Conditional Probability Calculations for the Nonlinear Schrödinger Equation with Additive Noise

I. S. Terekhov,1,2,* S. S. Vergeles,3 and S. K. Turitsyn4,2

1Budker Institute of Nuclear Physics of SB RAS, Novosibirsk 630090, Russia
2Novosibirsk State University, Novosibirsk 630090, Russia
3Landau Institute for Theoretical Physics, Moscow 119334, Russia
4Aston Institute of Photonic Technologies, Aston University, Aston Triangle, Birmingham B4 7ET, United Kingdom

(Received 16 February 2014; published 4 December 2014)

The nonlinear Schrödinger equation (NLSE) is one of the most general and fundamental models of nonlinear science, with a vast number of applications ranging from hydrodynamics, plasma physics, and biophysics to modern high-speed fiber-optic communications (see, e.g., Ref. [1–12] and references therein). The NLSE, in particular, describes in the main order high-frequency wave propagation in media with nonlinear and dispersive effects, making it a very generic mathematical model. The NLSE is also of special interest, because it presents an example of an integrable nonlinear system with an infinite number of degrees of freedom [8]. Here, without loss of generality, we consider the NLSE in the practically important context of optical communications; however, the obtained mathematical results are very general and may be applied in a wide range of physics problems. We present a method for computing a conditional probability density function (PDF) for the NLSE with additive white Gaussian noise. We apply the developed method for the derivation of an analytical expression for the PDF in the practically important limit of weak nonlinearity.

The Letter is organized as follows. First, we show that the NLSE for different noise realizations is done analytically. This is an important and nontrivial step change simplifying the overall numerical modeling of the probability density function for the equation of high practical importance (in the optical communication context) and of wide applicability across many areas of physics. Then, we demonstrate that in the case when the signal-to-noise power ratio is large, the main contribution to the path integral gives the “classical trajectory.” We found the equation for the classical trajectory, and the solution of this equation gives the main contribution to the conditional probability density function. Then, we demonstrate the application of our method to the calculation of the conditional probability density function for a specific example. We would like to stress that the proposed methodology of calculation of the general conditional probability is applied to an arbitrary field at a destination (received signal) and as such, cannot be practically obtained through direct Monte Carlo modeling of the NLSE for different noise realizations.

Consider the nonlinear Schrödinger equation for a field ψω(z) with additive noise ηω(z) presented here in the frequency domain,

\[ \partial_z \psi_\omega(z) = i \frac{\beta_2}{2} \omega^2 \psi_\omega(z) + \eta_\omega(z) \]

\[ + i \gamma \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \bar{\psi}_{\omega_1}(z) \psi_{\omega_2}(z) \bar{\psi}_{\omega}(z). \] (1)

Here and in what follows ω3 = ω1 + ω2 − ω. In the optical-fiber applications context, β2 is the group velocity dispersion parameter, γ is the Kerr nonlinearity coefficient, the bar means complex conjugation, ηω(z) is an additive complex white noise (resulting in fiber communication applications from optical amplification) with zero mean and correlation function [7,11]:

\[ \langle \eta_\omega(z) \eta_{\omega'}(z') \rangle = 2\pi Q \delta(z − z') \delta(\omega − \omega'). \]

Using the Martin-Siggia-Rose formalism [13–15], we can formally present the conditional probability density \( P[Y(\omega)|X(\omega)] \) to have \( \psi_\omega(L) = Y(\omega) \) if \( \psi_\omega(0) = X(\omega) \) in the form of the Feynman path integral, corresponding to the Eq. (1),

\[ P[Y(\omega)|X(\omega)] = \int_{\psi_\omega(0) = X(\omega)} \psi_\omega(L) = Y(\omega) \] \( \mathcal{D}\psi \exp \left\{ - \frac{S[\psi]}{Q} \right\}, \) (2)

where the action \( S[\psi] \) reads
\[ S[\psi] = \int_0^L dz \int \frac{d\omega}{2\pi} \mathcal{L}^{(0)}[\psi] - V[\psi], \]

\[ \mathcal{L}^{(0)}[\psi] = \partial_z \psi(z) - i \frac{\beta_2}{2} \omega^2 \psi(z), \]

\[ V[\psi] = i\gamma \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \psi(z_1)\psi(z_2)\psi(z), \]

with the measure \( D\psi \) defined as

\[ D\psi = \lim_{\Delta \to 0} \lim_{\Lambda \to 0} \left( \frac{\delta}{\Delta \pi Q} \right)^{-NM} \prod_{j=1}^N \prod_{i=1}^M d\psi_{i,j}; \]

Here \( \tilde{\Lambda} = (\Delta \pi Q/\delta)^{-NM} \), \( \psi_{i,j} = \psi_{i,j}(z_i) \), \( \delta \psi_{i,j} = \psi_{i,j} - \psi_{i-1,j}, \)

\[ z_i = \Delta i, \quad i = 0, 1, \ldots, N, \quad z_N = L, \quad \omega_j = \Omega_{\text{min}} + 2\pi(j - 1)\delta, \quad j = 1, 2, \ldots, M, \quad \omega_M = \Omega_{\text{max}}, \]

In Eq. (7), we take into account the boundary conditions \( \psi_{0,j} = X(\omega_j) = X_j, \quad \psi_{N,j} = \psi_{N,j}(\omega) = Y_j. \)

The conditional probability satisfies the standard condition \( \int D\psi \left[ \frac{\mathcal{L}^{(0)}[\psi]}{\mathcal{L}^{(0)}[\psi]} \right] = 1 \), where \( D\psi = \prod_{j=1}^M d\psi \). Equation (7) is the first important result of our work. The integral over all \( M \) can be calculated numerically with the required accuracy using Monte Carlo methods, see, e.g., Ref. [19]. Therefore, Eq. (7) provides a constructive way to compute the PDF for the NLSE in most general cases.

Moreover, the presentation (7) allows us to develop the perturbation theory using a small nonlinearity parameter and derive an analytical expression for the conditional probability. To do so, let us introduce two dimensionless parameters, \( \tilde{\gamma} = \gamma P_{\text{ave}} L \) and \( \tilde{\epsilon} = Q L W / (2 \pi P_{\text{ave}}) = 1/\text{SNR} \), where \( P_{\text{ave}} = T_{\text{total}}^{-1} \int (d\omega / 2\pi) |X(\omega)|^2 \) is the average power of the signal, \( T_{\text{total}} \) is the full time interval of a signal pattern, and \( W / (2\pi) \) is noise bandwidth [we imply that signal bandwidth is equal or less than \( W / (2\pi) \)].

The dimensionless parameter \( \tilde{\epsilon} \) is nothing more than the inverse power signal-to-noise-ratio (SNR). The dimensionless parameter \( \tilde{\gamma} \) describes the effective nonlinearity. Later, we impose that \( \tilde{\gamma} \ll 1 \) and develop the perturbation theory in the parameter \( \tilde{\gamma} \) for different values of \( \tilde{\epsilon} \). In the case \( \tilde{\gamma} / \epsilon \ll 1 \) we can expand the exponential function in Eq. (2). When the parameter \( \tilde{\gamma} \ll 1 \) and \( \tilde{\gamma} / \epsilon \sim 1 \), or even \( \tilde{\gamma} / \epsilon \gg 1 \), we use a method similar to that developed in quantum mechanics for finding the classical trajectory of the particle.

Let us start the consideration from the case \( \tilde{\gamma} / \epsilon \ll 1 \). Using standard methods of quantum field theory, see Refs. [13,20], we expand the exponent in Eq. (2) at small \( \gamma \). After that the function \( P[Y(\omega)|X(\omega)] \) can be represented as a series in \( \gamma \).

\[ P[Y(\omega)|X(\omega)] = \sum_{n=0}^{\infty} \frac{\gamma^n}{n!} P_{(n)}^{(\gamma)}(Y(\omega) | X(\omega)). \]

In zero order in \( \gamma \) we obtain an effective Gaussian channel approximation (see Ref. [15]).

\[ P_{(0)}^{(\gamma)}(Y(\omega) | X(\omega)) = \Lambda \exp \left\{ - \frac{1}{QL} \int \frac{d\omega}{2\pi} |B(\omega)|^2 \right\}, \]

where \( \Lambda \) is the normalization constant, \( \Lambda = P_{(0)}^{(\gamma)}[0|0] = (\pi Q L / \delta)^{-M/2} \), and \( B(\omega) = Y(\omega) e^{-i \beta_2 \omega^2 L / 2} - X(\omega) \). The function \( B(\omega) \) is proportional to the difference of the \( Y(\omega) \) and the solution \( \psi_{\omega}(L) \) of Eq. (1) with \( \gamma = 0 \), \( \eta = 0 \), and the boundary condition \( \psi_{\omega}(0) = X(\omega) \); therefore, \( P_{(0)}^{(\gamma)}(Y(\omega) | X(\omega)) \) is the Gaussian distribution of functions around \( \psi_{\omega}(L) \) in functional space. It is easy to see that \( P_{(0)}^{(\gamma)}(Y(\omega) | X(\omega)) \) is normalized as \( \int D\psi P_{(0)}^{(\gamma)}(Y(\omega) | X(\omega)) = 1 \). This means that all corrections in \( \gamma \) are normalized here to satisfy the condition

\[ \int D\psi P_{(\gamma)}^{(n\neq0)}(Y(\omega) | X(\omega)) = 0. \]

The higher-order corrections in \( \gamma \) can be calculated in any order from Eq. (7) (see the Supplemental Material [14] for details). As an example, we write down here the first
order correction $P_{(r)}^{(1)}[Y(\omega)|X(\omega)]$. Using the procedure described in the Supplemental Material [14] we obtain

$$P_{(r)}^{(1)}[Y(\omega)|X(\omega)] = P_{(r)}^{(0)}[Y(\omega)|X(\omega)] \text{Im} \left\{ \frac{WL}{2\pi} \int \frac{d\omega}{2\pi} e^{-i\beta_2\omega^2L/2} Y(\omega) \tilde{X}(\omega) + G \right\},$$

(10)

$$G = \frac{2}{Q} \frac{L}{L} \int_0^L \frac{dz}{L} \frac{d\omega_1 d\omega_2}{(2\pi)^2} e^{i\omega \tilde{\mu}_1(z) - i\omega \tilde{\mu}_2(z)},$$

$$\tilde{\mu}_1(z) = X(\omega) + \frac{\omega}{2} B(\omega)/L.$$  

(11)

Here $\mu = \beta_2(\omega - \omega_1)(\omega - \omega_2)L$. The result contains two different terms: the first one is proportional to the bandwidth $W$ and does not involve the parameter $Q$, the other one (function $G$) has $Q$ in the denominator. Therefore, in this limit (small $\gamma$) of the perturbation theory the noise parameter $Q$ is assumed to be not too small. In physical terms this is the limit of a weakly nonlinear and highly noisy system.

In the different and practically important limit of small $\epsilon$ or $Q$ (large SNR) the conditional probability can be computed using a method similar to the one used to calculate the classical trajectory in quantum mechanics [21]. In simple terms, this corresponds to finding the solution without the noise term and making a functional expansion around this solution due to the small noise (high signal-to-noise ratio). In the case under consideration we can use Laplace’s method, see, e.g., Ref. [22]. The main contribution to the path integral in Eq. (2) gives the region around the trajectory where the action $S[\psi]$ reaches the minimum. Let $S$ approach the minimum at the trajectory $\psi_0(z)$; Eq. (2) can be rewritten in the following form:

$$P[Y(\omega)|X(\omega)] = e^{-S[\psi_0(z)]/Q} \int \frac{\psi_0(L) = 0}{\psi_0(0) = 0} D\psi e^{-\delta S[\psi_0(z)]} \psi_0(z)/Q.$$

(12)

The explicit form of the path integral is shown in Ref. [15]. Thus, the problem of calculation of the conditional probability reduces to finding the function $\psi_0(z)$ and calculation of the path integral with zero boundary conditions. We would like to emphasize once more the important difference of the proposed approach and the direct Monte Carlo modeling of the NLSE with different realizations of noise. In the path-integral method we can constructively compute the probability density function for an arbitrary received signal $Y(\omega)$, even for a PDF with very rare events, while in the direct modeling of the NLSE it might be practically impossible to find trajectories with low probability that still can be important for system performance. Now we calculate the conditional probability in leading orders in $Q$.

To calculate the path integral we expand the expression in the exponent in the path integral at small $\psi$ to the series in $\psi$. Since $S$ reaches the minimum at $\psi_0(z)$, the series starts from second order in $\psi$. To calculate the path integral in leading order in $Q$ we keep terms only in the main order in $\psi$ in the series. Then, we calculate the integral using the perturbation theory in $\gamma$ developed in Ref. [15], and obtain the result in leading and next-to-leading order in $\gamma$,

$$P[Y(\omega)|X(\omega)] \approx \Lambda e^{-\delta S[\psi_0(z)]/Q} \left( \frac{2\gamma W}{\pi} \text{Im} \left\{ \int_0^L \frac{dz}{L} z(L - z) \int \frac{d\omega}{2\pi} \mathcal{L}^{(0)}[\psi_0(z)] \psi_0(z) \right\} \right).$$

(13)

It is seen that in order to calculate $P[Y(\omega)|X(\omega)]$ we first have to determine the function $\psi_0(z)$ (classical trajectory). The action approaches the minimum at $\psi_0(z)$; therefore, $\delta S[\psi] = 0$, where $\delta S$ is a variation of $S$. This last equation leads to the following equation for $\psi_0(z)$ (analogue of a classical trajectory):

$$\left( \frac{\partial}{\partial z} - i\beta_2\omega^2 \right)^2 \psi_0(z) - i\gamma \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \left\{ 4\psi_{\omega_1}(z) \psi_{\omega_2}(z) \right\} \psi_0(z) = -\frac{\mu}{L} \psi_0(z) \psi_0(z) \psi_0(z)$$

$$- 3\gamma^2 \int \frac{d\omega_1 d\omega_2 d\omega_3 d\omega_4}{(2\pi)^4} \delta(\omega_1 + \omega_2 + \omega_3 - \omega_4 - \omega_0) \psi_{\omega_1}(z) \psi_{\omega_2}(z) \psi_{\omega_3}(z) \psi_{\omega_4}(z) \psi_0(z) = 0,$$

(14)

with the boundary conditions $\psi_0(0) = X(\omega)$, $\psi_0(L) = Y(\omega)$. Equation (14) can be written in the time domain, see Ref. [15].

Calculating the action $S[\psi_0(z)]$ up to first order in $\gamma$ yields (see for details the Supplemental Material [15])

$$P[Y(\omega)|X(\omega)] \approx P^{(0)}_{(r)}[Y(\omega)|X(\omega)] e^{\text{Im}(G)} \left( 1 + \frac{\gamma WL}{3\pi} \text{Im} \left\{ \int \frac{d\omega}{2\pi} e^{-i\beta_2\omega^2L/2} Y(\omega) \tilde{X}(\omega) \right\} \right).$$

(15)
Note that the exponent \( e^{\text{Im}(G)} \) cannot be expanded at small \( \gamma \) in the general case, since we have assumed here that the parameter \( \varepsilon \ll 1 \). However, when \( \tilde{\gamma}/\varepsilon \ll 1 \) the result (15) coincides with \( P^{(0)}[Y(\omega)|X(\omega)] + \gamma P^{(1)}[Y(\omega)|X(\omega)] \), as it should. Equation (15) is the NLSE channel PDF in the limit of small \( \tilde{Q} \) in leading order in \( \gamma \).

Now, we illustrate the application of the derived general PDF (valid for arbitrary input \( X \)), considering some particular choices of initial signal. Let the signal in the time domain have the form

\[
X(t) = \sum_{k=-N}^{N} c_k F(t-kT), \quad F(t) = \alpha e^{-t^2/2\tau^2},
\]

where \( N \gg 1 \) is the number of pulses in the information pattern, \( c_k = e^{i\phi_k} \), where \( \phi_k \) is a value that is randomly chosen from \( \{0, i\pi/2, i\pi, -i\pi/2\} \), \( F(t) \) is the waveform of the carrier pulse, \( T \) is the time interval between pulses (baud rate), and \( \tau \) is a parameter related to the pulse width; we assume here that \( \tau \ll T \). The constant \( \alpha \) defines the signal average power \( P_{\text{ave}} = T^{-1} \int_{-\infty}^{\infty} F^2(t) dt = \alpha^2 \tau \sqrt{\pi}/T \). In the frequency domain the initial signal is presented as

\[
X(\omega) = \sqrt{2\pi\alpha} e^{-\omega^2\tau^2/2} \sum_{k=-N}^{N} c_k e^{i\omega k T}.
\]

Consider PDF distributions of \( c_k \) assuming that the received signal \( Y(\omega) \) can be approximated as

\[
Y(\omega) = \left\{ X(\omega) + \sqrt{2\pi\alpha} e^{-\omega^2\tau^2/2} \sum_{k=-N}^{N} \rho_k e^{i\phi_k} e^{i\omega k T} + i\gamma L \phi_\text{nl} (X(\omega)) \right\} e^{i\beta\omega L^2 T/2},
\]

where the average nonlinear phase shift (rotation of the phase, same to all pulses) is

\[
\phi_\text{nl} (X(\omega)) = \int \frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi)^3} X(\omega_1) X(\omega_2) \tilde{X}(\omega_3) \frac{1-e^{-\mu}}{\mu}.
\]

This choice of \( Y(\omega) \) implies that all coefficients \( c_k \) are changed to \( \tilde{c}_k = c_k + \rho_k e^{i\phi_k} \), which corresponds to corruption of the signal constellation points by noise and (weak) nonlinear effects. For the sake of clarity in this methodological Letter we imply that pulses do not broaden large \((|\beta L|/\tau \ll T) \). Then, we can use property (46) of Ref. [15] for the conditional probability. The substitution of Eqs. (18) and (17) into Eq. (15) yields the following conditional probability:

\[
P_k = \Lambda^{1/(2N+1)} \exp \left\{ -\frac{P_{\text{ave}} T}{QL} \rho_k^2 \right\} \times \left( 1 + \frac{\gamma WTLP_{\text{ave}}}{3\pi} \sum_{k=-N}^{N} \rho_k \sin(\tilde{\phi}_k - \phi_k) \right).
\]

One can see that the conditional probability of the whole signal pattern is the product of conditional probabilities for each pulse as it should be for noninterfering signals. Of course, the general PDFs derived above do include pulse-to-pulse interference that can be accounted for perturbatively. Since \( \tilde{\gamma}/\varepsilon \ll 1 \), \( P_k \) is the slightly deformed Gaussian distribution. Of course, the result (20) is formally written with excessive accuracy and should be used only in the first order of the parameter \( \tilde{\gamma}/\varepsilon \ll 1 \),

\[
P[Y(\omega)|X(\omega)] \approx \Lambda \exp \left\{ -\frac{P_{\text{ave}} T}{QL} \sum_{k=-N}^{N} \rho_k^2 \right\} \times \left( 1 + \frac{\gamma WTLP_{\text{ave}}}{3\pi} \sum_{k=-N}^{N} \rho_k \sin(\tilde{\phi}_k - \phi_k) \right).
\]

Note that we already took into account the overall phase shift \( \phi_\text{nl} (X(\omega)) \) in \( Y(\omega) \), see Eq. (18). We would like to stress that Eq. (20), of course, is just a particular example of using the general formulas for the NLSE PDF derived above. In the general case, one can use either the PDF [Eq. (7)] for numerical analysis with an arbitrary input signal or expressions (10) and (15) for simplified numerical or analytical analysis in practically important limits.

In conclusion, we have introduced a constructive method for numerical computation of the conditional probability for the nonlinear Schrödinger equation through multidimensional integrals. We have developed an analytical method for the conditional probability calculation for nonlinear noisy fiber optic communication channels in the case of weak nonlinearity and the arbitrary parameter \( \varepsilon = 1/\text{SNR} \), which is the inverse signal-to-noise power ratio. In the limit \( \varepsilon \sim 1 \) we derived an equation for calculating the conditional probability using the perturbation theory in \( \gamma \). In the limit \( \varepsilon \ll 1 \) we have derived the classical trajectory and developed a method similar to finding the quantum corrections to the classical trajectory in quantum mechanics. The path-integral method allows one to constructively compute PDFs for any received signal \( Y(\omega) \), even corresponding to very rare events, while in the direct modeling of the NLSE it might be practically impossible to find trajectories with such low probability. We believe that our results might find various applications ranging from statistical physics [23–25] to high capacity optical communications [11,26]. The approach provides a platform for optimization over initial signal distributions that is of critical importance for computation of the Shannon capacity of communication channels.
We acknowledge the financial support of the Engineering and Physical Sciences Research Council (Project UNLOC) and a grant from the Ministry of Education and Science of the Russian Federation (Agreement No. 14.B25.31.0003).

I.S.Terekhov@inp.nsk.su


