Channel Capacity Calculation at Large SNR and Small Dispersion within Path-Integral Approach

A V Reznichenko\textsuperscript{1,2}, I S Terekhov\textsuperscript{1,2}

\textsuperscript{1} Theory department, Budker Institute of Nuclear Physics of Siberian Branch Russian Academy of Sciences, Novosibirsk 630090, Russia
\textsuperscript{2} Novosibirsk State University, Novosibirsk 630090, Russia
E-mail: a.v.reznichenko@inp.nsk.su

Abstract. We consider the optical fiber channel modelled by the nonlinear Shr"{o}dinger equation with additive white Gaussian noise. Using Feynman path-integral approach for the model with small dispersion we find the first nonzero corrections to the conditional probability density function and the channel capacity estimations at large signal-to-noise ratio. We demonstrate that the correction to the channel capacity in small dimensionless dispersion parameter is quadratic and positive therefore increasing the earlier calculated capacity for a nondispersive nonlinear optical fiber channel in the intermediate power region. Also for small dispersion case we find the analytical expressions for simple correlators of the output signals in our noisy channel.

1. Introduction

We consider the fiber optical channel described by the nonlinear Schr"{o}dinger equation (NLSE) with additive white Gaussian noise for the case when the signal-to-noise ratio (SNR) is large:

\begin{equation}
\partial_z \psi + i \beta \partial_t^2 \psi - i \gamma |\psi|^2 \psi = \eta(z,t),
\end{equation}

here \(\psi(z,t)\) is the complex signal, \(\gamma\) is the Kerr nonlinerity, \(\beta = \beta_2/2\) is the dispersion parameter, \(\eta(z,t)\) is the white Gaussian noise: \(\langle \eta(z,t) \rangle = 0, \langle \eta(z,t) \bar{\eta}(z',t') \rangle = Q \delta(z - z') \delta(t - t')\), where \(Q\) is a power of the white Gaussian noise (per unit length and frequency). The bar here and hereafter means complex conjugation.

We consider the mutual information \[1\] of the channel for the given input (continuous) signal probability density function (PDF) \(P[X]\): \(\psi(z = 0, t) = X(t)\). In our approach we use the representation of the mutual information in the path-integral form

\begin{equation}
I_{P[X]} = \int DXY P[X] P[Y|X] \log \left[ \frac{P[Y|X]}{P_{\text{out}}[Y]} \right] = H[Y] - H[Y|X],
\end{equation}

and we assume that the input signal \(X\) has the fixed finite average power \(P_{\text{ave}}\). The function \(P[Y|X]\) in (2) is the conditional probability density function, that is the probability density of receiving output signal \(\psi(z = L, t) = Y(t)\) when the input signal is given. The output signal PDF reads \(P_{\text{out}}[Y] = \int DX P[X] P[Y|X]\). The quantity \(H[Y] = -\int DY P_{\text{out}}[Y] \log \left[ P_{\text{out}}[Y] \right]\) is the output signal entropy (responsible for the signal transmission), and the entropy
\[ H[Y|X] = - \int DXDY P[X]P[Y|X] \log \left[ P[Y|X] \right] \] is the conditional entropy responsible for the noise impact, i.e., it is the measure of the uncertainty of the output signal \( Y(t) \) for a given input signal \( X(t) \). The measures \( DX \) and \( DY \) in (2) are defined below.

In our approach we use the formulation for the conditional PDF \( P[Y|X] \) through the path-integral (Martin-Siggia-Rose formalism [2]). In Ref.[3] the following ("quasiclassical") representation for the conditional probability density function was obtained:

\[ P[Y|X] = \Lambda \exp \left\{ - \frac{S[\Psi_{cl}]}{Q} \right\} \approx e^{-\frac{S[\Psi_{cl}]}{Q}}, \quad \Lambda = \int \mathcal{D}\phi \exp \left\{ - \frac{S[\Psi_{cl} + \phi] - S[\Psi_{cl}]}{Q} \right\}, \quad (3) \]

where the action \( S[\psi] \) in (3) reads in the time domain as quadrated left hand side of NLSE (1):

\[ S[\psi] = \int_0^L dz \int_0^T dt \left| \mathcal{L}[\psi(z,t)] \right|^2, \quad \mathcal{L}[\psi] = \partial_z \psi(z,t) + i\beta \partial_t^2 \psi(z,t) - i\gamma \psi(z,t)\psi(z,t)^2. \quad (4) \]

The approximate ("saddle point" approximation) equation in Eq. (3) involves the one-loop theoretical field effective action \( \Gamma[\Psi_{cl}] \) corresponding to the "classical action" (4), see [4]. Our goal is to calculate both the action \( S[\Psi_{cl}] \) and the normalization factor \( \Lambda \).

We introduce \( T \) as a time interval containing both signals \( X(t) \) and \( Y(t) \). In what follows we use the discretization scheme in the time domain. This scheme uses the dense time grid with \( M' \) intervals and spacing \( \delta_t \) (i.e., \( t_j = -T/2 + j\delta_t, j = 0, 1, \ldots, M' - 1 \)) and the coarse time sub-grid with \( M \) intervals and spacing \( \tilde{\delta}_t: T = M'\delta_t = M\tilde{\delta}_t, \tilde{\delta}_t \propto 2\pi / W', \tilde{\delta}_t \propto 2\pi / W. \) Here \( W' \) is the bandwidth of the input signal \( X(t_j) = \rho(t_j) \exp[i\phi_0(t_j)] \) and \( W \) is the noise bandwidth (usually, \( W' \gg W \)). The measure in (3) in our scheme reads

\[ \mathcal{D}\phi(z,t) = \lim_{\delta_t \to 0} \lim_{\Delta \to 0} \left\{ \frac{\delta_t}{\Delta \pi Q} \right\}^{M'-1} \prod_{j=0}^{N-1} \prod_{i=1}^{\Delta'} \left\{ \frac{\delta_t}{\Delta \pi Q} \right\} d\text{Re}\phi(z_i,J) d\text{Im}\phi(z_i,J), \quad (5) \]

where \( \Delta = L/N \) is the coordinate grid spacing with \( N \) points on \( L \), \( L \) is the channel length.

The function \( \Psi_{cl} \) in Eq. (3) is the solution (referred to as "the classical solution") of the Euler-Lagrange equation \( \delta S[\Psi_{cl}] = 0 \) with the boundary conditions: \( \Psi_{cl}(z = 0) = X, \quad \Psi_{cl}(z = L) = Y \). In the time domain this equation for \( \Psi_{cl}(z,t) \) has a notedly simple form [5]

\[ (\partial_z + i\beta \partial_t^2 - 2i\gamma |\Psi_{cl}(z,t)|^2) \mathcal{L}[\Psi_{cl}(z,t)] + i\gamma \Psi_{cl}^3(z,t) \mathcal{L}[\Psi_{cl}(z,t)] = 0, \]

\[ \mathcal{L}[\Psi_{cl}(z,t)] = (\partial_z + i\beta \partial_t^2 - i\gamma |\Psi_{cl}(z,t)|^2) \Psi_{cl}(z,t). \quad (6) \]

It is convenient to introduce the function \( \Phi(z,t) \) which is the solution of the nonlinear Schrödinger equation (NLSE) with zero noise, i.e., \( \mathcal{L}[\Phi(z,t)] = 0 \), and with the boundary condition \( \Phi(0,t) = X(t) \). It is obvious that the function \( \Phi(z,t) \) obeys the equation Eq. (6) and the boundary condition at \( z = 0 \), but it does not obey the boundary condition at \( z = L \). It globally minimizes the action as well: \( S[\Phi(z,t)] = 0 \). Since we imply that the noise power is much less than the signal power we can present the solution of Eq. (6) in the form \( \Psi_{cl}(z,t) = \Phi(z,t) + \varkappa(z,t) \), where the function \( \varkappa(z,t) \) is of order of \( \sqrt{Q} \) for unsuppressed configurations \( \Psi(z,t) \), since in the leading order in \( 1/\text{SNR} \) the action \( S[\Psi_{cl}] \) is quadratic functional in \( \varkappa(z,t) \), see details in [5]. Therefore we substitute the function \( \Psi_{cl}(z,t) \) in this form to the Eq. (6), then linearizing Eq. (6) in \( \varkappa(z,t) \) we obtain the linear problem on \( \varkappa(z,t) \):

\[ (\partial_z + i\beta \partial_t^2 - 2i\gamma |\Phi(z,t)|^2) l[\varkappa(z,t)] + i\gamma \Phi^2(z,t) l[\Phi(z,t)] = 0, \]

\[ l[\varkappa(z,t)] = (\partial_z + i\beta \partial_t^2) \varkappa(z,t) - i\gamma \left( 2\varkappa(z,t)\Phi(z,t)^2 + \varkappa(z,t)\Phi^2(z,t) \right), \quad (7) \]
with the boundary conditions \( \varphi(z = 0, t) = 0 \), \( \varphi(z = L, t) = Y(t) - \Phi(L, t) \equiv \delta Y(t) \).

We will use the dimensionless dispersion parameter \( \beta = \beta L / \delta t^2 \). In our perturbative consideration this parameter is assumed to be small: \( \beta \ll 1 \). The dimensionless parameter \( \gamma = \gamma L P_{\text{ave}} / \sqrt{3} \) is the average nonlinearity parameter, where \( P_{\text{ave}} \) is the average power over the input signal PDF \( P[X] \): \( P_{\text{ave}} = \lim_{T \to \infty} \left| \int_X P[X(t)] \frac{1}{T} \int_T |X(t)|^2 \, dt \right| \).

2. **Conditional PDF \( P[Y|X] \) for small dispersion parameter**

In the leading order in \( 1/\text{SNR} \) the action difference in (3) reads in discrete time form [5]:

\[
S[\Psi_{cl} + \phi] - S[\Psi_{cl}] \approx S_2[\phi, \Phi] = \delta t \sum_{k=0}^{M'-1} \int_0^L dz \mathcal{L}_{cl}[\phi(z, t_k)],
\]

(8)

\[
\mathcal{L}_{cl}[\phi(z, t_k)] = \left( \partial_z \phi(z, t_k) + i \beta \partial_z^2 \phi(z, t_k) - i \gamma \left( 2 \phi(z, t_k) \Phi(z, t_k) + \bar{\phi}(z, t_k) \Phi^2(z, t_k) \right) \right)^2.
\]

(9)

We perform our calculation in the standard discretisation scheme where the second time derivative in the path-integral (3) is replaced with the second order discrete analog:

\[
\partial_t^2 \psi(z, t_i) \longrightarrow \Delta^2_{\delta t} \psi(z, t_i) = \psi(z, t_i + \delta t) - 2 \psi(z, t_i) + \psi(z, t_i - \delta t). \tag{10}
\]

Note that different discretisation schemes of the time second derivative correspond to the variation of time grid spacing (i.e., different renormalisation schemes). For instance, for the second difference derivation \( \Delta^2_{(3)} \) of the third order approximation we have \( \Delta^2_{(3)} \psi(z, t_i) \equiv -\psi(z, t_i + 2 \delta t) + 16 \psi(z, t_i + \delta t) - 30 \psi(z, t_i) + 16 \psi(z, t_i - \delta t) - \psi(z, t_i - 2 \delta t)/(12 \delta t^2) = (\frac{1}{3} \Delta^2_{3\delta t} - \frac{4}{3} \Delta^2_{2\delta t}) \psi(z, t_i) \), i.e., it is equivalent to the consideration of two combined grids \( \delta t \) and \( 2 \delta t \) with the same operator (10). In this sense, our fixed grid spacing \( \delta t \propto 2 \pi / W' \) is concerned with the bandwidth \( W' \) of the noise, and it places the role of the ultraviolet cutoff of our field theory with the action \( S[\psi] \), see (4).

Based on results of [6] for the conditional PDF \( P[Y|X] \) of a nondispersive nonlinear optical fiber channel we can present the following result for \( M' \) independent time channels governed by NLSE (1) with zero noise and \( \beta = 0 \):

\[
P[Y|X]^{(\beta = 0)} = \exp \left\{ -\delta t \sum_{j=0}^{M'-1} \frac{(1 + 4 \mu_j^2/3) x_{0,j}^2 - 2 \mu_j x_{0,j} y_{0,j} + y_{0,j}^2}{QL(1 + \mu_j^2/3)} \right\} \prod_{j=0}^{M'-1} \frac{\delta t}{\pi Q L \sqrt{1 + \mu_j^2/3}}. \tag{11}
\]

Here \( \mu(t_j) \equiv \mu_j = \gamma L \rho^2(t_j) \), and real functions \( x_{0,j} = x_0(t_j) \) and \( y_{0,j} = y_0(t_j) \) are defined through the output signal \( Y(t) \) and the signal \( \Phi^{(0)}(z = L, t) \), where \( \Phi^{(0)}(z = L, t) \) is the input signal \( X(t) \) evolved at \( z = L \) by the NLSE (1) with zero noise and \( \beta = 0 \):

\[
x_{0,j} + iy_{0,j} = \left[ Y(t_j) - \Phi^{(0)}(L, t_j) \right] \exp \left\{ -i \mu_j - i \phi_{0,j} \right\} = Y(t_j) \exp \left\{ -i \mu_j - i \phi_{0,j} \right\} - \rho_j. \tag{12}
\]

where we have used that for the zero noise and \( \beta = 0 \) the NLSE (1) with the input condition \( \Phi^{(0)}(z = 0, t_j) = X(t_j) = \rho_j \exp \{ i \phi_{0,j} \} \) has the solution \( \Phi^{(0)}(z, t_j) = \rho_j \exp \{ i \phi_{0,j} + i \mu_j z / L \} \).

For the PDF (11) the factor \( \Lambda \) denoted as \( \Lambda^{(0)} \) for the nondispersive channel reads

\[
\Lambda^{(0)} = \prod_{j=0}^{M'-1} \frac{\delta t}{\pi Q L \sqrt{1 + \mu_j^2/3}}. \tag{13}
\]
To present the first correction to $P[Y|X]$ in small dispersion parameter $\tilde{\beta}$ we first introduce
the real functions $x(t)$ and $y(t)$ from analogy with the definition (12):
\begin{equation}
\begin{aligned}
x(t_j) + iy(t_j) &= [Y(t_j) \Phi(z = L), t_j)]e^{-i\mu_j - i\phi_0 j},
\end{aligned}
\end{equation}
where $\Phi(z, t_j)$ is the solution of the NLSE equation (1) with zero noise, with nonzero $\beta$, and with
the input boundary condition $\Phi(z = 0, t_j) = X(t_j) = \rho_j \exp\{i\phi_0 j\}$. The solution $\Phi(z, t)$ has
the form: $\Phi(z, t) = e^{i\mu(t)z/L + i\phi_0(t)} \sum_{n=0}^{\infty} \Phi^{(n)}(z, t)$. Quantities $\Phi^{(n)}(z, t)$ up to the factors $(z\beta)^n$
are polynomials in nonlinear parameter $\mu(t)/\tilde{T}$. The function $\Phi^{(0)}(z, t) = \rho(t)$ corresponds to
the leading order in small $\tilde{\beta}$, and $\Phi^{(1)}(z, t)$ corresponds to the next-to-leading order, and so on.
As an example, the first correction reads $\tilde{\Phi}^{(1)}(z, t) = \beta z \sum_{k=0}^{2} \phi_k^{(1)}(t) \left(\mu(t)/\tilde{T}\right)^k$, where $\phi_k^{(1)}(t) = 2\hat{\beta} \phi_0 + \hat{\phi}_0 \rho + i \left(\rho \phi_0^2 - \hat{\beta}\right)$, $\phi_1^{(1)}(t) = \hat{\beta} + 3\hat{\beta}^2/\rho + i \left(\phi_0 \rho + 4\rho \phi_0\right)$, $\phi_2^{(1)}(t) = i \hat{\beta} \left[\hat{\rho} + 5\hat{\beta}^2/\rho\right]$. In
the perturbation expansion it is easy to find $\Phi^{(n)}(z, t)$ for any given $n$.

It is easy to find perturbatively $\Psi_{cl}(z, t)$ as well, and the corrections in $\tilde{\beta}$ for the action $S[\Psi_{cl}]$
are presented as $S[\Psi_{cl}] = S(0) + \delta S = S(0) + S(1) + S(2) + \mathcal{O}(\tilde{\beta}^3)$, here
\begin{equation}
S(0) = \delta_t \sum_{j=0}^{M-1} \left(1 + 4\mu^2/3\right) x_j^2 - 2\mu x_j y_j + y_j^2/L(1 + \mu^2/3),
\end{equation}
where $x_j \equiv x(t_j)$ and $y_j \equiv y(t_j)$ are given by (14). The first correction $S(1)$ reads simpler in
the continuous notations:
\begin{align*}
S(1) &= \int_{\mathcal{T}} \frac{\beta}{15\mu^2(3 + \mu^2)^4} \left\{ 2\mu^2 \left[ (154\mu^8 + 1848\mu^6 + 5805\mu^4 + 9990\mu^2 - 405) x^2 + \\
&\mu(200\mu^8 + 1947\mu^6 + 5841\mu^4 + 7101\mu^2 - 5265)xy - 3(50\mu^8 + 492\mu^6 + 1593\mu^4 + \\
&2322\mu^2 + 135) y^2 \right] + 2\mu(\mu^2 + 3) \rho \left[ \hat{\rho}(3(10\mu^6 + 66\mu^4 + 135\mu^2 - 135) x^2 + \mu(40\mu^6 + 243\mu^4 + \\
&288\mu^2 - 675) xy - 3(10\mu^6 + 66\mu^4 + 117\mu^2 - 45) y^2 \right] + 3\hat{\rho}(10(\mu^2 + 3) \phi_0((6\mu^3 + 33) \mu x^2 + \\
&4(2\mu^4 + 7\mu^2 - 3) xy - 3(2\mu^2 + 7) y^2) + \hat{\rho}(\mu(13\mu^4 + 102\mu^2 + 45) x - 6(3\mu^4 + 22\mu^2 + 15) y) + \\
&\hat{\rho} \hat{\dot{x}}(3(11\mu^2(\mu^2 + 6) - 45) y - 2(14\mu^6 + 99\mu^4 + 135) x) + \left( \mu^2 + 3 \right) \rho^2 \left[ -3\rho^2 (x(-30\hat{\phi}_0 + \\
&20\hat{\phi}_0 + 11\hat{\gamma}) - 40\rho x (\phi_0^2 + \rho x) + y (30\rho x \hat{\phi}_0) + 6\rho (5x (\phi_0 + 3x \hat{\phi}_0 + 2) y + \hat{x}) + y(80x \hat{\phi}_0 + \\
&10x \hat{\phi}_0 + 3\hat{\gamma}) - 15y^2 \hat{\phi}_0^2) - 45\rho^2 (x(-11x \hat{\phi}_0 + 6x \hat{\phi}_0 + 4) y + 2y \hat{\phi}_0 (\gamma - 4x \hat{\phi}_0) + 9y^2 \hat{\phi}_0) + \\
&90\rho (y(-2x \hat{\phi}_0 + 2x \hat{\phi}_0 + y) + x \hat{\phi}_0 (3x \hat{\phi}_0 + 2y) + x \hat{\phi}_0 (x \hat{\phi}_0 + 2y) - x \hat{\phi}_0 + 3y^2 \hat{\phi}_0^2) - 135(-y(\hat{x} - 2x \hat{\phi}_0) + \\
&x (x \hat{\phi}_0 + 2x \hat{\phi}_0 + y) + y^2 \hat{\phi}_0) + 8\rho^5 x (x + 15y \phi_0) \right] \right\}.
\end{align*}

We present the normalization factor $\Lambda$ in the form $\Lambda = \Lambda(0) + \delta \Lambda = \Lambda(0) \left\{ 1 + \frac{\Lambda(1)}{\Lambda(0)} + \frac{\Lambda(2)}{\Lambda(0)} + \mathcal{O}(\tilde{\beta}^3) \right\}$. The normalization condition for conditional PDF $P[Y|X]$ reads
\begin{equation}
\int_{\mathcal{T}} \mathcal{D}Y P[Y|X] = 1.
\end{equation}
It is obvious that up to the next order in $\tilde{\beta}$ one has
\begin{equation}
\frac{\Lambda(1)}{\Lambda(0)} = \langle S(1) \rangle_{P[Y|X]} = \frac{\int_{\mathcal{T}} \mathcal{D}Y P[Y|X] S(1)}{\int_{\mathcal{T}} \mathcal{D}Y P[Y|X] S(0)}.
\end{equation}
where \(P[Y|X]^{(0)} = \Lambda^{(0)} \exp \left[-S^{(0)}/Q\right]\) with \(S^{(0)}\) given by (15), and \(\Lambda^{(0)}\) given by (13). Note that \(DY = \prod_{j=0}^{M-1} dReY(t_j) dImY(t_j) = \prod_{j=0}^{M-1} dx_j dy_j\), see (14).

From the explicit discrete form of \(S^{(1)}/Q\) one can immediately obtain \(\Lambda^{(1)}/\Lambda^{(0)}\). Let us note that this quantity is singular in \(\delta t\): in the continuous notations it has the form

\[
\frac{\Lambda^{(1)}}{\Lambda^{(0)}} = \beta \int dt \left\{ -\delta_R''(0) \frac{2L\mu^3}{15(\mu^2 + 3)} - \delta_R(0) \frac{L\mu^2}{105(\mu^2 + 3)} \rho^2 \right\} \left\{ \frac{14}{3} (2\mu^2 + 15) \rho \dot{\phi}_0 + \right\}
\]

\[
\mu \left( 16\mu^2 + 189 \right) \rho^2 + 7\rho \left( 2\mu \dot{\phi}_0 + 15\ddot{\phi}_0 \right) \frac{105}{2} \rho \dot{\phi}_0 + 7\rho \right\}.
\]

The singularity is regularized by the renormalisation cutoff \(\delta t\): the integral \(\int dt\) is understood as the sum \(\delta t \sum_{j=0}^{M-1}, \int dt \delta_R(0) \ldots = \sum_{j=0}^{M-1} \ldots\). The label “R” means regularization of delta-function in the fixed discretization scheme. In our scheme (10) the second derivative \(\delta''(0)\) reads \(\delta''(0) = -2/\delta t^2\). The small dispersive parameter of the perturbation expansion is \(\tilde{\beta} = -\beta L\delta''(0)/(2\delta R(0))\). Another discretization scheme \(R'\) corresponds to another cutoff \(\delta t_{R'} \neq \delta t\) and another relation between the noise bandwidth \(W' \propto 1/\delta t_{R'}\).

In the second order in \(\beta\) we have

\[
\frac{\Lambda^{(2)}}{\Lambda^{(0)}} = \left( \frac{\Lambda^{(1)}}{\Lambda^{(0)}} \right)^2 + \left( \frac{S^{(2)}}{Q} - \frac{1}{2} \left( \frac{S^{(1)}}{Q} \right)^2 \right) P[Y|X]^{(0)}.
\]

We calculate this correction in the standard discretisation scheme. The result averaged over the angular variables is presented in the last section.

The conditional PDF \(P[Y|X]\) has the following perturbative form in the first and the second orders in \(\beta\)

\[
P[Y|X] \approx P[Y|X]^{(0)} \begin{pmatrix} 1 + \left( \frac{\Lambda^{(1)}}{\Lambda^{(0)}} - \frac{S^{(1)}}{Q} \right) + \left( \frac{\Lambda^{(2)}}{\Lambda^{(0)}} - \frac{\Lambda^{(1)}}{\Lambda^{(0)}} \frac{S^{(1)}}{Q} + \frac{1}{2} \left( \frac{S^{(1)}}{Q} \right)^2 - \frac{S^{(2)}}{Q} \right) \right).
\]

where the correction \(\Lambda^{(1)}/\Lambda^{(0)}\) is given by (18), (19) and the correction \(S^{(1)}\) to the action is given by (16).

The receiver (output signal detector) has bandwidth \(W_r \ll W'\).

\[
P_r[Y|X] = \int [DY]_{W'\setminus W_r} P[Y|X], \quad [DY]_{W'\setminus W_r} = \prod_{k, \omega_k \in W'\setminus W_r} dReY_{\omega_k} dImY_{\omega_k},
\]

where PDF \(P_r[Y|X]\) determines \(P_{out}[Y], H[Y], H[Y|X], I_{P[X]}, C, \) etc. To estimate the mutual information \(I_{P[X]}\) and the capacity \(C = \max_{P[X]} I_{P[X]}\) we simplify our calculation of \(P[Y|X]\) and \(P_r[Y|X]\) by assumption \(W' \sim W \sim W_r\). We will develop the perturbation expansion in \(\beta \ll 1\) on the base of \(\beta = 0\) results. In real channels \(W \ll W'\) but here we want only to demonstrate our method applied in the general case (of different \(W, W', W_r\)) as well.
3. Simple correlators

Let us consider the average of an arbitrary function \( F(x(t), y(t')) \), see Eq. (14), over the realizations of the output signal \( Y(t) \) when the input signal \( X(t) \) is fixed:

\[
\langle F(x_j, y_k) \rangle_{P[Y|X]} = \int DYP[Y|X]F(x_j, y_k) = \langle F(x_j, y_k) \rangle_{(0)} + \left\{ \frac{S^{(1)}}{Q} \right\}_{(0)}^{\alpha} \langle F(x_j, y_k) \rangle_{(0)} - \\
\left\{ \frac{S^{(1)}}{Q} \right\}_{(0)}^{\beta} \frac{\Lambda^{(1)}}{\Lambda^{(0)}} \left\{ \frac{S^{(1)}}{Q} \right\}_{(0)}^{\gamma} \langle F(x_j, y_k) \rangle_{(0)} - \langle F(x_j, y_k) \rangle_{(0)}^{\beta} \right\} + \\
\left\{ \frac{S^{(1)}}{Q} \right\}_{(0)}^{\delta} \langle F(x_j, y_k) \rangle_{(0)}^{\delta} - \langle F(x_j, y_k) \rangle_{(0)}^{\delta} \right\} - \frac{1}{2} \left\{ \frac{S^{(1)}}{Q} \right\}_{(0)}^{\epsilon} + \langle F(x_j, y_k) \rangle_{(0)}^{\epsilon} - \\
\langle F(x_j, y_k) \rangle_{(0)}^{\epsilon} \right\} + O(\beta^3),
\]

(23)

here we have used that the conditional PDF \( P[Y|X] \) has the form (21). Here we introduced the average over “nonperturbed” conditional PDF: \( \langle F(x_j, y_k) \rangle_{(0)} = \int DYP[Y|X]F(x_j, y_k) \), where \( P[Y|X]^{(0)} = \Lambda^{(0)} \exp \left\{ -S^{(0)}/Q \right\} \), see (13) and (15). Note that \( P[Y|X]^{(0)} \) does not correspond to the case \( \beta = 0 \), since \( x = Re \left( (Y(t_j) - \Phi(L, t_j)) e^{-i\mu_j - i\delta_{0}} \right) \neq x_0, y \neq y_0 \) in (15).

For the illustration of our method we consider the following simple simultaneous correlators:

\[
\alpha_{1,1}(t_j) = \langle x(t_j)^2 \rangle_{P[Y|X]}, \quad \alpha_{2,2}(t_j) = \langle y(t_j)^2 \rangle_{P[Y|X]}, \quad \alpha_{1,2}(t_j) = \langle x(t_j)y(t_j) \rangle_{P[Y|X]}.
\]

(24)

Now we use (23), where all averages \( \langle \ldots \rangle_{(0)} \) can be calculated using the Wick theorem [7] for pairing with Gaussian statistics of distribution \( P[Y|X]^{(0)} \). The machinery calculation is simple and we present the final result in the first order in \( \beta \):

\[
\alpha_{1,1}(t_j) \approx \frac{QL}{2 \delta_t} \left\{ 1 - \frac{4 \mu \beta L}{3 \delta_t^2} - \frac{2 \beta \mu L}{15 \rho^2} \left[ 10 (\rho^2 \phi_0^2 + \rho \bar{\rho}) + \mu^2 (10 \rho \bar{\rho} + 38 \rho^2) + 15 \mu^2 \rho \phi_0^2 + 45 \mu \rho \rho \phi_0^2 \right] \right\},
\]

(25)

\[
\alpha_{2,2}(t_j) \approx \frac{QL(4 \mu^2 + 3)}{6 \delta_t} \left\{ 1 - \frac{4 \mu (4 \mu^2 - 5)}{5 (4 \mu^2 + 3)} \beta \mu L \right\} + \frac{2 \beta \mu L}{105 \rho^2 (4 \mu^2 + 3)} \left[ 128 \mu^4 \rho^2 + 196 \mu^3 \rho \rho \phi_0 + 84 \mu^2 \rho \phi_0^2 + 504 \mu \rho \rho \phi_0 + 1932 \mu \rho \rho \phi_0^2 + 840 \mu \rho \rho \phi_0 + 1995 \mu \rho \rho \phi_0 + 210 \rho \rho \phi_0^2 - 210 \rho \rho \phi_0 - 84 \mu \rho \rho \phi_0 + 504 \mu \rho \rho \phi_0 + 1932 \mu \rho \rho \phi_0^2 + 210 \rho \rho \phi_0^2 - 210 \rho \rho \phi_0 \right\},
\]

(26)

\[
\alpha_{1,2}(t_j) \approx \frac{QL}{2 \delta_t} \left\{ 1 - \frac{4 \mu}{3 \delta_t^2} - \frac{\beta L}{45 \rho^2} \left[ 40 \mu^3 \rho \rho \phi_0 + 136 \mu^2 \rho \rho \phi_0 + 60 \mu^2 \rho \rho \phi_0 + 156 \mu^2 \rho \rho \phi_0 + 30 \mu \rho \rho \phi_0^2 - 75 \mu \rho \rho \phi_0 - 15 \mu \rho \rho \phi_0 - 120 \rho \rho \phi_0 \right] \right\},
\]

(27)

where all functions in the r.h.s. of (25)–(27) are taken at \( t = t_j \). Note that the correlators (25)–(27) are singular since they contain the noise power \( QL/\delta_t \). Here the dense grid spacing \( \delta_t \propto 2\pi W^{t-1} \) is related with the bandwidth \( W \) of the white noise. The dense grid spacing \( \delta_t \) plays the role of the smallest ultraviolet cutoff of our field theory with the action (9). From this point of view all of quantities (25)–(27) require the renormalization procedure [7].

6
4. Mutual information and capacity
From representation of [5] the mutual information \( I_{P[X]} \) in the leading order in 1/SNR reads:

\[
I_{P[X]} = M \log \left( \frac{P_{\text{ave}}}{P_{\text{noise}}} \right) + \left\langle \log \left( \frac{\Lambda}{\Lambda_{QL}} \right) \right\rangle_{P[X]},
\]

where \( \Lambda_{QL} = \left( \frac{\delta t}{\pi Q L} \right)^M \), and \( \left\langle F[X] \right\rangle_{P[X]} = \int DXP[X]F[X] \) is averaging over the input signal PDF \( P[X] \). The representation (28) was obtained in [5] on the assumption that PDF \( P[Y|X] \) obeys the factorization property (for meaning and remnant channels, see [5]). In our estimations we assume that \( W' \sim W \sim W_c \) (all time channels are meaning, \( M' = M \)) and (28) holds true in the leading order in 1/SNR (without any factorization assumption). For realistic case one has \( W' \gg W \) and for the concrete output signal detector \( W_c \neq W \). However, our estimations will give the general capacity behavior, and our method allows one to find the capacity for given realizations of noise \( (W') \) and output signal detector \( (W_c) \).

\[
I_{P[X]} = M \log \left( \frac{P_{\text{ave}}}{P_{\text{noise}}} \right) - \frac{M}{2} \left\langle \int \frac{dt}{T} \log \left[ 1 + \mu^2(t)/3 \right] \right\rangle_{P[X]} + \left\langle \log \left( 1 + \frac{\delta \Lambda}{\Lambda(0)} \right) \right\rangle_{P[X]},
\]

where the last term in (29) delivers the correction in \( \beta \) to the mutual information \( I_{P[X]}^{(\beta=0)} \) (the first two terms) of an nondispersive channel [6]. In the first order in \( \tilde{\beta} \) we arrive at the representation

\[
I_{P[X]} = I_{P[X]}^{(\beta=0)} + \left\langle \frac{\Lambda(1)}{\Lambda(0)} \right\rangle_{P[X]} + \left\langle \frac{\Lambda(2)}{\Lambda(0)} - \frac{1}{2} \left( \frac{\Lambda(1)}{\Lambda(0)} \right)^2 \right\rangle_{P[X]} + \mathcal{O}(\tilde{\beta}^3),
\]

where \( \Lambda(1)/\Lambda(0) \) is presented in (19) in the continuous notations. However for explicit averaging \( \left\langle \ldots \right\rangle_{P[X]} \) it is more convenient to use the discrete notation for \( \Lambda(1)/\Lambda(0) \) and \( \Lambda(2)/\Lambda(0) \).

Now it is easy to see that the first correction in \( \tilde{\beta} \) in (30) vanishes for wide class of distributions \( P[X] \). In particular, for an axially symmetric distribution \( P[X] = P[X] = P_{\text{ave}} \) the first correction in \( \tilde{\beta} \) is zero:

\[
\left\langle \frac{\Lambda(1)}{\Lambda(0)} \right\rangle_{P[X]} = \int P_{\text{ave}} \frac{M-1}{\Lambda(0)} \rho_j d\rho_j = 0.
\]

For the channel capacity correction \( \Delta C \) we need to know the second correction \( \Lambda(2) \) to the normalization factor, see Eq. (36) below. This correction \( \Lambda(2) \) has to be averaged over zero order optimal input signal distribution \( P_{X}^{\text{opt}(\beta=0)}[X] \) known in the intermediate power region \( QL \ll P \ll \left( Q\gamma^2L^3 \right)^{-1} \), see [6]. This distribution is angular symmetric, [6]:

\[
P_{X}^{\text{opt}(\beta=0)}[X] = \prod_{j=0}^{M-1} P_{X}^{\text{opt}(\beta=0)} \gamma[X(t_j)] = \prod_{j=0}^{M-1} N_0 \frac{\exp \left\{ -\lambda_0 |X(t_j)|^2 \right\}}{\sqrt{1 + \mu^2(t_j)/3}},
\]

where \( N_0 \) and \( \lambda_0 \) are functions depending on average input signal power \( P_{\text{ave}} = P/\delta t \) through the conditions, see Eqs. (46)–(47) in Ref. [6]:

\[
2\pi N_0 \int_0^\infty \frac{d\rho \rho e^{-\lambda_0 \rho^2}}{\sqrt{1 + \gamma^2L^2 \rho^2/3}} = 1, \quad 2\pi N_0 \int_0^\infty \frac{d\rho \rho^3 e^{-\lambda_0 \rho^2}}{\sqrt{1 + \gamma^2L^2 \rho^4/3}} = \frac{P}{\delta t}.
\]
The typical average for our calculation of the average $\langle \Delta^2 \rangle$ has the form of the integral

$$j^{\text{opt}}(\beta=0)(k,n) \equiv \int dR e X_j d\text{Im} X_j P^X_{\text{opt}(\beta=0),j} \{X(t_j)\} \frac{\mu_j(t_j)}{(3 + \mu_j(t_j)^2)} = n_0(\gamma) 3^{\delta_j/2} n_j(k, n; \alpha_0(\gamma)), \tag{33}$$

here $k$ and $n$ are integers, and $n_j(k, n; \alpha) \equiv \int_0^\infty dz e^{-\alpha z_j} \frac{e^{-\alpha z_j}}{(1 + e^{-\alpha z_j})^2}$. The average $j^{(\beta=0)}(k,n)$ can be expressed through the functions $\alpha_0(\gamma) \equiv P\lambda_0/(\delta \gamma^2)$ and $n_0(\gamma) \equiv \pi P \lambda_0/(\delta \gamma^2)$ which depends on dimensionless parameter $\gamma = \gamma LP/(\sqrt{3} \delta t)$ only. The properties of these functions were studied in details in Ref. [6]. For integer $k$ and $n$ and arbitrary $\alpha$ the integral $j^{(\beta=0)}(k,n;\alpha)$ can be recurrently expressed through $j(0,0;\alpha) = \int_0^\infty dz e^{-\alpha z_j}/\sqrt{1+z^2} = \frac{\pi}{2} (H_0(\alpha) - Y_0(\alpha))$ with $Y_0(\alpha)$ and $H_0(\alpha)$ being the Neumann and Struve functions of zero order, respectively. We will use this function $j(0,0;\alpha)$ for the calculation of the channel capacity corrections. The explicit form of the correction $\langle \Delta^2 \rangle$ is very cumbersome to be presented here, and we present the discrete form of the correction $\langle \Delta^2 \rangle$ averaged over the angular variables in the following form which is the most convenient for further numerical calculation of $\Delta C$:

$$\langle \Delta^2 \rangle / \langle \Delta^0 \rangle P[\|X\|] = 4M'\langle \Delta^2 \rangle \Re \left[ \frac{1}{M'} \sum_{j=0}^{M'-1} \left\{ 6j_{j+1,1,0,2,0,1,4} - 12j_{j+1,1,0,2,0,1,5} + 18j_{j+1,1,0,2,1,1,6} - 4ij_{j+1,1,0,3,0,1,4} + \right. \right.$$  

$$24iJ_{j+1,1}(0,3,0,1,5) - 48iJ_{j+1,1}(0,3,0,1,6) - iJ_{j+1,1}(0,3,1,1,5) + 54iJ_{j+1,1}(0,3,1,1,6) - 72iJ_{j+1,1}(0,3,1,1,7) - 36J_{j+1,1}(0,3,1,2,4) +$$  

$$72J_{j+1,1}(0,1,0,2,6) + 72J_{j+1,1}(1,0,2,6) - 432iJ_{j+1,1}(0,2,0,2,5) - 648iJ_{j+1,1}(0,2,0,2,6) + 432iJ_{j+1,1}(0,2,0,2,7) +$$  

$$18J_{j+1,1}(1,2,1,1,6) - 108iJ_{j+1,1}(1,2,1,1,7) + 16J_{j+1,1}(1,3,0,1,5) - 72J_{j+1,1}(1,3,0,1,6) + 96J_{j+1,1}(1,3,0,1,7) + 30J_{j+1,1}(1,3,0,2,4) -$$  

$$300J_{j+1,1}(1,3,0,2,5) + 1068J_{j+1,1}(1,3,0,2,6) - 1008J_{j+1,1}(1,3,0,2,7) - 1584J_{j+1,1}(1,3,0,2,8) + 6J_{j+1,1}(1,3,1,1,5) - 84J_{j+1,1}(1,3,1,1,6) +$$  

$$396J_{j+1,1}(1,3,1,1,7) - 504J_{j+1,1}(1,3,1,1,8) - 16iJ_{j+1,1}(1,4,0,1,5) + 176iJ_{j+1,1}(1,4,0,1,6) - 768iJ_{j+1,1}(1,4,0,1,7) +$$  

$$1334J_{j+1,1}(1,4,0,1,8) - 12J_{j+1,1}(1,4,0,2,4) - 24J_{j+1,1}(1,4,0,2,5) - 312J_{j+1,1}(1,4,0,2,6) + 1296J_{j+1,1}(1,4,0,2,7) -$$  

$$576J_{j+1,1}(1,4,0,2,8) - 3456J_{j+1,1}(1,4,0,2,9) + 32J_{j+1,1}(1,5,0,1,6) - 320J_{j+1,1}(1,5,0,1,7) + 1344J_{j+1,1}(1,5,0,1,8) -$$  

$$2304J_{j+1,1}(1,5,0,1,9) + 8J_{j+1,1}(1,5,0,2,4) - 96J_{j+1,1}(1,5,0,2,5) + 336J_{j+1,1}(1,5,0,2,6) - 96J_{j+1,1}(1,5,0,2,7) - 1356J_{j+1,1}(1,5,0,2,8) +$$  

$$6912J_{j+1,1}(1,5,0,2,10) + 16iJ_{j+1,1}(1,6,0,2,5) - 176iJ_{j+1,1}(1,6,0,2,6) + 768iJ_{j+1,1}(1,6,0,2,7) - 1344iJ_{j+1,1}(1,6,0,2,8) -$$  

$$32J_{j+1,1}(1,7,0,2,6) + 320J_{j+1,1}(1,7,0,2,7) - 1344J_{j+1,1}(1,7,0,2,8) + 2304J_{j+1,1}(1,7,0,2,9) + 18J_{j+1,1}(1,7,0,2,10) -$$  

$$18J_{j+1,1}(2,1,1,1,6) + 108iJ_{j+1,1}(2,1,1,1,7) - 6J_{j+1,1}(2,2,1,1,4) + 12J_{j+1,1}(2,2,1,1,5) + 24J_{j+1,1}(2,2,1,1,6) + 72J_{j+1,1}(2,2,1,1,7) -$$  

$$432J_{j+1,1}(2,2,1,1,8) - 2iJ_{j+1,1}(2,3,1,1,4) - 10iJ_{j+1,1}(2,3,1,1,5) - 48iJ_{j+1,1}(2,3,1,1,6) + 588iJ_{j+1,1}(2,3,1,1,7) -$$  

$$1728iJ_{j+1,1}(2,3,1,1,8) + 1728iJ_{j+1,1}(2,3,1,1,9) - 6iJ_{j+1,1}(3,0,1,1,5) + 18iJ_{j+1,1}(3,0,1,1,6) - 6J_{j+1,1}(3,0,1,1,7) +$$  

$$48J_{j+1,1}(3,0,1,1,8) - 108J_{j+1,1}(3,0,1,1,9) - 2iJ_{j+1,1}(3,2,1,1,5) + 14iJ_{j+1,1}(3,2,1,1,6) - 36iJ_{j+1,1}(3,2,1,1,7) + 12iJ_{j+1,1}(3,2,1,1,8) +$$  

$$20J_{j+1,1}(3,2,1,1,9) - 288J_{j+1,1}(3,3,1,1,8) + 1752J_{j+1,1}(3,3,1,1,9) - 608iJ_{j+1,1}(3,3,1,1,10) + 4608J_{j+1,1}(3,3,1,1,10) -$$  

$$12iJ_{j+1,1}(4,0,1,1,6) + 72J_{j+1,1}(4,0,1,1,7) - 72J_{j+1,1}(4,0,1,1,8) + 12iJ_{j+1,1}(4,1,1,1,7) - 144iJ_{j+1,1}(4,1,1,1,7) +$$  

$$576iJ_{j+1,1}(4,1,1,1,8) - 576iJ_{j+1,1}(4,1,1,1,9) + 4J_{j+1,1}(4,2,1,1,6) - 96J_{j+1,1}(4,2,1,1,7) + 744J_{j+1,1}(4,2,1,1,8) -$$  

$$2304J_{j+1,1}(4,2,1,1,10) \right\}, \tag{34}$$
The dimensionless nonlinearity parameter normalization factor $\Lambda$ with the accuracy and Fig. 2 for low powers ($\gamma < \tilde{\gamma}$) still zero. To find $\Delta C$ here $\exp_c$ we should calculate averages (35) for the zero order optimal input signal distribution $(k,n) = (1,1)$, taking into account $\gamma_0 = \gamma_0(\tilde{\gamma})$, $\delta_0 = \delta_0(\tilde{\gamma})$ to the capacity one should calculate the normalization factor $\Lambda$ with the accuracy $\beta^2$.

$$\Delta C = \int \mathcal{D}X P^{\text{opt}(\beta=0)}_{\text{x}}(X) \frac{\Lambda(2)}{\Lambda(0)}. \tag{36}$$

To find $\Delta C$ we should calculate averages (35) for the zero order optimal input signal distribution $P^{\text{opt}(\beta=0)}_{\text{x}}(X)$, see Eq. (31). It is easy to express averages (35) through the known functions $n_0(\tilde{\gamma})$, $\alpha_0(\tilde{\gamma})$ and integral $\tilde{j}(k,n;\alpha)$:

$$j_{(k_1,k_2,n_1,n_2,k)}^{j+1} \bigg|_{P[X]=P^{\text{opt}(\beta=0)}_{\text{x}}(X)} = 2^{k_1/2+k_2/2-n_1-n_2} n_0(\tilde{\gamma}) \int_0^1 dz^\gamma (1-z^\gamma)^{k-1} (k-1)! \times$$

$$\tilde{j}(k_1,n_1;\alpha_0(\tilde{\gamma}) - i\sqrt{3}\zeta')\tilde{j}(k_2,n_2;\alpha_0(\tilde{\gamma}) + i\sqrt{3}\zeta'), \tag{37}$$

where for given dimensionless parameter $\tilde{\gamma}$ the average $j_{(k_1,k_2,n_1,n_2,k)}^{j+1} \bigg|_{P[X]=P^{\text{opt}(\beta=0)}_{\text{x}}(X)}$ does not depend on the time $t_j$ and the integral (37) can be easily calculated numerically. This means that the formula (34) for $j_{(k_1,k_2,n_1,n_2,k)}^{j+1}$ from Eq. (37) determines the correction $\Delta C$ as the function of parameter $\tilde{\gamma}$ and we can omit $\frac{1}{M^p} \sum_{M=0}^{M-1}$ therein. Note that representations (34) is not unique.

Repeating the variation procedure introduced in [6] it is easy to see that taking into account the corrections proportional to $\beta$ in $\Lambda$ one obtains that the optimal input signal distribution $P^{\text{opt}}_{\text{x}}(X)$ acquires an angular dependance, but the correction to the capacity proportional to $\beta$ is still zero.

The numerical results of our estimations of the capacity correction $\Delta C$ are shown in Fig. 1 and Fig. 2 for low powers ($\tilde{\gamma}_\text{noise} < \tilde{\gamma} < 5$) and high powers ($1 < \tilde{\gamma} < 800$), correspondingly, both calculated within the intermediate power regime: $\tilde{\gamma}_\text{noise} \ll \tilde{\gamma} \ll (\tilde{\gamma}_\text{noise})^{-1}$. Here $\tilde{\gamma}$ is the dimensionless nonlinearity parameter $\tilde{\gamma} = \gamma LP/(\delta_0 \sqrt{3})$, $\tilde{\gamma}_\text{noise} = \gamma L^2 Q/(\delta_0 \sqrt{3})$, so that

Figure 1. The correction $\Delta C$ to the channel capacity as a function of parameter $\tilde{\gamma} = \gamma LP/(\delta_0 \sqrt{3})$ for low powers;

Figure 2. The correction $\Delta C$ to the channel capacity as a function of parameter $\tilde{\gamma} = \gamma LP/(\delta_0 \sqrt{3})$ for high powers.
SNR = \frac{\gamma}{\gamma_{\text{noise}}}$. In Fig. 1 and Fig. 2 we present the ratio $\Delta C/(M'\tilde{\beta}^2)$, where $\tilde{\beta} = \beta L/\delta^2$, and $M'$ is total time mesh points (independent time channels). The correction $\Delta C$ is positive in the intermediate power regime, i.e., it enhances the channel capacity of a nondispersive channel [6]. For typical fiber optical links [8] one has: $L = 1000$ km, $\gamma = 1.31$ (W/km)$^{-1}$, $2\pi/\delta t = 100$ GHz, $P_{\text{noise}} = QL/\delta t = 5.3 \times 10^{-4}$ mW. For these parameters one has $\gamma_{\text{noise}} \approx 4 \times 10^{-7}$, and our results are reliable if $\tilde{\beta} \ll 1$, i.e., $\beta \ll 1$ ps$^2$/km.

5. Concluding remarks
We present the path-integral approach [3, 5, 6, 9] to the calculation of the information theory quantities (conditional probability density function $P[Y|X]$, output signal correlators, mutual information $I_P[X]$, optimal input signal distribution $P_{\text{opt}}[X]$, channel capacity $C$, and so on) for the communication channel modelled by the nonlinear Shrödinger equation with additive white Gaussian noise at large signal-to-noise ratio. We apply this approach to the channel with the small dispersion parameter and find the first and the second corrections to $P[Y|X]$ perturbatively. Then we apply our method to the calculation of the simple output signal correlators. Finally on the simple model assumptions (the noise bandwidth $W'$ is of the same order as the input signal bandwidth $W$) we consider the estimations for the channel capacity correction $\Delta C$ in small dispersion parameter in the intermediate power regime. Of course, this model assumptions are far from the realistic channel but our approach allows one the calculate the correlators, the mutual information and the channel capacity corrections for small dispersion in general case (for various $W'$, $W$ and various output signal detectors with the bandwidth $W_r$) as well.

Acknowledgments

Acknowledgments Part of the work, i.e., the calculation of the conditional PDF $P[Y|X]$ and the channel capacity estimations for small dispersion, was supported by the Russian Foundation for Basic Research (RFBR), Grant No. 16-31-60031. The calculations of the small dispersion corrections to output signal correlators have been performed with support of the the Russian Science Foundation (RSF), Grant No. 16-11-10133.

References

[1] Shannon C E 1949 Communication in the presence of noise. Proc. Institute of Radio Engineers vol 37 p 1
[2] Martin P C, Sigia E D and Rose H A 1973 Phys. Rev. A 8 423
[3] Terekhov I S, Vergeles S S and Turitsyn S K 2014 Phys. Rev. Lett. 113 230602
[4] Hochberg D, Molina-Paris C, Perez-Mercader J and Visser M Phys. Rev. E 60 6343
[5] Terekhov I S, Reznichenko A V and Turitsyn S K 2016 Phys. Rev. E 94 042203
[6] Terekhov I S, Reznichenko A V, Kharkov Ya A and Turitsyn S K 2017 Phys. Rev. E 95 062133
[7] Itzykson C, Zuber J B 1980 Quantum Field Theory (New York: McGraw-Hill Book Company)
[8] Essiambre R-J, Kramer G, Winzer P J, Foschina G J and Goebel B 2010 J. of Lightwave Technol. 28 662
[9] Panarin A A, Reznichenko A V and Terekhov I S 2016 Phys. Rev. E 95 012127