



# Non-Gaussian error probability in optical soliton transmission

G. Falkovich<sup>a</sup>, I. Kolokolov<sup>b</sup>, V. Lebedev<sup>c</sup>, V. Mezentsev<sup>d,\*</sup>, S. Turitsyn<sup>d</sup>

<sup>a</sup> *Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel*

<sup>b</sup> *Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia*

<sup>c</sup> *Landau Institute for Theoretical Physics, Kosygina 2, Moscow 117940, Russia*

<sup>d</sup> *Aston University, Birmingham B4 7ET, UK*

Received 10 September 2001; received in revised form 14 October 2003; accepted 6 January 2004

Communicated by C.K.R.T. Jones

## Abstract

We find the probability distribution of the fluctuating parameters of a soliton propagating through a medium with additive noise. Our method is a modification of the instanton formalism (method of optimal fluctuation) based on a saddle-point approximation in the path integral. We first solve consistently a fundamental problem of soliton propagation within the framework of noisy nonlinear Schrödinger equation. We then consider model modifications due to in-line (filtering, amplitude and phase modulation) control. It is examined how control elements change the error probability in optical soliton transmission. Even though a weak noise is considered, we are interested here in probabilities of error-causing large fluctuations which are beyond perturbation theory. We describe in detail a new phenomenon of soliton collapse that occurs under the combined action of noise, filtering and amplitude modulation.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Soliton; Non-Gaussian statistics; Error probability; Optical communication

## 1. Introduction

Solitons play an important role in the dynamics and statistics of nonlinear systems in fields as diverse as hydrodynamics, plasmas, nonlinear optics, molecular biology, solid state physics, field theory, and astrophysics. Presumably the most impressive practical implementation of the fundamental soliton concept has been achieved in fiber optics, where soliton pulses are used as the information carriers (elementary “bits”) to transmit digital signal at high bit rates over long distances. Fiber optic applications of the soliton theory are governed by the integrable nonlinear Schrödinger equation (NLSE) and its modifications related to different control elements introduced into the optical line. The limitations on the error-free transmission distance are set mainly by the spontaneous emission noise added by in-line optical amplifiers. Even though the noise is weak one cannot generally use a perturbation approach to obtain the error probability because errors occur when signal changes substantially [2,3]. A priori it is not even clear whether one may still consider signal as a soliton-like or fluctuations with a substantial change of the waveform

\* Corresponding author.

*E-mail address:* [v.mezentsev@aston.ac.uk](mailto:v.mezentsev@aston.ac.uk) (V. Mezentsev).

determine the error probability. Large rare fluctuations in a nonlinear system are typically beyond the area of applicability of usual Gaussian statistics [2–4]. On the other hand, neither experiment nor direct numerical simulations are presently capable to provide an adequate statistics of such rare errors so that theoretical methods are of utmost importance here. The method to derive the probability of rare events was suggested in [2] as a maximum-likelihood approach which boils down to finding an optimal fluctuation that provides a given large deviation of soliton parameters. The method is technically a saddle-point approximation in the path integral for probabilities and is indeed known to describe the tails of the probability density function [5].

This paper presents a consistent development of the optimal fluctuation approach for soliton-bearing systems. The brief sketch of this theory and main results have been reported in [6] and here we explain in more detail the developed technique and present numerical evidence of the predicted effect of a soliton collapse. The conditions on the noise level and propagation distance will be formulated for an optimal fluctuation to be close to a soliton with slowly varying parameters. That makes it possible to reduce the formally infinite-dimensional problem to the analysis of the finite set of soliton parameters and effectively find the error probability for a single soliton transmission under different control schemes. The probability density function (PDF) is essentially Gaussian for timing jitter [8–11] in systems without control and may have substantially non-Gaussian tails in systems with in-line filtering and amplitude modulation. Here, we consider a single-soliton propagation, the effects related to soliton interaction are described in [2,3,12–14].

Amplified spontaneous emission (ASE) noise added at each amplifier changes randomly the amplitude, position, frequency and phase of the solitons. The fiber dispersion converts frequency variations in the arrival times jitter known as the Gordon–Haus effect [8] (see also [9] where the mathematical theory of the effect has been developed). Contribution of the Gordon–Haus effect to the total bit error rate (BER) is the probability that a soliton (corresponding to elementary “one”) will arrive outside the detection window. Depending on the detector construction it may also happen that the amplitude jitter will lead to a non-zero probability that some soliton with an amplitude lower than the detection threshold will be registered as “zero”. Such rare events, nevertheless, can be of importance in the modern communication systems operating with BER less than  $10^{-9}$ . Existing theoretical models of soliton parameters fluctuations under the impact of ASE noise are based on the assumption of the Gaussian statistics. In this case, one needs to know only standard deviations of moments under consideration to find BER. Widely employed numerical method of evaluation of fiber transmission systems performance (the so-called  $Q$ -factor method) in the basic formulation assumes a Gaussian noise distribution on both the zero and one levels. Qualitative justification of this assumption for a single pulse is based on the fact that in the NLS equation the first-order perturbations are captured by the soliton. Effectively first-order perturbations only renormalize soliton parameters leading as a result to the Gaussian statistics. It is not obvious, however, that statistics in the nonlinear models governing systems with soliton control elements will still be Gaussian. In fact, as we will show below some in-line control leads to substantially non-Gaussian statistics even for the distribution of soliton position (arrival time). Note also that to the best of our knowledge, the Gaussian statistics for pure (without soliton control elements) NLSE soliton under effect of additive noise has never been proved rigorously. We would like to emphasize once more that relatively small deviations from the Gaussian distribution tails can result in a wrong estimate of BER (that must be less than  $10^{-9}$ ). Evidently, at such low level estimate of BER is very sensitive to correct determination of the distribution tails. To be specific, we consider here a practical soliton transmission system, however, we would like to point out that our basic purpose in this work is to introduce a powerful and rigorous mathematical method to analyze non-Gaussian statistics in general soliton-bearing systems with additive noise. For this reason we also limit consideration by a scalar model and focus only on the effects of the distributed additive noise.

The structure of the paper is as follows: We introduce the models and present some general relations for the spectrum of soliton perturbations in [Section 2](#). We pass to the reduced description in terms of the soliton parameters and derive their probability distribution for a noisy nonlinear Schrödinger equation (NSE) in [Section 3](#). The role of

the continuous spectrum of perturbations is analyzed in Section 4 where we formulate the conditions for optimal fluctuation to be close to soliton. Section 5 is devoted to the statistical analysis of different control schemes, here we describe what is probably the most interesting physical phenomenon discovered in this work: soliton collapse under the action of noise in the system with in-line filtering and amplitude modulation. Derivations of the main results used through the paper can be found in Appendices A, B and C.

## 2. General relations

The evolution of nonlinear wave packet under the action of an additive noise is governed by the noisy NSE:

$$-i\partial_t\Psi = \partial_x^2\Psi + 2|\Psi|^2\Psi + \xi. \quad (1)$$

In particular, in the fiber optic applications  $t$  is the coordinate along the line,  $x$  is time, both expressed in dimensionless units. We assume that  $\xi$  is a white noise characterized by the correlation function

$$\langle \xi(t_1, x_1)\xi^*(t_2, x_2) \rangle = D\delta(t_1 - t_2)\delta(x_1 - x_2), \quad (2)$$

where  $D$  is the amplitude of the noise.

In this paper we describe the statistics of noise-induced perturbations of a single soliton. We assume that at  $t = 0$  the ideal soliton signal occurs

$$\Psi(0, x) = \frac{1}{\cosh(x)}. \quad (3)$$

We will examine a probability distribution of different distortions of the signal at a finite “time”  $T > 0$  (which is the length of the communication line). Another important problem is to estimate the probability to detect a signal at a finite ‘time’  $T$  provided there is no signal at  $t = 0$ . Answers can be obtained after studying statistics of noise-induced fluctuations of the field  $\Psi$  around the ideal signal form (2). We assume here that the “line length”  $T$  is long  $T \gg 1$  and that the noise is weak  $D \ll 1$ . More precise conditions will be introduced below.

To determine the probability of a wrong detection of digital bit, first one should specify a particular measuring procedure. For example, the presence of the signal at a time  $t = T$  can be decided by comparing with the threshold the value of the integral

$$\int_{-l}^l dx \Psi^*(T, x)\Psi(T, x), \quad (4)$$

taken over a window  $(-l, l)$  at  $t = T$ . If the value of  $l$  is large enough then the integral (4) (playing the role of energy) is close to 2 for the soliton (3). The value of the threshold is generally put smaller than 2 but larger than the noise level. Errors in detection are due to events with the value of the integral (4) essentially smaller than 2. There are two leading processes which can lead to such decay of the detected energy. The first process is a drastic decrease of the amplitude that can be characterized by the integral

$$Q = \frac{1}{2} \int_{-\infty}^{\infty} dx \Psi^*(T, x)\Psi(T, x), \quad (5)$$

that is equal to unity for the ideal signal (2). The second effect is a shift of the soliton position. In other words, shift of the soliton as a whole which is characterized by the quantity

$$Y = \frac{1}{2Q} \int_{-\infty}^{\infty} dx x\Psi^*(T, x)\Psi(T, x), \quad (6)$$

giving the “mass center” position of the soliton. For the ideal signal (2)  $Y = 0$ . Therefore, below we focus analysis on the joint probability distribution function (PDF)

$$\mathcal{P}(Q, Y) = \left\langle \delta \left[ Q - \frac{1}{2} \int dx \Psi^*(T, x) \Psi(T, x) \right] \delta \left[ Y - \frac{1}{2Q} \int dx x \Psi^*(T, x) \Psi(T, x) \right] \right\rangle. \quad (7)$$

The averaging here should be performed over the statistics of the noise.

### 2.1. Modifications of the basic model by in-line control elements

As it is well known, one of the important effects leading to a loss of the soliton signal is the so-called Gordon–Haus effect [8,9] that manifests itself as fluctuations of the soliton position  $Y$  (6). One can estimate the variance as  $\langle Y^2 \rangle \sim DT^3$ . To suppress this effect a number of techniques have been proposed that mathematically can be described by adding different terms (corresponding to specific in-line control elements) to the right-hand side of the basic equation (1). All the additional terms are assumed to be small. Therefore, below all coefficients  $\epsilon$  in the equations are treated as small parameters.

An attractive technique to suppress timing jitter is to use filtering control [15]. Under the presence of a shallow parabolic filter and weak amplification (to offset the energy lost due to filtering), the equation for the light envelope propagation down the line takes the following form [7]

$$-i\partial_t \Psi = \partial_x^2 \Psi + 2|\Psi|^2 \Psi + \xi - i\epsilon \Psi - 3i\epsilon \partial_x^2 \Psi. \quad (8)$$

We have chosen here the relation between filtering and amplification in a way to keep the amplitude of the stationary soliton equal to unity. To suppress the generation of background that is inherent for this scheme, it was suggested to use filters with sliding frequency [16]. The equation in this case takes the following form [7]

$$-i\partial_t \Psi = \partial_x^2 \Psi + 2|\Psi|^2 \Psi + \xi - i\epsilon \Psi - 3i\epsilon(\partial_x + 2i\lambda t)^2 \Psi. \quad (9)$$

Note that filtering only suppresses the timing jitter, but does not stop its growth. To stabilize the timing jitter one can apply in-line intensity modulation control which modifies the basic model in the following way:

$$-i\partial_t \Psi = \partial_x^2 \Psi + 2|\Psi|^2 \Psi + \xi - i\epsilon_1 \Psi - i\epsilon_2 \partial_x^2 \Psi + i\epsilon_3 x^2 \Psi. \quad (10)$$

An alternative way to suppress the Gordon–Haus effect is to use in-line phase modulation that mathematically can be described in terms of the following model

$$-i\partial_t \Psi = \partial_x^2 \Psi + 2|\Psi|^2 \Psi - \epsilon x^2 \Psi. \quad (11)$$

### 2.2. Separation of the discrete spectrum

An important issue in the consideration of the soliton statistics is to analyze the effect of radiation and a possibility to separate discrete spectrum from continuum. There are four degrees of freedom corresponding to the deformations of a single soliton. Let us introduce parameters  $\alpha, \beta, \eta, y$  describing those degrees of freedom:

$$\Psi = \eta \exp(i\beta x + i\alpha + i\tau) [\cosh^{-1}(z) + v], \quad (12)$$

$$z = \eta(x - y), \quad d\tau = \eta^2 dt, \quad (13)$$

where  $\alpha, \beta, \eta, y$  are arbitrary functions of time, and we defined the “internal time”  $\tau$ . We have also introduced the field  $v(z, t)$  which accounts for the continuous spectrum. This part of the perturbation can be expanded over  $f_k$  which are the eigenfunctions of the linearized equation introduced in Appendix A:

$$\begin{pmatrix} v \\ v^* \end{pmatrix} = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} [s_k f_{-k}(z) + s_k^* \bar{f}_{-k}(z)].$$

It is possible now to rewrite the original equation for  $\Psi$  in terms of the new variables  $\alpha, \beta, y, \eta, s_k, s_k^*$ . Substituting (12) and (13) into (1) one gets

$$U_1 = U_2 + U_\xi + U_\epsilon, \quad (14)$$

$$\begin{aligned} U_1 = & -i \int_{-\infty}^{+\infty} \frac{dk}{2\pi} [\partial_t s_k f_{-k} + \partial_t s_k^* \bar{f}_{-k}] + (\partial_t \beta x + \partial_t \alpha) \begin{pmatrix} v \\ -v^* \end{pmatrix} - i \partial_t \zeta (1 + z \partial_z) \begin{pmatrix} v \\ v^* \end{pmatrix} + i \partial_t y \partial_z \begin{pmatrix} v \\ v^* \end{pmatrix} \\ & + (\partial_t \beta y + \partial_t \alpha) f_0(z) + \eta^{-1} \partial_t \beta f_2(z) + i \partial_t \zeta f_3(z) - i \eta \partial_t y f_1(z), \end{aligned} \quad (15)$$

$$\begin{aligned} U_2 = & \eta^2 \hat{L}_z \begin{pmatrix} v \\ v^* \end{pmatrix} - \beta^2 f_0 - 2i\beta\eta f_1 - \beta^2 \begin{pmatrix} v \\ -v^* \end{pmatrix} + 2i\beta\eta \partial_z \begin{pmatrix} v \\ v^* \end{pmatrix} + \frac{2\eta^2}{\cosh(z)} \begin{pmatrix} v^2 + 2vv^* \\ -2vv^* - (v^*)^2 \end{pmatrix} \\ & + 2\eta^2 vv^* \begin{pmatrix} v \\ -v^* \end{pmatrix}, \end{aligned} \quad (16)$$

$$U_\xi = -\eta^{-1} \exp(i\alpha + ix\beta + i\tau) \xi^* \begin{pmatrix} \eta^{-1} \exp(-i\alpha - ix\beta - i\tau) \xi \\ -\eta^{-1} \exp(i\alpha - ix\beta + i\tau) \xi^* \end{pmatrix}, \quad (17)$$

where  $\zeta = \ln \eta$  and the functions  $f_0, f_1, f_2, f_3$  and the operator  $\hat{L}$  are introduced in [Appendix A](#). Here we denoted by  $U_1$  the sum of the terms with time derivatives, by  $U_2$  those generated by dispersion and nonlinearity and by  $U_\xi$  the sum of the noise-generated terms. For different control schemes, one must add to the right-hand side of (14) the terms that originate from the additional terms in the right-hand sides of [Eqs. \(8\), \(9\), \(11\) and \(67\)](#), we shall call the sum of such terms  $U_\epsilon$  (see section below).

Now, to derive the equations for  $\alpha, \beta, \eta, y, s_k, s_k^*$  one should find projections of [Eq. \(14\)](#) onto the discrete and continuous spectra using relations given in [Appendix A](#). This is done in the next sections.

### 3. Truncated model

Our main goal in the present paper is to compute probabilities of large deviations of the soliton parameters from their initial values. In general, this requires to take into account the complete set  $\alpha, \beta, \eta, y$  and  $v$ . However, at  $T \gg 1$  such probabilities are shown here to be determined by fluctuations of the discrete variables. This is because the discrete modes are localized on the “core” of soliton and the integral effect of their fluctuations (continued in time) can be significant. On the other hand, dynamics of the field  $v$  spread its fluctuations over the whole space and only weakly influences the soliton. In the subsequent section it will be shown that the influence of the fluctuations attributed to the continuous spectrum on the statistics of the “soft” variables  $\alpha, \beta, \eta$  and  $y$  is negligible in the limit  $DT^2 \ll 1$ . The variables  $\eta, y$  and  $\alpha, \beta$  are dynamically coupled (and this interaction plays the key role in the soliton statistics). This means that we can restrict our consideration by an analysis of the discrete modes given by  $\alpha, \beta, \eta$  and  $y$ . Those variables have to be introduced nonlinearly to avoid growing in time contribution from them into the dynamics of continuous spectrum (so that neglecting the latter is justified).

In this section we consider the situation described by the basic [equation \(1\)](#). An impact of the terms corresponding to in-line control elements will be examined in the subsequent sections.

### 3.1. Truncated action

First, we consider the case  $v = 0$ ,  $U_\epsilon = 0$ . In physical terms this corresponds to soliton propagation in system without in-line elements and we neglect interaction of the soliton with the continuum radiation. The later assumption is discussed and justified in [Appendix C](#). Then the projection of [Eq. \(14\)](#) onto the functions  $f_0$ ,  $f_1$ ,  $f_2$  and  $f_3$  (written in [Appendix A](#)) gives the following set of ODEs:

$$\partial_t \zeta = \mathcal{E}_0, \quad \partial_t \beta = \mathcal{E}_1, \quad \partial_t y - 2\beta = \mathcal{E}_2, \quad \partial_t \alpha + y \partial_t \beta + \beta^2 = \mathcal{E}_3, \quad (18)$$

$$\begin{aligned} \mathcal{E}_0 &= - \int dx \frac{1}{\cosh(z)} I, & \mathcal{E}_1 &= \eta \int dx \frac{\tanh(z)}{\cosh(z)} R, & \mathcal{E}_2 &= -\eta^{-1} \int dx \frac{z}{\cosh(z)} I, \\ \mathcal{E}_3 &= - \int dx \frac{z \tanh(z) - 1}{\cosh(z)} R, \end{aligned} \quad (19)$$

$$R = \text{Re}[\xi \exp(-i\alpha - i x \beta - i \tau)], \quad I = \text{Im}[\xi \exp(-i\alpha - i x \beta - i \tau)]. \quad (20)$$

The projection introduces new noise terms  $\mathcal{E}_0$ ,  $\mathcal{E}_1$ ,  $\mathcal{E}_2$ ,  $\mathcal{E}_3$  which depend both on our initial noise  $\xi$  and the soliton parameters. Unambiguous definition of the mean values and variances of the new noises requires regularization and is presented in [Appendix B](#). Here we directly apply the results obtained there with the only difference is that the summation over  $\mu$  should be replaced by the integration over  $x$ . Calculating the pair correlation functions in accordance with [Eq. \(B.13\)](#) we get zero cross-correlations of  $\mathcal{E}_0$ ,  $\mathcal{E}_1$ ,  $\mathcal{E}_2$ ,  $\mathcal{E}_3$  and

$$\begin{aligned} \langle \mathcal{E}_0(0) \mathcal{E}_0(t) \rangle &= \frac{D}{\eta} \delta(t), & \langle \mathcal{E}_1(0) \mathcal{E}_1(t) \rangle &= \frac{1}{3} D \eta \delta(t), & \langle \mathcal{E}_2(0) \mathcal{E}_2(t) \rangle &= \frac{\pi^2 D}{12 \eta^3} \delta(t), \\ \langle \mathcal{E}_3(0) \mathcal{E}_3(t) \rangle &= \frac{1}{3} \left( 1 + \frac{\pi^2}{12} \right) \frac{D}{\eta} \delta(t). \end{aligned} \quad (21)$$

Next, we calculate the average values of  $\mathcal{E}_0$ – $\mathcal{E}_3$  in accordance with [Eq. \(B.12\)](#). The average values of  $\mathcal{E}_1$ ,  $\mathcal{E}_2$  and  $\mathcal{E}_3$  are zero and for  $\mathcal{E}_0$  we get

$$\langle \mathcal{E}_0 \rangle = \frac{D}{2\eta}, \quad \langle \mathcal{E}_1 \rangle = 0, \quad \langle \mathcal{E}_2 \rangle = 0, \quad \langle \mathcal{E}_3 \rangle = 0. \quad (22)$$

It is interesting that there is a single non-zero average  $\langle \mathcal{E}_0 \rangle$  that means a systematic increase of the soliton amplitude due to the noise. Although, in physical terms this is not a surprise to have such a growing amplitude in the system with a permanent pump of the energy (through noise), to the best of our knowledge this mathematical fact has not yet been pointed out in the literature. Note that the corresponding amplitude increase factor is small in the limit  $DT^2 \ll 1$  considered here.

As a result of the previous analysis we obtain a closed description of the dynamics of  $\alpha$ ,  $\beta$ ,  $y$  and  $\eta$  with the relations (21) and (22). The correlation functions of those parameters can be examined applying the effective action constructed in accordance with [Eq. \(B.17\)](#)

$$\begin{aligned} i\mathcal{I} &= 2 \int dt \left[ -i\mu_0 \left( \partial_t \zeta - \frac{D}{2\eta} \right) + i\mu_5 \partial_t \beta - i\mu_6 (\partial_t y - 2\beta) - i\mu_3 (\partial_t \alpha + y \partial_t \beta + \beta^2) \right] \\ &\quad - D \int \frac{dt}{\eta} \left[ 2\mu_0^2 + \frac{2\eta^2}{3} \mu_5^2 + \frac{\pi^2}{6\eta^2} \mu_6^2 + \frac{2}{3} \left( 1 + \frac{\pi^2}{12} \right) \mu_3^2 \right], \end{aligned} \quad (23)$$

where  $\mu_0$ ,  $\mu_5$ ,  $\mu_6$ ,  $\mu_3$  are auxiliary fields and  $\zeta = \ln \eta$ . For instance, the pair correlation function of  $y$  can be written as the following functional integral

$$\langle y(t_1) y(t_2) \rangle = \int \mathcal{D}\alpha \mathcal{D}\beta \mathcal{D}\zeta \mathcal{D}y \mathcal{D}\mu_0 \mathcal{D}\mu_5 \mathcal{D}\mu_6 \mathcal{D}\mu_3 \exp(i\mathcal{I}) y(t_1) y(t_2). \quad (24)$$

In a similar way one can present the PDF (7). Next step is to drop  $v$  in Eq. (12). Then the integrals (5) and (6) are simply  $Q = \eta(T)$  and  $Y = y(T)$ . Therefore, we can rewrite Eq. (7) as

$$\mathcal{P}(Q, Y) = \int \mathcal{D}\alpha \mathcal{D}\beta \mathcal{D}\zeta \mathcal{D}y \mathcal{D}\mu_0 \mathcal{D}\mu_5 \mathcal{D}\mu_6 \mathcal{D}\mu_3 \exp[i\mathcal{I}]. \quad (25)$$

Due to causality, the integral over  $t$  in Eq. (23) can be taken from  $t = 0$  to  $t = T$ . The integral here is over the quantities which are arbitrary functions of time on the interval  $(0, T)$  with the following boundary conditions

$$t = 0: \quad \eta = 1, \quad y = 0, \quad \beta = 0, \quad \alpha = 0, \quad (26)$$

$$t = T: \quad Q = \eta(T), \quad Y = y(T), \quad \mu_6 = \mu_3 = 0. \quad (27)$$

The integration over  $\alpha$  in Eq. (25) and the boundary condition (27) give  $\mu_3 = 0$  for all times  $t$ . Therefore, this degree of freedom can be dropped in the following consideration. Next, the term with linear drift of  $\zeta$  in Eq. (23) does not play any essential role in the subsequent analysis (since we are interested in values  $\zeta \gg Dt$ ) and can be neglected. Then,  $\mathcal{P}(Q, Y)$  reads

$$\begin{aligned} \mathcal{P}(Q, Y) &= \int \mathcal{D}\beta \mathcal{D}\zeta \mathcal{D}y \mathcal{D}\mu_0 \mathcal{D}\mu_5 \mathcal{D}\mu_6 \exp(i\mathcal{I}_b), \\ \mathcal{I}_b &= \int_0^T dt \left\{ 2\mu_0 \partial_t \zeta - 2\mu_5 \partial_t \beta + 2\mu_6 (\partial_t y - 2\beta) + \frac{iD}{\eta} \left[ 2\mu_0^2 + \frac{2\eta^2}{3} \mu_5^2 + \frac{\pi^2}{6\eta^2} \mu_6^2 \right] \right\}. \end{aligned} \quad (28)$$

We can now integrate over  $\beta$  which yields the relation  $\partial_t \mu_5 = 2\mu_6$ . Substituting the relation into Eq. (28) we get an action for  $y$  and  $\eta$ . Recall that we are interested in times  $t \gg 1$ . This allows us to neglect in this integral the term  $(\partial_t \mu_5)^2$  compared to  $\mu_5^2$ . One should be careful since in this case we have to impose the condition  $\partial_t y = 0$  at  $t = 0$ . The resulting effective action is

$$\tilde{\mathcal{I}} = \int_0^T dt \left\{ [2\mu_0 \partial_t \zeta + \partial_t \mu_5 \partial_t y] + iD \left[ 2\mu_0^2 e^{-\zeta} + \frac{2}{3} e^{\zeta} \mu_5^2 \right] \right\}, \quad (29)$$

where  $e^{\zeta} = \eta$ . To find the PDF  $\mathcal{P}(T, Q, Y)$  one has to calculate the functional integral

$$\mathcal{P}(T, Q, Y) = \int \mathcal{D}\zeta \mathcal{D}y \mathcal{D}\mu_0 \mathcal{D}\mu_5 \exp(i\tilde{\mathcal{I}}), \quad (30)$$

where now the functions  $\zeta$  and  $y$  must satisfy the boundary conditions

$$\zeta(0) = 0, \quad y(0) = 0, \quad \partial_t y(0) = 0, \quad \exp[\zeta(T)] = Q, \quad y(T) = Y, \quad \mu_5(T) = 0. \quad (31)$$

### 3.2. Saddle-point calculations

We shall calculate the integral (30) in the saddle-point approximation. This approximation is justified by a smallness of the probability to observe deviations of order unity [5]. In this approximation, the probability of a rare event is determined by a single realization that realizes probability maximum:

$$\ln \mathcal{P} \approx i\tilde{\mathcal{I}}_{\text{saddle}}, \quad (32)$$

where  $\tilde{\mathcal{I}}_{\text{saddle}}$  is the saddle-point value of the effective action (29). Thus, the original problem is reduced to determination of a saddle-point, which has to be computed by extremum conditions for the effective action.

The extremum conditions for the action (29) are given by the following differential equations

$$i\partial_t \zeta + 2D\mu_0 e^{-\zeta} = 0, \quad i\partial_t \mu_0 + D\mu_0^2 e^{-\zeta} - \frac{1}{3} D e^{\zeta} \mu_5^2 = 0, \quad (33)$$

$$i\partial_t^2 y = \frac{4}{3} D e^{\zeta} \mu_5, \quad \partial_t^2 \mu_5 = 0. \quad (34)$$

The saddle-point value of the action (29) can be rewritten using these equations as

$$i\tilde{Z}_{\text{saddle}} = 2D \int_0^T dt \left[ \mu_0^2 \exp(-\zeta) + \frac{1}{3} \exp(\zeta) \mu_5^2 \right]. \quad (35)$$

A solution of the ODEs (34) with the boundary conditions (31) can be found as

$$\mu_5 = \frac{3i\lambda}{D}(T-t), \quad y = 4 \int_0^t dt' \int_0^{t'} dt'' \lambda(T-t'') \exp[\zeta(t'')], \quad (36)$$

where  $\lambda$  is a parameter to be determined from the other boundary conditions. This leads to

$$\partial_t^2 e^{\zeta/2} = -\lambda^2(T-t)^2 e^{\zeta/2}, \quad \mu_0 = -\frac{i}{2D} \partial_t e^{\zeta}. \quad (37)$$

A solution of Eq. (37) for  $e^{\zeta/2}$  is

$$e^{\zeta/2} = \sqrt{T-t} \{ C_1 J_{1/4}[\frac{1}{2}\lambda(T-t)^2] + C_2 J_{-1/4}[\frac{1}{2}\lambda(T-t)^2] \}, \quad (38)$$

where the constants  $C_1$  and  $C_2$  have to be determined from the boundary conditions. For example,

$$\sqrt{2}C_2 = \Gamma(\frac{3}{4})\lambda^{1/4}\sqrt{Q}.$$

Below we assume that  $\eta \sim 1$  (though the difference  $\eta - 1$  can be much larger than its average value). Then we can estimate from Eqs. (34) and (36)  $\lambda \sim Y/T^3$ . Therefore, if we are interested in the values  $Y \ll T$  (since we assumed  $T \gg 1$  then  $Y$  can be much larger unity), then  $\lambda T^2 \ll 1$ . Under this condition only first terms of the expansion of the Bessel functions in Eq. (38) are relevant and we can use in the analysis only first terms of this expansion. Moreover, under the condition  $Y \ll T$  an alternative perturbation scheme is applicable. Namely, at the condition  $Y \ll T$  the contribution to the effective action (29) associated with  $\mu_5, y$  can be treated as a perturbation. Therefore we can first solve the problem for  $\zeta, \mu_0$ , neglecting  $\mu_5, y$ , and then examine the dynamics of  $\mu_5, y$  on the background of this zero solution for  $\zeta, \mu_0$ . Then the correction to the action related to  $\mu_5, y$  can be found by substituting those solution for  $\mu_5, y$  into the  $\mu_5, y$ -dependent part of the effective action (29) or (35). Indeed this scheme is realized below.

After neglecting the term with  $\mu_5$  in Eq. (33) the set of equations for  $\zeta$  and  $\mu_0$  can be easily solved and we get for  $e^{\zeta/2}$  an explicit expression

$$\exp\left(\frac{\zeta}{2}\right) = 1 + \frac{t}{T}(\sqrt{Q} - 1), \quad (39)$$

which solves the zero approximation problem for the dynamics of  $\zeta, \mu_0$ . Next, substituting the expression (39) into the double integral (36) for  $y$  we get

$$Y = y(T) = 4\lambda T^3 \left( \frac{1}{5} + \frac{1}{10}\sqrt{Q} + \frac{1}{30}Q \right). \quad (40)$$

Eq. (37) allows us to rewrite  $e^{-\zeta} \mu_0^2$  as  $D^{-2}[\partial_t \exp(\zeta/2)]^2$ . Using the explicit expressions (36) for  $\mu_5$  and (39) for  $e^{\zeta/2}$  and expressing  $\lambda$  from Eq. (40) we obtain:

$$\ln \mathcal{P}(T, Q, Y) \approx -\frac{2}{DT} \left( \sqrt{Q} - 1 \right)^2 - R(Q) \frac{9Y^2}{8DT^3}, \quad (41)$$

$$R(Q) = \frac{[1/10 + (4/5)\sqrt{Q} + (1/10)Q]}{[3/5 + (3/10)\sqrt{Q} + (1/10)Q]^2}. \quad (42)$$

Recall that the expressions (41) and (42) are valid only for  $Y \ll T$ .



It is seen that for  $DT \ll 1$  the PDF (41) has a sharp maximum at  $\eta = 1$ . Considering fluctuations of  $y$  under condition  $\eta \approx 1$  yields the Gaussian statistics with

$$\langle Y^2 \rangle = \frac{4DT^3}{9}.$$

Besides for  $1 - Q \sim 1$  we get a non-trivial tail corresponding, roughly speaking, to a Gaussian distribution for  $\sqrt{Q} - 1$ . Distribution of  $Y$  at a given  $Q$  is qualitatively the same as for  $Q = 1$ . Namely, the same estimate  $Y \sim D^{1/2}T^{3/2}$  is correct, but the coefficients are different.

### 3.3. Probability to generate soliton from noise

Let us consider a probability to observe a soliton-like pulse (3) at a time  $t = T$  provided no signal exists at  $t = 0$ . Consider  $\Psi = 0$  at  $t = 0$ . Assume that the corresponding optimal fluctuation is close to a soliton. Then it is natural to start from the reduced action (29) derived in the previous section. Correspondingly, we can use the instanton (saddle-point) Eqs. (33) and (34). However, we should impose different boundary conditions. Namely,  $\eta = 0$  at  $t = 0$  and  $\eta = \eta_f$  at  $t = T$ . As to  $\mu_5$  we should take  $\mu_5 = 0$  since we are not interested in the initial position and the initial “velocity”  $\partial_t y$  of the soliton. Extrema conditions over the initial position  $y$  and over the initial “velocity”  $\partial_t y$  give  $\mu_5(0) = 0$ ,  $\partial_t \mu_5 = 0$ . Then  $\mu_5 = 0$  as a consequence of Eq. (34). Therefore the equation for  $\eta$  is Eq. (37) with  $\lambda = 0$ . It can be easily solved and we get a solution

$$\sqrt{\eta} = \frac{t}{T} \sqrt{\eta_f}, \quad (43)$$

satisfying the boundary conditions. Next, we obtain from Eq. (37) for  $\mu_0$

$$\mu_0 = -i \frac{t}{DT^2} \sqrt{\eta_f}. \quad (44)$$

Substituting Eqs. (43) and (44) into the action (29) we get

$$i\tilde{\mathcal{I}}_{\text{saddle}} = -\frac{2}{DT} \eta_f. \quad (45)$$

The probability to have a unit signal at zero input can be estimated as  $\exp(i\tilde{\mathcal{I}}_{\text{saddle}})$ , it grows with increasing  $T$ .

## 4. Role of continuous spectrum

The crucial point in the above analysis is that the contribution of the continuous spectrum to the PDF  $\mathcal{P}(T, Q, Y)$  is neglected provided  $T \gg 1$ . This is a manifestation of adiabaticity and of the fact that the first-order perturbations are incorporated into the discrete modes.

We describe here the role of the continuous spectrum  $s_k, s_k^*$  in terms of an effective action similar to the one examined above. The starting point of the description is the equation for  $s$  following from Eq. (14). Let us introduce the auxiliary fields  $v_k, v_k^*$  conjugated to  $s_k^*, s_k$ . Then the second-order term in the effective action determined by the terms  $U_1$  and  $U_2$  in Eq. (14) can be written as

$$i\mathcal{I}_{\text{cont}}^{(2)} = - \int_0^\infty d\tau \int_{-\infty}^{+\infty} \frac{dk}{2\pi} [i v_k \partial_\tau s_k^* - (k^2 + 1) v_k s_k^*] - \int_0^\infty d\tau \int_{-\infty}^{+\infty} \frac{dk}{2\pi} [i v_k^* \partial_\tau s_k + (k^2 + 1) v_k^* s_k]. \quad (46)$$

Here we used the relation

$$\tilde{L}_z \begin{pmatrix} v \\ v^* \end{pmatrix} = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} (k^2 + 1) [s_k f_{-k} - s_k^* \bar{f}_{-k}].$$

To analyze the role of the noise term  $U_\xi$  in Eq. (14) it is convenient to introduce the objects

$$\begin{pmatrix} m^* \\ m \end{pmatrix} = \hat{\sigma}_3 \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \left\{ \left( \frac{k+i}{k-i} \right)^2 v_k \varphi_k - \left( \frac{k-i}{k+i} \right)^2 v_k^* \bar{\varphi}_k \right\}, \quad (47)$$

$$\begin{pmatrix} \mu^* \\ \mu \end{pmatrix} = \hat{\sigma}_3 \left( \mu_0 f_0 + i\eta \mu_5 f_1 + \left( \frac{\mu_6}{\eta} \right) f_2 + i\mu_3 f_3 \right). \quad (48)$$

Then in the main approximation over  $v$  this noise term is written as

$$\begin{aligned} i\mathcal{I}_\xi = & -D \int_0^\infty \frac{d\tau}{\eta^3} \left[ 2\mu_0^2 + \frac{2\eta^2}{3} \mu_5^2 + \frac{\pi^2}{6\eta^2} \mu_6^2 + \frac{1}{3} \left( 1 + \frac{\pi^2}{6} \right) \mu_3^2 \right] \\ & -D \int_0^\infty \frac{d\tau}{\eta^3} \int_{-\infty}^\infty dz (m^* \mu + \mu^* m + m^* m) + \dots, \end{aligned} \quad (49)$$

where dots mean terms linear in  $m$  and  $\mu$ , origin of which is explained in Appendix B. These terms can be neglected due to the same reason as the linear drift term in Eq. (23).

#### 4.1. Linear coupling

Next we analyze the linear coupling between the discrete and the continuous spectrum, related to the cross terms in Eq. (49). Such coupling formally can change the saddle-point equations for the discrete modes. Let us consider the extended saddle-point equations that include both the discrete and continuous degrees of freedom.

Due to the presence of the contribution (46) in the effective action the saddle-point equations for  $s$  and  $v$  contain terms like  $(k^2 + 1)s_k$  and  $(k^2 + 1)v_k$ . There are terms of the order  $v$  and  $m$  in the saddle-point equations. In the equation for  $v$  one can identify the terms that can be estimated as  $v^2$ ,  $v^3$ ,  $\partial_t \eta$ ,  $Dm$ ,  $D\mu$ . The terms  $\partial_t \eta$ ,  $D\mu$  can be estimated as  $1/T$  that gives an estimate  $v \sim 1/T$  which justifies neglecting the high-order terms like  $v^2$ ,  $v^3$  in the equation for  $v$ . Recall that the initial condition for  $v$  at  $t = 0$  is  $v = 0$  that does not contradict estimate  $v \sim 1/T$ . Next we examine terms of the order of  $vm$  and  $m\partial_t \eta$  in the equation for  $m$ . The above estimate shows that these terms can be neglected in comparison with the linear term giving the oscillations of  $s_k$  and  $s_k^*$ . Thus we conclude that the only “dangerous” (or relevant) contribution is related to the term  $\propto Dm$  in the right-hand side of the equation for  $v$  since it can produce a secular growth compensating the smallness of  $1/T$  (see the analysis below). Therefore, one ought to check that  $m \ll \mu$ . Actually this inequality is guaranteed if  $m \ll \mu$  at  $t = T$ . In general, it is enough to check it only for a final boundary condition imposed on the auxiliary field  $m$ ,  $\mu$ .

To prove the assertion, one has to verify the condition  $m \ll \mu$  at  $t = T$ . The final boundary condition for the complete saddle-point equations (including both the discrete and continuous degrees of freedom) corresponding to the PDF (7) is as follows

$$\begin{pmatrix} m + \mu \\ m^* + \mu^* \end{pmatrix} = u_1 \eta \left[ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cosh^{-1}(z) + \begin{pmatrix} v \\ v^* \end{pmatrix} \right] + \frac{u_2}{Q} \eta x \left[ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cosh^{-1}(z) + \begin{pmatrix} v \\ v^* \end{pmatrix} \right]. \quad (50)$$

Here  $u_1$  and  $u_2$  are Lagrange multipliers which should be chosen to ensure the relations (5) and (6). Thus we get at  $t = T$

$$\mu_0 = \left( u_1 + \frac{u_2 y}{Q} \right) \eta, \quad \mu_6 = u_1, \quad \mu_5 \sim \mu_3 \sim m \sim \mu_0 v \quad (51)$$

It is seen that the condition  $v \ll 1$  is self-consistent and this justifies the application of the reduced description.

#### 4.2. Perturbation theory

In this section we examine perturbation effects related to the continuous spectrum. It is convenient to start from the formally exact expression for the effective action  $\tilde{\mathcal{I}}$ :

$$\exp(i\tilde{\mathcal{I}}) = \int \mathcal{D}v\mathcal{D}v^*\mathcal{D}m\mathcal{D}m^* \exp(i\mathcal{I}), \quad (52)$$

where  $\mathcal{I}$  is the complete effective action. It can be written as

$$\mathcal{I} = \mathcal{I}_{\text{disc}} + \mathcal{I}_{\text{int}} + \mathcal{I}_{\text{cont}},$$

where  $\mathcal{I}_{\text{disc}}$  depends on  $\alpha, \beta, \eta, \gamma, \mu_0, \mu_5, \mu_6, \mu_3$  only;  $\mathcal{I}_{\text{cont}}$  depends on  $s, s^*, v, v^*$  solely, and  $\mathcal{I}_{\text{int}}$  is a cross term which involves both sets of variables. Our strategy is to calculate the integral (52) expanding both  $\mathcal{I}_{\text{cont}}$  and  $\mathcal{I}_{\text{int}}$  over  $s, s^*$  and to use the perturbation theory. The main term in the expansion of  $\mathcal{I}_{\text{cont}}$  can be extracted from Eqs. (46) and (49).

Let us introduce the Green function

$$\hat{G}(\tau_1 - \tau_2, z_1, z_2) = \left\langle \left( \begin{array}{c} v(\tau_1, z_1) \\ v^*(\tau_1, z_1) \end{array} \right) \left( \begin{array}{cc} m^*(\tau_2, z_2) & m(\tau_2, z_2) \end{array} \right) \right\rangle. \quad (53)$$

In the first non-vanishing approximation, any average quantity can be calculated in accordance with the definition

$$\langle O \rangle \equiv \int \mathcal{D}v\mathcal{D}v^*\mathcal{D}m\mathcal{D}m^* \exp(i\mathcal{I}_{\text{cont}}^{(2)}) O, \quad (54)$$

where  $O$  is an arbitrary functional of our fields. Then we get from (46)

$$\begin{aligned} \hat{G}(\tau, z_1, z_2) = & i \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \left[ \left( \frac{k-i}{k+i} \right)^2 \exp(i\tau + ik^2\tau) f_{-k}(z_1) \bar{\varphi}_k^T(z_2) \right. \\ & \left. - \left( \frac{k+i}{k-i} \right)^2 \exp(-i\tau - ik^2\tau) \bar{f}_{-k}(z_1) \varphi_k^T(z_2) \right] \hat{\sigma}_3. \end{aligned} \quad (55)$$

The expression (55) is applied to the interval  $\tau > 0$ , since at  $\tau < 0$  the Green function  $G = 0$  due to causality.

Similarly, in the main order we get from (54) and (49)

$$\left\langle \left( \begin{array}{c} v(\tau_1, z_1) \\ v^*(\tau_1, z_1) \end{array} \right) \left( \begin{array}{cc} v(\tau_2, z_2) & v^*(\tau_2, z_2) \end{array} \right) \right\rangle = -D \int_0^\infty d\tau \int_{-\infty}^\infty dz \hat{G}(\tau_1 - \tau, z_1, z) \hat{\sigma}_1 \hat{G}^T(\tau_2 - \tau, z_2, z). \quad (56)$$

We look for the principal contribution to the correlation function (56) at  $\tau_1, \tau_2 \gg 1$ . Substitution of the expression (55) into (56) yields an expression containing the integrals like

$$\int_{-\infty}^\infty dz \bar{\varphi}_k^T(z) \hat{\sigma}_1 \varphi_q(z) = 2\pi \delta(k - q) + \dots,$$

where dots mean terms regular over  $k, q$ . Only such  $\delta$ -functions make the principal contribution to (56) since they produce terms which do not oscillate with  $\tau$ . The total contribution can be found as

$$\left\langle \left( \begin{array}{c} v(\tau_1, z_1) \\ v^*(\tau_1, z_1) \end{array} \right) \left( \begin{array}{cc} v(\tau_2, z_2) & v^*(\tau_2, z_2) \end{array} \right) \right\rangle \approx D\tau_* \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \{ \exp[i(k^2 + 1)(\tau_1 - \tau_2)] f_k(z_1) \bar{f}_k^T(z_2) + \text{c.c.} \}, \quad (57)$$

where the factor  $\tau_* = \min(\tau_1, \tau_2)$  originates from the integral over  $\tau$ .

The most important contribution to  $\tilde{\mathcal{I}}$  is related to the third-order interaction term which can be presented as

$$\begin{aligned} i\mathcal{I}_{\text{int}} = & -2 \int_0^\infty d\tau \int_{-\infty}^\infty dz \cosh^{-2}(z) \left\{ \left( \mu_0 + \frac{\mu_6 z}{\eta} \right) (v^2 - (v^*)^2) \right. \\ & \left. + i[\eta\mu_5 \tanh z + \mu_3(z \tanh z - 1)](v^2 + 4vv^* + (v^*)^2) \right\}, \end{aligned} \quad (58)$$

Let us, for instance, consider as an example the first contribution from Eq. (58)

$$i\mathcal{I}_{\text{int},1} = -2 \int_0^\infty d\tau \int_{-\infty}^\infty dz \cosh^{-2}(z) \mu_0 (v^2 - (v^*)^2). \quad (59)$$

In the second-order of the perturbation theory the term (59) produces the following contribution  $\Delta\mathcal{I}$  to  $\tilde{\mathcal{I}}$ :

$$i\Delta\mathcal{I} = \frac{1}{2} \langle (i\mathcal{I}_{\text{int},1})^2 \rangle,$$

where averaging is determined in accordance with (54). Substituting here the expression (59) we get

$$i\Delta\mathcal{I} = 4 \int_0^\infty d\tau_1 \int_0^\infty d\tau_2 \int_{-\infty}^\infty \frac{dz_1}{\cosh^2(z_1)} \int_{-\infty}^\infty \frac{dz_2}{\cosh^2(z_2)} K(\tau_1, \tau_2, z_1, z_2) \mu_0(\tau_1) \mu_0(\tau_2), \quad (60)$$

$$K = \langle v_1 v_2 \rangle^2 + \langle v_1^* v_2^* \rangle^2 - \langle v_1 v_2^* \rangle^2 - \langle v_1^* v_2 \rangle^2, \quad (61)$$

where  $v_1 = v(\tau_1, z_1)$ ,  $v_2 = v(\tau_2, z_2)$ . The explicit expression for the principal contribution to the kernel (61) can be extracted from (57). Such a contribution (60) can be estimated as

$$i\Delta\mathcal{I} \sim (DT)^2 \int d\tau \mu_0^2, \quad (62)$$

provided there are no divergences. The factors  $\cosh^{-2}$  in (60) ensure convergence of the integrals over  $z_1$  and  $z_2$ . Therefore the most “dangerous” divergence is related to a slow decrease of the kernel as  $\tau = \tau_1 - \tau_2$  increases. To examine such a possibility we calculate the integral

$$\int d\tau K = D^2 T^2 \int_{-\infty}^{+\infty} \frac{dk}{2\pi|k|} \{ \bar{f}_k^T(z_1) \hat{\sigma}_3 f_k(z_1) f_k^T(z_2) \hat{\sigma}_3 \bar{f}_k(z_2) + \bar{f}_{-k}^T(z_1) \hat{\sigma}_3 f_k(z_1) f_{-k}^T(z_2) \hat{\sigma}_3 \bar{f}_k(z_2) \}. \quad (63)$$

Note that only localized parts of the eigenfunctions  $f_k(z)$  contribute to (63). Using the expressions (A.4) and (A.5) one can easily check that the integral (63) converges both for small and for large  $k$ . The property confirms the estimate (62). This fluctuation contribution has to be compared with the bare term in (23) containing  $\mu_0$ . Then we conclude that the fluctuation contribution is negligible if

$$DT^2 \ll 1. \quad (64)$$

It is also worth estimating the modification of the effective action  $i\Delta\mathcal{I}$  produced by the  $\mu_5$ -dependent term in (58). It can be written in the form analogous to (61) where  $\mu_0$  is substituted by  $\mu_5$  and the new kernel  $\tilde{K}$  along with the terms of the same structure as in  $K$  contains the term  $|\langle v_1 v_2^* \rangle|^2$ . This is potentially “dangerous” because this term has non-oscillating part so that the corresponding integral over  $dk$  analogous to (63) is not convergent at small  $k$ . However, this diverging part is even with respect to two independent transformations  $z_1 \rightarrow -z_1$  and  $z_2 \rightarrow -z_2$  and its integration with the odd functions  $\tanh z_1$  and  $\tanh z_2$  vanishes. Physically it is transparent since small  $k$  correspond to large distances where the presence of the soliton is unimportant. In this limit the correlation function  $\langle v_1 v_2^* \rangle$  is simply the Green function of a free Schödinger particle with the factor  $D \min(\tau_1, \tau_2)$ . Its square modulus does not depend on  $z_1$  and  $z_2$  and is obviously even. The contribution to  $\int d\tau \tilde{K}$ , which is localized on the soliton,

has the same estimate as (63) and this confirms the criterion (64). We can see also that there are no such cancellations for the contribution  $\sim \int d\tau \mu_3^2$  to the effective action. However, the signal statistics does not depend on the dynamics of the variables  $\mu_3$  and  $\alpha$ .

## 5. Statistics of soliton in the modified models

In this section we analyze the additional terms describing different modifications of the basic equation (1) and their impact on soliton statistics. The modifications are described in Section 2.1.

Instead of (18), we get for the system with in-line filtering

$$\partial_t \eta = 2\epsilon\eta - 6\epsilon(\beta^2 + \frac{1}{3}\eta^2)\eta + \Xi_0, \quad \partial_t \beta = 4\epsilon\eta^2\beta = \Xi_1, \quad \partial_t y - 2\beta = \Xi_2, \quad \partial_t \alpha + y\partial_t \beta + \beta^2 = \Xi_3. \quad (65)$$

Without the noise term, the only possible steady-state values are  $\eta_s = 1$  and  $\beta_s = 0$ . The frequency and amplitude dispersions saturate:  $\langle(\eta - 1)^2\rangle, \langle\beta^2\rangle \sim D\epsilon^1[1 - \exp(-8\epsilon t)]$ —see (5.149–50) from [7]. Therefore,  $\langle y^2\rangle \propto DT$  so the Gordon–Haus effect is suppressed. Since the background grows exponentially  $\propto \exp(2\epsilon t)$  then it is this growth that sets the limit to the transmission distance.

In the case of sliding frequency filter control the equations governing evolution of the soliton parameters are

$$\partial_t \eta = 2\epsilon\eta - 6\epsilon[(\beta + 2\lambda t)^2 + \frac{1}{3}\eta^2]\eta + \Xi_0, \quad \partial_t \beta + 4\epsilon\eta^2(\beta + 2\lambda t) = \Xi_1, \\ \partial_t y - 2\beta = \Xi_2, \quad \partial_t \alpha + y\partial_t \beta + \beta^2 = \Xi_3. \quad (66)$$

Now, without the noise we have a soliton drifting over frequencies:

$$\beta_s = -2\lambda t + \frac{\lambda}{2\epsilon\eta_s}, \quad \eta_s^2 + 3\left(\frac{\lambda}{2\epsilon\eta_s}\right)^2 = 1 \quad (67)$$

with an acceleration  $y = -2\lambda t^2 + \lambda t/\epsilon\eta_s$ —see (5.226–30) from [7]. Such state is stable for  $\lambda^2 < 2\epsilon/3$  [7]. The dispersions of  $\eta$  and  $y$  are given by (5.276) and (5.284) from [7].

Below, we analyze two principal in-line control schemes: the phase modulation and the intensity modulation.

### 5.1. Phase modulation control

First, we consider the phase modulation case which is described by Eq. (11). An additional term with  $\epsilon$  leads to a modification of the basic equations for the discrete spectrum. To examine the corresponding changes in the PDF we can apply the same procedure as above.

Let us introduce a modified reduce action for the case of phase modulation control. Eq. (11) contains an extra term in comparison with (1) and that leads to the corresponding additional term  $U_\epsilon$  in the right-hand side of Eq. (14)

$$U_\epsilon = -\epsilon \left(\frac{z}{\eta} + y\right)^2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\cosh(z)} + \dots,$$

where dots mean terms proportional to  $v$ . Corresponding additional contribution to the reduced action reads

$$i\tilde{\mathcal{L}}_\epsilon = 4i\epsilon \int dt \mu_5 y + 2i\epsilon \int dt \mu_3 \left(\frac{\pi^2}{12\eta^2} - y^2\right).$$

Again, we can omit the term related to the dynamics of the phase  $\alpha$  so the only new contribution to the action will be  $\tilde{\mathcal{I}}_\epsilon = \int dt \mu_5 y$ . Varying the sum  $\tilde{\mathcal{I}} + \tilde{\mathcal{I}}_\epsilon$  we get the following saddle-point equations

$$\partial_t y = 2\beta + iD \frac{\pi^2}{6\eta^3} \mu_6, \quad \partial_t \beta = -2\epsilon y - \frac{2i}{3} D \eta \mu_5, \quad \partial_t \mu_5 = 2\mu_6, \quad \partial_t \mu_6 = -2\epsilon \mu_5. \quad (68)$$

Equations of motion for  $\eta$  and  $\mu_0$  are the same Eq. (33).

Eq. (68) for  $\mu_{5,6}$  have a harmonic solution. Taking into account the final condition  $\mu_5(T) = 0$  we obtain:

$$\mu_5(t) = \frac{3i\lambda}{2D} \sin[\Omega(T-t)], \quad (69)$$

where  $\Omega = 2\sqrt{\epsilon}$ , and the amplitude  $\lambda$  has to be defined from the final condition  $y = Y$ .

For small enough  $Y$  it is possible to treat the terms with  $\mu_5, \mu_6$  in the effective action (28) as a perturbation and to repeat the scheme developed in Section 3.2. First, we neglect the term with  $\mu_5$  in Eq. (33). Then we get the same answer (39), that is  $\eta(t) = (1+Bt)^2$ , where  $B = (\sqrt{Q} - 1)/T$ . The equations for  $y$  and  $\beta$  should be solved on this background. For small  $\epsilon$  we can neglect the term proportional to  $\mu_6$  in the right-hand side of the evolution equation (68) for  $y$ . Substituting then (69) into the equation of motion for  $\beta$  in (68) we readily solve the system of equations for  $\beta$  and  $y$  (it is harmonic oscillator with time-dependent external force having resonance frequency in its spectrum). Finally we get at  $T \gg \epsilon^{-1/2}$ :

$$y(T) = \frac{\lambda}{\Omega} T \left( 1 + BT + \frac{1}{3} B^2 T^2 \right). \quad (70)$$

Equating  $y(T)$  to  $Y$  we find the expression for  $\lambda$ :

$$\lambda = \frac{\Omega Y}{T} \frac{1}{1 + BT + B^2 T^2 / 3}. \quad (71)$$

A leading contribution to the logarithm of the PDF is given again by the expression (35). The  $\mu_0$ -dependent part expressed in terms of  $B$  and  $\lambda$  is not changed:

$$2D \int_0^T dt \mu_0^2 \exp(-\zeta) = -\frac{2}{DT} (BT)^2 + \frac{\lambda^2 BT}{DT} \left( 1 + \frac{1}{5} BT \right). \quad (72)$$

The term with  $\mu_5^2$  should be, however, recalculated. Using the expression (69) we obtain:

$$\frac{2D}{3} \int_0^T dt e^\zeta \mu_5^2 = -\frac{3\lambda^2}{4D} T \left( 1 + BT + \frac{B^2 T^2}{3} \right). \quad (73)$$

Rewriting  $\lambda$  via  $Y$  and  $Q$  according to (73), returning from  $\Omega$  and  $B$  to  $\epsilon$  and  $Q$ , and adding the term (72) we get finally:

$$\ln \mathcal{P}(T, Q, Y) \approx -\frac{2}{DT} \left( \sqrt{Q} - 1 \right)^2 - R_{\text{ph}}(Q) \frac{3\epsilon Y^2}{DT}, \quad (74)$$

$$R_{\text{ph}}(Q) = \frac{3}{1 + \sqrt{Q} + Q}. \quad (75)$$

Now we should establish an applicability condition for the expression (74). For that purpose, we estimate an effect related to the term with  $\mu_5$  in Eq. (33) omitted at treating the dynamics of  $\eta, \mu_0$ . We find a correction to  $\mu_0$ :

$$\Delta \mu_0 \sim D \int_0^T dt e^\zeta \mu_5^2 \sim \frac{\epsilon Y^2}{DT}. \quad (76)$$

Comparison of this expression with the terms containing  $\mu_0 \sim 1/(DT)$  (that follows from Eq. (37)) leads to the conclusion that this variation can be neglected if we consider fluctuations of  $Y$  less than  $1/\sqrt{\epsilon}$ .

Estimate of the impact of the continuous spectrum on the reduced dynamics proceeds in the same way as in Section 4.2. Actually, the perturbation  $\sim x^2\Psi$  of the evolution equation does not introduce any dissipation (in contrast to the case of intensity modulation control) and the linear in time Brownian growth of the mean square fluctuations amplitude remains. The modification affect long-distant properties of eigenfunctions of the linearized operator and slightly changes their frequencies. Namely, if we write the correlation function of the field  $v$  for  $t_1 > t_2$  as

$$\langle v_1 v_2^* \rangle = Dt_2 e^{i(t_1-t_2)} \mathcal{G}(t_1 - t_2; z_1, z_2) \quad (77)$$

then at distances  $x \sim \epsilon^{-1/2}$  where the perturbation becomes essential we can forget about the presence of the soliton and for  $\mathcal{G}$  we can use the expression (C.6) of the oscillator Green function with the real frequency  $\Omega = 2\sqrt{\epsilon}$ . For scales  $z \sim 1$  the perturbation of the evolution equation is negligible and we return to the formula (57) for  $\langle v_1 v_2^* \rangle$ ,  $\langle v_1 v_2 \rangle$ . Modification of dynamics of variables  $\mu_0$  and  $\eta$  is described by expressions (60) and (61) with the modified correlation function of the field  $v$ . Note that the structure of the kernel  $K$  is such that at distances  $z \gg 1$  it oscillates in time (see (77)) and an integration over  $\tau$  results in a negligible quantity. The only substantial contribution comes from  $z \sim 1$  for which the analysis made in (63) works. Thus, dynamics  $\mu_0$  and  $\eta$  remains unaffected by continuous spectrum fluctuations at  $DT^2 \ll 1$ . The contribution  $\sim \int d\tau \mu_0^2$  to the effective action is proportional to

$$\int d\tau \int dz_1 dz_2 \tanh z_1 \tanh z_2 |\mathcal{G}(\tau; z_1, z_2)|^2.$$

The distances  $z \sim \epsilon^{-1/2}$  do not contribute this integral because the square modulus  $|\mathcal{G}(\tau; z_1, z_2)|^2$  is even both in  $z_1$  and  $z_2$  (see (C.6)) while the functions  $\tanh z_{1,2}$  are odd. The only contribution which remains is that from the scales  $\sim 1$  where the perturbation  $\sim x^2$  can be neglected and we return to the analysis of Section 4.2. Thus, it is seen that the continuous spectrum is irrelevant under the condition  $DT^2 \ll 1$ .

## 5.2. Intensity modulation control

In this section, we analyze the statistics of soliton in systems with intensity modulation control which is governed by Eq. (10). Below we assume  $\epsilon_1 \sim \epsilon_2 \sim \epsilon_3 \ll 1$ .

In the intensity modulation scheme, the steady state without noise has  $\beta = y = 0$  and  $\epsilon_1 = \epsilon_2 \eta_s^2/3 + \epsilon_3 \pi^2/12\eta_s^2$ . It is linearly stable for  $4\epsilon_2 \eta_s^2 > \pi^2 \epsilon_3 \eta_s^{-2}$ . Another condition on parameters is imposed by requiring the stability of zero:  $\epsilon_3 > 2\epsilon_2^2 \eta_s^4/9$ —see (5.211) from [7]. Those two conditions are not contradictory. It is stated in [7] that the amplitude fluctuations saturate at the same level as for (8) while the timing variance is given by (5.192–3) and it saturates at

$$\langle Y^2 \rangle = \frac{D}{2\eta_s \epsilon_3} + \frac{3D\eta_s}{2\pi^2 \epsilon_3 \epsilon_2 \epsilon_1}. \quad (78)$$

Let us introduce a modified reduce action for the intensity modulation control case. There are additional terms in Eq. (10) comparing to Eq. (1) and they produce an additional term  $U_\epsilon$  in Eq. (14)

$$\begin{aligned} U_\epsilon = & -i\epsilon_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\cosh(z)} - i\epsilon_2 \eta^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \partial_z^2 \frac{1}{\cosh(z)} + i\epsilon_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{\beta^2}{\cosh(z)} + 2\epsilon_2 \eta \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} \partial_z \frac{1}{\cosh(z)} \\ & + i\epsilon_3 \left( \frac{z}{\eta} + y \right)^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\cosh(z)}. \end{aligned} \quad (79)$$

Extracting contributions to the equations for  $\alpha$ ,  $\beta$ ,  $y$  and  $\eta$  corresponding to Eq. (79) we get an additional contribution to the reduced action

$$\begin{aligned} i\tilde{\mathcal{L}}_\epsilon \equiv & - \int dt dz (\mu^* \mu) U_\epsilon = 4i\epsilon_1 \int dt \mu_0 - \frac{4}{3}i\epsilon_2 \int dt \mu_0 \eta^2 - 4i\epsilon_2 \int dt \mu_0 \beta^2 + \frac{8i}{3}\epsilon_2 \int dt \mu_5 \eta^2 \beta \\ & - 2i\epsilon_3 \int dt \mu_0 \left( \frac{\pi^2}{6\eta^2} + 2y^2 \right) - \frac{2\pi^2}{3}i\epsilon_3 \int dt \frac{\mu_6 y}{\eta^2}. \end{aligned} \quad (80)$$

The term (80) must be added to Eq. (23). As previously we can pass from Eq. (23) into Eq. (28). Varying  $\tilde{\mathcal{L}} + \tilde{\mathcal{L}}_\epsilon$  we get the following saddle-point equations

$$\partial_t y = 2\beta - \frac{\pi^2}{3}\epsilon_3 \frac{y}{\eta^2} + iD \frac{\pi^2}{6\eta^3} \mu_6, \quad \partial_t \beta = -\frac{4}{3}\epsilon_2 \eta^2 \beta - \frac{2i}{3}D\eta \mu_5, \quad (81)$$

$$\partial_t \zeta = 2 \left( \epsilon_1 - \frac{1}{3}\epsilon_2 \eta^2 - \epsilon_2 \beta^2 \right) - \epsilon_3 \left( \frac{\pi^2}{6\eta^2} + 2y^2 \right) + 2iD\eta^{-1} \mu_0. \quad (82)$$

Linearizing Eqs. (81) and (82) we conclude that  $\langle \zeta^2 \rangle \sim D/\epsilon$  due to the noise term in the right-hand side of Eq. (81). The probability of larger fluctuations of  $\zeta$  is related to the term  $\propto y^2$  in the right-hand side of Eq. (81). So, to investigate the deviations, we can omit the term with the noise in the right-hand side of Eq. (81). Then, fluctuations of  $\zeta$  will be forced by fluctuations of  $y$ . If  $y \ll 1$  then  $\zeta \ll 1$  also. The case is irrelevant for us if we are interested in a probability of the signal lost. If  $y \gg 1$  then a new phenomenon of soliton collapse takes place. Namely, in this case  $\eta \rightarrow 0$ . Let us reiterate that we describe here an optimal fluctuation which gives the maximum probability of a given  $y$ . If  $y$  is large enough, that is the soliton deviated too much into the region of high dissipation, then it indeed may disappear in a finite time. To describe the regime of decreasing  $\eta$  we can keep only the term  $\propto \eta^{-2}$  in the right-hand side of Eq. (81). Then we have

$$\partial_t \eta^2 = -\frac{1}{3}\epsilon_3 \pi^2,$$

what leads to  $\eta \rightarrow 0$  for a finite time. Therefore large  $y$  inevitably leads to the collapse. This effect can also be described in terms of the corresponding features of  $\mathcal{P}(Q, Y)$ . Namely, there is a critical value  $Y_{\text{cr}} \sim 1$  so that  $\mathcal{P}(Q, Y)$  falls into  $\delta(Q)$  if  $|Y| > Y_{\text{cr}}$ . Of course,  $Y_{\text{cr}}$  is a complicated function of  $Q$ ,  $\epsilon_1$ ,  $\epsilon_2$ ,  $\epsilon_3$ .

Collapse existence can be understood from the following simple analysis of a reduced system. The saddle-point equations (81) and (82) admit a solution with  $v_1 = v_2 = \beta = y = 0$ . This simplest case can be analyzed in more details. Since the equations follow from the minimal action principle, they can be rewritten in the usual Hamiltonian form

$$\partial_t v_0 = -\frac{\partial H}{\partial \zeta}, \quad \partial_t \zeta = \frac{\partial H}{\partial v_0}, \quad (83)$$

$$H = \frac{D}{2\pi} v_0 + \frac{D}{\eta} v_0^2 + 2\epsilon_1 v_0 - \frac{2}{3}\epsilon_2 v_0 \eta^2 - \tilde{\epsilon}_3 \frac{v_0}{\eta^2}, \quad (84)$$

where  $\tilde{\epsilon}_3 = \pi^2 \epsilon_3 / 6$ . In the explicit form the equations (83) read

$$\partial_t v_0 = \frac{D}{2\eta} v_0 + \frac{D}{\eta} v_0^2 + \frac{4}{3}\epsilon_2 v_0 \eta^2 - 2\tilde{\epsilon}_3 \frac{v_0}{\eta^2}, \quad (85)$$

$$\partial_t \zeta = \frac{D}{2\eta} + \frac{2D}{\eta} v_0 + 2\epsilon_1 - \frac{2}{3}\epsilon_2 \eta^2 - \tilde{\epsilon}_3 \frac{1}{\eta^2} = \frac{H}{v_0} + \frac{D}{\eta} v_0. \quad (86)$$



Note that for non-zero Hamiltonian and positive  $\eta$  the sign of the continuous function  $v_0(t)$  cannot be changed during evolution, because the integral of motion  $H = \text{const.}$  is directly proportional to  $v_0$ . Considering, initial condition with  $v_0(0) < 0$  we can easily estimate the derivative of the field  $\eta$  as

$$\partial_t \zeta = \frac{H}{v_0} + \frac{D}{\eta} v_0 < -2\sqrt{\frac{HD}{\eta}}.$$

The equation inevitably produces the collapse if  $H > 0$  since

$$\partial_t \zeta < -2\sqrt{\frac{HD}{\eta}} \rightarrow \sqrt{\eta} < 1 - \sqrt{HD}t.$$

We conclude from this inequality that the collapse time  $t_0$  is bounded from above as:  $t_0 < 1/\sqrt{DH}$ . Note that the case  $v_0 < 0$  corresponds to positive Hamiltonian  $H > 0$ . Numerically found collapsing solutions are presented below.

Let us examine now the region of parameters  $|Y| < Y_{\text{cr}}$ ,  $Q \sim 1$  that is  $y \sim 1$  and  $\zeta \sim 1$ . Then the life time of the corresponding instanton can be estimated as  $\epsilon^{-1}$ . Next, we come to estimates

$$\beta \sim \epsilon, \quad D\mu_5 \sim \epsilon^2, \quad \mu_0 \sim \mu_6 \sim \epsilon\mu_5.$$

It is a motivation to introduce rescaled fields and time:

$$\tilde{t} = \epsilon_1(T - t), \quad \beta = \epsilon_1\tilde{\beta}, \quad \mu_0 = -i\frac{\epsilon_1^3}{D}\tilde{\mu}_0, \quad \mu_5 = -i\frac{\epsilon_1^2}{D}\tilde{\mu}_5, \quad \mu_6 = -i\frac{\epsilon_1^3}{D}\tilde{\mu}_6.$$

Omitting irrelevant terms, we find an effective action

$$\begin{aligned} i\tilde{\mathcal{L}} = & \frac{\epsilon_1^3}{D} \int_0^\infty d\tilde{t} \left\{ 2\tilde{\mu}_0 \frac{\partial \zeta}{\partial \tilde{t}} - 2\tilde{\mu}_5 \frac{\partial \tilde{\beta}}{\partial \tilde{t}} + 2\tilde{\mu}_6 \left( \frac{\partial y}{\partial \tilde{t}} + 2\tilde{\beta} \right) + \frac{2\eta}{3} \tilde{\mu}_5^2 \right\} \\ & + \frac{\epsilon_1^3}{D} \int_0^\infty d\tilde{t} \left\{ 4\tilde{\mu}_0 \left( 1 - \frac{1}{3}\tilde{\epsilon}_2\eta^2 \right) - 2\tilde{\mu}_0\tilde{\epsilon}_3 \left( \frac{\pi^2}{6\eta^2} + 2y^2 \right) + \frac{8}{3}\tilde{\epsilon}_2\tilde{\mu}_5\eta^2\tilde{\beta} - \frac{2\pi^2}{3}\tilde{\epsilon}_3\tilde{\mu}_6\frac{y}{\eta^2} \right\}, \end{aligned} \quad (87)$$

where  $\tilde{\epsilon}_2 = \epsilon_2/\epsilon_1$ ,  $\tilde{\epsilon}_3 = \epsilon_3/\epsilon_1$ . Extrema of this effective action can be examined only numerically.

So, we can assert only that at  $|Y| < Y_{\text{cr}}$

$$\ln \mathcal{P}(Q, Y) = -\frac{\epsilon_1^3}{D} F \left( \frac{\epsilon_2}{\epsilon_1}, \frac{\epsilon_3}{\epsilon_1}, Q, Y \right), \quad (88)$$

where  $F$  is a function of order unity. Probably, the function  $F(Y)$  has a minimum at a  $Y \sim 1$ . Let us explain the existence of the minimum at  $1 - Q \sim 1$ . Remember, that  $Q, Y$  coincide practically with the final values of  $\eta$  and  $y$ . As is seen from Eq. (82) (where the term with  $\mu_0$  should be omitted in the approximation we use) to achieve a value  $1 - Q \sim 1$  at  $t = T$  the value of  $y^2$  should be of order unity during the evolution (since the life time of the instanton is  $\sim \epsilon^{-1}$ ). Therefore small  $Y$  means that the behavior of  $y$  is non-monotonic: First it grows from 0 to  $y \sim 1$  and then returns to a small value.

It is natural to expect, that such non-monotonic evolution costs more than a monotonic evolution at  $Y \sim 1$ . Therefore  $F$  seems to diminish at growing  $Y$ . Next, it is natural to expect that  $F$  grows when  $Y$  goes to  $Y_{\text{cr}}$ . Therefore there should be a minimum in between. Of course, the conclusion has to be confirmed by numerics.

Let us describe briefly the region  $|Y| > Y_{\text{cr}}$ . There is no instanton solution corresponding to the effective action (87) with the final conditions  $y(T) = Y$ ,  $\eta(T) = Q$ . Therefore to “draw”  $\eta$  up to  $Q$  we should take into account the

field  $\mu_0$  in the right-hand side of Eq. (82). Of course it costs much more (in terms of diminishing probability) than the instanton corresponding to the effective action (87). To estimate the corresponding contribution to the effective action one should take into account the term  $\propto \mu_0^2$  from Eq. (28) besides the action (87). The contribution can be estimated in the Gaussian approximation over variations near the instanton solution corresponding to the effective action (87). The answer is as follows

$$-\ln \mathcal{P}(Q, Y) \sim \frac{\epsilon}{D}(Y - Y_{\text{cr}})^2. \quad (89)$$

This expression leads to the conclusion that at  $Y > Y_{\text{cr}}$  the probability decreases fast with increasing  $Y$ . Therefore the region  $Y > Y_{\text{cr}}$  practically does not contribute to the probability of the signal lost.

The possibility of the collapse leads to the following interesting phenomenon. There appears to be a contribution to the probability distribution function  $\mathcal{P}(Q, Y)$  proportional to  $\delta(Q)$ . During the collapse  $\eta^2 \propto t_* - t$  (where  $t_*$  is the time when the collapse finishes). Substituting that into Eq. (81) for  $y$  we conclude that  $y \rightarrow 0$  also at  $t \rightarrow t_*$ . Thus the contribution we discuss is proportional to  $\delta(Y)$  as well. Of course the contribution must be proportional to the “observation time”  $T$  since we consider an escape from the stability region due to a leak through the barrier with a constant rate and no return. The probability to escape the barrier can be found balancing the noise  $\xi_1$  (which indirectly influences  $\eta$  through pumping  $\beta$  and  $y$ ) by other terms in Eq. (82) for  $\eta$ . The probability can be estimated in terms of the effective action (87). The result is as follows:

$$\mathcal{P}(Q, Y) = T\delta(Q)\delta(Y) \exp(-F_{\text{col}}) + \mathcal{P}_{\text{reg}}(Q, Y), \quad F_{\text{col}} \sim \frac{\epsilon^3}{D},$$

where  $\mathcal{P}_{\text{reg}}(Q, Y)$  is the regular contribution discussed above and we assume  $T\epsilon \gg 1$ . For large  $T$  the probability of loosing the signal is simply proportional to  $T \exp(-F_{\text{col}})$ .

The computations in Appendix C show that the probability to have a significant contribution of the continuous spectrum to the observable signal is strongly suppressed and thus negligible, if  $D \ll \epsilon^{1/4}$ . Another criterion follows from the estimation of influence of the continuous spectrum on dynamics of our reduced set of variables  $\beta, y, \eta$ . Repeating procedure which lead us to the formula (62) with the correlation functions from the Appendix C we obtain the estimation for the term modifying the reduced effective action due to interaction with the continuous spectrum:

$$i\Delta\mathcal{I} \sim \frac{D^2}{\sqrt{\epsilon}} \int d\tau \mu_0^2,$$

We can conclude that the continuous spectrum is irrelevant at the condition

$$D\epsilon^{-1/2} \ll 1. \quad (90)$$

## 6. Error probability

In this section we apply the above results to calculate error probability in soliton-based data transmission.

A common way (see [7, Section 5.5]) to make a decision concerning a received bit is to compare the “energy” of the signal in some window

$$E(l, \eta, y) = \int_{-l}^l dx |\Psi|^2 \approx \eta^2 \{ \tanh[\eta(l-y)] + \tanh[\eta(l+y)] \} \quad (91)$$

with some decision threshold  $2s \tanh l$  (usually  $s \lesssim 1$ , for instance,  $s = 1/2$ ). The probability of making an error (in this case, treating unity as zero)  $P_1(s, T)$  is the integral of  $\mathcal{P}(\eta, y)$  over the region defined by the inequality  $E(l, \eta, y) < 2s \tanh l$ .

This integral cannot be calculated analytically, but can easily be computed numerically. Comparing two terms in the PDF and substituting  $l$  for  $y$  one sees that for  $l \ll T$  the answer is the same as for Gaussian PDF as it is determined by  $y$ -dependence of  $\mathcal{P}(\eta, y)$  and does not depend on  $\eta$ -dependence. If, however,  $l$  is comparable with  $T$  (it is not clear if it is practical) then  $\eta$ -dependence is crucial. If we neglect  $y$  in  $\tanh$ , our region is given by  $\eta < s^{1/3}$

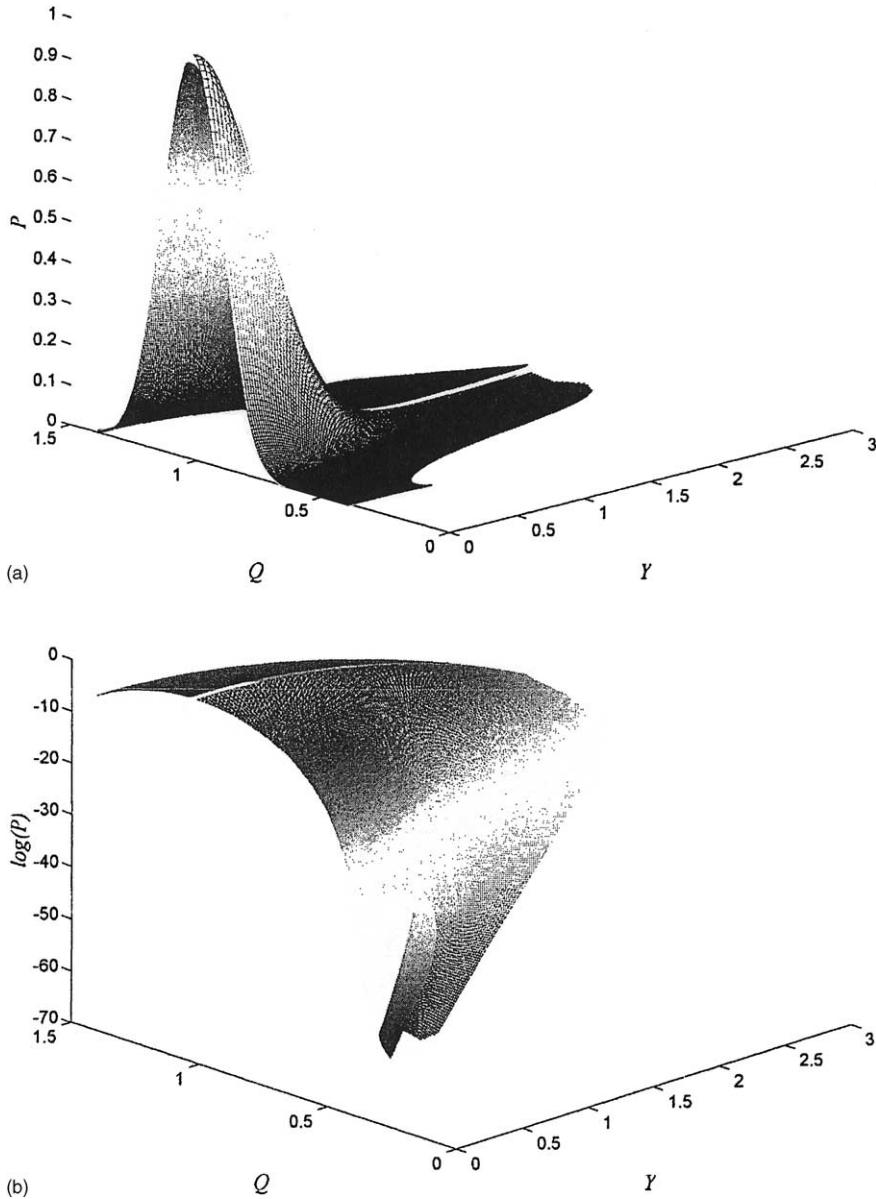


Fig. 1. Probability density function in normal (a) and logarithmic scale (b).

and the error probability is  $\exp[-(1 - s^{2/3})^2/DT]$ —similar to (45) for signal from noise. Using this simple model of the receiver one can easily apply derived PDFs to the evaluation of the corresponding contributions to BER.

## 7. Numerical simulations

In this section we present the results of numerical simulations. We focus on the most non-trivial case of in-line amplitude modulation and filtering. The parameters usually used in the numerical modeling are as follows:  $N = 2$ ;  $\epsilon_2 = 0.1$ ;  $\epsilon_1 = \epsilon_2(N + 1)/3N$ ;  $\epsilon_3 = 4\epsilon_2/\sqrt{\pi}N$ .

Fig. 1 shows two-dimensional PDF  $P(Y, Q)$  calculated for the system with in-line amplitude modulation and filtering. Fig. 1a depicts the PDF in the normal scale and Fig. 1b in the logarithmic scale. It is seen from Fig. 1b that probability does not cover the whole plane  $(Y, Q)$ .

There is a forbidden zone that is also presented in Fig. 2 showing a contour plot of  $P(Y, Q)$ . Existence of such zones where probability to observe soliton with corresponding parameters is zero is attributed to the stabilizing action of the amplitude modulators and in-line filters. Note that formal numerical solutions of the saddle-point equations can be double-valued in some regions in the plane  $(Y, Q)$  as it is illustrated by Figs. 3 and 4. These figures demonstrate that the same point in the plane  $(Y, Q)$  can be obtained with two different trajectories. Fig. 3 illustrates the dynamics of one instanton field  $\eta$  whereas Fig. 4 shows ‘phase portrait’ of the two different trajectories  $\eta$  versus  $y$  corresponding to the same parameters in double-valued zone. Single-valued PDF in such regions should be formed by taking the larger values of  $P(Y, Q)$ .

Fig. 5 describes collapse of the instanton trajectories under specific initial conditions. Figs. 4–6 have been obtained by integration of the reduced model (85) and (86). Recall that the reduced model corresponds to the particular case

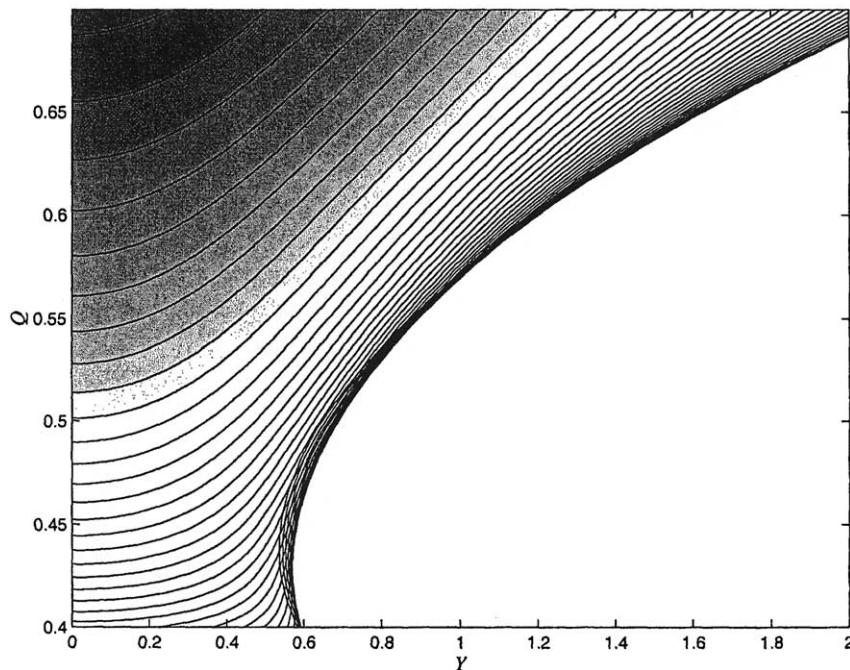


Fig. 2. Contour plot of PDF showing a fold and a forbidden zone.

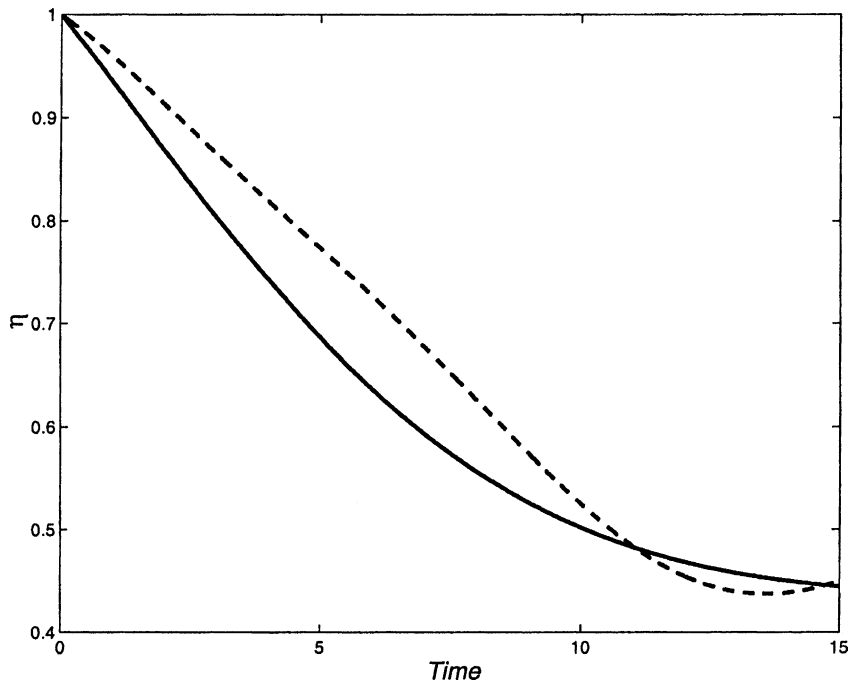


Fig. 3. Dynamics of the instanton field  $\eta$  corresponding to the two different trajectories.

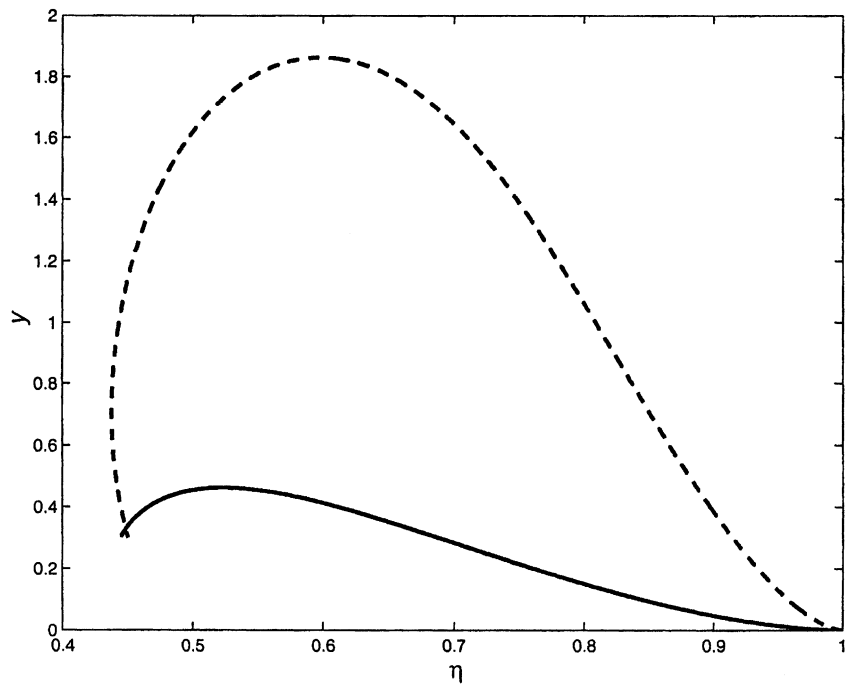


Fig. 4. Phase portrait  $y$  vs.  $\eta$ .

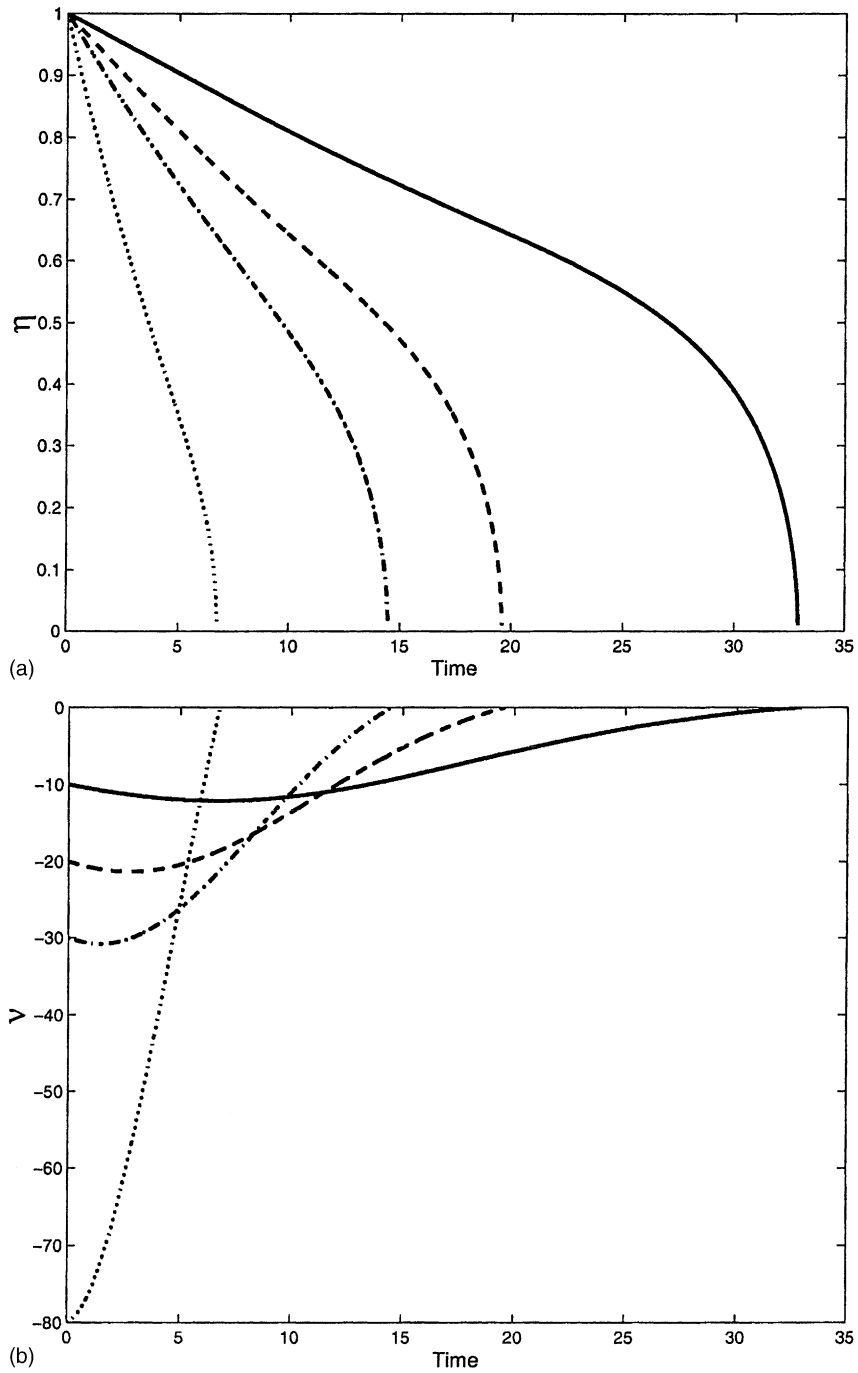
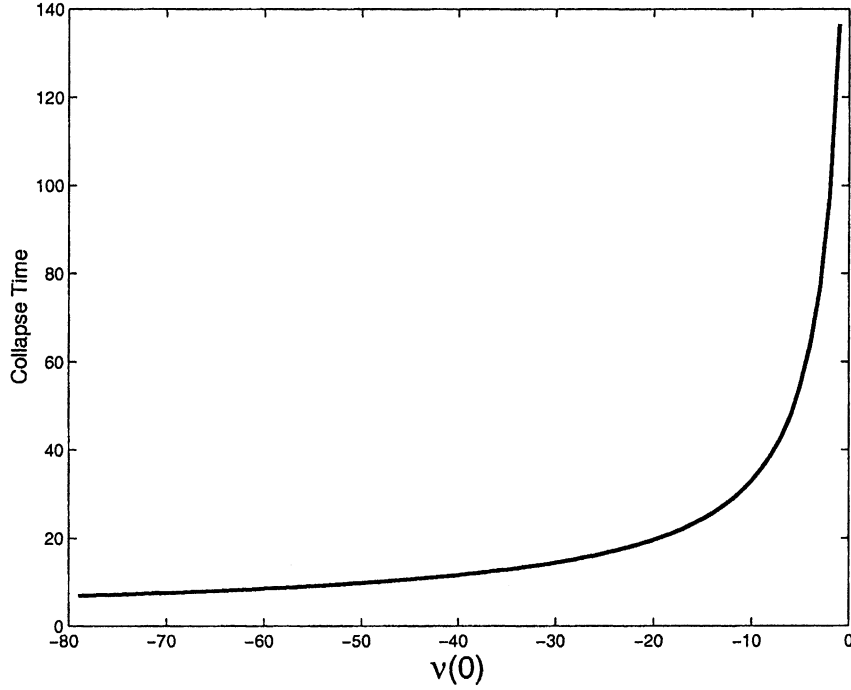


Fig. 5. Collapsing dynamics of the instanton fields  $\eta$  (a) and  $\nu$  (b) corresponding to the reduced model Eqs. (85) and (86). Different curves correspond to different values of  $\nu_0(0) = -10$  (solid);  $\nu_0(0) = -20$  (dashed);  $\nu_0(0) = -30$  (dot-dashed);  $\nu_0(0) = -80$  dotted.

Fig. 6. Collapse time vs.  $v(0)$ .

of instanton fields ( $v_1 = v_2 = \beta = \gamma = 0$ ). Results presented in Figs. 5 and 6 have been computed for the same set of parameters except different  $v_0(t = 0)$ . Fig. 5 depicts evolution of  $\eta$  (a) and  $v_0(t)$  (b) with time. Different curves in these two pictures correspond to different values of  $v_0(0) = -10$  (solid);  $v_0(0) = -20$  (dashed);  $v_0(0) = -30$  (dot-dashed);  $v_0(0) = -80$  (dotted). Collapse has been found for all the negative  $v_0(0) < 0$  and the collapse time has been defined as the point where the both fields become zero. Fig. 6 shows how the collapse time depends on  $v_0(0)$  suggesting that the collapse occurs only for negative  $v_0(0)$  (or positive Hamiltonian  $H$ ).

## Appendix A. Auxiliary relations

Recall some well-known [1] properties of the perturbations near a soliton described by the nonlinear Schrödinger equation. The perturbations can be examined in terms of the linear equation

$$i\partial_t \begin{pmatrix} v \\ v^* \end{pmatrix} + \hat{L} \begin{pmatrix} v \\ v^* \end{pmatrix} = 0, \quad (\text{A.1})$$

where the operator  $\hat{L}$  is

$$\hat{L} = (\partial_x^2 - 1)\hat{\sigma}_3 + \frac{2}{\cosh^2[x]}(2\hat{\sigma}_3 + i\hat{\sigma}_2). \quad (\text{A.2})$$

Evidently

$$\hat{L}^* = \hat{L}, \quad \hat{\sigma}_1 \hat{L} \hat{\sigma}_1 = -\hat{L}, \quad \hat{L}^T = \hat{\sigma}_3 \hat{L} \hat{\sigma}_3. \quad (\text{A.3})$$

The spectrum of the linear problem is determined by the equation  $\hat{L}f = \lambda f$ . A general solution of this equation reads

$$f_k = \exp[ikx] \left\{ 1 - \frac{2ik \exp[-x]}{(k+i)^2 \cosh[x]} \right\} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{\exp[ikx]}{(k+i)^2 \cosh^2[x]} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_k = k^2 + 1. \quad (\text{A.4})$$

Due to the property (A.3) the functions  $\bar{f}_k = \hat{\sigma}_1 f_k^*$  are also eigenfunctions of  $\hat{L}$ :

$$\bar{f}_k = \exp[-ikx] \left\{ 1 + \frac{2ik \exp[-x]}{(k-i)^2 \cosh[x]} \right\} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{\exp[-ikx]}{(k-i)^2 \cosh^2[x]} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_k = -(k^2 + 1). \quad (\text{A.5})$$

Another set of the eigenfunctions  $\hat{L}\varphi = \lambda\varphi$  can be written as

$$\varphi_k = \exp[-ikx] \left\{ 1 - \frac{2ik \exp[x]}{(k+i)^2 \cosh[x]} \right\} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{\exp[-ikx]}{(k+i)^2 \cosh^2[x]} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_k = -(k^2 + 1), \quad (\text{A.6})$$

and  $\bar{\varphi}_k = \hat{\sigma}_1 \varphi_k^*$ :

$$\bar{\varphi}_k = \exp[ikx] \left\{ 1 + \frac{2ik \exp[x]}{(k-i)^2 \cosh[x]} \right\} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{\exp[ikx]}{(k-i)^2 \cosh^2[x]} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_k = k^2 + 1. \quad (\text{A.7})$$

Here  $\varphi_k(x) = \bar{f}_{-k}(-x)$  and  $\bar{\varphi}_k(x) = f_{-k}(-x)$ .

There are also bound states corresponding to the marginally stable modes:

$$f_0 = \frac{1}{\cosh[x]} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \lambda_0 = 0; \quad f_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{\tanh[x]}{\cosh[x]}, \quad \lambda_1 = 0, \quad (\text{A.8})$$

Double poles at  $k = \pm i$  mean that two more functions must be added for closure, namely

$$f_2 = \frac{x}{\cosh[x]} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \hat{L}f_2 = -f_1; \quad (\text{A.9})$$

$$f_3 = \frac{x \tanh[x] - 1}{\cosh[x]} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \hat{L}f_3 = -f_0. \quad (\text{A.10})$$

Note that due to the property (A.3) the left eigenfunctions of the operator  $\hat{L}$  can be written as  $f_k^T \hat{\sigma}_3$ ,  $\bar{f}_k^T \hat{\sigma}_3$ ,  $\varphi_k^T \hat{\sigma}_3$ ,  $\bar{\varphi}_k^T \hat{\sigma}_3$ . That leads to a set of orthogonality conditions for the eigenfunctions. In an explicit form the conditions can be written as

$$\int_{-\infty}^{+\infty} dx \varphi_k^T \hat{\sigma}_3 \bar{f}_q = 2\pi \left( \frac{k-i}{k+i} \right)^2 \delta(k+q), \quad (\text{A.11})$$

$$\int_{-\infty}^{+\infty} dx \bar{\varphi}_k^T \hat{\sigma}_3 f_q = -2\pi \left( \frac{k+i}{k-i} \right)^2 \delta(k+q), \quad (\text{A.12})$$

$$\int_{-\infty}^{+\infty} dx f_2^T \hat{\sigma}_3 f_1 = 2, \quad \int_{-\infty}^{+\infty} dx f_0^T \hat{\sigma}_3 f_3 = -2. \quad (\text{A.13})$$



Let us give the values of some integrals

$$\begin{aligned} \int_{-\infty}^{+\infty} dz z^2 \frac{\sinh^2(z)}{\cosh^4(z)} &= \frac{1}{3} \left( 1 + \frac{\pi^2}{6} \right), & \int_{-\infty}^{+\infty} dz \frac{z^2}{\cosh^2(z)} &= \frac{\pi^2}{6}, \\ \int_{-\infty}^{+\infty} dz \frac{\exp(2\alpha z)}{\cosh(z)} &= \frac{\pi}{\cos(\pi\alpha)}, & \int_{-\infty}^{+\infty} dz \frac{\exp(2\alpha z)}{\cosh^2(z)} &= \frac{2\pi\alpha}{\sin(\pi\alpha)}, \\ \int_{-\infty}^{+\infty} dz \frac{\exp(2\alpha z)}{\cosh^3(z)} &= \frac{2\pi(1/4 - \alpha^2)}{\cos(\pi\alpha)}, & \int_{-\infty}^{+\infty} dz \frac{\exp(2\alpha z)}{\cosh^4(z)} &= \frac{4\pi\alpha(1 - \alpha^2)}{3 \sin(\pi\alpha)}. \end{aligned}$$

## Appendix B. Contact terms

Here, we explain an origin of the “contact terms” in the stochastic equations, discussed in the main body. Let us consider the following equation for an  $N$ -component field  $\phi = (\phi^1, \dots, \phi^N)$ :

$$\partial_t \phi^\alpha = F^\alpha(\phi) + A^{\alpha\mu}(\phi) \xi_\mu, \quad \langle \xi^\alpha(t_1) \xi^\beta(t_2) \rangle = 2D_{\alpha\beta} \delta(t_1 - t_2), \quad \langle \xi^\alpha \rangle = 0, \quad (\text{B.1})$$

where  $F$  is the “deterministic force” and  $\xi$  is the  $M$ -component white noise. Summation over repeated indices is assumed here and below. The matrix  $\hat{D}$  is assumed to have  $\phi$ -independent matrix elements  $D_{\alpha\beta}$ . Generally  $N \neq M$  and we can assume an arbitrary numbers  $N$  and  $M$ . Therefore the situation is generic. The case considered in the main body of the paper corresponds to a substitution of the summation over the components of the noise by an integration over the continuous variable.

If the matrix  $A^{\alpha\mu}(\phi)$  has a non-trivial  $\phi$ -dependence, then the continuous [equation \(B.1\)](#) is not unambiguously defined. It is completely fixed in the discretized form. We assume the following discretization:

$$\phi_{n+1}^a - \phi_n^a = \epsilon F^a(\phi_n) + \frac{1}{2} \epsilon (A_{n+1}^{a\mu} + A_n^{a\mu}) \xi_n^\mu, \quad (\text{B.2})$$

$$\langle \xi_n^\mu \xi_m^\nu \rangle = \frac{2\hat{D}}{\epsilon} \delta_{nm}, \quad \langle \xi_n \rangle = 0, \quad (\text{B.3})$$

where  $\epsilon$  is the value of time steps. Such a regularization has physical justification:  $\delta$ -function in the noise correlation function [\(B.1\)](#) is the limit of a narrow symmetrical distribution. That is the reason why the coefficient at  $\xi_n^\mu$  is taken at the middle of the interval  $t_n, t_{n+1}$  (see also an alternative scheme at the end of this appendix).

To recast [Eq. \(B.3\)](#) in the standard retarded form we expand:

$$\epsilon A^{a\mu} \left( \frac{\phi_{n+1} + \phi_n}{2} \right) \xi_n^\mu = \epsilon A^{a\mu}(\phi_n) \xi_n^\mu + \frac{\epsilon}{2} (\phi_{n+1}^b - \phi_n^b) \frac{\partial}{\partial \phi_n^b} A^{a\mu} \xi_n^\mu \quad (\text{B.4})$$

$$= \epsilon A^{a\mu}(\phi_n) \xi_n^\mu + \frac{\epsilon^2}{2} A^{bv} \frac{\partial}{\partial \phi_n^b} A^{a\mu} \xi_n^\mu \xi_n^\nu + \dots \quad (\text{B.5})$$

Here the equation of motion [\(B.3\)](#) was used in iterations to express the difference  $\phi_{n+1}^b - \phi_n^b$  in terms of  $\phi_n$  and  $\xi$  again. The term  $\sim \epsilon^2 \xi_n^\mu \xi_n^\nu$  in [\(B.5\)](#) should be taken into account because its expectation value is  $\sim \epsilon$ . One can easily verify that the omitted in [\(B.5\)](#) terms give negligible effect in the limit  $\epsilon \rightarrow 0$ . Thus, [Eq. \(B.3\)](#) can be rewritten as

$$\phi_{n+1}^a - \phi_n^a = \epsilon F^a(\phi_n) + \epsilon \Xi_n^a, \quad (\text{B.6})$$

$$\Xi_n^a = A^{a\mu}(\phi_n) \xi_n^\mu + \frac{\epsilon}{2} A^{bv} \frac{\partial}{\partial \phi_n^b} A^{a\mu} \xi_n^\mu \xi_n^\nu. \quad (\text{B.7})$$

After solution of the evolution equation for the variables  $\phi$  the new noises  $\mathcal{E}$  becomes complicated functionals of the initial one  $\xi$ . This makes statistics of  $\mathcal{E}$  non-trivial.

To avoid this difficulty we can formally substitute solving Eq. (B.6) by an integration over all functions  $\phi_n$  with the weight, which is the functional  $\delta$ -function ensuring that only solutions of Eq. (B.6) are taken into account. Then the value of any functional  $O$  of  $\phi$  on the solution of Eq. (B.6) is written as

$$\prod_j d\phi_j \delta(\phi_{n+1}^a - \phi_n^a - \epsilon F^a(\phi_n) - \epsilon \mathcal{E}_n^a) O. \quad (\text{B.8})$$

The retarded discretization of Eq. (B.6) leads to the absence of an additional Jacobian in Eq. (B.8). Then averaging of  $O$  with respect to  $\xi$  can be considered as the functional integration of the expression (B.8) over  $\xi$  with the measure:

$$\prod_j \int d\xi_j \exp(-\xi_j^T \hat{D}^{-1} \xi_j). \quad (\text{B.9})$$

We can change the order of integration over  $\phi$  and  $\xi$ . After that, integrating with respect to  $\xi$ , we can treat the field  $\phi$  as a parameter. Then this Gaussian integration over  $\xi$  can be treated as a Gaussian integration over  $\mathcal{E}$  with the averages

$$\langle \mathcal{E}_n^a \rangle = D_{\mu\nu} A^{bv}(\phi_n) \frac{\partial}{\partial \phi_n^b} A^{a\mu}(\phi_n), \quad (\text{B.10})$$

$$\langle \mathcal{E}_n^a \mathcal{E}_m^b \rangle_c = D_{\mu\nu} A^{a\mu}(\phi_n) A^{bv}(\phi_n) \frac{2}{\epsilon} \delta_{nm}. \quad (\text{B.11})$$

Here the designation  $\langle \dots \rangle_c$  means connected (irreducible) correlation functions. A continuous version of Eqs. (B.10) and (B.11) has the form:

$$\langle \mathcal{E}^a(t) \rangle = D_{\mu\nu} A^{bv} \frac{\partial}{\partial \phi^b} A^{a\mu}, \quad (\text{B.12})$$

$$\langle \mathcal{E}^a(t_1) \mathcal{E}^b(t_2) \rangle_c = 2D_{\mu\nu} A^{a\mu} A^{bv} \delta(t_1 - t_2). \quad (\text{B.13})$$

The non-zero expectation value of  $\mathcal{E}$  in (B.11)–(B.13) is just the “contact term”. It is worth noting that the last term in (B.7) at fixed  $\phi$  produces only non-zero  $\langle \mathcal{E} \rangle$  and does not affect any irreducible correlation function of  $\mathcal{E}$ .

Returning to the expression (B.8), it is convenient to rewrite the  $\delta$ -functions there as

$$\delta(\phi_{n+1}^a - \phi_n^a - \epsilon F^a(\phi_n) - \epsilon \mathcal{E}_n^a) = \int \frac{dp_n}{2\pi} \exp[ip_n(\phi_{n+1}^a - \phi_n^a - \epsilon F^a(\phi_n) - \epsilon \mathcal{E}_n^a)]. \quad (\text{B.14})$$

In the continuous limit the product of the  $\delta$ -functions in Eq. (B.8) is written as

$$\prod_j \delta(\phi_{n+1}^a - \phi_n^a - \epsilon F^a(\phi_n) - \epsilon \mathcal{E}_n^a) \rightarrow \int \mathcal{D}p \exp \left[ i \int dt p_a (\partial_t \phi^a - F^a - \mathcal{E}^a) \right]. \quad (\text{B.15})$$

Performing explicitly the Gaussian integration over  $\mathcal{E}$  we get finally

$$\langle O \rangle = \int \mathcal{D}\phi \mathcal{D}p \exp(i\mathcal{I}) O, \quad (\text{B.16})$$

$$i\mathcal{I} = i \int dt p_a \left( \partial_t \phi^a - F^a(\phi) - D_{\mu\nu} A^{bv} \frac{\partial}{\partial \phi^b} A_{a\mu} \right) - \int dt D_{\mu\nu} A_{a\mu} A^{bv} p_a p_b, \quad (\text{B.17})$$

Here  $\langle O \rangle$  designates the value of the functional  $O$  on the solutions of Eq. (B.1) averaged over the statistics of the noise  $\xi$ .

Let us give an alternative derivation of Eqs. (B.12) and (B.13). To obtain the ‘contact terms’ (B.12) one may consider the noise  $\xi$  in Eq. (B.1) to have a small but finite correlation time:

$$\langle \xi_\alpha(t_1) \xi_\beta(t_2) \rangle = 2\mathcal{D}_{\alpha\beta}(t_1 - t_2), \quad \mathcal{D}_{\alpha\beta} = \int_{-\infty}^{\infty} dt \mathcal{D}_{\alpha\beta}(t). \quad (\text{B.18})$$

We assume that  $\mathcal{D}(t)$  is an even function of  $t$ . Then, solving Eq. (B.1) on a time interval  $\tau$  which is much larger than the noise correlation time but much smaller than the characteristic time of varying  $\phi$ , we get (keeping only relevant terms)

$$\phi(t_0 + \tau) = \phi(t_0) + \int_{t_0}^{t_0 + \tau} dt' (F + \Xi), \quad (\text{B.19})$$

$$\Xi^\alpha(t) = A^{\alpha\mu}(t_0) \xi_\mu(t) + \int_{t_0}^t dt' A^{\beta\nu}(t_0) \xi_\nu(t') \frac{\partial}{\partial \phi^\beta} A^{\alpha\mu}(t_0) \xi_\mu(t). \quad (\text{B.20})$$

The relation (B.19) can be treated as an elementary step in time and then, using Eqs. (B.18) and (B.20) we get the relations (B.12) and (B.13).

### Appendix C. Estimations of the noise amplitude in continuous spectrum

Let us consider the linear equation

$$-i\partial_t \psi = \partial_x^2 \psi + i\xi x^2 \psi + \xi, \quad (\text{C.1})$$

which is a simplified version of the linearized equation of motion of fluctuations belonging to the continuous spectrum. It can be used to estimate their stationary averaged amplitude and, thus, their contribution to the error probability.  $\xi$  is our noise. Let us compute two-point simultaneous correlation function of the field  $\psi$ . This can be done using the formal solution of the evolution Eq. (C.1):

$$\psi(t, x) = \int_0^t d\tau e^{-i\hat{H}(t-\tau)} \xi(\tau, x), \quad (\text{C.2})$$

where the operator  $\hat{H}$  has the form of the Hamiltonian of the harmonic oscillator with the imaginary frequency:

$$\hat{H} = -\partial_x^2 - i\epsilon x^2. \quad (\text{C.3})$$

Averaging with respect to  $\xi$  (see the expression for its correlation function) we get:

$$\langle \psi(t, x_1) \psi^*(t, x_2) \rangle = D \int_0^t d\tau \int dy \mathcal{G}(\tau; x_1, y) \mathcal{G}^*(\tau; x_2, y), \quad (\text{C.4})$$

where  $\mathcal{G}(\tau; x, y)$  is the propagator satisfying

$$i\partial_\tau \mathcal{G}(\tau; x, y) = \hat{H} \mathcal{G}(\tau; x, y), \quad \mathcal{G}(0; x, y) = \delta(x - y). \quad (\text{C.5})$$

It can be taken from the book of Feynmann and Hibbs and has the form:

$$\mathcal{G}(\tau; x, y) = \left( \frac{\Omega}{4i\pi \sin \Omega\tau} \right)^{1/2} \exp \left[ \frac{i\Omega}{4 \sin \Omega\tau} ((x^2 + y^2) \cos \Omega\tau - 2xy) \right], \quad (\text{C.6})$$

where  $\Omega = (1 - i)\sqrt{2\epsilon}$ . To estimate  $\langle |\psi(0)|^2 \rangle$  we need only the asymptotics of  $\mathcal{G}(\tau; x, y)$  at large  $\tau$ :

$$\mathcal{G}(\tau; x, y) \approx \left( \frac{(1 - i)\sqrt{2\epsilon}}{\pi} \right)^{1/2} \exp \left[ -\tau(1 + i) \sqrt{\frac{\epsilon}{2}} - \frac{(1 - i)\sqrt{2\epsilon}}{4} (x^2 + y^2) \right]. \quad (\text{C.7})$$

Taking the final time  $t$  to infinity, we perform all the integrations in the formula (C.4) easily and obtain the desired estimation:

$$\langle |\psi(x)|^2 \rangle \sim \frac{D}{\epsilon^{1/4}} \exp\left(-x^2 \sqrt{\frac{\epsilon}{2}}\right). \quad (\text{C.8})$$

## References

- [1] D.J. Kaup, Perturbation theory for solitons in optical fibers, *Phys. Rev. A* 42 (1990) 5689.
- [2] C.R. Menyuk, *Opt. Lett.* 20 (1995) 285.
- [3] T. Georges, *Opt. Commun.* 123 (1996) 617.
- [4] F.Kh. Abdullaev, S.A. Darmanyan, F. Lederer, *Opt. Commun.* 126 (1996) 89.
- [5] G. Falkovich, I. Kolokolov, V. Lebedev, A. Migdal, *Phys. Rev. E* 54 (1996) 4896.
- [6] G. Falkovich, I. Kolokolov, V. Lebedev, S. Turitsyn, *Phys. Rev. E* 63 (2001) 025601(R).
- [7] E. Iannone, et al., *Nonlinear Optical Communication Neetworks*, Wiley, New York, 1998.
- [8] J.P. Gordon, H.A. Haus, *Opt. Lett.* 11 (1986) 665.
- [9] J. Elgin, *Phys. Lett. A* 110 (1985) 441.
- [10] H.A. Haus, Y. Lai, *JOSA B* 7 (1990) 386.
- [11] J.P. Gordon, *JOSA B* 9 (1992) 91.
- [12] G. Falkovich, M.G. Stepanov, *Opt. Lett.* 27 (1) (2002) 13–15.
- [13] G. Falkovich, M.G. Stepanov, S.K. Turitsyn, *Phys. Rev. E* 64 (6) (2001) 067602.
- [14] M. Chertkov, I. Gabbitov, I. Kolokolov, V. Lebedev, *JETP Lett.* 74 (2001) 357.
- [15] A. Mecozzi, J.D. Moores, H.A. Haus, Y. Lai, *Opt. Lett.* 16 (1991) 1841.
- [16] L.F. Mollenauer, J.P. Gordon, S.G. Evangelides, *Opt. Lett.* 17 (1992) 1575.