Channel Power Optimization of WDM Systems Following Gaussian Noise Nonlinearity Model in Presence of Stimulated Raman Scattering

Ian Roberts, *Student Member, IEEE, Member, OSA*, Joseph M. Kahn, *Fellow, IEEE*, James Harley, and David W. Boertjes, *Member, IEEE, Member, OSA*

***Abstract*—The impact of interchannel stimulated Raman scattering (SRS) on optimization of channel powers to maximize the minimum channel margin is examined using a discrete Gaussian noise model for the Kerr nonlinearity. The simultaneous consideration of these two nonlinear effects is found to be incompatible with the goal of a convex SNR expression that can be optimized globally. A sequence of convex optimizations is employed to obtain a locally optimal solution, along with a bound on the degree of suboptimality. Optimization results obtained are most accurate for Gaussian-distributed signals, such as probabilistically shaped high-order-modulated signals. In a dispersion-uncompensated 4000-km fiber system utilizing the full *C*-band with perfect per-span SRS gain compensation, power optimization yields benefits of 0.25 to 2 dB over optimal spectrally flat power allocations. In systems including both *C*- and *L*-band, an optimization method that accounts for both SRS and Kerr nonlinearity effects provides a 0.23 to 0.60 dB margin benefit over a method compensating for SRS gain alone. In a system spanning only the *C*-band, per-span SRS gain compensation is not critical, as the maximum benefit is a 0.14 dB gain in minimum margin for optimized power allocations. By contrast, in a system spanning both *C*- and *L*-band, per-span SRS gain compensation provides a gain of up to 1.23 dB with optimized power allocations and larger gains with suboptimal power allocations.**

***Index Terms*—Gaussian noise model, network optimization, optical communications, stimulated Raman scattering.**

I. Introduction

IN OPTICAL communication systems limited by amplified spontaneous emission (ASE) noise, increasing the signal power increases the signal-to-noise ratio (SNR), but the feasible power is constrained by the Kerr nonlinearity. The Gaussian noise (GN) model provides an efficient analytic form for this nonlinearity in dispersion-uncompensated transmission systems [1], [2]. In mesh networks or point-to-point fiber links combining multiple data rates on different channels via wavelength-division multiplexing (WDM), multi-dimensional power optimization increases margin by mediating the margin inequity between channels to mitigate the Kerr nonlinearity-induced noise [3]. By contrast, in the basic case of homogeneous point-to-point systems with the same SNR requirements on all channels, [3] found little benefit from power optimization.

Stimulated Raman Scattering (SRS) is another nonlinear effect present in optical fibers. High-power continuous-wave pumping produces high gain at the Stokes wavelength, enabling Raman optical amplifiers [4]. In a WDM communication system, shorter-wavelength channels act as low-power pump channels, providing gain to longer-wavelength channels, while being partially depleted.

Previous work has studied the combined effects of SRS and the Kerr nonlinearity on the dynamics of high-power pulses [5]. The system impact of the combined effects of SRS and cross-phase modulation-induced crosstalk has also been investigated [6]. Channel spacing optimization for dispersion-compensated IMDD systems has been studied in the presence of SRS and four-wave mixing [7]. Power or SNR equalization in the presence of arbitrary noise and gain spectra has been considered for many years [8].

Optical communication systems that expand capacity by utilizing both the C and L bands experience stronger nonlinear impairments from the increased total power. The availability of seamless amplification for C+L-band systems [9] brings such expanded-spectrum systems closer to commercialization. The increased SRS experienced by C+L-band systems raises the question of how to mitigate the resulting performance penalties.

This paper considers multi-dimensional channel power optimization using the Gaussian noise model for dispersion-uncompensated nonlinear transmission in the presence of SRS. This form of system optimization is enabled by the development of flexible transceivers [10], wavelength-selective switches (WSSs) and the Gaussian noise model [1] for uncompensated nonlinear transmission.

The Gaussian nonlinear noise model used in this paper is rigorously valid when propagating signals are Gaussian-distributed, in which case, our system performance estimates are highly accurate. Probabilistically shaped high-order modulation signals are well-approximated as Gaussian-distributed

Manuscript received September 12, 2017; accepted November 6, 2017. Date of publication November 7, 2017; date of current version November 20, 2017. This work was supported in part by Ciena Corporation, and in part by Stanford Rambus Corporation Graduate Fellowship. *(Corresponding author: Ian Roberts.)*

I. Roberts and J. M. Kahn 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).

J. Harley and D. W. Boertjes are with Ciena Corporation, Ottawa, ON K2H 8E9, Canada (e-mail: jharley@ciena.com; dboertje@ciena.com).

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.2017.2771719

[11], [12]. Unshaped or low-order-modulated signals can be modeled as Gaussian-distributed if they are subject to dispersion pre-compensation beyond that required to compensate the dispersion accumulated along the link. If subject to only partial pre-compensation, such signals appear Gaussian at the transmitter, but become non-Gaussian at some point along the link. The discussion in Section IV explains how the optimization methods of this paper can be applied in systems with significantly non-Gaussian signals.

In this paper we consider wide bandwidths up to 10 THz, so we depend upon the accuracy of the Gaussian noise model and Manakov equation at these wide bandwidths. Their accuracy has been verified up to 7.3 THz in uncompensated transmission [13].

The structure of this paper is as follows. Section II investigates properties of models of inter-channel SRS, formulates an expression for the channel SNR, and evaluates the impact of moderate distributed gain on the Gaussian noise nonlinearity model. Section III proposes a locally optimal power optimization method and provides loose bounds on the suboptimality of solutions obtained. SRS-aware power optimization is compared to SRS-equalized power allocation, SRS-unaware convex nonlinear power allocation, and optimized spectrally flat power allocation. For C+L-band systems, independent optimization of the amplifier gains in the C and L bands is also considered. Sections IV and V provide discussion of results and conclusions, respectively.

II. Stimulated Raman Scattering

A. Inter-Channel Raman Gain

The Raman effect is a wide-bandwidth gain process in which shorter wavelengths are depleted to provide gain at longer wavelengths. The width of the C band is small compared to the peak Raman shift, such that the Raman gain coefficients $g_R(\omega_m - \omega_n)$ remain small even at the greatest channel separation $\omega_m - \omega_n$. For highly dispersive fiber and systems with large numbers of channels, the variance of the SRS gain due to modulated signals is very small and is typically ignored [14].

The amount of gain or loss experienced by channel n due to the combination of fiber attenuation and inter-channel Raman gain is described by the following equation [15], [16]:

$$\frac{\partial P_n(z)}{\partial z} = -2\alpha P_n(z) + \sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{\text{eff}}} P_n(z) P_m(z) \quad (1)$$

$P_n(z)$ is the power in channel n as a function of z, the distance propagated along the fiber. α is the fiber attenuation coefficient. A_{eff} is the fiber effective area. Expression (1) assumes that the Raman gain coefficient $g_R(\omega)$ exhibits negligible variation over the bandwidth of each channel. An accurate solution to the system of (1) for different n is found through either a fine discretization over z, or an analytic solution available under the triangular Raman gain coefficient approximation [15]. In a system in which power is constrained by the Kerr nonlinearity, the channel power gain or loss due to SRS is typically -1 dB to 1 dB per span. This is small relative to the fiber losses, but can be significant in comparison to other sources of gain ripple in a well-designed line system.

In order to obtain an expression enabling convexity properties to be evaluated easily, we will first look at a coarse single-step-per-span discretization of (1). The error in the gain or loss predicted by the single-step-per-span solution is around 0.05 dB, as compared to a finely discretized model.

The Raman gain, following (1), on channel n with a single step per span is

$$r_{\text{one-step},n} = \exp\left(\sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{\text{eff}}} P_m(0) L_{\text{eff}}(L_s)\right). \quad (2)$$

For a single-step-per-span model, the step size is the span length L_s. For any step size L, the effective length L_{eff} depends on the step size as

$$L_{\text{eff}}(L) = \frac{1 - e^{-2\alpha L}}{2\alpha} \quad (3)$$

The Raman gain coefficient $g_R(\omega_m - \omega_n)$ may be positive, for ω_m larger than ω_n, leading to energy transfer from channel m to channel n. The coefficient may also be negative, for ω_m smaller than ω_n, leading to energy transfer from channel n to channel m. If a vector of linear powers $\boldsymbol{x}$ is substituted for the input powers $P_m(0)$ in (2) and a logarithm taken, the logarithmic gain expression is obtained:

$$\log(r_{\text{one-step},n}(\boldsymbol{x})) = \sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{\text{eff}}} L_{\text{eff}}(L_s) x_m \quad (4)$$

The right-hand side of (4) is linear in $\boldsymbol{x}$, from which we can conclude that $r_{\text{one-step},n}$ is both log-convex and log-concave in $\boldsymbol{x}$.

It was found in [3] that in order to obtain a convex expression for the channel margin in the presence of Kerr effects as described by the Gaussian noise model, logarithmic power variables $\boldsymbol{y} = \log(\boldsymbol{x})$ may be employed. Using these logarithmic power variables, (4) becomes

$$\log(r_{\text{one-step},n}(\boldsymbol{x})) = \sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{\text{eff}}} L_{\text{eff}}(L_s) e^{y_m} \quad (5)$$

Due to the combination of both positive and negative values for $g_R(\omega_m - \omega_n)$ in all channels n other than 1 and N, $r_{\text{one-step},n}$ is neither convex nor concave in $\boldsymbol{y}$ for $1 < n < N$. All components with $g_R(\omega_m - \omega_n)$ positive are convex and all components with $g_R(\omega_m - \omega_n)$ negative are concave; however, all channels n such that $1 < n < N$ contain components with both signs. The undepleted pump approximation would be required in order to obtain a convex expression where all negative values of $g_R(\omega_m - \omega_n)$ are removed from the summation. For inter-channel Raman gain, where the pump and Stokes wavelengths have similar powers, this is a poor approximation.

The single-step-per-span SRS model allows for easy calculation of gradients, but as it is not convex using logarithmic variables, it is not useful for optimization. In this paper, a discretized version of (1) with a step size $L_{\text{step}} = 100$ m is used because accuracy is desired and, as will be shown, a series of convex objectives can be formed assuming that the inter-channel Raman gain is locally power-independent.

In a multiple-step-per-span model, following (1), the Raman gain experienced in step k of the span is

$$r_n^{(k)} = \exp\left(\sum_{m=1}^{N} \frac{g_R(\omega_m - \omega_n)}{A_{\text{eff}}} P_m[k-1] L_{\text{eff}}(L_{\text{step}})\right), \quad (6)$$

where $P_m[k]$ is the power in channel m at distance $k \cdot L_{\text{step}}$ through the span. $P_m[0]$ is the power at the start of the span, and for $k > 0$ along that span:

$$P_m[k] = e^{-2\alpha L_{\text{step}}} r_m^{(k)} P_m[k-1]. \quad (7)$$

The Raman gain, on channel n in a given span, under this multiple-step-per-span model is

$$r_n = \prod_{k=1}^{\frac{L_s}{L_{\text{step}}}} r_n^{(k)}. \quad (8)$$

B. Cumulative Inter-Channel Raman Gain

In a multiple-span fiber communication system, power variation accumulates as signals propagate along multiple spans. Periodic gain compensation is generally required to maintain approximately constant power levels. With idealized lumped-element optical amplifiers, the fiber span loss and amplifier gain cancel fully. The residual SRS-induced gain ripple remains uncompensated.

We assume in this paper that channels fall on a traditional 50 GHz grid. To model flexible-grid system, the spectrum could be quantized into narrower slices, which could be grouped together to form variable-width channels.

Let $\boldsymbol{x}$ be a vector of initial input powers, and let $G_{c,n}^{(i)}$ be the residual compensation gain in channel n of span i. The input power to span s is

$$\boldsymbol{x}^{(s)} = R^{(s-1)}(\boldsymbol{x}) \circ \boldsymbol{x}, \quad (9)$$

where $\circ$ denotes element-wise vector multiplication, and $R_n^{(s)}$ is the cumulative Raman gain in channel n following the amplifier of span s:

$$R_n^{(s)}(\boldsymbol{x}) = \prod_{i=1}^{s} r_n(\boldsymbol{x}^{(i)}) G_{c,n}^{(i)}$$

$$R_n^{(0)}(\boldsymbol{x}) = 1. \quad (10)$$

The Raman gain experienced by channel n in span i, r_n, depends on the power vector input to span i, denoted by $\boldsymbol{x}^{(i)}$. If $G_{c,n}^{(i)}$ exactly compensates for the Raman gain in the fiber span, there is no cumulative power variation.

Fig. 1 shows the fiber attenuation-compensated output powers from a single span of propagation for a range of spectrally flat input powers. The channel numbering follows the ITU frequency grid ordering, with higher channel indices corresponding to shorter wavelengths, which are depleted due to SRS. The simulated fiber parameters are: one 100 km span of single mode fiber with a 72.8 μm^2 effective area providing a total of 21 dB attenuation, and an optical amplifier providing 21 dB gain. For higher input power levels there is a significant output power tilt, which accumulates between spans.

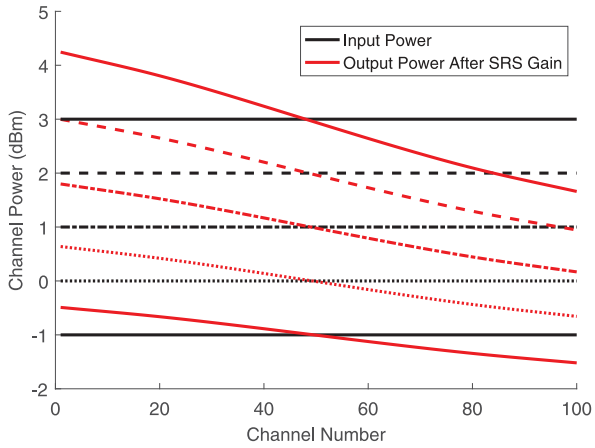

Fig. 1. Simulated output power profiles after a single span of optical fiber given different spectrally flat input power profiles. Each flat black line represents an input power level, and the red line with the same style represents the corresponding output power variation due to SRS gain after compensation of the fiber attenuation.

Fixed optical filters can help to equalize the channel powers, but the required spectral shape depends on the channel powers and spectral occupancy, so fixed filtering will be imperfect in any evolving or reconfigurable network. Even in the case of perfect fixed filtering, the distributed gain will lead to different noise levels on different channels. The operating point of the optical amplifiers can be shifted to provide an opposing gain tilt and somewhat compensate for the SRS gain spectrum. Both of these compensation methods are generally imperfect, leaving an error that accumulates over multiple spans. A WSS, whether a part of an add/drop multiplexer or isolated, can provide programmable wavelength filtering to compensate for the residual power tilt due to accumulated SRS gain.

Applying a gain adjustment at each amplifier to compensate for the tilted gain caused by SRS alters the noise introduced by the amplifiers. The amplifier noise power scales with $[G(\omega) - 1]$, which is approximately $G(\omega)$ at high gain. If $G(\omega)$ is not flat, then the noise spectrum will be similarly altered. If the amplifier gain is kept flat and an optical filter is used for gain balancing, the flat noise spectrum will also be filtered. In this paper we assume that a fraction $\psi \geq 0$ of the Raman gain (measured in dB units) that is experienced in each channel over one span is compensated at the subsequent amplifier, and after every five spans, the accumulated gain error is compensated perfectly. For spans s with only local compensation, the gain adjustment applied to channel n is a function of the single-span Raman gain (6):

$$G_{c,n}^{(s)} = [r_n(\boldsymbol{x}^{(s)})]^{-\psi}. \quad (11)$$

For spans s where the residual gain accumulation is compensated, the gain adjustment applied to channel n is the inverse of the cumulative Raman gain (10) for span s, which expands to

$$G_{c,n}^{(s)} = r_n(\boldsymbol{x}^{(s)})^{-1} \prod_{i=s_{\text{base}}}^{s-1} r_n(\boldsymbol{x}^{(i)})^{\psi-1}, \quad (12)$$

where $s \geq s_{\text{base}} \geq 1$ is the index of either the first span, or the first span following the previous cumulative power correction.

The compensation ratio ψ reflects the imperfect tuning that arises from fixed filters being used in a changing network, or the imprecision that results from applying a reverse tilt to an optical amplifier. Tunable filters are costly, making fixed filters or amplifier-based SRS gain compensation generally most economical. In a live system, it is much easier to apply a fixed tilt, which doesn't adjust for changing spectral loading, than dynamic SRS compensation. Transient events, such as fiber cuts, may abruptly change the SRS gain profile, making stable compensation a challenge. The inherent imprecision of a fixed filter is due to the necessity of providing a reasonable system spectral response for the full system lifetime ranging from initial setup with a few channels through towards full-fill.

C. Gaussian Noise Model in the Presence of Inter-Channel Raman Gain

The Gaussian noise Kerr nonlinearity model is expressed in (13)–(15) [1]. It expresses the nonlinear noise power spectrum $G_{NL}(f)$ in terms of the signal power spectrum $G(f)$ at three different frequencies and coefficients calculated from fiber parameters: attenuation coefficient α, dispersion propagation constant β_2, nonlinear coefficient γ, span length L_s, and effective span length L_{eff}.

$$G_{NL}(f) = \frac{16}{27}\gamma^2 L_{\text{eff}}^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(f_1)G(f_2) \times G(f_1+f_2-f)\cdot\rho(f_1,f_2,f) \times \chi(f_1,f_2,f)\partial f_1 \partial f_2 \tag{13}$$

$$\rho(f_1,f_2,f) = \left|\frac{1-e^{-2\alpha L_s + j4\pi^2\beta_2 L_s (f_1-f)(f_2-f)}}{2\alpha - j4\pi^2\beta_2(f_1-f)(f_2-f)}\right|^2 L_{\text{eff}}^{-2} \tag{14}$$

$$\chi(f_1,f_2,f) = \frac{\sin^2(2N_s\pi^2(f_1-f)(f_2-f)\beta_2 L_s)}{\sin^2(2\pi^2(f_1-f)(f_2-f)\beta_2 L_s)} \tag{15}$$

The expression for ρ in (14) assumes a lumped-amplification model. With the introduction of SRS, the fiber attenuation profile varies with wavelength, input signal power, and distance. The general expression for ρ is found in Appendix to be

$$\rho'(f_1,f_2,f) = \left|\int_0^z e^{\int_0^{z'}[\alpha(\zeta,f_1+f_2-f)-\alpha(\zeta,f_1)-\alpha(\zeta,f_2)-\alpha(\zeta,f)]\partial\zeta} \times e^{j[\beta(f_1+f_2-f)-\beta(f_1)-\beta(f_2)+\beta(f)]z'}\partial z'\right|^2 L_{\text{eff}}^{-2}. \tag{16}$$

The associated effective nonlinear length L_{eff} scaling becomes

$$L_{\text{eff}} = \int_0^z e^{\int_0^{z'} -2\alpha(\zeta,f)\partial\zeta}\partial z'. \tag{17}$$

Separating the fiber attenuation and SRS gain components (1) gives the spacial and frequency varying attenuation coefficient as

$$\alpha(z,f_n) = \alpha_{\text{fiber}} - \frac{1}{2}\sum_{m=1}^{N}\frac{g_R(2\pi(f_m-f_n))}{A_{\text{eff}}}P_m(z). \tag{18}$$

The ρ' given by (16) is thus a function of the input channel powers due to the inter-channel Raman gain. This is a problem for optimization, as (16) is not analytically integrable, and would yield a matrix of coefficients that are functions of the optimization variables.

Given the small magnitude of the inter-channel Raman gain, and the normalization of (16) by L_{eff}^{-2}, the coefficient array ρ' can be approximated by a constant matrix with a small error. Let us assume, for example, that the distributed Raman gain provides 1 dB of gain or loss per span, which is what is provided for the worst-case channels by a 2 dBm launch power on each of 100 channels in the scenario of Fig. 1. The coefficient error due to an assumption of constant coefficients and ignoring the SRS gain is less than 2.5% (0.11 dB) for the dominant self-phase modulation component, and less than 4% (0.17 dB) for the remaining coefficients. Given that the nonlinearity is cubic in channel power, the change in optimal power should follow the cube-root of this error (0.03 to 0.05 dB). For C+L-band systems, the increased SRS strength is mitigated by minimum-margin optimizing power allocations seeking to minimize channel inequalities, such that model errors are similar. Validation of this model is performed in Section III-C by comparison to split-step Fourier simulations.

Under the assumption of constant ρ' coefficients, the L_{eff}^2 (17) term of (13) retains the residual SRS-dependent variation in the Kerr nonlinearity model.

If the Gaussian noise model is discretized in the form of [3], the following function for the nonlinear Gaussian noise power in channel n is obtained:

$$NL_{R,n}(\boldsymbol{x}) = \gamma^2 \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=-1}^{1} x_i x_j x_{i+j-n+l} \times L_{\text{eff}}^2 D_l(i,j,n) \tag{19}$$

With the introduction of SRS, ρ takes the form of (14), which includes normalization by L_{eff}^{-2}. For a single span, the χ coefficient (15) is 1 and drops out. If a sequence of spans with varying input powers is considered, the powers of the nonlinear noise contributions from the spans vary with the input power of the spans, but the partially coherent accumulation of nonlinear noise products should still be considered. The ϵ-law approximation of [1] is a good approach for representing that accumulation. The discrete nonlinear coefficients D_l are calculated via [3]

$$D_l(n_1,n_2,n) = \frac{16}{27}\int_{-\frac{\Delta f}{2}}^{\frac{\Delta f}{2}}\int_{-\frac{\Delta f}{2}}^{\frac{\Delta f}{2}}\int_{-\frac{\Delta f}{2}}^{\frac{\Delta f}{2}} \rho(\xi_1+n_1\Delta f,\xi_2+n_2\Delta f,\xi+n\Delta f) \cdot\chi(\xi_1+n_1\Delta f,\xi_2+n_2\Delta f,\xi+n\Delta f)g(\xi_1) \cdot g(\xi_2)g(\xi_1+\xi_2-\xi+l\Delta f)R\,g(\xi)\partial\xi_1\partial\xi_2\partial\xi \tag{20}$$

D. Signal-to-Noise Ratio With Inter-Channel Raman Gain

Including inter-channel Raman gain using (10), the SNR of channel n is

$$SNR_n(\boldsymbol{x}) = \left(\sum_{s=1}^{N_s} \frac{\sigma_{s,n}^2 + r_n(\boldsymbol{x}^{(s)}) N_s^{\epsilon} NL_{R,n}(\boldsymbol{x}^{(s)})}{{G_{c,n}^{(s)}}^{-1} R_n^{(s)}(\boldsymbol{x}) x_n} \right)^{-1}, \tag{21}$$

where $\boldsymbol{x}^{(s)}$ is defined following (9).

In (21), N_s is the number of spans, $\sigma_{s,n}^2$ is the amplifier noise contribution due to compensation of fiber attenuation for channel n in span s, and $NL_{R,n}(\boldsymbol{x})$ is the Gaussian nonlinear noise produced following a model that includes the distributed SRS gain and loss across one span as defined in (19). The factor N_s^{ϵ} represents the partially coherent accumulation of the nonlinear noise components following the epsilon-law approximation [1]. The single-span inverse SNRs within the summation are evaluated with noises referred to the theoretical point where the post-fiber optical amplifier has precisely compensated for only fiber attenuation. The cumulative Raman gain following the amplifier of span s, $R_n^{(s)}(\boldsymbol{x})$, includes excess compensation gain $G_{c,n}^{(s)}$ beyond compensating for fiber attenuation.

The nonlinear Gaussian noise function $NL_{R,n}$ assumes the signal loss due to the combined fiber attenuation and Raman gain is precisely compensated by the optical amplifier at the end of the span. At the point along the fiber where only the fiber attenuation has been compensated, the nonlinear noise strength is increased by the SRS gain experienced during propagation.

Equation (21) exhibits complex interactions between the Kerr nonlinearity and the Raman-adjusted power vector. In order to work towards a convex formulation for the channel SNR, let us assume that the Raman power gain is perfectly compensated by a WSS after each span. If we also assume that the impact of Raman gain on the strength of the Kerr nonlinearity is negligible, the SNR expression becomes

$$SNR_n(\boldsymbol{x}) = \left(\sum_{s=1}^{N_s} \frac{\sigma_{s,n}^2 + r_n(\boldsymbol{x}) N_s^{\epsilon} NL_n(\boldsymbol{x})}{r_n(\boldsymbol{x}) x_n} \right)^{-1}, \tag{22}$$

where $NL_n(\boldsymbol{x})$ is the nonlinear Gaussian noise assuming zero Raman gain. Simplifying under the assumption of identical spans, we obtain

$$SNR_n(\boldsymbol{x}) = \frac{x_n}{\sigma_n^2 {r_n(\boldsymbol{x})}^{-1} + N_s^{1+\epsilon} NL_n(\boldsymbol{x})}. \tag{23}$$

In [3] it is found that the SNR without Raman gain was log-concave in logarithmic power vector $\boldsymbol{y} = \log(\boldsymbol{x})$. In terms of linear power vector $\boldsymbol{x}$, the SNR was found to be quasi-concave near the subspace where the powers in all channels are equal, but to have no general convexity properties. With the introduction of the Raman gain in (23), we find the SNR to depend on the sum of a log-convex function of $\boldsymbol{x}$, $r_n(\boldsymbol{x})$, and a locally convex one, $NL_n(\boldsymbol{x})$. For equal powers in all channels (23) is log-concave, but in the presence of Raman gain, any useful minimum-margin optimal solution will have unequal channel powers. The power inequality generally required to achieve equal margin in all channels is found through simulation to be sufficient to often push (23) out of the small log-concave region and into the remainder of the problem domain $\boldsymbol{x} \in \mathbf{R}_+^N$, which has many local optima.

If (23) is analyzed in terms of logarithmic power vector $\boldsymbol{y} = \log(\boldsymbol{x})$, the Gaussian noise nonlinearity portion is made convex, but the Raman gain factor becomes non-convex. We thus conclude that even in the most simple form, the combination of Raman with the Gaussian noise nonlinearity model leads to a non-convex SNR expression.

Relaxing the simplifying assumptions made above (22) and examining (21), we once again find that the SNR is non-convex. This is because the Raman gain takes the form of an exponential gain or loss that is log-convex in linear powers. In the frequency domain, the mechanism of the Kerr nonlinearity underlying the Gaussian noise model takes the form of multiplying frequency triples, which leads to posynomial expressions requiring logarithmic power variables for conversion into convex form. We conclude that the channel SNR in the presence of inter-channel Raman gain is incompatible with direct convex optimization.

III. Locally Optimal Margin Optimization

A. Optimization Method

In Section II-D, it was demonstrated that the general SNR expression with SRS gain (21) is non-convex. This means that all practical descent methods only find locally optimum solutions, and the level of global optimality is indeterminate. However, some descent methods may still obtain good results, depending upon the problem. Bounding the suboptimality allows determination of whether results are good enough for a particular application.

A minimum-margin optimization objective that is concave and differentiable was formulated in [3]. If that analysis is applied directly to the SNR expression including SRS, a significant problem arises. Attempting to use a quadratic approximation via Newton's method to obtain rapid convergence is no longer reliable, as the Hessian matrix may have both positive and negative eigenvalues. Following the calculated Newton step direction may increase or decrease the objective function in such a scenario. Saddle points become attractors for the quadratic steps of the Newton method [17]. Simple gradient descent does not have this problem and will eventually find a local minimum or inflection point, but takes many more iterations for convergence [18].

A better approach arises from assuming that the SRS gain is independent of changes to the channel power vector. If the SRS gain is assumed to be constant at some initial power level $\boldsymbol{x}_0$, the margin evaluated using (21) becomes

$$M_n(\boldsymbol{x}, \boldsymbol{x}_0) = SNR_{req,n}^{-1} \left(\sum_{s=1}^{N_s} \frac{A(\boldsymbol{x}, \boldsymbol{x}_0)}{{G_{c,n}^{(s)}}^{-1} R_n^{(s)}(\boldsymbol{x}_0) x_n} \right)^{-1}, \tag{24}$$

where $SNR_{req,n}$ is the required SNR in channel n, and

$$A(\boldsymbol{x}, \boldsymbol{x}_0) = \sigma_{s,n}^2 + r_n(\boldsymbol{x}_0^{(s)}) N_s^{\epsilon} NL_{R,n}(R^{(s-1)}(\boldsymbol{x}_0) \circ \boldsymbol{x}, \boldsymbol{x}_0). \tag{25}$$

In (25), $\boldsymbol{x}_0^{(s)}$ is the fixed input power to span s used only for calculating the Raman gain of that span given the fixed input power profile $\boldsymbol{x}_0$

$$\boldsymbol{x}_0^{(s)} = R^{(s-1)}(\boldsymbol{x}_0) \circ \boldsymbol{x}_0. \tag{26}$$

In order to fully extend the assumption of constant SRS gain, $NL_{R,n}$ defined in (19) needs to become

$$NL_{R,n}(\boldsymbol{x}, \boldsymbol{x}_0) = \gamma^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{l=-1}^{1} x_i x_j x_{i+j-n+l} \times L_{\text{eff},\boldsymbol{x}_0}^2 D_l(i,j,n) \tag{27}$$

where $L_{\text{eff},\boldsymbol{x}_0}$ follows from (17) with assumed fixed power profile $\boldsymbol{x}_0$.

Under the assumption of constant SRS gain, the Gaussian noise nonlinearity has been completely decoupled from the Raman gain function. Using a transformation to logarithmic gain variables $\boldsymbol{y} = \log(\boldsymbol{x})$, the margin (24) is log-concave in $\boldsymbol{y}$ for a fixed $\boldsymbol{x}_0$.

In order to maximize the minimum margin including the power-dependent effect of SRS, we thus wish to find some $\boldsymbol{x}_0$ such that maximizing the minimum-channel-margin given static SRS (24) finds an optimal power allocation $\boldsymbol{x}^* \approx \boldsymbol{x}_0$.

If we assume $\boldsymbol{x}_0$ takes some reasonable initial value $\boldsymbol{u}^{(1)}$, we can solve the convex minimum margin problem given $\boldsymbol{u}^{(1)}$ to obtain an optimal power allocation $\boldsymbol{x}^{*(1)}$. $\boldsymbol{x}^{*(1)}$ will generally be very different from $\boldsymbol{u}^{(1)}$. If we let $\boldsymbol{u}^{(i)} = \boldsymbol{x}^{*(i-1)}$ and solve a series of convex problems, we obtain a sequence of new optimal solutions $\boldsymbol{x}^{*(i)}$. Even though the channel SNR is bounded and hence the margin is limited, there is no guarantee that such a sequence of convex approximations to a non-convex function will converge.

In practice, performing such a sequence of convex approximations leads to $\boldsymbol{x}^{*(i)}$ converging to $\boldsymbol{u}^{(i)}$ for the specific problem of maximizing the minimum margin (24). The properties of the Raman gain are helpful here. The Raman gain bandwidth is much wider than that of the Gaussian noise nonlinearity, making the Raman gain insensitive to the powers of individual channels. The form of the Raman gain is an exponential of a linear function of the channel powers. Given the small magnitude of the loss or gain, it remains approximately linear in channel powers. The Kerr nonlinearity is cubic in channel powers, in addition to having a comparatively narrow bandwidth, making it much more sensitive to small changes in the vector of channel powers.

If this sequence of convex approximations converges, it is unlikely to converge to the singular optimal solution given the exponential number of sub-optimal solutions that arise in N dimensions. There may be multiple locally optimal fixed points $\boldsymbol{x}^*$ that satisfy

$$\left[\underset{\boldsymbol{x}}{\operatorname{argmax}} \min_{n\in[1,N]} M_n(\boldsymbol{x}, \boldsymbol{x}^*)\right] = \boldsymbol{x}^*. \tag{28}$$

Given any candidate solution $\boldsymbol{x}^*$, the convex approximation with a fixed SRS gain can be used to bound the sub-optimality of the solution within some region of the domain. If $\boldsymbol{x}^*$ is believed to be within $\pm\phi$ dB of the global optimum for all components, solving the two convex problems to maximize $\min M_n(\boldsymbol{x}, \boldsymbol{x}^* \cdot 10^{\pm\phi/10})$ provides bounds on the margin achievable at any fixed points near $\boldsymbol{x}^*$. This once again takes advantage of the insensitivity of the inter-channel Raman gain to individual channel power changes. While normally placing a bounding box around a vector variable in N dimensions would require 2^N points, the Raman gain is insensitive enough to individual channel powers that the two points where all channels are either increased or decreased in power dominate the error bound over all partial binary partitions. Additional bounding certainty is achieved by choosing ϕ larger than the expected range for the global optimum.

For the optimization results of this paper, $\phi = \pm 0.5$ dB is used to bound the optimized margins and sub-optimality upper bounds of a few hundredths of a dB were found. Increasing the range to $\phi = \pm 1$ dB maintains loose upper bounds on the sub-optimality below 0.1 dB, which is sufficient accuracy for a practical system implementation.

If a sequence of increasingly accurate barrier approximations is used to form an objective function out of a set of inequalities, as proposed in [3], then the solution method already solves a series of convex functions. After the solution for a given barrier accuracy has been obtained, the solution can be used to update the assumed fixed SRS power profile $\boldsymbol{x}_0$ before solving again with a more accurate barrier function. It should be noted that when $\boldsymbol{x}_0$ changes, the minimum margin may decrease and any slack variables should be loosened if necessary to avoid starting in an infeasible location. Updating the two different series of convex functions simultaneously saves a significant factor of computation time.

For C+L-band systems, one system design approach is to match the amplification gains provided to the two bands [9]. Without resorting to programmable filters for ideal SRS gain compensation, another system design approach independently assigns the channel gains for the two bands. If we represent the relative logarithmic gain for the C and L bands in span s as $\kappa_{C,s}$ and $\kappa_{L,s}$, respectively, then the power of channel n in those bands is $y_n + \kappa_{C,s}$ or $y_n + \kappa_{L,s}$. These new logarithmic variables maintain the convexity of the minimum margin optimization objective for a fixed SRS gain. Optimizing over the combined set of y and κ variables results in a solution that provides both a power profile and a cumulative gain for each band in each span. Optimized gain tuning for each band in each span allows for improved performance in C+L-band systems without the need for filter-based per-span SRS gain compensation.

B. C-Band Results and Analysis

We first look at power optimization for fully occupied C-band systems. We assume the following simulation parameters: 40 spans of single-mode fiber, each 100 km long, with loss $\alpha = 0.21$ dB/km plus fixed losses of 2 dB for a total of 23 dB span attenuation, dispersion $D = 17$ ps $\cdot$ nm^{-1} km^{-1}, nonlinear coefficient $\gamma = 1.4$ W^{-1} m^{-1}, a 72.8 μm^2 effective area, and amplifiers with 6 dB noise figures.

Fig. 2 shows two different optimized power allocations for a system with $\psi = 0\%$ per-span SRS gain compensation and

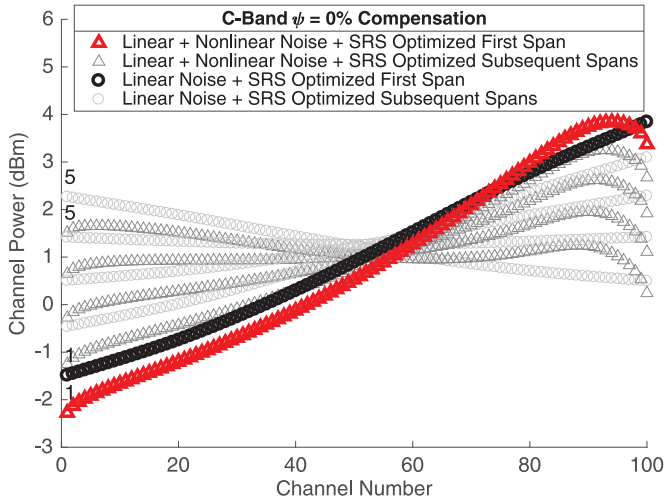

Fig. 2. The two highlighted power allocations are one optimized for the presence of SRS and nonlinear Gaussian noise, and one optimized for the presence of SRS with linear SNR and power bounded by the peak power of the nonlinear-aware power allocation. 100 C band channels are utilized. Simulations consider 40 spans of single-mode fiber, each 100 km long, with loss $\alpha = 0.21$ dB/km plus fixed losses of 2 dB for a total of 23 dB span attenuation, dispersion $D = 17$ ps $\cdot$ nm^{-1} km^{-1}, nonlinear coefficient $\gamma = 1.4$ W^{-1} m^{-1}, a 72.8 μm^2 effective area. Amplifiers have a 6 dB noise figure and are adjusted to compensate for fiber attenuation and $\psi = 0\%$ of the SRS-induced gain tilt. Every five spans, an amplifier plus WSS combination is assumed to perfectly compensate the signal power profile to that of the initial input. The pattern of power profiles is cyclical due to the assumed perfect compensation every five spans. The input power profile and the power profiles that are input to the subsequent four spans are shown.

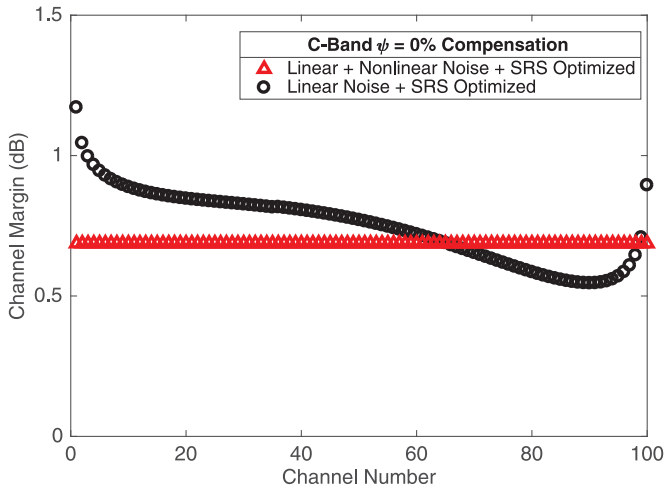

Fig. 3. Channel margins evaluated with partial SRS gain compensation for optimized SRS-aware nonlinear and linear SNRs as in Fig. 2. There is a 0.14 dB increase in minimum margin for the for the Kerr nonlinearity-aware power allocation. Simulation parameters are as in Fig. 2.

periodic power correction via WSS every five spans. The two optimization methods are the linear and nonlinear noise-plus-SRS optimization of this paper, and a linear noise-plus-SRS optimization similar to that of [8], with power capped by the peak power of the allocation produced by the nonlinear optimization. Fig. 3 shows the resulting channel margins for these two power allocations. A 0.14 dB gain in minimum margin is obtained from the Kerr nonlinearity-aware power optimization.

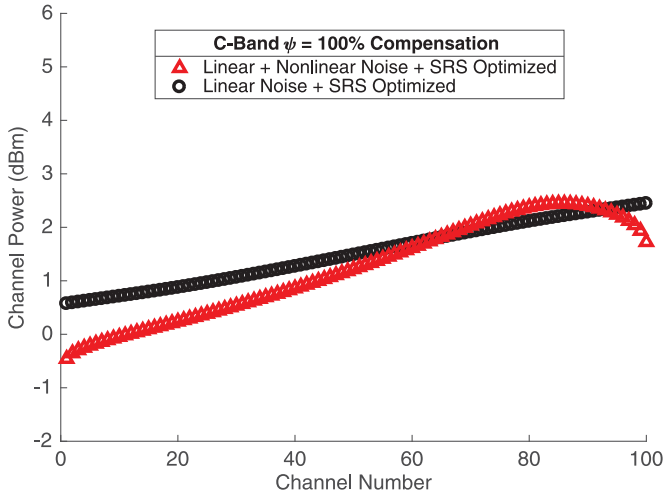

Fig. 4. The two highlighted power allocations are one optimized for the presence of SRS and nonlinear Gaussian noise, and one optimized for the presence of SRS with linear SNR and power bounded by the peak power of the nonlinear-aware power allocation. Simulation parameters are as in Fig. 2, except that $\psi = 100\%$ of the SRS-induced gain tilt is compensated at each amplifier.

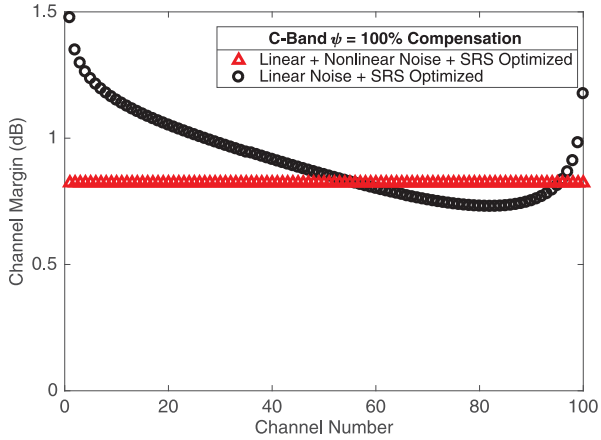

Fig. 5. Channel margins evaluated with partial SRS gain compensation for power allocations optimized with and without including SRS in the model and for an SRS optimized flat power allocation, as shown in Fig. 4. There is a 0.09 dB increase in minimum margin for the nonlinear power allocation. Simulation parameters are as in Fig. 2, except that $\psi = 100\%$ of the SRS-induced gain tilt is compensated at each amplifier.

Fig. 4 shows the same two power allocation methods applied to the same C-band system with $\psi = 100\%$ per-span SRS gain compensation. With no cumulative Raman gain, the same power profile is input to each span. The combination of the variable noise spectrum due to the gain compensation at each amplifier, and the power variations within each span due to SRS, once again lead to margin gains from optimization as shown in Fig. 5.

Fig. 6 shows minimum channel margins for varying values of the per-span SRS gain compensation parameter ψ. The power allocations compared are: SRS-aware power allocations considering and ignoring nonlinear noise as in Fig. 2, a SRS-unaware convex nonlinear noise-based power allocation, and a linear and nonlinear noise-plus-SRS-optimized spectrally flat power allocation. The SRS-unaware convex nonlinear optimization assumes that the amplifier compensation profiles of the SRS-aware

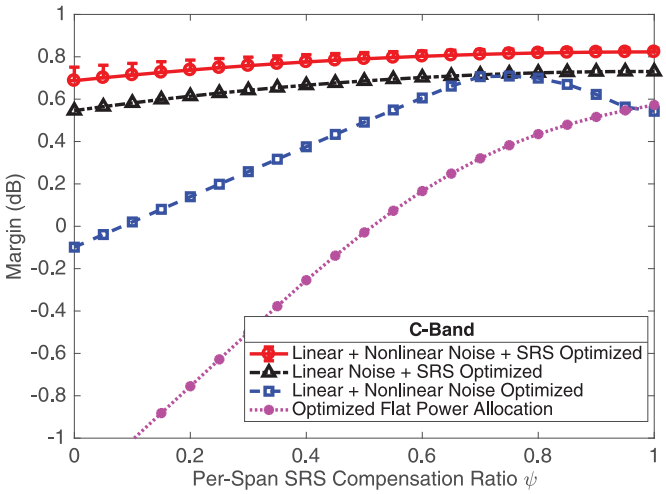

Fig. 6. Margins of different power allocation methods for varying levels of per-span SRS compensation. The optimized power allocation includes error bars indicating the loose upper bound on achievable margin. Power allocations included are: SRS-aware linear and nonlinear SNR-based power allocations, a SRS-unaware nonlinear power allocation, and an optimized spectrally flat power allocation. Simulation parameters are as in Fig. 2, except for varying levels of SRS compensation ratio ψ.

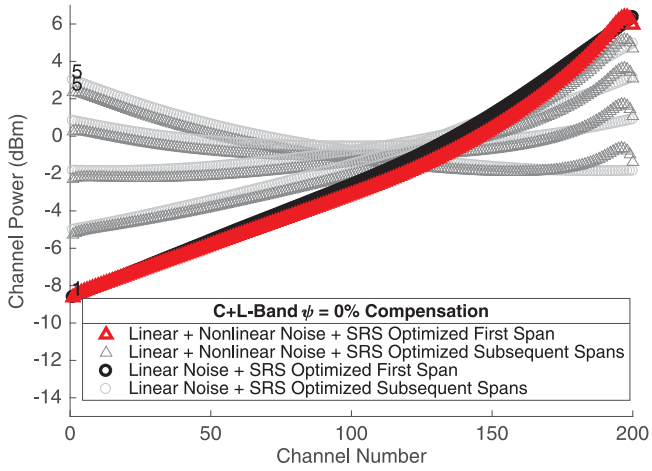

Fig. 7. The two highlighted power allocations are one optimized for the presence of SRS and nonlinear Gaussian noise, and one optimized for the presence of SRS with linear SNR and power bounded by the peak power of the nonlinear-aware power allocation. The pattern of power profiles is cyclical due to the assumed perfect compensation every five spans. The input power profile and the power profiles that are input to the subsequent four spans are shown. Simulation parameters are as in Fig. 2, except for including C+L-band channels.

linear SNR-based power allocation are held fixed, and performs a convex power optimization given the resulting channel dependent noise spectrum. This convex optimization is performed under the assumption that the SRS gain is zero and thus that the same channel power is input to all spans, and then the SNR is evaluated including SRS.

Fig. 6 shows that the spectrally flat power allocation has a margin penalty of 2 dB with $\psi = 0\%$ per-span SRS gain compensation, as compared to the nonlinear optimized power allocation. The linear noise-plus-SRS-optimized power allocation has a 0.09 to 0.14 dB penalty as compared to that including nonlinear noise. The zero-SRS nonlinear power allocation has variable performance that is highly dependent on the level of per-span SRS gain compensation. The linear and nonlinear noise-plus-SRS optimization exhibits a 0.14 dB increase in minimum margin when per-span SRS gain compensation is increased from $\psi = 0\%$ to $\psi = 100\%$.

C. C+L-Band Results and Analysis

We next look at power optimization for fully occupied 200 channel C+L-band systems. We assume that we can achieve seamless power-balanced amplification across both spectrum bands. We assume the following simulation parameters: 40 spans of single-mode fiber, each 100 km long, with loss $\alpha = 0.21$ dB/km plus fixed losses of 2 dB for a total of 23 dB span attenuation, dispersion $D = 17$ ps $\cdot$ nm^{-1} km^{-1}, nonlinear coefficient $\gamma = 1.4$ W^{-1} m^{-1}, a 72.8 μm^2 effective area, and amplifiers with 6 dB noise figures.

There are currently no known verifications of the GN model for the very wide bandwidths of C+L-band systems. We have performed a split-step Fourier method simulation of the first five spans of the power optimization for the most extreme scenario of $\psi = 0\%$ per-span SRS gain compensation in order to verify both the extension of the GN model to C+L-band systems and verify the simplifying assumptions used in our optimization. Due to the assumed WSS every five spans, the power profile repeats eight times in the full 40-span system. Gaussian signals are simulated with the power allocations of Fig. 7. Other system parameters remain the same. The most extreme scenario is selected for comparison with a split-step method due to the computational complexity of such simulations across the 10 THz of bandwidth of a 200 channel C+L-band system.

Fig. 7 shows two different optimized power allocations for a system with $\psi = 0\%$ per-span SRS gain compensation and periodic power correction via WSS every five spans. The two optimization methods are the same as those mentioned above for Fig. 2. The two power allocations appear similar for the input to the first span, but separate with the accumulation of uncompensated SRS gain across subsequent spans. Fig. 8 shows the resulting channel margins for these two power allocations. A 0.60 dB gain in minimum margin is obtained from the Kerr nonlinearity-aware power optimization under the GN+SRS model. For the two power allocations, a split-step Fourier simulation is performed across five spans between the assumed WSSs with the margin shown after SNRs reduced by a factor of 8 (~ 9 dB) to compare against the 40-span model results. There is close agreement between the simplified GN-based model and the split-step simulation across the entire band with maximum errors of about 0.15 dB, in this extreme scenario with 0% per-span SRS gain compensation. This serves to validate both the extension of the GN model to C+L-band systems, and the simplified model used for margin optimization in the presence of SRS. The minimum margin improvement is 0.61 dB between the two split-step simulation measurements.

Fig. 9 shows the same two power allocation methods applied to the same C+L-band system with $\psi = 100\%$ per-span SRS gain compensation. With no cumulative Raman gain, the same

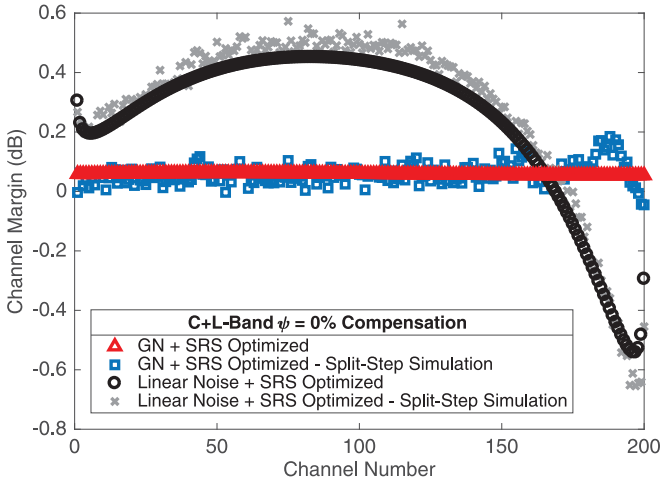

Fig. 8. Channel margins evaluated with partial SRS gain compensation for optimized SRS-aware nonlinear and linear SNRs as in Fig. 2. There is a 0.60 dB increase in minimum margin for the nonlinear power allocation. Simulation parameters are as in Fig. 2, except for including C+L-band channels. The margins obtained from a split-step Fourier simulation with the optimized power allocation are also shown for this most-extreme scenario to verify the simplified model used for optimization. There is agreement across the band with maximum error of about 0.15 dB. The minimum margin improvement is 0.61 dB between the two split-step simulation measurements.

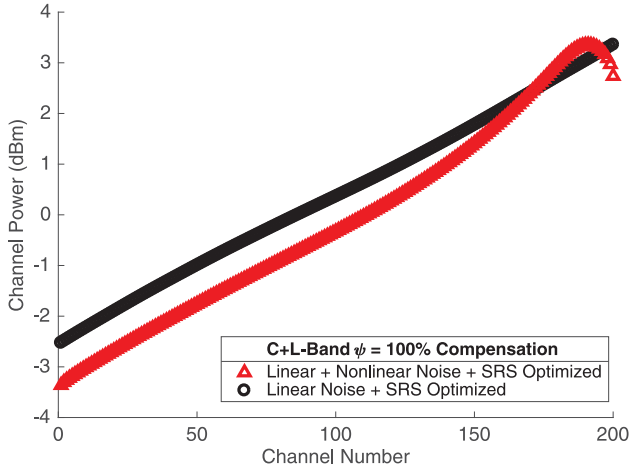

Fig. 9. The two highlighted power allocations are one optimized for the presence of SRS and nonlinear Gaussian noise, and one optimized for the presence of SRS with linear SNR and power bounded by the peak power of the nonlinear-aware power allocation. Simulation parameters are as in Fig. 2, except for including C+L-band channels, and having $\psi = 100\%$ of the SRS-induced gain tilt is compensated at each amplifier.

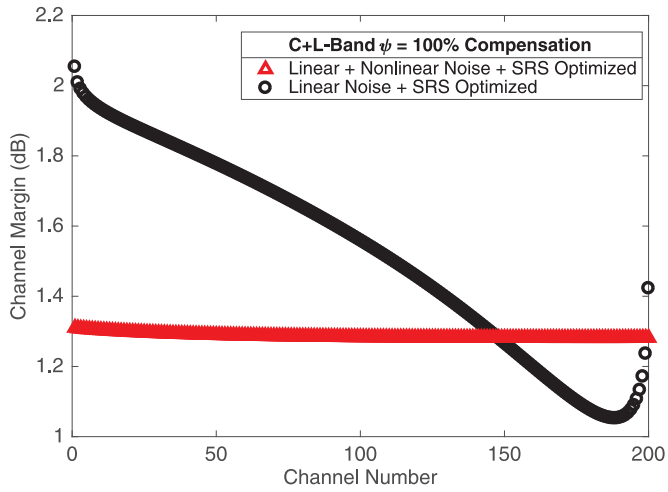

Fig. 10. Channel margins evaluated with partial SRS gain compensation for power allocations optimized with and without including SRS in the model and for an SRS optimized flat power allocation as shown in Fig. 4. There is a 0.23 dB increase in minimum margin for the nonlinear power allocation. Simulation parameters are as in Fig. 2, except for including C+L-band channels, and having $\psi = 100\%$ of the SRS-induced gain tilt is compensated at each amplifier.

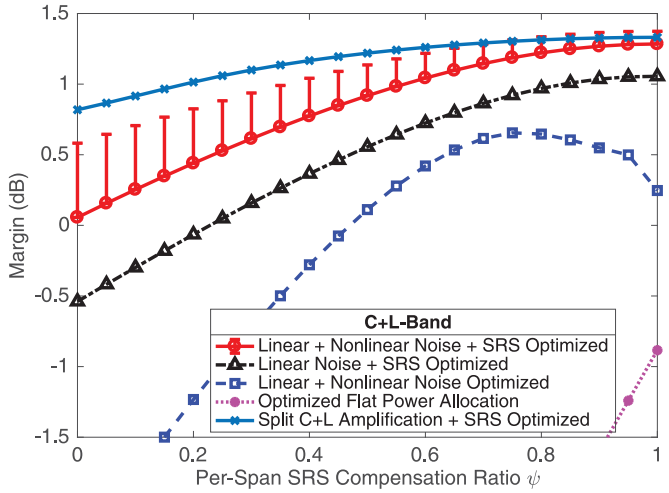

Fig. 11. Margins of different power allocation methods for varying levels of per-span SRS compensation. The optimized power allocation includes error bars indicating the loose upper bound on achievable margin. Power allocations included are: SRS-aware linear and nonlinear SNR-based power allocations, a SRS-unaware nonlinear power allocation, an optimized spectrally flat power allocation, and an SRS-aware nonlinear SNR-based optimization with additional degrees of freedom from independent per-band gain control. Simulation parameters are as in Fig. 2, except for varying levels of SRS compensation, and the inclusion of C+L-band channels.

power profile is input to each span. The increased nonlinearity strength, due to the increase in channels over the C-band simulations, leads to more significant minimum margin gains of 0.23 dB, as shown in Fig. 10.

Fig. 11 shows minimum channel margins for varying levels of per-span SRS gain compensation parameter ψ. The five power allocations compared are those four mentioned above for Fig. 6, and an SRS-aware nonlinear optimization across channel powers and additional dimensions of per-span and per-band gain. The error bars indicating the margin estimated by the loose upper margin bound, which assumes a fixed SRS profile, grow significant for small values of ψ. With low per-span compensation levels in a C+L-band system the cumulative SRS gain has a significant impact on the channel SNRs. Fixing the SRS gain artificially low based on a fixed power profile 0.5 dB below the optimized allocation allows a theoretical 0.5 dB margin gain to be obtained for the bound, but the actual minimum margin delivered by that bounding power allocation is 1 dB below the bound and thus 0.5 dB below that delivered by the optimized allocation.

For the C+L-band system of Fig. 11, a spectrally flat power allocation should be avoided even with ideal per-span SRS gain compensation. With $\psi = 0\%$ compensation, the minimum

margin penalty of a spectrally flat allocation as compared to an optimized power allocation is nearly 6 dB. The linear noise-plus-SRS-optimized power allocation performs significantly worse than that including nonlinear noise, with 0.23 to 0.60 dB of margin penalty. The minimum margin improvement from including per-span SRS gain compensation is significant for the simulated C+L-band system. The minimum margin delivered by the optimized power allocation with $\psi = 100\%$ per-span SRS gain compensation is 1.23 dB better than that delivered by the similar optimized power allocation for $\psi = 0\%$ per-span compensation. The best system performance is provided by the optimization that assumes independent control over the C-and L-band amplifiers, and adds additional optimization dimensions. The ability to independently control the C- and L-band gains provides most of the benefit of per-span SRS gain compensation, with only 0.51 dB of remaining minimum margin improvement from $\psi = 100\%$ per-span SRS gain compensation as compared to $\psi = 0\%$. The comparison optimization methods are not evaluated with split C+L-band amplifier control due to the necessity of using complex optimization methods to accommodate the large number of additional problem dimensions. Given that the combined Gaussian noise and SRS optimization method has been demonstrated to be superior, if a high-dimensional optimization method is necessary, it makes sense to also use the best high-dimensional power allocation method.

IV. Discussion

Channel margin variation arises when there is a mismatch between the delivered SNR distribution and the required SNR distribution. In [3], it is demonstrated that optimization of the channel powers is able to deliver significant minimum margin improvements in such mismatched situations. For a point-to-point link with a homogeneous set of SNR requirements, we have demonstrated here that the inter-channel Raman gain can lead to sufficient variation in the delivered SNR that valuable gains can be obtained from optimization.

In a live system, the variable noise spectrum due to the compensation of the inter-channel SRS gain will be directly measurable. An optimization method that is unaware of the Raman effect will still be able to take advantage of the corresponding non-flat ASE spectrum. Even if per-span SRS compensation is neglected ($\psi = 0$), the periodic power balancing that is generally required will lead to a non-flat ASE noise spectrum. Partial or full compensation of the inter-channel Raman gain will lead to a larger variation in the ASE across the channel range.

The scenarios investigated in this paper have 23 dB of span attenuation, which is at the high end of typical. A close amplifier spacing or low-attenuation fiber can significantly reduce span losses. In such a scenario, amplifiers contribute less ASE noise, leading to a lower optimal signal power, as the optimal single-channel SNR is achieved when the nonlinear noise power is half that of the ASE power. Lower signal power leads to lower inter-channel Raman gain, making the channel powers more homogeneous. Conversely, higher optimal powers from spans with higher loss will lead to greater inhomogeneity than presented here.

While 40 fiber spans were considered in this analysis, the assumption of perfect power compensation with the aid of a WSS or other programmable filter every five spans makes the interval of interest five spans. The subsequent repetition of the pattern eight times causes the SNR of all channels to decrease by a linear factor of eight, or about 9 dB. In a less idealized system without identical spans, or with a varying separation between ideal correcting filters, the SNR variation due to SRS will represent a weighted average over the different sections. The optimal input power profiles to differing sections will vary, requiring additional optimization variables.

The results presented in this paper come from simulations of systems occupying either a fully occupied C band or a fully occupied C and L bands. Partial wavelength occupancy reduces the total power in the fiber, which decreases the SRS strength. Partial wavelength occupancy may further lead to utilization of a portion of the channel band. As the SRS gain grows stronger with increasing separation, such wavelength clustering will reduce SRS strength, and can also change the location of the zero-gain wavelength.

The inclusion of partially compensated SRS gain results in a system with non-identical spans, even for systems with identical physical parameters for all spans. In such a system, amplifiers compensating for purely fiber attenuation may not be optimal, as the optimized average spectral power may be different for different spans. Changing the amplifier gain between spans provides additional degrees of freedom for optimization. For a C-band system, exploiting these additional degrees of freedom provides negligible gains for the scenarios with identical spans investigated in this paper. For a C+L-band system, the results are inconclusive. The assumption used in this paper to allow separation of the SRS and Kerr nonlinear effects is that any single channel or corresponding power variable has a negligible impact on the overall SRS gain spectrum. The addition of a variable corresponding to the total gain variation of an amplifier affects all of the channels in a span, and is directly correlated with the SRS strength. When the SRS strength grows significant, such as in a C+L-band system, this violates the assumption that enabled separation of the SRS and Kerr effects.

The fiber model used in this paper assumed a spectrally flat attenuation. The wavelength-dependent attenuation of optical fibers [19] is another source of channel inhomogeneity that can lead to benefits from channel power optimization. The slope of the wavelength dependence of the fiber attenuation, largely the result of Rayleigh scattering, provides additional loss to those shorter-wavelength channels depleted by SRS, compounding the two effects. The method described in Section II-C for adjusting the discretized Gaussian noise model in the presence of wavelength dependent gain can be similarly applied to the wavelength dependent loss scenario. This enables optimization with a feasible set of integrated coefficients with a small approximation error.

For C-band communication systems, the presence of SRS provides enough variation in per-channel SNRs to allow per-channel power optimization to significantly outperform the optimal flat power allocation, as shown in Fig. 6. A linear-SNR equalization for SRS captures most of the available performance

benefit, but there is a small benefit available from further optimization. Per-span SRS gain compensation provides negligible benefit for SRS-equalized power allocations, assuming periodic WSSs such as the every five spans assumed in this paper.

For C+L-band communication systems, spectrally flat power allocations are distinctly sub-optimal, even with perfect per-span SRS gain compensation. There remains a significant performance gap between a simple SRS-equalized power allocation, and a linear plus nonlinear noise and SRS-optimized power allocation. It is inadvisable to provide the same amplifier gains to both C and L bands as independent tuning allows a significant performance improvement for C+L-band systems without per-span SRS gain compensation. With amplifiers that compensate purely for attenuation in each span, there is a significant performance benefit from per-span SRS gain compensation, as shown in Fig. 11.

The simulation results presented here should be considered highly accurate only for systems with Gaussian signals, but the optimization methods and performance benefits are applicable to a much broader class of systems. The dependence of nonlinear noise on the modulation format [20] implies that channel powers optimized using the Gaussian noise model will not be precisely optimal when signals are non-Gaussian. The extended GN model [21] includes correction terms for non-Gaussian modulation formats, and agrees with split-step simulations within 0.2 dB across a variety of fiber types, with even greater accuracy for standard single-mode fiber. Our method's accuracy can be improved, at the price of increased complexity, by incorporating correction terms for the nonlinear noise spectrum [21] in the nonlinear coefficient calculation (20). Since the nonlinear noise is cubic in power, if nonlinear noise modeling is subject to an error of ϵ dB, the optimal channel powers will be incorrect by approximately $\epsilon/3$ dB.

Even with ideal Gaussian-distributed signals, imperfect system modeling can lead to errors in optimized channel powers. If such system modeling errors impact the nonlinear noise strength without significantly changing the frequency dependence of the nonlinear coefficients (which is determined primarily by path-integrated dispersion), the resulting error in optimized powers will be a uniform dB shift. Such a one-dimensional uniform power optimization error can be easily detected and corrected using measurements of the underlying modeled system. Only modeling errors that significantly change the spectral shape of the nonlinear interaction, and hence change the direction of the calculated SNR gradients, or those that change the ASE noise spectrum, will significantly impair system performance optimization.

V. Conclusion

An expression is formulated for the channel SNR in the presence of the Kerr nonlinearity, modeled using a Gaussian noise model, and inter-channel gain due to stimulated Raman scattering. Combining these two effects is found to result in non-convex expressions, even for simplified forms. An optimization approach using a sequence of convex objectives is able to find good locally optimal solutions, and provide loose bounds on the sub-optimality of those solutions. The extension of the GN model to the wide spectrum of C+L-band systems, along with the proposed simplifications for the incorporation of SRS in optimization, is verified via split-step Fourier simulation. Significant increases in minimum margin are obtained from wavelength-varying SRS-aware power allocations as compared to the optimal spectrally flat power allocation. For a C band system with a homogeneous required SNR, there is a 0.09 to 0.14 dB increase in minimum margin from using a Kerr nonlinearity-aware power allocation. For a similar C+L-band system the minimum margin improvement of an SRS and Kerr nonlinearity-aware power optimization increase to 0.23 to 0.60 dB, depending on the level of per-span SRS gain compensation. Per-span power SRS gain compensation provides small minimum margin increases for C band systems with SRS-aware power allocations, but provides a minimum margin increase up to 1.23 dB for C+L-band systems, even with optimized power allocations. When spectrally flat or SRS-unaware power allocations are used, per-span SRS gain compensation yields significant increases in minimum margin for both C- and C+L-band systems.

Appendix
Including SRS in Gaussian Noise Model

In the high-dispersion regime where the Gaussian noise model is valid [2], four-wave mixing products are small such that the majority of the Kerr nonlinearity does not redistribute power between channels. In this way, given an input power profile and a known fiber plant, the SRS gain and loss profile can be determined from channel powers under linear propagation. With the inclusion of SRS, the effective attenuation profile of the fiber system becomes both spatially and frequency varying. Assuming this attenuation profile $\alpha(z,f)$ has been determined, we insert it into the Gaussian noise model derivation and follow it through to the nonlinear noise spectral density expression, as in [22].

We start with [22, eq. (45)]. The mean nonlinear rotation term is modified to integrate out the spatial and frequency variations of the attenuation profile:

$$\frac{\partial}{\partial z}E(z,f) = \left[-j\beta(f) - j2\gamma\int_{-\infty}^{\infty} P_{TX}(f')e^{\int_0^z -2\alpha(\zeta,f')\partial\zeta}\partial f' - \alpha(z,f)\right]E(z,f) + Q_{NLI,\tilde{A}_i}(z,f) \tag{29}$$

This has solution

$$E(z,f) = e^{\Gamma(z,f)}\cdot\int_0^z e^{-\Gamma(z',f)}Q_{NLI,\tilde{A}_i}(z',f)\partial z' + e^{\Gamma(z,f)}E(0,f), \tag{30}$$

where

$$\Gamma(z,f) = \int_0^z\left[-j\beta(f) - j2\gamma\int_{-\infty}^{\infty} P_{TX}(f')e^{\int_0^{z'} -2\alpha(\zeta,f')\partial\zeta}\partial f' - \alpha(z,f)\right]\partial z'. \tag{31}$$

If we let

$$\Psi(z,f) = e^{-j2\gamma \int_{-\infty}^{\infty} P_{TX}(f') e^{\int_0^{z'} -2\alpha(\zeta,f')\partial\zeta} \partial f'}, \quad (32)$$

then we obtain the following expression for the Kerr nonlinear term:

$$\begin{aligned} Q_{NLI}(z,f) = &-j\gamma\Psi(z,f) \\ &\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-j[\beta(f_1)-\beta(f_1-f_2)+\beta(f-f_2)]z} \\ &\times e^{-\int_0^z [\alpha(z',f_1)+\alpha(z',f_1-f_2)+\alpha(z',f-f_2)]\partial z'} \\ &\times E(0,f)E^*(0,f_1-f_2)E(0,f-f_2)\partial f_1 \partial f_2. \end{aligned} \quad (33)$$

Using the discrete-frequency representation ξ_k of [22], we obtain

$$\begin{aligned} Q_{NLI}(z,f) = &-j\gamma f_0^{\frac{3}{2}} \Psi(z,f) \\ &\times \sum_{m,n,k=-\infty}^{\infty}\sum^{\infty}\sum^{\infty} \delta(f-[m-n+k]f_0)\xi_m\xi_n^*\xi_k \\ &\times e^{-\int_0^z [\alpha(z',mf_0)+\alpha(z',nf_0)+\alpha(z',kf_0)]\partial z'} \\ &\times \sqrt{G_{TX}(mf_0)G_{TX}(nf_0)G_{TX}(kf_0)} \\ &\times e^{-j[\beta(mf_0)-\beta(nf_0)+\beta(kf_0)]z}. \end{aligned} \quad (34)$$

The nonlinear noise electric field component is expressed as

$$\begin{aligned} E_{NLI}(z,f) &= e^{-j\beta(f)z}\Psi(z,f) e^{\int_0^z -\alpha(\zeta,f)\partial\zeta} \\ &\times \int_0^z e^{j\beta(f)z'} \Psi^{-1}(z',f) e^{\int_0^{z'} \alpha(\zeta,f)\partial\zeta} Q_{NLI}(z',f)\partial z' \\ &= -j\gamma f_0^{\frac{3}{2}} e^{-j\beta(f)z}\Psi(z,f) e^{\int_0^z -\alpha(\zeta,f)\partial\zeta} \sum_{i=-\infty}^{\infty} \delta(f-if_0) \\ &\times \sum_{m,n,k\in\tilde{A}_i} \sqrt{G_{TX}(mf_0)G_{TX}(nf_0)G_{TX}(kf_0)}\xi_m\xi_n^*\xi_k \\ &\times \int_0^z e^{\int_0^{z'} [\alpha(\zeta,[m-n+k]f_0)-\alpha(\zeta,mf_0)-\alpha(\zeta,nf_0)-\alpha(\zeta,kf_0)]\partial\zeta} \\ &\times e^{j[\beta([m-n+k]f_0)-\beta(mf_0)+\beta(nf_0)-\beta(kf_0)]z'}\partial z' \end{aligned} \quad (35)$$

The power of the nonlinear noise component is the expected value of the magnitude-squared electric field

$$G_{NLI}(z,f) = \mathcal{E}\left\{|E_{NLI}(z,f)|^2\right\}. \quad (36)$$

The time averaging of the various cross-terms of the discrete frequency terms remains unchanged from [22], giving the following expression for the nonlinear noise power spectrum, which is asymptotically exact for $f_0 \to 0$:

$$\begin{aligned} G_{NLI}(f) = &2\gamma^2 f_0^3 e^{\int_0^z -2\alpha(\zeta,f)\partial\zeta} \sum_{i=-\infty}^{\infty} \delta(f-if_0) \\ &\times \sum_m \sum_k G_{TX}(mf_0)G_{TX}(kf_0)G_{TX}([m+k-i]f_0) \\ &\times \left| \int_0^z e^{\int_0^{z'} [\alpha(\zeta,[m+k-i]f_0)-\alpha(\zeta,mf_0)-\alpha(\zeta,if_0)-\alpha(\zeta,kf_0)]\partial\zeta} \right. \\ &\left. \times e^{j[\beta([m+k-i]f_0)-\beta(mf_0)+\beta(if_0)-\beta(kf_0)]z'}\partial z' \right|^2 \end{aligned} \quad (37)$$

Taking the limit of $f_0 \to 0$, we converge to a continuous-time expression

$$\begin{aligned} G_{NLI}(f) = &2\gamma^2 e^{\int_0^z -2\alpha(\zeta,f)\partial\zeta} \\ &\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_{TX}(f_1)G_{TX}(f_2)G_{TX}(f_1+f_2-f) \\ &\times \left| \int_0^z e^{\int_0^{z'} [\alpha(\zeta,f_1+f_2-f)-\alpha(\zeta,f_1)-\alpha(\zeta,f_2)-\alpha(\zeta,f)]\partial\zeta} \right. \\ &\left. \times e^{j[\beta(f_1+f_2-f)-\beta(f_1)-\beta(f_2)+\beta(f)]z'}\partial z' \right|^2 \partial f_1 \partial f_2 \end{aligned} \quad (38)$$

This analysis is done for a single polarization, but it can similarly be done for dual polarizations as in [22]. For dual polarizations, the expression (38) has leading coefficient $\frac{16}{27}$ rather than 2.

For the dominant XPM terms of (38), $f_1 = f$ or $f_2 = f$ such that $f_1 + f_2 - f = f_2$ or $f_1 + f_2 - f = f_1$. Thus $\alpha(\zeta,[f_1+f_2-f]f_0) - \alpha(\zeta,f_1) - \alpha(\zeta,f_2) - \alpha(\zeta,f)$ reduces to $-2\alpha(\zeta, f_{\text{XPM}})$, where f_{XPM} is the frequency of the XPM-interfering channel.

REFERENCES

[1] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[2] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN model of non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[3] I. Roberts, J. M. Kahn, and D. W. 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] J. Bromage, "Raman amplification for fiber communications systems," *J. Lightw. Technol.*, vol. 22, no. 1, pp. 79–93, Jan. 2004.

[5] D. Schadt, B. Jaskorzynska, and U. sterberg, "Numerical study on combined stimulated raman scattering and self-phase modulation in optical fibers influenced by walk-off between pump and stokes pulses," *J. Opt. Soc. Amer. B*, vol. 3, pp. 1257–1262, Oct. 1986.

[6] A. Sharma and S. Arya, "Improved analysis for SRS- and XPM-induced crosstalk in SCM-WDM transmission link in the presence of HOD," *Optik-Int. J. Light Electron. Opt.*, vol. 120, no. 15, pp. 773–781, 2009.

[7] I. Djordjevic, "Transmission limitations of WDM transmission systems with dispersion compensated links in the presence of fiber nonlinearities," in *Proc. 5th Int. Conf. Telecommun. Modern Satell., Cable Broadcast. Serv.*, Sep. 2001, vol. 2, pp. 496–499.

[8] 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.

[9] B. Zhu *et al.*, "70 nm seamless band transmission of 17.3 tb/s over 40 × 100km of fiber using complementary raman/edfa," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper W3G.4. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2015-W3G.4
[10] K. Roberts and C. Laperle, "Flexible transceivers," in *Proc. Eur. Conf. Exhib. Opt. Commun.*, 2012, Paper We.3.A.3.
[11] T. Fehenberger, G. B"ocherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper Th2A.23.
[12] J. X. Cai *et al.*, "70.4 tb/s capacity over 7,600 km in c+l band using coded modulation with hybrid constellation shaping and nonlinearity compensation," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Paper Th5B.2.
[13] G. Saavedra *et al.*, "Experimental investigation of nonlinear signal distortions in ultra-wideband transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Paper W1G.1.
[14] K.-P. Ho, "Statistical properties of stimulated Raman crosstalk in WDM systems," *J. Lightw. Technol.*, vol. 18, no. 7, pp. 915–921, Jul. 2000.
[15] D. N. Christodoulides and R. B. Jander, "Evolution of stimulated raman crosstalk in wavelength division multiplexed systems," *IEEE Photon. Technol. Lett.*, vol. 8, no. 12, pp. 1722–1724, Dec. 1996.
[16] G. P. Agrawal, *Fiber-Optic Communication Systems*. Hoboken, NJ, USA: Wiley, 2010.
[17] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, "Identifying and attacking the saddle point problem in high-dimensional non-convex optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2933–2941.
[18] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[19] G. P. Agrawal, *Nonlinear Fiber Optics*, 5th ed. Oxford, U.K.: Academic, 2013.
[20] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.
[21] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, 2014.
[22] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A detailed analytical derivation of the GN model of non-linear interference in coherent optical transmission systems," arXiv:1209.0394, 2012.

Ian Roberts received the Bachelor's degree from the University of Ottawa, Ottawa, ON, Canada, the Master's degree from Stanford University, Stanford, CA, USA, both in electrical engineering, and is currently working toward the Ph.D. degree at Stanford University. He worked at Ciena on DSP hardware development for WaveLogic 3 Extreme prior to beginning graduate studies. His areas of interest 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 University of California, Berkeley, Berkeley, CA, USA, in 1981, 1983, and 1986, respectively. From 1987 to 1990, he was at 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 on the faculty of the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering at Stanford University, Stanford, CA, USA, where he heads the Optical Communications Group. In 2000, he helped found StrataLight Communications, where he served as the Chief Scientist from 2000 to 2003. StrataLight was acquired by Opnext, Inc., in 2009. 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 the recipient of the National Science Foundation Presidential Young Investigator Award in 1991. From 1993 to 2000, he served as 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.

James Harley received the M.A.Sc. degree in electrical engineering as an N.S.E.R.C. Scholar in digital communication theory from the University of Manitoba, Winnipeg, MB, Canada, in 1995. Since 1995, he has been working with the creative team of Nortel, Mississauga, ON, Canada, and currently Ciena, Hanover, MD, USA, in many different areas including coherent modem design (soft forward error correction, carrier recovery, clock recovery, dispersion compensation), digital network synchronization methods, DSP-based optical monitoring methods, optical line control, and novel cheap optical modems. He has been granted 32 U.S. patents to date. His currently a member of the principal coherent modem systems design team on the next generation of coherent modems and an Advisor on new optical line control.

David W. Boertjes received the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, where he worked on applications of active and passive polymer optics. He joined Nortels Optical Networks division in 1998, working on photonic components in fiber optic systems. He has several patents in the area of photonic transport and has been involved in several conferences including OFC/NFOEC and the Supercomputing Conference series. He was the recipient of the Cienas Distinguished Engineer Award, in 2014. He is currently a Technical Advisor working on Cienas next generation packet optical transport systems.