Measurement-Based Optimization of Channel Powers With Non-Gaussian Nonlinear Interference Noise

Ian Roberts, *Student Member, OSA* and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—We present two enhancements that expand the utility of convex optimization of channel powers in wavelength-division-multiplexed systems with nonlinear interference noise (NLIN). First, we present a first-order perturbation theory-based model that can quantify NLIN without assuming any particular noise distribution, such as circular Gaussian. The model supports a convex form for the problem of maximizing the minimum channel SNR in the presence of NLIN. Second, we describe how to perform the channel power optimization based on the measurements by receivers in a live system, rather than the numerical modeling of the system. Such measurement-based optimization avoids impairments caused by the modeling errors or incorrect choice of model parameters. We verify our approach through split-step Fourier simulations of a system with significantly noncircular NLIN. We obtain a gain of 0.25 dB in minimum margin over an optimized one-dimensional power allocation in a system with 24 channels with alternating 16-QAM and 32-QAM modulation and a tilted noise spectrum.**

***Index Terms*—Optical communications, nonlinear propagation, network optimization.**

I. Introduction

In wavelength-division-multiplexed optical fiber communication systems, interference due to the Kerr nonlinearity limits achievable system performance. Channel power optimization allows system performance to be maximized in the presence of nonlinear interference. Inter-channel nonlinear interactions make such a power optimization challenging in systems with wavelength-dependent impairments.

In dispersion-uncompensated transmission systems, the Gaussian noise nonlinearity model has been a valuable tool for system analysis [1], [2]. Using such a model, previous work found that the problem of optimizing an array of channel powers to maximize the minimum channel SNR had a convex form [3]. In homogeneous spectrally flat scenarios, a spectrally flat, one-dimensional power optimization was found to be close to optimal, while multi-dimensional power optimization yielded significant improvement over one-dimensional optimization in scenarios in which different channels have disparate noise levels. The diversity of path lengths in a mesh network lead to such disparity, and multi-dimensional power optimization yields significant benefits [3]. Such multi-dimensional power optimization also yields significant benefits when disparate discrete modulation formats or data rates are used across the wavelength band [4].

Other works have investigated different aspects of the power allocation problem in optical communication networks. Equalization of channel SNRs in WDM communication systems has been used to combat amplifier gain variation [5], [6]. Using a linearized model of nonlinear interference, the problem of optical network power allocation has been modeled as a Nash game [7], [8], and as a constrained optimization problem [9], [10]. These formulations incorporate the interesting dynamics introduced by mesh networks, but are limited by the linearized interference model. Power pre-emphasis has been investigated for OFDM systems to mitigate intra-channel four-wave mixing [11]. That work is limited to power distributions featuring a super-Gaussian hole in the OFDM spectrum.

In this paper, we address two limitations that remain from previous work.

The first limitation we address is the assumption of Gaussian nonlinear noise. We demonstrate that channel power optimization for SNR has a convex form under a more general perturbation-based model of the nonlinear noise. This serves to characterize the shape of the optimization space, whether or not the particular model is used to analytically evaluate the strength of the nonlinear noise. Perturbative approaches to nonlinear interference modeling have an extensive history in the literature for intensity modulated systems [12]–[20], and more recently for coherent systems [21], [22], with significant recent works in [23] and [24].

As dispersion-uncompensated transmission reduces nonlinear interference contributions compared to dispersion-compensated transmission, uncompensated architectures are the choice for new fiber deployments [25]. Despite this, dispersion-compensated fiber plants remain, particularly in existing undersea cables where dispersion compensating fiber cannot be removed. Such systems are outside of the region of accuracy of the Gaussian noise model, and often feature noticeably noncircular nonlinear noise. Performing useful power optimization for these and other systems requires a nonlinear noise model that is accurate for such systems.

The second limitation we address is the dependence of power optimization on an analytic channel model. In system

Manuscript received November 28, 2017; revised March 1, 2018 and March 23, 2018; accepted March 31, 2018. Date of publication April 3, 2018; date of current version May 15, 2018. This work was supported in part by Ciena Corporation and in part by a Stanford Rambus Corporation Graduate Fellowship. *(Corresponding author: Ian Roberts.)*

The authors are with the E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: iroberts@stanford.edu; jmk@ee.stanford.edu).

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

Digital Object Identifier 10.1109/JLT.2018.2822719

optimization based on a model, if there are errors in the system information used by the model, or an accurate model does not exist, then the optimization performance is impaired. Conservative margins are used for system performance predictions in order to accommodate modeling errors, at the expense of significant unrealized capacity. A channel power optimization method that is measurement-based rather than exclusively model-based is necessary to extract the best performance.

Even in the regime where the Gaussian nonlinear noise assumption is valid, a variety of modeling errors can introduce suboptimality into model-based optimization. Incorrect accounting of the amplified spontaneous emission noise level will shift optimal nonlinear noise levels. Differences in fiber length, or attenuation profile, due to connectors or splices, between the real-world system and the model will introduce error into the nonlinear noise modeling. Wavelength dependencies in the fiber parameters, amplifier gain, or due to stimulated Raman scattering, will introduce inhomogeneities that can be exploited by multi-dimensional power optimization to increase performance. Spectral filtering due to wavelength selective switches, and variations in channel spacing, will impact the nonlinear noise spectrum. The level of dispersion pre-compensation can also impact the generation of nonlinear noise [26].

We demonstrate the optimization of a multi-channel system, simulated via the split-step Fourier method, using measurements of the nonlinear channel performed by the receiver. This serves to demonstrate the practicality of multi-dimensional power optimization of real-world systems without the constraint of reliance on a theoretical channel model.

The remainder of this paper is as follows. In Section II, we characterize the nonlinear noise contributions using the time-domain perturbation approach of [23] and [24]. In Section III, we describe how power optimization can be performed using measurements obtained from a live system. In Section IV, we present power optimizations of multi-channel systems simulated by the split-step Fourier method. We discuss our results in Section V and present conclusions in Section VI.

II. First-Order Perturbation Nonlinearity Model

In order to perform power optimization on systems that fall outside of the Gaussian noise model regime, we require a nonlinear noise model that supports more general nonlinear noise distributions. The time-domain first-order perturbation model analyzed in [23] provides a very general nonlinear interference noise model. Arbitrary dispersion and attenuation profiles are supported, with limited assumptions on the symmetry of message symbols that are satisfied by common modulation formats. The limitation of [23] is the single-polarization basis of the derivation. An intermediate result of the dual-polarization derivation is presented in [24], which we use to investigate the nonlinear interference noise properties in a manner similar to [23].

In this section, we examine the components of the nonlinear interference noise variance from a dual-polarization perspective. The analysis of this section does not seek to produce a convenient analytic model to quantify the nonlinear interference noise, as in Section III we describe a power optimization method that obviates the need for an analytical model by using received signal measurements. Instead, through examining the nonlinear interference noise components we seek to determine the form of the analytic expressions, given that constants can be obtained from measurements while avoiding system model inaccuracies. In this way, we find that the problem of nonlinear power optimization introduced for the Gaussian noise model in [3] retains a convex form using a more general nonlinear noise model.

A. Nonlinear Interference Noise

The Manakov equation (1) describes dual-polarization fiber propagation where the nonlinear operator has been averaged across polarization under the assumption of rapid evolution of the state of polarization

$$\frac{\partial}{\partial z}\bar{E} = \frac{g(z)-\alpha}{2}\bar{E} + i\frac{\beta_2}{2}\frac{\partial^2 \bar{E}}{\partial t^2} - i\frac{8}{9}\gamma|\bar{E}|^2\bar{E}. \tag{1}$$

In (1), $g(z)$ is the position-dependent gain coefficient, α is the fiber attenuation coefficient, β_2 is the dispersion coefficient, γ is the nonlinear coefficient, and $\bar{E}$ is the complex vector-valued electric field.

Let s be the channel index within the fiber. We consider the nonlinear interference noise on a channel of interest, $s = 0$, due to interactions with itself, referred to as the SPM component, and XPM interactions with other channels, $s \neq 0$. We assume for the example calculated, identical normalized pulse shapes $u(z,t)$ that are functions of position z and time t in each polarization, multiplied by complex vector-valued amplitudes $\bar{a} = [a_X, a_Y]^T$. With a time offset of nT_0, where T_0 is the symbol period, we have pulse shape $u_n(z,t) = u(z, t - nT_0)$, and amplitude $\bar{a}_{n,s}$ in channel s. For the channel of interest, we drop the $s = 0$ subscript from the amplitude variables. Variable $\mathcal{E}$ is the pulse energy, and $f(z) = \Psi(z)^2$ is a result of the re-scaling of E into U to remove the attenuation term, where $\Psi(z)$ is defined via

$$\frac{\partial \Psi(z)}{\partial z} = \frac{g(z)-\alpha}{2}\Psi(z). \tag{2}$$

The inter-channel XPM component of the nonlinear interference noise is [24]

$$\begin{aligned}\Delta\bar{a}_{0,\mathrm{XPM},s} = i\frac{8}{9}\gamma\mathcal{E}\int_0^L f(z)\mathrm{d}z \int \mathrm{d}t u_f^*(z,t) \\ \times \sum_{n,k,m}\left(\bar{a}_{m,s}^H\bar{a}_{k,s}\mathrm{I} + \bar{a}_{k,s}\bar{a}_{m,s}^H\right)\bar{a}_n \\ \times u_k(z,t_s)u_m^*(z,t_s)u_n(z,t)\end{aligned} \tag{3}$$

where I is a 2×2 identity matrix, t_s is the retarded time index tracking channel s and which is a function of t, as shown in (4) [23]. Equation (3) is the simplified form assuming that the initial polarization alignment between channel s and channel $s = 0$ is maintained during propagation. The NLIN variance was found to be invariant to the relative polarization rotation of the interfering channel when each channel is power balanced

between polarizations [27]

$$t_s = t - \omega_s \int_0^z \mathrm{d}z' \beta_2 - \delta T_0. \tag{4}$$

Equation (3) can be re-written as

$$\Delta\bar{a}_{0,\mathrm{XPM},s} = i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k,m}\left(\bar{a}_{m,s}^H\bar{a}_{k,s}\mathrm{I} + \bar{a}_{k,s}\bar{a}_{m,s}^H\right)\bar{a}_n C_{n,k,m} \tag{5}$$

where

$$C_{n,k,m} = T_0\int_0^L f(z)\mathrm{d}z \times \int u_f^*(z,t)u_n(z,t)u_k(z,t_s)u_m^*(z,t_s)\mathrm{d}t. \tag{6}$$

The $u_f^*(z,t)$ term in (6) comes from incorporating assumed receiver matched filtering.

The inter-channel XPM nonlinear interference component can be separated into a constant mean contribution, with expected value denoted by $\langle\cdot\rangle$, and a varying noise contribution

$$\delta\bar{a}_{0,\mathrm{XPM},s} = \Delta\bar{a}_{0,\mathrm{XPM},s} - \langle\Delta\bar{a}_{0,\mathrm{XPM},s}\rangle. \tag{7}$$

We assume that messages in each polarization are proper complex random variables, i.e.,

$$\langle\bar{a}_k^T\bar{a}_{k'}\rangle = 0 \tag{8}$$

$$\langle\bar{a}_k^H\bar{a}_{k'}\rangle = \langle|a_{k,X}|^2 + |a_{k,Y}|^2\rangle\,\delta_{k,k'}. \tag{9}$$

We further assume that symbols on different channels are uncorrelated. Assuming average transmission power is static, $\langle|\bar{a}_{k,s}|^2\rangle = \langle|\bar{a}_s|^2\rangle$.

Using these assumptions, we find that

$$\langle\Delta\bar{a}_{0,\mathrm{XPM},s}\rangle = i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k}\left(\langle|\bar{a}_s|^2\rangle\mathrm{I} + \begin{bmatrix}\langle|a_{s,X}|^2\rangle & 0\\ 0 & \langle|a_{s,Y}|^2\rangle\end{bmatrix}\right) \times \bar{a}_n C_{n,k,k}. \tag{10}$$

We find $\delta\bar{a}_{0,\mathrm{XPM},s}$ by substituting (10) and (5) into (7).

We find the variance of the XPM nonlinear interference noise due to channel s interfering with the channel of interest reduces to

$$\begin{aligned}\langle|\delta\bar{a}_{0,\mathrm{XPM},s}|^2\rangle = {}& \frac{64}{81}\gamma^2\frac{\mathcal{E}^2}{T_0^2}\sum_n\left[\sum_{k\neq m}2|C_{n,k,m}|^2\right.\\ &\times\left(\langle|\bar{a}_s|^2\rangle^2\mathrm{I} + \begin{bmatrix}\langle|a_{s,X}|^2\rangle^2 & 0\\ 0 & \langle|a_{s,Y}|^2\rangle^2\end{bmatrix}\right)\\ &+\sum_k|C_{n,k,k}|^2\left(\left(\langle|\bar{a}_s|^4\rangle - \langle|\bar{a}_s|^2\rangle^2\right)\mathrm{I}\right.\\ &\left.\left.+3\begin{bmatrix}\langle|a_{s,X}|^4\rangle - \langle|a_{s,X}|^2\rangle^2 & 0\\ 0 & \langle|a_{s,Y}|^4\rangle - \langle|a_{s,Y}|^2\rangle^2\end{bmatrix}\right)\right]\\ &\times\begin{bmatrix}\langle|a_{n,X}|^2\rangle\\ \langle|a_{n,Y}|^2\rangle\end{bmatrix}.\end{aligned} \tag{11}$$

Equation (11) is the dual-polarization derived version of [23, eq. (70)]. In (11), we observe that $\langle|X|^4\rangle \geq \langle|X|^2\rangle^2$ for any random variable X, by Jensen's inequality [28]. Given scalar b, $\langle|b\bar{a}_s|^4\rangle = |b|^4\langle|\bar{a}_s|^4\rangle$ and $\langle|b\bar{a}_s|^2\rangle^2 = |b|^4\langle|\bar{a}_s|^2\rangle^2$. Thus, for any given modulation format, the fourth moment is linearly proportional to the square of the second moment: $\langle|\bar{a}_s|^4\rangle = d\langle|\bar{a}_s|^2\rangle^2$ for some scalar d.

When polarization powers are unequal, such as when there is polarization-dependent loss in the fiber, nonlinear interference generation is unequal per polarization. In such a scenario, the optimal powers will vary per polarization, but taking advantage of that is a challenge with a dynamic fiber polarization state. For this paper we assume that both polarizations have identical powers in each channel. This assumption also satisfies the requirement for the NLIN to be invariant to the relative polarization rotation between the interfering channel and the channel of interest [27].

When the X- and Y-polarization powers are equal, the variance of the XPM nonlinear interference noise due to channel s interfering with the channel of interest, (11), is a function of the total powers of the two channels. This nonlinear interference noise experienced by the channel of interest, $s = 0$, can be reduced to a single term with a non-negative real coefficient $D_{s,0}$ that combines all of the pulse collision $|C_{n,k,k}|^2$ coefficients and the modulation format dependence of the fourth moment's relative proportionality to the square of the second moment. Thus, when polarizations are power-balanced, (11) reduces to

$$\langle|\delta\bar{a}_{0,\mathrm{XPM},s}|^2\rangle = D_{s,0}\langle|\bar{a}_s|^2\rangle^2\langle|\bar{a}_0|^2\rangle \tag{12}$$

with equal nonlinear noise in each polarization.

As $|C_{n,k,k}|^2$ is non-negative and $\langle|\bar{a}_s|^4\rangle - \langle|\bar{a}_s|^2\rangle^2 \geq 0$, $D_{s,0}$ of (12) is non-negative.

The self-interference component of the nonlinear interference noise is [23] [24]

$$\Delta\bar{a}_{0,\mathrm{SPM}} = i\frac{8}{9}\gamma\mathcal{E}\int_0^L f(z)\mathrm{d}z\int\mathrm{d}t u_f^*(z,t)\sum_{h,m,n}\bar{a}_h^H\bar{a}_m\bar{a}_{n,0} \times u_0^*(z,t-hT_0)u_0(z,t-mT_0)u_0(z,t-nT_0). \tag{13}$$

The variance of the self-channel interference (13) can be evaluated similarly to the XPM case, and depends upon breaking down a six-dimensional summation as described in [29] for a frequency-domain first-order perturbation approach.

The self-interference component is a function of the cube of the power of the channel of interest, and its variance is non-negative. When the X- and Y-polarization powers are identical, this leads to the simple expression (14) for the self-induced nonlinear interference noise with a non-negative real coefficient D_0 that describes the process efficiency

$$\langle|\delta\bar{a}_{0,\mathrm{SPM}}|^2\rangle = D_{0,0}\langle|\bar{a}_0|^2\rangle^3. \tag{14}$$

The expressions for $D_{0,0}$ and $D_{s,0}$ have complicated expressions that are infeasible to evaluate analytically, which are omitted here, as their values can be obtained via measurement without any theoretical channel model assumptions, as described in Section III. The important result regarding $D_{0,0}$ and $D_{s,0}$ is that

they are non-negative real values, allowing the analysis of [3] to apply with a more general nonlinearity model.

B. Phase Noise Component

Nonlinear phase noise can be a significant impairment in dispersion-compensated systems, unlike in dispersion-uncompensated systems [30]. Accounting for differences in amplitude and phase noise is important to allow a nonlinear noise model to accommodate legacy systems. Such systems may include undersea cables, where the dispersion profile cannot be adjusted. Dispersion-uncompensated systems also exhibit non-circular noise under certain circumstances.

For modern dispersion-uncompensated systems, the Gaussian noise model [1] provides a variance for complex circular nonlinear noise. The enhanced Gaussian noise model [2] incorporates corrections from the higher-order field moments due to non-Gaussian modulation, and also considers nonlinear phase noise, which is found to be insignificant for dispersion-uncompensated systems. Higher-order field moments are also considered in [24], where nonlinear phase noise is found to be significant.

We also seek to verify that the phase noise component of the nonlinear noise, which is often greater than the amplitude noise component, supports a convex form for power optimization in the same manner as the overall SNR.

Using (7), the time-varying component of the inter-channel nonlinear noise due to channel $s \neq 0$ is found to be

$$\begin{aligned}\delta\bar{a}_{0,\mathrm{XPM},s} = {} & i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k,k'}\left(\bar{a}_{k,s}^H\bar{a}_{k',s}\mathrm{I} + \bar{a}_{k',s}\bar{a}_{k,s}^H\right)\bar{a}_{n,0}C_{n,k,k'} \\ & - i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k}\left(\langle|\bar{a}_s|^2\rangle\mathrm{I} + \begin{bmatrix}\langle|a_{s,X}|^2\rangle & 0 \\ 0 & \langle|a_{s,Y}|^2\rangle\end{bmatrix}\right) \\ & \times \bar{a}_{n,0}C_{n,k,k}. \end{aligned} \tag{15}$$

Looking at only the X-polarization component of (15), we find that

$$\begin{aligned}\delta\bar{a}_{0,\mathrm{XPM},s,X} = {} & i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k,k'}\left(\bar{a}_{k,s}^H\bar{a}_{k',s}a_{n,X}\right. \\ & \left. + a_{k',s,X}\bar{a}_{k,s}^H\bar{a}_{n,0}\right)C_{n,k,k'} \\ & - i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0}\sum_{n,k}\left(\langle|\bar{a}_s|^2\rangle + \langle|a_{s,X}|^2\rangle\right)a_{n,X}C_{n,k,k}. \end{aligned} \tag{16}$$

We are interested in the phase noise of the X-polarization component of symbol $\bar{a}_0$. If we restrict our analysis to QPSK or higher-order modulation formats, which is reasonable due to the small nonlinearity assumed for the originating first-order perturbation approach, the only pure phase noise components are those terms containing $ia_{0,X}$ multiplied by real values. As we assumed in (8) and (9) that our messages are proper complex random variables, terms containing $a_{n\neq 0,X}$ or $a_{n,Y}$ have symmetric phase distributions that are uncorrelated with $a_{0,X}$.

We thus find that the pure phase noise component of (16) is

$$\begin{aligned}\delta\bar{a}_{0,\mathrm{XPM},s,X,\text{phase-only}} = {} & i\frac{8}{9}\gamma\frac{\mathcal{E}}{T_0} \\ & \times a_{0,X}\left(\sum_k\left(|\bar{a}_{k,s}|^2 + |a_{k,s,X}|^2 - \langle|\bar{a}_s|^2\rangle - \langle|a_{s,X}|^2\rangle\right)C_{0,k,k}\right. \\ & \left. + \sum_{k\neq k'}\times\left(\bar{a}_{k,s}^H\bar{a}_{k',s} + a_{k,s,X}^*a_{k',s,X}\right)C_{0,k,k'}\right). \end{aligned} \tag{17}$$

The coefficients $C_{0,k,k}$, defined in (6), of the first summation of (17), are purely real with matched filtering. For the second summation, each term is present with its complex conjugate obtained by swapping k and k', as (6) gives $C_{0,k,k'}^* = C_{0,k',k}$ assuming matched filtering, so the overall sum is purely real.

Due to the phase symmetry of the proper random messages, the remaining terms that combine phase and amplitude noise contribute equally to both. The average phase variance is equal to the variance from the pure phase component (17) plus half of the variance from the remaining components and cross terms. Equivalently, the phase variance can be calculated as half of the X-polarization variance (16) plus the remaining half of the variance due to the pure phase component (17), giving

$$\begin{aligned}\langle|\delta\bar{a}_{0,\mathrm{XPM},s,X,\text{phase}}|^2\rangle = {} & \frac{1}{2}\langle|\delta\bar{a}_{0,\mathrm{XPM},s,X}|^2\rangle \\ & + \frac{1}{2}\langle|\delta\bar{a}_{0,\mathrm{XPM},s,X,\text{phase-only}}|^2\rangle. \end{aligned} \tag{18}$$

Evaluating the variance of (17), we find

$$\begin{aligned}\langle|\delta\bar{a}_{0,\mathrm{XPM},s,X,\text{phase-only}}|^2\rangle = {} & \frac{64}{81}\gamma^2\frac{\mathcal{E}^2}{T_0^2}\langle|\bar{a}_{0,X}|^2\rangle \\ & \times\left[\sum_k\left(\langle|\bar{a}_s|^4\rangle + 2\langle|\bar{a}_s|^2|a_{s,X}|^2\rangle + \langle|\bar{a}_{s,X}|^4\rangle\right)|C_{0,k,k}|^2\right. \\ & \left. - \sum_k\frac{9}{4}\langle|\bar{a}_s|^2\rangle^2|C_{0,k,k}|^2 + \sum_{k\neq m}\frac{5}{4}\langle|\bar{a}_s|^2\rangle^2|C_{0,k,m}|^2\right]. \end{aligned} \tag{19}$$

The third summation term of (19) with $k \neq m$ originates from terms of (17) where $n + k - k' \neq 0$, which are found to be zero in [23] for sinc pulses, due to lack of pulse overlap in the C coefficient integral (6). Assuming that the X-polarization power is half of the total channel power, we observe the pure phase noise variance (19) being small or zero for QPSK modulation of interfering channels, where $\langle|\bar{a}_s|^4\rangle = \langle|\bar{a}_s|^2\rangle^2$. For complex Gaussian modulation of interfering channels, where $\langle|\bar{a}_s|^4\rangle = 2\langle|\bar{a}_s|^2\rangle^2$, the pure phase noise term is observed. These observations agree with the single-polarization analysis and simulation of [31].

The long-time correlations of the nonlinear phase noise contributions observed in [31] can be partially compensated through carrier recovery and adaptive equalization in a receiver. This supports the idea that optimization based upon statistics of the theoretical optical channel model is inferior to receiver measurements of the actual nonlinear impairment that a channel experiences after equalization.

For equal power in each polarization, looking at the summation terms of (19) we find $\langle|\bar{a}_s|^4\rangle + 2\langle|\bar{a}_s|^2|a_{s,X}|^2\rangle + \langle|\bar{a}_{s,X}|^4\rangle - \frac{9}{4}\langle|\bar{a}_s|^2\rangle^2 \geq 0$ for each k. Thus, we find that for each k, we have a non-negative coefficient on the power variable $\langle|\bar{a}_s|^2\rangle$, supporting a convex form for power optimization [3].

III. Live System Optimization

Channel power optimization of SNR has been demonstrated to have a convex form under the Gaussian noise model [3]. The main limitation of such optimization is any discrepancy between the theoretical model and the underlying system. Model-based optimization can only solve for the optimum of the model, which may fail to optimize the underlying system in the presence of modeling errors.

Model errors, and the resulting imperfection of optimized channel powers, can be avoided if the underlying system is optimized directly. This requires that gradients of nonlinear noise power or SNR are obtained through system measurements. Such gradients can be measured directly at each optimization step, or can be used to calculate the nonlinear interference strength coefficients between different channels, as used in (12). These coefficients can be subsequently used to calculate the series of gradients during the optimization process.

A. Obtaining Nonlinear Coefficients From SNR Gradients

Let us assume that the total nonlinear noise is due entirely to SPM or XPM contributions, neglecting four-wave-mixing (FWM). As XPM dominates FWM outside of fully dispersion-shifted fibers [32], this error is small. Given a convex optimization objective, the gradient defines a tangential hyper-plane to the optimization surface [28] such that errors in the gradient direction need to be greater than 90 degrees in order for the approximate gradient to avoid improving the objective. As the total nonlinear noise variance is measurable, the error introduced from accounting for all inter-channel nonlinear interference noise contributions as XPM will slightly change gradient directions, but will not mis-characterize the total nonlinear noise and optimal power allocation. The benefit of neglecting FWM is that the number of perturbation measurements required to learn the full set of nonlinear interference coefficients is reduced from quadratic to linear in the number of channels.

In Section II, the analysis was performed for the nonlinear interference on the channel of interest, $s = 0$. This analysis generalizes for arbitrary channel of interest $s = n \in [1 \dots N]$ for a system with N channels. The power of channel n is denoted by e^{y_n}, as logarithmic power variables are necessary to maintain a convex form and avoid having local minima in the channel SNR expression [3]. Assuming only SPM and XPM nonlinearity components, the nonlinearity in channel n is

$$NL_n(e^{\boldsymbol{y}}) = \sum_{s=1}^{N} D_{s,n} e^{2y_s + y_n} \tag{20}$$

where $D_{s,n}$ is the nonlinear interference coefficient for channel n, experiencing interference due to SPM or XPM from channel s, as in (12) and (14).

Using $\boldsymbol{y}$ to denote the vector of logarithmic channel powers, the optical SNR of channel n is a function of (20), given by

$$\text{SNR}_n(e^{\boldsymbol{y}}) = \frac{e^{y_n}}{\sigma_n^2 + NL_n(e^{\boldsymbol{y}})}. \tag{21}$$

While referred to here as SNR, this is more rigorously a signal to noise-and-uncompensated-distortions ratio, where the nonlinear distortions that remain uncompensated are treated as noise. In order to obtain the optical SNR from measurements at the receiver, the transceiver noise contributions will need to be quantified and removed. Alternative formulations are possible based upon received electrical SNR, but the optimal optical power depends only on the optical SNR component of that electrical SNR.

Using $\mathbf{e}_n$ as the unit vector for dimension n, the gradient of the SNR of channel n (21) is

$$\begin{aligned}\nabla\text{SNR}_n(e^{\boldsymbol{y}}) &= \frac{\mathbf{e}_n e^{y_n}}{(\sigma_n^2 + NL_n)} - \frac{\sum_{s=1}^{N} D_{s,n} e^{2y_s + 2y_n}(2\mathbf{e}_s + \mathbf{e}_n)}{(\sigma_n^2 + NL_n)^2} \\ &= \mathbf{e}_n \text{SNR}_n - \text{SNR}_n^2 \sum_{s=1}^{N} D_{s,n} e^{2y_s}(2\mathbf{e}_s + \mathbf{e}_n) \\ &= \mathbf{e}_n \text{SNR}_n - \text{SNR}_n^2 B^{(n)} D_n \end{aligned} \tag{22}$$

where D_n is a vector of nonlinear interference coefficients for channels interfering via XPM or SPM with channel n, the matrix

$$B^{(n)} = \begin{bmatrix} e^{2y_1} & & 0 \\ & \ddots & \\ 0 & & e^{2y_N} \end{bmatrix} (2\mathrm{I} + \mathbf{e}_n \mathbf{1}^T) \tag{23}$$

and $\mathbf{1}$ is an N-dimensional vector of ones.

Equation (22) can be inverted when $B^{(n)}$ (23) is full-rank, allowing the SPM and XPM interference components to be obtained from measurements of the SNR gradient. $B^{(n)}$ is always full-rank when all channels have non-zero power. When a channel power is zero, it is logical that no measurement of the nonlinear interference strength involving that channel is possible. Excluding channels with zero powers leads to a smaller $B^{(n)}$ matrix that is invertible, and a lower-dimensional problem. Inverting (22) gives the vector of interference coefficients D_n as

$$D_n = {B^{(n)}}^{-1} \frac{\text{SNR}_n \mathbf{e}_n - \nabla\text{SNR}_n}{\text{SNR}_n^2}. \tag{24}$$

The SNR gradient can be approximated as the slope of SNR changes for small power changes, following the limit definition of the derivative. This leads to a training sequence-based method for measuring the matrix of nonlinear interference coefficients $D = [D_1 \cdots D_N]$. First, the optical SNRs are measured for each channel given their initial power allocations. Next, each logarithmic channel power variable y_n is slightly increased and optical SNRs are measured for each channel before the power change is reverted, and the next channel is changed. This training sequence set has length N for a system with N channels.

In a live system, the extremely high symbol rates allow very large sample sets to be accumulated, minimizing sampling errors in the noise variance. Testing this approach with split-step Fourier simulations is a challenge, as the standard deviation of the SNR measurements is proportional to $\frac{1}{\sqrt{M}}$ given M samples. For the amplified spontaneous emission component of the noise which is Gaussian, the sample variance of the noise is itself a chi-squared random variable with M degrees of freedom, given the known distribution, and variance $\frac{2\sigma^4}{M}$. Using (24), we find that errors in the nonlinear interference coefficients D are proportional to the SNR measurement error.

B. System Optimization

The optimization objective of maximizing the minimum channel margin by adjusting channel powers is not differentiable due to the use of the minimum function [3]. This objective can be transformed into a differentiable form through the addition of a slack variable and use of barrier functions [3]. For theoretical model-based optimization, the repeatable noiseless model predictions allow descent methods to be used for rapid convergence. Systems using real-world measurements of nonlinear interference, or measurements obtained from simulations, break the conditions required for methods such as gradient descent to act as descent methods. Stochastic subgradient optimization methods are able to guarantee convergence for a convex objective with unbiased noise [33]. Convergence is guaranteed for such objectives when optimization step sizes sum to infinity, but the sum of the squares of the step sizes is bounded, and where the subgradient variance is bounded [33] [34]. When the objective is differentiable and noiseless, the subgradient direction is the same as the direction of the gradient.

While convergence of the stochastic subgradient method is guaranteed in the limit of optimization steps, the addition of noise to the subgradient steps further increases the number of optimization steps required beyond that of the already slow subgradient method [3] [34] [35]. Given the very high symbol rate of optical communication modems, accurate nonlinearity measurements can be taken rapidly, and hundreds of iteration steps can be performed in a relatively short time period. By contrast, simulating the optimization of a system using the split-step Fourier method requires extensive computation, even for a few channels.

Quantization and transceiver noise will be present in a receiver measurement of SNR. Assuming proper analog gain control, such noise components are independent of signal power. These noise components need to be quantified through calibration in order to isolate the optical SNR due to amplifier noise and nonlinearity, upon which the power optimization operates. When the optical SNR is very high, such that the receiver measurement of SNR contains a significant proportion of transceiver noise, the error in the calculated optical SNR will grow, making longer averaging times necessary. The trade-off between optimization convergence time and accuracy will dictate the constraints on measurement accuracy.

The nonlinear coefficients (24) complete the nonlinear interference noise model (20). When performing optimization, the gradient of the channel SNRs is used for the stochastic subgradient method. The SNR gradient can be obtained from perturbation measurements, but the fastest method, lending itself to optimization simulations, is to use (22).

Optimizing the minimum channel margin requires coordination at least at the level of communication modems. Nonlinear interference noise gradients must be measured at each optimization step, or nonlinear interference coefficients must first be determined through a training sequence of power variations described in Section III-A. Optimization requires identifying the minimum-margin channel at each step, and calculating the subgradient given that minimum-margin channel. This coordination can be achieved with communication between system modems for a point-to-point link. Performing such optimization in a mesh scenario requires knowledge of the system layout and channel paths, lending itself to a higher-level control system.

The nonlinear interference noise observed contains time correlations between symbols [24]. If these correlations are not taken advantage of through joint symbol decoding, blocks of independent symbols for error correction can be obtained by using interleavers with lengths longer than the correlation time.

A typical receiver design features final bitwise decoding under an additive white Gaussian noise channel assumption [36]. In operating regimes where the nonlinear interference noise is independent identically distributed Gaussian or well approximated as Gaussian [1], maximizing the channel SNR maximizes decoding performance at high SNR. At low SNR, the particular modulation format labeling becomes extremely significant. For arbitrary nonlinear interference distributions, maximizing the SNR is generally a good indicator of system performance, but there exist distributions where that is not the case, as the Generalized Mutual Information (GMI) for bitwise decoded receivers is not always a concave function of SNR [37, 4.4.1]. GMI gives an accurate performance metric that accounts for coded modulation and labeling impacts [36]. The GMI in this case operates under the assumption that uncompensated nonlinear distortions are treated as noise. In general, the signal-derived nonlinear distortions are deterministic, but it is often infeasible to compensate for them.

Analytic calculation of the GMI of an optical link is only tractable when the nonlinear interference noise follows a well-defined distribution, such as Gaussian. For more general application, the GMI can be estimated empirically from received symbol measurements obtained from a real system or a propagation simulation. The GMI is estimated as [36]

$$GMI \approx m - \frac{1}{n_s}\sum_{k=1}^{m}\sum_{l=1}^{n_s}\log_2\left(1+e^{(-1)^{c_{k,l}}L_{k,l}}\right) \tag{25}$$

where the bitwise decoder features m bits, there are n_s sample symbols, $c_{k,l}$ is the transmitted bit number k of symbol l, and $L_{k,l}$ is the log-likelihood ratio of the received symbol l for bit k given by (26). For a receiver assuming an additive white Gaussian noise channel, the log-likelihood ratios can be calculated

as [36]

$$L_{k,l} = \log \frac{\sum_{x \in \mathcal{X}^1_{k,l}} \exp(-\frac{1}{\sigma^2}|y_l - x|^2)}{\sum_{x \in \mathcal{X}^0_{k,l}} \exp(-\frac{1}{\sigma^2}|y_l - x|^2)} \tag{26}$$

where σ^2 is the noise variance, y_l is received symbol l, $\mathcal{X}^1_k$ is the set of constellation points where bit k decodes to 1, and $\mathcal{X}^0_k$ is the set of constellation points where bit k decodes to 0.

The predictive ability of empirical GMI measurements is limited by the number of symbols used in the measurement. The optimization objective of this paper is formed using the SNR rather than empirical GMI, leaving the GMI as an observational metric. The combination of empirical GMI measurement noise, the lack of concavity of the GMI, and the inability to calculate efficient gradients of the empirical GMI are all reasons why SNR was instead used to form the optimization objective. Sample noise in the empirical GMI measurements dominates the fine power optimization in the results of this paper, but the metric is included in this paper, as it is the best predictor of system performance under non-Gaussian noise given a large number of empirical samples.

IV. Results

We demonstrate the process of power optimization to maximize the SNR of the minimum-SNR channel using measurements of signals by receivers. Split-step Fourier propagation simulations are used to model the nonlinear signal propagation of the signals. Only eight channels can be simulated due to the high level of computational complexity presented by the large number of optimization steps and the large (2^{17}) symbol count per channel required for accurate SNR measurements, despite the benefit of GPU acceleration. The SNR is measured from the middle half of the symbols in each polarization following residual dispersion compensation, carrier recovery, and adaptive equalization. A full floating-point receiver is used, avoiding the need to quantify transceiver and quantization noise.

We simulate a dispersion-compensated system with low-dispersion fiber in order to emphasize a non-circular non-Gaussian nonlinear noise distribution. The simulated system features five 80 km spans of fiber with sloped dispersion ranging from 2 to 5 ps/nm/km over the C-band, each followed by a realistically matched dispersion-compensating module [38]. Fiber with attenuation coefficient $\alpha = 0.21$ dB/km and nonlinear coefficient $\gamma = 1.4\ \mathrm{W}^{-1}\ \mathrm{km}^{-1}$ is used. Amplifiers following each dispersion compensating module feature 6 dB noise figures and a 0.64 dB tilt to the noise spectrum across the eight simulated channels.

Fig. 1 shows the convergence of channel powers while optimizing the minimum channel SNR using a stochastic subgradient optimization method. As the subgradient method is not a descent method, there is no smooth convergence to the optimal power levels. This is exacerbated by the noise present in this stochastic optimization using measurements from split-step Fourier method propagation simulations. There is a sharp increase in a channel's power whenever that channel becomes the one with the minimum SNR, and otherwise a slow decay in channel power is observed as the SNR in the minimum channel is improved. The step sizes decay throughout the optimization process to converge toward the optimal power levels.

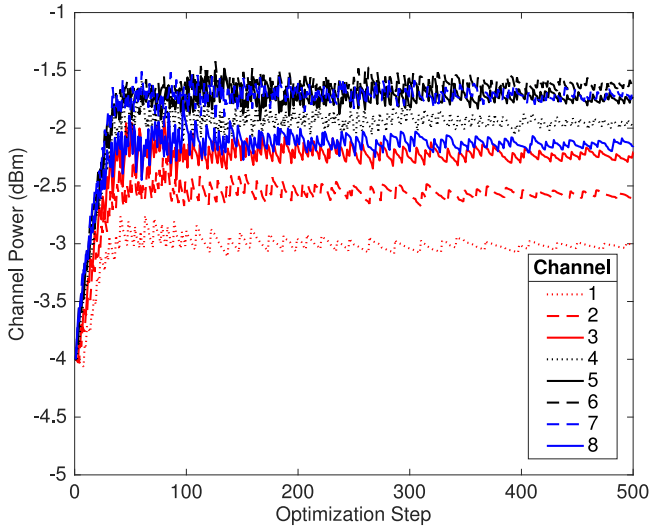

Fig. 1. Evolution of channel powers versus optimization step in order to maximize the minimum SNR. Initial channel powers of −4 dBm separate and converge towards optimal levels through optimization. The combination of inter-channel nonlinear interactions, a 0.64 dB ASE noise tilt across the eight channels, and dispersion slope in the simulated fiber system lead to differing optimal powers per channel. Simulation parameters are: dual-polarization 16-QAM modulation, five 80 km spans of fiber with dispersion ranging from 2 to 5 ps/nm/km over the C-band, each followed by dispersion compensating fiber sections. The fiber has attenuation coefficient $\alpha = 0.21$ dB/km, and nonlinear coefficient $\gamma = 1.4\ \mathrm{W}^{-1}\ \mathrm{km}^{-1}$. Amplifiers with a 6 dB noise figure are used.

Fig. 1 shows the system channel powers at each step of the stochastic subgradient optimization. Significant variation can be observed in the optimal channel powers. The dominant contribution is due to the 0.64 dB tilt in the amplifier noise spectrum across the eight channels. Dispersion slope in the simulated low dispersion fiber and the dispersion compensating module represents another contribution. The last observable contribution to the optimal channel powers is the number of neighboring channels leading to nonlinear interference. Channels numbered 1 and 8 are at the two edges of the band and only have neighbors on one side. This leads to channel 8 having an optimal power level between that of channels 3 and 4, despite having a higher amplifier noise level.

Fig. 2 shows the SNR convergence during the process of power optimization shown in Fig. 1. After about 50 optimization steps, most of the SNR gains are achieved as approximately optimal channel powers are obtained. There remains a wide range of channel SNR levels after 50 optimization steps, with the minimum channel SNR varying from 16 dB to 16.2 dB while the maximum is around 16.5 dB. Further optimizing the minimum channel SNR as the step sizes decrease allows the minimum channel SNR to increase to a consistent 16.3 dB. The excess SNR of the high-SNR channels around step 50 is given up to raise the minimum-SNR channel, as can be observed from the decrease in maximum SNR in Fig. 2 from optimization step 50 to step 500.

Fig. 3 shows the improvement in generalized mutual information through the optimization of the minimum channel SNR.

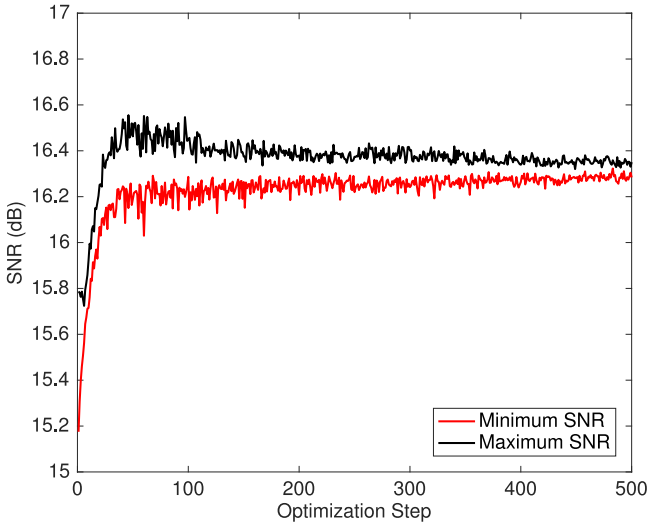

Fig. 2. Minimum and maximum SNRs among the band of channels is plotted versus optimization step or order to maximize the minimum SNR. Optimization maximizing the minimum channel SNR raises the minimum SNR while taking away from non-minimum channels. The convergence is not monotonic due to the subgradient optimization method required and sample noise due to the split-step Fourier simulation. The separation of the minimum and maximum SNR lines at the end of optimization are separated by sample noise limitations. Simulation parameters are as in Fig. 1.

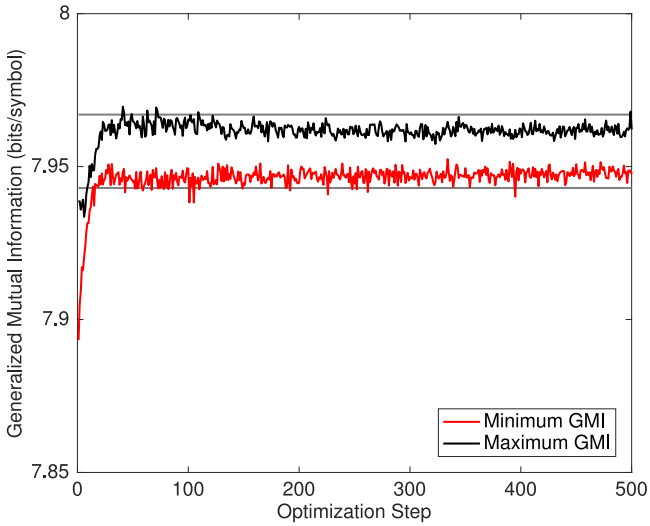

Fig. 3. GMI versus optimization step while maximizing the minimum SNR. GMI serves as a good metric for system performance in a soft decision bitwise decoded receiver [36]. Optimizing the minimum channel SNR based on overall SNR in this dispersion-compensated system with non-circular non-Gaussian nonlinear noise optimizes the GMI. Sample noise in the empirical GMI leads to the residual separation between minimum and maximum GMI. Simulation parameters are as in Fig. 1.

Empirical estimates of the generalized mutual information are obtained via (25). Following the first 50 optimization steps, the approximate powers are obtained and the maximum GMI is at its highest. The remaining steps optimize the minimum SNR, which leads to a small further improvement in the minimum GMI while the maximum GMI decreases. The level of noise in the empirical GMI measurement remains high at the symbol counts measured, limiting its predictive ability for small power changes.

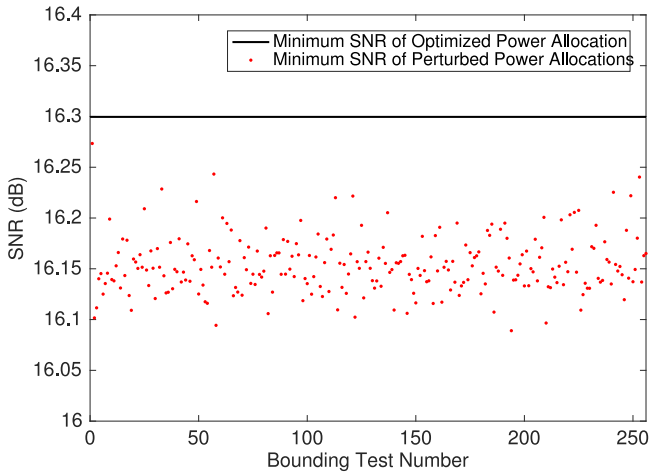

Fig. 4. Minimum channel SNR versus test iteration for power perturbations of the optimized power system power allocation. All permutations of ± 0.1 dB channel power perturbations are tested to verify optimized power allocation. With eight channels, the $2^8 = 256$ binary perturbations provide a bounding box on the optimized power allocation. Optimized power allocation and simulation parameters are as in Fig. 1.

The analysis of Section II indicates that optimizing the minimum channel SNR is convex when the first-order perturbation approach is valid, which is the case when the signal power dominates the nonlinear noise power. In order to verify that our optimization obtains the true optimum, and that sufficient iterations have been taken with the stochastic subgradient method, which does not monotonically improve the objective, we test the SNR of a surrounding box of perturbed power allocations. Fig. 4 shows the minimum channel SNR of the optimized power allocation and the SNR observed at all 2^8 permutations of ± 0.1 dB channel power perturbations from the optimized allocation. As the optimized power allocation improves the minimum SNR relative to the surrounding box of perturbed power allocations, we confirm the optimality of the optimized allocation.

The low-dispersion fiber used with dispersion compensating modules leads to notably non-circular nonlinear noise. Fig. 5 shows the evolution of the amplitude and phase noise SNRs through the optimization process. While the overall SNR converges, as shown in Fig. 2, the ratio of amplitude to phase noise varies for each channel. A constellation plot of the received optimized signal for channel 7, containing both Gaussian amplifier noise and nonlinear noise contributions, is shown in Fig. 6. The circular Gaussian amplifier noise is about 3 dB stronger than the nonlinear noise at the optimal power allocation, due to the cubic nature of the nonlinearity, implying that small differences in the overall amplitude and phase noise are due to proportionally larger changes in the amplitude and phase noise ratios of the nonlinear noise.

Fig. 7 shows the margin improvement from nonlinear power optimization for a 24-channel system with alternating 16-QAM and 32-QAM channels. Due to the increased channel count, a smaller number of symbols are simulated per channel, leading to increased sample noise. An optimized power allocation gives a 0.25 dB improvement in minimum margin by taking mar-

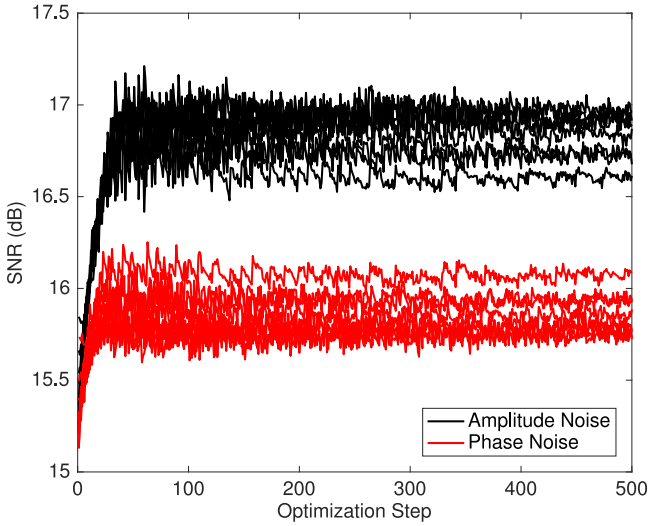

Fig. 5. Amplitude and phase SNR components of the band of channels plotted versus optimization step while optimizing the minimum SNR. The phase and amplitude noise components maintain distinct magnitudes during optimization. While the overall SNR converges, the two noise components do not converge between channels. Simulation parameters are as in Fig. 1.

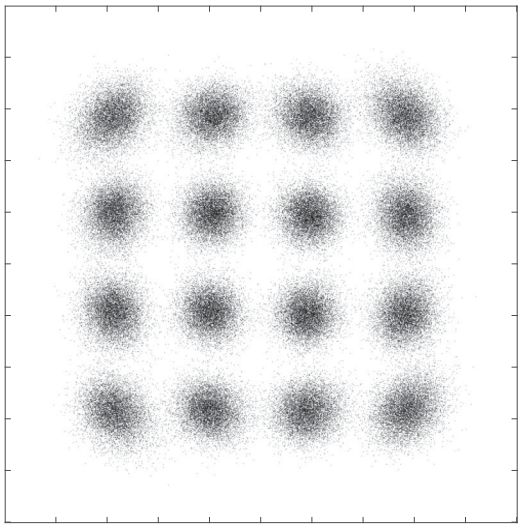

Fig. 6. Noisy 16 QAM constellation plot for one of the channels at the optimized power level demonstrating signal-dependent phase noise and noncircular noise components, particularly for the highest-power symbols. The figure shows the raw constellation points from the termination of the simulation of Fig. 1.

gin away from channels with excess margin. The fixed-ratio one-dimensional power optimization comparison maintains a linear relationship between the powers assigned to the channels with the two different modulation formats proportional to their ratio of required SNRs. Such a fixed-ratio optimization is one-dimensional in a point-to-point link, but the dimensionality scales with the number of fiber paths in a mesh network due to varying total path lengths per channel. The best spectrally flat power allocation has minimum margin 0.9 dB below the optimized power allocation. This scenario assumes standard single-mode fiber and a 1.92 dB amplifier noise tilt that provides significant excess margin to the channels with less noise. This particular alternating allocation of rates would naturally arise

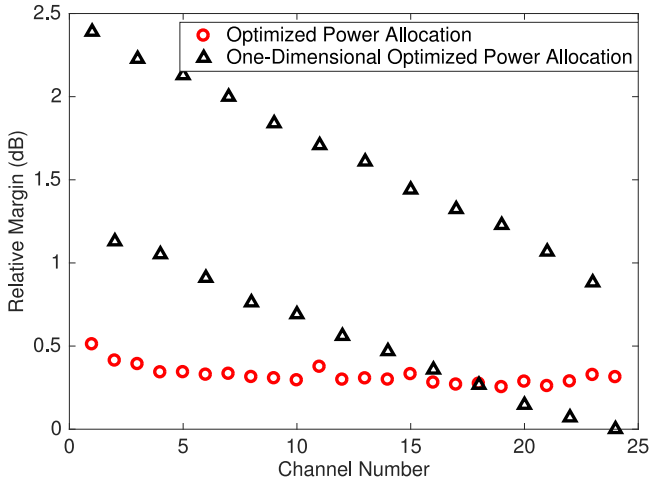

Fig. 7. Channel margins relative to the minimum margin of a one-dimensional power optimization for a 24-channel system with alternating 16-QAM and 32-QAM modulated channels. A minimum margin improvement of 0.25 dB is obtained from nonlinear power optimization over a one-dimensional optimized power allocation with a fixed ratio between 16-QAM and 32-QAM channel powers, corresponding to the ratio of their required SNRs. Simulation fiber parameters are: five 80 km spans of fiber with 17 ps/nm/km of dispersion, attenuation coefficient $\alpha = 0.21$ dB/km, and nonlinear coefficient $\gamma = 1.4$ W^{-1} km^{-1}. Amplifiers with a 6 dB noise figure are used. Tilted ASE noise is used to provide an SNR inequality in the system which can be compensated for via nonlinear power optimization.

when maximizing data throughput in a system that delivers an SNR level between the required SNR levels for 16-QAM and 32-QAM modulation formats.

V. Discussion

In Section IV, we demonstrated the optimization of a particular system with non-Gaussian nonlinear noise. The optimality of the solution was demonstrated by testing a bounding box of small power changes, which verified the optimization approach. Within the small-nonlinearity (high-SNR) regime where the first-order perturbation approach is accurate, SNR optimization has a convex form, yielding a unique global optimum. Whether real measurements, or a theoretical model, are used for determining optimization steps, the shape of the nonlinear interference noise function determines the structure of the SNR optimization problem.

While the uniqueness of the optimal solution and convergence guarantee are provided by the nonlinear noise model, this method can be applied using direct system measurements, avoiding any biases from imperfect modeling. Channel inequalities due to inter-channel stimulated Raman scattering, amplifier gain tilt or ripple, frequency-dependent fiber parameters, and others, can be sources of system modeling errors. Such variations in channel performance can be alleviated with power optimization, leading to performance gains from optimization.

A 0.25 dB minimum-margin gain was observed for a 24-channel system with a significant level of channel inhomogeneity. The available SNR gains for a fully filled C-band are expected to be larger following the observations of [3]. In particular, in mesh networks, where different channels propagate different distances over different paths, significant channel inho-

mogeneity is observed, leading to substantial gains from nonlinear power optimization. Amplifier tilt and ripple, wavelength-dependent fiber parameters, and stimulated Raman scattering [39], are all sources of channel inhomogeneity that will become more significant with more channels, increasing gains available from optimization. The tilted noise used in this paper allows the performance gains in the presence of such inhomogeneities to be demonstrated at smaller channel counts.

GMI serves as a definitive performance metric for a bitwise-decoded receiver. As the GMI cannot be quantified analytically without a well-defined nonlinear noise distribution, we depend upon the empirical estimation of the GMI (25). Through extensive A/B testing, gradients of the empirical GMI could be obtained, but as the GMI is not always a concave function of SNR [37, 4.4.1], power optimization of GMI is very likely to converge to a local minimum rather than the global optimum.

The use of the first-order perturbation based nonlinear model generalizes the class of systems for which convex channel power optimization is feasible beyond the domain of the Gaussian noise nonlinearity model. Arbitrary fiber types and dispersion profiles are supported, with the limitation of small nonlinearity restricting application to high spectral efficiency systems. The assumption made in Section III-A, that all nonlinear noise components are due to SPM or XPM in order to reduce the number of perturbation tests required from N^2 to N, further limits the application to systems that are not FWM-dominated.

The first system simulated here employs low-dispersion fiber and dispersion-compensating modules, while compensating residual dispersion via digital equalization. This fiber link was chosen as an exemplary link with poor nonlinear performance and to emphasize the non-circular non-Gaussian noise components. High-dispersion systems and those without dispersion-compensating modules provide more ideal Gaussian nonlinear noise components. Such systems can be optimized either using the methods of this paper, as shown in Fig. 7, or those based on the Gaussian noise model [3].

The optimization methods are demonstrated for small channel counts due to computational complexity constraints on the split-step Fourier simulations at each optimization step. The optimization methods are fully scalable to any number of channels. The SNR gradient measurement time scales linearly with the number of channels when a power perturbation approach is used. The optimization problem size increases with the number of channels, decreasing convergence speed, but the subgradient-based optimization techniques described in [3] can be used to improve convergence speed for the power optimization problem.

While the stochastic subgradient method guarantees convergence with unbiased noise with a bounded variance [34], practical convergence speeds require small noise levels. The middle 2^{16} dual-polarization symbols over which the SNRs used in the optimization of Figs. 1 and 2 are calculated provides accurate convergence in a few hundred iterations. With smaller SNR measurement windows, the noise in the SNR measurement is greater, and convergence is slower. In a real-world system, such measurements require only a few microseconds of measurement time, assuming the receiver is set up to take continuous SNR measurements.

VI. Conclusion

Utilizing a first-order perturbation approach, we demonstrate that power optimization to improve the nonlinear channel SNR retains a convex form beyond the regime of circular Gaussian nonlinear noise. We observe that power optimization of a theoretical channel model is limited in applicability by the accuracy of the model for the underlying physical system. We propose a method of power optimization for SNR that utilizes physical system measurements and avoids performance penalties due to model inaccuracies. We demonstrate this optimization via a series of split-step Fourier simulations of an eight-channel system with significantly non-circular nonlinear noise. The optimality of the optimized power allocation is verified by testing a surrounding box of perturbed power allocations. The power optimization of a 24-channel system with alternating 16-QAM and 32-QAM channels and ASE noise tilt provides 0.25 dB of minimum margin improvement as compared to a fixed-ratio one-dimensional power allocation.

References

[1] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN model of nonlinear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 32, pp. 694–721, Feb. 2014.

[2] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of nonlinear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, 2014.

[3] 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. 2016.

[4] 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. 2017.

[5] A. R. Chraplyvy, J. A. Nagel, and R. W. Tkach, "Equalization in amplified WDM lightwave transmission systems," *IEEE Photon. Technol. Lett.*, vol. 4, no. 8, pp. 920–922, Aug. 1992.

[6] E. Ciaramella, L. Giorgi, A. D'Errico, F. Cavaliere, G. Gaimari, and G. Prati, "A highly effective technique for setting the power preemphasis in WDM optical systems," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 342–356, Jan. 2006.

[7] Y. Pan and L. Pavel, "A Nash game approach for OSNR optimization with capacity constraint in optical links," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1919–1928, Nov. 2008.

[8] Y. Pan and L. Pavel, "OSNR optimization in optical networks: Extension for capacity constraints," in *Proc. Amer. Control Conf.*, 2005, pp. 2379–2384.

[9] L. Pavel, "OSNR optimization in optical networks: Modeling and distributed algorithms via a central cost approach," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 54–65, 2006.

[10] Y. Pan, T. Alpcan, and L. Pavel, "A system performance approach to OSNR optimization in optical networks," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1193–1200, Apr. 2010.

[11] S. T. Le, K. Blow, and S. Turitsyn, "Power pre-emphasis for suppression of FWM in coherent optical OFDM transmission," *Opt. Express*, vol. 22, no. 6, pp. 7238–7248, 2014.

[12] D. Kaup, "Perturbation theory for solitons in optical fibers," *Phys. Rev. A*, vol. 42, no. 9, pp. 5689–5694, 1990.

[13] V. P. Tzolov, N. Godbout, S. Lacroix, and M. Fontaine, "Nonlinear self-phase-modulation effects: A vectorial first-order perturbation approach," *Opt. Lett.*, vol. 20, no. 5, pp. 456–458, 1995.

[14] R. Hui, M. O'Sullivan, A. Robinson, and M. Taylor, "Modulation instability and its impact in multispan optical amplified IMDD systems: Theory and experiments," *J. Lightw. Technol.*, vol. 15, no. 7, pp. 1071–1082, Jul. 1997.

[15] R. Hui, K. R. Demarest, and C. T. Allen, "Cross-phase modulation in multispan WDM optical fiber systems," *J. Lightw. Technol.*, vol. 17, no. 6, pp. 1018–1026, Jun. 1999.

[16] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.
[17] M. J. Ablowitz and T. Hirooka, "Resonant nonlinear intrachannel interactions in strongly dispersion-managed transmission systems," *Opt. Lett.*, vol. 25, no. 24, pp. 1750–1752, 2000.
[18] P. Johannisson, D. Anderson, A. Berntson, and J. Mårtensson, "Generation and dynamics of ghost pulses in strongly dispersion-managed fiber-optic communication systems," *Opt. Lett.*, vol. 26, no. 16, pp. 1227–1229, Aug. 2001.
[19] E. E. Narimanov and P. Mitra, "The channel capacity of a fiber optics communication system: Perturbation theory," *J. Lightw. Technol.*, vol. 20, no. 3, pp. 530–537, Mar. 2002.
[20] S. Kumar and D. Yang, "Second-order theory for self-phase modulation and cross-phase modulation in optical fibers," *J. Lightw. Technol.*, vol. 23, no. 6, pp. 2073–2080, Jun. 2005.
[21] A. Mecozzi, "A unified theory of intrachannel nonlinearity in pseudolinear phase-modulated transmission," *IEEE Photon. J.*, vol. 2, no. 5, pp. 728–735, Oct. 2010.
[22] Z. Tao *et al.*, "Simple fiber model for determination of XPM effects," *J. Lightw. Technol.*, vol. 29, no. 7, pp. 974–986, Apr. 2011.
[23] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.
[24] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Interchannel nonlinear interference noise in WDM systems: Modeling and mitigation," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.
[25] G. Gavioli *et al.*, "Nrz-pm-qpsk 16×100 gb/s transmission over installed fiber with different dispersion maps," *IEEE Photon. Technol. Lett.*, vol. 22, no. 6, pp. 371–373, Mar. 2010.
[26] A. Ghazisaeidi, J. Renaudier, M. Salsi, P. Tran, G. Charlet, and S. Bigo, "System benefits of digital dispersion precompensation for nondispersion-managed PDM-WDM transmission," in *Proc. Eur. Conf. Opt. Commun.*, 2013, Paper We.4.D.4.
[27] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 593–607, Jan. 2016.
[28] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[29] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 31, no. 8, pp. 1273–1282, Apr. 2013.
[30] S. Kumar, "Effect of dispersion on nonlinear phase noise in optical transmission systems," *Opt. Lett.*, vol. 30, no. 24, pp. 3278–3280, 2005.
[31] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25685–25699, 2013.
[32] S. Ten, K. Ennser, J. Grochocinski, and S. Burtsev, "Comparison of four-wave mixing and cross phase modulation penalties in dense WDM systems," in *Proc. Optical Fiber Commun. Conf.*, 1999, Paper ThC4.
[33] N. Z. Shor, *Nondifferentiable Optimization and Polynomial Problems* (ser. Nonconvex Optimization and its Applications). Norwell, MA, USA: Kluwer, 1998.
[34] S. Boyd and A. Mutapcic, "Stochastic subgradient methods," Notes for EE364b, Stanford Univ., Stanford, CA, USA. Apr. 2014. [Online]. Available: web.stanford.edu/class/ee364b/lectures/stoch_subgrad_notes.pdf
[35] S. Boyd and J. Park, "Subgradient methods," Notes for EE364b, Stanford Univ., Stanford, CA, USA. May 2014. [Online]. Available: web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf
[36] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *J. Lightw. Technol.*, vol. 33, no. 20, pp. 4338–4352, Jan. 2015.
[37] L. Szczecinski and A. Alvarado, *Bit-Interleaved Coded Modulation: Fundamentals, Analysis and Design*. Hoboken, NJ, USA: Wiley, 2015.
[38] "Specification for dispersion slope compensating module for LEAF transmission fiber," OFS Fitel, Norcross, GA, USA. [Online]. Available: http://fiber-optic-catalog.ofsoptics.com/Asset/OFS_LEAFDK.pdf
[39] I. Roberts, J. M. Kahn, J. Harley, and D. 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. 2017.

Ian Roberts received the Bachelor's degree from the University of Ottawa, Ottawa, ON, Canada, and the Master's degree from Stanford University, Stanford, CA, USA, both in electrical engineering. He is currently working toward the Ph.D. degree at Stanford University. He worked with Ciena on DSP hardware development for WaveLogic 3 Extreme prior to beginning graduate studies. His research interests include optical communication, optimization, and compensation of nonlinear propagation.

Joseph M. Kahn (F'00) received the A.B., M.A., and Ph.D. degrees in physics from the University of California, Berkeley (U.C. Berkeley), Berkeley, CA, USA, in 1981, 1983, and 1986, respectively. From 1987 to 1990, he was with AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ, USA. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990 to 2003, he was a faculty with the Department of Electrical Engineering and Computer Sciences, U.C. Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering with Stanford University, where he heads the Optical Communications Group. His current research interests include fiber-based imaging, spatial multiplexing, rate-adaptive and spectrally efficient modulation and coding methods, coherent detection and associated digital signal processing algorithms, digital compensation of fiber nonlinearity, and free-space systems. He was a recipient of the National Science Foundation Presidential Young Investigator Award in 1991. From 1993 to 2000, he was a Technical Editor for the IEEE PERSONAL COMMUNICATIONS MAGAZINE. From 2009 to 2014, he was an Associate Editor for the IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING. In 2000, he helped found StrataLight Communications, where he served as a Chief Scientist from 2000 to 2003. StrataLight was acquired by Opnext, Inc., in 2009.