Received January 24, 2022, accepted February 22, 2022. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.3154414

Joint Power and Gain Allocation in MDM-WDM Optical Communication Networks Based on Enhanced Gaussian Noise Model

MOHAMMAD ALI AMIRABADI[1], MOHAMMAD HOSSEIN KAHAEI[1], S. ALIREZA NEZAMALHOSSEINI[1], AND LAWRENCE R. CHEN[2], (Senior Member, IEEE)

[1]School of Electrical Engineering, Iran University of Science and Technology, Tehran 1684613114, Iran
[2]Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada

Corresponding author: S. Alireza Nezamalhosseini (nezam@iust.ac.ir)

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

ABSTRACT Achieving reliable communication over different wavelength channels and modes is one of the main goals of Mode Division Multiplexing-Wavelength Division Multiplexing (MDM-WDM) transmission. The reliability can be described by the minimum Signal to Noise Ratio (SNR) margin which depends on launch power, the gain of Few-Mode Erbium-Doped Fiber Amplifiers (FM-EDFA), and the nonlinear impairments of Few-Mode Fiber (FMF). In this paper, we develop the Enhanced Gaussian Noise (EGN) nonlinear model for FMF, which can be used in both weak and strong coupling regimes. We validate the model by comparing simulation results with those obtained through the Split-Step Fourier Method. Based on our proposed EGN model, we address the problem of joint optimized power and gain allocation based on minimum SNR margin maximization when accounting for practical FM-EDFA constraints such as saturation power and maximum gain. The problem is solved using a convex optimization approach and considering different scenarios such as the best equal power, optimized power, and joint optimized power and gain. Results demonstrate that the minimum SNR margin improvement for the joint optimized power and gain allocation compared to the best equal power allocation is 1.4 *dB* and 1.7 *dB* for MDM-single channel and single-mode fiber-WDM systems, respectively.

INDEX TERMS Enhanced Gaussian noise model, few-mode fiber, power allocation, gain allocation.

I. INTRODUCTION

Mode Division Multiplexing (MDM) over Few-Mode Fibers (FMF) or multi-core fibers has emerged as a possible solution for overcoming the data-rate crunch in optical communication networks. Theoretically, deploying FMF that supports D spatial modes can increase capacity D times [1]. The combination of MDM with Wavelength Division Multiplexing (WDM) and polarization division multiplexing schemes further increases total capacity. However, MDM-WDM systems suffer from both linear and nonlinear transmission impairments in FMF. The linear impairments include attenuation, chromatic and modal dispersion, and linear coupling [2]–[6]. The linear coupling between spatial (polarization) modes results in a power transfer from one spatial (polarization) mode to another one [3]. When linear mode coupling is insignificant compared to the linear polarization coupling, it is called weak coupling which is more prone in short-range links [7], [8]. On the other hand, in long-range links, strong coupling appears wherein the linear mode coupling is significant compared to the linear polarization coupling [3], [5]. For compensating FMF linear effects in the weak coupling regime, each mode is processed separately without using complex Multiple-Input-Multiple-Output (MIMO) Digital Signal Processing (DSP) [8]. However, for compensating FMF linear effects in the strong coupling regime, MIMO DSP is required [2]. It has been shown that the MIMO DSP complexity can be used only in the case of nearly equal group delays between the propagating modes [9]–[11]. Graded index fibers, especially those with a nearly parabolic index profile, can minimize the differential mode group delay and are the most commonly employed in current FMF-based MDM-WDM transmission [4]. The nonlinear impairments include Kerr-based nonlinearity and

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Tang.

nonlinear coupling [1]; these represent the main limitations towards realizing practical MDM-WDM systems.

A. RELATED WORKS

During the last decade, several studies have investigated the analyzed the impact of FMF nonlinearities; these include those that focus mainly on numerical simulations [3] and analytical predictions [12] combined with experimental verification [13]. The first step towards analyzing the impact of FMF nonlinearity is to solve the Manakov equation [3]. While the Split Step Fourier Method (SSFM) can solve the Manakov equation through many successive numerical simulation steps, it involves high computational complexity. On the other hand, perturbation-based methods can solve approximately the Manakov equation [14] and result in analytical formulations that predict FMF nonlinearity. The Gaussian Noise (GN) model is the most practical perturbation-based model which describes the nonlinear effects by an additive Gaussian noise source [26]. Existing publications on the GN model for FMF system can be separated in two main categories: those involving an integral-form [4],[10],[15],[20],[21] and those that are closed-form [6],[22]. These formulations all involve and Incoherent GN (IGN) model, which is based on the assumption that the nonlinear interference noise created at each span is accumulated incoherently at the receiver. This results in an under-estimation of the Nonlinear Interference (NLI) noise power spectral density [23]. Previously, we compared the complexity performance of the closed-form IGN model, integral-form IGN model, and SSFM for FMF [24]. Although the closed-form IGN model is very fast, it can only be applied to rectangular shaped Nyquist WDM [17], [18] and is accurate when the channel spacing is close to the symbol-rate [19]. Recently, we derived an integral-form coherent GN model for FMF [25] and compared its complexity performance with the IGN model. The GN model is only accurate in Gaussian-shaped constellations for multi-span links [26]. The Gaussian distribution assumption of the transmitted signal leads to an over-estimation of the NLI noise power in practical applications with modulation formats such as Phase Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM). This over-estimation, which is greater in the first spans (several *dB*) [16], in turn results in a 0.5 *dB* error on predicting the optimum launched power. As a consequence, the obtained gains with respect to the best equal power allocation can be conservative and fall within the accuracy of the GN model [16]. In Single Mode Fiber (SMF) systems, the Enhanced GN (EGN) model, presented by [16]–[18], removes the signal Gaussian distribution assumption and does not have the above mentioned limitations.

One of the main goals of MDM-WDM systems is to achieve reliable communication which is usually expressed in terms of the Signal to Noise Ratio (SNR) margin between the existing situation and the required error correction threshold. Channels and modes with the lowest SNR margins have the most likely failure. This failure can be minimized by the minimum SNR margin maximization. The SNR margin is directly related to the NLI noise, which is a dominating issue for achieving reliable communication [10]. The NLI noise is related to system parameters such as transmitted power and FM-EDFA gain; therefore, power allocation [27]-[31] and gain allocation [32] play essential roles in achieving reliable communication.

B. NOVELTIES AND CONTRIBUTIONS

Previous researches on FMF systems have led to the development of a closed-form IGN model, an integral-form IGN model, and an integral-form coherent GN model. Despite the improvements offered by the EGN model for simulating and analyzing SMF systems, no such EGN model exists for FMF systems. Therefore, in the first part of the paper, we derive an EGN model for FMF. This model, which can be used in both weak and strong coupling regimes, includes the first four dispersion terms as well as Carrier Phase Estimation (CPE). The significance of the first part of this paper include the following:

- Presenting the EGN model for MDM-WDM system, which can provide very accurate estimates of NLI noise power for different numbers of spans, launched power, and modulation format. Results from the ENG model show that we can remove the 0.1 *dB* and 0.9 *dB* SNR margins predicted by the integral-form IGN model [4] in the weak and strong coupling regimes, respectively, for the optimal (best equal) launched power per channel-mode. This is important in applications such as resource allocation, quality of transmission estimation, and optical performance monitoring.
- Proposing a formulation for NLI noise power wherein different system and link parameters can be selected independently (e.g., launch power of different channels and modes, FM-Erbium Doped Fiber Amplifier (FM-EDFA) gain of different spans, and fiber parameters of different spans). This then allows for the formulation of different marginal (or joint) resource allocation problems at the physical and network layers (e.g., joint optimized power and gain allocation).

The above referenced works use convex optimization for power allocation in SMF links and networks. In our recent work [24], we solved a power allocation problem for FMF links using convex optimization without considering any practical constraints, e.g., FM-EDFA saturation power. However, the joint optimized power and gain allocation has not been investigated in FMF links and networks. Therefore, in the second part of this paper, we use the proposed EGN model to address the problem of joint optimized power and gain allocation considering minimum SNR margin maximization. The contributions of the second part of this paper include the following:

- Deploying power allocation in a multi-node linear MDM-WDM network by considering constraints such as FM-EDFA saturation power and FM-EDFA

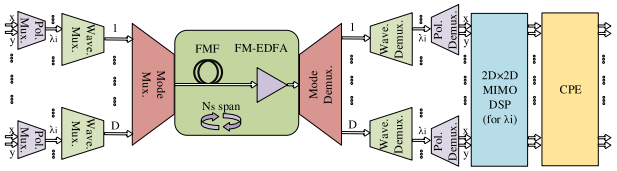

FIGURE 1. Schematic diagram of the MDM-WDM system.

maximum gain. We show that the results are adaptable to practical scenarios.

- Providing convex optimization solutions along with discussions on convergence speed and computational complexity. This demonstrates that the proposed algorithms can be implemented for joint optimized power and gain allocation.
- Preparing a comprehensive investigation over joint optimized power and gain allocation, optimal power allocation, and best equal power allocation in SMF-WDM and MDM-single channel systems. We show that the joint optimized power and gain allocation achieves 1.4 *dB* and 1.7 *dB* higher minimum SNR margin compared with the best equal power allocation in SMF-WDM and MDM-single channel scenarios, respectively.

Note that in *joint optimized power and gain allocation*, by *optimized power allocation*, we refer to the adjustment of the power of different channels and modes at the source node, and by *optimized gain allocation* we refer to adjusting the FM-EDFA gain of different spans considering equal FM-EDFA gain for different modes per span. In other words, we do not adjust the FM-EDFA gain for different channels and modes, rather we adjust the FM-EDFA gain for different spans, since according to the current progress of FM-EDFA, optimizing the FM-EDFA gain for different modes to a specific gain and power is still difficult to realize technically. In the case of SM-EDFAs, the change in gain of one polarization state will result in a change in gain of another polarization state (at the same time); such problems also exist in FM-EDFAs. The rest of this paper is organized as follows. In Section II, we present the system and signal models. Next, we derive the EGN model in Section III and present the problem statements in Section IV. We show the results and our analysis in Section V and conclude the paper in Section VI.

II. SYSTEM AND SIGNAL MODEL

A. SYSTEM MODEL

Fig. 1 shows the MDM-WDM system investigated in which the input data is a multiplexing of N_{ch} channels, D spatial modes, and 2 polarization modes. This link has N_s spans with length L_s, combined by an FM-EDFA at the end of each span for compensating the optical fiber loss. To minimize the FMF nonlinear effects, neither modal nor chromatic dispersion is compensated [4]. The Kerr nonlinearity produced by inter/intra channel and mode interactions is considered. We consider both weak and strong linear coupling among spatial modes to be applicable in both short-haul and long-haul links [3]–[6]. Moreover, a MIMO DSP at the receiver is considered for compensating FMF linear effects, and a CPE is used for carrier phase recovery [14].

B. SIGNAL MODEL

The following Ket notation represents the time domain of the optical signal launched into the FMF link as

$$|A(0,t)\rangle = \sum_{i_1=-\infty}^{\infty} \sum_{i_2=1}^{N_{ch}} \sum_{i_3=1}^{2D} \zeta_{\mathbf{i}} W_{Tx}^{i_2,i_3}(t - i_1 T_{i_2}) e^{j2\pi f_{i_2} t} |i_3\rangle, \tag{1}$$

where $\zeta_{\mathbf{i}}$ is the digital symbol (for example, Quadrature PSK (QPSK)) at time index i_1, WDM channel index i_2, polarization-spatial mode index i_3; $i_3 = 1, \ldots, 2D$, and $\mathbf{i} = [i_1, i_2, i_3]$. Moreover, $W_{Tx}^{i_2,i_3}(t - i_1 T_{i_2})$ is the transmitted pulse at time index i_1. Moreover, i_2 and i_3 represent the WDM channel and polarization-spatial mode index, respectively. T_{i_2} shows the symbol duration and f_{i_2} is the carrier frequency. $|i_3\rangle$ represents a one-hot vector wherein the i_3th element is one and the other elements are zero. Actually, i_3 is used to denote the polarization-spatial mode of the propagated signal. In other words, $i_3 \triangleq p', p$, where $p'; p' = x, y$ is the polarization mode index and $p; p = 1, \ldots, D$ represents the spatial mode index. It is obvious that i_3 takes values between 1 and $2D$, since each spatial mode is a multiplexing of 2 polarization modes. Therefore, the time domain representation of the optical signal launched into the FMF can be expressed as

$$|A(0,t)\rangle = \sum_{i_3=1}^{2D} A_{i_3}(0,t)|i_3\rangle \triangleq \begin{bmatrix} A_1(0,t) \\ A_2(0,t) \\ \ldots \\ A_{2D-1}(0,t) \\ A_{2D}(0,t) \end{bmatrix} \triangleq \begin{bmatrix} A_{x,1}(0,t) \\ A_{y,1}(0,t) \\ \ldots \\ A_{x,D}(0,t) \\ A_{y,D}(0,t) \end{bmatrix}, \tag{2}$$

and

$$\begin{aligned} &\langle A(0,t)| \\ &\triangleq \left[A_1^*(0,t)\ A_2^*(0,t)\ \ldots\ A_{2D-1}^*(0,t)\ A_{2D}^*(0,t)\right] \\ &\triangleq \left[A_{x,1}^*(0,t)\ A_{y,1}^*(0,t)\ \ldots\ A_{x,D}^*(0,t)\ A_{y,D}^*(0,t)\right], \end{aligned} \tag{3}$$

where $A_{i_3}(0,t) \triangleq A_{p',p}(0,t)$ represents the time domain of the propagated signal in i_3th polarization-spatial mode (i.e., p'th polarization mode and pth spatial mode) and can be expressed as

$$|A_{i_3}(0,t)\rangle = \sum_{i_1=-\infty}^{\infty} \sum_{i_2=1}^{N_{ch}} \zeta_{\mathbf{i}} W_{Tx}^{i_2,i_3}(t - i_1 T_{i_2}) e^{j2\pi f_{i_2} t} |i_3\rangle, \tag{4}$$

Considering (1), the propagated signal in the frequency domain can be expressed by

$$|\tilde{A}(0,f)\rangle = \sum_{\mathbf{i}} \zeta_{\mathbf{i}} |\tilde{W}_{Tx}^{\mathbf{i}}(f)\rangle, \tag{5}$$

where

$$|\tilde{W}^{\mathbf{i}}_{Tx}(f)\rangle \triangleq \tilde{W}^{\mathbf{i}}_{Tx}(f)|i_3\rangle = \tilde{W}^{i_2,i_3}_{Tx}(f-f_{i_2})e^{-j2\pi(f-f_{i_2})i_1T_{i_2}}|i_3\rangle. \quad (6)$$

III. EGN MODEL FORMULATION

The Manakov equation for the considered MDM-WDM link can be expressed as [14]

$$\frac{\partial|A_{i_3}(z,t)\rangle}{\partial z} = \mathcal{L}+\mathcal{N}, \quad (7)$$

where $\mathcal{L}$ and $\mathcal{N}$ represent the linear and nonlinear effects in the FMF. The linear effects $\mathcal{L}$ are given by [3]–[6]

$$\mathcal{L} = -\frac{\alpha_{i_3}}{2}|A_{i_3}(z,t)\rangle + j\beta_{0_{i_3}}|A_{i_3}(z,t)\rangle - \beta_{1_{i_3}}\frac{\partial|A_{i_3}(z,t)\rangle}{\partial t} - j\frac{\beta_{2_{i_3}}}{2}\frac{\partial^2|A_{i_3}(z,t)\rangle}{\partial t^2} - \frac{\beta_{3_{i_3}}}{6}\frac{\partial^3|A_{i_3}(z,t)\rangle}{\partial t^3}, \quad (8)$$

where α_{i_3} is the attenuation, $\beta_{m_{i_3}}$ is the mth order Taylor coefficient of the i_3th polarization-spatial mode propagation constant. The nonlinear effects $\mathcal{N}$ are comprised of [3]–[6]

$$\mathcal{N} = j\sum_{k_3,m_3,n_3=1}^{2D}\tilde{\gamma}_{i_3k_3m_3n_3}\Big(\langle A_{n_3}(z,t)|A_{m_3}(z,t)\rangle|A^*_{k_3}(z,t)\rangle + \langle A^*_{k_3}(z,t)|A_{m_3}(z,t)\rangle|A_{n_3}(z,t)\rangle\Big), \quad (9)$$

where

$$\tilde{\gamma}_{i_3k_3m_3n_3} = \begin{cases}\frac{4}{3}(\frac{2}{3})^{\delta_{i_3k_3}\delta_{i_3m_3}\delta_{i_3n_3}}f_{i_3k_3m_3n_3}\gamma, & \text{weak coupling}\\ \kappa\gamma, & \text{strong coupling},\end{cases}$$

with

$$\kappa = \sum_{\substack{k_3,m_3,n_3\le i_3\\ i_3,k_3,m_3,n_3\in\{1,2,\dots,2D\}}}\frac{32}{2^{\delta_{i_3k_3}\delta_{i_3m_3}\delta_{i_3n_3}}}\frac{f_{i_3k_3m_3n_3}}{6D(2D+1)},$$

and

$$f_{i_3k_3m_3n_3} = \frac{A_{eff}}{\sqrt{I_{i_3}I_{k_3}I_{m_3}I_{n_3}}}\times\iint F_{i_3}(x,y)F_{k_3}(x,y)F_{m_3}(x,y)F_{n_3}(x,y)dxdy,$$

γ is the Kerr nonlinearity coefficient, $I_{i_3}=\iint F^2_{i_3}(x,y)dxdy$, $F_{i_3}(x,y)$ is the spatial profile of i_3th mode, and A_{eff} is the effective area of the fundamental mode [3]–[6]. The first-order perturbation approximation of the solution to the Manakov equation expresses the received signal as [14]

$$|A(z,t)\rangle \simeq e^{\mathcal{L}z}|A(0,t)\rangle + \int_0^z e^{\mathcal{L}(z-\xi)}\mathcal{N}\Big(e^{\mathcal{L}\xi}|A(0,t)\rangle\Big)d\xi. \quad (10)$$

After compensating the linear effects of the FMF using MIMO DSP, the received signal can be expressed as

$$|A_R\rangle \simeq |A(0,t)\rangle + \int_0^z e^{-\mathcal{L}\xi}\mathcal{N}(e^{\mathcal{L}\xi}|A(0,t)\rangle)d\xi, \quad (11)$$

The $\mathcal{L}$ can be described in the frequency domain by its Fourier transform $\mathcal{F}(e^{\mathcal{L}z}) = |e^{\nu(z,f)}\rangle$ where

$$\nu_{i_3}(z,f) = -\int_0^z(\alpha_{i_3}(\xi)+j\beta_{i_3}(\xi,f))d\xi. \quad (12)$$

Moreover, $\beta_{i_3}(z,f)$ can be calculated as

$$\beta_{i_3}(z,f) = \beta_{0_{i_3}} + \beta_{1_{i_3}}(2\pi f) + \frac{\beta_{2_{i_3}}}{2}(2\pi f)^2 + \frac{\beta_{3_{i_3}}}{6}(2\pi f)^3. \quad (13)$$

The second term in (11) represents the NLI noise which, by considering (12), can be simplified in the frequency domain as

$$|\tilde{n}(f)\rangle = -j\iint_{-\infty}^{-\infty}|\eta(f,f_1,f_2)\rangle\langle\tilde{A}(f+f_1+f_2)|\tilde{A}(f+f_2)\rangle|\tilde{A}(f+f_1)\rangle df_1df_2, \quad (14)$$

where

$$\begin{aligned}\eta_{i_3}(f,f_1,f_2) &\triangleq \sum_{s=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\,L_n)^{1/2}\sum_{k_3,m_3,n_3}\tilde{\gamma}_{i_3,k_3,m_3,n_3}\\ &\times\int_{z_{s-1}}^{z_s}e^{\nu_{n_3}(\xi,f+f_1)+\nu_{m_3}(\xi,f+f_2)+\nu^*_{k_3}(\xi,f+f_1+f_2)-\nu_{i_3}(\xi,f)}d\xi\\ &\triangleq \sum_{s=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^3\prod_{n=s}^{N_s}(G_n\,L_n)\eta_{i_3,s}(f,f_1,f_2),\end{aligned} \quad (15)$$

with G_n being FM-EDFA amplifier gain and L_n being the loss of the nth fiber span. Using matched filtering on the received signal gives

$$\int_{-\infty}^{\infty}\langle g^{\mathbf{i}}(f)|\tilde{A}(f)\rangle R_{\mathbf{i}}df, \quad (16)$$

where $|g^{\mathbf{i}}(f)\rangle = g^{\mathbf{i}}(f)|i_3\rangle$ is the spectral shape of transmitted pulse on the i_2th channel and i_3th polarization-spatial mode, which has been normalized such that $\int_{-\infty}^{+\infty}g^{\mathbf{i}}(f)df = 1$. $R_{\mathbf{i}}$ is the symbol rate of transmitted pulse on the i_2th channel and i_3th polarization-spatial mode. Accordingly, the NLI noise takes the form of (17), as shown at the bottom of the page, at the receiver which can be written in (18), as shown at the bottom of the next page.

$$n_{\mathbf{i}}(f) = -j\sum_{s=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\,L_n)^{1/2}\sum_{\mathbf{k},\mathbf{m},\mathbf{n}}\zeta^*_{\mathbf{k}}\zeta_{\mathbf{m}}\zeta_{\mathbf{n}}\iiint_{-\infty}^{\infty}|\eta_s(f,f_1,f_2)\rangle\langle\tilde{W}^{\mathbf{k}}_{Tx}(f+f_1+f_2)|\tilde{W}^{\mathbf{m}}_{Tx}(f+f_2)\rangle\langle g^{\mathbf{i}}(f)|\tilde{W}^{\mathbf{n}}_{Tx}(f+f_1)\rangle R_{\mathbf{i}}df_1df_2df \quad (17)$$

The variance of the NLI noise is derived in (19), as shown at the bottom of the page, which depends on six infinite integration/summations where the following Poisson summation helps to drop some of the interactions/summations:

$$\sum_{k=-\infty}^{\infty} e^{jkf2\pi T} = T \sum_{k=-\infty}^{\infty} \delta(f - k/T). \tag{20}$$

By considering the *sinc* pulses with finite bandwidth interacting with only one Dirac delta function, summation over the time index can be given by (19). Furthermore, equating the arguments with equal atoms results in

$$\begin{aligned} f + f_1 + f_2 &= v + v_1 + v_2 \\ f + f_2 &= v + v_2 \\ f + f_1 &= v + v_1, \end{aligned} \tag{21}$$

and accordingly

$$\begin{aligned} f &= v \\ f_1 &= v_1 \\ f_2 &= v_2, \end{aligned} \tag{22}$$

which yields to simplify three integrals with v, v_1, and v_2.

The Kerr nonlinearity is cubic; therefore, the product $E[n_{\mathbf{i}}(f)n_{\mathbf{i}}^*(f)]$ depends on the product of six atoms. Note that only combinations with an equal number of conjugate/non-conjugate pairs are non-zero. In addition, it should be noted that $\zeta_{\mathbf{i}}$ are independent and identically distributed random variables with zero mean and unit variance. The considered CPE removes the average phase at the receiver (ϕ) [14]. In a perturbative frame, this corresponds to work with the following nonlinear interference noise

$$n'_{\mathbf{i}} = n_{\mathbf{i}} + j\phi\zeta_{\mathbf{i}}. \tag{23}$$

The NLI noise variance in the EGN model can be interpreted as a summation over the Second-Order Noise (SON) which is usually called the GN contribution, Fourth-Order Noise

$$\begin{aligned} n_{\mathbf{i}}(f) = & -j\sum_{s=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\sum_{\mathbf{k},\mathbf{m},\mathbf{n}}\zeta_{\mathbf{k}}^*\zeta_{\mathbf{m}}\zeta_{\mathbf{n}}\iiint_{-\infty}^{\infty}|\eta_s(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_2)g^{\mathbf{i}^*}(f) \\ & \tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle k_3|m_3\rangle\langle i_3|n_3\rangle R_{\mathbf{i}}df_1df_2df \end{aligned} \tag{18}$$

$$\begin{aligned} E[n_{\mathbf{i}}(f)n_{\mathbf{i}}^*(f)] = & \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{\mathbf{kmnljo}}E[\zeta_{\mathbf{k}}^*\zeta_{\mathbf{m}}\zeta_{\mathbf{n}}\zeta_{\mathbf{l}}\zeta_{\mathbf{j}}^*\zeta_{\mathbf{o}}^*]\int\cdots\int_{\infty}^{\infty} \\ & \langle\eta_s(f,f_1,f_2)|\eta_{s'}(v,v_1,v_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1) \\ & \tilde{W}_{Tx}^{\mathbf{j}*}(v+v_2)\tilde{W}_{Tx}^{\mathbf{l}}(v+v_1+v_2)\tilde{W}_{Tx}^{\mathbf{o}*}(v+v_1)g^{\mathbf{i}}(v)\langle k_3|m_3\rangle\langle i_3|n_3\rangle\langle j_3|l_3\rangle\langle o_3|i_3\rangle R_{\mathbf{i}}df_1df_2dfdv_1dv_2dv \end{aligned} \tag{19}$$

$$\begin{aligned} \sigma^2_{EGN,i_2,p} = & \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\Bigg[3/4\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}P_{k_2,q}P_{m_2,q} \\ & \times P_{n_2,p}X^a_{i_2,p}(k_2,m_2,n_2,q) + 1/4\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(P^2_{k_2,q}P_{n_2,p}5X^b_{i_2,p}(k_2,k_2,n_2,q) + P_{k_2,p}P_{k_2,q}P_{n_2,q} \\ & \times X^c_{i_2,p}(k_2,n_2,k_2,q)) + 1/4\sum_{n_2}\kappa_3^{(n_2)}P^2_{n_2,q}P_{n_2,p}X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg] \end{aligned} \tag{25}$$

$$\begin{aligned} M_{i_2,p} = & \Bigg(P_{i_2,p}\prod_{n=1}^{N_s}(G_n L_n)/\Bigg(\prod_{n=1}^{N_s}(G_n L_n)(F(G_{BA}-1)h\nu B_{i_2}) + \sum_{s=1}^{N_s}[(F(G_s-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(G_n L_n)] + \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2} \\ & \prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\Bigg[3/4\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}P_{k_2,q}P_{m_2,q}P_{n_2,p}X^a_{i_2,p}(k_2,m_2,n_2,q) \\ & +1/4\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(P^2_{k_2,q}P_{n_2,p}5X^b_{i_2,p}(k_2,k_2,n_2,q) + P_{k_2,p}P_{k_2,q}P_{n_2,q}X^c_{i_2,p}(k_2,n_2,k_2,q)) + 1/4\sum_{n_2}\kappa_3^{(n_2)}P^2_{n_2,q}P_{n_2,p} \\ & \times X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg] + \sigma^2_{RxN}\Bigg)\Bigg)/SNR^{req}_{i_2,p} \end{aligned} \tag{28}$$

(FON), and Higher-Order Noise (HON) variances, *i.e.*,

$$\sigma^2_{EGN} = \sigma^2_{GN} + \sigma^2_{FON} + \sigma^2_{HON}. \tag{24}$$

It is shown in Appendix A that the NLI noise variance of the whole link becomes equal to (25), as shown at the bottom of the previous page, where $P_{i_2,p} \triangleq P_{i_2,p,1}$ is the launch power at the first span and $P_{i_2,p,s} = P_{i_2,p,s-1}(G_s\, L_s)$ is the launch power at the sth span. The power at each span input is the multiplication of the power of the previous span by the loss and FM-EDFA gain of that span. Note that (25) represents both self- and cross-channel as well as intra- and inter-modal interactions.

IV. PROBLEM STATEMENT

Based on the proposed notations in the previous section, the SNR of i_2th channel and pth mode can be expressed as [26]

$$SNR_{i_2,p} = \frac{P_{i_2,p}\prod_{n=1}^{N_s}(G_n\, L_n)}{\sigma^2_{ASE} + \sigma^2_{EGN,i_2,p} + \sigma^2_{RxN}}, \tag{26}$$

where σ^2_{RxN} is the receiver noise power. Moreover, the variance of the Amplified Spontaneous Emission (ASE) noise in the receiver can be expressed as

$$\begin{aligned}\sigma^2_{ASE} = &\prod_{n=1}^{N_s}(G_n\, L_n)(F(G_{BA}-1)h\nu B_{i_2}) \\ &+\sum_{s=1}^{N_s}[(F(G_s-1)h\nu B_{i_2}) \\ &\prod_{n=s+1}^{N_s}(G_n\, L_n)],\ ]\end{aligned} \tag{27}$$

where F is the amplifier noise figure, G_{BA} is the booster amplifier gain, h is Plank's constant, and ν is the central frequency. The SNR margin of the i_2th channel and pth mode can be defined in (28), as shown at the bottom of the previous page, where $M_{i_2,p}$ denotes SNR margin of the i_2th channel and pth mode, and $SNR^{req}_{i_2,p}$ is the required SNR of the i_2th channel and pth mode. Therefore, the minimum SNR margin maximization problem can be expressed in (29), as shown at the bottom of the page, where the second constraint means that the total power at the s-th FM-EDFA should be less than the saturation power of the s-th FM-EDFA, and the third constraint means that the s-th FM-EDFA gain should be less than the maximum possible gain. The optimization problem (29) is equivalent to the optimization problem (30), as shown at the bottom of the next page, as the min-max of a function is equivalent to the max-min of its inverse. (30) is a non-convex optimization problem. To solve this problem, we replace $P_{i_2,p}$, G_i with $e^{\hat{P}_{i_2,p}}$, e^{g_i} in (30) and note that $log(x)$ is a monotonic function in x; thus, we get in (31), as shown at the bottom of the next page, with the same minimum as (30). By defining the slack variable β, (31) can be rewritten in (32), as shown at the bottom of the next page.

We use the gradient descent algorithm in vector form to solve (32). 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 as $\mathbf{B} = [B_1, B_1, \ldots, B_1, B_2, B_2, \ldots, B_2, \ldots, B_{N_{ch}}, B_{N_{ch}}, \ldots, B_{N_{ch}}]$ in which each B_n has been repeated D times. Also, let X be a $N_{ch} \times D \times N_{ch} \times N_{ch} \times N_{ch} \times D$ dimensional tensor with elements $X^{(.)}_{l_2,p}(k_2, m_2, n_2, q)$. To match the latter dimensions with $\mathbf{p}$, we define a $N_{ch}D \times N_{ch}D \times N_{ch}D \times N_{ch}D$ tensor, H, whose elements, $H^{(.)}_l(l_1, l_2, l_3)$, are equal to $X^{(.)}_{l_2,p}(k_2, m_2, n_2, q)$ in different subscripts. Thus, (32) can be expressed as (33). (33) is a convex optimization problem (see Appendix B) and can be solved using many different methods, e.g., the Bisection method [33]. The Bisection method converts the main problem into a feasibility problem by selecting a region and choosing a candidate for the objective function. The feasibility problem can then be solved using the Lagrangian method [34]. In each step, the region boundaries are updated based on the obtained solution

$$\begin{aligned}\max_{G_i,P_{i_2,p}}\ \min_{i_2,p,i}\ & P_{i_2,p}\prod_{n=1}^{N_s}(G_n\, L_n)/\Bigg(\prod_{n=1}^{N_s}(G_n\, L_n)(F(G_{BA}-1)h\nu B_{i_2}) + \sum_{s=1}^{N_s}[(F(G_s-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(G_n\, L_n)] + \sum_{s,s'=1}^{N_s} \\ & \prod_{n=1}^{s-1}(G_n\, L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\, L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n\, L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n\, L_n)^{1/2}\sum_{q=1}^{D}\Bigg[3/4\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}P_{k_2,q}P_{m_2,q}P_{n_2,p} \\ & X^a_{i_2,p}(k_2,m_2,n_2,q) + 1/4\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(P^2_{k_2,q}P_{n_2,p}5X^b_{i_2,p}(k_2,k_2,n_2,q) + P_{k_2,p}P_{k_2,q}P_{n_2,q}X^c_{i_2,p}(k_2,n_2,k_2,q)) \\ & +1/4\sum_{n_2}\kappa_3^{(n_2)}P^2_{n_2,q}P_{n_2,p}X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg] + \sigma^2_{RxN}\Bigg)\frac{1}{SNR^{req}_{i_2,p}}\end{aligned}$$

$$s.t.\ \begin{cases} P_{i_2,p,s} = P_{i_2,p,s-1}G_s\, L_s \\ \sum_{i_2,p}P_{i_2,p}\prod_{n=1}^{s-1}(G_n\, L_n) \le P_{sat}^{FM-EDFA_s} \\ G_s \le G_{max}^{FM-EDFA_s} \end{cases} \tag{29}$$

for the feasibility problem from the previous step. In this manner, the Bisection method converges to the optimum objective.

Algorithm 1 summarizes the Bisection method for solving (33). The first step in this algorithm is to define appropriate upper (u) and lower (l) bounds for the search region

$$\min_{P_{i_2,p},G_i} \max_{i_2,p,i} SNR^{req}_{i_2,p}\Bigg(\prod_{n=1}^{N_s}(G_n\,L_n)(F(G_{BA}-1)h\nu B_{i_2})+\sum_{s=1}^{N_s}[(F(G_s-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(G_n\,L_n)]+\sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}$$
$$(G_n\,L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n\,L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n\,L_n)^{1/2}\sum_{q=1}^{D}\Bigg[3/4\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}P_{k_2,q}P_{m_2,q}P_{n_2,p}X^a_{i_2,p}(k_2,m_2,n_2,q)$$
$$+1/4\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(P^2_{k_2,q}P_{n_2,p}5X^b_{i_2,p}(k_2,k_2,n_2,q)+P_{k_2,p}P_{k_2,q}P_{n_2,q}X^c_{i_2,p}(k_2,n_2,k_2,q))+1/4\sum_{n_2}\kappa_3^{(n_2)}P^2_{n_2,p}$$
$$\times P_{n_2,q}X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg]+\sigma^2_{RxN}\Bigg)\Big/\Bigg(P_{i_2,p}\prod_{n=1}^{N_s}(G_n\,L_n)\Bigg)$$
$$s.t.\ \begin{cases} P_{i_2,p,s}=P_{i_2,p,s-1}G_s\,L_s \\ \sum_{i_2,p}P_{i_2,p}\prod_{n=1}^{s-1}(G_n\,L_n)\le P^{FM-EDFA_s}_{sat} \\ G_s\le G^{FM-EDFA_s}_{max}\end{cases} \tag{30}$$

$$\min_{\hat{P}_{i_2,p},g_i} \max_{i_2,p,i} log(SNR^{req}_{i_2,p})+log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{i_2})+\sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)]+\sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}$$
$$(e^{g_n}\,L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}\,L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}\,L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}\,L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{q=1}^{D}\sum_{k_2,m_2,n_2=1}^{N_{ch}}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}e^{\hat{P}_{k_2,p}+\hat{P}_{m_2,q}+\hat{P}_{n_2,q}}$$
$$3X^a_{i_2,p}(k_2,m_2,n_2,q)+\frac{1}{4}\sum_{q=1}^{D}\sum_{k_2,n_2=1}^{N_{ch}}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(e^{2\hat{P}_{k_2,q}+\hat{P}_{n_2,p}}5X^b_{i_2,p}(k_2,k_2,n_2,q)+e^{\hat{P}_{k_2,p}+\hat{P}_{k_2,q}+\hat{P}_{n_2,q}}$$
$$X^c_{i_2,p}(k_2,n_2,k_2,q))+\frac{1}{4}\sum_{q=1}^{D}\sum_{n_2=1}^{N_{ch}}\kappa_3^{(n_2)}e^{2\hat{P}_{n_2,q}+\hat{P}_{n_2,p}}X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg]+\sigma^2_{RxN}\Bigg)-(\hat{P}_{i_2,p}+\sum_{n=1}^{N_s}g_n log(L_n))$$
$$s.t.\ \begin{cases} \hat{P}^s_{i_2,p}=\hat{P}^{s-1}_{i_2,p}+g_s+log(L_s) \\ \sum_{i_2,p}e^{\hat{P}_{i_2,p}}\prod_{n=1}^{s-1}(e^{g_n}L_n)\le P^{FM-EDFA_s}_{sat} \\ g_s\le log(G^{FM-EDFA_s}_{max})\end{cases} \tag{31}$$

$$\min_{\beta,\hat{P}_{i_2,p},g_i}\ \beta$$
$$s.t.\ \begin{cases} log(SNR^{req}_{i_2,p})+log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{i_2})+\sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)]+\sum_{s,s'=1}^{N_s} \\ \prod_{n=1}^{s-1}(e^{g_n}\,L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}\,L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}\,L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}\,L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{q=1}^{D}\sum_{k_2,m_2,n_2=1}^{N_{ch}}\kappa_1^{(k_2)}\kappa_1^{(m_2)} \\ \kappa_1^{(n_2)}e^{\hat{P}_{k_2,p}+\hat{P}_{m_2,q}+\hat{P}_{n_2,q}}3X^a_{i_2,p}(k_2,m_2,n_2,q)+\frac{1}{4}\sum_{q=1}^{D}\sum_{k_2,n_2=1}^{N_{ch}}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(e^{2\hat{P}_{k_2,q}+\hat{P}_{n_2,p}}5X^b_{i_2,p}(k_2,k_2,n_2,q) \\ +e^{\hat{P}_{k_2,p}+\hat{P}_{k_2,q}+\hat{P}_{n_2,q}}X^c_{i_2,p}(k_2,n_2,k_2,q))+\frac{1}{4}\sum_{q=1}^{D}\sum_{n_2=1}^{N_{ch}}\kappa_3^{(n_2)}e^{2\hat{P}_{n_2,q}+\hat{P}_{n_2,p}}X^d_{i_2,p}(n_2,n_2,n_2,q)\Bigg]+\sigma^2_{RxN}\Bigg) \\ -(\hat{P}_{i_2,p}+\sum_{n=1}^{N_s}g_n log(L_n))\le\beta \\ \hat{P}^s_{i_2,p}=\hat{P}^{s-1}_{i_2,p}+g_s+log(L_s) \\ \sum_{i_2,p}e^{\hat{P}_{i_2,p}}\prod_{n=1}^{s-1}(e^{g_n}L_n)\le P^{FM-EDFA_s}_{sat} \\ g_s\le log(G^{FM-EDFA_s}_{max})\end{cases} \tag{32}$$

for β. The defined upper bound is assigned to β, and the problem (30) is converted to a feasibility problem. The feasibility problem is solved using the Lagrangian method. If the defined upper bound is lower than the optimal solution for β, the feasibility problem does not have a solution. In other words, the feasible set is empty and a higher upper bound should be used. The defined lower bound can be tested and adjusted in the same manner.

In the second step, the upper bound $(u + l)/2$ is assigned to β. As with the first step, we test whether $(u + l)/2$ is the upper or lower bound of the feasible set. Then, the upper and lower bounds are updated. The second step is repeated until we reach convergence.

At each iteration of Algorithm 1, the feasibility problem is solved by the Lagrange duality method as summarized in Algorithm 2. Furthermore, the second constraint can be relaxed since it is satisfied by the objective function. The Lagrangian function of (33), as shown at the bottom of the page, is given by (34), as shown at the bottom of the page, where λ_l, μ_s, and $\nu_s \in R^+$ are the Lagrangian multipliers. Accordingly, the Lagrangian dual function of (33) can be expressed as (35), as shown at the bottom of the page.

$$\min_{\beta,\hat{P}_l,g_s} \ \beta$$

$$s.t. \begin{cases} \Bigg[log(SNR_l^{req}) + log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_l) + \sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{l_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)] + \sum_{s,s'=1}^{N_s} \\ \prod_{n=1}^{s-1}(e^{g_n}L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_1^{\prime(l_1)}\kappa_1^{\prime(l_2)}\kappa_1^{\prime(l_3)}e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}} \\ 3H_l^a(l_1,l_2,l_3) + \frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_2^{\prime(l_1)}\kappa_1^{\prime(l_2)}(e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}5H_l^b(l_1,l_1,l_2) + e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}H_l^c(l_1,l_2,l_3)) + \frac{1}{4}\sum_{l_1,l_2=1}^{DN_{ch}}\kappa_3^{\prime(l_1)} \\ e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}H_l^d(l_1,l_1,l_2)\Bigg] + \sigma_{RxN}^2\Bigg) - (\hat{P}_l + \sum_{n=1}^{N_s} g_n log(L_n))\Bigg] \le \beta \\ \hat{P}_l^s = \hat{P}_l^{s-1} + g_s + log(L_s) \\ \sum_l e^{\hat{P}_l}\prod_{n=1}^{s-1}(e^{g_n}L_n) \le P_{sat}^{FM-EDFA_s} \\ g_s \le log(G_{max}^{FM-EDFA_s}) \end{cases} \tag{33}$$

$$\beta + \sum_{l=1}^{DN_{ch}}\lambda_l\Bigg(\Bigg[log(SNR_l^{req}) + log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{l_2}) + \sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{l_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)] + \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}$$
$$(e^{g_n}L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_1^{\prime(l_1)}\kappa_1^{\prime(l_2)}\kappa_1^{\prime(l_3)}e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}3H_l^a(l_1,l_2,l_3) + \frac{1}{4}$$
$$\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_2^{\prime(l_1)}\kappa_1^{\prime(l_2)}(e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}5H_l^b(l_1,l_1,l_2) + e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}H_l^c(l_1,l_2,l_3)) + \frac{1}{4}\sum_{l_1,l_2=1}^{DN_{ch}}\kappa_3^{\prime(l_1)}e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}H_l^d(l_1,l_1,l_2)\Bigg] + \sigma_{RxN}^2\Bigg)$$
$$-(\hat{P}_l + \sum_{n=1}^{N_s} g_n log(L_n))\Bigg] - \beta\Bigg) + \sum_{s=1}^{N_s}\mu_s\Bigg(\sum_l e^{\hat{P}_l}\prod_{n=1}^{s-1}(e^{g_n}L_n) - P_{sat}^{FM-EDFA_s}\Bigg) + \sum_{s=1}^{N_s}\nu_s(g_s - log(G_{max}^{FM-EDFA_s})) \tag{34}$$

$$\inf_{\hat{P}_l,g_i} \beta + \sum_{l=1}^{DN_{ch}}\lambda_l\Bigg(\Bigg[log(SNR_l^{req}) + log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{l_2}) + \sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{l_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)] + \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}$$
$$(e^{g_n}L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_1^{\prime(l_1)}\kappa_1^{\prime(l_2)}\kappa_1^{\prime(l_3)}e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}3H_l^a(l_1,l_2,l_3) + \frac{1}{4}$$
$$\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_2^{\prime(l_1)}\kappa_1^{\prime(l_2)}(e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}5H_l^b(l_1,l_1,l_2) + e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}H_l^c(l_1,l_2,l_3)) + \frac{1}{4}\sum_{l_1,l_2=1}^{DN_{ch}}\kappa_3^{\prime(l_1)}e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}H_l^d(l_1,l_1,l_2)\Bigg] + \sigma_{RxN}^2\Bigg)$$
$$-(\hat{P}_l + \sum_{n=1}^{N_s} g_n log(L_n))\Bigg] - \beta\Bigg) + \sum_{s=1}^{N_s}\mu_s\Bigg(\sum_l e^{\hat{P}_l}\prod_{n=1}^{s-1}(e^{g_n}L_n) - P_{sat}^{FM-EDFA_s}\Bigg) + \sum_{s=1}^{N_s}\nu_s(g_s - log(G_{max}^{FM-EDFA_s})) \tag{35}$$

Algorithm 1: Bisection Method to Solve Convex Optimization Problem (33)

Initialization: upper bound $u = 100$, and lower bound $l = -10$;
$\beta \leftarrow u$;
Solve convex problem (33) by Lagrangian method;
if $\hat{P}^{*(t)} == NAN$ **then** break;
$\beta \leftarrow l$;
Solve convex problem (33) by Lagrangian method;
if$\hat{P}^{*(t)} == NAN$ **then** break;
while $u - l \leq \epsilon$ **do**
  $\beta \leftarrow (u + l)/2$;
  Solve convex problem (33) by Lagrangian method;
  if $\hat{P}^{*(t)} == NAN$ **then** $l \leftarrow \beta$ **else** $u \leftarrow \beta$
end

Algorithm 2: Lagrangian Duality Method to Solve the Convex Problem (35)

Initialization: iteration counter $t = 0$, step size parameter $a > 0$, $b > 0$, $c > 0$, and $\lambda^{(0)} \succeq 0$, $\mu^{(0)} \succeq 0$, $\nu^{(0)} \succeq 0$;
while *achieving convergence* **do**
  Solve convex problem (35) with fixed λ, μ, and ν values, and obtain optimized power $\hat{P}^{*(t)}$, and optimal gain $g^{*(t)}$;
  $\lambda^{(t+1)} = \left[\lambda^{(t)} - a\Delta_\lambda\right]^+$;
  $\mu^{(t+1)} = \left[\mu^{(t)} - b\Delta_\mu\right]^+$;
  $\nu^{(t+1)} = \left[\nu^{(t)} - c\Delta_\nu\right]^+$;
  update $t = t + 1$
end

(35) is a convex problem with respect to $\hat{P}_l$, g_i, since the dual problem is a convex optimization problem [34]. Note that at each iteration of Algorithm 2, λ_l, μ_s, and ν_s are updated based on the derivative of (35) with respect to λ_l, μ_s, and ν_s which are shown in (36).

The convergence proof and complexity analysis of Algorithms 1 and 2 are presented in Appendices C and D, respectively.

V. SIMULATION RESULTS

In the first part of this section, the accuracy of the proposed EGN model is examined for both weak and strong linear coupling regimes using the well-known SSFM. The second part of this section presents the simulation results of the proposed joint optimized power and gain allocation problem. Note that in the second part, we consider strong linear coupling as the link ranges are long. The simulation parameters and their values are presented in Tables 1, 2, and 3.

TABLE 1. Simulation parameters.

Coefficient	Symbol	Value
Symbol rate	R_n	$32\ GBaud$
Channel spacing	B_w	$50\ GHz$
Channel bandwidth	B_n	$32\ GHz$
Center frequency	ν	$1550\ nm$
Noise figure	F	$5\ dB$
FM-EDFA maximum gain	G_{max}^{EDFA}	$30\ dB$
FM-EDFA saturation power	P_{sat}^{EDFA}	$25\ dBm$
Receiver noise power	σ^2_{Rxn}	$-28\ dBm$
Booster amplifier gain	G_{BA}	$20\ dB$

A. ACCURACY OF THE PROPOSED EGN MODEL FORMULATION

In this section, the signal propagation in weak and strong linear coupling is simulated by approximating the output of the Manakov equation (7) using the well-known SSFM with

$$
\begin{aligned}
\Delta_{\lambda_l} = \Bigg[& log(SNR_l^{req}) + log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{l_2}) + \sum_{s=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{l_2}) \prod_{n=s+1}^{N_s}(e^{g_n}L_n)] + \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(e^{g_n}L_n)^{3/2} \\
& \prod_{n=s}^{N_s}(e^{g_n}L_n)^{1/2}\prod_{n=1}^{s'-1}(e^{g_n}L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_1^{\prime(l_1)}\kappa_1^{\prime(l_2)}\kappa_1^{\prime(l_3)}e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}3H_l^a(l_1,l_2,l_3) + \frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_2^{\prime(l_1)}\kappa_1^{\prime(l_2)} \\
& (e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}5H_l^b(l_1,l_1,l_2) + e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}H_l^c(l_1,l_2,l_3)) + \frac{1}{4}\sum_{l_1,l_2=1}^{DN_{ch}}\kappa_3^{\prime(l_1)}e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}H_l^d(l_1,l_1,l_2)\Bigg] \\
& +\sigma^2_{RxN}\Bigg) - (\hat{P}_l + \sum_{n=1}^{N_s}g_n log(L_n))\Bigg] \\
& -\beta \\
\Delta_{\mu_s} = & \sum_l e^{\hat{P}_l}\prod_{n=1}^{s-1}(e^{g_n}L_n) - P_{sat}^{FM-EDFA_s} \\
\Delta_{\nu_s} = & g_s - log(G_{max}^{FM-EDFA_s})
\end{aligned}
\tag{36}
$$

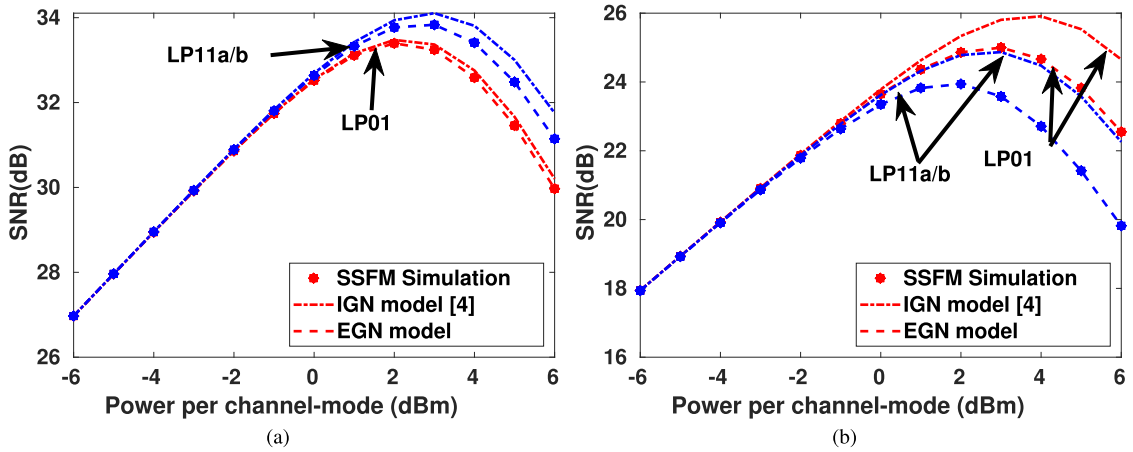

FIGURE 2. SNR versus launched power per channel-mode, based on the proposed EGN model, IGN model [4], and the SSFM simulation, for a) weak coupling ($N_s = 1, L_s = 80\ km$) and b) strong coupling ($N_s = 8, L_s = 80\ km$), considering MDM-WDM system ($D = 3, N_{ch} = 3$).

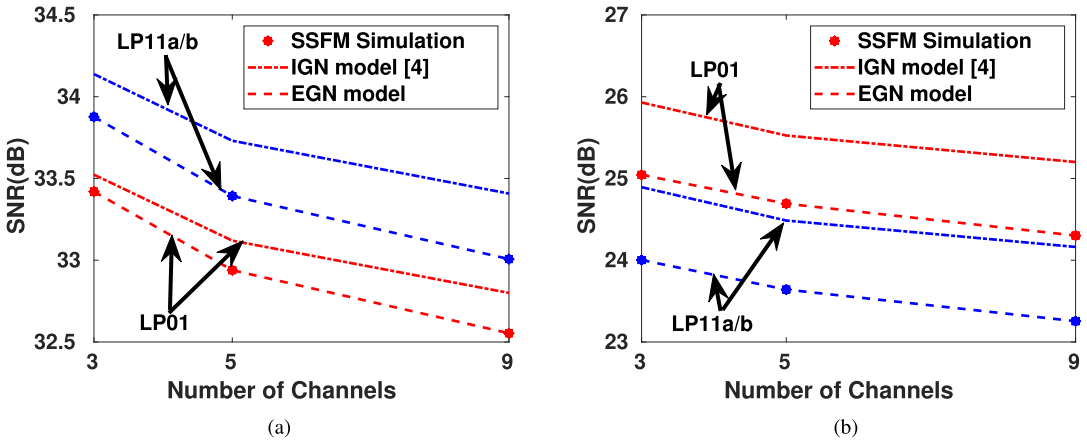

FIGURE 3. SNR versus number of channels for a) weak ($N_s = 1, L_s = 80\ km$) and b) strong ($N_s = 8, L_s = 80\ km$) coupling, considering MDM-WDM system ($D = 3$).

TABLE 2. Nonlinear coupling coefficient ($\gamma f_{pq}(1/W/km)$) [36].

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 3. 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)) [36].

	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

logarithmic step-size [35] in the Python/Tensorflow environment.

As an example, we consider an MDM-WDM system with ($D = 3, N_{ch} = 3$). Fig. 2 compares the SNR versus launch power per channel-mode, calculated using the proposed EGN model, the IGN model presented in [4], and the SSFM simulation, for weak coupling ($N_s = 1, L_s = 80\ km$), Fig. 2a, and strong coupling ($N_s = 8, L_s = 80\ km$), Fig. 2b. Note that only the SNR of the central channel is plotted, and that QPSK modulation is considered in the SSFM simulation. The results highlight the accuracy of the proposed EGN model in both weak and strong coupling regimes at all power levels. The IGN model [4] matches with the SSFM in the linear regime (i.e., at low power) and pseudo-linear regime (i.e., moderate power). However, the accuracy of the IGN model decreases when the power is increased to the nonlinear regime (high power). Considering the optimum launch power per channel-mode, the difference between the IGN and SSFM is almost 0.1 dB and 0.9 dB for weak coupling and strong coupling regimes, respectively.

Figs. 3a and 3b show the SNR versus number of channels for weak ($N_s = 1, L_s = 80\ km$) and strong ($N_s = 8, L_s = 80\ km$) coupling, respectively. In both plots, we consider the optimum (best equal) launched power per channel-mode in an MDM-WDM system with $D = 3$ spatial modes. The

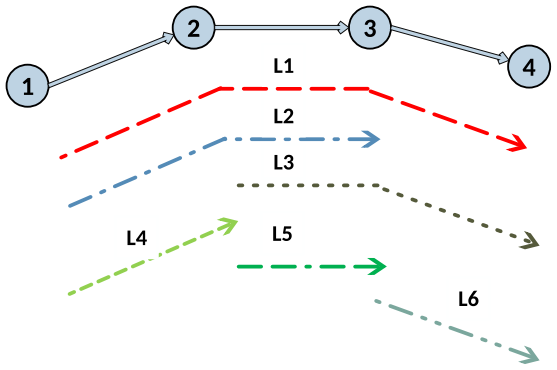

FIGURE 4. A 4-node linear network with 6 lightpaths.

proposed EGN model and SSFM simulation are in agreement for differing number of channels. However, the difference between the IGN model and the SSFM simulation is 0.1 *dB* and 0.9 *dB* in the weak and strong coupling regimes, respectively.

B. MINIMUM SNR MARGIN MAXIMIZATION

In this section, the results of joint optimized power and gain allocation based on minimum SNR margin maximization are presented. Three scenarios are compared including a) best equal power, b) optimized power, and c) joint optimized power and gain. In the first scenario, equal powers are considered for the different channels and modes with equal FM-EDFA gain in all spans. It is worth mentioning that the FM-EDFA gain is equal to the span loss. In the second scenario, different powers are allocated to the different channels and modes with equal FM-EDFA gain in all spans. In the third case, the allocated powers to each channel and mode are different. Moreover, the FM-EDFA gain for each span is optimized separately. To maximize the minimum SNR margin based on the optimized power and gain allocation, the 4-node linear network [37] with 6 lightpaths shown in Fig. 4 is considered. Moreover, the SMF-WDM ($D = 1$, $N_{ch} = 11$) and MDM-single channel ($D = 4, N_{ch} = 1$) systems are considered [3], [4]. For the aforementioned systems, the lightpath number propagated by each channel and mode is presented in Table 4. Furthermore, Binary PSK (BPSK) modulation with 5.5 *dB* the corresponding required SNR is considered [38]. There are 1, 3 and 4 spans respectively between nodes 1 to 2, 2 to 3 and 3 to 4, and the span length is $L_s = 100$ *km*. Therefore, the 4-node linear network we consider is composed of 8 spans where at the end of each span is an FM-EDFA. We consider strong coupling in all of the simulations shown in this section as all of the channels and

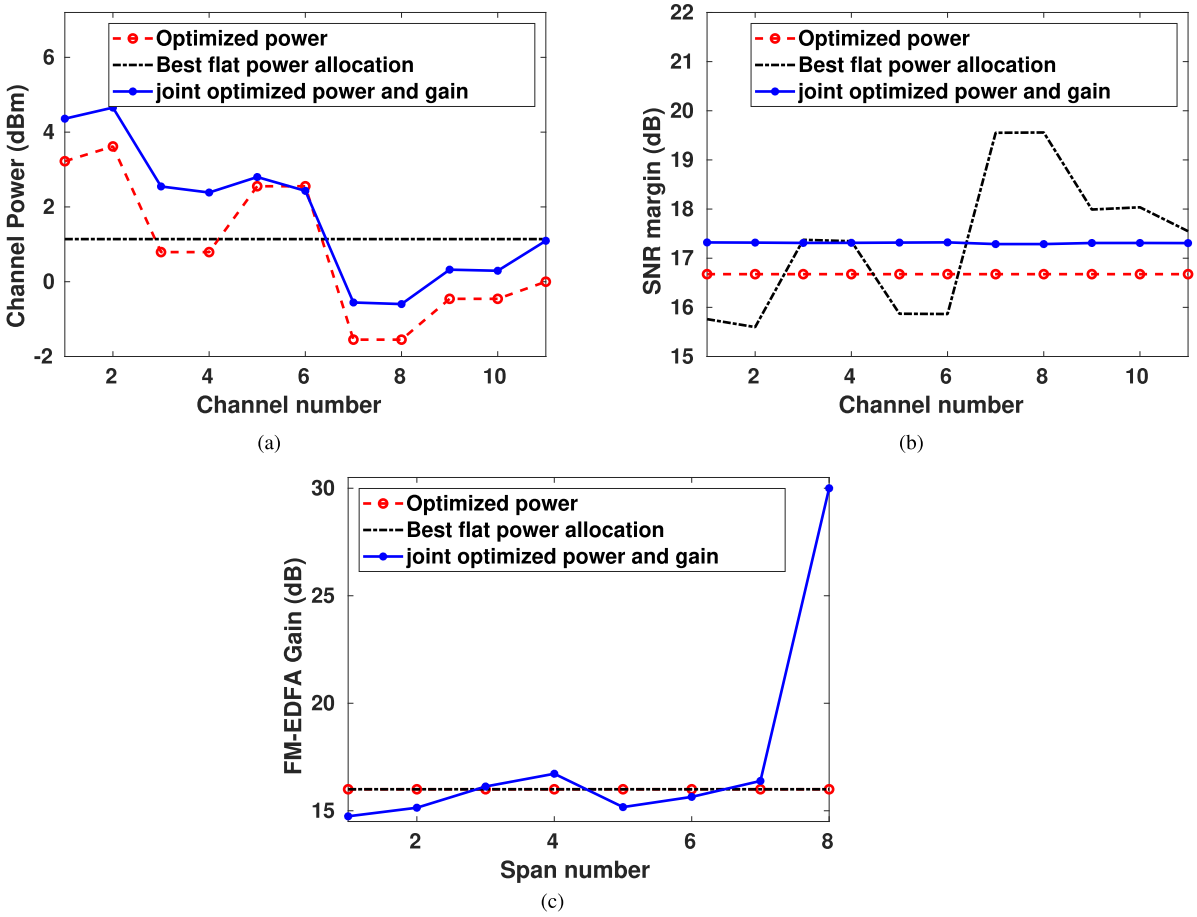

FIGURE 5. a) Channel power and b) SNR margin versus channel number, and c) FM-EDFA gain versus span number, for joint optimized power and gain allocation, optimized power allocation and best equal power allocation, considering SMF-WDM system.

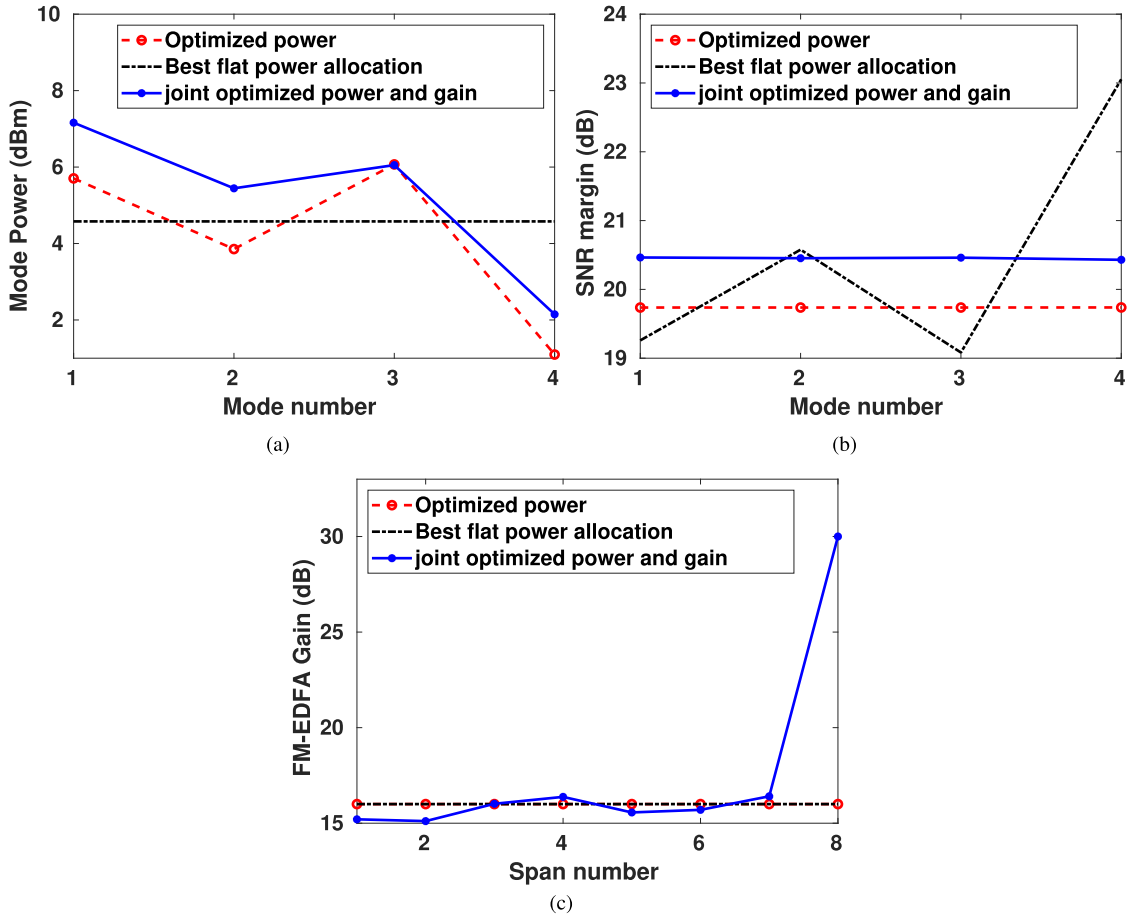

FIGURE 6. a) Mode power and b) SNR margin versus mode number, and c) FM-EDFA gain versus span number, for joint optimized power and gain allocation, optimized power allocation and best equal power allocation, considering MDM-single channel system.

TABLE 4. Lightpath number propagated by each channel and mode.

SMF-WDM	
Channel	lightpath #
$1, 2$	$L1$
$3, 4$	$L2$
$5, 6$	$L3$
$7, 8$	$L4$
$9, 10$	$L5$
11	$L6$

MDM-single channel	
Mode	lightpath #
LP01	$L1$
LP11a	$L2$
LP11b	$L3$
LP02	$L4$

modes propagate through long-range links. It is worth noting that the point-to-point links often have a homogeneous set where different channels/modes experience the same interacting channels/modes. However, in multi-node linear networks, the channels/modes may propagate different distances so that they accumulate different NLI noise, experience fragmentation/partial utilization (and hence see different interacting channels/modes), and observe different FM-EDFA gains.

The results for the SMF-WDM system are summarized in Fig. 5. Figs. 5a and 5b depict respectively the channel power and SNR margin versus channel number, while Fig. 5c shows the FM-EDFA gain versus span number for joint optimized power and gain allocation, optimized power allocation, and best equal power allocation. The best equal power allocation, optimized power allocation, and joint optimized power and gain allocation achieve 15.6 *dB*, 16.7 *dB*, and 17.3 *dB* minimum SNR margin improvement, respectively. The joint optimized power and gain allocation obtains improvements of 0.6 and 1.7 *dB* in minimum SNR margin compared with the optimized power allocation and best equal power allocation, respectively. This is due to the joint optimized power and gain allocation having more degrees of freedom. Note that the central channel indices have higher NLI noise (lower SNR margins); correspondingly, they should be allocated higher power to have a reliable link. Moreover, channels that propagate through longer lightpaths should be allocated higher power, since longer lightpaths involve more ASE/NLI noise power. In the joint optimized power and gain scenario, the last FM-EDFA gain is set to its maximum possible value. This result can be deduced from the SNR formulation where all terms except for receiver noise are scaled with G_{N_s} (to minimize the contribution of the receiver noise term, the maximum possible value should be chosen for G_{N_s}).

The results for the MDM-single channel system are summarized in Fig. 6: 6a and 6b demonstrate the mode power and SNR margin versus mode number, respectively, while 6c plots

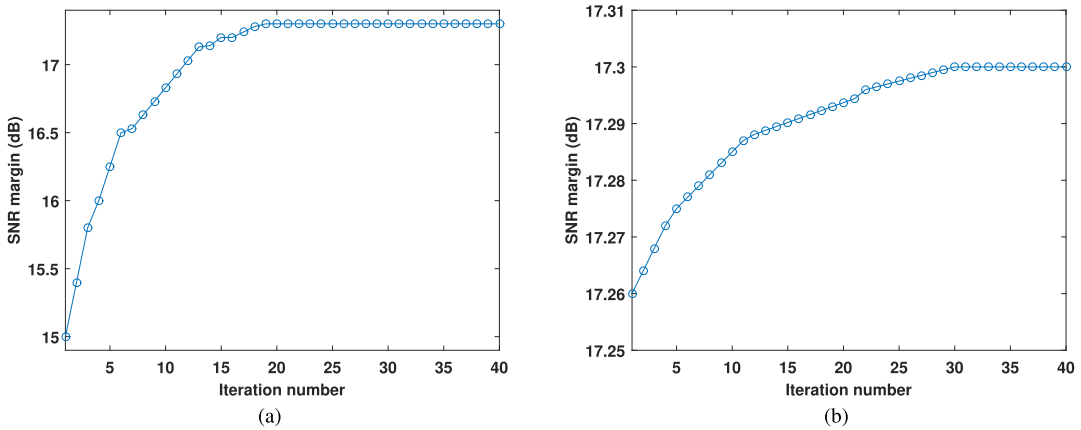

FIGURE 7. The SNR margin of the central channel in SMF-WDM scenario as a function of iteration number of a) Algorithms 1 and b) Algorithm 2.

the FM-EDFA gain versus span number, for joint optimized power and gain allocation, optimized power allocation, and best equal power allocation. The minimum SNR margins are 19.1 *dB* for best equal power allocation, 19.7 *dB* for optimized power allocation, and 20.5 *dB* for joint optimized power and gain allocation showing that the latter results in improvements of 0.8 and 1.4 *dB* compared with optimized power allocation and best equal power allocation, respectively. The main difference between the different modes of the same channel is their spatial profile. The LP11a/b mode has a larger spatial profile and therefore, a higher overlap with the other modes. Therefore, this mode has more NLI noise power and lower SNR margin. Accordingly, it should be allocated higher power than the other modes. Neither the allocated powers nor the FM-EDFA gains are not symmetric, since different modes have different NLI noise power which is not symmetric due to the nonlinear coupling.

Figs. 7a and 7b plot the SNR margin of the central channel in the SMF-WDM scenario as a function of iteration number of Algorithms 1 and 2. Although we consider the SNR margin of the central channel in the SMF-WDM scenario, the convergence speeds for the other channels and for the MDM-single channel case are essentially the same. Moreover, due to the fact that Algorithm 2 is inside the while loop of Algorithm 1, we consider the last iteration of Algorithm 1 for plotting the convergence speed of Algorithm 2, but the same convergence speed can be observed for other iterations within Algorithm 1. Algorithms 1 and 2 converge after 20 and 30 iterations, respectively. It should be noted that since we consider the FMF nonlinear channel as an AWGN channel with deterministic noise variance, we only need to deploy the joint optimized power and gain allocation once (and offline). Thus, the computational complexity and convergence speed do not impact practical implementations at scale.

VI. CONCLUSION

The GN model for FMF nonlinearity overestimates the NLI noise power in practical scenarios where the signal does not have a Gaussian distribution. For instance, in power allocation applications, this inaccuracy results in around 0.5 *dB* error on predicting the optimum launch power so that the obtained gains with respect to the best equal power allocation are too conservative and fall within the inaccuracy of the GN model. In the first part of this paper, we derived the EGN model for FMF nonlinearity. Compared with the GN model and based on SSFM simulations, our proposed EGN model is capable of providing very accurate estimates of NLI noise power at different number of spans in both the weak and strong coupling regimes, as well as for any launch power and modulation format. Achieving reliable communication over different channels and modes is one of the main goals in MDM-WDM networks, and the reliability is generally quantified through the minimum SNR margin. Based on the proposed EGN model, we formulated and solved the joint optimized power and gain allocation problem using the minimum SNR margin maximization in a multi-node linear network. The joint optimized power and gain allocation compared with best equal power allocation achieved 1.4 *dB* and 1.7 *dB* minimum SNR margin improvement in MDM-single channel and SMF-WDM systems, respectively.

Two main directions can be continued in future MDM-WDM investigations: (1) nonlinearity modelling and (2)

TABLE 5. Valid combinations yielding non-zero $E[\zeta_{\mathbf{k}}^{*}\zeta_{\mathbf{m}}\zeta_{\mathbf{n}}\zeta_{\mathbf{l}}\zeta_{\mathbf{j}}^{*}\zeta_{\mathbf{o}}^{*}]$.

$\mathbf{k}^*$	$\mathbf{m}$	$\mathbf{n}$	$\mathbf{l}$	$\mathbf{j}^*$	$\mathbf{o}^*$	
$*$	$*$	$*$	$*$	$*$	$*$	HON
$*$	$*$	$**$	$*$	$*$	$**$	$FONa$
$*$	$*$	$**$	$*$	$**$	$*$	$FONa$
$*$	$**$	$*$	$*$	$*$	$**$	$FONa$
$*$	$**$	$*$	$*$	$**$	$*$	$FONa$
$*$	$*$	$*$	$**$	$*$	$**$	$FONb$
$*$	$*$	$*$	$**$	$**$	$*$	$FONb$
$**$	$*$	$**$	$*$	$*$	$*$	$FONc$
$**$	$*$	$*$	$**$	$*$	$*$	$FONc$
$**$	$*$	$*$	$**$	$*$	$*$	$FONd$
$*$	$*$	$**$	$+$	$+$	$**$	GNa
$*$	$*$	$**$	$+$	$**$	$+$	GNa
$*$	$**$	$*$	$+$	$+$	$**$	GNa
$*$	$**$	$*$	$+$	$**$	$+$	GNa
$*$	$+$	$**$	$*$	$+$	$**$	GNb
$*$	$+$	$**$	$*$	$**$	$+$	GNb

TABLE 6. Computational complexity of Algorithms 1 and 2.

	Number of multiplications	Number of additions
Algorithm 1	$(N_{iter1}+1)\Big[N_{iter2}\Big(N_{iter3}\Big\{DN_{ch}[N_s\{2N_s+n_1(9+6N_s+19DN_{ch})\}+DN_{ch}\{10n_1+5\}]+N_s[N_s\{n_1(6+6N_s+25DN_{ch})\}+2+N_s]\Big\}+3\Big)+1\Big]$	$(N_{iter1}+1)\Big[N_{iter2}\Big(N_{iter3}\Big\{DN_{ch}[N_s\times\{14+11DN_{ch}+n_1(5+4N_s+46DN_{ch})\}]+N_s\times[N_s\{9+10DN_{ch}+n_1(5+7N_s+25DN_{ch})\}]\Big\}+3\Big)+1\Big]$
Algorithm 2	$N_{iter2}\Big(N_{iter3}\Big\{DN_{ch}[N_s\{2N_s+n_1(9+6N_s+19DN_{ch})\}+DN_{ch}\{10n_1+5\}]+N_s\times[N_s\{n_1(6+6N_s+25DN_{ch})\}+2+N_s]\Big\}+3\Big)$	$N_{iter2}\Big(N_{iter3}\{DN_{ch}[N_s\{14+11DN_{ch}+n_1(5+4N_s+46DN_{ch})\}]+N_s[N_s\{9+10DN_{ch}+n_1(5+7N_s+25DN_{ch})\}]\Big\}+3\Big)$

resource allocation. First, in terms of modelling, it should be noted that we provided a series of investigations over IGN, GN, and the EGN models for MDM-WDM systems. These works are based on the Manakov equation which considers the average of the FMF linear and nonlinear interactions. Therefore, the derivation of IGN, GN, and EGN models based on the nonlinear Schrodinger equation is a topic for future work. Moreover, we derived the EGN formulation for the observed NLI noise after CPE; it will also be important to derive the EGN model for the NLI noise before CPE. Finally, the existing IGN, GN, and EGN models only estimate the amplitude of NLI noise and should be extended to account for the phase. Second, different resource allocation problems can be formulated based on our proposed IGN, GN, and EGN models, including marginal (or joint) routing, mode-wavelength assignment, and power allocation. These problems can be solved by iterative optimization algorithms or by deep learning methods. We also only considered a linear multi-node network while nonlinear multi-node networks, dynamic networks, and flex-grid networks can also be explored.

APPENDIX A: THE NLI NOISE VARIANCE OF THE i_2TH CHANNEL AND pTH MODE

The GN, FON, and HON terms in the EGN model can be expressed in (37), (38), and (39), as shown at the top of the next page, respectively.

$$\begin{aligned}\kappa_1 &= \mu_2,\\ \kappa_2 &= \mu_4 - 2\mu_2^2,\\ \kappa_3 &= \mu_6 - 4\mu_4\mu_2 + 12\mu_2^3,\\ \mu_n &= E[|\zeta_{\mathbf{k}}|^n]. \end{aligned} \tag{40}$$

Table 5 shows valid combinations yielding non-zero $E[\zeta_{\mathbf{k}}^*\zeta_{\mathbf{m}}\zeta_{\mathbf{n}}\zeta_{\mathbf{l}}\zeta_{\mathbf{j}}^*\zeta_{\mathbf{o}}^*]$ where FONb, FONc, and GNa are the removed terms from the EGN model formulation due to the CPE assumption. Moreover, $\sigma_{EGN}^{(i_3,odd)}(f) = \sigma_{EGN}^{(i_3,even)}(f)$, and $\sigma_{EGN}^{(p)}(f) = \sigma_{EGN}^{(i_3,odd)}(f) + \sigma_{EGN}^{(i_3,even)}(f)$. Therefore, the GN, FON, and HON contributions of the NLI noise variance of the i_2th channel and pth mode can be written in (41), (42), and (43), as shown at the bottom of the 15th page, respectively.

The power spectral density of the optical signal launched into the fiber can be written as $\tilde{G}_{Tx}^{i_2,p}(f) = P_{i_2,p}g_{i_2,p}(f)$. Accordingly, the GN, FON, and HON contributions of the NLI noise variance of the i_2th channel and pth mode can be expressed in (44), (45), and (46), as shown at the bottom of the 15th page, respectively. Therefore, the NLI noise variance can be written in (47), as shown at the top of the 17th page.

APPENDIX B: CONVEXITY PROOF OF OPTIMIZATION PROBLEM (33)

The expression (48), as shown at the top of the 17th page is convex in $\hat{P}_l$, g_s, since $log-sum-exp(x)$ is a convex function in x [34]. The constraint function of (33) is the summation of some convex functions, therefore, it is convex [34]. The objective and constraint functions of (33) are convex, therefore, (33) is a convex optimization problem [34].

APPENDIX C: CONVERGENCE OF ALGORITHMS 1 AND 2

In Algorithm 2, (35) is solved at each iteration as a function of $\hat{P}_l$ and g_n using the gradient descent algorithm which will converge to its optimum solution due to the convexity of the dual problem [34]. This procedure is repeated by Algorithm 1 in the "While loop", by which the minimum SNR margin is improved successively until convergence to the maximum value. Note that Algorithm 1 will stop searching while the difference between the upper and lower bounds becomes less than ϵ.

APPENDIX D: COMPUTATIONAL COMPLEXITY OF ALGORITHMS 1 AND 2

For computational complexity analysis of Algorithms 1 and 2, we will count the total number of additions and multiplications per iteration. Note that $exp(x) = \sum_{i=0}^{n_1} x^i/i!$ and $log(x) = \sum_{i=1}^{n_1}(-1)^{i+1}(x-1)^i/i$ can be computed by $2n_1$ multiplications and n_1 additions, with n_1 as an integer where we get better accuracy for a larger n_1. The computational complexity of Algorithms 1 and 2 is presented in Table 6. Each iteration of Algorithm 1 is composed of updating the search boundaries and solving the convex problem (33) by the Lagrangian method described in Algorithm 2 in a while loop. Therefore, the complexity of Algorithm 1 is $(N_{iter1}+1)\times(complexity\ of\ Algorithm2 + 1\ addition + 1 multiplication)$ with N_{iter1} as total iterations. Each iteration of Algorithm 2 is composed of solving (35) by the gradient descent algorithm and updating the Lagrangian variables in a while loop. Therefore, the complexity of

$$
\begin{aligned}
&\sigma^2_{GN,i_2,i_3} \\
&= \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\,L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n\,L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n\,L_n)^{1/2}\sum_{\mathbf{k},\mathbf{m},\mathbf{n}}\kappa_1^{(\mathbf{k})}\kappa_1^{(\mathbf{m})}\kappa_1^{(\mathbf{n})}\Bigg(\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle k_3|m_3\rangle\langle i_3|n_3\rangle\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle m_3|k_3\rangle \\
&\quad \langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle k_3|m_3\rangle\langle i_3|n_3\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{n}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_1)g^{\mathbf{i}}(f)\langle n_3|k_3\rangle\langle m_3|i_3\rangle df_1df_2df\Bigg)+\Bigg(\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle k_3|n_3\rangle\langle i_3|k_3\rangle\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle m_3|m_3\rangle \\
&\quad \langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle k_3|n_3\rangle\langle i_3|k_3\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{n}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_1)g^{\mathbf{i}}(f)\langle n_3|m_3\rangle\langle m_3|i_3\rangle df_1df_2df\Bigg)+\Bigg(\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|m_3\rangle\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle m_3|m_3\rangle \\
&\quad \langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|m_3\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{m}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle m_3|m_3\rangle\langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|m_3\rangle\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle m_3|m_3\rangle \\
&\quad \langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|m_3\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{n}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{m}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{m}*}(f+f_1)g^{\mathbf{i}}(f)\langle n_3|m_3\rangle\langle m_3|i_3\rangle df_1df_2df\Bigg)
\end{aligned}
\tag{37}
$$

$$
\begin{aligned}
&\sigma^2_{FON,i_2,i_3} \\
&= \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\,L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n\,L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n\,L_n)^{1/2}\sum_{\mathbf{k},\mathbf{m},\mathbf{n}}\kappa_2^{(\mathbf{k})}\kappa_1^{(\mathbf{n})}\Bigg(\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \Big|\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|n_3\rangle+\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1) \\
&\quad \langle k_3|n_3\rangle\langle i_3|k_3\rangle\Big|^2df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\Big|\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle n_3|k_3\rangle \\
&\quad \langle i_3|k_3\rangle\Big|^2df_1df_2df\Bigg)+\Bigg[\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle k_3|k_3\rangle \\
&\quad \langle i_3|k_3\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1)g^{\mathbf{i}}(f)\langle n_3|k_3\rangle\langle n_3|i_3\rangle df_1df_2df+\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle k_3|k_3\rangle\langle i_3|k_3\rangle\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1)g^{\mathbf{i}}(f)\langle n_3|n_3\rangle \\
&\quad \langle i_3|k_3\rangle df_1df_2df\Bigg]+\Bigg[\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1)\langle n_3|n_3\rangle\langle i_3|k_3\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{k}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1)g^{\mathbf{i}}(f)\langle k_3|k_3\rangle\langle i_3|k_3\rangle df_1df_2df\Bigg]+\Bigg[\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle k_3|n_3\rangle\langle i_3|n_3\rangle\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_2)\tilde{W}_{Tx}^{\mathbf{k}}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{k}*}(f+f_1)g^{\mathbf{i}}(f) \\
&\quad \langle k_3|k_3\rangle\langle i_3|k_3\rangle df_1df_2df\Bigg]
\end{aligned}
\tag{38}
$$

$$
\begin{aligned}
\sigma^2_{HON,i_2,i_3} &= \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n\,L_n)^{3/2}\prod_{n=s}^{N_s}(G_n\,L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n\,L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n\,L_n)^{1/2}\sum_{\mathbf{n}}\kappa_3^{(\mathbf{n})}\iiint_{-\infty}^{\infty}\langle\eta_s(f,f_1,f_2)|\eta_{s'}(f,f_1,f_2)\rangle \\
&\quad \Big|\tilde{W}_{Tx}^{\mathbf{n}*}(f+f_1+f_2)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_2)g^{\mathbf{i}*}(f)\tilde{W}_{Tx}^{\mathbf{n}}(f+f_1)\langle n_3|n_3\rangle\langle i_3|n_3\rangle\Big|^2df_1df_2df
\end{aligned}
\tag{39}
$$

Algorithm 2 is $N_{iter2} \times$ (*complexity of gradient descent*+ 3 *additions* +3*multiplications*) with N_{iter2} as total iterations. To compute the complexity of the gradient descent algorithm, we should note that it updates $\hat{P}_l$ and g_n within N_{iter3} iterations according to the following formula

$$\hat{P}_l^{(i+1)} = \hat{P}_l^{(i)} - \mu \nabla_{\hat{P}_l},$$

$$g_n^{(i+1)} = g_n^{(i)} - \mu \nabla_{g_n}, \tag{49}$$

with μ as the step size, $\nabla_{\hat{P}_l}$ and ∇_{g_n} as the gradient of objective function of (35) regarding $\hat{P}_l$ and g_n, respectively. Therefore, each iteration of the gradient descent algorithm has $DN_{ch}[N_s\{2N_s + n_1(9 + 6N_s + 19DN_{ch})\} + DN_{ch}\{10n_1 +5\}]+N_s[N_s\{n_1(6+6N_s+25DN_{ch})\}+2+N_s]$ multiplications

$$\sigma^2_{GN,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\frac{3}{4}\sum_{q=1}^{D}\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)$$
$$\eta^*_{s'}(f,f_1,f_2)\Big(\tilde{G}_{Tx}^{m_2,q}(f+f_2)\tilde{G}_{Tx}^{k_2,q}(f+f_1+f_2)\tilde{G}_{Tx}^{n_2,p}(f+f_1)g^{i_2,p}(f)+\tilde{G}_{Tx}^{n_2,q}(f+f_2)\tilde{G}_{Tx}^{k_2,q}(f+f_1+f_2)\tilde{G}_{Tx}^{m_2,p}(f+f_1)$$
$$g^{i_2,p}(f)\Big)df_1df_2df \tag{41}$$

$$\sigma^2_{FON,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\frac{1}{4}\sum_{q=1}^{D}\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}\Big(\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)$$
$$\eta^*_{s'}(f,f_1,f_2)5\tilde{G}_{Tx}^{k_2,q}(f+f_1+f_2)\tilde{G}_{Tx}^{k_2,q}(f+f_2)g^{i_2,p}(f)\tilde{G}_{Tx}^{n_2,p}(f+f_1)df_1df_2df+\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)$$
$$\tilde{G}_{Tx}^{n_2,q}(f+f_1+f_2)\tilde{G}_{Tx}^{k_2,q}(f+f_2)g^{i_2,p}(f)\tilde{G}_{Tx}^{k_2,p}(f+f_1)df_1df_2df\Big) \tag{42}$$

$$\sigma^2_{HON,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\frac{1}{4}\sum_{q=1}^{D}\sum_{n_2}\kappa_3^{(n_2)}\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)$$
$$\tilde{G}_{Tx}^{n_2,q}(f+f_1+f_2)\tilde{G}_{Tx}^{n_2,q}(f+f_2)g^{i_2,p}(f)\tilde{G}_{Tx}^{n_2,p}(f+f_1)df_1df_2df \tag{43}$$

$$\sigma^2_{GN,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\frac{3}{4}\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)$$
$$\eta^*_{s'}(f,f_1,f_2)\Big(P_{k_2,q}P_{m_2,q}P_{n_2,p}g^{m_2,q}(f+f_2)g^{k_2,q}(f+f_1+f_2)g^{n_2,p}(f+f_1)g^{i_2,p}(f)+P_{k_2,q}P_{m_2,p}P_{n_2,q}g^{n_2,q}(f+f_2)$$
$$g^{k_2,q}(f+f_1+f_2)g^{m_2,p}(f+f_1)g^{i_2,p}(f)\Big)df_1df_2df \tag{44}$$

$$\sigma^2_{FON,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\frac{1}{4}\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}\Big(P^2_{k_2,q}P_{n_2,p}5\iiint_{-\infty}^{\infty}$$
$$\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{k_2,q}(f+f_1+f_2)g^{k_2,q}(f+f_2)g^{i_2,p}(f)g^{n_2,p}(f+f_1)df_1df_2df+P_{k_2,p}P_{k_2,q}P_{n_2,q}\iiint_{-\infty}^{\infty}$$
$$\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{n_2,q}(f+f_1+f_2)g^{k_2,q}(f+f_2)g^{i_2,p}(f)g^{k_2,p}(f+f_1)df_1df_2df\Big) \tag{45}$$

$$\sigma^2_{HON,i_2,p} = \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\frac{1}{4}\sum_{n_2}\kappa_3^{(n_2)}P^2_{n_2,q}P_{n_2,p}\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)$$
$$\eta^*_{s'}(f,f_1,f_2)g^{n_2,q}(f+f_1+f_2)g^{n_2,q}(f+f_2)g^{i_2,p}(f)g^{n_2,p}(f+f_1)df_1df_2df \tag{46}$$

$$
\begin{aligned}
&\sigma^2_{EGN,i_2,p} \\
&= \sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(G_n L_n)^{3/2}\prod_{n=s}^{N_s}(G_n L_n)^{1/2}\prod_{n=1}^{s'-1}(G_n L_n)^{3/2}\prod_{n=s'}^{N_s}(G_n L_n)^{1/2}\sum_{q=1}^{D}\frac{3}{4}\sum_{k_2,m_2,n_2}\kappa_1^{(k_2)}\kappa_1^{(m_2)}\kappa_1^{(n_2)}P_{k_2,q}P_{m_2,q} \\
&P_{n_2,p}X^a_{i_2,p}(k_2,m_2,n_2,q)+\frac{1}{4}\sum_{k_2,n_2}\kappa_2^{(k_2)}\kappa_1^{(n_2)}(P^2_{k_2,q}P_{n_2,p}5X^b_{i_2,p}(k_2,k_2,n_2,q)+P_{k_2,p}P_{k_2,q}P_{n_2,q}X^c_{i_2,p}(k_2,n_2,k_2,q))+\frac{1}{4}\sum_{n_2} \\
&\kappa_3^{(n_2)}P^2_{n_2,q}P_{n_2,p}X^d_{i_2,p}(n_2,n_2,n_2,q) \\
&\textit{where}\begin{cases}
X^a_{i_2,p}(k_2,m_2,n_2,q)=\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{m_2,q}(f+f_2)g^{k_2,q}(f+f_1+f_2)g^{n_2,p}(f+f_1)g^{i_2,p}(f)df_1df_2df\\
X^b_{i_2,p}(k_2,k_2,n_2,q)=\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{k_2,q}(f+f_1+f_2)g^{k_2,q}(f+f_2)g^{i_2,p}(f)g^{n_2,p}(f+f_1)df_1df_2df\\
X^c_{i_2,p}(k_2,n_2,k_2,q)=\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{n_2,q}(f+f_1+f_2)g^{k_2,q}(f+f_2)g^{i_2,p}(f)g^{k_2,p}(f+f_1)df_1df_2df\\
X^d_{i_2,p}(n_2,n_2,n_2,q)=\iiint_{-\infty}^{\infty}\eta_s(f,f_1,f_2)\eta^*_{s'}(f,f_1,f_2)g^{n_2,q}(f+f_1+f_2)g^{n_2,q}(f+f_2)g^{i_2,p}(f)g^{n_2,p}(f+f_1)df_1df_2df
\end{cases}
\end{aligned}
\tag{47}
$$

$$
\begin{aligned}
&log\Bigg(\prod_{n=1}^{N_s}(e^{g_n}L_n)(F(G_{BA}-1)h\nu B_{i_2})+\sum_{s,s'=1}^{N_s}[(F(e^{g_s}-1)h\nu B_{i_2})\prod_{n=s+1}^{N_s}(e^{g_n}L_n)]+\sum_{s,s'=1}^{N_s}\prod_{n=1}^{s-1}(e^{g_n}L_n)^{3/2}\prod_{n=s}^{N_s}(e^{g_n}L_n)^{1/2}\prod_{n=1}^{s'-1} \\
&(e^{g_n}L_n)^{3/2}\prod_{n=s'}^{N_s}(e^{g_n}L_n)^{1/2}\Bigg[\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_1^{\prime(l_1)}\kappa_1^{\prime(l_2)}\kappa_1^{\prime(l_3)}e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}3H^a_l(l_1,l_2,l_3)+\frac{1}{4}\sum_{l_1,l_2,l_3=1}^{DN_{ch}}\kappa_2^{\prime(l_1)}\kappa_1^{\prime(l_2)}(e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}5 \\
&H^b_l(l_1,l_1,l_2)+e^{\hat{P}_{l_1}+\hat{P}_{l_2}+\hat{P}_{l_3}}H^c_l(l_1,l_2,l_3))+\frac{1}{4}\sum_{l_1,l_2=1}^{DN_{ch}}\kappa_3^{\prime(l_1)}e^{2\hat{P}_{l_1}+\hat{P}_{l_2}}H^d_l(l_1,l_1,l_2)\Bigg]+\sigma^2_{RxN}\Bigg)-(\hat{P}_l+\sum_{n=1}^{N_s}g_n log(L_n))
\end{aligned}
\tag{48}
$$

$DN_{ch}[N_s\{14+11DN_{ch}+n_1(5+4N_s+46DN_{ch})\}]+N_s[N_s\{9+10DN_{ch}+n_1(5+7N_s+25DN_{ch})\}]$ additions.

REFERENCES

[1] F. M. Ferreira, C. S. Costa, S. Sygletos, and A. D. Ellis, "Nonlinear performance of few-mode fiber links with intermediate coupling," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 989–999, Feb. 1, 2019.

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

[3] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: Generalization of the Manakov equations," *J. Lightw. Technol.*, vol. 31, no. 3, pp. 398–406, Feb. 1, 2013.

[4] G. Rademacher and K. Petermann, "Nonlinear Gaussian noise model for multimode fibers with space-division multiplexing," *J. Lightw. Technol.*, vol. 34, no. 9, pp. 2280–2287, May 1, 2016.

[5] A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," *Opt. Exp.*, vol. 20, no. 21, pp. 23436–23441, 2012.

[6] A. E. Elfiqi, A. A. I. Ali, Z. A. El-Sahn, K. Kato, and H. M. H. Shalaby, "Theoretical analysis of long-haul systems adopting mode-division multiplexing," *Opt. Commun.*, vol. 445, pp. 10–18, Aug. 2019.

[7] J. Li, Z. Wu, D. Ge, J. Zhu, Y. Tian, Y. Zhang, J. Yu, Z. Li, Z. Chen, and Y. He, "Weakly-coupled mode division multiplexing over conventional multi-mode fiber with intensity modulation and direct detection," *Frontiers Optoelectron.*, vol. 12, no. 1, pp. 31–40, Mar. 2019.

[8] P. Sillard, "Few-mode-fiber developments and applications," in *Proc. 23rd Opto-Electron. Commun. Conf. (OECC)*, Jul. 2018, pp. 1–2.

[9] B. Inan, B. Spinnler, F. Ferreira, D. van den Borne, A. Lobato, S. Adhikari, V. Sleiffer, M. Kuschnerov, N. Hanik, and S. Jansen, "DSP complexity of mode-division multiplexed receivers," *Opt. Exp.*, vol. 20, no. 10, pp. 10859–10869, 2012.

[10] G. Rademacher, S. Warm, and K. Petermann, "Nonlinear interaction in differential mode delay managed mode-division multiplexed transmission systems," *Opt. Exp.*, vol. 23, no. 1, pp. 55–60, 2015.

[11] F. Ye, S. Warm, and K. Petermann, "Differential mode delay management in spliced multimode fiber transmission systems," in *Proc. Opt. Fiber Commun. Conf./Nat. Fiber Optic Engineers Conf.*, 2013, pp. 1–3.

[12] D. Kroushkov, G. Rademacher, and K. Petermann, "Cross mode modulation in multimode fibers," *Opt. Lett.*, vol. 38, no. 10, pp. 1642–1644, 2013.

[13] A. D. Ellis, N. Mac Suibhne, F. C. G. Gunning, and S. Sygletos, "Expressions for the nonlinear transmission performance of multi-mode optical fiber," *Opt. Exp.*, vol. 21, no. 19, pp. 22834–22846, 2013.

[14] P. Serena and A. Bononi, "A time-domain extended Gaussian noise model," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1459–1472, Apr. 1, 2015.

[15] G. Rademacher, R. S. Luis, B. J. Puttnam, H. Furukawa, R. Maruyama, K. Aikawa, Y. Awaji, and N. Wada, "Investigation of intermodal four-wave mixing for nonlinear signal processing in few-mode fibers," *IEEE Photon. Technol. Lett.*, vol. 30, no. 17, pp. 1527–1530, Sep. 1, 2018.

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

[17] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A simple and effective closed-form GN model correction formula accounting for signal non-Gaussian distribution," *J. Lightw. Technol.*, vol. 33, no. 2, pp. 459–473, Jan. 15, 2015.

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

[19] P. Johannisson and E. Agrell, "Modeling of nonlinear signal distortion in fiber-optic networks," *J. Lightw. Technol.*, vol. 32, no. 23, pp. 4544–4552, Dec. 1, 2014.
[20] G. Rademacher, S. Warm, and K. Petermann, "Analytical description of cross-modal nonlinear interaction in mode multiplexed multimode fibers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1929–1932, Nov. 1, 2012.
[21] G. Rademacher, R. S. Luis, B. J. Puttnam, R. Maruyama, K. Aikawa, Y. Awaji, H. Furukawa, K. Petermann, and N. Wada, "Investigation of intermodal nonlinear signal distortions in few-mode fiber transmission," *J. Lightw. Technol.*, vol. 37, no. 4, pp. 1273–1279, Feb. 15, 2019.
[22] A. A. I. Ali, A. E. El-Fiqi, Z. A. El-Sahn, H. M. H. Shalaby, and R. K. Pokharel, "Analytical formula of nonlinear interference in few-mode fibers in strong coupling regime," in *Proc. 17th Int. Conf. Transparent Opt. Netw. (ICTON)*, Jul. 2015, pp. 1–4.
[23] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 15, 2014.
[24] M. A. Amirabadi, M. H. Kahaei, S. A. Nezamalhosseini, and R. Lawrence Chen, "Optimal power allocation in nonlinear MDM-WDM systems using Gaussian noise model," *IET Optoelectron.*, to be published.
[25] M. A. Amirabadi, M. H. Kahaei, S. A. Nezamalhosseini, and R. Lawrence Chen, "Improving MDM-WDM optical network performance via optimal power allocation using Gaussian noise model," *Submitted*, submitted for publication.
[26] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 15, 2012.
[27] I. Roberts, J. M. Kahn, and D. Boertjes, "Convex channel power optimization in nonlinear WDM systems using Gaussian noise model," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3212–3222, Jul. 1, 2016.
[28] I. Roberts, J. M. Kahn, J. Harley, and D. W. Boertjes, "Channel power optimization of WDM systems following Gaussian noise nonlinearity model in presence of stimulated Raman scattering," *J. Lightw. Technol.*, vol. 35, no. 23, pp. 5237–5249, Dec. 1, 2017.
[29] I. Roberts and J. M. Kahn, "Measurement-based optimization of channel powers with non-Gaussian nonlinear interference noise," *J. Lightw. Technol.*, vol. 36, no. 13, pp. 2746–2756, Jul. 1, 2018.
[30] I. Roberts and J. M. Kahn, "Efficient discrete rate assignment and power optimization in optical communication systems following the Gaussian noise model," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4425–4437, Oct. 15, 2017.
[31] G. Rademacher, F. Schmidt, and K. Petermann, "Optimum capacity utilization in space-division multiplexed transmission systems with multimode fibers," in *Proc. 42nd Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2016, pp. 1–3.
[32] R. Hashemi, H. Beyranvand, and H. Rabbani, "Joint channel power and amplifier gain optimization in coherent DWDM systems," *Opt. Commun.*, vol. 475, Nov. 2020, Art. no. 126212.
[33] F. Daneshfar and E. Hosseini, "Load-frequency control in a deregulated environment based on bisection search," *Iranian J. Elect. Electron. Eng.*, vol. 8, no. 4, pp. 303–310, 2012.
[34] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[35] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Suppression of spurious tones induced by the split-step method in fiber systems simulation," *IEEE Photon. Technol. Lett.*, vol. 12, no. 5, pp. 489–491, May 2000.
[36] F. M. Ferreira, C. S. Costa, S. Sygletos, and A. D. Ellis, "Overcoming degradation in spatial multiplexing systems with stochastic nonlinear impairments," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, Dec. 2018.
[37] H. Rabbani, L. Beygi, S. Ghoshooni, H. Rabbani, and E. Agrell, "Quality of transmission aware optical networking using enhanced Gaussian noise model," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 831–838, Feb. 1, 2019.
[38] A. Raeesi, H. Rabbani, L. Beygi, and S. Zokaei, "Low-complexity physical layer impairment aware spectrum assignment based on discretized Gaussian model for nonlinear noise in elastic optical networks," *Opt. Commun.*, vol. 474, Nov. 2020, Art. no. 126011.

MOHAMMAD ALI AMIRABADI was born in Zahedan, Iran, in 1993. He received the B.Sc. degree in optics and laser engineering from the Malek-e-Ashtar University of Technology, Isfahan, Iran, in 2015, and the M.Sc. degree in communication engineering from the Iran University of Science and Technology, Tehran, Iran, in 2017, where he is currently pursuing the Ph.D. degree in communication engineering. His research interests include multimode fiber optic communication, free space optical communication, and deep learning.

MOHAMMAD HOSSEIN KAHAEI received the B.Sc. degree in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 1986, the M.Sc. degree in adaptive signal processing from the University of the Ryukyus, Okinawa, Japan, in 1994, and the Ph.D. degree in signal processing from the Queensland University of Technology, Brisbane, Australia, in 1998. Since 1999, he has been with the School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran, where he is currently an Associate Professor and the Head of the Signal and System Modeling Laboratory. His research interests include array signal processing with primary emphasis on compressed sensing, sparse optimization problems, data science, DOA estimation, blind source separation, learning-based techniques, communication systems, and wireless sensor networks.

S. ALIREZA NEZAMALHOSSEINI received the B.Sc. degree in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2006, and the M.Sc. and Ph.D. degrees in electrical engineering from the Sharif University of Technology (SUT), Tehran, in 2008 and 2013, respectively. He is currently an Assistant Professor with the Iran University of Science and Technology (IUST), Tehran. His research interests include underwater wireless optical communications, mode-division multiplexing in optical fibers, and visible light communications.

LAWRENCE R. CHEN (Senior Member, IEEE) received the B.Eng. degree in electrical engineering and mathematics from McGill University, Montreal, QC, Canada, in 1995, and the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 1997 and 2000, respectively. Since 2000, he has been with the Department of Electrical and Computer Engineering, McGill University. His research interests include optical communications, silicon photonics, and microwave photonics, and engineering education (particularly learning systems and learning mechanisms). He was the Guest Co-Editor of the IEEE/OSA JOURNAL OF LIGHTWAVE TECHNOLOGY Special Issue on Microwave Photonics published, in 2020, the Principal Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS Issue on Optical Signal Processing published, in 2021, and the Principal Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS Issue on Photonic Signal Processing to be published, in 2023. He is a fellow of Optica.