega$. As a result, dispersion slope must be included in ultra wide-band MDM–WDM systems.

We should mention that we had access to an Intel Xeon CPU with 32 cores and 64 GB RAM by which we were limited to consider 3 spatial modes with 9 WDM channels (0.45 THz) bandwidth [4,16,17] in model validation of Figs. 2, 3, and 4. Because of the limitations of our computing resources, the scenarios we examined did not allow for the excitation of β_3-induced phase matching. Thus, Figs. 2, 3, and 4 highlight only the difference between the GN and IGN models. However, as shown in Fig. 5, β_3 is expected to contribute additional phase matching conditions and needs to be accounted for when designing and analyzing the performance of MDM–WDM systems.

5.2. Optimized power allocation

The presented optimized power allocation results are based on "minimum SNR margin maximization" and "total capacity maximization" problems, and included two case scenarios; (a) the best equal power allocation, and (b) the optimized power. In the first scenario, equal powers are allocated to each channel and each mode. In the second situation, the powers assigned to various channels and modes may differ.

The DTG multi-node nonlinear network shown in Fig. 6 is considered in this paper for testing the advantages of optimized power allocation [49]. Any potential scaling of the minimum SNR margin improvement available from power allocation is detected by optimizing over a random lightpath of the network by selecting a random node pair as source–destination [29]. Random traffic demand is considered with a

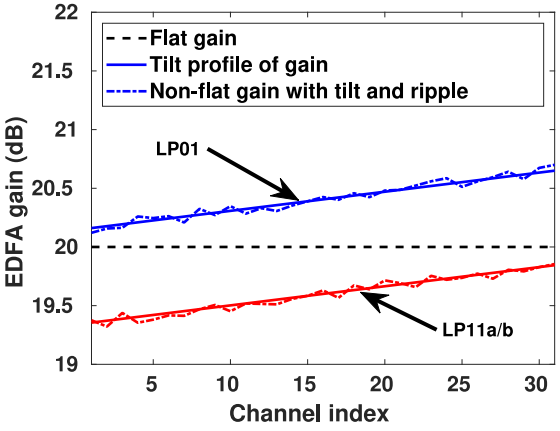


Fig. 7. The considered EDFA gain profile to model the gain tilt and ripple in MDM–WDM system.

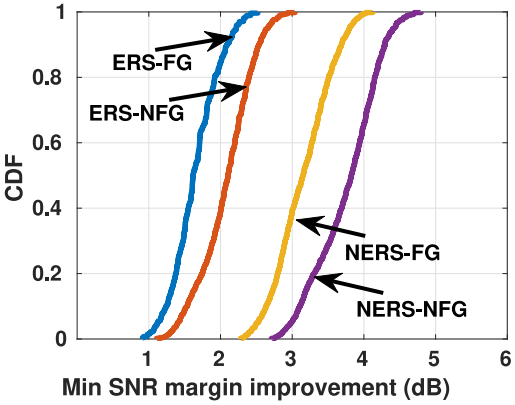


Fig. 8. CDF of obtained SNR margin improvement of optimized power allocation over best flat power allocation in MDM–WDM system, for ERS-FG, ERS-NFG, NERS-FG, and NERS-NFG scenarios.

uniform demand distribution between each node pair, satisfied by using random channel/mode indices. Routes are assigned by shortest-path computation to the destination node [29]. At each node, an add/drop section is considered wherein the dropped channel/mode is not added again [37]. In multi-node nonlinear networks, different channels-modes may propagate at different distances thus accumulate different NLI noises, experience fragmentation/partial utilization thereby see different interacting channels-modes with different modulation formats and may observe different EDFA gains [29].

The SNR margin and the total capacity formulations are dependent on the ASE noise that depends on the EDFA gain. In the Flat EDFA Gain (FG) scenario, all channels and modes are amplified by the same gains at EDFA, whereas in the Non-Flat EDFA Gain (NFG) scenario, various gains are allocated [48,50,51]. Moreover, the SNR margin depends on the required SNR of each channel and mode. In the Equal Required SNR (ERS) scenario, all channels and modes have the same modulation format, whereas, in the Non-Equal Required SNR (NERS) scenario, different modulation formats are employed [29]. Note that the EDFA gain is practically non-flat and a flattening filter should be used to make it flat [52]. Although the focus of this paper is on nonlinearity, the EDFA gain flatness influences NLI noise power; for instance, at higher ASE noise powers, the allocated power is higher which enhances the NLI noise power. Thereby, to have a better discussion and conclusion, we investigate both FG and NFG scenarios. The employment of a discrete data-rates set that falls below and above the highest feasible continuous communication data rates, as well as the integration of different communication hardware generations, are all practical aspects of using different SNR requirements [32–55].

The simulation results for the minimum SNR margin maximization problem include the FG-ERS, FG-NERS, NFG-ERS, and NFG-NERS. Moreover, we consider the FG and NFG scenarios for the total capacity maximization problem. In the FG and NFG, we include the EDFA gain profiles shown in Fig. 7. In the NFG profiles, 0.4 dB tilt is considered along 31 WDM channels [48] and 0.8 dB tilt for 3 FMF modes [51]. The gain ripples are randomly selected within $[-0.05, 0.05]$ dB [48]. In the ERS, PM-QPSK modulation is used, and in NERS, PM-QPSK and PM-16QAM are respectively used for odd and even channels of LP01 mode and vise versa for odd and even channels of LP11a/b. The required SNRs of PM-QPSK and PM-16QAM are 7 and 11.48 dB, respectively [44]. We consider $D = 3$ modes and $N_{ch} = 31$ channels [4,9]. These are the maximum possible mode-channel numbers with our accessible hardware.

Fig. 8 represents Cumulative Distribution Function (CDF) of SNR margin improvement obtained by optimized power allocation over best flat power allocation in MDM–WDM system, for ERS-FG, ERS-NFG, NERS-FG, and NERS-NFG scenarios. Each plot is based on the results of 200 different sets of randomized traffic demands (random channel-mode distributions), which have been defined at the beginning of this section. The obtained minimum SNR margin improvements are in 99% more than 1 dB, 1.2 dB, 2.4 dB and 2.8 dB for ERS-FG, ERS-NFG, NERS-FG, and NERS-NFG scenarios, respectively. The best flat power allocation is a 1 dimensional optimization method while in this case, the optimized power allocation is a 93 dimensional optimization method. The obtained gains are due to the higher degree of freedom of optimized power allocation. The ASE/NLI noise spectrum is not flat in the NFG and NERS scenarios, and the required SNR (modulation format) is not equal among channels/modes in the NERS scenario. As a result, ERS-NFG, NERS-FG, and NERS-NFG scenarios show a much more efficient deployment of power allocation by min SNR margin maximization. This is due to higher optimization degrees of freedom of ERS-NFG, NERS-FG, and NERS-NFG scenarios comparing with FG-ERS.

The total capacity formulation does not the include required SNR term, therefore, only FG and NFG scenarios are applicable. The obtained total capacity improvements are on average 15 Gbits/s and 23 Gbits/s for FG and NFG scenarios, respectively, which is negligible. Therefore, a flat power allocation can effectively maximize the total capacity for a flat/non-flat noise spectrum. This is further supported by (27) which shows that capacity and SNR have a logarithmic relationship. Thus, different channels/modes have almost the same capacity values even if the SNRs of channels and modes become different from each other which in turn results in low optimization degrees of freedom in optimized power allocation.

6. Conclusion

The FMF nonlinearity has been modeled by an additive GN source. The available GN models in MDM–WDM investigations assume that the NLI noise created at each span is summed up in power incoherently at the receiver which leads to underestimating the NLI noise power. Furthermore, these models neglect the dispersion slope, which is important in dealing with large bandwidths. In the first part of this paper, a coherent GN model has been derived considering the dispersion slope and compared with SSFM simulations and the IGN model considering both weak coupling and strong coupling regimes. The proposed GN model provides an acceptable accuracy with much less complexity than the SSFM. On the other hand, the IGN model overestimates the SNR by around 1 dB and 2 dB in weak coupling and strong coupling, respectively, considering the optimum launched power per channel and mode. Therefore, the significance of this paper in the

first part include; presenting a coherent GN model for MDM–WDM system with better accuracy than IGN model in weak and strong coupling, considering dispersion slope in GN model formulation for MDM–WDM system and describing the impact of dispersion slope induced intermodal phase-matching on NLI noise in MDM–WDM system, presenting a GN model formulation with a clear relation between NLI noise and per channel-mode launched power, making it proper for power allocation in link, lightpath and network level. Note that the proposed GN model formulation can be applied beyond the addressed power allocation problems, in applications such as quality of transmission estimation, optical performance monitoring, routing, wavelength and mode allocation, etc.

In the second part of this paper, based on the obtained GN model, the total capacity maximization, as well as the minimum SNR margin maximization problems have been presented and solved. An MDM–WDM multi-node nonlinear network under different scenarios including ERS/NERS, and FG/NFG has been considered in the simulation. Considering 200 different sets of randomized traffic demands (random channel-mode distributions) the CDF of minimum SNR margin improvement in optimized power allocation compared with the best flat power allocation was in 99% more than 1 dB, 1.2 dB, 2.4 dB, and 2.8 dB, for ERS-FG, ERS-NFG, NERS-FG, and NERS-NFG scenarios, respectively. Thereby, the significance of the second part of this paper contains; deploying optimized power allocation in MDM–WDM network, providing comprehensive investigation over optimized power allocation for different MDM–WDM network scenarios.

CRediT authorship contribution statement

M.A. Amirabadi: Conception and design of study, Writing – original draft, Writing – review & editing. **S.A. Nezamalhosseini:** Conception and design of study, Writing – original draft, Writing – review & editing. **M.H. Kahaei:** Conception and design of study, Writing – original draft, Writing – review & editing. **Lawrence R. Chen:** Conception and design of study, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgment

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada. All authors approved the version of the manuscript to be published.

Appendix A. NLI noise variance of *n*th channel and *m*th mode

The PSD of the NLI noise of x polarization of mth mode can be expressed as

$$G_{GN}^{(m,x)}(f)=\int_{-\infty}^{\infty}E[n_{m,x}(z,t_1)n_{m,x}^{*}(z,t_2)]e^{-j2\pi f\tau}d\tau, \tag{A.1}$$

where $\tau=t_1-t_2$. By substituting (14) in (A.1) the PSD of the NLI noise of x polarization of mth mode can be written as in Eq. (A.2) given in Box I. (A.2) can be expressed as

$$G_{GN}^{(m,x)}(f)=G_{GN,1}^{(m,x)}(f)+G_{GN,2}^{(m,x)}(f)+G_{GN,3}^{(m,x)}(f)+G_{GN,4}^{(m,x)}(f), \tag{A.3}$$

where $G_{GN,i}^{(m,x)}(f);i=1,2,3,4$ is the integration by considering G_i terms. The terms of $G_{GN,1}^{(m,x)}(f)$, $G_{GN,2}^{(m,x)}(f)$, $G_{GN,3}^{(m,x)}(f)$, and $G_{GN,4}^{(m,x)}(f)$ contain summations over six indices of k,l,h,k',l',h'. However, most of the combinations of these indices can be removed, since $\xi_{i,q},\zeta_{i,q}$ are independent Gaussian random variables with $E[\xi_{i,q}]=0, E[\zeta_{i,q}]=0$, and $E[|\xi_{i,q}|^2]=1, E[|\zeta_{i,q}|^2]=1$, and $E[\xi_{i,q}\xi_{j,q}^{*}]=0, E[\zeta_{i,q}\zeta_{j,q}^{*}]=0$. Furthermore, as explained earlier, the GN modeling only considers the FWM-like interactions between frequency components, not the SPM-like and XPM-like terms [35]. Note that here, the "SPM-like" and "XPM-like" refer to phase shifts between frequency components of the transmitted signal, not between WDM channels. The transmitted signal in GN modeling is treated as a single broadband signal without any division into channels and the individual channels are never seen in the final results. Accordingly, in calculating $G_{GN,1}^{(m,x)}(f)$ only index combinations $k=k',l=l',h=h'$ and $k=l',l=k',h=h'$ remain which results in

$$G_{GN,1}^{(m,x)}(f)=2\int_{-\infty}^{\infty}\left[f_0^3\sum_{q=1}^{D}\sum_{k,l,h}\tilde{\gamma}_{mq}^2|\eta(kf_0,lf_0,hf_0)|^2e^{j2\pi(kf_0+lf_0-hf_0-f)\tau}\times G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0)\right]d\tau. \tag{A.4}$$

In calculating $G_{GN,2}^{(m,x)}(f)$ and $G_{GN,3}^{(m,x)}(f)$ no index combination remains, therefore, these terms are not parts of the final results. In calculating $G_{GN,4}^{(m,x)}(f)$ only the index combination $k=k',l=l',h=h'$ remains which results in

$$G_{GN,4}^{(m,x)}(f)=\int_{-\infty}^{\infty}\left[f_0^3\sum_{q=1}^{D}\sum_{k,l,h}\tilde{\gamma}_{mq}^2|\eta(kf_0,lf_0,hf_0)|^2e^{j2\pi(kf_0+lf_0-hf_0-f)\tau}\times G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0)\right]d\tau. \tag{A.5}$$

Accordingly, by substituting (A.4) and (A.5) in (A.3), the PSD of the NLI noise of x polarization of mth mode can be written as

$$\begin{aligned}G_{GN}^{(m,x)}(f)&=\int_{-\infty}^{\infty}\left[3f_0^3\sum_{q=1}^{D}\sum_{k,l,h}\tilde{\gamma}_{mq}^2|\eta(kf_0,lf_0,hf_0)|^2e^{j2\pi(kf_0+lf_0-hf_0-f)\tau}\times G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0)\right]d\tau\\&=3f_0^3\sum_{q=1}^{D}\sum_{k,l,h}\tilde{\gamma}_{mq}^2|\eta(kf_0,lf_0,hf_0)|^2\delta(j2\pi(kf_0+lf_0-hf_0-f))\times G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0).\end{aligned} \tag{A.6}$$

As discussed earlier, the obtained PSD of the NLI noise is discrete which is related to the discrete assumption of the transmitted signal. However, the transmitted signal has a continuous spectrum, and $f_0\to 0$, therefore, the PSD of the NLI noise of x polarization of mth mode becomes equal to

$$G_{GN}^{(m,x)}(f)=3\sum_{q=1}^{D}\iint_{-\infty}^{\infty}\tilde{\gamma}_{mq}^2|\eta(f_1,f_2,f)|^2G_{Tx}^{(m,x)}(f_1)\times G_{Tx}^{(q,x)}(f_2)G_{Tx}^{(q,x)}(f_1+f_2-f)df_1df_2. \tag{A.7}$$

Note that in the MDM–WDM link, $G_{GN}^{(m,x)}(f)=G_{GN}^{(m,y)}(f)$, and $G_{GN}^{(m)}(f)=G_{GN}^{(m,x)}(f)+G_{GN}^{(m,y)}(f)$. Accordingly, the PSD of the NLI noise of mth mode becomes equal to

$$G_{GN}^{(m)}(f)=3/4\sum_{q=1}^{D}\iint_{-\infty}^{\infty}\tilde{\gamma}_{mq}^2|\eta(f_1,f_2,f)|^2G_{Tx}^{(m)}(f_1)G_{Tx}^{(q)}(f_2)\times G_{Tx}^{(q)}(f_1+f_2-f)df_1df_2. \tag{A.8}$$

$$
\begin{aligned}
&G_{GN}^{(m,x)}(f) = \\
&\int_{-\infty}^{\infty} f_0^3 \sum_{q=1}^{D} \sum_{k,l,h,k',l',h'} \tilde{\gamma}_{mq}^2 \eta(kf_0, lf_0, hf_0)\eta^*(k'f_0, l'f_0, h'f_0) e^{j2\pi(k+l-h)f_0 t_1} e^{-j2\pi(k'f_0+l'f_0-h'f_0)t_2} e^{-j2\pi f\tau} \\
&\Bigg[\underbrace{E[\xi_{k,q}\xi^*_{k',q}\xi_{l,q}\xi^*_{h,q}\xi_{l',q}\xi^*_{h',q}]\sqrt{G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(m,x)}(k'f_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0)G_{Tx}^{(q,x)}(l'f_0)G_{Tx}^{(q,x)}(h'f_0)}}_{G_1} + \\
&\underbrace{E[\xi_{k,q}\xi^*_{k',q}\xi_{l,q}\xi^*_{h,q}\zeta_{l',q}\zeta^*_{h',q}]\sqrt{G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(m,x)}(k'f_0)G_{Tx}^{(q,x)}(lf_0)G_{Tx}^{(q,x)}(hf_0)G_{Tx}^{(q,y)}(l'f_0)G_{Tx}^{(q,y)}(h'f_0)}}_{G_2} + \\
&\underbrace{E[\xi_{k,q}\xi^*_{k',q}\zeta_{l,q}\zeta^*_{h,q}\xi_{l',q}\xi^*_{h',q}]\sqrt{G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(m,x)}(k'f_0)G_{Tx}^{(q,y)}(lf_0)G_{Tx}^{(q,y)}(hf_0)G_{Tx}^{(q,x)}(l'f_0)G_{Tx}^{(q,x)}(h'f_0)}}_{G_3} + \\
&\underbrace{E[\xi_{k,q}\xi^*_{k',q}\zeta_{l,q}\zeta^*_{h,q}\zeta_{l',q}\zeta^*_{h',q}]\sqrt{G_{Tx}^{(m,x)}(kf_0)G_{Tx}^{(m,x)}(k'f_0)G_{Tx}^{(q,y)}(lf_0)G_{Tx}^{(q,y)}(hf_0)G_{Tx}^{(q,y)}(l'f_0)G_{Tx}^{(q,y)}(h'f_0)}}_{G_4} \Bigg] d\tau.
\end{aligned}
\tag{A.2}
$$

Box I.

(A.8) represents different types of nonlinear interaction between different channels/modes which can be categorized into inter and intra channel nonlinearity, considering frequency channel interactions, and inter and intra modal nonlinearity, considering spatial mode interactions.

At the receiver side, the detector performs demodulation, matched filtering, zero-forcing equalization, and sampling. The matched filtering operation can be described by

$$\int_{-\infty}^{\infty} \tilde{A}(f) R_n g^*_{n,m}(f) df, \tag{A.9}$$

where $g_{n,m}(f)$ is the spectral shape of transmitted pulse on the nth channel and the mth mode which has been normalized such that $\int_{-\infty}^{\infty} g_{n,m}(f)df = 1$, and R_n is the symbol rate of the nth channel. Therefore, the variance of the NLI noise of the nth channel and the mth mode can be written as

$$
\begin{aligned}
\sigma^2_{GN_{m,n}} = 3/4 \sum_{q=1}^{D} \iiint_{-\infty}^{\infty} \tilde{\gamma}^2_{mq} |\eta(f_1, f_2, f)|^2 G_{Tx}^{(m)}(f_1) G_{Tx}^{(q)}(f_2) \\
\times G_{Tx}^{(q)}(f_1 + f_2 - f) R_n g^*_{n,m}(f) df_1 df_2 df.
\end{aligned}
\tag{A.10}
$$

The three-dimensional integration in (A.10) should be solved to obtain the NLI noise variance which has high computational complexity. Following, consideration of such an integral-form formulation in power allocation problems is not effective, as the iterative optimizer algorithm should calculate the whole integration at each iteration once [29]. In this situation, a discrete-form GN model is appropriate [29]. Discrete-form GN model can be obtained by rewriting the PSD of the transmitted WDM signal in terms of its discrete frequency components. Therefore, the launched powers can be factored out from the integral. The first step towards discretization of GN model is describing the PSD of the transmitted signal in the nth channel and the mth mode in the following form

$$G_{Tx}^{(m)}(f) = \sum_{n=1}^{N_{ch}} P_{n,m} g_{n,m}(f), \tag{A.11}$$

where $P_{n,m}$ is the transmitted power on the nth channel and the mth mode, and is fixed versus f. $g_{n,m}(f)$ is variable versus f, and can take different shapes in f, its center frequency is f_{c_n} and its bandwidth is B_n.

Substituting (A.11) in (A.10), the variance of NLI noise of the nth channel and the mth mode takes the form of (17).

Algorithm 1: Bisection method to solve convex optimization problem (26).

Initialization: upper bound $u = 100$, and lower bound $l = -10$;
$\beta \leftarrow u$;
Solve convex problem (26) by Lagrangian method;
if $\hat{P}^{*(t)} == NAN$ **then**
  break;
end
$\beta \leftarrow l$;
Solve convex problem (26) by Lagrangian method;
if $\hat{P}^{*(t)} == NAN$ **then**
  break;
end
while $u - l \le \epsilon$ **do**
  $\beta \leftarrow (u + l)/2$;
  Solve convex problem (26) by Lagrangian method;
  if $\hat{P}^{*(t)} == NAN$ **then**
    $l \leftarrow \beta$;
  else
    $u \leftarrow \beta$;
  end
end

Appendix B. Convexity proof of optimization problem (26)

The expression

$$log\left(e^{log(N_s F(G-1)\hbar\nu B_l)} + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1, l_2, l_3)\right), \tag{B.1}$$

is convex in $\hat{P}_l$, since $log(\sum_l e^{x_l})$ is a convex function in x_l [56]. The constraint function of (26) the summation of some convex functions, therefore, it is convex. The objective and constraint functions of (26) are convex, therefore, (26) is a convex optimization problem.

Appendix C. Bisection method to solve convex optimization problem (26)

In the Bisection method, a region is selected for the objective function of (26), and a candidate for β is selected and fixed. Then,

Algorithm 2: Lagrangian duality method to solve convex problem (26).

initialization: iteration counter $t = 0$, step size parameter $a > 0$, and $\lambda_l^{(0)} \geq 0$;
while *achieving convergence* **do**
  Solve convex problem (26) with fixed λ_l, and obtain optimal power $\hat{P}^{*(t)}$;
  $\lambda_l^{(t+1)} = \left[\lambda_l^{(t)} - a\Delta_l\right]^+$;
  Update $t = t + 1$;
end

Algorithm 3: SCA algorithm.

Inputs: iteration counter $t = 0$, $\xi_l^{(0)} = 1$,$\zeta_l^{(0)} = 0$, $\epsilon = 10^{-10}$;
Output: optimal power $\hat{P}_l^*$;
while *achieving convergence* **do**
  Solve (D.1) by the gradient ascent method and obtain $\hat{P}^{*(t)}$;
  Calculate $x_l^{*(t+1)}$ using (D.3);
  if $x_l^{*(t+1)} \leq \epsilon$ **then**
    Set $\zeta_l^{(t+1)} = 0$ and $\xi_l^{(t+1)} = 0$;
  else
    Update $\zeta_l^{(t+1)}$ and $\xi_l^{(t+1)}$;
  end
  Update $t = t + 1$;
end

the feasibility problem is solved using the Lagrangian method [38]. It should search about the optimum β in the defined region, accordingly, at each step, it updates the region boundaries based on the obtained solution for the feasibility problem in the last step. The same procedure repeats until achieving a convergence.

Algorithm 1 summarizes the Bisection method used for solving the convex optimization (26). The first step is defining a search region for β by choosing the upper (u) and lower (l) bounds. The upper bound is assigned to β, and the problem (26) is converted to a feasibility problem. Next, the feasibility problem is solved using the Lagrangian method. If the feasibility problem does not have a solution, it means that the feasible set is empty and the selected upper bound is lower than the optimal solution, therefore, a higher upper bound should be used. The defined lower bound should be tested/adjusted in the same way. Now, the search region boundaries are finalized and the main loop of the Bisection method starts.

The second step is solving the feasibility problem using the Lagrangian method. In this step, $(u + l)/2$ value is assigned to β and the feasibility problem is solved. If the feasibility problem does not have a solution, it means that the feasible set is empty, and the assigned value to β is the lower bound of the feasible set. If the feasibility problem has a solution, it means that the assigned value to β is the upper bound of the feasible set. Therefore, the updated upper/lower bound is updated and the second step is repeated until the convergence. Algorithm 2 summarizes the Lagrangian method used for solving the feasibility problem, the explanation is provided in the following.

The Lagrangian function for (26) is given by

$$\beta + \sum_{l=1}^{DN_{ch}} \lambda_l \left(log(SNR_l^{req}) + log\left(N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3) \right) - \hat{P}_l - \beta \right), \tag{C.1}$$

where $\lambda_l \in R^+$ is the Lagrangian multiplier. Hence, the Lagrangian dual function of (26) takes the following form

$$\inf_{\hat{P}_l} \beta + \sum_{l=1}^{DN_{ch}} \lambda_l \left(log(SNR_l^{req}) + log\left(N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3) \right) - \hat{P}_l - \beta \right). \tag{C.2}$$

The dual problem is a convex optimization problem [38], therefore, the optimization problem (C.2) is a convex optimization problem with respect to $\hat{P}_l$ for a fixed λ_l. Note that at each iteration of Algorithm 2, λ_l is updated based on the derivative of (C.2) with respect to λ_l which can be expressed as

$$\Delta_l = log(SNR_l^{req}) + log\left(N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3) \right) - \hat{P}_l - \beta. \tag{C.3}$$

Remark 1.

In Algorithm 2, (C.2) is solved at each iteration as a function of $\hat{P}_l$ using the gradient descent algorithm which will converge to its optimum solution due to the convexity of the problem. This procedure is repeated by Algorithm 1 in the "While loop" by which the minimum SNR margin is improved until convergence to the maximum value. Note that Algorithm 1 will stop when the difference between upper and lower bounds becomes less than ϵ.

Appendix D. SCA method to solve convex optimization problem (30)

To deal with optimization problem (30), we use SCA as summarized in Algorithm 3. More specifically, we first define an auxiliary variable $\hat{P}_l = ln(P_l)$, by which (30) can be expressed as

$$\max_{\hat{P}_l} 2 \sum_{l=1}^{DN_{ch}} R_l log_2\left(1 + \frac{e^{\hat{P}_l}}{N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3)}\right). \tag{D.1}$$

The objective function of (D.1) is of the form $\sum_l log_2(1 + x_l)$ which can be approximated by the concave function $\sum_l \zeta_l + \xi_l log_2(x_l)$ where $\zeta_l = log_2(1 + x_l^*) - x_l^*/(x_l^* + 1)log_2(x_l^*)$, and $\xi_l = x_l^*/(x_l^* + 1)$, with x_l^* denoting the optimal point computed in the previous iteration [40]. In this way, (D.1) can be approximated by a convex model as

$$\max_{\hat{P}_l} 2 \sum_{l=1}^{DN_{ch}} R_l \left[\zeta_l + \xi_l \left(\hat{P}_l - log_2\left(N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3) \right) \right) \right], \tag{D.2}$$

where

$$x_l^* = \frac{e^{\hat{P}_l^*}}{N_s F(G-1)h\nu B_l + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3)}, \tag{D.3}$$

where $\hat{P}_l^*$ is the optimal power obtained by solving (D.2) using the gradient ascent method.

Note that the expression

$$log_2\left(e^{log(N_s F(G-1)h\nu B_l)} + \sum_{l_1,l_2,l_3=1}^{DN_{ch}} e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} D_l(l_1,l_2,l_3) \right), \tag{D.4}$$

is convex in $\hat{P}_l$, since $log(\sum_l e^{x_l})$ is a convex function in x_l [38]. Accordingly, the objective function of (D.1) is concave, since it is a

sum of some concave functions, and thus problem (D.1) is a convex optimization problem.

Remark 2.

In Algorithm 3, (D.1) is solved at each iteration as a function of $\hat{P}_l$ using the gradient ascent algorithm which will converge to its optimum solution due to the convexity of the problem. This procedure is repeated by the SCA algorithm in the "While loop", by which the total capacity is improved successively until convergence to the maximum value [38]. Note that, since the power allocation is limited due to the nonlinearity effects, the SCA algorithm will stop searching if by increasing the power, the total capacity starts decreasing.

References

[1] R. Ryf, S. Randel, A.H. Gnauck, C. Bolle, R.-J. Essiambre, P.J. Winzer, D.W. Peckham, A. McCurdy, R. Lingle, Space-division multiplexing over 10 km of three-mode fiber using coherent 6×6 MIMO processing, in: Proc. Opt. Fiber Commun. Conf. Expo./ Nat. Fiber Opt. Eng. Conf, 2011, pp. 1–3.

[2] G. Rademacher, S. Warm, K. Petermann, Analytical description of cross-modal nonlinear interaction in mode multiplexed multimode fibers, IEEE Photonics Technol. Lett. 24 (21) (2012) 1929–1932.

[3] P.J. Winzer, G.J. Foschini, MIMO capacities and outage probabilities in spatially multiplexed optical transport systems, Opt. Express 19 (17) (2011) 16680–16696.

[4] G. Rademacher, K. Petermann, Nonlinear Gaussian noise model for multimode fibers with space-division multiplexing, J. Lightwave Technol. 34 (9) (2016) 2280–2287.

[5] P. Sillard, Few-mode-fiber developments and applications, in: 2018 23rd Opto-Electronics and Communications Conference, OECC, 2018, pp. 1–2.

[6] C. Koebele, M. Salsi, L. Milord, R. Ryf, C. Bolle, P. Sillard, S. Bigo, G. Charlet, 40km transmission of five mode division multiplexed data streams at 100Gb/s with low MIMO-DSP complexity, in: Proc. 37th Eur. Conf. Opt. Commun., 2011, pp. 1–3.

[7] B. Inan, B. Spinnler, F. Ferreira, D. van den Borne, A. Lobato, S. Adhikari, V. Sleiffer, M. Kuschnerov, N. Hanik, S. Jansen, DSP complexity of mode-division multiplexed receivers, Optics Express 20 (10) (2012) 10859–10869.

[8] F. Ye, S. Warm, K. Petermann, Differential mode delay management in spliced multimode fiber transmission systems, in: Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, 2013, pp. 1–3.

[9] S. Mumtaz, R.J. Essiambre, G.P. Agrawal, Nonlinear propagation in multimode and multicore fibers: Generalization of the Manakov equations, J. Lightwave Technol. 31 (3) (2012) 398–406.

[10] D. Kroushkov, G. Rademacher, K. Petermann, Cross mode modulation in multimode fibers, Optcs Lett. 38 (10) (2013) 1642–1644.

[11] A.D. Ellis, N. Mac Suibhne, F.C.G. Gunning, S. Sygletos, Expressions for the nonlinear transmission performance of multi-mode optical fiber, Opt. Express 21 (19) (2013) 22834–22846.

[12] P. Serena, C. Lasagni, A. Bononi, The enhanced Gaussian noise model extended to polarization-dependent loss, J. Lightwave Technol. 38 (20) (2020) 5685–5694.

[13] P. Poggiolini, The GN model of non-linear propagation in uncompensated coherent optical systems, J. Lightwave Technol. 30 (24) (2012) 3857–3879.

[14] G. Rademacher, R.S. Luis, B.J. Puttnam, H. Furukawa, R. Maruyama, K. Aikawa, Y. Awaji, N. Wada, Investigation of intermodal four-wave mixing for nonlinear signal processing in few-mode fibers, IEEE Photonics Technol. Lett. 30 (17) (2018) 1527–1530.

[15] G. Rademacher, S. Warm, K. Petermann, Nonlinear interaction in differential mode delay managed mode-division multiplexed transmission systems, Opt. Express 23 (1) (2015) 55–60.

[16] G. Rademacher, R.S. Luis, B.J. Puttnam, R. Maruyama, K. Aikawa, Y. Awaji, H. Furukawa, K. Petermann, N. Wada, Investigation of intermodal nonlinear signal distortions in few-mode fiber transmission, J. Lightwave Technol. 37 (4) (2019) 1273–1279.

[17] G. Rademacher, R.S. Luís, B.J. Puttnam, Y. Awaji, M. Suzuki, T. Hasegawa, H. Furukawa, N. Wada, Wideband intermodal nonlinear signal processing with a highly nonlinear few-mode fiber, IEEE J. Sel. Top. Quantum Electron. 26 (4) (2020) 1–7.

[18] A.E. Elfiqi, A.A. Ali, Z.A. El-Sahn, K. Kato, H.M. Shalaby, Theoretical analysis of long-haul systems adopting mode-division multiplexing, Opt. Commun. 445 (2019) 10–18.

[19] A.A. Ali, A.E. El-Fiqi, Z.A. El-Sahn, H.M. Shalaby, R.K. Pokharel, Analytical formula of nonlinear interference in few-mode fibers in strong coupling regime, in: 2015 17th International Conference on Transparent Optical Networks, ICTON, 2015, pp. 1–4.

[20] M.A. Amirabadi, M.H. Kahaei, S.A. Nezamalhosseini, Lawrence R. Chen, Optimal power allocation in nonlinear MDM-WDM systems using Gaussian noise model, IET Optoelectron. (2022).

[21] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, F. Forghieri, The GN-model of fiber non-linear propagation and its applications, J. Lightwave Technol. 32 (4) (2013) 694–721.

[22] P. Poggiolini, Y. Jiang, A. Carena, F. Forghieri, Analytical modeling of the impact of fiber non-linear propagation on coherent systems and networks, in: Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks, 2016.

[23] M.R. Zefreh, F. Forghieri, S. Piciaccia, P. Poggiolini, Accurate closed-form real-time EGN model formula leveraging machine-learning over 8500 thoroughly randomized full C-band systems, J. Lightwave Technol. 38 (18) (2020) 4987–4999.

[24] M.R. Zefreh, P. Poggiolini, Characterization of the link function in GN and EGN methods for nonlinearity assessment of ultrawideband coherent fiber optic communication systems with Raman effect, 2020, arXiv preprint arXiv:2009.12687.

[25] M.R. Zefreh, A. Carena, F. Forghieri, S. Piciaccia, P. Poggiolini, A GN/EGN-model real-time closed-form formula tested over 7, 000 virtual links, in: 45th European Conference on Optical Communication, 2019, pp. 1–4.

[26] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, F. Forghieri, A simple and effective closed-form GN model correction formula accounting for signal non-Gaussian distribution, J. Lightwave Technol. 33 (2) (2015) 459–473.

[27] P. Poggiolini, Y. Jiang, A. Carena, F. Forghieri, A simple and accurate closed-form EGN model formula, 2015, arXiv:1503.04132.

[28] P. Johannisson, E. Agrell, Modeling of nonlinear signal distortion in fiber-optic networks, J. Lightwave Technol. 32 (23) (2014) 4544–4552.

[29] I. Roberts, J.M. Kahn, D. Boertjes, Convex channel power optimization in nonlinear WDM systems using Gaussian noise model, J. Lightwave Technol. 34 (13) (2016) 3212–3222.

[30] I. Roberts, J.M. Kahn, J. Harley, D.W. Boertjes, Channel power optimization of WDM systems following Gaussian noise nonlinearity model in presence of stimulated Raman scattering, J. Lightwave Technol. 35 (23) (2017) 5237–5249.

[31] I. Roberts, J.M. Kahn, Measurement-based optimization of channel powers with non-Gaussian nonlinear interference noise, J. Lightwave Technol. 36 (13) (2018) 2746–2756.

[32] I. Roberts, J.M. Kahn, Efficient discrete rate assignment and power optimization in optical communication systems following the Gaussian noise model, J. Lightwave Technol. 35 (20) (2017) 4425–4437.

[33] S.O. Arik, J.M. Kahn, K.P. Ho, MIMO signal processing for mode-division multiplexing: An overview of channel models and signal processing architectures, IEEE Signal Process. Mag. 31 (2) (2014) 25–34.

[34] A. Mecozzi, C. Antonelli, M. Shtaif, Coupled Manakov equations in multimode fibers with strongly coupled groups of modes, Opt. Express 20 (21) (2012) 23436–23441.

[35] A. Carena, V. Curri, G. Bosco, P. Poggiolini, F. Forghieri, Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links, J. Lightwave Technol. 30 (10) (2012) 1524–1539.

[36] P. Serena, A. Bononi, A time-domain extended Gaussian noise model, J. Lightwave Technol. 33 (7) (2015) 1459–1472.

[37] H. Rabbani, L. Beygi, S. Ghoshooni, H. Rabbani, E. Agrell, Quality of transmission aware optical networking using enhanced Gaussian noise model, J. Lightwave Technol. 37 (3) (2019) 831–838.

[38] S. Boyd, S.P. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[39] F. Daneshfar, E. Hosseini, Load-frequency control in a deregulated environment based on bisection search, Iran. J. Electr. Electron. Eng. 8 (4) (2012) 303–310.

[40] R. Jiang, Q. Wang, H. Haas, Z. Wang, Joint user association and power allocation for cell-free visible light communication networks, IEEE J. Sel. Areas Commun. 36 (1) (2017) 136–148.

[41] F.M. Ferreira, C.S. Costa, S. Sygletos, A.D. Ellis, Overcoming degradation in spatial multiplexing systems with stochastic nonlinear impairments, Sci. Rep. 8 (1) (2018) 1–10.

[42] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, S. Benedetto, Suppression of spurious tones induced by the split-step method in fiber systems simulation, IEEE Photonics Technol. Lett. 12 (5) (2000) 489–491.

[43] S. Gaiarin, F. Daros, R.T. Jones, D. Zibar, End-to-end optimization of coherent optical communications over the split-step Fourier method guided by the nonlinear Fourier transform theory, J. Lightwave Technol. (2020).

[44] P. Poggiolini, A. Carena, V. Curri, G. Bosco, F. Forghieri, Analytical modeling of nonlinear propagation in uncompensated optical transmission links, IEEE Photonics Technol. Lett. 23 (11) (2011) 742–744.

[45] D. Semrau, R.I. Killey, P. Bayvel, A closed-form approximation of the Gaussian noise model in the presence of inter-channel stimulated Raman scattering, J. Lightwave Technol. 37 (9) (2019) 1924–1936.

[46] D. Semrau, E. Sillekens, R.I. Killey, P. Bayvel, A modulation format correction formula for the Gaussian noise model in the presence of inter-channel stimulated Raman scattering, J. Lightwave Technol. 37 (19) (2019) 5122–5131.

[47] R. Dar, M. Feder, A. Mecozzi, M. Shtaif, Accumulation of nonlinear interference noise in fiber-optic systems, Opt. Express 22 (12) (2014) 14199–14211.

[48] R. Hashemi, H. Beyranvand, H. Rabbani, Joint channel power and amplifier gain optimization in coherent DWDM systems, Opt. Commun. 475 (2020) 126212.

[49] D.J. Ives, P. Bayvel, S.J. Savory, Routing, modulation, spectrum and launch power assignment to maximize the traffic throughput of a nonlinear optical mesh network, Photon. Netw. Commun. 29 (3) (2015) 244–256.
[50] J.K. Perin, J.M. Kahn, J.D. Downie, J. Hurley, K. Bennett, Importance of amplifier physics in maximizing the capacity of submarine links, J. Lightwave Technol. 37 (9) (2019) 2076–2085.
[51] Y. Jung, E.L. Lim, Q. Kang, T.C. May-Smith, N.H.L. Wong, R. Standish, F. Poletti, J.K. Sahu, S.U. Alam, D.J. Richardson, Cladding pumped few-mode EDFA for mode division multiplexed transmission, Opt. Express 22 (23) (2014) 29008–29013.
[52] M. Yucel, G. Yenilmez, The comparison of the gain flattening techniques EDFA configurations in the C/L band, J. Optoelectron. Adv. Mater. 17 (1450) (2015) 9–10–1457.
[53] G.-H. Gho, L. Klak, J.M. Kahn, Rate-adaptive coding for optical fiber transmission systems, J. Lightwave Technol. 29 (2) (2011) 222–233.
[54] M. Karlsson, E. Agrell, Multidimensional modulation and coding in optical transport, J. Lightwave Technol. 35 (4) (2017) 876–884.
[55] K. Roberts, Q. Zhuge, I. Monga, S. Gareau, C. Laperle, Beyond 100 gb/s: Capacity, flexibility, and network optimization [invited], J. Opt. Commun. Netw. 9 (4) (2017) C12–C24.
[56] D.J. Ives, S.J. Savory, Transmitter optimized optical networks, in: National Fiber Optic Engineers Conference Optical Society of America, 2013, pp. JW2A–64, March.