Log-log growth of channel capacity for nondispersive nonlinear optical fiber channel in intermediate power range: Extension of the model

A. V. Reznichenko,[1,2,*] A. I. Chernykh,[3,2,†] S. V. Smirnov,[2,‡] and I. S. Terekhov[1,2,§]

[1]*Budker Institute of Nuclear Physics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia*

[2]*Novosibirsk State University, Novosibirsk 630090, Russia*

[3]*Institute of Automation and Electrometry of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia*

(Received 2 October 2018; published 17 January 2019)

In our previous paper [Terekhov *et al.*, Phys. Rev. E **95**, 062133 (2017)] we considered the optical channel modeled by the nonlinear Schrödinger equation with zero dispersion and additive Gaussian noise. We found per-sample channel capacity for this model. In the present paper we extend the per-sample channel model by introducing the initial signal dependence on time and the output signal detection procedure. The proposed model is a closer approximation of the realistic communications link than the per-sample model where there is no dependence of the initial signal on time. For the proposed model we found the correlators of the output signal both analytically and numerically. Using these correlators we built the conditional probability density function. Then we calculated an entropy of the output signal, a conditional entropy, and the mutual information. Maximizing the mutual information we found the optimal input signal distribution, channel capacity, and their dependence on the shape of the initial signal in the time domain for the intermediate power range.

DOI: 10.1103/PhysRevE.99.012133

I. INTRODUCTION

Nonlinear communication channels have received a lot of attention in the last twenty years due to the development of fiber optical communication systems. In these communication systems the Kerr nonlinearity in the optical fiber becomes important when one increases the power of the transmitted signal. The problem of capacity finding was considered analytically and numerically in a series of papers; see, e.g., [1–11] and references therein. In spite of a lot of publications this problem has not been solved for the case of arbitrary Kerr nonlinearity and the second dispersion parameter of an optical fiber. The nondispersive model is much simpler than the case with nonzero dispersion but it catches-up the main features connected with nonlinearity. Also this model is more convenient for understanding the dependence of the capacity on the channel nonlinearity. So the analytical form of the conditional probability density function $P[Y|X]$, i.e., the probability density function (PDF) to receive the output signal Y if the input signal is X, for nondispersive per-sample channel was first obtained in Ref. [12]. The upper bound for the capacity at very large input signal power for the model was obtained in Refs. [13,14]. The capacity of the channel was found in Ref. [1] in the intermediate power range implying both large signal-to-noise ratio (SNR) and the condition for the next-to-leading corrections in the noise power to be small; see Eq. (23) in Ref. [1].

The per-sample model assumes that the input signal does not depend on time. In a realistic communication channel the transmitted signal does depend on time. In the recent paper [15] the influence of the receiver, signal, and noise bandwidth on the autocorrelation function and the capacity was discussed within the filter-and-sample model for the channel with zero dispersion. In our opinion, one of the important results of the paper [15] is understanding that the conditional PDF depends significantly on the properties of the receiver.

In this paper we consider the nondispersive channel in the intermediate power range in the case where the initial signal depends on time and has a bandwidth much less than the noise bandwidth. We also introduce a detection procedure which takes into account the time resolution characteristics of the detector and we demonstrate the influence of the detector and the noise bandwidth on statistical properties of the channel. Therefore this paper is a generalization of the previous results of Refs. [1,2] for the per-sample model to the time-dependant signal.

The paper is organized in the following way. In Sec. II we present the model of the signal propagation, the input signal, and the receiver model. In Sec. III we obtain the conditional probability density function for the introduced model. In Sec. IV we present numerical results for the correlators and compare these results with analytical ones. And in Sec. V we calculate the optimal input signal distribution and the channel capacity in the intermediate power range. In the conclusion we discuss our results.

II. MODEL OF THE SIGNAL PROPAGATION AND DETECTION

In our model the propagation of the signal $\psi(z,t)$ is described by the stochastic nonlinear Schrödinger equation

*a.v.reznichenko@inp.nsk.su
†chernykh@iae.nsk.su
‡smirnov@lab.nsu.ru
§i.s.terekhov@gmail.com

(NLSE) with zero dispersion:

$$\partial_z \psi - i\gamma|\psi|^2\psi = \eta(z,t), \tag{1}$$

where γ is the Kerr nonlinearity coefficient, the function $\psi(z,t)$ obeys the input and output conditions $\psi(z=0,t) = X(t)$ and $\psi(z=L,t) = Y(t)$, respectively; L is the length of the signal propagation; $\eta(z,t)$ is an additive complex noise with zero mean, $\langle\eta(z,t)\rangle_\eta = 0$, and the correlation function in the frequency domain,

$$\langle\eta(z,\omega)\bar{\eta}(z',\omega')\rangle_\eta = 2\pi Q\delta(\omega-\omega')\theta\left(\frac{W'}{2} - |\omega|\right) \times \delta(z-z'), \tag{2}$$

where the bar means complex conjugation; Q is a power of the noise per unit length and per unit frequency, $\theta(\omega)$ is the Heaviside theta function, $\delta(\omega)$ is the Dirac delta function, and W' is the bandwidth of the noise. The noise $\eta(z,\omega)$ is not white due to limited bandwidth. In the time domain this correlator has the form

$$\langle\eta(z,t)\bar{\eta}(z',t')\rangle_\eta = Q\frac{W'}{2\pi}\mathrm{sinc}\left(\frac{W'(t-t')}{2}\right)\delta(z-z'). \tag{3}$$

One can see that if the time difference $t - t' = 2n\pi/W'$ then the correlator (3) is equal to zero; here n is an integer. Thus we can solve Eq. (1) independently for parameters $t_j = j\Delta$ for different integer j, where $\Delta = 2\pi/W'$ is the time grid spacing. Therefore instead of the continuous time model (1) we will consider the following discrete model:

$$\partial_z\psi(z,t_j) - i\gamma|\psi(z,t_j)|^2\psi(z,t_j) = \eta(z,t_j) \tag{4}$$

for any time moment t_j. This means that we obtain the set of independent time channels since the noise in these moments is not correlated. We present the input and output conditions in the discrete form as well: $\psi(z=0,t_j) = X(t_j)$ and $\psi(z=L,t_j) = Y(t_j)$. Note that the solution $\Phi(z,t_j)$ of Eq. (4) with zero noise which obeys the input condition $\Phi(z=0,t_j) = X(t_j)$ has the form

$$\Phi(z,t_j) = X(t_j)e^{i\gamma z|X(t_j)|^2}. \tag{5}$$

Below we assume that the frequency bandwidth W' of the noise is much broader than the frequency bandwidth W of the input signal $X(t)$ and the frequency bandwidth $\widetilde{W}$ of the function $\Phi(z=L,t)$.

In our model the input signal $X(t)$ has the form

$$X(t) = \sum_{k=-N}^{N} C_k\, f(t-kT_0), \tag{6}$$

where C_k are complex random coefficients with some probability density function $P_X[\{C\}]$, $\{C\} = \{C_{-N},\ldots,C_N\}$; the pulse envelope $f(t)$ is the real function which is normalized as $\int_{-\infty}^{\infty}\frac{dt}{T_0}f^2(t) = 1$. The pulse envelope $f(t)$ has the following properties: the overlapping of the functions $f(t-kT_0)$ and $f(t-mT_0)$ for $k\neq m$ is negligible: $\int_{-\infty}^{\infty}dtf(t-kT_0)f(t-mT_0)\approx 0$. This means that the function $f(t)$ has almost the finite support $[-T_0/2, T_0/2]$, and the input signal $X(t)$ is defined on the interval $T = (2N+1)T_0$. Thus the frequency support of the function $X(t)$ is infinite. But we imply that $\int_W|X(\omega)|^2d\omega \approx \int_{W'}|X(\omega)|^2d\omega$, where $X(\omega)$ is the Fourier transformation of $X(t)$. The last relation means that $T_0W \gg 1$.

In our consideration the average input signal power P is fixed:

$$P = \int\left(\prod_{k=-N}^{N}d^2C_k\right)P_X[\{C\}]\int_{-\infty}^{\infty}\frac{dt}{T}|X(t)|^2, \tag{7}$$

where $d^2C_k = d\,\mathrm{Re}\,C_k d\,\mathrm{Im}\,C_k$, and the input signal probability density function $P_X[\{C\}]$ is normalized as follows:

$$\int\left(\prod_{k=-N}^{N}d^2C_k\right)P_X[\{C\}] = 1. \tag{8}$$

Using the properties of the function $f(t-kT_0)$ we can rewrite Eq. (7):

$$P = \int d^2C_m P_X^{(m)}[C_m]|C_m|^2, \tag{9}$$

where

$$P_X^{(m)}[C_m] = \int\left(\prod_{k=-N,k\neq m}^{N}d^2C_k\right)P_X[\{C\}], \tag{10}$$

and we imply that the distribution $P_X^{(m)}[C_m]$ does not depend on m.

Let us describe the output signal detection procedure. Our detector recovers the information which is carried by the coefficients $\{C_k\}$. First, the detector receives the signal $\psi(z=L,t_j)$ at the discrete time moments $t_j = j\Delta$; here $j = -M,\ldots,M-1$, where $M = T/(2\Delta) \gg N$. This means that the time resolution of the detector coincides with the time discretization Δ. Since $\Delta \ll 1/\widetilde{W}$ our detector can completely recover the input signal in the noiseless case. Second, the detector removes the nonlinear phase to obtain the recovered input signal $\widetilde{X}(t)$ in the following form:

$$\widetilde{X}(t_j) = \psi(z=L,t_j)e^{-i\gamma L|\psi(z=L,t_j)|^2}. \tag{11}$$

And finally, using $\widetilde{X}(t)$ the detector recovers the coefficients $\widetilde{C}_k$ by projecting on the basis functions $f(t-kT_0)$:

$$\widetilde{C}_k = \frac{1}{T_0}\int_{-\infty}^{\infty}dtf(t-kT_0)\widetilde{X}(t) \approx \frac{\Delta}{T_0}\sum_{j=-M}^{M-1}f(t_j-kT_0)\widetilde{X}(t_j). \tag{12}$$

One can check that in the case of zero noise $\widetilde{X}(t) = X(t)$ and $\widetilde{C}_k = C_k$.

III. STATISTICS OF $\widetilde{C}_k$

In the previous paper [1] we obtained the conditional probability function $P[Y|X]$ for the case where input X and output Y signals do not depend on time (per-sample conditional PDF). In the previous section we extend our model [1] by including the detector procedure and time dependence of the input signal $X(t)$. Our goal is to obtain the conditional probability function $P[\{\widetilde{C}\}|\{C\}]$, i.e., the probability to detect the set of coefficients $\{\widetilde{C}\}$ if the transmitted set is $\{C\}$. Using

the function $P[\{\widetilde{C}\}|\{C\}]$ we can calculate the probability density function $P_{\rm out}[\{\widetilde{C}\}]$ as

$$P_{\rm out}[\{\widetilde{C}\}] = \int \prod_{k=-N}^{N} d^2C_k P[\{\widetilde{C}\}|\{C\}]P_X[\{C\}]. \quad (13)$$

Since the propagation of the signal in the different time moments t_j is independent and noise is not correlated, the conditional probability function $P[Y(t)|X(t)]$, i.e., the probability density to obtain the output signal $Y(t)$ for the given input signal $X(t)$, can be presented in the factorized form:

$$P[Y(t)|X(t)] = \prod_{j=-M}^{M-1} P_j[Y_j|X_j], \quad (14)$$

where $X_j = X(t_j)$, $Y_j = Y(t_j)$, and $P_j[Y_j|X_j]$ is per-sample conditional PDF obtained in Ref. [1]. The function $P_j[Y_j|X_j]$ in the leading and next-to-leading order in parameter $\sqrt{Q}$ can be deduced from the results of Ref. [1], where we have to replace parameter Q by Q/Δ:

$$\begin{aligned} P_j[Y_j|X_j] = \Delta \frac{\exp\left\{-\Delta \dfrac{(1+4\mu_{(j)}^2/3)x_{(j)}^2 - 2\mu_{(j)}x_{(j)}y_{(j)} + y_{(j)}^2}{QL(1+\mu_{(j)}^2/3)}\right\}}{\pi QL\sqrt{1+\mu_{(j)}^2/3}} \\ \times \left\{1 - \frac{\mu_{(j)}/\rho_{(j)}}{15(1+\mu_{(j)}^2/3)^2}\left[\mu_{(j)}(15+\mu_{(j)}^2)x_{(j)} - 2(5-\mu_{(j)}^2/3)y_{(j)}\right] - \frac{\mu_{(j)}\Delta}{135QL\rho_{(j)}(1+\mu_{(j)}^2/3)^3}\right. \\ \times \left[\mu_{(j)}(4\mu_{(j)}^4 + 15\mu_{(j)}^2 + 225)x_{(j)}^3 + (23\mu_{(j)}^4 + 255\mu_{(j)}^2 - 90)x_{(j)}^2 y_{(j)}\right. \\ \left.\left. + \mu_{(j)}(20\mu_{(j)}^4 + 117\mu_{(j)}^2 - 45)x_{(j)}y_{(j)}^2 - 3(5\mu_{(j)}^4 + 33\mu_{(j)}^2 + 30)y_{(j)}^3\right]\right\}. \end{aligned} \quad (15)$$

Here $\rho_{(j)} = |X_j|$, $X_j = \rho_{(j)}e^{i\phi_{(j)}}$, $\mu_{(j)} = \gamma L\rho_{(j)}^2$, $x_{(j)} = {\rm Re}[Y_j e^{-i\phi_{(j)} - i\mu_{(j)}} - \rho_{(j)}]$, and $y_{(j)} = {\rm Im}[Y_j e^{-i\phi_{(j)} - i\mu_{(j)}} - \rho_{(j)}]$. The expression (15) was obtained in Ref. [1] on the condition that the average input signal power P lies in the intermediate power range:

$$\frac{QL}{\Delta} \ll P \ll \Delta/(QL^3\gamma^2), \quad (16)$$

where $P = 2\pi\int_0^\infty d\rho\rho^3 P[\rho]$, and $P[\rho]$ is the distribution function of the quantity ρ; see Ref. [1]. Therefore, our consideration is restricted by the condition (16). The factorization of $P[Y(t)|X(t)]$ in the form (14) means that there are $2M$ independent "subchannels." Note that the signal $X(t)$ is completely defined by $2N+1$ coefficients C_k; i.e., there are only $2N+1$ independent X_j, but all $2M$ quantities Y_j are independent. However, our detector reduces the function $Y(t)$ to the set of $2N+1$ coefficients $\{\widetilde{C}_k\}$ by the procedure (11) and (12). Therefore we have to reduce the function $P[Y(t)|X(t)]$ to the function $P[\{\widetilde{C}_k\}|\{C_k\}]$ by integrating over $2M-2N-1$ redundant degrees of freedom. Using the conditional PDF $P[Y(t)|X(t)]$ in the form (14) one can calculate all correlators of the coefficients $\widetilde{C}_k$: $\langle\widetilde{C}_{k_1}\rangle$, $\langle\widetilde{C}_{k_1}\widetilde{C}_{k_2}\rangle$, $\langle\widetilde{C}_{k_1}\ldots\widetilde{C}_{k_n}\rangle$. Here

$$\langle\widetilde{C}_{k_1}\ldots\widetilde{C}_{k_n}\rangle = \int \prod_{j=-M}^{M-1} d^2Y_j P[Y(t)|X(t)]\widetilde{C}_{k_1}\ldots\widetilde{C}_{k_n}, \quad (17)$$

where $d^2Y_j = d\,{\rm Re}\,Y_j d\,{\rm Im}\,Y_j$; $\widetilde{C}_k$ is defined in Eq. (12), and in the discrete form it reads

$$\widetilde{C}_k = \frac{\Delta}{T_0}\sum_{j=-M}^{M-1} f(t_j - kT_0)Y_j e^{-i\gamma L|Y_j|^2}. \quad (18)$$

To recover the function $P[\{\widetilde{C}_k\}|\{C_k\}]$ in the leading approximation in parameter Q it is necessary to know only three correlators: $\langle\widetilde{C}_k\rangle$, $\langle\widetilde{C}_k\widetilde{C}_m\rangle$, $\langle\widetilde{C}_k\overline{\widetilde{C}}_m\rangle$. After substitution of Eqs. (14), (15), and (18) into Eq. (17) and performing the integration we obtain in the leading order in the noise parameter Q

$$\langle\widetilde{C}_k\rangle = C_k - \frac{iC_k\,QL^2\gamma}{\Delta}\left(1 - \frac{i\gamma L|C_k|^2 n_4}{3}\right), \quad (19)$$

$$\begin{aligned} &\langle(\widetilde{C}_m - \langle\widetilde{C}_m\rangle)(\widetilde{C}_n - \langle\widetilde{C}_n\rangle)\rangle \\ &\quad = -i\delta_{m,n}\frac{C_m^2 QL^2\gamma}{T_0}\left(n_4 - \frac{2in_6}{3}\gamma L|C_m|^2\right), \quad (20) \\ &\langle(\widetilde{C}_m - \langle\widetilde{C}_m\rangle)\overline{(\widetilde{C}_n - \langle\widetilde{C}_n\rangle)}\rangle \\ &\quad = \delta_{m,n}\frac{QL}{T_0}\left(1 + \frac{2n_6}{3}\gamma^2L^2|C_m|^4\right), \quad (21) \end{aligned}$$

where $\delta_{m,n}$ is Kronecker symbol and

$$n_s = \int_{-T_0/2}^{T_0/2}\frac{dt}{T_0}f^s(t). \quad (22)$$

Note that for the first correlator $(\langle\widetilde{C}_k - C_k\rangle)$ is proportional to $QL/\Delta = QLW'/(2\pi)$, i.e., it is proportional to the total noise power, whereas the correlators (20) and (21) are proportional to QL/T_0 and do not depend on the discretization parameter Δ only in leading order in parameter Q and depend on the parameter Δ in higher order corrections in parameter Q; see the Appendix.

Using the correlators (19)–(21) we obtain the conditional PDF $P[\widetilde{C}|C]$ in the leading order in parameter Q:

$$P[\widetilde{C}|C] = \prod_{m=-N}^{N} P_m[\widetilde{C}_m|C_m], \tag{23}$$

where

$$P_m[\widetilde{C}_m|C_m] \approx \frac{T_0}{\pi QL\sqrt{1+\xi^2\mu_m^2/3}} \exp\left[-T_0 \frac{\left[1+4n_6\mu_m^2/(3)\right]x_m^2 + 2x_m y_m \mu_m n_4 + y_m^2}{QL\left(1+\xi^2\mu_m^2/3\right)}\right]. \tag{24}$$

Here we have introduced the notations

$$x_m = \mathrm{Re}\left\{e^{-i\phi_m}\left[\widetilde{C}_m - C_m + \frac{iC_m\,QL^2\gamma}{\Delta}\left(1 - \frac{i\gamma L|C_m|^2 n_4}{3}\right)\right]\right\}, \tag{25}$$

$$y_m = \mathrm{Im}\left\{e^{-i\phi_m}\left[\widetilde{C}_m - C_m + \frac{iC_m\,QL^2\gamma}{\Delta}\left(1 - \frac{i\gamma L|C_m|^2 n_4}{3}\right)\right]\right\}, \tag{26}$$

$$\phi_m = \arg C_m, \quad \mu_m = \gamma L|C_m|^2, \tag{27}$$

$$\xi^2 = \left(4n_6 - 3n_4^2\right). \tag{28}$$

The parameter ξ^2 obeys the inequality $\xi^2 > n_6 > 0$ due to the Cauchy-Schwarz-Buniakowski inequality. Note that the function $P[\widetilde{C}|C]$ has the factorized form (23) only in the leading approximation in the parameter Q. Equation (23) means that we have $2N+1$ independent information channels, and the channel corresponding to the time slot m is described by the function $P_m[\widetilde{C}_m|C_m]$. The function $P_m[\widetilde{C}_m|C_m]$ obeys the normalization condition

$$\int d^2\widetilde{C}_m\, P_m[\widetilde{C}_m|C_m] = 1. \tag{29}$$

Since there are $2N+1$ independent channels, we can choose the input signal distribution $P_X[\{C_m\}]$ in the factorized form:

$$P_X[\{C\}] = \prod_{k=-N}^{N} P_X^{(k)}[C_k], \tag{30}$$

and we can consider only one channel, say the mth channel. For this channel we can calculate the probability distribution function of the coefficients $\widetilde{C}_m$:

$$P_{\mathrm{out}}^{(m)}[\widetilde{C}_m] = \int d^2C_m\, P_m[\widetilde{C}_m|C_m] P_X^{(m)}[C_m]. \tag{31}$$

We imply that the function $P_X^{(m)}[C_m]$ is a smooth function that changes on a scale $|C_m|^2 \sim P$ which is much greater than QL/Δ:

$$P \gg QL/\Delta \gg QL/T_0. \tag{32}$$

In other words, the signal power is much greater than the noise power in the channel. The variation scale of the function $P_m[\widetilde{C}_m|C_m]$ in the variable C_m is of order $\sqrt{QL/T_0}$; therefore we can use Laplace's method [16] for the calculation of the integral (31). Performing the integration in the leading order in parameter Q we obtain

$$P_{\mathrm{out}}^{(m)}[\widetilde{C}_m] \approx P_X^{(m)}[\widetilde{C}_m]; \tag{33}$$

for details see Appendix C in Ref. [1]. The result (33) implies that the statistics of the coefficients $\widetilde{C}_m$ coincides with the statistics of the coefficients C_m.

IV. NUMERICAL CALCULATIONS OF THE CORRELATORS

In order to verify the analytical results we performed numerical simulations of pulse propagation through nonlinear nondispersive optical fibers and calculated correlators (19), (20), and (21). For these purposes we solve numerically Eq. (1) for fixed input signal $X(t)$ and for different realizations of the noise $\eta(z,t)$. Then we numerically perform the detection procedure described by Eqs. (11) and (12). Finally, we average the coefficients $\widetilde{C}_k$ and their quadratic combinations over noise realizations. In our simulations we use two numerical methods of the solution of Eq. (1): the split-step Fourier method and Runge-Kutta method of the fourth order. The results are presented in the following subsections. We have checked that the numerical results do not depend on the numerical method and these results are consistent with analytical ones for different realizations of the form $f(t)$ of the input pulse.

For numerical simulation we choose the following realistic channel parameters. The duration of one pulse is $T_0 = 10^{-10}$ sec, the fiber length is equal to $L = 800$ km, and the Kerr nonlinearity parameter is $\gamma = 1.25$ (km W)$^{-1}$.

A. Split-step Fourier method

Equation (1) was integrated numerically over z from 0 up to communication line length L using the split-step Fourier method [17,18]:

$$\psi(z+h,t) = \psi(z,t)\exp[i\gamma|\psi(z,t)|^2 h] + \hat{F}_-\{\delta Q_h\}, \tag{34}$$

where $\psi(z,t)$ stands for the numerical solution of (1), h is a step size of the z mesh, and $\hat{F}_-$ denotes the discrete inverse Fourier transform. The quantity δQ_h stands for the noise addition per step h which is made in the frequency domain according to

$$\psi(z,\omega_j) \to \psi(z,\omega_j) + \sqrt{\frac{hQ}{T}}\frac{\eta_X + i\eta_Y}{\sqrt{2}}, \tag{35}$$

where $j = 0, \ldots, 2M-1$ stands for the index of the ω mesh, $2M$ is the number of t- and ω-mesh points, T is the total width of the t mesh [we choose $T = 63T_0$; see Eq. (36) below], η_X and η_Y are independent standard Gauss random numbers with zero mean and $\sigma^2 = 1$, and the additive noise level is $Q = 10^{-21}$ W/(km Hz).

The input signal for $z = 0$ has the form

$$\psi(z=0,t) = X(t) = \sum_{k=-31}^{31} C_k\, f(t - kT_0); \qquad (36)$$

here we use the pulse envelope of the Gaussian form:

$$f(t) = \sqrt{\frac{T_0}{T_1\sqrt{\pi}}}\exp\left(-\frac{t^2}{2T_1^2}\right), \qquad (37)$$

where $T_1 = T_0/10 = 10^{-11}$ sec stands for the characteristic timescale of the function $f(t)$. The pulse intersection is negligible. For such pulses coefficients n_s defined in Eq. (22) are $n_4 = \frac{T_0/T_1}{\sqrt{2\pi}} \approx 3.989$, $n_6 = \frac{(T_0/T_1)^2}{\pi\sqrt{3}} \approx 18.38$, $n_8 = \frac{(T_0/T_1)^3}{2\pi\sqrt{\pi}} \approx 89.79$; $\xi \approx 5.08$.

In the numerical simulation we vary the average power $\frac{1}{63}\sum_{k=-31}^{31}|C_k|^2$ of the input signal from 0.0177 mW up to 4.43 mW. This corresponds to the variation of the peak power $[|C_k|^2 f^2(0)]$ from 0.1 mW up to 25 mW.

Simulations are performed for different t meshes (different grid spacing Δ), i.e., for different noise bandwidths and fixed noise parameter Q. These meshes differ from each other by time grid spacing $\Delta = T/(2M)$: $\Delta_1 = 9.77 \times 10^{-14}$ sec, $\Delta_2 = 1.95 \times 10^{-13}$ sec, and $\Delta_3 = 3.91 \times 10^{-13}$ sec. These grid spacings determine the widths of conjugated ω meshes: $1/\Delta_1 = 10.26$ THz, $1/\Delta_2 = 5.12$ THz, and $1/\Delta_3 = 2.56$ THz.

For each average power of the signal and each mesh step size Δ we simulate propagation of the signal for different realizations of the noise and then average the obtained results for correlators over realizations. The total number of noise realizations for fixed $X(t)$, see Eq. (36), is determined by the necessary statistical relative error and is chosen as 5.0×10^4. This number of the realizations corresponds to the statistical relative error for correlators (20) and (21) on the level of 0.2% (since the total number of pulses is $63 \times 5.0 \times 10^4 \approx 3.2 \times 10^6$). We performed simulations on z meshes with different numbers of points (100, 200, 400, 800) and checked that the results do not depend on step size h.

In Figs. 1–5 the numerical and analytical results for correlators (19)–(21) are presented for different time grid spacings Δ as a function of input signal power. In Fig. 5 the results are presented for the grid spacings Δ_1 and Δ_3 because the results for Δ_1 and Δ_2 almost coincide. One can see that numerical and analytical results are in a good agreement up to 3 mW at least. However the difference between numerical and analytical results for the smallest time grid spacing Δ_1 is maximal. Decreasing of the parameter Δ means the increasing of the spectral bandwidth of the noise. This increasing results in the growth of the total noise power received by the detector. Note that the analytical expressions for correlators were obtained using the conditional PDF $P[Y(t)|X(t)]$ in the form (15). This form was derived in the approximation of large signal-to-noise ratio SNR $= P\Delta/(QL)$. Decreasing parameter Δ we diminish the parameter SNR and, as a consequence, the accuracy of our approximation. The difference between numerical and analytical results can be explained by taking into account the next-to-leading order (NLO) corrections in noise power parameter QL/Δ. The analytical results in Figs. 3–5 are shown with taking into account both leading order results (20)–(21) and NLO corrections presented in the Appendix; see Eqs. (A1) and (A3).

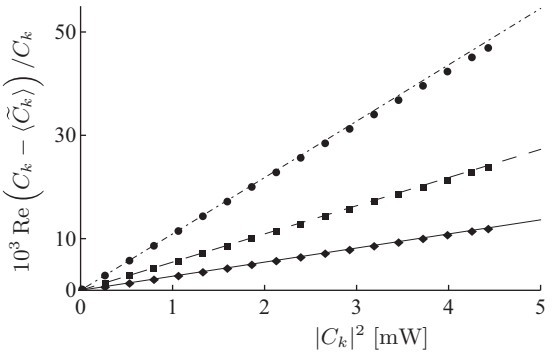

FIG. 1. The real part of the relative difference of the coefficient C_k and the correlator (19) in units of 10^{-3} as a function of input signal power $|C_k|^2$ for $f(t)$ from Eq. (37). The noise power parameter is $Q = 10^{-21}$ W/(km Hz). Dashed-dotted, dashed, and solid lines correspond to analytic representation (19) for time grid spacings Δ_1, Δ_2, Δ_3, respectively. Circles, squares, and diamonds correspond to numerical results for time grid spacings Δ_1, Δ_2, Δ_3, respectively.

B. Runge-Kutta method

For Eq. (1) the time t is the incoming parameter. Thus the simulation consists of the solution of the ordinary differential equation with various initial conditions determined by the real pulse shape $f(t)$, the amplitude C_m, and independent random

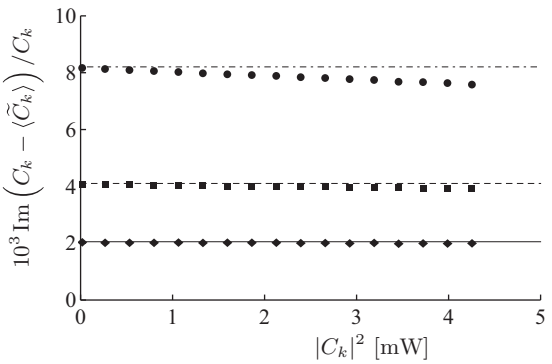

FIG. 2. The imaginary part of the relative difference of the coefficient C_k and the correlator (19) in units of 10^{-3} as a function of input signal power $|C_k|^2$ for $f(t)$ from Eq. (37). The noise power parameter is $Q = 10^{-21}$ W/(km Hz). Dashed-dotted, dashed, and solid lines correspond to analytic representation (19) for time grid spacings Δ_1, Δ_2, Δ_3, respectively. Circles, squares, and diamonds correspond to numerical results for time grid spacings Δ_1, Δ_2, Δ_3, respectively.

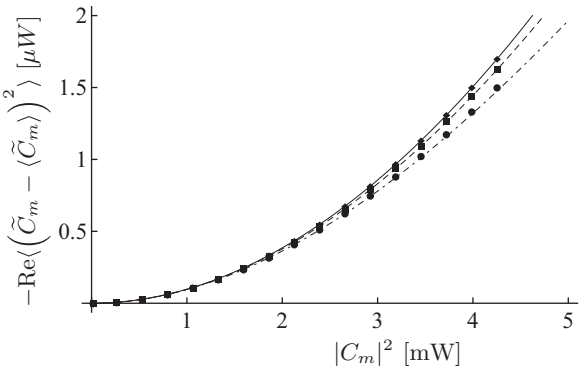

FIG. 3. The real part of the correlator (20) multiplied by (-1) as a function of input signal power $|C_m|^2$ for $f(t)$ from Eq. (37). The noise power parameter is $Q = 10^{-21}$ W/(km Hz). Dashed-dotted, dashed, and solid lines correspond to analytic representation (20) with NLO corrections (A1) for time grid spacings Δ_1, Δ_2, Δ_3, respectively. Circles, squares, and diamonds correspond to numerical results for time grid spacings Δ_1, Δ_2, Δ_3, respectively.

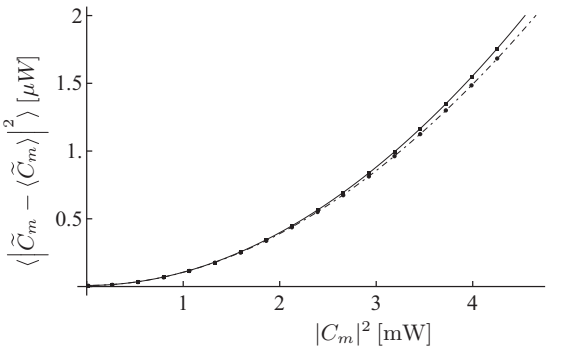

FIG. 5. The correlator (21) as a function of input signal power $|C_m|^2$ for $f(t)$ from Eq. (37). The noise power parameter is $Q = 10^{-21}$ W/(km Hz). Dashed-dotted and solid lines correspond to analytic representation (21) with NLO corrections (A3) for time grid spacings Δ_1, Δ_3, respectively. Circles and squares correspond to numerical results for time grid spacings Δ_1, Δ_3, respectively.

noise functions $\eta(z,t)$. In the second method we used pulse envelopes of the form

$$f_n(t) = A_n \cos^n(\pi t/T_0) \tag{38}$$

for $n = 2, 4$ and $t \in [-T_0/2,\ T_0/2]$: $A_2 = \sqrt{\frac{8}{3}}$ and $A_4 = \sqrt{\frac{128}{35}}$. We choose the time discretization parameter $\Delta = T_0/64$. The random noise was realized as the telegraph process with a step of the length $\Delta_z = 10^{-4}L$ and of the random height with zero average and with the dispersion $\sigma^2 = 2.38 \times 10^{-8}$ W/(km^2) both for real and imaginary parts. The noise power parameter reads as $Q = 2\sigma^2\Delta\Delta_z \approx 5.94 \times 10^{-21}$ W/(km Hz) and it is almost six times greater than that in the previous method. We independently control this parameter Q by using the leading order contribution to the correlator (21) numerically simulated for $\gamma = 0$. The noise η is constant within the step. Within the step Eq. (1) was solved by the Runge-Kutta method of the fourth order with the step $h = \Delta_z/50$. The recovered input signal $\widetilde{X}(t_j)$ was calculated using Eq. (11) at the equidistant points t_j. The coefficients $\widetilde{C}_k$ were calculated using Eq. (12). The average (19) and correlators (20) and (21) were calculated over 16 384 values of $\widetilde{C}_k$ which were found for various noise realizations.

To control the accuracy of the method we solved Eq. (1) with zero noise from $z = 0$ to $z = L$ with the step h and then we performed the backward propagation from $z = L$ to $z = 0$ with the found solution as the initial condition. In the procedure the input signal was recovered with the relative precision equal to 10^{-6}.

The analytical results in comparison with the numerical results are presented in Figs. 6–10 for different pulse envelopes. The numerical results are presented by dots with statistical errors on the level of three standard deviations. And the curves correspond to the analytical expressions obtained from Eqs. (19)–(21) with corrections (A1) and (A3). In Figs. 6 and 7 we plot the real and imaginary parts of the relative

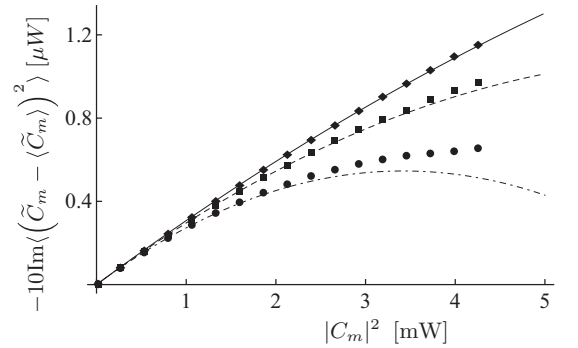

FIG. 4. The imaginary part of the correlator (20) multiplied by (-10) as a function of input signal power $|C_m|^2$ for $f(t)$ from Eq. (37). The noise power parameter is $Q = 10^{-21}$ W/(km Hz). Dashed-dotted, dashed, and solid lines correspond to analytic representation (20) with NLO corrections (A1) for time grid spacings Δ_1, Δ_2, Δ_3, respectively. Circles, squares, and diamonds correspond to numerical results for time grid spacings Δ_1, Δ_2, Δ_3, respectively.

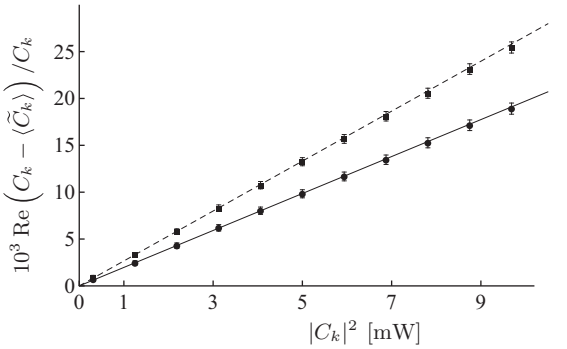

FIG. 6. The real part of the relative difference of the coefficient C_k and the correlator (19) in units of 10^{-3} as a function of input signal power $|C_k|^2$ for $f_2(t)$, see black solid line, and for $f_4(t)$, see black dashed line. The noise power parameter is $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles and rectangles correspond to numerical results with statistical error on the level of three standard deviations for the functions f_2 and f_4, respectively.

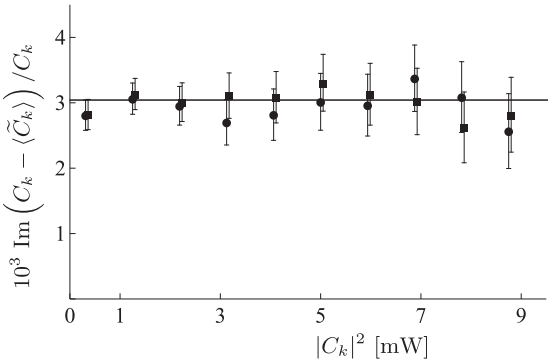

FIG. 7. The imaginary part of the relative difference of the coefficient C_k and the correlator (19) in units of 10^{-3} as a function of input signal power $|C_k|^2$ for $f_2(t)$ and $f_4(t)$; see black solid line. The noise power parameter is $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles and rectangles correspond to numerical results with statistical error on the level of three standard deviations for the functions f_2 and f_4, respectively.

difference $(C_k - \langle\widetilde{C}_k\rangle)/C_k$ as a function of input signal power $|C_k|^2$ for two envelope functions $f_2(t)$ and $f_4(t)$. One can see that in Fig. 6 the curve corresponding to the envelope $f_4(t)$ is disposed above one corresponding to the envelope $f_2(t)$. The reason is follows: the envelope $f_4(t)$ is more picked than the envelope $f_2(t)$; this results in the integrals n_4 and n_6 for $f_4(t)$ being greater than those for the envelope $f_2(t)$. In Fig. 7 the results for the both envelopes coincide as predicted by Eq. (19). In Figs. 8–10 we plot the real, imaginary parts of $-\langle(\widetilde{C}_m - \langle\widetilde{C}_m\rangle)^2\rangle$ and the absolute value of $\langle|\widetilde{C}_m - \langle\widetilde{C}_m\rangle|^2\rangle$, respectively. Similarly to Fig. 6 the curves corresponding to the envelope $f_4(t)$ are disposed above ones corresponding to envelope $f_2(t)$ for the same reason. One can see that numerical and analytical results are in a good agreement.

We have checked that for the same envelope function $f(t)$ both numerical methods (the split-step Fourier method and Runge-Kutta method) give the coinciding results with the

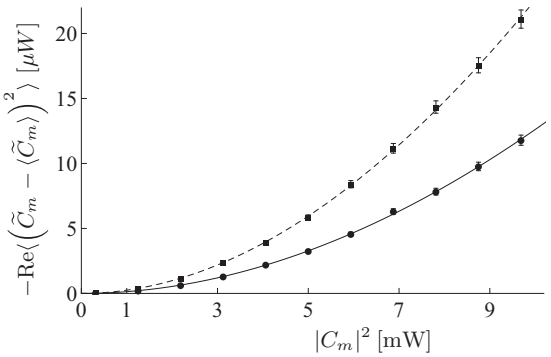

FIG. 8. The real part of the correlator (20) multiplied by (-1) as a function of input signal power $|C_m|^2$ for $f_2(t)$, see black solid line, and for $f_4(t)$, see black dashed line. Solid and dashed lines correspond to the real part of leading order contribution (20) with the next-to-leading order corrections; see Eq. (A1). The noise power parameter is $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles and rectangles correspond to numerical results with statistical error on the level of three standard deviations for the functions f_2 and f_4, respectively.

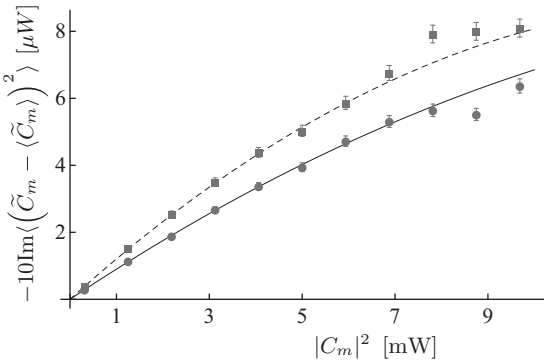

FIG. 9. The imaginary part of the correlator (20) multiplied by (-10) as a function of input signal power $|C_m|^2$ for $f_2(t)$, see black solid line, and for $f_4(t)$, see black dashed line. Solid and dashed lines correspond to the imaginary part of leading order contribution (20) with the next-to-leading order corrections; see Eq. (A1). The noise power parameter is $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles and rectangles correspond to numerical results with statistical error on the level of three standard deviations for the functions f_2 and f_4, respectively.

accuracy of the statistical errors. Therefore we have presented the numerical calculation using the split-step Fourier method for the Gaussian envelope function $f(t)$, see Eq. (37), and the Runge-Kutta method for envelope functions $f_2(t)$ and $f_4(t)$, see Eq. (38).

V. ENTROPIES AND MUTUAL INFORMATION

Now we proceed to the calculation of the output signal entropy

$$H[\widetilde{C}_m] = -\int d^2\widetilde{C}_m P^{(m)}_{\mathrm{out}}[\widetilde{C}_m]\log P^{(m)}_{\mathrm{out}}[\widetilde{C}_m], \tag{39}$$

conditional entropy

$$H[\widetilde{C}_m|C_m] = -\int d^2\widetilde{C}_m d^2C_m\, P_m[\widetilde{C}_m|C_m] \times P^{(m)}_X[C_m]\log P_m[\widetilde{C}_m|C_m], \tag{40}$$

and the mutual information

$$I_{P^{(m)}_X} = H[\widetilde{C}_m] - H[\widetilde{C}_m|C_m]. \tag{41}$$

Our calculations of the entropies (39) and (40) and the mutual information (41) are similar to calculations of the entropies and the mutual information for the per-sample channel; see Secs. III and IV of Ref. [1]. Therefore we will not repeat the similar calculations here and present only the final results:

$$H[\widetilde{C}_m] = \quad H[C_m] = -\int d^2C_m P^{(m)}_X[C_m]\log P^{(m)}_X[C_m], \tag{42}$$

$$H[\widetilde{C}_m|C_m] = 1 + \log\left[\pi\frac{QL}{T_0}\right] + \frac{1}{2}\int d^2C_m P^{(m)}_X[C_m] \times \log\left[1+\xi^2\frac{\gamma^2L^2|C_m|^4}{3}\right]. \tag{43}$$

To calculate the optimal input signal distribution $P_{\rm opt}^{(m)}[C_m]$ we calculate the mutual information substituting Eqs. (42) and (43) into Eq. (41); then we vary the mutual information over $P_X^{(m)}[C_m]$ with taking into account the normalization condition (8) and the fixed average power (9). Assuming the variation of the mutual information to be zero, we obtain the equation for the optimal input signal distribution $P_{\rm opt}^{(m)}[C_m]$. We solve the equation and obtain (for details of the similar calculations for the per-sample channel see Sec. III of Ref. [1])

$$P_{\rm opt}^{(m)}[C_m] = N_0 \frac{e^{-\lambda_0 |C_m|^2}}{\sqrt{1+\xi^2\gamma^2 L^2 |C_m|^4/3}}, \tag{44}$$

where parameters $N_0 = N_0(P, \xi\gamma)$ and $\lambda_0 = \lambda_0(P, \xi\gamma)$ are functions of the power P and modified nonlinearity parameter $\xi\gamma$ by virtue of the relations (compare with Eqs. (46) and (47) of Ref. [1])

$$\int d^2 C_m P_{\rm opt}^{(m)}[C_m] = \int_0^\infty d\rho \frac{2\pi N_0 \rho\, e^{-\lambda_0 \rho^2}}{\sqrt{1+\xi^2\gamma^2 L^2 \rho^4/3}} = 1, \tag{45}$$

$$P = \int d^2 C_m P_{\rm opt}^{(m)}[C_m] |C_m|^2 = \int_0^\infty d\rho \frac{2\pi N_0 \rho^3 e^{-\lambda_0 \rho^2}}{\sqrt{1+\xi^2\gamma^2 L^2 \rho^4/3}}. \tag{46}$$

Note that in Ref. [14] the authors used the half-Gaussian distribution for the input signal as the optimal one for the per-sample channel. One can see that the result (44) for the channel under consideration and the optimal input signal distribution (45) in Ref. [1] for the per-sample channel differ from the half-Gaussian distribution.

The capacity of one channel m, i.e., the mutual information calculated using the optimal input signal distribution (44), reads

$$C = I_{P_{\rm opt}^{(m)}} = \log\left(\frac{PT_0}{\pi e QL}\right) + P\lambda_0 - \log[PN_0]. \tag{47}$$

One can see that the first term on the right-hand side of Eq. (47) corresponds to Shannon's result [19] for the linear channel at large signal-to-noise ratio; the second and third terms are related to the nonlinearity impact. The result (47) is similar to that obtained for the per-sample model in Ref. [1] but with modification of the Kerr nonlinearity parameter γ for the per-sample model to parameter $\xi\gamma$ for the present model, where $\xi = \sqrt{4n_6 - 3n_4^2}$. There is no simple analytical form for N_0 and λ_0, see Secs. III and IV of Ref. [1], and therefore we present below the analytical results for the asymptotics of the mutual information for small and large dimensionless nonlinearity parameters $\xi\gamma LP$ and the numerical calculations in Fig. 11.

Performing the substitution $\gamma \to \xi\gamma$ in the results of Secs. III and IV of Ref. [1] we arrive at the following asymptotics of the mutual information for small and large dimensionless nonlinearity parameters γLP:

$$I_{P_X^{\rm opt}[X]} \approx \log\left(\frac{PT_0}{QL}\right) - \frac{\xi^2\gamma^2 L^2 P^2}{3}, \tag{48}$$

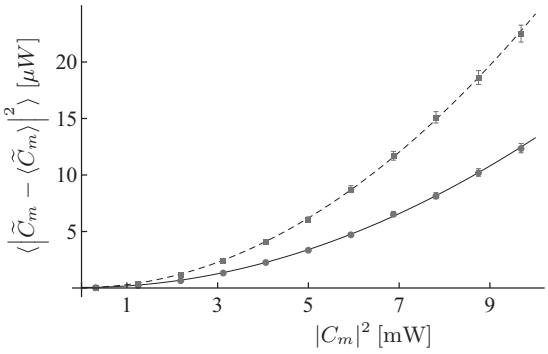

FIG. 10. Correlator (21) as a function of input signal power $|C_m|^2$ for $f_2(t)$, see black solid line, and for $f_4(t)$, see black dashed line. Solid and dashed lines correspond to leading order contribution (21) with the next-to-leading order corrections; see Eq. (A3). The noise power parameter is $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles and rectangles correspond to numerical results with statistical error on the level of three standard deviations for the functions f_2 and f_4, respectively.

for $\xi\gamma LP \ll 1$, and

$$I_{P_X^{\rm opt}[X]} = \log\log(B\xi\gamma LP/\sqrt{3}) - \log(QL^2\xi\gamma e/\sqrt{3}) + \frac{1}{\log(B\xi\gamma LP/\sqrt{3})}\left[\log\log(B\xi\gamma LP/\sqrt{3}) + 1 - \frac{\log\log(B\xi\gamma LP/\sqrt{3})}{\log(B\xi\gamma LP/\sqrt{3})}\right], \tag{49}$$

for $\log\xi\gamma LP \gg 1$ and $P \ll \Delta/(QL^3\xi^2\gamma^2)$. Here $B = 2e^{-\gamma_E}$, and $\gamma_E \approx 0.5772$ is the Euler constant. Note that the asymptotics (49) is obtained with accuracy $1/\log^2(\xi\gamma LP)$; see Sec. IV of Ref. [1].

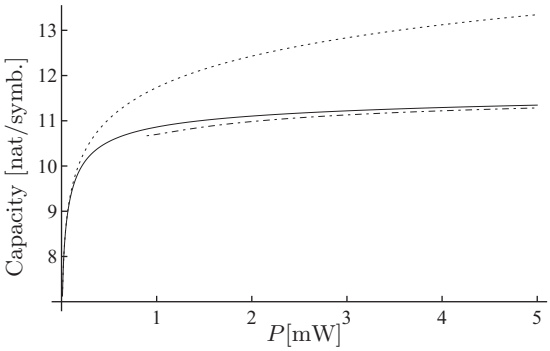

FIG. 11. Shannon capacity and the mutual information $I_{P_{\rm opt}^{(m)}}$ for the parameters $Q = 10^{-21}$ W/(km Hz), $L = 800$ km, $\gamma = 1.25$ (km W)$^{-1}$, $T_0 = 10^{-10}$ sec, and for the Gaussian shape (37) of $f(t)$. The black dotted line corresponds to the Shannon limit $\log(\frac{PT_0}{QL})$, the black solid line corresponds to $I_{P_{\rm opt}^{(m)}}$, see Eq. (47), and the black dashed-dotted line corresponds to the asymptotics (49) for large γLP.

VI. CONCLUSION

In the present paper we use results obtained in Ref. [1] for the per-sample model to calculate the informational characteristics of the channel where the input signal $X(t)$ depends on time; see Eq. (6). For this channel the information is carried by coefficients C_k. In the process of the signal propagation the input signal is transformed by the Kerr nonlinearity and the noise in the channel. To recover the transmitted information we introduce the detection procedure which removes the nonlinearity effects, see Eq. (11), and then projects $\widetilde{X}(t)$ on the basis functions, see Eq. (12), to obtain the coefficients $\widetilde{C}_k$. Using the conditional probability density function for the per-sample model obtained in Ref. [1] we calculate the correlators of the coefficients $\widetilde{C}_k$; see Eqs. (19)–(21). We demonstrate that these correlators depend on the noise bandwidth parameter Δ. We also perform the numerical calculations of these correlators using two different methods and show that the numerical and analytical results are in agreement. Using the obtained results for correlators we find the conditional probability density function $P[\{\widetilde{C}_k\}|\{C_k\}]$ in the leading and next-to-leading orders in parameter $QL/(\Delta P)$. Then we calculate the informational entropies and the mutual information for the channel in leading order in the parameter $QL/(T_0P)$. We perform variation of the mutual information over the input signal distribution function and obtain the optimal input signal distribution function which maximizes the mutual information. We calculate the channel capacity in the leading order in parameter $QL/(T_0P)$ and demonstrate that the capacity depends on the pulse envelope through one parameter ξ; see Eq. (28). The capacity grows as $\log\log P$ for sufficiently large average power P: $(\xi\gamma QL)^{-1} \ll P \ll \Delta/(QL^3\xi^2\gamma^2)$. Note that the same asymptotics was obtained for the per-sample model; therefore taking into account the time dependence of the pulse envelope does not change the asymptotics behavior and modifies only the nonlinearity parameter γ to $\xi\gamma$.

ACKNOWLEDGMENTS

All authors would like to thank the Russian Science Foundation (RSF), Grant No. 16-11-10133. A.V.R. would like to thank the Russian Foundation for Basic Research (RFBR), Grant No. 16-31-60031. I.S.T. would like to thank the Russian Science Foundation (RSF), Grant No. 17-72-30006, and the Ministry of Education and Science of the Russian Federation (14.Y26.31.0017).

APPENDIX: CORRELATORS (20) AND (21) WITH NLO CORRECTIONS

Let us present the correlator (20) with next-to-leading order (NLO) corrections in the noise power:

$$\begin{aligned}\langle(\widetilde{C}_m-\langle\widetilde{C}_m\rangle)(\widetilde{C}_n-\langle\widetilde{C}_n\rangle)\rangle &= \delta_{m,n}\left(\langle(\widetilde{C}_m-C_m)(\widetilde{C}_m-C_m)\rangle-\left(\frac{QL^2\gamma}{\Delta}\right)^2 C_m^2\left[-1+\frac{n_4^2}{9}\gamma^2L^2|C_m|^4+i\frac{2n_4}{3}\gamma L|C_m|^2\right]\right)\\ &= \delta_{m,n}\left(\frac{QL^2\gamma}{T_0}C_m^2\left[-\frac{2n_6}{3}\gamma L|C_m|^2-in_4\right]+\left(\frac{QL^2\gamma}{T_0}\right)^2\frac{T_0}{\Delta}C_m^2\right.\\ &\quad\left.\times\left[-\frac{9n_4}{2}+\frac{2n_8}{3}\gamma^2L^2|C_m|^4+i\frac{58n_6}{15}\gamma L|C_m|^2\right]\right).\end{aligned}\tag{A1}$$

Here we have used the relation

$$\langle(\widetilde{C}_m-\langle\widetilde{C}_m\rangle)(\widetilde{C}_m-\langle\widetilde{C}_m\rangle)\rangle=\langle(\widetilde{C}_m-C_m)(\widetilde{C}_m-C_m)\rangle-\langle\widetilde{C}_m-C_m\rangle^2,\tag{A2}$$

the result (19) for $\langle\widetilde{C}_m-C_m\rangle$, and the calculation of $\langle(\widetilde{C}_m-C_m)(\widetilde{C}_m-C_m)\rangle$ on the base of next-to-leading order result for $P[Y|X]$ in Ref. [2].

In a similar manner it is easy to calculate the following corrections to correlator (21) from the results obtained in Ref. [2]:

$$\begin{aligned}\langle(\widetilde{C}_m-\langle\widetilde{C}_m\rangle)\overline{(\widetilde{C}_n-\langle\widetilde{C}_n\rangle)}\rangle &= \delta_{m,n}\left(\langle(\widetilde{C}_m-C_m)\overline{(\widetilde{C}_m-C_m)}\rangle-\left(\frac{QL^2\gamma}{\Delta}\right)^2|C_m|^2\left[1+\frac{n_4^2}{9}\gamma^2L^2|C_m|^4\right]\right)\\ &= \delta_{m,n}\left(\frac{QL}{T_0}\left[1+\frac{2n_6}{3}\gamma^2L^2|C_m|^4\right]+\left(\frac{QL^2\gamma}{T_0}\right)^2\frac{T_0}{\Delta}|C_m|^2\left[n_4-\frac{2n_8}{9}\gamma^2L^2|C_m|^4\right]\right).\end{aligned}\tag{A3}$$

Note that these NLO results (A1) and (A3) contain the time discretization parameter Δ related to the noise bandwidth $W'=2\pi/\Delta$. The relative importance of the NLO corrections in correlators (A1) and (A3) is governed by the dimensionless parameter $(\frac{QL}{\Delta}\gamma L)\gamma LP$; i.e., it increases linearly for large and increasing P. To demonstrate the importance of these corrections for our numerical results we present Fig. 12, where for the noise power parameter $Q=5.94\times10^{-21}$ W/(km Hz) the imaginary part of the leading order contribution (20) and the next-to-leading order corrections (A1) are presented together with the numerical results (Runge-Kutta method) for the envelope form $f_2(t)=\sqrt{\frac{8}{3}}\cos^2(\pi t/T_0)$. One can see that our calculations, i.e., Eq. (24)

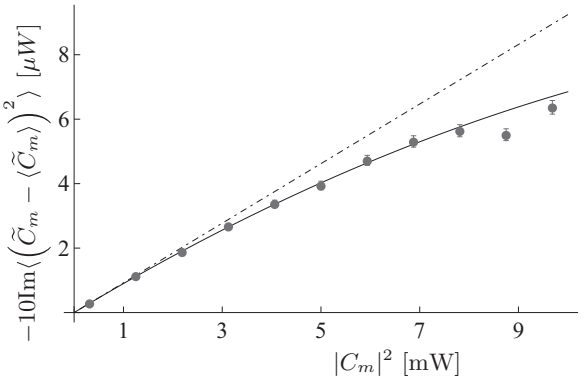

FIG. 12. The imaginary part of the correlator (20) multiplied by (-10) as a function of input signal power $|C_m|^2$ for $f_2(t) = \sqrt{\frac{8}{3}}\cos^2(\pi t/T_0)$ in the leading order (20), see black dashed-dotted line, and with the next-to-leading order corrections (A1), see the solid line. The noise power parameter $Q = 5.94 \times 10^{-21}$ W/(km Hz). Circles represent the numerical results for Runge-Kutta method.

and formulas of Sec. V based on the leading order results (19)–(21), are in good agreement with the numerical calculations up to the average power of order 4 mW for given noise and channel parameters.

[1] I. S. Terekhov, A. V. Reznichenko, Ya. A. Kharkov, and S. K. Turitsyn, Phys. Rev. E **95**, 062133 (2017).

[2] A. A. Panarin, A. V. Reznichenko, and I. S. Terekhov, Phys. Rev. E **95**, 012127 (2017).

[3] P. P. Mitra and J. B. Stark, Nature (London) **411**, 1027 (2001).

[4] E. E. Narimanov and P. Mitra, J. Lightwave Technol. **20**, 530 (2002).

[5] J. M. Kahn and K.-P. Ho, IEEE. J. Sel. Topics Quant. Electron. **10**, 259 (2004).

[6] R.-J. Essiambre, G. J. Foschini, G. Kramer, and P. J. Winzer, Phys. Rev. Lett. **101**, 163901 (2008).

[7] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, J. Lightwave Technol. **28**, 662 (2010).

[8] R. Killey and C. Behrens, J. Mod. Opt. **58**, 1 (2011).

[9] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, IEEE/OSA J. Lightwave Technol. **32**, 2862 (2014).

[10] M. A. Sorokina and S. K. Turitsyn, Nat. Commun. **5**, 3861 (2014).

[11] I. S. Terekhov, A. V. Reznichenko, and S. K. Turitsyn, Phys. Rev. E **94**, 042203 (2016).

[12] A. Mecozzi, J. Lightwave Technol. **12**, 1993 (1994).

[13] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, Phys. Rev. Lett. **91**, 203901 (2003).

[14] M. I. Yousefi and F. R. Kschischang, IEEE Trans. Inf. Theory **57**, 7522 (2011).

[15] G. Kramer, IEEE Trans. Inf. Theory **64**, 5131 (2018).

[16] M. A. Lavrentiev and B. V. Shabat, *Method of Complex Function Theory* (Nauka, Moscow, 1987) [in Russian]; M. Lavrentiev and B. Chabot, *Methodes de la Theorie des Fonctions d'une Variable complexe* (Mir, Moscou, 1977) [in French].

[17] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. (Academic Press, 2007).

[18] R. A. Fisher and W. K. Bischel, J. Appl. Phys. **46**, 4921 (1975).

[19] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).