

Information Capacity of Optical Fiber Channels with Zero Average Dispersion

K. S. Turitsyn,¹ S. A. Derevyanko,^{2,3} I. V. Yurkevich,⁴ and S. K. Turitsyn²

¹Landau Institute for Theoretical Physics, Moscow, Kosygina 2, 117940, Russia

²Photonics Research Group, Aston University, Birmingham B4 7ET, United Kingdom

³Institute for Radiophysics and Electronics, Kharkov 61085, Ukraine

⁴The Birmingham University, Birmingham B15 2TT, United Kingdom

(Received 14 March 2003; published 13 November 2003)

We study the statistics of optical data transmission in a noisy nonlinear fiber channel with a weak dispersion management and zero average dispersion. Applying analytical expressions for the output probability density functions both for a nonlinear channel and for a linear channel with additive and multiplicative noise we calculate in a closed form a lower bound estimate on the Shannon capacity for an arbitrary signal-to-noise ratio.

DOI: 10.1103/PhysRevLett.91.203901

PACS numbers: 42.81.-i, 05.10.Gg, 42.65.-k, 42.79.Sz

Introduction.—Information theory (IT) is an interdisciplinary science with applications in areas as diverse as communications, self-organizing complex systems, data compression, genetics, molecular nanotechnology, computer science, biomolecular recognition, thermal physics, quantum computation, economics, and other fields. The most developed part of modern IT is a theory of linear communication channels with additive white Gaussian noise (AWGN). The classical theorem of IT [1] states that the capacity of a power-constrained transmission in an AWGN channel grows logarithmically with the increase of the signal to noise ratio (SNR). Thus, an improvement of the capacity (maximum average information per symbol that can be transmitted through the channel) in such systems can be achieved by the increase of the signal power assuming that the noise level is not affected. The Gaussian statistics of noise that is typical for linear channels is a fundamental assumption in the derivation of this widely known Shannon's result. It should be emphasized that the AWGN linear channel model is not just a simple mathematical construction, but is applied directly to many practical problems such as deep-space communication. However, in some applications, the nonlinear response of a transmission medium must be taken into account. Evidently, properties of nonlinear information channel can be significantly different from that for AWGN models. The interaction of noise and signal in a nonlinear transmission channel can result in non-Gaussian statistics of received signals. The theory of non-Gaussian information channels, though being an evident challenge for many decades, is not yet well established compared to the success of AWGN models. Studies in this fundamental research area are additionally motivated by practical technological achievements and growing demand for efficient, high-speed, high-quality communications. Recent progress in fiber optics has attracted fresh interest to the information theory of non-Gaussian nonlinear communication channels [2–7]. Optical fiber waveguides made of silica present low loss, ultrahigh capacity, cost-efficient transmission media with

many attractive features. Using optical amplifiers to recover signal power simultaneously at many carrier frequencies (channels) within the fiber bandwidth it is possible to transmit optical information data over thousands of kilometers. It is already well recognized, however, that the nonlinear response of the transmission medium (Kerr effect nonlinearity) plays a crucial role in limiting the aggregate capacity of optical fiber systems. Accumulation of nonlinear interactions with propagation along the transmission line makes fiber information channels essentially nonlinear. Evidently, nonlinear impairments (or in other words, a level of signal corruption due to nonlinearity) depend on the signal power. Therefore, in nonlinear channels an increase of the signal power would not necessarily improve the system capacity. Recently, in their pioneering work Mitra and Stark suggested that from the information theory perspective, under certain conditions one can treat essentially nonlinear noisy channels as linear ones with effective multiplicative noise [2]. Applying this idea to multichannel optical fiber transmission systems they derived a heuristic linear model with multiplicative noise that presumably approximates some features of the original nonlinear channel. Though a connection between the statistical properties of such an effective “nonlinear noise” and the system/signal characteristics is still a subject of further research and justification, this intuitive approach outlines a possible way to treat nonlinear transmission channels. In order to compute the Shannon capacity it is necessary to make one more step beyond the determination of a conditional probability. Namely, one has to find the optimal input signal statistics (that is an even more complicated functional problem). The channel capacity C defined by Shannon is a maximum of the following functional (called mutual information) with respect to the statistics of an input signal X , given by a distribution function $p(X)$:

$$C = \max_{p(X)} \int DXDY P(X, Y) \log_2 \frac{P(X, Y)}{P_{\text{out}}(Y)p(X)}. \quad (1)$$

Here $P(X, Y) = P(Y|X)p(X)$ is the joint distribution function of input X and output Y ; $P_{\text{out}}(Y) = \int \mathcal{D}X P(Y|X) \times p(X)$, and all the specific properties of a communication channel are given by the conditional probability $P(Y|X)$. To the best of our knowledge the only case for which there exists an explicit analytical solution of the corresponding functional optimization problem is when the joint distribution of input and output signals is Gaussian. In this case the Shannon capacity can be explicitly expressed through an input-output pair correlation matrix introduced by Pinsker [8]. Difficulties in the analysis of non-Gaussian nonlinear channels are to some extent caused by a relatively limited number of appropriate mathematical methods. Therefore, to practically estimate the capacity of the nonlinear fiber channel, most authors [5–7] apply the Pinsker formula which for the Gaussian noise coincides with Shannon's definition. However, this gives only the lower bound on the capacity [2]. An interesting open problem is the capability of the Pinsker formula to mimic the true behavior of the capacity of nonlinear information channels, especially in the case of large input signal power. Computation of the Shannon capacity for any realistic optical amplifier transmission system is a very complicated problem which is unlikely to be solved analytically. Therefore, it is of crucial importance for further progress in this area to find basic simplified models of fiber nonlinear channels that can be treated analytically. Such solvable models can provide guidance to the analysis of much more complicated general problems in the information theory of nonlinear fiber channels.

In this Letter we present a theoretical analysis of a physical model which describes the transmission of light signals in a noisy nonlinear fiber channel with zero average dispersion. To examine the similarity and difference between the effects of nonlinearity and multiplicative noise, in parallel, we study a linear model of the channel with both additive and multiplicative noise. We calculate analytically the probability density function (PDF) of the channel output for both models. Using our derived conditional probabilities we analyze the capacity of the corresponding transmission systems. We compare here two approaches to the estimation of system capacity: first, based on Pinsker's formula for the input-output correlation matrix and, second, directly applying Shannon's definition of the capacity.

The average propagation of a complex light envelope $E(z, t)$ in a noisy optical fiber line with the so-called weak dispersion management (see for detail, e.g., [9,10]) in the main order is described by the stochastic nonlinear Schrödinger equation

$$\frac{\partial E}{\partial z} = i \frac{\langle d \rangle}{2} \frac{\partial^2 E}{\partial t^2} + i|E|^2 E + n. \quad (2)$$

Here $n(z, t)$ is an additive complex white noise with zero mean and correlation function (see for notation [9])

$$\langle n(z, t) n^*(z', t') \rangle = \langle n_0 \rangle \delta(z - z') \delta(t - t'). \quad (3)$$

The average model (2) is valid when the so-called strength of the map parameter (proportional to the local fiber dispersion and the dispersion oscillation period and inversely proportional to the square of the pulse width) that characterizes the effect of the variation of local dispersion is small (see for detail [10], and references therein). This can be realized either by a short-scale dispersion management [10] or for not very high bit rate allowing for broad enough carrier pulses.

Channel models and calculation of a conditional probability.—In the present Letter we restrict consideration to the two key channel models: First is a weakly dispersion-managed fiber system with zero average dispersion $\langle d \rangle = 0$. The propagation Eq. (2) is then effectively reduced to the Langevin equation for the regularized field $u(z) \equiv E(z, 0)$, with the regularized noise $\eta(z) \equiv n(z, 0)$. We call this model a nonlinear channel (NLCH).

$$\begin{aligned} \frac{du}{dz} - i|u|^2 u &= \eta, & u(z=0) &= u_0, \\ \langle \eta^*(z') \eta(z) \rangle &= D \delta(z - z'). \end{aligned} \quad (4)$$

The second model is a linear channel with multiplicative noise (LMNCH):

$$\frac{du}{dz} - i v u = \eta, \quad u(z=0) = u_0, \quad (5)$$

$$\langle \eta^*(z') \eta(z) \rangle = D \delta(z - z'), \quad (6)$$

$$\langle v(z) v(z') \rangle = D' \delta(z - z'). \quad (7)$$

Here $\eta(z)$ is a white noise with zero mean and $D = 2W \langle n_0 \rangle$ is the regularized noise intensity. To restore the capacity for a bandwidth limited signal one simply has to multiply all the corresponding results by the channel bandwidth W . To calculate the conditional probability $P(u, z|u_0)$ we apply here the so-called Martin-Siggia-Rose formalism [11], that presents the conditional PDF of the output as the following functional integral:

$$P(u, z|u_0) = \int_{q(0)=u_0}^{q(z)=u} \mathcal{D}q e^{-\int_0^z dz' \mathcal{L}[q(z')]}, \quad (8)$$

where the effective Lagrangian is defined as $\mathcal{L}[q] = (D)^{-1} |q' - i|q|^2 q|^2$. Integral (8) can be calculated analytically. The substitution $q(z) = \tilde{q}(z) \exp[i \int_0^z dz' |\tilde{q}(z')|^2]$ brings the Lagrangian to its free form. The Jacobian of this transform is unity and in the new variables the integral becomes Gaussian. After simple straightforward algebra it can be reduced to

$$P(u, z|u_0) = \sum_{m=-\infty}^{+\infty} e^{im\phi} \int \frac{d\phi'}{2\pi} e^{-im\phi'} P'(r, \phi', z|r_0, \phi_0), \quad (9)$$

where the auxiliary “partition function” is

$$P'(r, \phi', z|r_0, \phi_0) \equiv \int_{q(0)=r_0 e^{i\phi_0}}^{q(z)=r e^{i\phi'}} \mathcal{D}q e^{-\int_0^z dz' \{im|q|^2 + D^{-1}|q'|^2\}} \quad (10)$$

(here $u = re^{i\phi}$, $u_0 = r_0 e^{i\phi_0}$). The effective action decomposes into a sum of the classical part and a fluctuating part that does not depend on the limits. The fluctuating field is calculated by expanding over the complete set of eigenfunctions of the operator $-\partial_z^2 + k_m^2$ satisfying zero boundary conditions at $z' = 0$ and $z' = z$. Omitting details of these operations we present a final expression for the conditional probability of our nonlinear channel in polar coordinates (r, ϕ) :

$$P(r, \phi, z|r_0, \phi_0) = \frac{1}{2\pi} \sum_{m=-\infty}^{+\infty} e^{im(\phi-\phi_0)} P_m(r, z|r_0) \\ = \frac{r}{\pi D} \sum_{m=-\infty}^{+\infty} \frac{e^{im(\phi-\phi_0)} e^{-\frac{r^2+r_0^2}{D} k_m \coth k_m z}}{\sinh k_m z} \\ \times k_m I_{|m|}(q_m), \quad (11)$$

here $q_m = 2k_m r r_0 / [D \sinh(k_m z)]$, $k_m = \sqrt{imD}$, and $I_{|m|}$ is the modified Bessel function. Applying a similar procedure to the above, we derive the conditional probability function for the LMNCH of the form Eq. (11) with the replacement $P_m \rightarrow \tilde{P}_m$ where

$$\tilde{P}_m(r, z|r_0) = \frac{2r}{Dz} e^{-m^2 D'z/2} I_{|m|} \left(\frac{2r r_0}{Dz} \right) e^{-[(r^2+r_0^2)/Dz]}. \quad (12)$$

Note that if the information is transmitted using only signal power (the so-called intensity modulation-direct detection systems) $r = |u|$ the conditional probability takes the form [after integration in polar coordinates (r, ϕ) over phase ϕ : $\int d\phi P(r, \phi, z|r_0, \phi_0) = P_0(r, z|r_0)$]:

$$P_0(r, z|r_0) = \tilde{P}_0(r, z|r_0) = \frac{2r}{Dz} I_0 \left(\frac{2r r_0}{zD} \right) e^{-[(r^2+r_0^2)/zD]}. \quad (13)$$

Note that in both cases (nonlinear and effective multiplicative noise channels) formulas (11) and (12) yield the same result. The PDF for NLCH was previously obtained using a different approach by Mecozzi [12].

Channel capacity.—First we revise the procedure commonly used in the recent literature for the channel capacity estimation. We demonstrate here that the consideration based on pair correlation functions [8] can lead to results very different from the Shannon capacity and, therefore, should be used with caution. Some authors [5,6] instead of using the original Shannon definition calculate capacity by exploiting Pinsker formula, which is much more simple, and is based on a complex self-conjugate *input-output correlation matrix* $C_{\alpha\beta}$:

$$C_G \equiv \log_2 \frac{\text{Det diag}(C_{\alpha\beta})}{\text{Det } C_{\alpha\beta}}, \quad C_{\alpha\beta} \equiv \langle u_\alpha u_\beta^* \rangle. \quad (14)$$

Here indices $\alpha, \beta = \text{input, output}$; and brackets stand for the average over noise (η for the nonlinear problem and η and ν for the double noise model) and over the statistics of the input signal u_0 (which is assumed to be Gaussian). Defined in this way the Gaussian capacity

C_G coincides with the Shannon capacity for the case of Gaussian joint input-output distributions, which corresponds to the linear channel with additive noise [8]. For nonlinear channels or channels with multiplicative noise, the Gaussian capacity (14) represents the lower estimate for the true Shannon capacity C (see [2]).

We start from the calculation of the correlation matrix. To perform noise averages we use either PDF (11) or (12). It is easy to find that $C_{\text{in,in}} = \langle |u_0|^2 \rangle \equiv S$, $C_{\text{out,out}} = \langle |u(z)|^2 \rangle = (S + N)$, $N \equiv Dz$ regardless of the model. However, the cross correlations $C_{\text{in,out}} = \langle u_0 u^*(z) \rangle$ are different

$$C_{\text{in,out}} = \left\{ \begin{array}{l} \frac{S \text{sech}^2 k_1 z}{[1+(S/N)k_1 z \tanh k_1 z]^2}, \text{ NLCH} \\ S e^{-D'z/2}, \text{ LMNCH} \end{array} \right\}, \quad (15)$$

where $k_1 = \sqrt{iD}$. Recall that here $\text{SNR} = S/N = s$ changes only due to a variation of S , while $N = Dz$ is fixed as we consider a fixed transmission distance. Substitution of the correlation matrix into the definition Eq. (14) yields the final result

$$C_G = \log_2 \left[1 + \frac{s}{(1+s)a(z)|1+sb(z)|^4 - s} \right], \quad (16)$$

where

$$a(z) = \left\{ \begin{array}{l} |\cosh k_1 z|^4, \text{ NLCH} \\ e^{D'z}, \text{ LMNCH} \end{array} \right\}, \quad (17)$$

$$b(z) = \left\{ \begin{array}{l} k_1 z \tanh k_1 z, \text{ NLCH} \\ 0, \text{ LMNCH} \end{array} \right\}. \quad (18)$$

It is seen from Eq. (16) that with the increase of SNR, C_G decays to zero for the case of the nonlinear channel (similar to the conclusions made in [5,6]) and tends to a constant for the case of the multiplicative noise channel. However, below we will show that in both cases the true Shannon capacity C is unbounded and grows logarithmically with increase of S/N similar to the linear channel.

Direct estimate of the Shannon capacity.—Following Shannon [1] we consider now the channel capacity C defined as a maximum of the mutual information with respect to the statistics of input, u_0 , given by the distribution function $p(u_0)$ under the fixed average input power S :

$$C = \max_{p(u_0)} \int d^2 u d^2 u_0 P(u, u_0) \log_2 \frac{P(u, u_0)}{P_{\text{out}}(u) p(u_0)}. \quad (19)$$

The conditional probability $P(u|u_0)$ connecting output and input probabilities: $P_{\text{out}}(u) = \int d^2 u_0 P(u|u_0) p(u_0)$ is given either by (11) or (12). Note that the Shannon definition allows one to obtain directly an estimate of the capacity. Any arbitrary trial distribution $p(u_0)$ provides for a certain low boundary estimate of the capacity C . The closer a trial function is to the optimal distribution of $p(u_0)$ the better is our approximation of the true capacity. Applying the so-called Klein inequality for two arbitrary probability distribution functions P and \mathcal{P}

$$\int d^2u d^2u_0 P(u, u_0) \log_2 \frac{P(u, u_0)}{\mathcal{P}(u, u_0)} \geq 0 \quad (20)$$

we obtain the following chain of inequalities:

$$\begin{aligned} C &\geq \int d^2u d^2u_0 P(u, u_0) \log_2 \frac{P(u, u_0)}{P_{\text{out}}(u)p(u_0)} \\ &\geq \int d^2u d^2u_0 P(u, u_0) \log_2 \frac{\mathcal{P}(u, u_0)}{P_{\text{out}}(u)p(u_0)} \end{aligned} \quad (21)$$

where \mathcal{P} is an arbitrary PDF (by this we mean that it is non-negative and normalized) and $p(u_0)$ is an arbitrary (not optimal) initial signal distribution. Next we exploit an arbitrariness of \mathcal{P} and $p(u_0)$ in (21) by choosing $p(u_0) = (2\pi r_0)^{-1} p(r_0)$, $\mathcal{P}(u, u_0) = (2\pi)^{-2} (r r_0)^{-1} \times P_0(r|r_0) p(r_0)$. Here we assumed that both an input distribution $p(u_0)$ and \mathcal{P} are phase independent and we have transformed the PDFs to polar coordinates, with $P_0(r|r_0)$ being the radial conditional probability given by Eq. (13). Substitution of these trial functions into inequality Eq. (21) and writing it in polar coordinates yields

$$C \geq \int dr dr_0 P_0(r|r_0) p(r_0) \log_2 \frac{P_0(r|r_0)}{\int dr' P_0(r|r') p(r')}. \quad (22)$$

Evaluation of the right-hand side of this inequality for any trial function leads to an estimate of a lower bound for the Shannon capacity. In what follows we choose the trial input distribution function to be Gaussian which, after the transformation to polar coordinates, yields a Rayleigh distribution for the modulus: $p(r_0) = (2r_0/S) \exp(-r_0^2/S)$. Using this trial function and substituting $P(r, z|r_0)$ from (13), after simple algebra we obtain

$$C \geq C_0(s) = \log_2(1+s) - 2s/\ln 2 + F_1(s), \quad (23)$$

$$F_1(s) = s^{-1} \int_0^\infty dx x K_0(x\sqrt{1+s^{-1}}) I_0(x) \log_2 I_0(x), \quad (24)$$

where I_0 and K_0 are modified Bessel functions and $s = S/N$ is the SNR. Then the main contribution from the integral $F_1(s)$ to the asymptotic behavior of $C_0(s)$ for large s comes from the region $x \gg 1$. Using the asymptotic expansion of modified Bessel functions we get

$$C \geq \frac{1}{2} \log_2 s + O(1). \quad (25)$$

This proves that C_0 and hence the Shannon capacity C are both unbounded as $S/N \rightarrow \infty$.

Our result, in particular, shows that a naive straightforward application of the Pinsker formula for evaluation of the capacity of a nonlinear channel as, for instance, in [5,6], can lead to wrong conclusions regarding the asymptotic behavior of the capacity with $S/N \rightarrow \infty$. Note that the nonlinearity in Eq. (4) does not affect directly the amplitude of the signal. The best way to see this is to make use of the substitution we employed to bring the Lagrangian in the integral (8) to its free form. As a result,

one can expect, at least asymptotically for $s \rightarrow \infty$, that system (4) will be equivalent to the linear intensity-modulated system with direct detection considered in [3]. The result of [3] that asymptotically the Shannon capacity of such a system is roughly one half of that for the system with AWGN is perfectly consistent with Eq. (25).

Discussion and conclusions.—We have examined the statistics of optical data transmission in a noisy nonlinear fiber channel with a weak dispersion management and zero average dispersion. We have also studied the similarity and difference between effects of nonlinearity and multiplicative noise, considering in parallel a linear channel with multiplicative (and additive) noise. Using the analytically calculated conditional PDF we analyzed the Shannon transmission capacity for both models. We managed to find analytically a lower bound estimate for the Shannon capacity of the nonlinear fiber channel considered here. We revise the Pinsker formula which has been used without justification in recent works [5,6] and show that the Gaussian capacity defined through the pair correlation functions should be used with caution in the case of nonlinear transmission channels. Another important result of our analysis is that nonlinearity and multiplicative noise do not necessarily degrade input-output correlations in the same way. Therefore, relating the nonlinear problem to a linear one with multiplicative noise has to be carefully justified for each specific transmission system model.

This work was supported by INTAS Grant No. YS 2002-165 (S.D.) and by the Liverhulme Trust Project No. A/20010049 (S.D., S.K.T.). I.V.Y. gratefully acknowledges support by the Leverhulme Trust under Contract No. F/94/BY and by the EPSRC Grant No. GR/R95432.

-
- [1] C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).
 - [2] P. Mitra and J. Stark, *Nature (London)* **411**, 1027 (2001).
 - [3] A. Mecozzi and M. Shtaif, *IEEE Photonics Technol. Lett.*, **13**, 1029 (2001).
 - [4] E. Desurvire, *Opt. Lett.* **25**, 701 (2000).
 - [5] J. Tang, *J. Lightwave Technol.* **19**, 1104 (2001).
 - [6] J. Tang, *J. Lightwave Technol.* **19**, 1110 (2001).
 - [7] A. Green, P. Littlewood, P. Mitra, and L. Wegener, *Phys. Rev. E* **66**, 46627 (2002).
 - [8] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes* (Holden Day, San Francisco, 1964).
 - [9] E. Iannoe, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks* (John Wiley & Sons, New York, 1998).
 - [10] S. K. Turitsyn, S. B. Medvedev, M. P. Fedoruk, and E. G. Turitsyna, *Phys. Rev. E* **61**, 3127 (2000).
 - [11] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, Oxford, 2002).
 - [12] A. Mecozzi, *J. Lightwave Technol.* **12**, 1993 (1994).