Quantum Electronics **47** (11) 1049–1052 (2017)

OPTICAL INFORMATION TRANSMISSION

https://doi.org/10.1070/QEL16394

Influence of nonlinear interaction on the capacity of an optical dispersion-compensating channel

E.G. Shapiro, D.A. Shapiro

Abstract. **A channel with weight coefficients of nonlinear finite memory is simulated numerically. Areas of the signal power are found where the capacity is higher than that of a regular Gaussian channel. It is shown that an increase in the number of accountable adjacent symbols has little effect on the channel capacity. The numerical calculation reveals that the negative mean dispersion of the link reduces the interaction of neighbouring pulses.**

Keywords: *fibre-optic communication lines, capacity, channel with memory, nonlinear noise.*

1. Introduction

The growing need for rate and volume of transmitted information requires an increase in the capacity of communication lines. Coherent detection and new modulation formats provide an increase in the information capacity of optical communication systems. The problem of estimating the throughput of optical links is complex and urgent. The complexity of the problem is due to the presence of nonlinearity, which accompanies the propagation of pulses in fibre-optic systems. In communication lines without dispersion compensation, sets of transmitted symbols become overlapped, with the nonlinear interaction effect decreasing. In such systems, the combined action of the Kerr nonlinearity and chromatic dispersion manifests itself as additive Gaussian noise [1–4]. A number of papers published experimental results confirming the adequacy of the nonlinear Gaussian noise (GN) model [5–9].

In a regular GN model, the detected symbols are independent on the receiving device and have a Gaussian distribution with the same variance. Unlike a linear channel, the nonlinearity limits the signal-to-noise ratio. This circumstance is often called the nonlinear Shannon limit.

In this paper we consider fibre links that transmit symbols of different power. In this case, the variances of the symbols are different, in contrast to the GN model. Examples of such optical systems include dispersion-compensating fibre links. In this paper, we numerically simulate signal propagation along the communication link with controlled dispersion. It is shown that from the point of view of the communication link capacity, the data transmission formats where the nonlinear self-interaction of the pulse prevails over the interaction with neighbouring pulses are more preferable. It is also found that an increase in the amount of the accountable memory from one nearest bit on the right and left to two adjacent bits on both sides has little effect on the information capacity of the optical system.

E.G. Shapiro, D.A. Shapiro Institute of Automation and Electrometry, Siberian Branch, Russian Academy of Sciences, prosp. Akad. Koptyuga 1, 630090 Novosibirsk, Russia; Novosibirsk State University, ul. Pirogova 2, 630090 Novosibirsk, Russia; e-mail: elena.shap@gmail.com

Received 4 April 2017; revision received 1 August 2017
Kvantovaya Elektronika **47** (11) 1049–1052 (2017)
Translated by I.A. Ulitkin

2. Channels with memory

The mutual information [10] of a discrete time-invariant channel with memory is given by the formula

$$I(X;Y) = \lim_{N\to\infty}\frac{1}{N}\,I(X_1,\dots,X_N;Y_1,\dots,Y_N), \quad (1)$$

where $(X_1,\dots,X_N)$ and $(Y_1,\dots,Y_N)$ are input and output sequences of symbols, respectively. The signal X_i for each number i is a random variable that takes values from a certain set of numbers; this set of numbers is called the input alphabet. The signal values recorded at the end of the link are also the values of the random variable Y_i and are called the output alphabet.

In Refs [1–9], the nonlinear interaction distorting the signal is considered to be Gaussian noise, which is cubic dependent on the signal power. Mathematically, this is written as follows:

$$Y_k = X_k + Z_k,\; Z_k = \tilde{Z}_k\sqrt{\sigma^2_{\rm ASE} + \mu P^3}. \quad (2)$$

Here, X_k is the symbol transmitted in the time slot with the number k; Y_k is the value recorded at the receiver; $\tilde{Z}_k$ is the zero-mean unit-variance Gaussian random variable; $\sigma^2_{\rm ASE}$ and μ (memory parameter) are the nonnegative constants; and P is the average signal power. The constant $\sigma^2_{\rm ASE}$ corresponds to the noise of spontaneous emission of amplifiers.

A model of a Gaussian channel with finite memory was proposed by Agrell et al. [11]. In this model, the average signal power P in (2) is replaced by the empirical power, i.e., by the average power of the symbol $|X_k|^2$ and of the $2L$ symbols around it:

$$Z_k = \tilde{Z}_k\left[\sigma^2_{\rm ASE} + \mu\left(\frac{1}{2L+1}\sum_{i=k-L}^{k+L}|X_i|^2\right)^3\right]^{1/2}. \quad (3)$$

Formula (3) means that the symbols with the numbers $k-L,\dots,k+L$ contribute equally to the signal distortion. If we introduce the weight coefficients $e_{k-L}, e_{k-L+1},\dots, e_{k+L}$, then formula (3), with equality of all coefficients to the value $(2L+1)^{-1}$, is a particular case of the expression

$$Z_k = \tilde{Z}_k\left[\sigma_{\rm ASE}^2 + \mu\left(\sum_{i=k-L}^{k+L} e_i\,|X_i|^2\right)^3\right]^{1/2}, \ \sum_i e_i = 1. \quad (4)$$

The value of e_i corresponds to the contribution of the symbol X_i to nonlinear noise.

We considered the following cases: (i) the same values of the weight coefficients e_i, $L = 1$; (ii) the weight coefficient of the central bit is two times larger than the coefficients of the neighbouring bits, $L = 1$; (iii) the limiting case $L = 0$, when adjacent bits do not cause nonlinear distortion of the signal. In addition, we compared the channel capacity for $L = 1$ and $L = 2$, i. e., one symbol on the right and left and two adjacent symbols respectively in the case of the same weight coefficients were taken into account. Optical pulses are broadened and overlapped when propagating along the optical fibre, with the greatest distortion being caused by the interaction of neighbouring pulses. Therefore, the case $L = 1$ was mainly considered.

According to Shannon's theorem, the capacity of a channel without memory is given by the formula

$$C = \sup I(X; Y), \quad (5)$$

where maximisation (5) is satisfied over all distributions p_X of the input alphabet at a given level of the average signal power $\int |x|^2 p_X \,dx = P$.

The capacity of a regular channel (2) is given by the formula

$$C = \lg\left(1 + \frac{P}{\sigma_{\rm ASE}^2 + \mu P^3}\right)$$

in the case of a complex channel and complex Gaussian noise (the lg function here and below denotes the base-2 logarithm). If the Gaussian noise Z_k in the channel and also the values of X_k are real, then

$$C = \frac{1}{2}\lg\left(1 + \frac{2P}{\sigma_{\rm ASE}^2 + \mu P^3}\right). \quad (6)$$

3. Lower bandwidth limit

To calculate the capacity of the communication link, we used the calculation method [12], which is a modification of the method presented in Ref. [13]. This method consists in calculating the capacity of an auxiliary channel, the bandwidth of which approximates the desired value from below with an accuracy of $O(\mu^2)$.

For convenience, we describe the auxiliary channel. Consider a fixed bit slot with the number k. The value transmitted in this bit slot is a random value, just like the value being recorded. Let q_i denote the probability of an event $X_k = x_i$. In the finite-memory model (4), the detection probability $Y_k = y_j$ depends on the nearest L bits on the left, $X_{k-L}, \ldots, X_{k-1}$, and L bits on the right, $X_{k+1}, \ldots, X_{k+L}$.

We denote by $p_{ji}(x_{t_{k-L}}, \ldots, x_{t_{k-1}}, x_{t_{k+1}}, \ldots, x_{tk+L})$ the probability of the event $Y_k = y_j$, $X_k = x_i$, $X_{k-L} = x_{t_{k-L}}, \ldots, X_{k+L} = x_{t_{k+L}}$.

Then the conditional probability Q_{ji} of the detection $Y_k = y_j$ at the transmitted value $X_k = x_i$ is defined by the formula

$$Q_{ji} = \sum_{t_{k-L}, \ldots, t_{k+L}} p_{ji}(x_{t_{k-L}}, \ldots, x_{t_{k-1}}, x_{t_{k+1}}, \ldots, x_{t_{k+L}})\, q_{t_{k-L}} \ldots q_{t_{k+L}}.$$

Consider the function

$$F(q_1, \ldots, q_n) = \sum_{j,i} Q_{ji} q_i \lg \frac{Q_{ji}}{\sum_{i'} Q_{ji'} q_{i'}}, \quad (7)$$

which specifies the mutual information in the auxiliary channel. As shown in [12], the mutual information in the finite-memory channel (4) coincides with an accuracy of $O(\mu^2)$ with $F(q_1, \ldots, q_n)$.

4. Numerical experiment

To verify the applicability of models (2) and (4), we have performed a numerical simulation of the signal propagation along a 1000-km-long dispersion-controlled link. We have considered an optical communication link consisting of 25 spans with a configuration

SMF(L_1) + EDFA + DCF(L_2) + EDFA.

Here, SMF is a standard single-mode fibre, and DCF is a dispersion-compensating fibre. Erbium-doped fibre amplifiers EDFAs had a noise figure of 4.5 dB and completely compensated for the attenuation of the signal in the fibre span between the amplifiers. The lengths of the SMF and DCF spans are denoted by L_1 and L_2, respectively. In numerical simulation, it was assumed that $L_1 = 40$ km and $L_2 = 6.8$ or 7.18 km. Thus, the average dispersion of the span is zero (at $L_2 = 6.8$ km) or -0.8 ps nm^{-1} km^{-1} (at $L_2 = 7.18$ km). Gaussian pulses with a width of 6.67 ps at half-maximum peak power propagated along the optical link. The average power in the bit slot varied from 0.3 to 1 mW. The simulation was performed for the on/off keying (OOK) format. The probabilities of unit and zero bits were 1/2. To describe the dynamics of optical pulses, the nonlinear Schrödinger equation was used. The calculated complex function $A(z,t)$ depends on the time t and the distance z. At the receiving device (at $z = 1000$ km), the function $A(z,t)$ was averaged over a 3/4 bit slot in the central part.

The values of zero and unit bits obtained in numerical simulation form samples of random variables. We denote by σ_0^2 and σ_1^2 the variances of the detected zero and unit bits, respectively. The results of numerical simulation show that these variances differ significantly, with the dispersion of unit bits being greater. This is not consistent with model (2), in which the variances are the same for all transmitted symbols.

Let us consider model (4) under the assumption of equal probability of zeros and ones. We denote $e_{k-1} = \beta$. Since in the nonlinear Schrödinger equation the time t can be replaced by $-t$, then $e_{k-1} = e_{k+1} = \alpha$. Simple calculations show that the equalities

$$\sigma_0^2 = \sigma_{\rm ASE}^2 + \mu\alpha P^3, \quad \sigma_1^2 = \sigma_{\rm ASE}^2 + \mu(\alpha + \beta) P^3,$$

$$\mu = (\sigma_1^2 + \sigma_0^2 - 2\sigma_{\rm ASE}^2)/P^3,$$

$$\alpha = (\sigma_0^2 - \sigma_{\rm ASE}^2)/(\sigma_1^2 + \sigma_0^2 - 2\sigma_{\rm ASE}^2),$$

are fulfilled, where P is the power of units. The parameter β is given by the relation $\beta = 1 - 2\alpha$. Thus, the variances of zero and unit bits are cubically increasing with increasing P, while the unit variance demonstrates a greater growth. A similar behaviour is observed in the statistics of units and zeros in the numerical experiment. Figure 1 shows the dependence of μ and α on the power of the original unit bits. The accumulated noise of the amplifier is $\sigma_{\rm ASE}^2 = 3.66 \times 10^{-3}$ mW. It can be

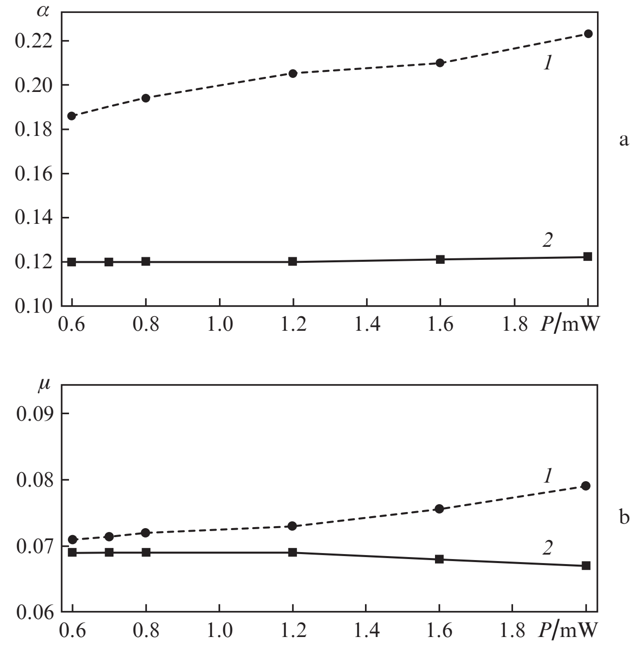

Figure 1. Dependences of the nonlinear memory coefficients (a) α and (b) μ on the power of the initial unit bits for zero dispersion (*1*) and dispersion equal to 0.8 ps nm^{-1} km^{-1} (*2*).

seen that the coefficients μ and α vary only slightly in the power range 0.6–1.6 mW, especially weakly in the case of nonzero dispersion. In addition, it is seen that in the presence of an average negative dispersion, the coefficient α can be reduced.

Next, we estimate model (4) from the point of view of the capacity in the framework of information theory.

5. Information capacity of the channel

To calculate the information capacity of communication lines, it is convenient to pass to dimensionless quantities. In formulas (2)–(4), the amplifiers' noise $\sigma^2_{\rm ASE}$ is measured in power units. Let A_i correspond to the transmitted symbol with the number i and the set $\{A_i\}$ be the alphabet of the optical system. We replace A_i by $x_i = A_i/\sigma_{\rm ASE}$; then, the average dimensionless signal power is given by expression

$$S = \frac{1}{\sigma^2_{\rm ASE}} \sum_i p_i |A_i|^2,$$

the dimensionless gain noise is equal to 1, and the nonlinear memory coefficient is $\mu\sigma^4_{\rm ASE}$. For numerical simulation, we assume that $\mu\sigma^4_{\rm ASE} \approx 9.4 \times 10^{-7}$, all quantities are dimensionless and $\sigma^2_{\rm ASE} = 1$. We have considered the cases when $\mu = 6.75 \times 10^{-3}$ and 2.7×10^{-8}, that is, the values of μ differ by several orders of magnitude.

In numerical simulation, the realisation of infinite alphabets is impossible. Consider the finite input alphabet

$$x_i = \Delta x(i-1) - L_x/2, \quad i = 1, \ldots, n, \quad \Delta x = L_x/(n-1),$$

where L_x is some given interval. We set $x_0 = -\infty$ and $x_{n+1} = \infty$. The output alphabet is given by the formulas $y_j = \Delta y(j-1) - L_y/2$, $[-L_y/2, L_y/2]$ is the region containing $y_1, \ldots, y_m$, $j = 1, \ldots, m$, $\Delta y = L_y/(m-1)$, $y_0 = -\infty$, $y_{m+1} = \infty$. In this paper, the parameters were: $n = 15$, $m = 31$, $L_x = 7$, $L_y = 8$ for $\mu = 6.75 \times 10^{-3}$ and $n = 41$, $m = 41$, $L_x = 50$, $L_y = 50$ for $\mu = 2.7 \times 10^{-8}$.

We calculated the capacity of the link for both the values of the memory parameter and the following triplets of weight coefficients:

1) $e_{k-1} = e_k = e_{k+1} = 1/3$;
2) $e_{k-1} = e_{k+1} = 0.25$, $e_k = 0.5$;
3) $e_{k-1} = e_{k+1} = 0$, $e_k = 1$.

Figure 2 shows the dependence of the capacity on the signal power for these three variants. It is seen that with a smaller contribution of neighbouring bits to nonlinear noise compared to the central bit (triplets 2 and 3), the capacity is greater. In addition, the optimal power area is shifted towards increasing the power and is wider in comparison with the case of the same weight coefficients. It also follows from Fig. 2 that the qualitative behaviour of the curves is the same for both values of the coefficient μ.

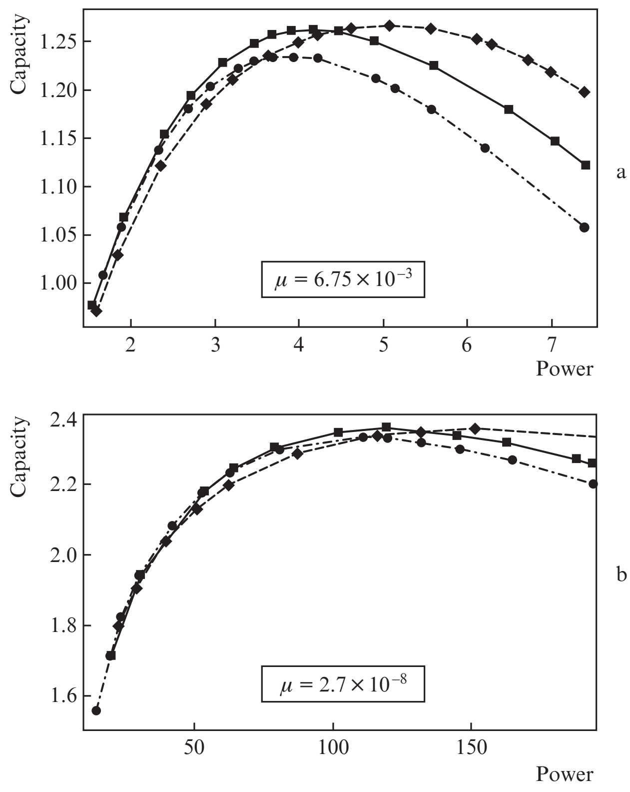

Figure 2. Dependences of the capacity on the signal power for the indicated triplets of the weight coefficients (1 – dash-dotted curve, 2 – solid curve, 3 – dashed curve).

The model of a regular Gaussian channel (2) takes into account the averaged noise and is a model of a channel without memory. Formula (6) specifying the capacity of an ideal Gaussian channel is valid if the real straight line is the domain of variation of X_k and Y_k. Limiting the power of the input signal practically does not reduce the bandwidth at a relatively low signal power. However, in the case of high power, the limitation reduces the information capacity of the channel. Figure 3 shows the graphs of the capacity of a regular Gaussian channel without power limitation [curve (*1*)] and with power limitation [curve (*2*)] at $\mu = 6.75 \times 10^{-3}$. It is seen that in the ideal case, without power limitations, the bandwidth is higher. To clarify the influence of the power of the central bit on the bandwidth, we compared the capacity of an ideal Gaussian channel, a channel with a contribution of only the central bit and a nonideal Gaussian channel with a limited signal power. It follows from the figure that the channel, in which the noise is set by the noise of the amplifiers and the Kerr nonlinearity of only the transmitted bit, demonstrates a wider range of optimum powers. At high powers, the capacity in this channel is higher than that in the Gaussian channel. In addition, a greater signal power is required to achieve the maximum capacity. Thus, the communication link can be optimised by choosing the weight coefficients of nonlinear finite memory.

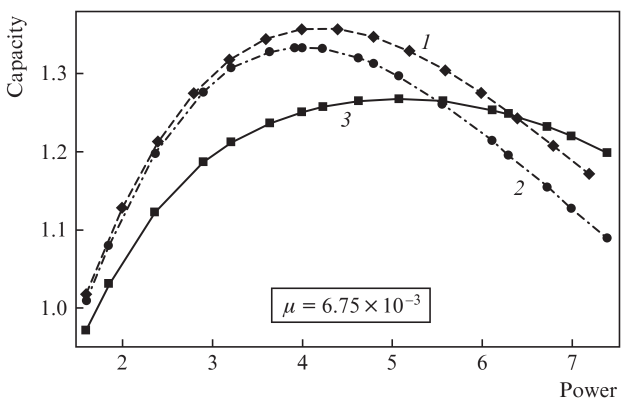

Figure 3. Capacity vs. signal power for (*1*) an ideal Gaussian channel, (*2*) a Gaussian channel with power limitation and (*3*) a channel with only central bit noise.

To calculate the channel capacity (see Fig. 3), the Arimoto–Blahut method was used [14]. For $L = 1$ and $L = 2$, that is, if one symbol is taken into account on the right and left and two symbols, respectively, the weight coefficients were assumed to be the same: $e_{k-1} = e_k = e_{k+1} = 1/3$ for $L = 1$ and $e_{k-2} = e_{k-1} = e_k = e_{k+1} = e_{k+2} = 1/5$ for $L = 2$. Figure 4 shows that an increase in the number of accountable symbols slightly changes the bandwidth, i.e. at $L = 2$ the information capacity is slightly larger than at $L = 1$.

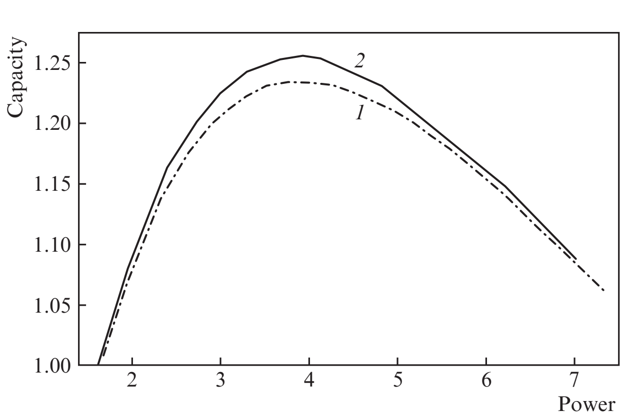

Figure 4. Dependences of the capacity on the signal power for (*1*) $L = 1$ and (*2*) $L = 2$.

6. Conclusions

We have calculated for the first time the channel capacity taking into account the pattern effect, i.e., a different contribution of the interaction of neighbouring pulses and self-interaction to the nonlinear noise. The capacity of the channels with different variants of the weight coefficients of nonlinear finite memory has been compared. It has been shown that for large signal powers and with the same total noise, the design of an optical system limiting nonlinear interactions with neighbouring bits is preferable. It has been found that an increase in the volume of accountable memory does not virtually affect the capacity of the channel with the same contribution of neighbouring symbols to nonlinear noise.

The results obtained can be useful in choosing the design of the communication line. The change in the dispersion of transmitted symbols with the help of various components of optical systems is an additional possibility of optimising the fibre link.

Acknowledgements. The work was supported by the Russian Science Foundation (Grant No. 16-11-10133, 2013–2016). E.G. Shapiro thanks the Ministry of Education for support (Grant No. 14. B25.31.0003). The work of D.A. Shapiro was supported by the RF President's Grants Council (State Support to the Leading Scientific Schools Programme, Grant No. NSh-6898.2016.2).

References

1. Splett A., Kurtzke C., Petermann K. *Proc. 19th ECOC* (Montereux, Switzerland, 1993) Vol. 2, pp 41–44.
2. Tang J. *J. Lightwave Technol.*, **20** (7), 1095 (2002).
3. Poggiolini P., Carena A., Curri V., Bosco G., Forghieri F. *IEEE Photon. Technol. Lett.*, **23**, 742 (2011).
4. Beygi L., Irukulapati N.V., Agrell E., Johanisson P., Karlsson M., Wymeersch H., Serena P., Bononi A. *Opt. Express*, **21** (22), 26376 (2013).
5. Gurkin N.V., Nanii O.E., Novikov A.G., Plaksin S.O., Treshchikov V.N., Ubaydullaev R.R. *Quantum Electron.*, **43** (6), 550 (2013) [*Kvantovaya Elektron.*, **43** (6), 550 (2013)].
6. Gurkin N.V., Konyshev V.A., Nanii O.E., Novikov A.G., Treshchikov V.N., Ubaydullaev R.R. *Quantum Electron.*, **45** (1), 69 (2015) [*Kvantovaya Elektron.*, **45** (1), 69 (2015)].
7. Yushko O.V., Nanii O.E., Redyuk A.A., Treshchikov V.N., Fedoruk M.P. *Quantum Electron.*, **45** (1), 75 (2015) [*Kvantovaya Elektron.*, **45** (1), 75 (2015)].
8. Konyshev V.A., Leonov A.V., Nanii O.E., Novikov A.G., Treshchikov V.N., Ubaydullaev R.R. *Opt. Commun.*, **349**, 19 (2015).
9. Konyshev V.A., Leonov A.V., Nanii O.E., Novikov A.G., Treshchikov V.N., Ubaydullaev R.R. *Quantum Electron.*, **46**, 1121 (2016) [*Kvantovaya Elektron.*, **46**, 1121 (2016)].
10. Shannon C. *Bell Syst. Tech. J.*, **27**, 379, 623 (1948).
11. Agrell E., Alvarado A., Durisi G., Karlsson G. *J. Lightwave Technol.*, **32**, 2862 (2014).
12. Shapiro E.G., Shapiro D.A. *Quantum Electron.*, **46** (12), 1117 (2016) [*Kvantovaya Elektron.*, **46** (12), 1117 (2016)].
13. Shapiro E.G., Shapiro D.A., Turitsyn S.K. *Opt. Express*, **23**, 15119 (2015).
14. Arimoto S. *IEEE Trans. Inf. Theory*, **18**, 14 (1972).