Open access • Journal Article • DOI:10.1109/JLT.2017.2668474

Capacity Analysis of Signaling on the Continuous Spectrum of Nonlinear Optical Fibers — Source link

Iman Tavakkolnia, Majid Safari

Institutions: University of Edinburgh

Published on: 01 Jun 2017 - Journal of Lightwave Technology (IEEE)

Topics: Channel capacity, Time domain, Communication channel, Signal and Noise (electronics)

Capacity Analysis of Signaling on the Continuous Spectrum of Nonlinear Optical Fibers

Capacity Analysis of Signaling on the Continuous Spectrum of Nonlinear Optical Fibers

Iman Tavakkolnia, *Student Member, IEEE/OSA* and Majid Safari, *Member, IEEE,*

***Abstract*—This paper investigates the capacity of signaling on the continuous spectrum (CS) of optical fibers as defined by nonlinear Fourier transform (NFT). The channel model of such NFT-based optical fiber communication systems is studied based on the behavior of the propagated optical signal in time domain for an asymptotically long fiber length. The derived channel model is validated using simulation for different scenarios of practical interest. An important characteristic of the channel in nonlinear spectral domain is the strong dependency of noise to signal. A variance normalizing transform is applied as a tool to obtain an estimate on the capacity of the underlying channel. The results predict a remarkable capacity for signaling on CS which can be utilized as an extra degree of freedom along with discrete spectrum (i.e., soliton transmission). It is further observed that the channel capacity of signaling on CS saturates at high signal power despite the linearizing capability of NFT.**

***Index Terms*—Optical fiber communication, nonlinear Fourier transform, inverse scattering, optical fiber capacity, continuous spectrum, noise statistics, variance normalizing transform.**

I. Introduction

In recent years, due to drastic growth in demands for high-speed data access, extensive efforts have been made to increase the information transmission rate over optical fibers. The long-haul links are specifically in the center of attention since they are the back-bone for almost all communication networks. Two different approaches can be taken to address this issue, designing more efficient fibers/fiber-based devices (e.g., fibers with lower nonlinearity or more efficient amplifiers) or optimizing data transmission and processing techniques in order to operate closer to the capacity of current optical fiber systems. However, the capacity of current long-haul fiber optic channels, impaired by nonlinear effects, cannot be simply exploited using conventional signal processing and transmission techniques developed for linear channels.

Despite many attempts in the literature (See [1]–[4] and the references therein), the optimal signaling and capacity limits for nonlinear optical fiber channels is still an open problem. Several techniques have been proposed to suppress nonlinearity along with other degrading effects such as dispersion while using traditional signaling schemes. Digital back propagation (DBP) [5] and phase conjugation [6] are two well-known examples, in which fiber impairments are compensated using signal processing techniques. Another possible approach is to design novel signal processing and transmission schemes that are optimized to exploit the capacity of the nonlinear fiber optic channel. This has been recently addressed in the context of inverse scattering [7]–[9].

Inverse scattering or nonlinear Fourier transform (NFT) is a well-established method for solving the integrable nonlinear Schrödinger (NLS) equation [10]. Using this method, the continuous spectrum (also known as radiation) and discrete spectrum (also known as solitons) of input signal are obtained which evolve linearly along the fiber. In [11], signaling on eigenvalues as part of the discrete spectrum were nominated as a suitable tool for data communication since they remain unchanged along the fiber in the absence of noise and other perturbations. It should be noted that this technique was different from the soliton transmission techniques in which soliton pulses are encoded/decoded in time domain [12]. More recently [7]–[9], nonlinear frequency division multiplexing (NFDM) was proposed as a general technique for data communication for an optical fiber link using nonlinear Fourier transform. In an NFDM system, data can be mapped on continuous and discrete spectra, defined in a generalized domain called nonlinear spectral domain (NSD). Different aspects of NFDM have became subjects for many research projects recently.

In the recent literature, either continuous spectrum (CS) or discrete spectrum (DS) are commonly considered as the data transmission medium. In [13]–[16], the performance of data transmission using only the CS was investigated assuming distributed and lumped amplification, and the non-ideal distributed amplification was taken into account in [17]. Recently, an experimental work [18] confirmed the feasibility of effective data transmission over CS for transoceanic distances. Periodic nonlinear Fourier transform was proposed recently in [19], [20]. In some other works [21]–[26], DS is considered, and performance of single-soliton or multisoliton transmission is investigated by simulation and experiment. Fast NFT and Inverse NFT (INFT) were proposed in [27], [28], and DBP using NFT was proposed in [29]. Multiuser NFDM was recently compared to wavelength division multiplexing (WDM) in [30]. Bounds for the capacity of single eigenvalue communication were derived in [31], [32]. In [33], an effective dispersion pre-compensation method was introduced for maximizing the transmission bit rate. In [34], both spectra were taken into account simultaneously, and it was observed that the noise variance is dependent on the signal amplitude.

In this paper, a channel model is developed for signaling on CS considering the signal-dependency of noise. The channel model is developed based on asymptotic analysis and validated

Manuscript received August 24, 2016; revised January 16, 2017; Accepted February 8, 2017.

Authors are with the Institute for Digital Communications, University of Edinburgh, Edinburgh, EH9 3JL, UK. e-mail:(i.tavakkolnia@ed.ac.uk)

using simulation. Based on this model, an estimate on the channel capacity is obtained using a variance normalizing transform (VNT). It is shown that CS provides an extra degree of freedom with remarkable achievable data rates; however, its capacity ultimately saturates at high signal power. Performance evaluations in literature agree with this result as degradation is observed when launch power is increased.

It should be noted that a related work, [35], was published while the current paper was under review. In [35], using a different approach, an estimated lower bound was reported when the data is mapped only on CS. The general findings of [35] are consistent with the results of this paper.

The outline of this paper is as follows: In the next section, basics of NFT, numerical methods, and general NFT-based communication systems are reviewed. In section III, different aspects of channel are studied and a channel model is developed. In section IV, an estimate on the channel capacity is obtained, followed by discussion and conclusion in section V and VI.

II. Problem Formulation

In this section theoretical basics of NFT-based systems are briefly elaborated [34]. More details can be found in [7], [10], [14]. The propagation of the complex envelope of a narrowband optical field in a standard single-mode fiber can be described by the stochastic nonlinear Schrödinger (NLS) equation [12]. Assuming the fiber loss to be perfectly compensated by ideal distributed amplification, the NLS equation can be described as

$$\frac{\partial Q(T,l)}{\partial l} = -\frac{j\beta_2}{2}\frac{\partial^2 Q(T,l)}{\partial T^2} + j\gamma Q(T,l)\left|Q(T,l)\right|^2 + N(T,l),\ \ 0 \le l \le L, \tag{1}$$

where $Q(T,l)$, l, and T respectively represent the complex envelope of the optical field, distance in km, and retarded time in seconds. For ordinary time τ we have $T = \tau - \beta_1 l$. Typical values of chromatic dispersion coefficient β_2 and nonlinearity parameter γ are $22 \times 10^{-24}[\mathrm{s}^2/\mathrm{km}]$ and $1.27[\mathrm{W}^{-1}\mathrm{km}^{-1}]$ respectively. $N(T,l)$ represents the amplified spontaneous emission (ASE) noise added by the amplifiers which is a white Gaussian process with autocorrelation

$$\mathsf{E}[N(T,l)N^*(T',l')] = \frac{N_{ASE}}{L}\,\delta(T-T')\delta(l-l'),$$

where $N_{ASE} = \alpha L h\nu_s K_T$ is the accumulated spectral density of noise at fiber length L. $\alpha = 0.046[\mathrm{km}^{-1}]$ is the fiber loss, $h\nu_s = 1.28 \times 10^{-19}[\mathrm{J}]$ is the photon energy , and $K_T = 1.13$ is the photon occupancy factor. Throughout this paper we consider the focusing case $\beta_2 < 0$ and no dispersion compensation. It is also assumed that the pulses are wide enough, so that higher order dispersions can be ignored. The NLS equation can be normalized in the form of

$$jq_z(t,z) = q_{tt}(t,z) + 2|q(t,z)|^2 q(t,z) + n(t,z), \tag{2}$$

by normalization rules

$$q = \sqrt{\gamma L_D}Q, \quad z = \frac{l}{2L_D}, \quad t = \frac{T}{T_0}, \tag{3}$$

where the normalizing parameters T_0 and $L_D = T_0^2/|\beta_2|$ can be chosen independent of other parameters. For isolated pulses, T_0 and L_D indicate the pulse width in time domain and dispersion length. For more complex signal shapes, these normalization parameters do not correspond to the same physical attributes, but they can be modified to control the physical characteristics of the system such as bandwidth, temporal duration, or induced dispersion. Therefore, the autocorrelation of the normalized noise can be described as

$$\mathsf{E}[n(t,z)n^*(t',z')] = \sigma^2\delta(t-t')\delta(z-z'),$$

where based on (3)

$$\sigma^2 = \frac{N_{ASE}}{L}\frac{2\gamma L_D^2}{T_0}.$$

Equations (1) and (2) are integrable when noise is absent, and inverse scattering can be applied on (2) independent of the channel parameters.

A. Inverse scattering for nonlinear Schrödinger equation

The integrable NLS equation in (2) can be described by a couple of auxiliary operators with invariant eigenvalues called Lax pairs [36]. Thus, NLS is solvable using inverse scattering [37], [38]. The M invariant eigenvalues of these operators (denoted by λ_m's), which can be expressed in terms of $q(t,z)$, remain unchanged even when $q(t,z)$ evolves in z. Using inverse scattering, nonlinear Fourier transform (NFT) can be defined which transforms the time domain optical signal into scattering data consist of discrete and continuous spectra $\{\rho(\lambda,z), \{C_m(z)\}_{m=1}^M, \{\lambda_m\}_{m=1}^M\}$ which evolve linearly along the fiber in nonlinear spectral domain. The continuous spectrum $\rho(\lambda,z)$ is defined on the real axis $\lambda \in \mathbb{R}$, and the discrete spectrum $C_m(z)$ is defined on the upper half complex plane $\lambda_m \in \mathbb{C}^+$. It can be shown that, in a noise-free scenario, the eigenvalues are preserved during the evolution along the fiber, and the nonlinear fiber only imposes a phase shift on the continuous and discrete spectra expressed as

$$\lambda_m(z) = \lambda_m(0) = \lambda_m, \quad m = 1, \cdots, M, \tag{4a}$$

$$C_m(z) = C_m(0)e^{-4j\lambda_m^2 z}, \quad m = 1, \cdots, M, \tag{4b}$$

$$\rho(\lambda,z) = \rho(\lambda,0)e^{-4j\lambda^2 z}. \tag{4c}$$

The signal $q(t,z)$ is fully determined by the full spectra defined in (4) at any distance z. As a consequence, the signal space of the optical fiber channel is fully described by continuous and discrete spectra, and therefore, the data can be mapped onto multiple degrees of freedom, namely, eigenvalues, DS, and CS. The forward and inverse transforms are defined here.

1) NFT: The NFT is a transformation that decomposes a signal $q(t)$ into its discrete and continuous spectra in nonlinear spectral domain. It is performed by solving the so-called Zakharov-Shabat eigenvalue problem formulated as

$$\frac{d\boldsymbol{v}}{dt} = \begin{bmatrix} -i\lambda & q(t) \\ -q^*(t) & i\lambda \end{bmatrix}\boldsymbol{v}, \quad \boldsymbol{v}(t \to -\infty, \lambda) \to \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{-j\lambda t}, \tag{5}$$

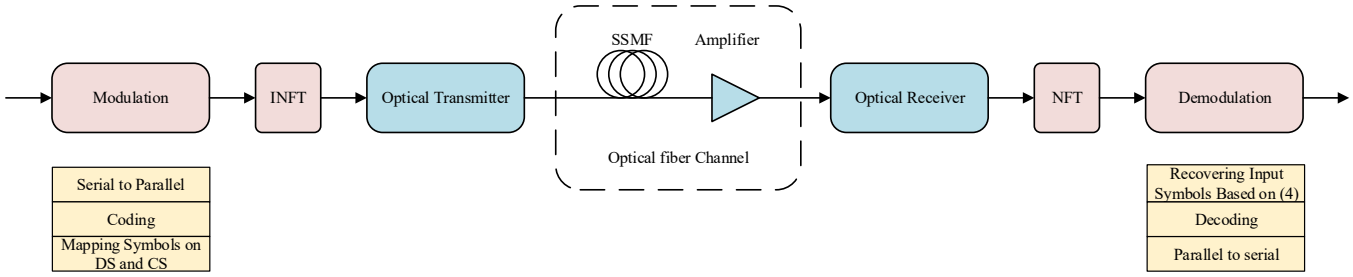

Fig. 1. General block diagram of a NFT-based communication system. SSMF: standard single mode fiber.

$$a(\lambda) = \lim_{t\to\infty} v_1 e^{j\lambda t}, \tag{6}$$

$$b(\lambda) = \lim_{t\to\infty} v_2 e^{-j\lambda t}. \tag{7}$$

Here it is assumed that $q(t) \in L^1(\mathbb{R})$, and $|q(t)| \to 0$ as $|t| \to \infty$. Zeros of $a(\lambda)$ on $\mathbb{C}^+$ are the eigenvalues, and the spectra are defined as

$$\rho(\lambda) = \frac{b(\lambda)}{a(\lambda)}$$

$$C_m = \frac{b(\lambda_m)}{\mathrm{d}a(\lambda)/\mathrm{d}\lambda|_{\lambda=\lambda_m}}$$

In above equations the z-dependency is omitted because the equation is solved respect to the variable t in any needed z. The same argument applies to the INFT below.

2) INFT: When the spectra are known, the signal $q(t)$ can be calculated using inverse NFT. INFT is performed by solving either the Riemann-Hilbert system [7], or the Gelfand-Levitan-Marchenko (GLM) integral equations [10]. In this paper we consider GLM. The GLM equations are defined for $y \geq x$ as

$$K_2^*(x,y) + \int_x^\infty K_1(x,s)F(s+y)ds = 0, \tag{8a}$$

$$K_1(x,y) - \int_x^\infty K_2^*(x,s)F^*(s+y)ds = F^*(x+y), \tag{8b}$$

where the kernel for the above integrals is given by

$$F(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \rho(\lambda)e^{j\lambda x}\, d\lambda - j\sum_{m=1}^{M} C_m e^{j\lambda_m x}.$$

Having $K_1(x,y)$, by solving the GLM equations in (8), the signal in time domain is reconstructed as

$$q(t) = -2K_1(t,t).$$

B. NFT-based communication systems

Fig. 1 demonstrates a general NFT-based optical fiber communication system. The data is directly encoded on CS and/or DS in the nonlinear spectral domain. Data is first converted to a sequence of digital symbols of a specific modulation format (e.g., QAM) via serial to parallel conversion and constellation mapping. The sequence of digital symbols associated to CS are then converted to an oversampled continuous waveform using a pulse shaping filter (e.g., raised-cosine or sinc pulses) to enhance spectral efficiency. Note that since the integral equations of INFT are solved digitally in practice, oversampling is essential for minimizing numerical errors. Then, INFT is applied to the generated CS waveform along with the encoded DS to generate the signal in time domain. The optical transmitter and receiver consist of all operations needed for launching the signal into the fiber and detecting it as a digital signal, such as filtering and analog/digital conversion. As indicated before, standard single mode fiber (SSMF) and ideal distributed amplification are assumed. After photo-detection at the receiver, NFT is performed, and the noisy outputs are obtained. The linear dispersion removal is then performed based on equations (4).

There are several numerical methods proposed in the literature for NFT and INFT [8], [14], [15], [27], [34], [39]. We use the Ablowitz method for NFT, and discretization of Marchenko equations and using the symmetry of the resulted matrix for INFT [34]. In this work we consider only the CS, or in other words, the DS is zero throughout this paper.

III. Channel model

In this section, a channel model will be developed for the CS channel, i.e, signaling only on the continuous spectrum. First, a noise model is derived based on an asymptotic analysis which signifies the fundamental features of the channel. Asymptotic analysis here refers to the scenario in which normalized fiber length z defined in (3) is large. This asymptotic scenario can be satisfied for long fiber (L large) or highly dispersive case (L_D small). Simulation results are then presented to validate the channel model in practical scenarios.

A. Asymptotic solution of NLS equation

For relatively long fiber length, it can be shown that signal can be expressed in a slow-varying similarity solution form [40]

$$q(t,z) = \frac{1}{\sqrt{z}}A(t,z)e^{jz\theta(t,z)} \tag{9a}$$

$$A(t,z) = f\left(\frac{t}{z}\right) + \sum_{n=1}^{\infty}\sum_{k=0}^{n}\frac{(\log z)^k}{z^n} f_{n,k}\left(\frac{t}{z}\right) \tag{9b}$$

$$\theta(t,z) = \frac{1}{4}\left(\frac{t}{z}\right)^2 + \sum_{n=1}^{\infty}\sum_{k=0}^{n}\frac{(\log z)^k}{z^n}\theta_{n,k}\left(\frac{t}{z}\right) \tag{9c}$$

where

$$\theta_{1,0} = g$$
$$\theta_{1,1} = 2f^2$$
$$f_{1,0} = fg^{''} + 2g^{'}f^{'} + 4f(2(f^{'})^2 + ff^{''})$$
$$f_{1,1} = 4f(3(f^{'})^2 + ff^{''})$$

where $f(\frac{t}{z}) \geq 0$ and $g(\frac{t}{z})$ are real valued functions. For simplicity, terms $(\frac{t}{z})$ are omitted in above equations. The order of remaining terms would be $O((\frac{\log z}{z})^2)$, so for large z higher orders can be ignored without introducing significant error. It should be recalled that $z = L/2L_D$. Thus, asymptotic scenario (large z) can be satisfied for long fiber or high signal bandwidth (small T_0) or a combination of both.

There are different ways to find $f(\frac{t}{z})$ and $g(\frac{t}{z})$ in terms of spectra [40]–[42]. For the case that DS is zero, the asymptotic solution for the NLS equation can be related to the CS as follows:

$$f^2(\frac{t}{z}) = \frac{1}{4\pi}\log\{1 + |\rho(\frac{-t}{4z})|^2\},$$
$$g(\frac{t}{z}) = -\arg\{\rho(\frac{-t}{4z})\} + \frac{3\pi}{4} - \arg\{\Gamma(1 + 2jf^2(\frac{t}{z}))\}$$
$$- 2\log 2f^2(\frac{t}{z}) - 2\int_{t/z}^{\infty}\log(y - \frac{t}{z})\frac{\partial f^2(y)}{\partial y}\mathrm{d}y$$
$$+ 2\int_{-\infty}^{t/z}\log|y + \frac{t}{z}|\frac{\partial f^2(y)}{\partial y}\mathrm{d}y.$$

Therefore, the asymptotic temporal signal can be expressed in terms of $\rho(\cdot)$. Now by only considering the leading factor in (9b), we have

$$A(t,z) \approx f\left(\frac{t}{z}\right), \tag{10}$$

and therefore the amplitudes of temporal signal $q(t,z)$ at length z can be expressed in terms of amplitude of CS by

$$|q(t,z)|^2 = \frac{1}{z}f^2(\frac{t}{z}) = \frac{1}{4\pi z}\log\{1 + |\rho(\frac{-t}{4z})|^2\} \tag{11}$$

or inversely

$$|\rho(\lambda)|^2 = e^{4\pi z|q(-4z\lambda,z)|^2} - 1. \tag{12}$$

Fig. 2 compares the simulated amplitude of the temporal signal propagated over the fiber at $z = 10$ with the corresponding signal predicted by the asymptotic equation in (11). We first generate a CS continuous waveform by applying raised-cosine filtering to a sequence of 16 random symbols from the 16QAM constellation. By applying INFT on this CS signal, the temporal signal is generated at the input of the fiber which occupies a bandwidth of $W = 30$ GHz measured based on 99% energy of the signal $Q(T,0)$. It can be seen that the temporal signal can be well approximated by asymptotic equation (11) while increasing z would further reduces the error.

Note that inverse scattering implies that this asymptotic solution for propagation over optical fiber can be also used to describe the behavior of nonlinear Fourier transform at the end of the fiber.

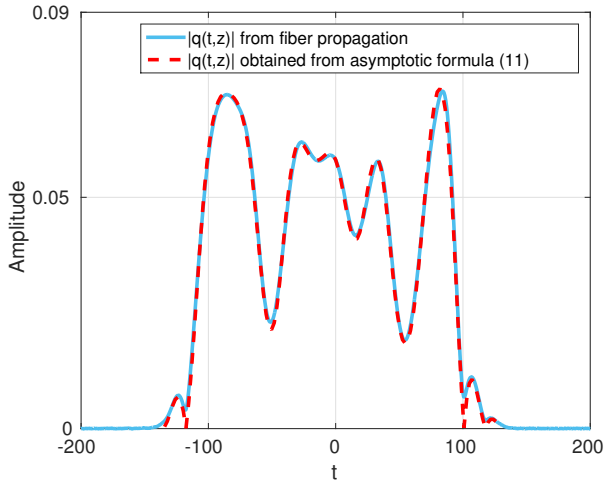

Fig. 2. Accuracy of the asymptotic formula for temporal signal in (11) at $z = 10$ where t is the normalized time defined in (3) for $T_0 = 0.05$ ns.

B. Channel statistics based on the asymptotic analysis

The nonlinear Fourier transform allows us to describe nonlinear optical fiber channels in a linear form based on signaling over the nonlinear spectrum. For CS, this linear channel model is give by [7]

$$\rho(\lambda, z) = \rho(\lambda, 0)e^{-4j\lambda^2 z} + \eta(\lambda), \tag{13}$$

where $\eta(\lambda)$ is the noise added to CS in nonlinear spectral domain. Letting $\rho(\lambda) = \rho(\lambda, z)$ and $\rho_0(\lambda) = \rho(\lambda, 0)\exp(-4j\lambda^2 z)$ respectively denote the received noisy CS and the CS in the noise free scenario, (13) is simplified as

$$\rho(\lambda) = \rho_0(\lambda) + \eta(\lambda). \tag{14}$$

It can be shown that the noise is zero-mean, i.e., $\mathbf{E}\{\rho(\lambda)\} = \rho_0(\lambda)$. This can be done readily by considering that $\mathsf{E}\{n(t,z)\} = 0$ in (2) and taking an ensemble average from the Zakharov-Shabat eigenvalue problem (5) which are basically two coupled differential equations.

In order to obtain the second order statistics of the noise in CS domain, we will use the asymptotic solution introduced in III-A to describe NFT at the end of the fiber. Let $q(t,z) = q_0(t,z) + n(t)$ denote the propagated signal in time at the end of fiber where $n(t) = n_r(t) + jn_i(t)$ is the noise and $q_0(t,z) = q_0^r(t,z) + jq_0^i(t,z)$ is the signal when noise is absent. Using (12), the intensity of CS signal after applying NFT can be expressed asymptotically as

$$|\rho(\lambda)|^2 = e^{4\pi z|q_0(t,z)+n(t)|^2} - 1.$$

Accordingly, the second moment of the noise in CS can be written as

$$\mathsf{E}\{|\eta|^2\} = \mathsf{E}\{|\rho|^2\} - |\rho_0|^2 \tag{15}$$
$$= e^{4\pi z|q_0|^2}\left(\mathsf{E}\{e^{4\pi z\left(n_r^2 + n_i^2 + 2q_0^r n_r + 2q_0^i n_i\right)}\} - 1\right).$$

Here, for simplicity, the dependence on λ and time is omitted and $q_0 = q_0(t,z)$. Assuming circularly symmetric Gaussian noise $n(t) \sim \mathcal{N}(0,\sigma_0^2)$ at the end of fiber, the part corresponding to the real noise in time can be described as

$$\mathsf{E}\{e^{4\pi z\left(n_r^2+2q_0^r n_r\right)}\} = \frac{e^{\frac{(4\pi z q_0^r)^2}{c}}}{\sqrt{\pi\sigma_0^2}} \int_{-\infty}^{\infty} e^{-ct^2}\mathrm{d}t,$$

where $c = \frac{1}{\sigma_0^2} - 4\pi z$. For practical ranges of fiber-optic channel parameters, the parameter c remains positive[1] for which we have

$$\mathsf{E}\left\{e^{4\pi z\left(n_r^2+2q_0^r n_r\right)}\right\} = \frac{1}{\sqrt{c\sigma_0^2}} e^{\frac{(4\pi z q_0^r)^2}{c}}. \tag{16}$$

Similar derivations can me made for imaginary part of (15) and thus, the noise variance in CS can be asymptotically calculated as

$$\begin{aligned}\mathsf{E}\{|\eta|^2\} &= e^{4\pi z|q_0|^2}\left(\frac{e^{\frac{|4\pi z q_0|^2}{c}}}{c\sigma_0^2} - 1\right) \\ &= \left(|\rho_0|^2+1\right)\left(\frac{\left(|\rho_0|^2+1\right)^{\frac{4\pi z}{c}}}{c\sigma_0^2} - 1\right).\end{aligned} \tag{17}$$

It can be seen that the noise variance depends on the signal amplitude $|\rho_0|$, and the order of dependence is $2 + 8\pi z/c$ which is larger than two. In the following sections, it will be demonstrated that this characteristic plays an important role in determining the capacity of the channel. We also expect that the distribution of noise not to be normal in general. In the next section, the behavior of the CS channel will be studied numerically and these main characteristics (i.e. signal dependency of noise, order of dependence larger than 2, and non-Gaussian distribution) are verified even at not necessarily asymptotic regimes.

C. Numerical study of the channel model

In this section, simulation results are presented to study the characteristics of noise in CS. It is essential to minimize the numerical error in NFT and INFT calculation as well as split-step Fourier method for fiber propagation, so that only noise remains as the source of distortion. This would increase the computational complexity, and since a large number of samples are needed for estimating the noise statistics, the simulations would be highly time consuming. To overcome this problem, the simulations were all performed using more than 500 CPUs available for parallel computing at the Institute for Digital Communication at the University of Edinburgh without which the time needed for these simulations would have been escalated.

It should be noted that the complexity of numerical methods increases when the input power increases as observed in [18], [34]. For a larger signal amplitude of CS a higher time resolution is needed for INFT in order to achieve the same error as for smaller amplitudes of signal. Therefore, it is essential to adaptively increase complexity for noise analysis to keep the numerical errors constant and below a desired value. However, downsampling is performed after calculation of $q(t,0)$ for launching the signal into the fiber which will not affect the numerical errors in the next steps.

[1]The normalized noise power at the end of the fiber can be expressed as $\sigma_0^2 = \alpha h\nu_s K_T \gamma L_D LW$ where W is the linear signal bandwidth. Then, the criteria for positive c can be obtained as $W < 1/(2\pi\alpha h\nu_s K_T\gamma L^2)$. For example, for the parameters used in this paper and $L = 3000$ km we should have $W < 2$ THz.

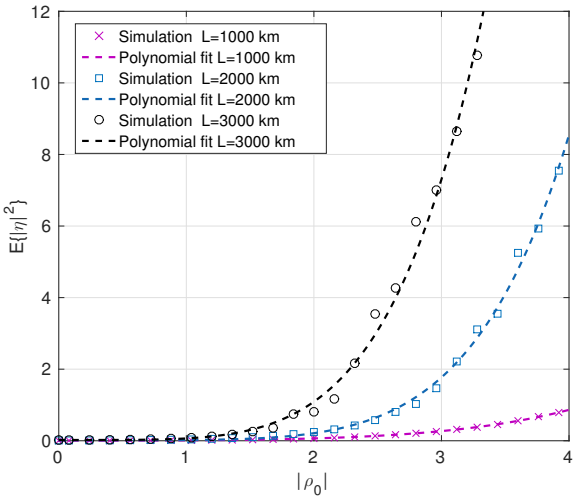

Fig. 3. Variance of noise $\mathsf{E}\{|\eta|^2\}$ for different signal amplitudes $|\rho_0|$.

In order to obtain the noise statistics, several streams of random symbols in various nonlinear frequencies λ with different noise realizations were analyzed, truncated sinc pulses were utilized for pulse shaping, and $T_0 = 0.1$ ns was chosen. Overall about 1.6×10^7 random samples were used. Split-step Fourier method was used for propagation of signal along the fiber, and ASE noise was added in each step. Linear signal bandwidth was $W = 50$ GHz. After NFT, noise samples were first extracted and then their statistic were calculated.

The dependence of noise variance on the amplitude of signal in CS is shown in Fig. 3. In order to plot this figure, several combination of real and imaginary signals are considered for each amplitude, resulting in the same noise statistics for all of them. It can be seen in Fig. 3 that the variance of noise drastically increases for longer fibers. The Simulation results also show that the variance of noise has a dependency to signal amplitude with an order larger than two if the variance is approximated by a polynomial. For example, in Fig. 3, the orders of dependence for different fiber lengths are about 4. This result numerically confirms that the asymptotic statistical behavior of the noise in CS as shown in equation (17) is even true for these lower values of z which correspond to some practical scenarios.

In Fig. 4, the PDF of the received signal amplitude in CS $\rho(\lambda)$ is demonstrated. Many noise samples are extracted at different nonlinear frequencies for which $\rho_0(\lambda) = 1$, 2, or 3. As expected from III-A, it can be seen that the distribution is not normal, and moreover the variance is increasing for higher values of input signal.

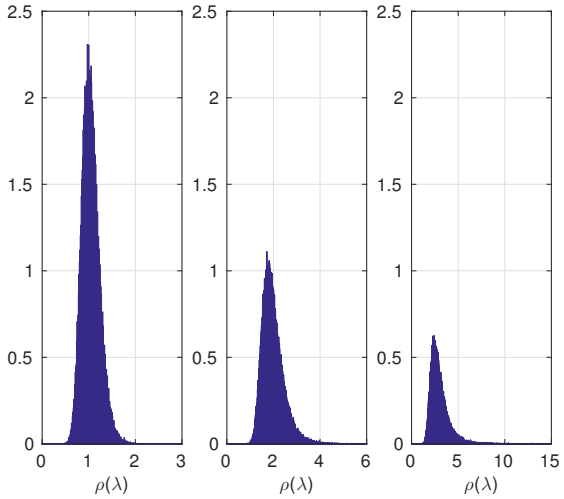

Fig. 4. PDFs of received noisy CS for different input signal $\rho_0(\lambda) = 1, 2, 3$.

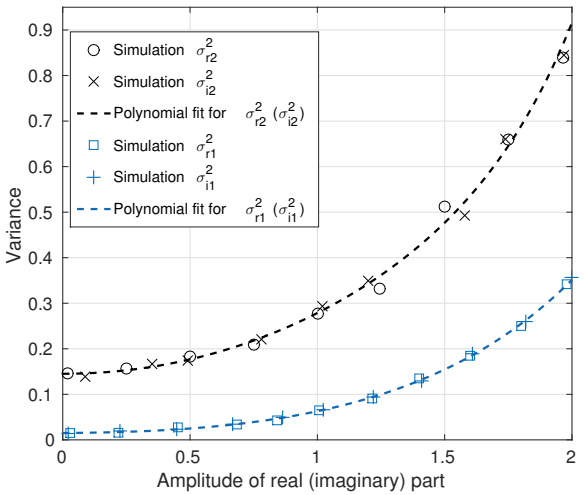

Fig. 5. Dependence of variance of noise in real (imaginary) channel on the amplitude of signal in imaginary (real) channel.

Note that the asymptotic model described in section III-A does not deliver explicit information about the noise on real and imaginary parts of CS. Here, we introduce the additional variables below in order to investigate the behavior of the noise for real and imaginary part of CS.

$$\rho_0(\lambda) = \rho_0^r(\lambda) + j\rho_0^i(\lambda), \quad \rho_0^r(\lambda), \rho_0^i(\lambda) \in \mathbb{R}, \tag{18a}$$

$$\rho(\lambda) = \rho^r(\lambda) + j\rho^i(\lambda), \quad \rho^r(\lambda), \rho^i(\lambda) \in \mathbb{R}, \tag{18b}$$

$$\eta(\lambda) = \eta^r(\lambda) + j\eta^i(\lambda), \quad \eta^r(\lambda), \eta^i(\lambda) \in \mathbb{R}. \tag{18c}$$

Furthermore, the conditional noise variance of real and imaginary parts of CS are defined below to demonstrate the dependency between real and imaginary parts.

$$\sigma^2_{ry}(x) = \mathsf{E}\{|\eta^r(\lambda)|^2 | \rho_0^r(\lambda) = x, \rho_0^i(\lambda) = y\}, \tag{19a}$$

$$\sigma^2_{iy}(x) = \mathsf{E}\{|\eta^i(\lambda)|^2 | \rho_0^i(\lambda) = x, \rho_0^r(\lambda) = y\}. \tag{19b}$$

As depicted in Fig. 5, the simulation results of the conditional variances, $\sigma^2_{r1}(x)$ and $\sigma^2_{r2}(x)$, show that the noise on real and imaginary parts of signal are dependent. Hence, real and imaginary channels on CS cannot be considered as independent channels, but nonetheless, these two channel have identical noise behavior. This is concluded from the fact that $\sigma^2_{ry}(x) = \sigma^2_{iy}(x)$ as demonstrated in Fig. 5. It should be emphasized that despite dependence of real and imaginary channels, the variance of noise $\mathsf{E}\{|\eta(\lambda)|^2\}$ perfectly depends only on the amplitude of signal $|\rho_0(\lambda)|$ as shown in Fig. 3. Therefore, it can be concluded that if signally is only performed on the real channel ($\rho_0^i(\lambda) = 0$), the variance of the real noise would be perfectly dependent on the signal amplitude.

IV. Channel Capacity

In this section, the capacity of the CS channel is investigated based on the channel model introduced in the last section and using a variance normalizing transform, which will be defined next.

A. Variance Normalizing transform

In this section, a variance normalizing transform (VNT) is introduced that can be applied to any random variable with a variance that is a function of its mean to generate an approximately Gaussian random variable with a variance independent of its mean [43]–[45]. VNT can be essentially used to convert an additive (possibly non-Gaussian) noise channel with signal-dependent noise to a conventional additive Gaussian noise channel where signal and noise are independent [43]–[45]. In effect, the conventional coding and signal processing techniques would be sufficient for efficient communication over the transformed channel. Shot noise is an example of a signal-dependent noise and VNT has been recently used for efficient signaling and capacity approximation over shot-noise-limited channels [46]–[48].

The variance normalizing transform that normalizes the random variable W, with mean μ_W and variance

$$\sigma^2_W = f^2(\mu_W), \tag{20}$$

is defined as [43]

$$T(s) = \int \frac{1}{f(s)} \mathrm{d}s. \tag{21}$$

The normalized (i.e., transformed) random variable $Y = T(W)$ has then the statistics of $\sigma^2_Y \simeq 1$ and $\mu_Y \simeq T(\mu_W)$ for sufficiently large values of μ_W. In [45], it was also shown that the probability distribution of the normalized random variable tends to Gaussian distribution for a family of originally non-Gaussian probability distribution.

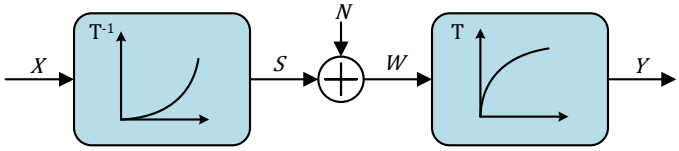

Fig. 6. Normalizing communication channel with signal-dependent noise using VNT.

Let random variable W be the output of a channel where the signal $S = \mu_W$ is corrupted by a zero-mean signal dependent noise N with variance as in (20). Applying VNT as defined in (21) to the noisy signal $W = S + N$, we will have

$$Y = T(W) = T(S + N) = T(S) + N_T,$$

where N_T is a zero mean Gaussian noise with unit variance independent of the transformed signal $T(S)$. Consequently, a communication channel can be defined as shown in Fig. 6, in which the signal is originally generated in a *transformed* domain and then mapped into the original signal domain S using inverse VNT. After transmission through the channel and addition of signal-dependent noise N, VNT is applied, and the output signal can be expressed as

$$Y = X + N_T. \tag{22}$$

which defines a conventional additive Gaussian noise channel. The following lemma shows that the capacity of this transformed channel is equal to the capacity of the original channel $W = S + N$ with signal-dependent noise.

Lemma. *$I(X;Y) = I(S;W)$ for the communication system in Fig. 6.*

Proof. For the communication system in Fig. 6 we have the Markov chain $X \leftrightarrow S \leftrightarrow W$. Thus, based on the data processing inequality [49], $I(S;W) \geq I(X;W)$. Since VNT is a deterministic injective function (i.e., $f(s) > 0$ and consequently $T(s_1) > T(s_2)$ for $s_1 > s_2$), the Markov chain $S \leftrightarrow X \leftrightarrow W$ also exists, and therefore $I(X;W) \geq I(S;W)$. Consequently, $I(X;W) = I(S;W)$. Similarly, for the two Markov chains $X \leftrightarrow W \leftrightarrow Y$ and $X \leftrightarrow Y \leftrightarrow W$, it can be shown similarly that $I(X;W) = I(X;Y)$. Hence, the equation $I(X;Y) = I(S;W)$ is proved. □

Based on this lemma, the capacity of the original channel with signal dependent noise can be expressed in terms of transformed input and output as

$$C = \sup_{\mathsf{E}\{|T^{-1}(X)|^2\} \leq P} I(X;Y). \tag{23}$$

Note that the mutual information maximization is performed subject to a transformed average power constraint where P denotes the average power limit in the original channel.

The following section will study the capacity of the CS channel defined in (14) based on the assumption that the signal bandwidth is not greater than the bandwidth for which the spectral density of noise remains constant, and consequently noise can be considered as white. In effect, using VNT, the CS channel can be transformed into an additive white Gaussian noise (AWGN) channel. We will further discuss the effect of correlation of noise in CS later in section V.

B. Capacity of signaling on CS

In section III, the statistical behavior of the noise in non-linear spectral domain has been investigated concluding that the noise becomes signal-dependent and non-Gaussian in CS. In addition, real and imaginary parts of the noise in CS are dependent. In this section, we first consider real CS signaling ($\Im\{\rho(\lambda)\} = 0$), for which the noise variance is only dependent on the signal amplitude, and employ VNT as a tool to estimate the capacity of CS channel defined in (14). We can then provide a bound for the capacity of complex signaling on CS.

We first start with the following theorem to show that if the dependency of the noise variance to signal is defined by a polynomial with an order higher than two, the signal space in the transformed channel (i.e., $X = T(S)$) is not only limited by the transformed average power constraint in (23) but also by an imposed peak power constraint. Note that both asymptotic analytical result in equation (17) and simulation results in Fig. 3 confirm that the variance of noise in CS is dependent to signal's amplitude by a polynomial of order higher than 2.

Theorem. *Let the dependency of the variance of noise to the signal amplitude be approximated by a polynomial $f(|s|) = a_n|s|^{b_n} + a_{n-1}|s|^{b_{n-1}} + ... + a_1|s|^{b_1} + a_0$, where $a_i \geq 0$ for $i = 0, ..., n$ and $b_n > b_{n-1} > ... > b_1$ are not necessarily integers. Then, the transform function, $T(\cdot)$ defined in (21) is bounded if $b_n > 2$.*

Proof. It should be noted that the transform properties imposes $T(0) = 0$. For large values of $|s| > \Gamma$ the polynomial can be approximated by its largest order, and thus for $\nu > 0$

$$\begin{aligned} T(\nu) &= \int_0^{\nu} \frac{1}{\sqrt{f(s)}} \mathrm{d}s \\ &\leq M + \int_{\Gamma}^{\nu} \frac{1}{\sqrt{a_n s^{b_n}}} \mathrm{d}s \\ &= M + \frac{1}{(1 - b_n/2)\sqrt{a_n}} t^{\frac{-b_n}{2}+1} |_{\Gamma}^{\nu} \\ &= A + \frac{1}{(1 - b_n/2)\sqrt{a_n}} \nu^{\frac{-b_n}{2}+1}. \end{aligned} \tag{24}$$

where M and A are finite values, and

$$M = \int_0^{\Gamma} \frac{1}{\sqrt{f(s)}} \mathrm{d}s,$$

$$A = M - \frac{1}{(1 - b_n/2)\sqrt{a_n}} \Gamma^{\frac{-b_n}{2}+1}.$$

It can be further shown that

$$\lim_{\nu \to \infty} \nu^{\frac{-b_n}{2}+1} = \begin{cases} \infty & b_n \leq 2 \\ 0 & b_n > 2 \end{cases},$$

Considering the limit above and noting that the second term in (24) is negative for all $b_n > 2$, we can write $T(\nu) \leq A$.

Replacing variable s with $-s$ in the equations above, it can be also shown that $T(\nu) \geq -A$ for $\nu < 0$. Finally, we can conclude that

$$|T(\nu)| \leq A$$

□

Note that VNT reveals an important consequence of the signal-dependency of the noise with an order higher than two which renders the signal space further limited by a peak power constraint. In other words, the signal dependency in nonlinear spectral domain is translated to peak-power-limited signal space in the transformed domain.

Applying the theorem above to the asymptotic statistics of the CS channel in (17), the signal amplitude in the transformed channel is limited by

$$A = \int_0^\infty \frac{1}{\sqrt{(s^2+1)\left(\frac{(s^2+1)\frac{4\pi z}{c}}{c\sigma_0^2} - 1\right)}} \mathrm{d}s, \qquad (25)$$

which is determined by channel parameters z and c which are related to fiber length, signal bandwidth, and ASE noise spectral density. Note that although the value of A in (25) is only accurate at asymptotic regimes but it gives some insight about how the maximum value of signal in the transformed channel and consequently the capacity depends on the channel parameters. For instance, for longer fiber links the limiting amplitude would be smaller because the noise variance is higher.

Here, some numerical results are presented on the effectiveness of the VNT approach before finalizing the capacity analysis. Fig. 7 depicts the variance normalizing transforms required to normalize the signal-dependent noise in CS for different fiber lengths. These are obtained directly from the simulation results, shown in Fig. 3, without any approximation. It can be seen that the transforms are limited to certain values. For instance, maximum amplitude A for fiber lengths $L = 1000, 2000$, and 3000 km are respectively 28.63, 14.35, and 9.11. Therefore, in the AWGN channel resulted from the VNT, the signal amplitude cannot be higher than a specific value determined by VNT.

In order to demonstrate the effectiveness of the VNT, the transform is performed on a symbol with amplitude $\rho_0 = 2$ for a 2000 km fiber. Applying the VNT the PDF of transformed samples tend to Gaussian distribution with mean $T(2) = 11.6$ and variance of approximately equal to one as demonstrated in Fig. 8a. The Q-Q plot in Fig. 8b confirms this statement.

Now, it is time to establish some results on the capacity of the CS channel defined in (14). Letting $W = \rho(\lambda)$, $S = \rho_0(\lambda)$, and $N = \eta(\lambda)$ and considering Fig. 6, the capacity of the CS channel can be derived based on the capacity of the equivalent AWGN channel in (22) as in (23). Assuming real signaling on CS, section III shows that the variance of the noise N is dependent to the amplitude of signal S with a polynomial of an order higher than two. Therefore, based on the theorem presented earlier, the real signal space of the transformed CS channel is limited not only by an average power constraint but also by a peak amplitude constraint.

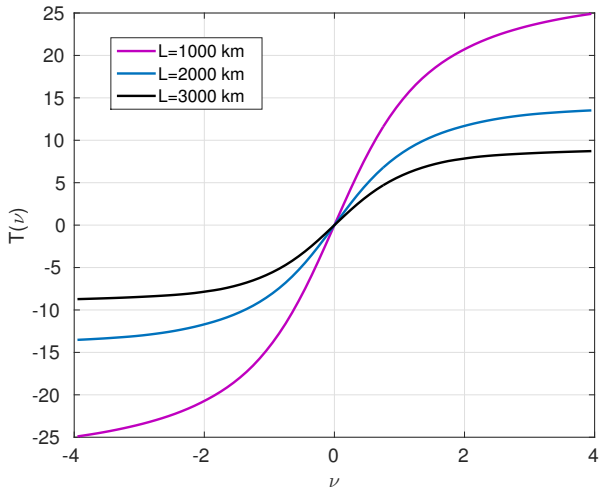

Fig. 7. Variance normalizing transform for different fiber lengths.

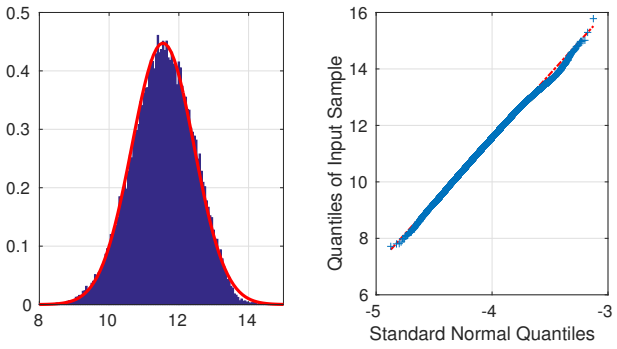

Fig. 8. (a) Probability density of transformed received signal when the noise-free amplitude is equal to 2. (b) Q-Q plot for the 2 samples.

Hence, (23) needs to be modified to determine the capacity C_R of the real CS channel by maximizing mutual information for the transformed AWGN channel with both peak amplitude and average power constraint as

$$C_R(P, A) = \sup_{\mathsf{E}|T^{-1}(X)|^2 \leq P, \quad |X| \leq A} I(X;Y). \qquad (26)$$

As proved in [50], the introduction of the peak power constraints in problem (26) leads to a capacity-achieving input distribution that is a set of discrete points. This optimal distribution and the resulting capacity can therefore be calculated numerically using the algorithm proposed in [50]. Since the signaling is performed in nonlinear spectral domain, the maximum average signal power in CS (i.e., P) is considered as the constraint. However, the actual power in time domain cannot be readily expressed as a function of P. The relationship between P and signal power in time domain will be discussed later in section V.

Fig. 9 shows the capacity of the real CS channel, C_R, for different link lengths calculated by solving the problem (26) using the algorithm provided in [50]. Note that average power constraint needs to be converted to a constraint in the

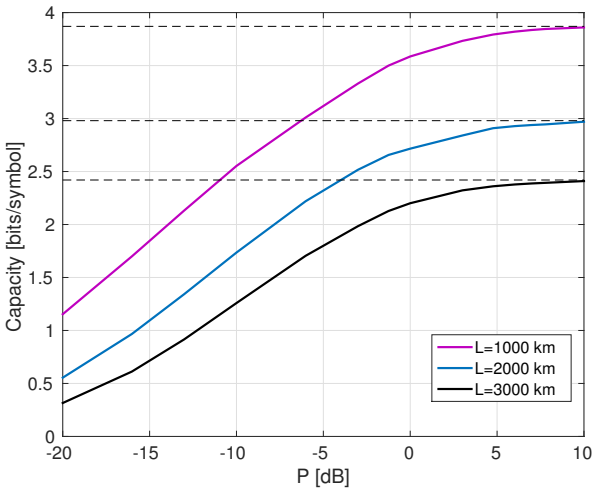

Fig. 9. Capacity for different fiber lengths.

transformed domain. This imposes a modification to the algorithm in [50]. The results clearly demonstrate that the capacity saturates for high signal power for different fiber lengths. Similar saturation effects have been previously observed for achievable data rates on nonlinear optical fiber channels [1]. The maximum capacity for each fiber length is equal to the capacity without an average power constraint or $P \to \infty$. These values are respectively 3.873, 2.981, and 2.425 bits per symbol for fiber lengths 1000 km, 2000 km and 3000 km. It should be noted that these values for the capacity are achieved for 50 GHz bandwidth of optical fiber, and higher capacities in bits per symbol are expected for larger bandwidth [35].

As shown in [50], for high values of A and P, the capacity-achieving distribution tends to a set of uniformly distributed discrete points for which the capacity can be derived in closed-form as

$$C_R = \log(2A/\sqrt{2\pi e}). \tag{27}$$

For instance, for the 1000 km fiber, the capacity achieving distribution for high average powers is approximately uniform and consequently the capacity tends to $\log(2A/\sqrt{2\pi e}) = 3.79$ which is close to the value (i.e., $C_R = 3.873$) numerically calculated in Fig. 9 at high signal power.

In section III, it was demonstrated that the noise on real and imaginary parts of CS ($\eta^r(\lambda)$ and $\eta^i(\lambda)$) are correlated, and increasing real or imaginary amplitude increases the noise variance in both channels. Since, the details of this correlation is not yet known, we use the capacity results of real singling on CS to present bounds on the capacity of the complex CS signaling in bits per complex symbol as

$$C_R(P, A) \leq C(P, A) \leq 2\ C_R(P/2, A). \tag{28}$$

where we have used the fact that the real and imaginary channels are identical as demonstrated in section III.

V. Discussions on achievable data rates

In the previous section, we derived an estimate on the capacity of CS channel in bits per symbol or, in other words, the capacity of the *discrete-λ channel*. It would be also beneficial to estimate achievable data rates by taking into account available time-frequency resources and addressing *continuous-λ channel* capacity. Due to nonlinear nature of the system, however, we cannot simply apply Nyquist theorem to relate these two quantities. Here, we briefly discuss how this relationship looks like for the nonlinear fiber channel considering several properties of the NFT-based system.

Note that the simulation results presented in this section are derived based on the same simulation method explained in section III. Here, the width of signal $\rho(\lambda)$ is called spectral width denoted by Λ. Unless otherwise stated, 64 random symbols are considered for CS with $\Lambda = 16$ and $T_0 = 0.1$ ns. Temporal width, bandwidth, and spectral width are calculated as the interval consisting 99% of energy.

TABLE I
Bandwidth for different CS powers

P_ρ [dB]	-10	-2	4	10	13.5	16
B	5.09	5.11	5.16	5.43	5.62	5.89

A. Bandwidth

Considering a low power regime, $F(x)$ can be replaced by $\epsilon F(x)$ ($\epsilon << 1$) as the Fourier transform of $\rho(\lambda)$ in GLM equation (8). As a result, it can be easily concluded that $q(t) = \text{INFT}\{\rho(\lambda)\} \simeq -2\epsilon F(2t)$. This shows that, for low signal energy, INFT behaves similar to ordinary inverse Fourier transform (IFT). Therefore, the bandwidth of signal after INFT is related to the spectral width as $B \simeq \Lambda/\pi$. For higher power, the INFT causes the bandwidth B to increase beyond Λ/π.

Table I shows simulation results for bandwidth for different values of signal power in nonlinear spectral domain (P_ρ). It can be seen that for the lowest power value, the measured bandwidth is very close to the actual value of $\Lambda/\pi = 5.09$ and for higher power values B increases. It should be noted that B is the normalized bandwidth, and the actual bandwidth of the signal would be $W = B/T_0$.

Moreover, the bandwidth of signal may increase during propagation along the fiber due to nonlinear spectral broadening [12]. Therefore, both INFT and fiber propagation result in increased signal bandwidth, and consequently, the nonlinear spectral width in CS available for information transmission is limited by the available bandwidth as $\Lambda \leq \pi B$.

B. Temporal width

Equation (12) shows that the width of signal in time domain at the receiver asymptotically equals to $\Delta t_\infty = 4z\Lambda$. This is based on the asymptotic scenario in which the dispersion effect is high and $T(z) >> T(0)$ where $T(z)$ is the temporal width of signal at length z. In non-asymptotic z, this approximation can be used as the added temporal width due to dispersion. Similarly, in linear fiber, the temporal width of signal can be estimated as $T(z) = T(0) + 2\pi\beta_2 LW$ [12]. Table II presents the simulation results of the signal temporal width after propagation over different lengths of

TABLE II
APPROXIMATION FOR TEMPORAL BROADENING

L [km]	z	$T(z)$ [ns]	$T(0)+4z\Lambda T_0$	$T(0)+2\pi\beta_2 LW$
1000	1.05	8.62	10.57	11.04
3000	3.15	22.12	23.98	25.36
6000	6.30	42.82	44.09	46.84

TABLE III
TIME DOMAIN POWER AND WIDTH FOR DIFFERENT CS POWERS

P_ρ [dB]	P_q [dBm]	$T(0)$ [ns]	r
-10	-12.12	1.26	1
-2	-5.99	1.47	1.17
4	-3.26	1.94	1.54
10	-1.73	2.59	2.06
13.5	-1.09	2.97	2.36
16	-0.76	3.26	2.56

fiber for $P_\rho = 10$ dB. It is observed that the approximation $T(z) = T(0) + 4z\Lambda T_0$ is slightly better than the linear approximation in every fiber length. Therefore,this estimate can be used to estimate the temporal width at the receiver for an NFT-based communication system.

Following the linear approximation argument in V-A, it can be concluded that the temporal width for low power signals can be obtained as $T_l(0) = T_0\pi K/\Lambda$ based on Nyquist theorem and denormalizing by T_0. Here, K is the number of symbols. However, when the signal power increases in nonlinear spectral domain, the temporal width after INFT increases beyond this linear estimate. Table III presents a comparison between the actual measured temporal width of signal after INFT and the linear estimate. Parameter $r = T(0)/T_l(0)$ is defined which demonstrates the excess temporal width due to INFT. It can be seen that r increases beyond 1 as the signal power increases. It is also observed in Table III that the power in spectral domain (P_ρ) and time domain (P_q) are not linearly related. For instance, P_ρ increases 6 dB for twice the amplitude, but this is not the case for P_q. As a matter of fact, the power in time domain increases slowly compared to the power in nonlinear spectral domain since the temporal width after INFT, $T(0)$, increases for higher CS power. The Parseval's equality for NFT [7] when considering only CS gives an insight on this effect as the square of temporal signal is related to the logarithm of the square of CS.

C. Achievable rates

In order to estimate the effective data rates, we need to take into account the dispersion effect and the corresponding guard interval required for signal transmission using NFT based systems. We thus introduce a new parameter W_d which determines the number of symbols received per second as

$$W_d = \frac{K}{T(0)+\Delta T} \quad [\text{symbol/sec}],$$

where ΔT is the increase in temporal width because of dispersion and K is number of symbols. Then, data rate in [bits/sec] can be estimated by $R = CW_d$, where C is the complex capacity bound derived in section IV as $C = 2C_R$ and the capacity of real signaling, C_R, is shown in Fig. 9. Parameter

TABLE IV
THE EFFECT OF NUMBER OF SYMBOLS ON THE DATA RATE. $T(0)$ [ns], $T(z)$ [ns], W_d [Gsymbol/sec], $\widetilde{W}_d$ [Gsymbol/sec] ,R [Gbps]

K	$T(0)$	$T(z)$	W_d	$\widetilde{W}_d$	$R = CW_d$
16	0.92	15.64	1.02	1.14	6.14
32	1.3	15.04	2.13	2.18	12.77
64	2.59	15.29	4.19	4.01	25.12
128	5.51	17.06	7.5	6.93	45.02
256	9.56	20.38	12.56	10.90	75.36

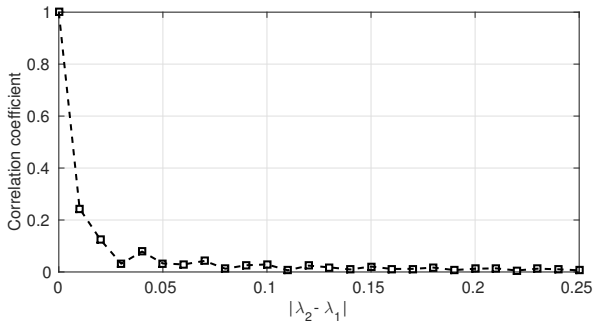

Fig. 10. Average correlation coefficient plotted for frequency differences.

W_d can be derived by simulation considering different number of random symbols in a fixed nonlinear frequency interval. For instance, in Table IV, W_d is calculated for a 2000 km fiber link and $P_\rho = 10$ dB. It is also possible to estimate W_d based on the approximations for bandwidth and temporal width derived in previous subsections. The temporal width at the transmitter after INFT is approximated by $T(0) = rT_0\pi K/\Lambda$. Therefore, we have

$$\widetilde{W}_d \simeq \frac{K}{T_0\left(r\frac{\pi K}{\Lambda} + 4z\Lambda\right)}, \tag{29}$$

where r is given by Table III, for instance, $r \simeq 2$ for $P_\rho = 10$ dB and $K = 64$. Table IV compares W_d and $\widetilde{W}_d$ and also presents achievable data rate, For instance, $R = 75.36$ Gbps is achieved for $K = 256$. Note that this data rate is achieved taking into account the needed guard interval due to dispersion. This table also shows that increasing K results in higher data rate.

Using (27-29), a closed-form expression can be written for achievable data rate based on complex signaling on CS as

$$R \simeq C\widetilde{W}_d = \frac{2K\log(2A/\sqrt{2\pi e})}{T_0\left(r\frac{\pi K}{\Lambda} + 4z\Lambda\right)},$$

which can be simplified for large K as

$$R \simeq \frac{2\Lambda\log(2A/\sqrt{2\pi e})}{rT_0\pi}. \tag{30}$$

Considering the low power regime, we have $r = 1$ and $W = \Lambda/(\pi T0)$, and therefore (30) would tend to $R = WC$ where C is in bits per complex symbol which can be originally predicted by Nyquist theorem.

D. Correlation of noise

In section IV-B, we assumed that the noise in CS is white over the signal spectrum. Here, simulation results are presented

to investigate the noise correlation. Fig. 10 shows the average of absolute value of correlation coefficients for noise samples with their neighboring samples for the fiber length of 2000 km. similar results have been observed for different fiber lengths. It can be seen that as the difference between the frequencies increases, the correlation decreases substantially. For instance, if the distance between two data samples is 0.03 (corresponding to approximately 95 MHz), the noise on the samples can be regarded as uncorrelated. Since the results presented here are based on simulations, part of the observed correlation can be caused by numerical errors in NFT and INFT. Since $\Lambda = 16$, the maximum number of symbols without correlation can be estimated as $K \simeq 512$ which lead to $R \simeq 92$ Gbps based on the capacity and achievable rate results.

VI. Conclusion

In this paper, the noise characteristics of the channel for CS in NFT-based optical fiber communication systems was studied. The channel characteristics were derived analytically and confirmed by simulation. The noise for CS channel was observed to be dependent on the signal amplitude by a polynomial with an order larger than two. Applying a variance normalizing transform, it is observed that the signal dependency of noise leads to peak-power-limited signal space. As a consequence, the CS channel capacity is saturated for high powers. However, it was demonstrated that remarkable data rates can be achieved by only signaling on CS.

References

[1] R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Lightwave Technology, Journal of*, vol. 28, no. 4, pp. 662–701, 2010.

[2] E. Agrell, A. Alvarado, and F. R. Kschischang, "Implications of information theory in optical fibre communications," *Phil. Trans. R. Soc. A*, vol. 374, no. 2062, p. 20140438, 2016.

[3] P. Bayvel, R. Maher, T. Xu, G. Liga, N. A. Shevchenko, D. Lavery, A. Alvarado, and R. I. Killey, "Maximizing the optical network capacity," *Phil. Trans. R. Soc. A*, vol. 374, no. 2062, p. 20140440, 2016.

[4] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *Information Theory Workshop (ITW), 2015 IEEE*, April 2015, pp. 1–4.

[5] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, Oct 2008.

[6] S. Watanabe and M. Shirasaki, "Exact compensation for both chromatic dispersion and kerr effect in a transmission fiber using optical phase conjugation," *Lightwave Technology, Journal of*, vol. 14, no. 3, pp. 243–248, 1996.

[7] M. Yousefi and F. Kschischang, "Information transmission using the nonlinear fourier transform, part i: Mathematical tools," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4328, July 2014.

[8] ——, "Information transmission using the nonlinear fourier transform, part ii: Numerical methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, July 2014.

[9] ——, "Information transmission using the nonlinear fourier transform, part iii: Spectrum modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, July 2014.

[10] M. J. Ablowitz, B. Prinari, and A. D. Trubatch, *Discrete and continuous nonlinear Schrödinger systems*. Cambridge University Press, 2004.

[11] A. Hasegawa and T. Nyu, "Eigenvalue communication," *Lightwave Technology, Journal of*, vol. 11, no. 3, pp. 395–399, 1993.

[12] G. P. Agrawal, *Nonlinear fiber optics*, 5th ed. Academic Press, 2013.

[13] J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, "Nonlinear spectral management: Linearization of the lossless fiber channel," *Optics express*, vol. 21, no. 20, pp. 24 344–24 367, 2013.

[14] J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," *Phys. Rev. Lett.*, vol. 113, p. 013901, Jul 2014. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.113.013901

[15] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," *Optics express*, vol. 22, no. 22, pp. 26 720–26 741, 2014.

[16] ——, "Nonlinear inverse synthesis technique for optical links with lumped amplification," *Optics express*, vol. 23, no. 7, pp. 8317–8328, 2015.

[17] S. T. Le, J. E. Prilepsky, P. Rosa, J. D. Ania-Castan, and S. K. Turitsyn, "Nonlinear inverse synthesis for optical links with distributed raman amplification," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1778–1786, April 2016.

[18] S. T. Le, I. D. Philips, J. E. Prilepsky, P. Harper, A. D. Ellis, and S. K. Turitsyn, "Demonstration of nonlinear inverse synthesis transmission over transoceanic distances," *Journal of Lightwave Technology*, vol. 34, no. 10, pp. 2459–2466, May 2016.

[19] M. Kamalian, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn, "Periodic nonlinear fourier transform for fiber-optic communications, part i: theory and numerical methods," *Opt. Express*, vol. 24, no. 16, pp. 18 353–18 369, Aug 2016.

[20] ——, "Periodic nonlinear fourier transform for fiber-optic communications, part ii: eigenvalue communication," *Opt. Express*, vol. 24, no. 16, pp. 18 370–18 381, Aug 2016.

[21] H. Bülow, "Experimental demonstration of optical signal detection using nonlinear fourier transform," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1433–1439, 2015.

[22] V. Aref, H. Bulow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in *Optical Communication (ECOC), 2015 European Conference on*, Sept 2015, pp. 1–3.

[23] H. Bulow, V. Aref, K. Schuh, and W. Idler, "Practical implementation of nonlinear fourier transform based optical nonlinearity mitigation," in *Optical Communication (ECOC), 2015 European Conference on*, Sept 2015, pp. 1–3.

[24] Z. Dong, S. Hari, T. Gui, K. Zhong, M. I. Yousefi, C. Lu, P.-K. A. Wai, F. R. Kschischang, and A. P. T. Lau, "Nonlinear frequency division multiplexed transmissions based on nft," *IEEE Photonics Technology Letters*, vol. 27, no. 15, pp. 1621–1623, 2015.

[25] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue communication," *Journal of Lightwave Technology*, vol. 34, no. 13, pp. 3110–3117, July 2016.

[26] S. Hari and F. R. Kschischang, "Bi-directional algorithm for computing discrete spectral amplitudes in the nft," *Journal of Lightwave Technology*, vol. 34, no. 15, pp. 3529–3537, Aug 2016.

[27] S. Wahls and H. V. Poor, "Fast numerical nonlinear fourier transforms," *Information Theory, IEEE Transactions on*, vol. 61, no. 12, pp. 6957–6974, 2015.

[28] S. Wahls and V. Vaibhav, "Fast inverse nonlinear fourier transforms for fiber bragg grating design and related problems," *arXiv preprint arXiv:1607.01305*, 2016.

[29] S. Wahls, S. T. Le, J. E. Prilepsk, H. V. Poor, and S. K. Turitsyn, "Digital backpropagation in the nonlinear fourier domain," in *Signal Processing Advances in Wireless Communications (SPAWC), 2015 IEEE 16th International Workshop on*. IEEE, 2015, pp. 445–449.

[30] M. I. Yousefi and X. Yangzhang, "Nonlinear frequency-division multiplexing," *arXiv preprint arXiv:1603.04389*, 2016.

[31] Q. Zhang and T. H. Chan, "Achievable rates of soliton communication systems," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 605–609.

[32] N. A. Shevchenko, S. A. Derevyanko, J. E. Prilepsky, A. Alvarado, P. Bayvel, and S. K. Turitsyn, "A lower bound on the capacity of the noncentral chi channel with applications to soliton amplitude modulation," *arXiv preprint arXiv:1609.02318*, 2016.

[33] I. Tavakkolnia and M. Safari, "Dispersion pre-compensation for nft-based optical fiber communication systems," in *Conference on Lasers and Electro-Optics*. Optical Society of America, 2016, p. SM4F4.

[34] ——, "Signalling over nonlinear fibre-optic channels by utilizing both solitonic and radiative spectra," in *Networks and Communications (EuCNC), 2015 European Conference on*, June 2015, pp. 103–107.

[35] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, "Capacity estimates for optical transmission based on the nonlinear fourier transform," *Nature Communications*, vol. 7, p. 12710, 2016.

[36] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," *Communications on pure and applied mathematics*, vol. 21, no. 5, pp. 467–490, 1968.
[37] A. Shabat and V. Zakharov, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Soviet Physics JETP*, vol. 34, no. 1, p. 62, 1972.
[38] M. J. Ablowitz, D. J. Kaup, and A. C. Newell, "The inverse scattering transform-fourier analysis for nonlinear problems," *Studies in Applied Mathematics*, vol. 53, no. 4, pp. 249–315, 1974.
[39] S. Wahls and H. V. Poor, "Fast inverse nonlinear fourier transform for generating multi-solitons in optical fiber," in *Information Theory (ISIT), 2015 IEEE International Symposium on*, June 2015, pp. 1676–1680.
[40] M. J. Ablowitz and H. Segur, *Solitons and the inverse scattering transform*. SIAM, 1981, vol. 4.
[41] H. Segur and M. J. Ablowitz, "Asymptotic solutions and conservation laws for the nonlinear schrödinger equation. i," *Journal of Mathematical Physics*, vol. 17, no. 5, pp. 710–713, 1976.
[42] V. Zakharov and S. Manakov, "Asymptotic behavior of non-linear wave systems integrated by the inverse scattering method," *Soviet Journal of Experimental and Theoretical Physics*, vol. 44, p. 106, 1976.
[43] M. Bartlett, "The square root transformation in analysis of variance," *Supplement to the Journal of the Royal Statistical Society*, pp. 68–78, 1936.
[44] J. H. Curtiss, "On transformations used in the analysis of variance," *The Annals of Mathematical Statistics*, vol. 14, no. 2, pp. 107–122, 1943.
[45] P. R. Prucnal and B. E. Saleh, "Transformation of image-signal-dependent noise into image-signal-independent noise," *Optics letters*, vol. 6, no. 7, pp. 316–318, 1981.
[46] M. Safari, "Efficient optical wireless communication in the presence of signal-dependent noise," in *Communication Workshop (ICCW), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1387–1391.
[47] A. Tsiatmas, F. M. Willems, and C. P. Baggen, "Square root approximation to the poisson channel," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 1695–1699.
[48] I. Tavakkolnia and M. Safari, "Variance normalizing transform for performance improvement in radio-over-fiber systems," in *Networks and Communications (EuCNC), 2016 European Conference on*, June 2016.
[49] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[50] J. G. Smith, "The information capacity of amplitude-and variance-constrained sclar gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203–219, 1971.

Iman Tavakkolnia (S'15) received his B.Sc. degree in Telecommunications from University of Tehran, Iran, in 2011, and his M.Sc. degree in Communication Systems from Sharif University of Technology, Iran, in 2013. He is currently working towards his Ph.D. degree in Digital Communications at the University of Edinburgh, United Kingdom. His research interests include fibre-optic communications, optical wireless communication, and digital signal processing.

Majid Safari (S08-M11) received B.Sc. in Electrical and Computer Engineering from the University of Tehran, Iran, in 2003, M.Sc. in Electrical Engineering from Sharif University of Technology, Iran, in 2005 and Ph.D. in Electrical and Computer Engineering from the University of Waterloo, Canada in 2011.

He is currently an assistant professor in the Institute for Digital Communications at the University of Edinburgh. Before joining Edinburgh in 2013, He held postdoctoral fellowship at McMaster University, Canada. He is the recipient of a number of research and academic awards including best student paper award in 21st Canadian Conference on Electrical and Computer Engineering in 2008, the University of Waterloo doctoral thesis completion award in 2010 and MITACS elevate strategic fellowship in 2012. Dr. Safari is currently an associate editor of IEEE Communication letters and was the TPC co-chair of the 4th International Workshop on Optical Wireless Communication in 2015. His research interests include fiber-optic communications, free-space optical communications, visible light communications, and signal processing for optical communication systems.