# Optimal input signal distribution and capacity for nondispersive nonlinear optical fiber channel at large signal to noise ratio

I. S. Terekhov, A. V. Reznichenko, S. K. Turitsyn

**SPIE.**

# Optimal input signal distribution and capacity for nondispersive nonlinear optical fiber channel at large signal-to-noise ratio

I. S. Terekhov[a,b], A. V. Reznichenko[a,b], and S. K. Turitsyn[c,a]

[a]Budker Institute of Nuclear Physics of Siberian Branch Russian Academy of Sciences, Novosibirsk, 630090 Russia
[b]Novosibirsk State University, Novosibirsk, 630090 Russia
[c]Aston Institute of Photonics Technologies, Aston University, Aston Triangle, Birmingham, B4 7ET, UK

## ABSTRACT

We consider a model nondispersive nonlinear optical fiber channel with additive Gaussian noise at large SNR (signal-to-noise ratio) in the intermediate power region. Using Feynman path-integral technique we find the optimal input signal distribution maximizing the channel's per-sample mutual information. The finding of the optimal input signal distribution allows us to improve previously known estimates for the channel capacity. We show that in the intermediate power regime the per-sample mutual information for the optimal input signal distribution is greater than the per-sample mutual information for the Gaussian and half-Gaussian input signal distributions.

**Keywords:** Channel capacity, path-integral

## 1. INTRODUCTION

The channel capacity $C$ introduced by Shannon in Ref.[1] is related to the maximum amount of information that can be reliably transmitted over a noisy communication channel. Shannon calculated the capacity of the linear channel with additive white Gaussian noise and found the famous logarithmic dependence of the channel's capacity on the signal power:

$$C \propto \log_2 \left(1 + \text{SNR}\right),\tag{1}$$

where $\text{SNR} = P/N$ is the signal-to-noise power ratio, $P$ is the signal power, and $N$ is the noise power. This means that for fixed noise power $N$, it is necessary to increase the signal power $P$ in order to increase the capacity of the channel.

The interest in nonlinear communication channels increases because the power of the signals in the fiber optics communication channels grows, which leads to appearence of the nonlinear effects connected with Kerr's nonlinearity of the fiber. Fiber optic nonlinear channels have been studied both analytically and numerically in numerous papers, see e.g.[2–14] and references therein. The calculation of the capacity of nonlinear optical fiber channels is a difficult problem therefore new techniques and methods are highly desirable to advance these studies.[3, 13–17] In present paper we consider a simple model nonlinear channel with is far from modern optical communication channels, but it takes into account main features connected with nonlinearity. Also, methods developed for and tested on such model channels might be useful for much more complex and challenging nonlinear fiber communication problems.

The channel capacity $C$ is defined as the maximum of the mutual information $I_{P_X[X]}$ with respect to the probability density function $P_X[X]$ of the input signal $X$:

$$C = \max_{P_X[X]} I_{P_X[X]},\tag{2}$$

Ivan S. Terekhov: E-mail: i.s.terekhov@gmail.com

where the maximum value of $I_{P_X[X]}$ should be found subject to the condition of fixed average signal power:

$$P = \int \mathcal{D}X |X|^2 P_X[X]. \tag{3}$$

The mutual information of a memoryless channel is defined in terms of the output signal entropy $H[Y]$ and conditional entropy $H[Y|X]$:

$$I_{P_X[X]} = H[Y] - H[Y|X], \tag{4}$$

with

$$H[Y|X] = -\int \mathcal{D}X \mathcal{D}Y P_X[X] P[Y|X] \log P[Y|X], \tag{5}$$

$$H[Y] = -\int \mathcal{D}Y P_{out}[Y] \log P_{out}[Y], \tag{6}$$

$$P_{out}[Y] = \int \mathcal{D}X P_X[X] P[Y|X], \tag{7}$$

where $P[Y|X]$ is the conditional probability density function (PDF) for an output signal $Y$ when the input signal is $X$, and $P_{out}[Y]$ is the PDF for an output signal $Y$. The measure $\mathcal{D}Y$ is defined as $\int \mathcal{D}Y P[Y|X] = 1$, and $\mathcal{D}X$ is defined as $\int \mathcal{D}X P_X[X] = 1$. The capacity (2), as defined by (4)-(7), is measured in units of $(\log 2)^{-1}$ bits per symbol (nats per symbol). The input and output signals are functions of time with given bandwidths, but here we consider only per-sample quantities. It means than input and output signals do not depend on time.

The channel's mutual information (4) depends on the probability distribution $P_X[X]$ of the input signal. The input signal PDF, that maximizes the channel's per-sample mutual information is called optimal PDF $P_X^{opt}[X]$. Obviously, the problem of finding the optimal PDF of the input signal for nonlinear optical channels is of great practical importance.

In the previous studies of nondispersive nonlinear optical channels, see, e.g.,[11,13,14] the Gaussian and half-Gaussian input signal PDF's were used as trial functions in order to put low bound constraint on the channel capacity, or to provide asymptotic estimate of the capacity in the regime of large SNR. The authors of[14] argued, that the half-Gaussian PDF which we denote as $P_X^{(1)}[X]$,

$$P_X^{(1)}[X] = \frac{\exp\left\{-|X|^2/(2P)\right\}}{\pi |X| (2\pi P)^{1/2}}, \tag{8}$$

provides the best approximation for the capacity-approaching or optimal input signal distribution at large SNR. In the present paper we show that it is not the case. We find a optimal distribution $P_X^{opt}[X]$ in the regime of large SNR for intermediate power range. We show, that in this regime the mutual information (4) for our optimal input signal PDF is larger than the mutual information for the Gaussian and half-Gaussian input signal distributions.

The estimates for the capacity of nonlinear fiber channels with zero dispersion and additive Gaussian noise in the regime of large SNR were obtained in Refs.[13,14] The lower bound for capacity of the channel, based on trial Gaussian input signal PDF, reads:[13]

$$C \geq \frac{1}{2} \log(\text{SNR}) + o(1) . \tag{9}$$

In[14] capacity estimates were also given in the intermediate power range $N \ll P \ll 6\pi^2 \left(N\gamma^2 L^2\right)^{-1}$. Where $\gamma$ is the Kerr nonlinearity coefficient and $L$ is the fiber link length, see below. For such a power $P$ the following estimate of the lower bound for the capacity, based on the half-Gaussian input signal PDF, was derived:[14]

$$C \geq -\log(\gamma N L) + \frac{\gamma_E - 1 + \log(3\pi)}{2} + \mathcal{O}\left(\frac{1}{\sqrt{\text{SNR}}}\right), \tag{10}$$

where $\gamma_E \approx 0.577$ is the Euler-gamma constant, and instead of $\mathcal{O}\left(\frac{1}{\sqrt{\text{SNR}}}\right)$ the authors presented the explicit function of the parameter SNR which decreases at large SNR, see (40) in.[14] However, the authors of[14] did not take into account the $1/\sqrt{\text{SNR}}$ corrections in the output signal entropy $H[Y]$, therefore, using these explicit functions in the capacity inequality is beyond the calculation accuracy.

The analytical expression for the conditional probability density function of the channel was obtained in the complex form of an infinite series[10,13,14] within the Martin-Siggia-Rose formalism based on quantum field theory methods.[18] In the present paper we adopt the Martin-Siggia-Rose formalism and develop a new method for the approximate computation of the conditional probability density function $P[Y|X]$. Using this method we obtain the simple analytical expression for the function $P[Y|X]$ in the leading order in the parameter 1/SNR for the intermediate power regime

$$N \ll P \ll \left(N\gamma^2 L^2\right)^{-1}. \tag{11}$$

Our method allows us first to derive the analytical expression for the mutual information and then the optimal input signal distribution $P_X^{opt}[X]$ which is different from the half-Gaussian.

In[17] a method to calculate the conditional PDF for a nonlinear optical fiber channel with nonzero dispersion in the large SNR limit was introduced. Here we illustrate this general approach in application to a simpler nondispersive nonlinear optical fiber channel as considered in.[10,13,14] Since the channel is dispersionless, the temporal signal waveform does not change during propagation (note, though, that the signal bandwidth will grow due to the fiber nonlinearity and signal modulation). Therefore, instead of considering the evolution of $\psi(z,t)$ we can consider a set of independent scalar channels[10,14] (per-sample channels) governed by the following model:

$$\partial_z \psi(z) - i\gamma|\psi(z)|^2\psi(z) = \eta(z), \tag{12}$$

where $\psi(z)$ is the signal function that is assumed to obey the boundary conditions $\psi(0) = X$, $\psi(L) = Y$. The noise $\eta(z)$ has zero mean $\langle\eta(z)\rangle_\eta = 0$ and a correlation function $\langle\eta(z)\bar{\eta}(z')\rangle_\eta = Q\delta(z-z')$, so that the SNR $= P/QL$, where $P$ and $N = QL$ are the per-sample signal power and the per-sample noise power, respectively. The connection between the differential model (12) and the conventional information-theoretic presentation in the form of an explicit input-output probabilistic model and appropriate sampling has been discussed in detail in.[10,13,14]

The paper is organized as follows. In Section 2 we develop the quasi-classical method for the calculation of the conditional PDF $P[Y|X]$ for arbitrary nonlinearity in the intermediate power regime (11) in the leading order in 1/SNR. Using this method we find a simple representation for $P[Y|X]$. This allows us to calculate the output signal distribution $P_{out}[Y]$. The optimal signal distribution $P_X^{opt}[X]$ is found in Section 3. Section 4 is focused on the calculation and the comparison of the mutual information for various input signal distributions. We discuss our results in Section 5.

## 2. THE CONDITIONAL PDF $P[Y|X]$ AND OUTPUT SIGNAL PDF $P_{OUT}[Y]$ AT LARGE SNR

### 2.1 "Quasiclassical" method for the conditional PDF $P[Y|X]$ calculation

The conditional probability density function can be written via the path-integral form[13,18,19] in a retarded discretization scheme, see e.g. Supplemental Materials of Ref.[17]

$$P[Y|X] = \int_{\psi(0)=X}^{\psi(L)=Y} \mathcal{D}\psi \exp\left\{-\frac{S[\psi]}{Q}\right\}, \tag{13}$$

and can be reduced to the quasi-classical form, see Ref.:[19]

$$P[Y|X] = e^{-\frac{S[\Psi_{cl}(z)]}{Q}} \int_{\tilde{\psi}(0)=0}^{\tilde{\psi}(L)=0} \mathcal{D}\tilde{\psi}\, e^{-\frac{S[\Psi_{cl}(z)+\tilde{\psi}(z)]-S[\Psi_{cl}(z)]}{Q}}, \tag{14}$$

where the effective action $S[\psi] = \int_0^L dz \left| \partial_z \psi - i\gamma |\psi|^2 \psi \right|^2$, the function $\Psi_{cl}(z)$ is the "classical" solution of the equation $\delta S[\Psi_{cl}] = 0$, where $\delta S$ is the variation of our action $S[\psi]$. The equation $\delta S[\Psi_{cl}] = 0$ (Euler-Lagrange equation) has the form

$$\frac{d^2\Psi_{cl}}{dz^2} - 4i\gamma |\Psi_{cl}|^2 \frac{d\Psi_{cl}}{dz} - 3\gamma^2 |\Psi_{cl}|^4 \Psi_{cl} = 0, \tag{15}$$

with the boundary conditions $\Psi_{cl}(0) = X$, $\Psi_{cl}(L) = Y$.

In order to find $P[Y|X]$ one should calculate the exponent $e^{-\frac{S[\Psi_{cl}(z)]}{Q}}$ and the path-integral in (14). All details of calculations are presented in.[20] Here we present only final result for the function $P[Y|X]$:

$$P[Y|X] = \frac{1}{\pi Q L \sqrt{1 + \mu^2/3}} \exp \left\{ -\frac{(1 + 4\mu^2/3)x_0^2 - 2\mu x_0 y_0 + y_0^2}{QL(1 + \mu^2/3)} \right\}, \tag{16}$$

where $x_0$ and $y_0$ are the real functions of $X$ and $Y$:

$$x_0 + iy_0 = \frac{Ye^{-i\mu}\bar{X}}{|X|} - |X|, \tag{17}$$

$\mu = \gamma L |X|^2$. Since we consider here only the result (16) for the large SNR limit we imply that $|X|^2 \gg QL$. Note that the conditional PDF $P[Y|X]$ was already derived in[13] in the form of an infinite series. Our result (16) for the function $P[Y|X]$ is the analytic summation of this series in the limit of large SNR and intermediate power region

$$QL \ll P \ll \left( QL^3\gamma^2 \right)^{-1}. \tag{18}$$

One can show that the normalization condition $\int DY P[Y|X] = 1$ is fulfilled. Also one can check that the distribution (16) obeys the following important property

$$\lim_{Q\to 0} P[Y|X] = \delta\left( Y - \Psi_0(L) \right). \tag{19}$$

The expression (19) is nothing else, but the deterministic limit of $P[Y|X]$ in the absence of noise. Also Eq. (16) has the correct limit for the linear channel ($\gamma \to 0$):

$$P^{(0)}[Y|X] = \frac{e^{-|Y-X|^2/QL}}{\pi QL}, \tag{20}$$

that is nothing else but the conditional PDF for the linear nondispersive channel with additive Gaussian noise.

## 2.2 Output signal PDF $P_{out}[Y]$

Let us consider the integral, see Eq. (7),

$$P_{out}[Y] = \int \mathcal{D}X P[Y|X] P_X[X], \tag{21}$$

where the function $P_X[X]$ is a smooth function with a scale of variation $P$ which is much greater than $QL$. In that case using the Laplace's method[21] we can calculate the integral (21) up to terms of order of 1/SNR and $\gamma^2 L^3 QP$. The result has the form:

$$P_{out}[Y] = \int \mathcal{D}X P[Y|X] P_X[X] = P_X\left[ Ye^{-i\gamma|Y|^2 L} \right]. \tag{22}$$

The result (22) can be easily understanding because the function $P[Y|X]$ varies on a scale of order $QL$, the function $P_X[X]$ varies on a scale of order of $P$ which is which is much larger than $QL$ therefore instead of function $P[Y|X]$ we can substitute the delta-function limit (19). Note that to obtain the result (22) we do not require the limit $Q \to 0$ but only the relation between the scales $P$ and $QL$ to be satisfied, for details see.[20]

## 3. OPTIMAL INPUT SIGNAL DISTRIBUTION AT LARGE SNR

The optimal input signal distribution at large SNR can be found calculating the mutual information (4) and then maximizing the result with respect to the input signal distribution function $P_X[X]$. Let us start from the calculation of the output signal entropy $H[Y]$, see Eq. (6), at large SNR. When the parameter SNR $\gg 1$ we can substitute $P_X\left[Y\exp\left\{-i\gamma|Y|^2L\right\}\right]$ instead of $P_{out}[Y]$ due to the relation (22):

$$H[Y] = -\int_0^{2\pi} d\phi \int_0^\infty d\rho'\rho' P_X\left[\rho'e^{i\phi}\right]\log P_X\left[\rho'e^{i\phi}\right]. \tag{23}$$

In order to obtain Eq. (23) we have performed the change of the integration variable $\phi = \phi^{(Y)} + \gamma|Y|^2L$. One can check that the output signal entropy (23) coincides with the input signal entropy in the leading order in $1/\text{SNR}$ and $\gamma^2L^3QP$.

The conditional entropy $H[Y|X]$ can be calculated by substitution of $P[Y|X]$ in the form of (16) into (5). After integration over $x_0$, $y_0$ we obtain

$$H[Y|X] = \log(e\pi QL) + \frac{1}{2}\int_0^{2\pi} d\phi^{(X)}\int_0^\infty d\rho\rho P_X\left[\rho e^{i\phi^{(X)}}\right]\log\left(1 + \frac{\gamma^2L^2}{3}\rho^4\right). \tag{24}$$

To find the optimal distribution $P_X^{opt}[X]$ normalized to unity and having a fixed average power $P$ one should solve the variational problem for the functional $J[P_X, \lambda_1, \lambda_2]$

$$J[P_X, \lambda_1, \lambda_2] = H[Y] - H[Y|X] - \lambda_1\left(\int \mathcal{D}XP_X[X] - 1\right) - \lambda_2\left(\int \mathcal{D}XP_X[X]|X|^2 - P\right), \tag{25}$$

where $\lambda_{1,2}$ are Lagrange multipliers. We substitute $H[Y]$ and $H[Y|X]$ from Eqs. (23) and (24) to (25), perform the variation of the functional $J[P_X, \lambda_1, \lambda_2]$ over $P_X[X]$, $\lambda_1$, $\lambda_2$, obtain the Euler-Lagrange equations $\delta J[P_X, \lambda_1, \lambda_2] = 0$, then solve it, and finally obtain $P_X^{opt}[X]$:

$$P_X^{opt}[X] = N_0(P)\frac{\exp\left\{-\lambda_0(P)|X|^2\right\}}{\sqrt{1 + \gamma^2L^2|X|^4/3}}, \tag{26}$$

where $N_0(P)$ and $\lambda_0(P)$ are determined from the conditions $\int \mathcal{D}P_X^{opt}[X] = 1$, and $\int \mathcal{D}|X|^2P_X^{opt}[X] = P$ and depend on signal power $P$. In a parametric form this dependence reads

$$\lambda_0(P) = \frac{\gamma L}{\sqrt{3}}\alpha, \qquad N_0(P) = \frac{\gamma L}{\pi\sqrt{3}\,G(\alpha)}, \tag{27}$$

here $G(\alpha) = \frac{\pi}{2}\left\{H_0(\alpha) - Y_0(\alpha)\right\}$ with $Y_0(\alpha)$ and $H_0(\alpha)$ being the Neumann and Struve functions of zero order, respectively. The parameter $\alpha(P) > 0$ emerges as the real solution of the nonlinear equation $\frac{d}{d\alpha}\log G(\alpha) = -\gamma LP/\sqrt{3}$.

## 4. THE MUTUAL INFORMATION

Substituting the expression (26) for $P_X^{opt}[X]$ in equations (23)-(24) and using the definition (4) we obtain the mutual information for the optimal distribution in the leading order in $1/\text{SNR}$:

$$I_{P_X^{opt}[X]} = P\lambda_0(P) - \log N_0(P) - \log(\pi eQL), \tag{28}$$

which gives a capacity estimate in the region (18) of the average power $P$. The mutual information (28) is depicted by the black solid line in Fig. 1 as a function of power $P$ for the following parameters: $Q = 1.5\times10^{-7}\,\text{mW}\,\text{km}^{-1}$, $\gamma = 10^{-3}\,\text{mW}^{-1}\text{km}^{-1}$, $L = 1000\,\text{km}$. These parameters are usual for the modern optical fiber links. For these parameters, the power range (18) is actually very wide:

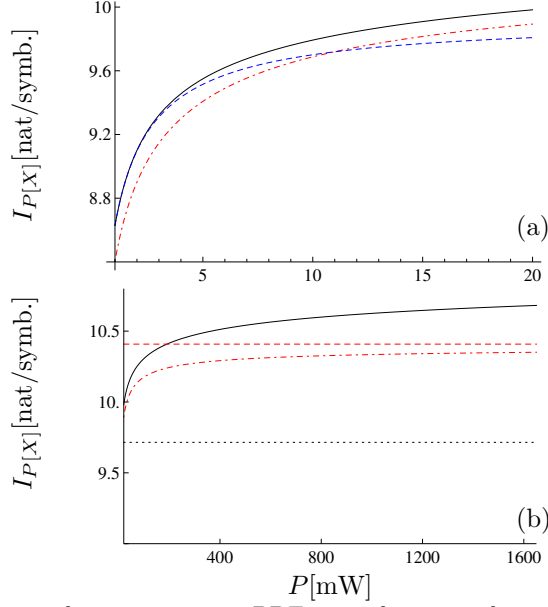$$1.5\times10^{-4}\text{mW} \ll P \ll 0.66\times10^4\text{mW}. \tag{29}$$

Figure 1. The mutual information for various input PDFs as a function of input average power $P$ for the parameters $Q = 1.5 \times 10^{-7}\,\text{mW}\,\text{km}^{-1}$, $\gamma = 10^{-3}\,\text{mW}^{-1}\text{km}^{-1}$, $L = 1000\,\text{km}$. (a): The solid black line, blue dashed line, red dashed dotted line correspond to the optimal PDF $P_X^{opt}[X]$, Gaussian PDF $P_X^{(2)}[X]$ (see[20]), and half-Gaussian PDF $P_X^{(1)}[X]$, respectively. (b): The solid black line corresponds to $I_{P_X^{opt}[X]}$, see Eq. (28); the red dashed dotted line corresponds to the mutual information for the half-Gaussian distribution $I_{P_X^{(1)}[X]}$, see Eq. (31); the red dashed horizontal line corresponds to our limit (32) at $\tilde{\gamma} \gg 1$ for the half-Gaussian distribution; the black dotted horizontal line corresponds to the result,[14] see Eq. (10).

There is no simple analytical form for $N_0(P)$ and $\lambda_0(P)$ therefore to plot Fig. 1 and Fig 2 we calculated $\lambda_0(P)$ and $N_0(P)$ numerically. For large and small values of the parameter $\tilde{\gamma}$ we found analytical expressions for parameters $\lambda_0(P)$ and $N_0(P)$, see.[20] At small $\tilde{\gamma} = \gamma L P / \sqrt{3}$ we obtain

$$I_{P_X^{opt}[X]} = \log(\text{SNR}) - \tilde{\gamma}^2, \tag{30}$$

which is simply the Shannon capacity, $\log(1 + \text{SNR})$, of the linear additive Gaussian noise channel (1) with the first nonlinear correction at large SNR. In the power sub-interval $(\gamma L)^{-1} \ll P \ll (Q L^3 \gamma^2)^{-1}$, we obtain that the mutual information increases as (loglog) with $P$:

$$I_{P_X^{opt}[X]} \approx -\log\left(\frac{Q L^2 \gamma e}{\sqrt{3}}\right) + \log\log\left(\frac{C \gamma L P}{\sqrt{3}}\right).$$

Note that mutual information for half-Gaussian-like has constant behaviour in the sub-interval. To show this we perform an analysis of the mutual information for the distribution $P_X^{(1)}[X]$, see Eq. (8). In the leading order in $1/\text{SNR}$ using (23), (24) and (4) we obtain the mutual information $I_{P_X^{(1)}[X]}$:

$$I_{P_X^{(1)}}[X] \approx -\frac{1}{2\sqrt{\pi}} \int_0^\infty d\tau e^{-\tau} \tau^{\beta/2 - 1} \log\left(1 + \frac{4\tilde{\gamma}^2}{\beta^2} \tau^2\right) + \log \text{SNR} + \log\left(2\sqrt{\pi} e^{\gamma_E - 1}\right). \tag{31}$$

The mutual information is depicted in Fig. 1(a) by the red dashed dotted line. One can see that at small $P$ the mutual information for the Gaussian distribution, see Fig. 1(a), is greater than that of the half-Gaussian, whereas at $P > 11$mW the mutual information is greater for the half-Gaussian distribution. Note that $I_{P_X^{opt}[X]}$ is greater than $I_{P_X^{(1)}[X]}$ for all values of $P$, as it should be. At $\tilde{\gamma} \gg 1$ the mutual information $I_{P_X^{(1)}[X]}$ has the form

$$I_{P_X^{(1)}[X]} = -\log\left(\frac{Q L^2 \gamma}{2}\right) + \frac{\log 3\pi - 1 + \gamma_E}{2}. \tag{32}$$

Since we have found $P_X^{opt}[X]$ in the power region (18), we can calculate an approximation for the capacity of the considered per-sample nonlinear channel. By definition it coincides with the mutual information expression (28):

$$C = I_{P_X^{opt}[X]}. \tag{33}$$

Note that the result for the capacity is calculated in the leading order in the parameters $1/\text{SNR}$ and $\gamma^2 L^3 QP$. The capacity (33) and the Shannon capacity of the linear channel with additive Gaussian noise is depicted in Fig. 2. One can see that the Shannon capacity is always greater than the approximation (33) for the nondispersive
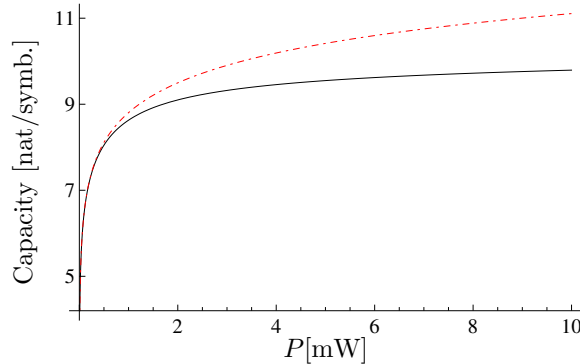


Figure 2. Shannon capacity and the capacity of the nonlinear channel $I_{P_X^{opt}[X]}$ for the parameters $Q = 1.5 \times 10^{-7}\,\text{mW}\,\text{km}^{-1}$, $\gamma = 10^{-3}\,\text{mW}^{-1}\text{km}^{-1}$, $L = 1000\,\text{km}$. The black solid line corresponds to $I_{P_X^{opt}[X]}$, see Eq. (28), the red dashed-dotted line corresponds to the Shannon limit $\log[1 + \text{SNR}]$.

nonlinear fiber channel for the considered region of $P$.

## 5. CONCLUSION

We have developed a new method for the calculation of the conditional probability density function using the path-integral representation. Using this method we obtain compact analytical expressions for the conditional PDF, for the per-sample nondispersive nonlinear fiber channel. The simple analytical form of conditional PDF allow us to calculate the conditional and the output signal entropies for different input signal PDFs $P_X[X]$. Using these entropies we found $P_X^{opt}[X]$ maximizing the mutual information in the leading order in the parameter $1/\text{SNR}$ in the power region (18). The found optimal PDF allows us to find the channel capacity (28) up to corrections of order of $1/\text{SNR}$ and $\gamma^2 L^3 QP$ in the power region $QL \ll P \ll (\gamma^2 QL^3)^{-1}$. The found distribution $P_X^{opt}[X]$ is different from the half-Gaussian one, and at the zero nonlinearity $P_X^{opt}[X]$ approaches the Gaussian distribution. We demonstrated that the found capacity of the channel is always greater than the mutual information calculated for Gaussian and half-Gaussian distributions, and lower than the Shannon capacity of the linear channel with additive Gaussian noise.

## REFERENCES

[1] C. Shannon, "A mathematical theory of communication", *Bell System Techn. J.,* vol. 27, no. 3, pp. 379–423, 1948; vol. 27, no. 4, pp. 623–656, 1948.

[2] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications", *Nature*, vol. 411, pp. 1027-1030, 2001.

[3] E. E. Narimanov and P. Mitra, "The channel capacity of a fiber optics communication system: Perturbation theory", *J. Lightw. Technol.,* vol. 20, no. 3, pp. 530537, 2002.

[4] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation detection techniques for DWDM systems", *IEEE. J. Sel. Topics Quant. Electron.,* vol. 10, no. 2, pp. 259272, 2004.

[5] R.-J. Essiambre, G. J. Foschini, G. Kramer, and P. J. Winzer, "Capacity Limits of Information Transport in Fiber-Optic Networks", *Phys. Rev. Lett.*, vol. 101, p. 163901, 2008.

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

[7] R. Killey and C. Behrens, "Shannon's theory in nonlinear systems", *J. Mod. Opt.,* vol. 58, no. 1, pp. 1–10, 2011.

[8] E. Agrell, A. Alvarado, G. Durisi, M. Karlsson, "Capacity of a Nonlinear Optical Channel with Finite Memory", *arXiv:1403.3339.*

[9] M. A. Sorokina and S. K. Turitsyn, "Regeneration limit of classical Shannon capacity", *Nat. Comm.,* vol. 5, p. 3861, 2014.

[10] A. Mecozzi, "Limits to long-haul coherent transmission set by the Kerr nonlinearity and noise of the in-line amplifiers", *J. Lightwave Technol.,* vol. 12, no. 11, pp. 1993 - 2000, 1994.

[11] A. Mecozzi and M. Shtaif, "On the capacity of intensity modulated systems using optical amplifiers", *IEEE Photonics Technol. Lett.,* vol. 13, no. 9, pp. 1029 - 1031, 2001.

[12] J. Tang, "The Shannon channel capacity of dispersion-free nonlinear optical fiber transmission", *J. Lightwave Technol.,* vol. 19, no. 8, pp. 1104 - 1109, 2001.

[13] K.S. Turitsyn, S.A. Derevyanko, I.V. Yurkevich, and S.K. Turitsyn, "Information Capacity of Optical Fiber Channels with Zero Average Dispersion", *Phys. Rev. Lett.,* vol. 91, p. 203901, 2003.

[14] M. I. Yousefi,F. R. Kschischang, "On the Per-Sample Capacity of Nondispersive Optical Fibers", *IEEE transactions on information theory*, vol. 57, no. 11, pp. 7522 - 7541, 2011.

[15] E. Agrell, "The Channel Capacity Increases with Power", *arXiv: 1108.0391v3.*

[16] E. Agrell, "Nonlinear Fiber Capacity", *Eur. Conf. Opt. Commun. London U.K.,* paper We.4.D.3, 2013.

[17] I. S. Terekhov, S. S. Vergeles, and S. K. Turitsyn, "Conditional Probability Calculations for the Nonlinear Schrödinger Equation with Additive Noise", *Phys. Rev. Lett.,* vol. 113, p. 230602, 2014.

[18] J. Zinn-Justin, "Quantum Field Theory and Critical Phenomena", Oxford University Press, Oxford, 2002.

[19] R. P. Feynman, A. R. Hibbs, "Quantum mechanics and path integrals", McGraw-Hill Book Company, New York, 1965.

[20] I. S. Terekhov, A. V. Reznichenko, Ya. A. Kharkov, S. K. Turitsyn, "Log-log growth of channel capacity for nondispersive nonlinear optical fiber channel in intermediate power range", *Phys. Rev. E*, vol. 95, p. 062133, 2017.

[21] M. A. Lavrentiev, B. .V. Shabat, "Method of Complex Function Theory". Nauka, Moscow, 1987 (in Russian) or M. Lavrentiev, B. Chabot, "Methodes de la Theorie des fonctions d'une variable complexe", Mir, Moscou, 1977 (in French).