On the Capacity of the Noncentral Chi-Channel with Applications to Soliton Amplitude Modulation

Nikita A. Shevchenko, Stanislav A. Derevyanko, Jaroslaw E. Prilepsky, Alex Alvarado, Polina Bayvel, and Sergei K. Turitsyn

Abstract—The channel law for amplitude-modulated solitons transmitted through a nonlinear optical fibre with ideal distributed amplification and a receiver based on the nonlinear Fourier transform is a noncentral chi distribution with $2n$ degrees of freedom, where $n = 2$ and $n = 3$ correspond to the single- and dual-polarisation cases, respectively. In this paper, we study the capacity of this channel under an average power constraint in bits per channel use. We develop an asymptotic semi-analytic approximation for a capacity lower bound for arbitrary $n$ and a Rayleigh input distribution. It is shown that this lower bound grows logarithmically with signal-to-noise ratio (SNR), independently of the value of $n$. Numerical results for other continuous input distributions are also provided. A half-Gaussian input distribution is shown to give larger rates than a Rayleigh input distribution for $n = 1, 2, 3$. At an SNR of 25 dB, the best lower bounds we developed are approximately 3.68 bit per channel use. The practically relevant case of amplitude shift-keying (ASK) constellations is also numerically analysed. For the same SNR of 25 dB, a 16-ASK constellation yields a rate of approximately 3.45 bit per channel use.

Index Terms—Achievable information rates, channel capacity, mutual information, nonlinear optical fibres, nonlinear Fourier transform, optical solitons.

I. INTRODUCTION

Optical fibre transmission systems carrying the overwhelming bulk of the world’s telecommunication traffic have undergone a long process of increasing engineering complexity and sophistication [1]–[3]. However, the key physical effects affecting the performance of these systems remain largely the same. These are: chromatic dispersion, fibre nonlinearity due to the optical Kerr effect, and optical noise. Although the bandwidth of optical fibre transmission systems is large, these systems are ultimately band-limited. This bandwidth limitation combined with the ever-growing demand for data rates is expected to result in a so-called "capacity crunch" [4], which caps the rate increase of error-free data transmission

Up until recently, the common belief among some researchers in the field of optical communication was that nonlinearity was always a nuisance that necessarily degrades the system performance. This led to the assumption that the capacity of the optical channel had a peaky behaviour when plotted as a function of the transmit power [5]. Partially motivated by the idea of improving the data rates in optical fibre links, a multitude of nonlinearity compensation methods have been proposed (see, e.g., [15]–[20]), each resulting in different discrete-time channel models. Recently, a paradigm-shifting approach for overcoming the effects of nonlinearity has been receiving increased attention. This approach relies on the fact that both the ME and NSE in the absence of losses and noise are exactly integrable [21], [22].

One of the consequences of integrability is that the signal evolution can be represented using nonlinear normal modes. While the pulse propagation in the ME and NSE is nonlinear, the evolution of these nonlinear modes in the so-called nonlinear spectral domain is essentially linear [23], [24]. The decomposition of the waveform into the nonlinear modes (and the reciprocal operation) is often referred to as nonlinear Fourier transform (NFT), due to its similarity with the application of the conventional Fourier decomposition in

The precise mathematical expressions for both channel models are given in Sec. II-A. However, nondecaying lower bounds can be found in the literature, e.g., in [10], [12]–[14].
linear systems \cite{25}. The linear propagation of the nonlinear modes implies that the nonlinear cross-talk in the NFT domain is theoretically absent, an idea exploited in the so-called nonlinear frequency division multiplexing \cite{23}, \cite{26}. In this method, the nonlinear interference can be greatly suppressed by assigning users different ranges in the nonlinear spectrum, instead of multiplexing them using the conventional Fourier domain.

Integrability also leads to several nonlinearity compensation and transmission schemes \cite{27}–\cite{37}. These can be seen as a generalisation of soliton-based communications \cite{8}, \cite{9}, \cite{38} Chapter 5, which follow the pioneering work by Hasegawa and Nuy \cite{39}, and where only the discrete eigenvalues were used for communication. The development of efficient and numerically stable algorithms has also attracted a lot of attention \cite{40}. Furthermore, there have also been a number of experimental demonstrations and assessments for different NFT-based systems \cite{32}–\cite{37}. However, for systems governed by the ME, the only results available come from the recent theoretical work of Maruta and Matsuda \cite{31}.

Two nonlinear spectra (types of nonlinear modes) exist in the NSE and the ME. The first one is the so-called continuous spectrum, which is the exact nonlinear analogue of the familiar linear FT \cite{30}, corresponding to the dispersive nonlinear modes. The unique feature of the NFT is, however, that apart from the continuous spectrum, it can support a set of nonlinear modes. The unique feature of the NFT is, however, of the familiar linear FT \cite{30}, corresponding to the dispersive continuous spectrum, which is the exact nonlinear analogue of the NFT methods have been recently investigated in \cite{52}. Integrability has also lead to several nonlinearity compensation and transmission schemes \cite{27}–\cite{37}. These can be seen as a generalisation of soliton-based communications \cite{8}, \cite{9}, \cite{38} Chapter 5, which follow the pioneering work by Hasegawa and Nuy \cite{39}, and where only the discrete eigenvalues were used for communication. The development of efficient and numerically stable algorithms has also attracted a lot of attention \cite{40}. Furthermore, there have also been a number of experimental demonstrations and assessments for different NFT-based systems \cite{32}–\cite{37}. However, for systems governed by the ME, the only results available come from the recent theoretical work of Maruta and Matsuda \cite{31}.

In mathematics and physics literature, the name inverse scattering method for the NFT is more commonly used.

3Since the imaginary part of a single discrete eigenvalue is proportional to the soliton amplitude.

4In mathematics and physics literature, the name inverse scattering method for the NFT is more commonly used.

where data rate gains of the continuous soliton modulation versus an on-off-keying (OOK) system were also shown. Yousefi and Kschischang \cite{28} addressed the question of achievable spectral efficiency for single- and multi-eigenvalue transmission systems using a Gaussian model for the nonlinear spectrum evolution. Some results on the continuous spectrum modulation were also presented. Later in \cite{41}, the spectral efficiency of a multi-eigenvalue transmission system was studied in more detail. In \cite{42}, the same problem was studied by considering the correlation functions of the spectral data obtained in the quasi-classical limit of large number of eigenvalues. Achievable information rates for multi-eigenvalue transmission systems utilising all four degrees of freedom of each scalar soliton in NSE were analytically obtained in \cite{33}. These results were obtained within the framework of a Gaussian noise model provided in \cite{28}, \cite{44} and assuming a continuous uniform input distribution subject to peak power constraints. The spectral efficiency for the NFT continuous spectrum modulation was considered in \cite{49}–\cite{51}. Periodic NFT methods have been recently investigated in \cite{52}.

In \cite{46}, we used a non-Gaussian model for the evolution of a single soliton amplitude and the NSE. Our results showed that a lower bound for the capacity per channel use of such a model grows unbounded with the effective signal-to-noise ratio (SNR). In this paper, we generalise and extend our results in \cite{46} to the ME. To this end, we use the exact channel laws for soliton amplitudes previously reported in \cite{47}, \cite{48} (for the NSE) and \cite{53} (for the ME). Both channel laws are a noncentral chi distribution with degrees of freedom, where \(n = 2\) and \(n = 3\) correspond to the NSE and ME, respectively. Motivated by the similarity of the channel models mentioned above, in this paper we study asymptotic lower bound approximations on the capacity (in bit per channel use) of a general noncentral \(\chi\) channel arbitrary (even) number of degrees of freedom. To the best of our knowledge, this has not been previously reported in the literature.

The first contribution of this paper is to numerically obtain lower bounds for the channel capacity for three continuous input distributions, as well as for amplitude shift-keying (ASK) constellations with discrete number of constellation points. For all the continuous inputs, the lower bounds are shown to be nondecreasing functions of the SNR under an average power constraint. The second contribution of this paper is to provide an asymptotic closed-form expression for the MI of the noncentral \(\chi\) channel with an arbitrary (even) number of degrees of freedom. This asymptotic expression shows that the MI grows unbounded and at the same rate, independently of the number of degrees of freedom.

II. CONTINUOUS-TIME CHANNEL MODEL

A. The Propagation Equations

The propagation of light in optical fibres in the presence of amplified spontaneous emission (ASE) noise can be described by a stochastic partial differential equation which captures the effects of chromatic dispersion, nonlinear polarisation mode dispersion, optical Kerr effect, and the generation of ASE noise from the optical amplification process. Throughout the propagation of light in optical fibres in the presence of amplified spontaneous emission (ASE) noise can be described by a stochastic partial differential equation which captures the effects of chromatic dispersion, nonlinear polarisation mode dispersion, optical Kerr effect, and the generation of ASE noise from the optical amplification process. Throughout the propagation of light in optical fibres in the presence of amplified spontaneous emission (ASE) noise can be described by a stochastic partial differential equation which captures the effects of chromatic dispersion, nonlinear polarisation mode dispersion, optical Kerr effect, and the generation of ASE noise from the optical amplification process. Throughout the propagation of light in optical fibres in the presence of amplified spontaneous emission (ASE) noise can be described by a stochastic partial differential equation which captures the effects of chromatic dispersion, nonlinear polarisation mode dispersion, optical Kerr effect, and the generation of ASE noise from the optical amplification process. Throughout
this paper we assume that the fibre loss is continuously compensated along the fibre by means of (ideal) distributed Raman amplification (DRA) [55], [56]. In this work we consider the propagation of a slowly varying 2-component envelope \( E(\ell, \tau) = [E_1(\ell, \tau), E_2(\ell, \tau)] \in \mathbb{C}^2 \) over a nonlinear birefringent optical fibre, where \( \tau \) and \( \ell \) represent time and propagation distance, respectively. Our model also includes the 2-component ASE noise \( N(\ell, \tau) = [N_1(\ell, \tau), N_2(\ell, \tau)] \) due to the DRA. We also assume a uniform change of polarised state on the Poincaré sphere [57] and the guiding centre approximation [8].

The resultless lossless ME is then given by [7] eq. (1.26), [8] Sec. 10.3.1], [53], [58]
\[
i E_\ell + \frac{\beta_2}{2} E_{\tau \tau} + \frac{8 \gamma}{9} |E|^2 E = N(\ell, \tau),
\]
where the retarded time \( \tau \) is measured in the reference frame moving with the optical pulse average group velocity, \( E \equiv E(\ell, \tau) \) represents the slowly varying 2-component envelope of electric field, \( \beta_2 \) is the group velocity dispersion coefficient characterising the chromatic dispersion, and \( \gamma \) is the fibre nonlinearity coefficient. The pre-factor \( 8/9 \) in [1] comes from the averaging of the fast polarisation rotation [8] Sec. 10.3.1], [57]. For simplicity we will further work with the effective averaged nonlinear coefficient \( \gamma^* \equiv \frac{8 \gamma}{9} \) when addressing the ME. In the case of a single polarisation state, the propagation equation above reduces to the resultless generalised scalar NSE [6], [9]
\[
i E_\ell + \frac{\beta_2}{2} E_{\tau \tau} + \gamma |E|^2 E = N(\ell, \tau).
\]

In this paper we consider the case of anomalous dispersion (\( \beta_2 < 0 \)), i.e., the focusing case. In this case, both the ME in [1] and the NSE in [2] permit bright solitons solutions ("particle-like waves"), which will be discussed in more detail in Sec. [10].

It is customary to re-scale [1] to dimensionless units. We shall use the following normalisation: The power will be measured in units of \( P_0 = 1 \text{ mW} \) since it is a typical power level used in optical communications. The normalised (dimensionless) field then becomes \( q = E/\sqrt{P_0} \). For the distance and time, we define the dimensionless variables \( z \) and \( t \) as \( z = \ell/\ell_0 \) and \( t = \tau/\tau_0 \), where
\[
\ell_0 = (\gamma^* P_0)^{-1}, \quad \tau_0 = \sqrt{|\ell_0|/\beta_2} = \sqrt{|\beta_2|/\gamma^* P_0}.
\]
For the scalar case [2], we use the same normalisation but we replace \( \gamma^* \) by \( \gamma \). Then, the resulting ME reads
\[
i q_z + \frac{1}{2} q_{tt} + (q, q) q = n(z, t),
\]
while the NSE becomes
\[
i q_z + \frac{1}{2} q_{tt} + |q|^2 q = n(z, t).
\]

The ASE noise \( n(z, t) = [n_1(z, t), n_2(z, t)] \) in [4] is assumed to have the following correlation properties
\[
E[n_1(z, t)] = E[n_2(z, t)] n_j(z', t') = 0,
E[n_1(z, t) n_j(z', t')] = D \delta_{ij} \delta (z - z') \delta (t - t'),
\]
with \( i, j \in \{1, 2\} \), with \( \delta_{ij} \) being a Kronecker symbol, \( E[.] \) is the mathematical expectation operator, and \( \delta (\cdot) \) is the Dirac delta function. The correlation properties [6] mean that each noise component \( n_i(z, t) \) is assumed to be a zero-mean, independent, white circular Gaussian noise. The scalar case follows by considering a single noise component only.

The noise intensity \( D \) in [6] is (in dimensionless units)
\[
D = \alpha^2_0 \frac{\ell_0}{P_0 \tau_0} = \frac{\alpha^2_0}{\gamma^* |\beta_2| P_0},
\]
where \( \alpha^2_0 \) is the power spectral density (PSD) of the noise, with real world units \([W/\text{km} \cdot \text{Hz}]\). For ideal DRA, this PSD can be expressed through the optical fibre and transmission system parameters as follows: \( \alpha^2_0 = \ell_{\text{fibre}} K_T \cdot h_0 \), where \( \ell_{\text{fibre}} \) is the fibre attenuation coefficient, \( h_0 \) is the average photon energy, \( K_T \) is a temperature-dependent phonon occupancy factor [6].

From now on, all the quantities in this paper are in normalised units unless specified otherwise. Furthermore, we define the continuous time channel as the one defined by the normalised the ME and the NSE. This is shown schematically in the inner part of Fig. [1] where the transmitted and received waveforms are \( x(t) \equiv q(0, t) \) and \( y(t) \equiv q(Z, t) \), respectively, where \( Z \) is the propagation distance.

\section*{B. Fundamental Soliton Solutions}

It is known that the noiseless \( (n(z, t) = 0) \) ME [4] possesses a special class of solutions, the so-called fundamental bright solitons [6]. In general, the Manakov fundamental soliton is fully characterised by 6 parameters [53] (4 in the NSE case): frequency (also having the meaning of velocity in some physical applications), phase, phase mismatch, centre-of-mass position, polarisation angle, and amplitude (the latter is inversely proportional to the width of the soliton). In this paper we consider amplitude-modulated solitons, and thus, without loss generality, the soliton frequency, phase, phase mismatch, and centre-of-mass position, are all set to zero. Under these assumptions, the soliton solution at \( z = 0 \) is given by [53], [58]
\[
q(0, t) = [q_1(0, t), q_2(0, t)] = [\cos \beta_0, \sin \beta_0] A \text{sech}(\beta_0 t),
\]
where \( A \) is the soliton amplitude and \( 0 < \beta_0 < \pi/2 \) is the polarisation angle. The value of \( \beta_0 \) can be used to control how the signal power is split across the two polarisations.

\footnote{Fundamental solitons are "bright" only for the focusing case we consider in this paper, i.e., for anomalous dispersion.}
For any $\beta_0$, the Manakov soliton solution after propagation over a distance $Z$ with the initial condition given by (8), is expressed as

$$q(Z, t) = [\cos \beta_0, \sin \beta_0] A \text{sech}(AT) \exp \left(\frac{iA^2 Z}{2} \right)$$

(9)

$$= q(0, t) \exp \left(\frac{iA^2 Z}{2} \right).$$

(10)

The soliton solution for the NSE in (3) can be obtained by using $\beta_0 = 0$ in (8)–(10) which gives

$$q(0, t) = A \text{sech} (AT),$$

(11)

and

$$q(Z, t) = A \text{sech} (AT) \exp \left(\frac{iA^2 Z}{2} \right) = q(0, t) \exp \left(\frac{iA^2 Z}{2} \right).$$

(12)

As shown by (10) and (12), the solitons in (8) and (11) only acquire a phase rotation after propagation. When the noise is not zero, however, these solutions will change. This will be discussed in detail in the following section.

III. DISCRETE-TIME CHANNEL MODEL

A. Amplitude-modulated Solitons: One and Two Polarisations

We consider a continuous-time input signal $x(t) = [x_1(t), x_2(t)]$ of the form

$$x(t) = \sum_{k=1}^{\infty} s_k(t),$$

(13)

where $s_k(t) = [s_{k,1}(t), s_{k,2}(t)]$ and $k$ is the discrete-time index. Motivated by the results in Sec. II-B the pulses $s_k(t)$ are chosen to be

$$s_k(t) = [\cos \beta_0, \sin \beta_0] A_k \text{sech}[A_k(t - kT_s)],$$

(14)

where $T_s$ is the symbol period. In principle, it is also possible to encode information by changing the polarisation angle $\beta_0$ from slot to slot. However, in this paper, we fix its value to be the same for all the time slots corresponding to a fixed (generally elliptic) degree of polarisation. Thus, the transmitted waveform corresponds to soliton amplitude modulation, which is schematically shown in Fig. 2 for the scalar (NSE) case.

At the transmitter, we assume that symbols $X_k$ are mapped to soliton amplitudes $A_k$ via $A_k = X_k^{\frac{1}{2}}$. This normalisation is introduced only to simplify the analytical derivations in this paper. To avoid soliton-to-soliton interactions, we also assume that the separation $T_s$ is large, i.e., $\exp(-A_k T_s) \ll 1, \forall k$.

The receiver in Fig. 1 is assumed to process the received waveform during a window of $T_s$ via the forward NFT [21], [31] and returns the amplitude of the received soliton, which we denoted by $R_k = Y_k^2$.

Having defined the transmitter and receiver, we can now define a discrete-time channel model, which encompasses the transmitter, the optical fibre, and the receiver, as shown in Fig. 1. Due to the assumption on solitons well-separated in time, we model the channel as memoryless, and thus, from now on we drop the time index $k$. This memoryless assumption is supported by additional numerical simulations we performed, which are included in Appendix A. Note that even in cases where the memoryless assumption would not hold, considering a memoryless channel model would still result in a lower bound on the channel capacity, which is the main focus of this paper. This capacity lower bound will hold for the channel formed by the transmitter, the waveform channel, and a memoryless receiver (such as the one we consider in Appendix A). This follows from the fact that the receiver ignores the memory in the channel, and thus, is an instance of a mismatched decoder. This can be rigorously proven using mismatched decoding theory, as explained in [65, Sec. II-D] and references therein.

The conditional probability density function (PDF) for the received soliton amplitude $R$ given the transmitted amplitude $A$ was obtained in [53, eq. (15)] using standard perturbative approach and the Fokker-Planck equation method. The result can be expressed as

$$p_{R|A}(r|a) = \frac{1}{\sigma_N^2} \frac{r}{a} \exp \left(\frac{-a + r}{\sigma_N^2} \right) I_2 \left(\frac{2\sqrt{a\sigma}}{\sigma_N^2} \right),$$

(15)

where

$$\sigma_N^2 = D \cdot \frac{Z}{2}$$

(16)

is the normalised variance of accumulated ASE noise, and $I_2(\cdot)$ is the modified Bessel function of the first kind of order two. The expression in (15) is a noncentral chi-squared distribution with six degrees of freedom (see, e.g., [59], eq. (29.4)) providing non-Gaussian statistics for Manakov soliton amplitudes. By making the change of variables $Y = \sqrt{R}$, and using $X = \sqrt{A}$, the PDF in (15) can be expressed as

$$p_{Y|X}(y|x) = \frac{2}{\sigma_N^2} \frac{y^2}{x^3} \exp \left(\frac{-x^2 + y^2}{\sigma_N^2} \right) I_2 \left(\frac{2xy}{\sigma_N^2} \right),$$

(17)

which corresponds to the noncentral $\chi$ distribution with six degrees of freedom.

This corresponds to the case where all the signal power is transmitted in the first polarisation.
B. Generalised Discrete-time Channel Model

The results in the previous section show that both scalar and vector soliton channels can be modelled using the same class of the noncentral \( \chi \) distribution with an even number of degrees of freedom \( 2n \), with \( n = 2,3 \). The simplest channel of this type corresponds to \( n = 1 \), which describes a fibre optical communication channel with zero-dispersion [12] as well as the noncoherent phase channel studied in [60]. Motivated by this, here we consider a general communication channel described by the noncentral \( \chi \) distribution with an arbitrary (even) degrees of freedom \( 2n \). Although we are currently not aware of any physically-relevant communication system that can be modelled with \( n \geq 4 \), we present results for arbitrary \( n \) to provide an exhaustive treatment for channels of this type.

The channel in question is therefore modelled via the PDF

\[
p_{Y|X}(y|x) = \frac{2}{\sigma_N^2} \frac{y^n}{x^{n-1}} \exp\left(-\frac{x^2 + y^2}{\sigma_N^2}\right) I_{n-1} \left(\frac{2xy}{\sigma_N^2}\right),
\]

with \( n \in \mathbb{N} \) and where \( \mathbb{N} \triangleq \{1,2,3,\ldots\} \). This channel law corresponds to the following input-output relation

\[
Y^2 = \frac{1}{2} \sum_{i=1}^{2n} \left(\frac{X_i}{\sqrt{n}} + N_i\right)^2,
\]

where \( \{N_i\}_{i=1}^{2n} \) is a set of independent and identically distributed Gaussian random variables with zero mean and variance \( \sigma_N^2 \). The above input-output relationship is schematically shown in Fig. 3, which particularizes to (17) and (18), for \( n = 3 \) and \( n = 2 \), respectively.

IV. MAIN RESULTS

In this section, we study the capacity of the channel in (19). We will show results as a function of the effective SNR defined as \( \rho \triangleq \sigma_0^2/\sigma_N^2 \), where \( \sigma_0^2 \) is the second moment of the input distribution \( p_X \) and \( \sigma_N^2 \) is given by (16). The value of \( \sigma_0^2 \) also corresponds to the average soliton amplitude, i.e., \( \sigma^2_0 = \mathbb{E}[X^2] = \mathbb{E}[A] \). It can be shown that for a given system parameters, the noise power (in real world units) is constant and proportional to \( \sigma_0^2 \), and the signal power (in real world units) is proportional to \( \sigma_N^2 \). The parameter \( \rho \) therefore indeed corresponds to an effective SNR.

As previously explained, the inter-symbol interference due to pulse interaction can be neglected due to the large enough soliton separation assumed, and thus, the channel can be treated as a memoryless (see Appendix A for more details). The channel capacity, in bits per channel use, is then given by

\[
C(\rho) \triangleq \max_{p_X(x): \mathbb{E}[X^2] \leq \rho} I_{X,Y}(\rho),
\]

where

\[
I_{X,Y}(\rho) \triangleq \mathbb{E} \left[ \log_2 \frac{p_X(X,Y)}{p_X(X) \cdot p_Y(Y)} \right] = h_Y(\rho) - h_{X|Y}(\rho),
\]

and where \( h_Y(\rho) \triangleq -\mathbb{E} \left[ \log_2 p_Y(Y) \right] \) and \( h_{X|Y}(\rho) \triangleq -\mathbb{E} \left[ \log_2 p_{X|Y}(Y|X) \right] \) are the output and conditional differential entropies, respectively. The optimisation in (21) is performed over all possible statistical distributions \( p_X(x) \) that satisfy the power constraint. In our case this constraint corresponds to a fixed second moment of the input symbol...
distribution or, equivalently, to a fixed average \( \chi \) channel in a given symbol period.

The exact solution for the power-constrained optimisation problem \([21]\) with the channel law \([19]\) is unknown. For the noncentral \( \chi \) distribution with 2 degrees of freedom (i.e., to the noncoherent additive noise channel), it was shown \([60]\) that the capacity achieving distribution is discrete with an infinite number of mass points. To the best of our knowledge, that problem has not been extended to higher number of degrees of freedom, however, we expect that will be the case for \([19]\) too.

In this paper, we do not aim at finding the capacity-achieving distribution, but instead, we study lower bounds on the capacity. We do this because the capacity problem is in general very difficult, but also because of the relevance of having nondecreasing lower bounds on the capacity for the optical community. To obtain a lower bound on the capacity, we will simply choose an input distribution \( p_X(x) \) (as done in, e.g., \([5]\), \([46]\)). Without claiming the generality, we, however, consider four important candidates for the input distribution. First, following \([46]\), we use symbols drawn from a Rayleigh distribution

\[
p_X(x) = \frac{x}{\sigma_S^2} \exp \left(-\frac{x^2}{2\sigma_S^2}\right), \quad x \in [0, \infty).
\]  

(24)

This input distribution—although not being the tightest one around—has one important advantage: it allows some analytical results for the mutual information (as we will see below). The other three distributions are considered later in this section as numerical examples.

The next two Lemmas provide an exact closed-form expression for the conditional differential entropy \( h_{Y|X}(\rho) \) and an asymptotic expression for the output differential entropy \( h_Y(\rho) \).

**Lemma 1:** For the channel in \([19]\) and the input distribution \([24]\)

\[
h_{Y|X}(\rho) = \left(2\rho + n - \frac{n}{2} \psi(n) \right) \log_2 e - \frac{n-1}{2} (\log_2 \rho + \psi(1) \log_2 e) + \frac{n \log_2 e}{2} \rho - \frac{n}{2} \Phi\left(\frac{\rho}{\rho + 1}, 1, n\right)
\]

\[
- \rho^{-1} \left(\frac{\rho + 1}{\rho}\right)^{(n-1)/2} F_n(\rho) \log_2 e,
\]

(25)

where \( \psi(x) \triangleq d \log \Gamma(x)/dx \) is the digamma function and \( \Phi(\alpha, 1, n) \) is the special case of the Lerch transcendent function \([64] \text{ eq. (9.551)}\]

\[
\Phi(\alpha, 1, n) \triangleq - \frac{\log(1 - \alpha)}{\alpha^n} - \sum_{k=0}^{n-2} \frac{\alpha^{k+1-n}}{k+1}.
\]

(26)

The function \( F_n(\rho) \) is defined as

\[
F_n(\rho) \triangleq \int_0^\infty \xi K_{n-1}(\sqrt{1 + \rho^{-1} \xi}) I_{n-1}(\xi) \log [I_{n-1}(\xi)] \, d\xi,
\]

(27)

and \( K_n(x) \) is the modified Bessel function of the second kind of order \( n \).

**Proof:** See Appendix B.

**Lemma 2:** For the channel in \([19]\) and the input distribution \([24]\)

\[
h_Y(\rho) = \frac{1}{2} \log_2 \rho + \left(1 - \frac{\psi(1)}{2}\right) \log_2 e - 1 + O\left[\rho^{-1}\right], \quad \rho \to \infty
\]

(28)

**Proof:** See Appendix C.

The next theorem is one of the main results of this paper.
Theorem 3: The MI for the channel in (19) and the input distribution (24) admits the following asymptotic expansion
\[ I_{X,Y}(\rho) = \frac{1}{2} \log_2 \frac{e^{1 - \psi(1)}}{4\pi} \rho + O(\rho^{-1}), \quad \rho \to \infty. \tag{29} \]

Proof: We expand the function \( F_n(\rho) \) in (27) defining the conditional entropy in Lemma [1]. Using a large argument expansion of both Bessel functions, we obtain
\[ F_n(\rho) = 2\rho^2 \log_2 e + \frac{\rho}{2} \log_2 \frac{1}{\rho} + \frac{\rho}{2} (1 - \log 4\pi - \psi(1)) \log_2 e + O[1], \tag{30} \]
which used in (25) gives the asymptotic expression
\[ h_{Y|X}(\rho) = \frac{1}{2} \log_2 \pi e + O[\rho^{-1}], \quad \rho \to \infty. \tag{31} \]

The proof is completed by combining (31) and (28) with (23).

The result in Theorem 3 is a universal and \( n \)-independent expression. The expression in (29) shows that the capacity lower bound is asymptotically equivalent to half of logarithm of SNR plus a constant which is order-independent. Fig. 4 shows the numerical evaluation of \( I_{X,Y}(\rho) \) for \( n = 1, 2, 3, 12 \) as well as the asymptotic expression in Theorem 3. Interestingly, we can see that even in the medium-SNR region, the influence of the number of degrees of freedom on the MI is minimal, and the curves are quite close to each other.

![Fig. 4](image_url)

The MI \( I_{X,Y}(\rho) \) in (29) (numerically calculated) for the \( \chi \)-distribution with different degrees of freedom and the channel model (19). The asymptotic estimate given by Theorem 3 is also shown.

The main reason for considering a Rayleigh input distribution was that it yields a semi-analytical lower bound on the capacity. In the following example, we consider three other input distributions and numerically calculate the resulting MI.

Example 1: Consider the geometric (exponential), half-Gaussian, and Maxwell-Boltzmann distributions given by
\[
p_X(x) = \sqrt{\frac{2}{\sigma_S}} \exp\left(-\frac{\sqrt{2}x}{\sigma_S}\right), \quad x \in [0, \infty), \tag{32}\]
\[
p_X(x) = \sqrt{\frac{2}{\pi\sigma_S}} \exp\left(-\frac{x^2}{2\sigma_S^2}\right), \quad x \in [0, \infty), \tag{33}\]
and
\[
p_X(x) = 3\sqrt{\frac{6}{\pi\sigma_S^3}} \exp\left(-\frac{3x^2}{2\sigma_S^2}\right), \quad x \in [0, \infty), \tag{34}\]
respectively. The MIs for these three distributions for \( n = 1, 2, 3 \) are shown in Fig. 5 and show that the lower bound given by the geometric input distribution in (32) displays high MI in the low SNR regime (\( \rho < 10 \text{ dB} \)), whereas the half-Gaussian input distribution in (33) is better for medium and large SNR. On the other hand, the Maxwell-Boltzmann distribution in (34) gives the lowest MI for all SNR. Numerical results also indicate that all the presented MIs asymptotically exhibit an equivalent growth irrespective of the number of the degrees of freedom \( 2n \).

The following example considers the use of discrete constellations. In particular, we assume that the soliton amplitudes take values on a set \( \mathcal{X} \triangleq \{x_1, ..., x_M\} \), where \( M \triangleq |\mathcal{X}| = 2^m \) is the cardinality of the constellation, and \( m \) is a number of bits per symbol. The MI (23) in this case can be evaluated as
\[
I_{X,Y}(\rho) = \frac{1}{M} \sum_{x \in \mathcal{X}} \int_0^\infty p_{Y|X}(y|x) \log_2 \frac{p_{Y|X}(y|x)}{\frac{1}{M} \sum_{x' \in \mathcal{X}} p_{Y|X}(y|x')} \, dy,
\tag{35}
\]
where we assumed the symbols are equally likely.

Example 2: Consider ASK constellations \( \mathcal{X} = \{0, 1, ..., M - 1\} \) with \( m = 1, 2, 3, 4 \), which correspond to OOK, 4-ASK, 8-ASK, and 16-ASK, respectively. The MI numerically evaluated for these constellations is shown in Fig. 6 for \( \chi \)-channel with \( n = 1, 2, 3 \). As a reference, in this figure we also show (black lines) the MI for the (continuous) half-Gaussian input distribution. The results in this figure show that in the low SNR regime, the use of binary modulation is in fact better than the half-Gaussian distribution. This can, however, be remedied by using a geometric distribution, which, as shown in Fig. 5, outperforms the half-Gaussian distribution in the low SNR regime. In the high SNR regime, however, this is not the case. The results for 16-ASK in the high SNR regime seem to indicate that—in analogy to the well-known ultimate shaping gain of 1.53 dB for the Gaussian channel—a constant gap of approximately 1.35 dB appears at high SNR and for large constellation cardinalities (see 16-ASK at 3.5 [bit/sym]).

V. CONCLUSIONS

A non-Gaussian channel model for the conditional PDF of well-separated (in time) soliton amplitudes was used to study lower bounds on the channel capacity. Results for propagation of signals over a nonlinear optical fibre using one and two polarisations were presented. The results in this paper demonstrated both analytically and numerically that there exist lower bounds on the channel capacity that display an unbounded growth with the effective SNR, similarly to the linear Gaussian channel. All the results in this paper are given in bit per channel use. The more practically relevant (and also more challenging) problem of the channel capacity in bit per second per unit bandwidth is left for further investigation.

This paper studied lower bounds on the capacity of an abstract general noncentral chi channel with arbitrary number
Fig. 5. MI estimates (numerically calculated) for different trial continuous input distributions and different values of $n$ (different line types). Different distributions are shown with different colours.

Fig. 6. MI estimates (numerically calculated) for equally-spaced $M$-ASK constellations with $M = \{2, 4, 8, 16\}$ constellation points.
of degrees of freedom. Similar channel models appear in the study of relatively general systems of noise-driven coupled nonlinear oscillators \cite{54}. Therefore, we believe the results in this paper could also have applications beyond the scope of soliton transmission.

**APPENDIX A**

**MEMORYLESS PROPERTY OF THE DISCRETE-TIME CHANNEL MODEL**

In this section, we present numerical simulations to verify the memoryless assumption for the discrete channel model in Sec. III. To this end, we simulated the propagation of sequences of \( N = 10 \) soliton symbols through the scalar waveform channel given by (5) and a propagation distance of 500 km via the standard split-step Fourier method. The soliton amplitudes were generated as i.i.d. samples from a Rayleigh input distribution (see (24)) with unit mean value \((\sigma_x^2 = 1)\). The transmitted waveform \( x(\tau) \) was created using (13) at a symbol rate of 1.7 GBd and an average launch power of \(-0.53\) dBm. To guarantee an accurate simulation, the time-domain samples were taken every 4.6 ps and the step size was 0.1 km. White Gaussian noise was added at each step to model the ideal DRA process. The simulation parameters are similar to those used in \cite{63} and are summarised in Table I.

![Fig. 7](image-url)

Fig. 7. Continuous-time input \( x(\tau) \) and output \( y(\tau) \) soliton waveforms for 10 solitons and distributed noise due to IRA, and a propagation distance of 500 km.

Fig. 7 shows the waveforms before and after propagation through the channel given in (3). As expected, the received signal is a noisy version of the transmitted waveform. This noisy waveform was then used to obtain soliton amplitudes \( Y \triangleq [Y_1, Y_2, \ldots, Y_{10}] \) via the forward NFT. Each amplitude is obtained by processing the corresponding symbol period via the spectral matrix method \cite{27} Sec. IV-B.

To test the memoryless assumption, we perform a simple correlation test. In particular, we consider the normalised output symbol correlation matrix, whose entries are defined as

\[
\rho_{kk'} = \frac{\mathbb{E}[(Y_k - \mathbb{E}[Y_k])(Y_{k'} - \mathbb{E}[Y_{k'}])]}{\mathbb{E}[Y_k] \mathbb{E}[Y_{k'}]},
\]

(36)

![Fig. 8](image-url)

Fig. 8. The normalised output symbol correlation matrix for an intersymbol distance 0.6 ns (1.7 GBd) and the simulation parameters in Table I.

The obtained correlation matrix is shown in Fig. 8 where statistics were gathered by performing \( 10^3 \) Monte-Carlo runs of the signal propagation. As we can see from Fig. 8 the matrix is almost diagonal. Since our communication channel is believed to be non-Gaussian, the absence of correlation does not of course necessarily imply the memoryless property (understood here as the statistical independence). However, it does constitute an important quantification of the qualitative criterion \( \exp(-At/T_s) \ll 1 \) as given in Sec. III-A.

**APPENDIX B**

**PROOF OF LEMMA 1**

The MI is invariant under a simultaneous linear re-scaling of the variables \( x \rightarrow x/\sigma_N \) and \( y \rightarrow y/\sigma_N \). For notation simplicity, and without loss of generality, throughout this proof we thus assume \( \sigma_N^2 = 1 \). Furthermore, we study the conditional entropy as a function of \( \rho = \sigma_N^2 \) and all the results will be given in nats.
We express the conditional differential entropy as

$$h_{Y|X}(\rho) = -\int p_{X,Y}(x,y) \log p_{Y|X}(y|x) \, dy \, dx \quad (37)$$

$$= -\log 2 - n \mathbb{E} \left[ \log Y \right] + \frac{n-1}{2} \mathbb{E} \left[ \log X \right] + \mathbb{E} \left[ X^2 \right] + \mathbb{E} \left[ Y^2 \right] - \mathbb{E} \left[ \log I_{n-1}(2XY) \right], \quad (38)$$

where (38) follows from (19). In what follows, we will compute the 5 expectations in (38).

The third and fourth terms in (38) can be readily obtained using (24)

$$\mathbb{E} \left[ \log X \right] = \frac{1}{2} \left( \log \rho + \psi(1) \right), \quad (39)$$

$$\mathbb{E} \left[ X^2 \right] = \rho. \quad (40)$$

To compute the second and fifth terms in (38), we first calculate the output distribution as

$$p_Y(y) = \int p_{X,Y}(x,y) \, dx \quad (41)$$

$$= \frac{2y}{\rho \alpha^{n-2}} e^{-\frac{\rho y^2}{2}} \left( 1 - e^{-\alpha y^2} \sum_{k=0}^{n-2} \frac{(\alpha y^2)^k}{k!} \right), \quad (42)$$

where the joint distribution $p_{X,Y}(x,y)$ can be expressed using (19) and (24) as

$$p_{X,Y}(x,y) = \frac{4}{\rho x^{n-2}} y^n \exp \left( -\frac{x^2 + \alpha y^2}{\alpha} \right) I_{n-1}(2xy), \quad (43)$$

with

$$\alpha = \frac{\rho}{\rho + 1} < 1, \quad (44)$$

and where (42) can be obtained using a symbolic integration software. Using (42), we obtain (using a symbolic integration software)

$$\mathbb{E} \left[ \log Y \right] = \frac{1}{2} \left( \alpha \Phi(\alpha,1,n) + \psi(n) \right), \quad (45)$$

where $\psi(n)$ is the digamma function, $\Phi(\alpha,1,n)$ is given by (26). The second moment of the output distribution is obtained directly from the channel input-output relation (20), yielding

$$\mathbb{E} \left[ Y^2 \right] = \rho + n. \quad (46)$$

Substituting (39), (40), (45) and (46) into (38), we have

$$h_{Y|X}(\rho) = -\log 2 - \frac{n}{2} \alpha \Phi(\alpha,1,n) - \frac{n}{2} \psi(n)$$

$$+ \frac{n-1}{2} \left( \log \rho + \psi(1) \right) + 2\rho + n - h_{Y|X}^{(6)}(\rho),$$

where

$$h_{Y|X}^{(6)}(\rho) \triangleq \int \int p_{X,Y}(x,y) \, \mathbb{E} \left[ \log I_{n-1}(2xy) \right] \, dx \, dy \quad (48)$$

The last step is to compute the term $h_{Y|X}^{(6)}(\rho)$, which using (43) can be expressed as

$$h_{Y|X}^{(6)}(\rho) = \frac{4}{\rho} \int \int \frac{y^n}{x^{n-2}} \exp \left( -\frac{x^2 + \alpha y^2}{\alpha} \right)$$

$$I_{n-1}(2xy) \, \log \left[ I_{n-1}(2xy) \right] \, dx \, dy. \quad (49)$$

We then make the change of variables $\xi = 2xy$, $\eta = y^2$, with the Jacobian $\partial(x,y)/\partial(\xi,\eta) = (4y^2)^{-1}$, yielding

$$h_{Y|X}^{(6)}(\rho) = \frac{2^{n-2}}{\rho} \int I_{n-1}(\xi) \log \left[ I_{n-1}(\xi) \right]$$

$$\int \frac{\eta^{n-2}}{\xi} \exp \left( -\frac{\xi^2}{4\eta \alpha} - \eta \right) \, d\eta \, d\xi. \quad (50)$$

The integration over $\eta$ can be performed analytically, yielding

$$\int \frac{\eta^{n-2}}{\xi} \exp \left( -\frac{\xi^2}{4\eta \alpha} - \eta \right) \, d\eta$$

$$= 2^{n-2} \frac{\alpha^{(1-n)/2}}{\alpha^{1/2}} \xi K_{n-1} \left( \frac{\xi}{\alpha^{1/2}} \right), \quad (51)$$

where $K_n(x)$ is the modified Bessel function of the second kind of order $n$. Using (51) in (50) gives

$$h_{Y|X}^{(6)}(\rho) = \frac{\alpha^{(1-n)/2}}{\rho} \int \frac{\xi^{(1-n)/2}}{\alpha^{1/2}} I_{n-1}(\xi) \log \left[ I_{n-1}(\xi) \right] \, d\xi$$

$$= \frac{\alpha^{(1-n)/2}}{\rho} F_n(\rho). \quad (52)$$

The proof is completed by using (53) in (47), the definition of $\alpha$ in (44), and by returning to logarithm base 2.

APPENDIX C

PROOF FOR LEMMA 2

From (42), it follows that the output entropy can then be expressed as

$$h_Y(\rho) = \log \left( \frac{\rho \alpha^{n-2}}{2} \right) - \mathbb{E} \left[ \log Y \right] + \frac{1}{\rho + 1} \mathbb{E} \left[ Y^2 \right] + h_Y^{(4)}(\rho), \quad (54)$$

where $\alpha$ is given by (44).

$$h_Y^{(4)}(\rho) \triangleq \int p_X(x) \int p_{Y|X}(y|x) g_Y^{(4)}(y) \, dy \, dx, \quad (55)$$

and

$$g_Y^{(4)}(y) \triangleq -\log \left( 1 - e^{-\alpha y^2} \sum_{k=0}^{n-2} \frac{(\alpha y^2)^k}{k!} \right). \quad (56)$$

Using the small-$x$ approximation $\log(1 + x) \approx x$ we have

$$g_Y^{(4)}(y) \approx -e^{-\alpha y^2} \sum_{k=0}^{n-2} \frac{(\alpha y^2)^k}{k!}, \quad (57)$$

Similarly to Appendix 6, the results in this proof are in nats.
which in our case holds because as $\rho \to \infty$, $\alpha \to 1$, and thus, the argument of the logarithm tends to 1. Using (57) in (55) and moving the summation out of the integrals gives

$$h_Y^{(4)}(\rho) \approx \sum_{k=0}^{n-2} \frac{\alpha k}{k!} \int_0^\infty p_X(x) \int_0^\infty y^{2k e^{-\alpha y^2}} p_Y|X(y|x) \, dy \, dx.$$  

(58)

For a given $k$ and $x$, we only consider the fast variation of the exponential term obtaining

$$\int_0^\infty y^{2k e^{-\alpha y^2}} p_Y|X(y|x) \, dy \approx (x+1)^{-n} x^{2k e^{-\alpha x^2 / x^2}}$$  

(59)

with $p_Y|X(y|x)$ given by (19). Using (59) in (58), we obtain

$$h_Y^{(4)}(\rho) \approx \sum_{k=0}^{n-2} \frac{\alpha k}{k!} \int_0^\infty x^{2k e^{-\alpha x^2 / x^2}} p_X(x) \, dx$$  

(60)

$$= \sum_{k=0}^{n-2} \frac{\rho^2}{\rho + \beta} \frac{k}{k!} \left( \frac{1}{2} \right)^{k} \partial_{\alpha}^{k} f(\alpha)$$  

(61)

$$= \sum_{k=0}^{n-2} \frac{2k+1}{2} \partial_{\alpha}^{k} f(\alpha) + O\left(\rho^2\right),$$  

(62)

where (61) follows from using (24), and (62) by keeping only the leading terms in $\rho$. The asymptotic expression for the output entropy can be written by combining (62), (55), and (54), which yields

$$h_Y(\rho) = \frac{1}{2} \log \rho + 1 - \frac{\psi(1)}{2} - \log 2 + O\left(\rho^{-1}\right).$$  

(63)

The proof is completed by returning to logarithm base 2.

REFERENCES

[1] M. Cvijetic and I. B. Djordjevic, Advanced Optical Communication Systems and Networks, Artech House, Norwood, MA, 2013.

[2] P. J. Winzer, “Scaling Optical Fiber Networks: Challenges and Solutions,” Optics & Photonics News, vol. 26, pp. 28–35, Mar. 2015.

[3] P. Bayvel, R. Maher, T. Xu, G. Lina, N. A. Shevchenko, D. Lavery, A. Alvarado, and R. I. Killey “Maximizing the Optical Network Capacity,” Phil. Trans. R. Soc. A, vol. 374, art. no. 20140440, Mar. 2016.

[4] D. J. Richardson, “Filling the Light Pipe,” Science, vol. 330, pp. 327–328, Oct. 2010.

[5] P. Mitra and J. Stark, “Nonlinear Limits to the Information Capacity of Optical Fibre Communications,” Nature, vol. 411, pp. 1027–1030, June 2001.

[6] R. Essiambre, G. Kramer, P. J. Winzer, G. I. Foschini, and B. Goebel, “Capacity Limits of Optical Fiber Networks,” J. Lightwave Technol., vol. 28, no. 4, pp. 662–701, Feb. 2010.

[7] R. J. Essiambre, R. W. Tkach, and R. Ryf, “Fiber Nonlinearity and Capacity: Single Mode and Multimode Fibers,” in Optical Fiber Telecommunications, Vol. VIII: Systems and Networks, 6th ed., eds. I. Kaminow, T. Li, and A. E. Willner, pp. 1–43, Academic Press, 2013.

[8] A. Hasegawa and Y. Kodama, Solitons in Optical Communication, Clarendon Press, Oxford, UK, 1995.

[9] E. Iannone, F. Matera, A. MecoZZi, and M. Settembre, “Nonlinear Optical Communication Networks,” Wiley, New York, 1998.

[10] E. Agrati, A. Alvarado, and F. R. Kschischang “Implications of Information Theory in Optical Fibre Communications,” Phil. Trans. R. Soc. A, vol. 374, art. no. 20140438, Mar. 2016.

[11] E. Forestieri and M. Secondini, “The Nonlinear Fiber-Optic Channel: Modeling and Achievable Information Rate,” in Progress In Electromagnetics Research Symposium (PIERS), Prague, Czech Republic, July 2015.
Multiplexed Transmissions based on NFT,” IEEE Photon. Tech. Lett., vol. 27, no. 15, pp. 1621–1623, May 2015.

[35] A. Maruta, A. Toyya, Y. Matsuda, and Y. Ikeda, “Experimental Demonstration of Long Haul Transmission of Eigenvalue Modulated Signals,” Tyrrhenian International Workshop on Digital Communications (TIWDC), Florence, Italy, Sep. 2015.

[36] A. Maruta, “Eigenvalue Modulated Optical Transmission System,” in The 20th Opto-Electronics and Communications Conference (OECC), Shanghai, China, Paper JThA.21, Jun.-Jul. 2015.

[37] 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,” J. Lightwave Technol., vol. 34, no. 10, pp. 2459–2466, May 2016.

[38] L. F. Mollenauer and J. P. Gordon, Solitons in Optical Fibers: Fundamentals and Applications (Academic Press, 2006).

[39] A. Hasegawa and T. Niu, “Eigenvalue Communication,” J. Lightwave Technol., vol. 11, pp. 395–399, Mar. 1993.

[40] S. Wahls and H. V. Poor, “Fast Numerical Nonlinear Fourier Transforms,” IEEE Trans. Inform. Theory, vol. 61, no. 12, pp. 6957–6974, Dec. 2015.

[41] S. Hari, M. Yousefi, and F. Kschischang, “Multieigenvalue Communication,” J. Lightwave Technol., vol. 34, pp. 3110–3117, July 2016.

[42] P. Kazakopoulos and A. L. Moustakas, “On the Soliton Spectral Efficiency in Non-Linear Optical Fibers,” in 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, July 2016.

[43] Q. Zhang, and T. H. Chan, “Achievable Rates of Soliton Communication Systems,” in 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, July 2016.

[44] Q. Zhang, and T. H. Chan, “A Gaussian Noise Model of Spectral Amplitudes in Soliton Communication Systems,” in IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, June 2015.

[45] E. Meron, M. Feder, and M. Shilat, “On the Achievable Communication Rates of Generalized Soliton Transmission Systems,” Dec. 2012. Available: [http://arxiv.org/abs/1207.0297v2](http://arxiv.org/abs/1207.0297v2)

[46] N. A. Shevchenko, J. E. Prilepsky, S. A. Derevyanko, A. Alvarado, P. Bayvel, S. K. Turitsyn, “A Lower Bound on the per Soliton Capacity of the Nonlinear Optical Fibre Channel,” in IEEE Information Theory Workshop (IFW), Jeju Island, Korea, pp. 104–108, Oct. 2015.

[47] S. A. Derevyanko, S. K. Turitsyn, and D. A. Yakushev, “Non-Gaussian Statistics of an Optical Soliton in the Presence of Amplified Spontaneous Emission,” Opt. Lett., vol. 28, no. 21, pp. 2097–2099, Nov. 2003.

[48] S. A. Derevyanko, S. K. Turitsyn, and D. A. Yakushev, “Fokker-Planck Equation Approach to the Description of Soliton Statistics in Optical Fiber Transmission Systems”, J. Opt. Soc. Am. B, vol. 22, no. 4, pp. 743–752, Apr. 2005.

[49] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, “Capacity Estimates for Optical Transmission Based on the Nonlinear Fourier Transform”, Nat. Commun., vol. 7, Art. No. 12710, Sep. 2016.

[50] M. I. Yousefi and X. Yangzhang, “Linear and Nonlinear Frequency-Division Multiplexing,” May 2016. Available: [http://arxiv.org/abs/1603.04389v2](http://arxiv.org/abs/1603.04389v2)

[51] I. Tavakkolnia and M. Safari, “Capacity Analysis of Signaling on the Continuous Spectrum of Nonlinear Optical Fibers,” J. Lightwave Technol., 2017 (to appear).

[52] M. K. Kopae, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn, “Spectral Efficiency Estimation in Periodic Nonlinear Fourier Transform Based Communication Systems,” in Optical Fiber Communication Conference (OFC), Los Angeles, USA, paper Th2A.54, Mar. 2017.

[53] S. A. Derevyanko, J. E. Prilepsky, and D. A. Yakushev, “Statistics of a Noise-Driven Manakov Soliton,” J. Phys. A: Math. Gen. vol. 39, no. 6, pp. 1297–1309, Jan. 2006.

[54] J. E. Prilepsky and S. A. Derevyanko, “Statistics of Noise-Driven Coupled Nonlinear Oscillators: Applications to Systems with Kerr Nonlinearity,” Physica D, vol. 203, no. 3-4, pp. 249–269, Apr. 2005.

[55] J. D. Ania-Castañón, T. J. Ellingham, R. Ibbotson, X. Chen, L. Zhang, and S. K. Turitsyn, “Ultralong Raman Fibre Lasers as Virtually Lossless Optical Media,” Phys. Rev. Lett., vol. 96, art. no. 023902, Jan. 2006.

[56] T. J. Ellingham, J. D. Ania-Castañon, R. Ibbotson, X. Chen, L. Zhang, and S. K. Turitsyn, “Quasi-Lossless Optical Links for Broad-Band Transmission and Data Processing,” IEEE Photon. Technol. Lett., vol. 18, no. 1, pp. 268 – 270, Jan. 2006.

[57] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, “Stability of Solitons in Randomly Varying Birefringent Fibers,” Opt. Lett., vol. 16, no. 16, pp. 1231–1233, Aug. 1991.

[58] T. I. Lakoba and D. J. Kaup, “Perturbation Theory for the Manakov Soliton and its Applications to Pulse Propagation in Randomly Birefringent Fibers,” Phys. Rev. E, vol. 56, pp. 6147–6165, Nov. 1998.

[59] N. L. Johnson, S. Kotz and N. Balakrishnan, Continuous Univariate Distributions. Volume 2, 2nd ed., John Wiley & Sons, New York, 2006.

[60] M. Katz and S. Shamai, “On the Capacity-Achieving Distribution of the Discrete-Time Noncoherent and Partially Coherent AWGN Channels,” IEEE Trans. Inform. Theory, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.

[61] C. E. Shannon, “A Mathematical Theory of Communication,” Bell Syst. Tech. J., vol. 27, no. 3, pp. 379–423, Jul. 1948.

[62] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed., Hoboken, NJ: Wiley, 2006.

[63] O. V. Yushko, A. A. Redyuk, M. P. Fedoruk, and S. K. Turitsyn, “Coherent Soliton Communication Lines,” J. Exp. Theor. Phys., vol. 119, no. 5, pp. 787–794, Nov. 2014.

[64] I. S. Gradshteyn, I. M. Ryzhik, Table of Integrals, Series, and Products, 7th ed., New York, Academic Press, 2007.

[65] L. Schmalen, A. Alvarado, and R. Rios-Müller, “Performance Prediction of Nonbinary Forward Error Correction in Optical Transmission Experiments,” J. Lightwave Technol., vol. 35, pp. 1015–1027, Feb. 2017.