Schrödinger equation with a spatially and temporally random potential: Effects of cross-phase modulation in optical communication

A. G. Green,[1] P. B. Littlewood,[2] P. P. Mitra,[3] and L. G. L. Wegener[2]
[1]*Theoretical Physics, Oxford University, Oxford, United Kingdom*
[2]*TCM, Cavendish Laboratory, Cambridge University, Cambridge, United Kingdom*
[3]*Lucent Technologies, Bell Laboratories Innovations, Murray Hill, New Jersey 07974*
(Received 23 April 2002; revised manuscript received 2 August 2002; published 30 October 2002)

We model the effects of cross-phase modulation in frequency (or wavelength) division multiplexed optical communications systems, using a Schrödinger equation with a spatially and temporally random potential. Green's functions for the propagation of light in this system are calculated using Feynman path-integral and diagrammatic techniques. This propagation leads to a non-Gaussian joint distribution of the input and output optical fields. We use these results to determine the amplitude and timing jitter of a signal pulse and to estimate the system capacity in analog communication.

DOI: 10.1103/PhysRevE.66.046627 PACS number(s): 42.81.Dp, 42.65.Tg, 05.40.Ca

I. INTRODUCTION

A. General

The ability to transmit information is ultimately limited by signal distortion. Information theory [1,2] quantifies the extent to which such distortions inhibit communication; it was originally developed for the study of radio communication or electrical communication along copper wires. In these cases the signal propagation is linear— the received signal is a linear function of the transmitted signal. Distortion arises due to the addition of extraneous signal fluctuations, which propagate linearly alongside the original signal. These fluctuations may come from noisy amplifiers or other circuit elements, or from cross-talk with other messages. Such linear systems with additive noise are very well characterized from an information theoretical perspective.

In modern communication systems, the situation is rather more complicated. The very high transmission rates— particularly in optical communication— require that the transmission medium be operated in regimes where the signal propagation is substantially nonlinear [3]. These nonlinearities can lead to a variety of new mechanisms of signal distortion [4–9] and are often very difficult to characterize analytically. This has limited the understanding of nonlinear channels, particularly from an information theoretical perspective.

Faced with these difficulties, there are two ways in which progress may be made. One approach is to perform detailed numerical simulations of the underlying partial differential equation. This has been the approach of much of the work in the literature. However, it is very difficult to use the results of such analyses to understand the information theoretical aspects of communication in non-linear media. An alternative approach is to harness physical understanding of the signal propagation to motivate simple phenomenological models that capture some aspects of the nonlinear propagation, but which are nevertheless analytically tractable. This approach has been used to understand a number of aspects of nonlinear propagation in optical fibers [4–9]. Only recently have such approaches been applied to study the information theoretical aspects of the problem [10–12].

When one is interested in optimizing the design of a communications system, it is often useful to characterize the various mechanisms of signal distortion according to their physical origin in the system. This is not always the best way of classifying signal distortion from an information theoretical point of view. A more convenient classification is to divide signal distortions into additive and multiplicative noise. The signal distortions in optical fibers may be classified in this manner [12]. It is the multiplicative noise that presents the main difficulty in understanding nonlinear channels. It lies behind some of the classic unsolved problems in information theory [2]. It is to this aspect of the problem that path-integral and diagrammatic techniques are particularly suited. In the system studied in this paper, multiplicative noise appears as a random potential in a Schrödinger equation.

The tension between additive and multiplicative noise is similar to the tension between disorder and interactions in many-body physics. In the latter case, the consequence of this tension is that a full analytical characterization is not possible. Instead, one must resort to approximation schemes that are appropriate under different regimes of system parameters. It is likely that for the same reasons one will be forced to use similar approximate schemes to study the propagation of signals in nonlinear media. Physicists have worked hard to develop such schemes and some of them are directly applicable to the study of nonlinear signal propagation; for example, Falkovich *et al.* [13] have used the theory of optimal fluctuation to understand the fluctuations in soliton amplitude and timing that occur in optical communication. In this paper, we use Feynman path-integral [14] and diagrammatic techniques [15] in order to understand some of the effects of nonlinearity upon optical communication. Most of this paper will concentrate upon characterizing the propagation of light in the fiber. Towards the end of the paper, we will harness this characterization to understand some information theoretical aspects of the problem.

B. The wavelength division multiplexed optical fiber

In this paper, we will consider some of the effects of nonlinearity upon communication using frequency division

multiplexed optical fibers [3]. Optical fibers have an enormous transparent bandwidth. It is far beyond current electronics to modulate at these frequencies. In order to utilize as much of the bandwidth as possible, it is divided into a number of nonoverlapping frequency bands, or subbands, each of which is modulated independently. This is called frequency division multiplexing or alternatively wavelength division multiplexing (WDM).

Light propagating in a WDM optical fiber experiences several sources of signal distortion and nonlinearity. The first of these is due to dispersion; the different frequency components of a signal travel at different velocities along the fiber, leading to a spreading of a signal pulse. This spreading is a linear operation, however, and does not reduce the capacity to communicate information. The effects of dispersion *could* be taken into account by a suitable linear transformation of the received signal during processing. Usually, however, dispersion is accounted for by introducing dispersion compensating elements into the transmission link; spans of standard, dispersive fiber are interleaved with spans of dispersion compensating fiber in which the sign of dispersion has been reversed, so that more slowly moving frequency components catch up with the faster moving components.

As light propagates in an optical fiber, some of it is scattered out of the fiber by Raman scattering leading to a loss of signal power. This is compensated for by the periodic insertion of short spans of lasing fiber. As the light passes through the lasing fiber, it is amplified by stimulated emission. It is impossible to avoid a certain amount of spontaneous emission in this process. This is amplified alongside the signal and provides a source of additive noise called amplified spontaneous emission (ASE) noise.

Although we take full account of the effects of dispersion, dispersion compensation and ASE noise in our analysis, these are not the main focus of this work. A system with only these effects is a linear system with additive noise. This is precisely the type of system understood by Shannon many years ago [1]. The additional aspects of optical communication are the nonlinearities present in the propagation of light in an optical fiber. The most important of these nonlinearities is the optical Kerr nonlinearity, by which the refractive index of the fiber depends upon the intensity of light in the fiber [3].

The Kerr nonlinearity causes scattering between different frequency components of light in the fiber. Two incoming frequency components scatter into two outgoing frequency components. In WDM systems the effects of the Kerr nonlinearity are grouped according to the origin of the frequency components which scatter off one another. When the incoming and outgoing frequency components all lie in the same subband, the effect is known as self-phase modulation (SPM). It is this nonlinearity that is responsible for the possibility of soliton propagation within a subband. The combination of SPM and ASE noise leads to several mechanisms of signal distortion: fluctuations in the amplitude and arrival times of soliton pulses, known as Gordon-Haus jitter [4]; nonlinear amplification of the phase noise [5]; and, depending upon the sign of dispersion, a nonlinear instability of amplitude fluctuations about a constant signal called the modulation instability [9]. When the incoming frequency components are in different subbands and scatter back into their original subbands, the effect is called cross-phase modulation (XPM). This is the effect that we will concentrate upon here. We will show explicitly that this leads to multiplicative noise— to a spatially and temporally random potential due to the signal in the other subbands. We will find a superdiffusive spreading of the pulse shape and a superdiffusive spreading of arrival times. Finally, the effect of the Kerr nonlinearity, when all of the incoming and scattered frequency components lie in different subbands is known as four-wave mixing (FWM). From the point of view of the signal in any particular subband, it has an identical effect to additive amplifier noise [16].

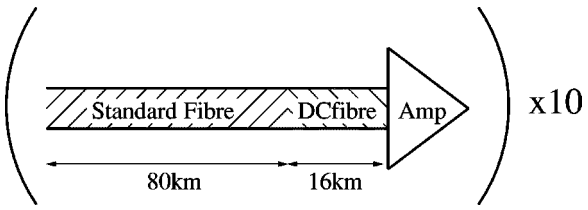

FIG. 1. Schematic diagram of an optical fiber system.

A complete WDM optical communications system consists of a series of spans of standard, dispersive fiber interleaved with dispersion compensating fibers and loss compensating amplifiers. The precise ordering of these elements is known as the dispersion map. The performance of the system can depend quite sensitively upon this dispersion map. However, for concreteness, we will restrict ourselves to one particular map, shown in Fig 1.

Our approach will be to calculate Feynman path integrals describing the propagation of light in a single subband, taking account of scattering from signals in other subbands through XPM. Green's functions for this propagation must be averaged over realizations of the signal in other subbands. Green's functions for propagation in the optical fiber turn out to be Green's functions of a Schrödinger equation with a spatially and temporally random potential. The main calculations in this paper are the evaluation of the single- and two-particle Green's functions of this Schrödinger equation. This calculation is similar to calculations carried out in two other contexts: Turbulent flow or the growth of interfaces may both be described by the Kardar, Parisi, and Zhang equation [17–19]. This may be mapped to a diffusion equation with a spatially and temporally random potential. Calculating the quadratic statistics for these processes and averaging over the turbulent flow or noise in the growth process is very similar to calculating the two-particle Green's function in our problem. In particular, short-range temporal correlations allow the calculation to be reduced to an effective single-body problem for which the Feynman path integral may be readily calculated. We will find the same structure in our solution to the present problem. The second problem to which our work bears similarity, is the calculation of Green's function in the presence of a quenched random potential [15]. In this context, it is usual to truncate the diagrammatic series for the single- and two-particle Green's functions to the Born and Ladder series, respectively. Here we will find

that this truncation is justified by the short-range temporal correlations of the random potential.

Our approach contrasts with other methods of including the effects of XPM in WDM systems [6], in which individual scattering events between soliton pulses in different subbands are considered. Our results are very similar to those obtained for Gordon-Haus jitter [4], although of a different physical origin.

Before going on to a detailed calculation, let us first state our results and the physics underlying them. As a signal pulse propagates along a particular subband it passes through regions of higher and lower intensity of the total signal in the other subbands. This causes the pulse to speed up and slow down due to the Kerr nonlinearity. Since the signals in the other subbands are unknown to the users of any particular subband, this speeding up and slowing down of the pulse has a statistical uncertainty. The result is a diffusive spreading of the phase velocity and a corresponding superdiffusive spreading of the distribution of pulse arrival time; $\langle \delta T^2 \rangle \propto L^3$ (to see this requires integrating the diffusively spreading phase velocity over the length of propagation). This is the main result of our calculation. The functional form is very similar to that obtained previously for Gordon-Haus jitter [4], although its physical origin is slightly different; it is due to XPM rather than a combination of SPM and ASE noise. In addition to the uncertainty in the pulse arrival time, the pulse shape itself is distorted due to the effect of XPM; since they travel at different speeds, the different frequency components of the pulse sample slightly different portions of the signal in the other channels and arrive at slightly different times. Therefore, the pulse width itself will show a superdiffusive spreading due to the effects of XPM. The effect of XPM upon the pulse power is much weaker than its effect upon pulse timing. The processes leading to fluctuations in pulse power are of higher order in the (weak) nonlinearity. The timing errors induced by XPM cause the pulse power to spread over a larger interval leading to the possibility of overlap between adjacent pulses. This intersymbol interference allows timing errors to be converted into amplitude errors. Numerical simulation, however, shows that the dominant contribution to amplitude errors comes from SPM [16], the effects of which are not studied here.

II. THE MODEL

The propagation of light in an optical fiber is described by Maxwell's equations with a dielectric constant dependent upon the energy density of the electromagnetic field. The optical fibers used in communications systems support only one transverse mode. Expanding the Maxwell's equation for the low frequency envelope of the electric field and taking the long wavelength limit, one obtains a nonlinear Schrödinger equation with the roles of space and time interchanged [3]:

$$i\partial_z E(z,t) = (\beta \partial_t^2 - i\alpha) E(z,t) - \gamma |E(z,t)|^2 E(z,t). \quad (1)$$

The term proportional to α on the right-hand side of this equation describes the loss of optical power from the channel and the term proportional to γ encodes the effects of the Kerr nonlinearity. Dividing the electric field into its components in each subband of the WDM system, the low frequency envelope of the electric field in the i^{th} subband, $E_i(z,t)$, is given by

$$i\partial_z E_i(z,t) = (\beta \partial_t^2 - i\alpha) E_i(z,t) + V(z,t) E_i(z,t),$$

$$V(z,t) = -2\gamma \sum_{j \neq i} |E_j(z,t)|^2. \quad (2)$$

In writing down this expression we have neglected self-phase modulation, $\gamma |E_i(z,t)|^2 E_i(z,t)$, and four-wave mixing, $\gamma \Sigma_{j,k \neq i} E_j^*(z,t) E_k(z,t) E_i(z,t)$. This neglect of self-phase modulation and four-wave mixing may be justified on two grounds. First, there are regimes where numerical simulation confirms this to be a good approximation [16]. Studying the system in the absence of the effects of self-phase modulation and four-wave mixing allows us to deduce the effects of XPM. Our aim is to determine what constraints are imposed upon communication by these effects. Other sources of signal distortion will impose additional constraints in a real system, but these are not of direct concern in this paper. Equation (2) corresponds to a nonlinear Schrödinger equation with a spatially and temporally random potential due to nonlinear interaction with the signals in the other subbands. The effects of loss in this system may be accounted for by using rescaled electrical fields that maintain their amplitude during propagation; $E_{eff,i}(z,t) = e^{\alpha z} E_i(z,t)$. The propagation of these rescaled fields is described by Eq. (2) with an exponential rescaling of the strength of nonlinearity, $\gamma_{eff} = \gamma(1 - e^{-\alpha L})/\alpha L$, where L is the length of a single span of fiber, and without the loss term proportional to α. Henceforth, we will assume the use of these rescaled fields and drop the subscript "eff" for brevity.

We follow Ref. [10] and model the stochastic potential by a Gaussian distribution with short-ranged correlations in space and time; the signals in the separate subbands are statistically independent and have short-ranged temporal correlations on time scales of order 1/bandwidth. Relative dispersion of signals in different subbands leads to short-ranged spatial correlations, which may be approximated by a delta function if the dispersion is sufficiently large. The strength of the potential, $\eta = \int dz \langle \delta V(z,t) \delta V(0,t) \rangle = [(\gamma P)^2/(2\beta \Delta \delta W)] \ln[n_c/2]$, is obtained by summing the contributions from n_c channels with separation δW, bandwidth Δ, and signal power P. The logarithmic dependence upon the number of channels is due to the suppression of the effects of widely separated channels by dispersion [10]. In the following section, we will often write $\langle \delta V(z,t) \delta V(0,0) \rangle = 2\pi\eta/\Delta\, \delta(t)\delta(z)$, with an explicit δ function in time. This simplifies our calculations and elucidates which contributions are most important for propagation in a potential with short temporal correlations. It is important to realize, however, that the potential and signal are both band limited, so that the correlations are not strictly δ functional. We account for this by regularizing the δ function of zero argument to $\delta(t=0) = \Delta/2\pi$.

The system considered here consists of $10\times$ [80 km standard fiber followed by 16 km dispersion compensating (dc)

fiber and a loss compensating amplifier]. The input signal has power, bandwidth and channel separation $P=5$ mW, $\Delta=10$ GHz, and $\delta W=15$ GHz, respectively. The standard (or dc) fiber loss, nonlinearity, and dispersion parameters are $\alpha=0.048(0.115)$ km^{-1}, $\gamma=1.2(5.1)$ W^{-1} km^{-1}, and $\beta=11.0(-66.9)$ ps^2 km^{-1}. Comparing $\gamma P/\alpha$ for the standard and dc fibers gives a measure of their relative strengths of nonlinearity. The power entering the dc fiber is a factor of ~ 50 lower than that entering the standard fiber. This, combined with the greater loss in the dc fiber, leads to an effective strength of nonlinearity ~ 150 times greater in the standard fiber. We ignore nonlinearity in the dc fiber in our analytical work. It may be included in precisely the same way as in the standard fiber and does not lead to any qualitative change in our results. Three important lengthscales may be determined from these system parameters; the total fiber length $L_{tot}=10\times 80$ km, the nonlinear length $\eta^{-1}=223$km, and the dispersion length $L_D=(\beta\Delta^2)^{-1}=747$ km. The response of the system is completely determined by the ratios of these three length scales.

III. INPUT OUTPUT STATISTICS

The output electric field $E_{out}(t)$, after propagating along a single span of standard fiber, may be expressed in terms of the input electric field $E_{in}(t)$ as

$$E_{out}(t)=\int dt'\mathcal{G}(t,t')E_{in}(t')+n(t), \tag{3}$$

where $n(t)$ is additive Gaussian white noise, with variance N, representing ASE noise. $\mathcal{G}(t,t')$ is the Green's function of Eq. (2) for propagation from $z=0$ at time t to $z=L$ at time t', with a fixed realization of $V(z,t)$. More generally, we will use the notation $\mathcal{G}(z_1,z_2;t_1,t_2)$ to indicate the Green's function for propagation from $t=t_1$, $z=z_1$ to $t=t_2$, $z=z_2$. Using this notation, $\mathcal{G}(t_1,t_2)\equiv\mathcal{G}(0,L;t_1,t_2)$. Suppression of one of the spatial indices implies propagation from $z=0$; $\mathcal{G}(z;t_1,t_2)\equiv\mathcal{G}(0,z;t_1,t_2)$. Similarly, $\mathcal{G}(z,t)\equiv\mathcal{G}(0,z;0,t)$. It is particularly important to note that $\mathcal{G}(t_1,t_2)\neq\mathcal{G}(t_1-t_2)$, as would be the case for a static potential.

Dispersion compensation is included by replacing $\mathcal{G}(t,t')$ in Eq. (3) with a dispersion compensated Green's function $\mathcal{G}_d(t,t')$. This is obtained by convolving the Green's function for the standard fiber, $\mathcal{G}(t,t')$, with the Green's function for the dispersion compensating fiber, $\mathcal{G}_{comp}(t,t')$; $\mathcal{G}_d(t,t')=\int dt''\mathcal{G}_{comp}(t,t'')\mathcal{G}(t'',t')$. $\mathcal{G}_{comp}(t,t')$, has the opposite sign of β to $\mathcal{G}(t,t')$ and describes propagation with no potential, V.

Propagation along a series of interleaved standard and dc spans is modeled by convolving a string of dispersion compensated Green's functions. This convolution is independent of the ordering of the Green's functions and the effects of XPM are, therefore, independent of the dispersion map. Real systems are sensitively dependent upon the choice of the map due to the effects of SPM, which is ignored here. Finally, due to the unitarity of $\mathcal{G}_d(t,t')$, adding ASE noise at each amplifier is equivalent to adding Gaussian white noise at the receiver. Analytical calculations may, therefore, be performed using a simplified dispersion map consisting of N_s standard spans followed by N_s dispersion compensating spans and adding Gaussian white noise at the receiver.

We now use this simple model of the system to calculate the effects of XPM upon the input and output signals. We shall refer to this as the input-output statistics for brevity from now on. The quadratic statistics are given by

$$\begin{aligned}\langle E^*_{in}(t)E_{in}(t')\rangle&=P\,\delta(t-t'),\\ \langle E^*_{in}(t)E_{out}(t')\rangle&=P\langle\mathcal{G}_d(t,t')\rangle,\\ \langle E^*_{out}(t)E_{out}(t')\rangle&=(P+N)\,\delta(t-t'),\end{aligned} \tag{4}$$

and the quartic statistics are given by

$$\begin{aligned}\langle\delta|E_{in}(t)|^2\delta|E_{in}(t')|^2\rangle&=\Delta P^2\delta(t-t')/2\pi,\\ \langle\delta|E_{in}(t)|^2\delta|E_{out}(t')|^2\rangle&=P^2\mathcal{G}^{II}_d(t-t'),\\ \langle\delta|E_{out}(t)|^2\delta|E_{out}(t')|^2\rangle&=\Delta(P+N)^2\delta(t-t')/2\pi.\end{aligned} \tag{5}$$

The angular brackets indicate averages over the signal in the subband of interest, the signals in the other subbands, and the amplifier noise. $\langle\mathcal{G}_d(t,t')\rangle$ is the average of the dispersion compensated, single-particle Green's function over the potential (or realizations of the signal in the other subbands) and $\mathcal{G}^{\mathrm{II}}_d(t-t')=\langle|\mathcal{G}_d(t,t')|^2\rangle$ is the average of the two-particle Green's function over the random potential. The input signal is assumed to have a Gaussian distribution with power P and the amplifier noise to have a Gaussian distribution with power N. We have used the unitarity of the single-point Green's function for a particular realization of the potential in deriving Eqs.(4) and (5); $\mathcal{G}(t,t)=1$. The delta function $\delta(t-t'=0)$ has been regularized as discussed in Sec. II to allow for the fact that the subband has a bandwidth Δ; $\delta(t-t'=0)=\Delta/2\pi$. The averaged Green's functions $\langle\mathcal{G}_d(t,t')\rangle$ and $\mathcal{G}^{\mathrm{II}}_d(t-t')$ are invariant under temporal translations and are therefore, functions of $t-t'$.

The statistics embodied by Eqs. (4) and (5) are non-Gaussian, since $\mathcal{G}^{\mathrm{II}}_d(t-t')\neq|\langle|\mathcal{G}_d(t,t')\rangle|^2$ in general. The corrections to Gaussian statistics are given by the vertex corrections to the two-particle Green's function. Alternatively, one may say that linear propagation in the presence of a random potential leads to nonlinear propagation on average. Calculation of the quadratic and quartic statistics in terms of Green's functions allows a simple characterization of this nonlinearity. In the following section, we will give an explicit calculation of the averaged Green's functions and follow this in the subsequent sections by a discussion of their consequences for optical communication. The quadratic statistics will be useful for our discussion of coherent communication and the quartic statistics will be useful for our discussion of incoherent communication.

IV. CALCULATIONS

The average Green's functions $\langle\mathcal{G}\rangle$ and $\mathcal{G}^{\mathrm{II}}$ may be calculated from their Feynman path-integral representations [14]. If we assume that the stochastic potential is δ function correlated in time, these path integrals may be evaluated in a closed form. These results correspond to resumming the Born and ladder series [15] in the diagrammatic expansions of $\langle\mathcal{G}\rangle$ and $\mathcal{G}^{\mathrm{II}}$, respectively. These series may be used to

approximate the effects of a non-δ-function correlated potential. The diagrams that are neglected in this approximation are small for a potential with short-ranged temporal correlations; the use of a δ-function correlated potential is a calculational tool that emphasizes the range of validity of this approximation. In the following, we will calculate equations of motion for $\langle \mathcal{G} \rangle$ and $\mathcal{G}^{\mathrm{II}}$ from their path integrals, under the assumption that V is δ function correlated in time. We will show how the solutions of these equations correspond to the resummed perturbative, ladder and Born series. Finally, the average Green's functions $\langle \mathcal{G} \rangle$ and $\mathcal{G}^{\mathrm{II}}$ will be calculated allowing for the band limit of the subband and potential due to the other subbands.

A. Single-particle Green's function

Recalling that the roles of space and time have been interchanged here compared with their more familiar roles in quantum mechanics, the single-particle Green's function is written in the form

$$\begin{aligned}\langle \mathcal{G}(L,t) \rangle &= \left\langle \int_{t(0)=0}^{t(L)=t} Dt(z) \exp\left[-\int_0^L dz \frac{i}{2\beta} (\partial_z t(z))^2 \right.\right. \\ &\quad \left.\left. -i\int_0^L dz V(z,t(z)) \right] \right\rangle \\ &= \int_{t(0)=0}^{t(L)=t} Dt(z) \exp\left[-\int_0^L dz \frac{i}{2\beta} [\partial_z t(z)]^2 \right. \\ &\quad \left. -\frac{1}{2}\int_0^L \int_0^L dz dz' \langle V((z,t(z)) V(z',t(z')) \rangle \right],\end{aligned} \tag{6}$$

where the angular brackets denote averages over realizations of V. We have carried out the average over V in order to obtain the second line, assuming that V is a Gaussian random variable.

Taking the derivative of Eq. (6) with respect to L [14], we obtain the following equation of motion for $\langle \mathcal{G} \rangle$:

$$\begin{aligned}&\left(-i\frac{\beta}{2}\partial_t^2 - \partial_L \right) \langle \mathcal{G}(L,t) \rangle \\ &\quad = \delta(L)\delta(t) + \int_0^L \int_{-\infty}^{\infty} dt' dz \langle V(L,t) V(z,t') \rangle \\ &\qquad \times \langle \mathcal{G}(L-z,t-t') \rangle \langle \mathcal{G}(z,t') \rangle .\end{aligned} \tag{7}$$

This equation is local in time for a stochastic potential that is δ function correlated in time; $\langle V(z,t)V(0,0) \rangle = (2\pi\eta/\Delta)\delta(t)\delta(z)$. The solution of this local equation of motion is $\langle \mathcal{G}(L,t) \rangle = e^{-\eta L/2} \mathcal{G}_0(L,t)$, where $\mathcal{G}_0(L,t)$ is the Green's function of the operator $(-i\beta\partial_t^2/2 - \partial_L)$, i.e., the Green's function for propagation in the absence of V. This result was first obtained in Ref. [10]. Equation (7) may also be obtained diagrammatically by resumming the Born series; it is nothing but the Dyson's equation for $\langle \mathcal{G}(L,t) \rangle$ obtained by resumming this series [15]. The temporal δ-function correlation of V implies that the propagators of V must not cross, thus restricting the contributing diagrams to the Born series. Figure 2 shows a symbolic representation of Eq. (7) after multiplying on the left by $\mathcal{G}_0$. The dashed line indicates propagators of V, the solid line indicates $\mathcal{G}_0$, and the thick solid line indicates $\langle \mathcal{G} \rangle$. This diagram is to be interpreted in real space or momentum space in the usual way. The Green's function in frequency and momentum space is

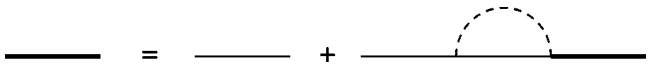

FIG. 2. Born series for $\langle \mathcal{G} \rangle$.

$$\langle \mathcal{G}(p,\omega) \rangle = i(p - \beta\omega^2 + i\eta/2)^{-1}, \tag{8}$$

where p is the wave vector conjugate to z and ω is the angular frequency conjugate to t. For later calculations, it is useful to have $\langle \mathcal{G} \rangle$ in a mixed, position and frequency notation, where it is given by

$$\langle \mathcal{G}(z,\omega) \rangle = \exp[-(\eta/2 - i\beta\omega^2)z]. \tag{9}$$

Dispersion compensation is achieved by temporally convolving $\mathcal{G}(L,t)$ for the standard fiber with that of the compensating fiber (or alternatively multiplying together the position and frequency space Green's functions). Ignoring nonlinearity in the compensating fiber, we find $\mathcal{G}_{comp}(L,t) = \mathcal{G}_0^*(L,t)$. The dispersion compensated Green's function is then given by

$$\langle \mathcal{G}_d \rangle (L,t) = e^{-\eta L/2} \delta(t). \tag{10}$$

On average, therefore, the effects of XPM nonlinearity separate from those of dispersion. Moreover, the effect of XPM is to dephase the propagating electric field.

B. Calculation of $\mathcal{G}^{\mathrm{II}}$

The calculation of $\mathcal{G}^{\mathrm{II}}$ from its path integral proceeds in a very similar manner to the calculation of $\langle \mathcal{G} \rangle$. After writing down the path-integral representation of $\mathcal{G}^{\mathrm{II}}$, we perform the average over V and deduce an equation of motion for $\mathcal{G}^{\mathrm{II}}$ by differentiating with respect to L. This differential equation is local in time when the correlation function of V is a temporal δ function. The equation of motion is most easily solved in the frequency domain, where it is recognizable as the resummed ladder series. As before, the set of contributing diagrams is restricted due to the requirement that the propagators of V may not cross. We will first of all derive an expression for $\mathcal{G}^{\mathrm{II}}(\omega,q)$ in the absence of dispersion compensation from its path-integral representation. We will verify that this result is identical to that obtained by summing the ladder series of diagrams. In fact, closed expressions for $\mathcal{G}^{\mathrm{II}}(L,t)$ are most readily obtained by resumming this series directly in the length and frequency domain. We explicitly carry out this resummation to obtain expressions for $\mathcal{G}^{\mathrm{II}}(L,\omega)$ and its dispersion compensated counterpart $\mathcal{G}_d^{\mathrm{II}}(L,\omega)$.

We will first derive a path integral and equation of motion for the Green's function, $\mathcal{G}^{\mathrm{II}}(L;\tau_1,\tau_2) = \langle \mathcal{G}(L,\tau_1)\mathcal{G}^*(L,\tau_2) \rangle$. In the end we will be interested in $\mathcal{G}^{\mathrm{II}}(L,t) = \mathcal{G}^{\mathrm{II}}(L;\tau_1=t,\tau_2=t)$. The path-integral representation of $\mathcal{G}^{\mathrm{II}}$ is

$$\mathcal{G}^{II}(L;\tau_1,\tau_2)=\langle \mathcal{G}(L,\tau_1)\mathcal{G}^*(L,\tau_2)\rangle$$
$$=\left\langle \int_{\tau_1(0)=0,\tau_2(0)=0}^{\tau_1(L)=\tau_1,\tau_2(L)=\tau_2} D\tau_1(z)D\tau_2(z)\exp\left[\int dz\left(i\frac{(\partial_z\tau_1)^2}{2\beta}-i\frac{(\partial_z\tau_2)^2}{2\beta}\right)\right]\right.$$
$$\left.\times\exp\left[-\frac{1}{2}\int dz[-iV(z,\tau_1(z))+iV(z,\tau_2(z))]\right]\right\rangle$$
$$=\int_{\tau_1(0)=0,\tau_2(0)=0}^{\tau_1(L)=\tau_1,\tau_2(L)=\tau_2} D\tau_1(z)D\tau_2(z)\exp\left[\int dz\left(i\frac{(\partial_z\tau_1)^2}{2\beta}-i\frac{(\partial_z\tau_2)^2}{2\beta}\right)\right]$$
$$\times\exp\left[-\frac{1}{2}\int dzdz'\langle V(z,\tau_1(z))V(z',\tau_1(z'))\rangle\right]$$
$$\times\exp\left[-\frac{1}{2}\int dzdz'\langle V(z,\tau_2(z))V(z',\tau_2(z'))\rangle\right]$$
$$\times\exp\left[\int dzdz'\langle V(z,\tau_1(z))V(z',\tau_2(z'))\rangle\right]. \tag{11}$$

The average over V has been carried out in passing from the penultimate to final line of Eq. (11). The equation of motion for $\mathcal{G}^{\mathrm{II}}$ is derived by differentiating both sides of this equation with respect to L. The resulting differential equation is local in time when V has δ-function correlations in time. It is slightly easier to make the assumption of temporally δ-function correlations for V in Eq. (11) before differentiating. The path integral then reduces to

$$\mathcal{G}^{\mathrm{II}}(L;\tau_1,\tau_2)=\int_{\tau_1(0)=0,\tau_2(0)=0}^{\tau_1(L)=\tau_1,\tau_2(L)=\tau_2} D\tau_1(z)D\tau_2(z)$$
$$\times\exp\left[\int dz\left(i\frac{(\partial_z\tau_1)^2}{2\beta}-i\frac{(\partial_z\tau_2)^2}{2\beta}+\eta\right.\right.$$
$$\left.\left.-\eta\frac{2\pi}{\Delta}\delta(\tau_1(z)-\tau_2(z))\right)\right]. \tag{12}$$

This result is very similar to that obtained in the analysis of the diffusion equation with a spatially and temporally random potential [17]. In that case, the path integral for the square of the wave function reduced to that of two interacting random walks. Here the path integral for $\mathcal{G}^{\mathrm{II}}$ reduces to that of two interacting Schrödinger particles, or waves, propagating in opposite directions. Differentiating Eq. (12) with respect to L, we obtain the following local equation of motion for $\mathcal{G}^{\mathrm{II}}$:

$$\left(\partial_L-i\frac{\beta}{2}(\partial_{\tau_1}^2-\partial_{\tau_2}^2)+\eta\right)\mathcal{G}^{\mathrm{II}}(L;\tau_1,\tau_2)$$
$$=\delta(L)\delta(\tau_1)\delta(\tau_2)+\eta\frac{2\pi}{\Delta}\delta(\tau_1-\tau_2)\mathcal{G}^{\mathrm{II}}(L;\tau_1,\tau_2). \tag{13}$$

This equation is most easily solved by first changing coordinates to the mean time, $t=(\tau_1+\tau_2)/2$, and time difference, $T=(\tau_1-\tau_2)/2$. After Fourier transforming in space and time we find

$$-i\left(q-\frac{\beta}{2}\omega\Omega+i\eta\delta(0)\right)\mathcal{G}^{\mathrm{II}}(q;\omega,\Omega)$$
$$=1+\frac{2\pi\eta}{\Delta}\int\frac{d\Omega}{2\pi}\mathcal{G}^{\mathrm{II}}(q;\omega,\Omega), \tag{14}$$

where ω and Ω are the angular frequencies conjugate to t and T, respectively and q is the wave-vector conjugate to L. Equation (14) is equivalent to the ladder series shown in Fig. 2. In order to make this correspondence clear, the effects of nonlinearity have been divided into two parts; the term proportional to η on the left-hand side of Eq. (14) takes account of self-energy correction to the single-particle Green's function and the term proportional to η on the right-hand side takes account of vertex corrections. Following the usual convention, we denote the solution of Eq. (14) in the absence of vertex corrections by $\Pi(q;\omega,\Omega)=-i(q-\beta\omega\Omega/2+i\eta)^{-1}$. One may check by direct substitution of $\langle\mathcal{G}(q,\omega)\rangle$ from Eq. (8) that $\Pi(q;\omega,\Omega)=\int dp/2\pi\langle\mathcal{G}(q+p,\omega+\Omega)\rangle\langle\mathcal{G}^*(p,\Omega)\rangle$. Finally, using the notation $\Pi(q,\omega)=\int d\Omega/2\pi\Pi(q;\omega,\Omega)$, the solution of Eq. (14) is

$$\mathcal{G}^{\mathrm{II}}(q,\omega)=\frac{\Pi(q,\omega)}{1-\frac{2\pi\eta}{\Delta}\Pi(q,\omega)}$$
$$=\Pi(q,\omega)\left(1+\frac{2\pi\eta}{\Delta}\Pi(q,\omega)\right.$$
$$\left.+\left(\frac{2\pi\eta}{\Delta}\right)^2\Pi(q,\omega)^2+\cdots\right), \tag{15}$$

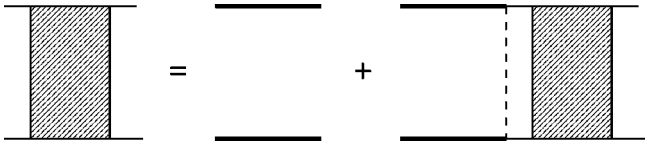

FIG. 3. Ladder series for $\mathcal{G}^{\mathrm{II}}$.

which corresponds to the ladder diagram shown in Fig. 3.

The modification to these results due to dispersion compensation is not quite as simple as the modification to $\langle \mathcal{G} \rangle$. Although it is possible to derive a path-integral expression for $\mathcal{G}_d^{\mathrm{II}}$ akin to Eq. (13), it is easiest to derive the dispersion compensated $\mathcal{G}^{\mathrm{II}}$ using diagrammatic perturbation theory in the length and frequency domain. The appropriate diagrammatic series is shown in Fig. 4. The result of summing this series is

$$\mathcal{G}_d^{\mathrm{II}}(L,\omega) = \Pi_d(L,\omega) + \frac{2\pi\eta}{\Delta}\int_0^L dx \Pi(x,\omega)\Pi_d(L-x,\omega) + \left(\frac{2\pi\eta}{\Delta}\right)^2 \int_0^L dx \int_0^x dy \Pi(x,\omega)\Pi(y,\omega)\Pi_d(L-y,\omega) + \ldots . \tag{16}$$

$\Pi(z,\omega)$ is given by

$$\Pi(z,\omega) = \int \frac{d\omega'}{2\pi} \langle \mathcal{G}(z,\omega'+\omega/2)\rangle \langle \mathcal{G}^*(z,\omega'-\omega/2)\rangle = e^{-\eta z}\int \frac{d\omega'}{2\pi} e^{i\beta\omega'\omega z} \tag{17}$$

and $\Pi_d(z,\omega)$ is the dispersion compensated analogue of $\Pi(z,\omega)$, given by

$$\Pi_d(z,\omega) = \int \frac{d\omega'}{2\pi} \langle \mathcal{G}(z,\omega'+\omega/2)\rangle \mathcal{G}_0^*(L,\omega'+\omega/2) \times \langle \mathcal{G}^*(z,\omega'-\omega/2)\rangle \mathcal{G}_0(L,\omega'-\omega/2) = e^{-\eta z}\int \frac{d\omega'}{2\pi} e^{i\beta\omega'\omega(z-L)}. \tag{18}$$

The calculation of $\mathcal{G}^{\mathrm{II}}$ proceeds by substituting $\Pi_d(z,\omega)$ and $\Pi(z,\omega)$ into Eq. (16), carrying out the frequency integrals in each term, followed by the integrals over length, and finally summing the contributions from each term. In carrying out this procedure, one must take care over the support for the various frequency integrals. The band limitation of the subband implies that the angular frequency carried by each internal Green's function must lie between $-\pi\Delta$ and $+\pi\Delta$. Carrying out the frequency integrals with this restriction leads to very complicated expressions. An alternative approach, which yields much more manageable expressions, is to take into account the fact that the stochastic potential is not strictly δ function correlated in time, but is band limited to lie between $-\pi\Delta$ and $+\pi\Delta$. We use the fact that V has short-range correlations in time to justify neglecting terms other than those of the Born and ladder series in our diagrammatic calculation, but allow for the band limitation of V in calculations of internal frequency integrals.

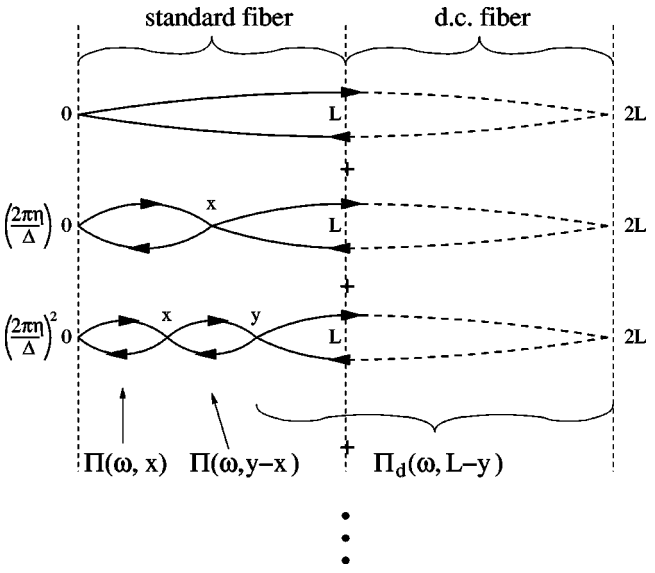

FIG. 4. Position-space diagrammatic series for $\mathcal{G}_d^{\mathrm{II}}(L,\omega)$.

Let us illustrate this by explicitly calculating the second-order term in Eq. (16). This is given by

$$\int_0^L dx \int_0^x dy \Pi(x,\omega)\Pi(y,\omega)\Pi_d(L-y,\omega) = \int_0^L dx \int_0^x dy \int \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\frac{d\omega_3}{2\pi} \times \exp[i\beta\omega(\omega_3-\omega_2)x + i\beta\omega(\omega_2-\omega_1)y] = \left(\frac{\Delta}{2\pi}\right)^3 e^{-\eta L}\int_0^{\pi\beta\omega\Delta L} dX \int_0^X dY \frac{\sin X}{X}\frac{\sin Y}{Y} = \left(\frac{\Delta}{2\pi}\right)^3 L e^{-\eta L}\frac{1}{2}\left(\frac{\mathrm{Si}[X]}{X}\right)\Bigg|_{X=\pi\beta\omega\Delta L}, \tag{19}$$

where $Si[X] = \int_0^X du \sin u/u$. We have carried out the frequency integrals in passing between the second, third, and fourth lines. The integration variables have been changed from ω_1, ω_2, and ω_3 to ω_1, $\omega_3-\omega_2$, and $\omega_2-\omega_1$. Integrations have been carried out over the domains $-\Delta/2 \leqslant \omega_1, \omega_3-\omega_2, \omega_2-\omega_1 \leqslant \Delta/2$, the first corresponding to the bandlimitation of the signal in the subband and the second and third to bandlimitation of the stochastic potential leading to a restriction on the transfer of frequency at each vertex.

Calculating each term in the series in this way and resumming, we obtain our final result

$$\mathcal{G}_d^{\mathrm{II}}(L,\omega) = \frac{\Delta}{2\pi}\exp\left[-\eta L\left(1-\frac{\mathrm{Si}[X]}{X}\right)\right]\Bigg|_{X=\pi\beta\omega\Delta L} \tag{20}$$

for dispersion compensated propagation. The Green's function for uncompensated propagation is calculated similarly replacing Π_d with Π in Eq. (16). The result is

$$\mathcal{G}^{\mathrm{II}}(L,\omega)=\mathcal{G}_d^{\mathrm{II}}(L,\omega)\frac{\sin X}{X}\bigg|_{X=\pi\beta\omega\Delta L}. \tag{21}$$

The prefactor of $\sin X/X$ describes the dispersive spreading of a pulse propagating in the absence of nonlinearity. The distortion of the pulse due to XPM is described by $\mathcal{G}_d^{\mathrm{II}}(L,\omega)$. There are two important limits of Eq. (20). At small frequencies, $\mathrm{Si}[X]/X\to 1-X^2/18$. As a consequence $\mathcal{G}_d^{\mathrm{II}}(L,\omega)$ tends to a Gaussian of width $\sqrt{\langle\omega^2\rangle}=2\pi\Delta_{eff}=(3\Delta/\pi)\sqrt{L_D^2/\eta L_{tot}^3}$. The interpretation of this width as an effective bandwidth will be discussed below. At large frequencies, $\mathcal{G}_d^{\mathrm{II}}(L,\omega)$ tends to a constant. These limits imply that $\mathcal{G}_d^{\mathrm{II}}(L,t)$ has Gaussian tails at large times and approaches a δ function at short times. We will discuss the consequences of these results for communication in the following sections.

V. AMPLITUDE AND TIMING JITTER

The results of Secs. III and IV give the input and output statistics of the WDM optical fiber in terms of its single- and two-particle Green's functions. In this and the following section, we shall use these results to discuss the effectct of XPM upon communication.

First, we discus two operational measures of the efficiency of an optical communications system. Our discussion will be quite brief, since these considerations are not central to the main message of this paper. Nevertheless, it is important to demonstrate that analytical results such as Eqs. (20) and (21) can be used to enhance understanding from this operational perspective. A common way of characterizing signal distortion in optical communications is by amplitude and timing jitter. These measures of system performance are specific to digital communication. A digital message is made up of a string of pulses or marks representing 1's, and spaces representing zeros. As a pulse propagates along a noisy channel, it speeds up and slows down in a random way leading to a distribution of arrival times (defined as the centroid in time of the power distribution at the output) at the receiver known as timing jitter. Similarly, the total power of the pulse fluctuates leading to a distribution of the amplitudes of the received pulse known as amplitude jitter. The distributions of the amplitude and timing fluctuations need not be independent [13]. In fact, the terms amplitude and timing jitter are often reserved for the variance in the amplitude and arrival times of a pulse. This is the sense in which we shall uses these terms from now on. In addition to jitter, the shape of the individual pulses may be distorted during propagation. We do not discus these distortions here.

The effects of SPM and ASE upon the amplitude and timing jitter of soliton, or solitonlike, pulses has been considered in a number of works, starting with Ref. [4]. All of these works expand in small fluctuations— induced by the additive ASE noise— about a soliton solution. Such analyses lead to the conclusion that amplitude jitter is not very significant and that timing jitter leads to an L^3 dependence of the variance of arrival times. This latter result is known as Gordon-Haus jitter after its discovery in Ref. [4]. A recent work of Falkovich *et al.* [13] has used a saddle point approximation within a Martin-Siggia-Rose field theory for this propagation to obtain a full joint distribution of amplitude and timing errors. This work shows in a particularly transparent way, the connection between the method of optimal fluctuations in field theory and large deviations in statistics. Here we use Green's function techniques to consider the contribution of XPM to jitter. Despite its different physical origin, the resulting jitter has a broadly similar dependence upon system parameters to Gordon-Haus jitter.

In principle, the output for an arbitrary input pulse shape may be calculated, accounting for the effects of XPM, in terms of Green's functions for propagation along the fiber. The amplitude and timing jitter may then be obtained by taking appropriate averages. This calculation may in practice prove rather cumbersome. However, the functional dependence of the jitter upon system parameters is expected to be broadly independent of the input pulse shape (the magnitude of the jitter will show a weak dependence upon the pulse shape). For simplicity, therefore, we present approximate expressions for the amplitude and timing jitter due to XPM, by considering the output for constant and delta-function inputs, respectively. The simple expressions obtained in this way demonstrate the functional dependence of the jitter upon the system parameters. In particular, we find that the variance of arrival times is proportional to L^3, whereas the amplitude jitter is largely independent of XPM. These functional dependences are the same as those due to Gordon-Haus jitter although the physical origin is somewhat different [4].

Timing jitter may be found by considering the response to an input δ -function in time. A real signal pulse is band limited, but it may be thought of as being made up of many δ functions. The use of a δ function input allows us to calculate the functional form of the timing jitter in an analytically tractable way. The dependence of the resulting jitter upon system parameters is expected to be broadly similar for arbitrary input pulse shapes. Without XPM, the output electric field resulting from a δ-function input is $E_{out}(t)=\mathcal{G}_d(t)\propto\delta(t)$ and the modulus of the electric field is given by $|E_{out}(t)|^2=|\mathcal{G}_d(t)|^2\propto\delta(t)$. Introducing XPM broadens this response. Different frequencies making up the pulse see a slightly different stochastic potential and arrive at slightly different times. Since the signals in the neighboring subbands are unknown, this broadening leads to uncertainty in the arrival time. Averaging the response to a δ-function input over realizations of the signals in the other subbands, therefore, we find that $\mathcal{G}_d^{\mathrm{II}}(t)=\langle|\mathcal{G}_d(t)|^2\rangle$ gives a measure of the distribution of arrival times. Using this interpretation and the results of the preceding section on the calculation of $\mathcal{G}_d^{\mathrm{II}}$, we find $\langle t^2\rangle=(2\pi\Delta_{eff})^{-2}\propto L_{tot}^3\eta$; XPM causes a superdiffusive spreading of arrival times. This functional dependence of the timing jitter upon the fiber length is identical to that found by Gordon and Haus for the combined effects of ASE noise and SPM [4]. Notice that this contribution to the timing jitter arises entirely from vertex corrections to the two-particle Green's function, $\mathcal{G}_d^{\mathrm{II}}(t)$; if we ignore vertex corrections by making the approximation $\mathcal{G}_d^{\mathrm{II}}(t)=|\langle\mathcal{G}_d(t)\rangle|^2$ and substituting Eq. (9), we find no contribution to the timing jitter.

Using Eq. (3) to calculate the statistics of the output for

inputs held constant at $E_{in}(t)=0$ and $E_{in}(t)=A$ gives an estimate of the amplitude statistics of a 1 and 0 in the output bit stream [7]. For a zero input, the output has the same Gaussian distribution as the noise. For an input of constant amplitude A,

$$\langle E_{out}(t)\rangle_1 = A\langle \mathcal{G}_d(\omega=0)\rangle,$$

$$\langle |E_{out}(t)|^2\rangle_1 = (A^2+N)\,\delta(0),$$

$$\begin{aligned}\langle \delta|E_{out}|^2(t)\,\delta|E_{out}|^2(t)\rangle_1 &= A^4[\langle |\mathcal{G}_d(\omega=0)|^4\rangle - \delta(0)^2] \\ &\quad + N(2A^2+N)\,\delta(0)^2 \\ &\approx N(2A^2+N)\,\delta(0)^2, \end{aligned} \tag{22}$$

where the δ function of zero argument is to be understood as $\Delta/2\pi=\delta(0)$ for our band-limited subband. The evaluation of these results is outlined in the Appendix. The output for a 1 input is not Gaussian: a Gaussian fit to $\langle E_{out}(t)\rangle_1$ and $\langle |E_{out}(t)|^2\rangle_1$ does not reproduce $\langle \delta|E_{out}|^2(t)\,\delta|E_{out}|^2(t)\rangle_1$. These results also predict that the contribution of XPM to the amplitude jitter is small. Equation (22) depends upon the evaluation of $\langle |\mathcal{G}(\omega=0)|^4\rangle$, which we have not carried out here. A full evaluation of this four-particle Green's function would be required to give a full expression for the amplitude jitter. If, however, we approximate $\langle |\mathcal{G}(\omega=0)|^4\rangle \approx \langle |\mathcal{G}(\omega=0)|^2\rangle^2 = \delta(0)^2$, then the amplitude jitter shows no dependence upon XPM. Corrections to this result arise from higher-order vertex corrections to the four-particle Green's function and are small for weak nonlinearity. In particular, these vertex corrections are of higher order than those that give rise to timing jitter. The predominant effect of XPM, therefore, is to induce timing errors. This conclusion is in accord with other works that have attempted to model the effects of XPM as a modification to the Gordon-Haus results for the soliton jitter [6]. In these works, individual scattering events between solitonlike pulses in different subbands of the WDM system are considered and the amplitude jitter due to scattering is ignored.

VI. INFORMATION THEORY

In this section, we will discuss how the input-output statistics calculated in Secs. III and IV may be used to construct general arguments about the maximum rate of communication via a WDM system. This is still an open problem. The only systems for which exact answers about information capacity may be obtained are those with Gaussian input-output statistics. We will use the results for Gaussian systems to obtain lower bounds upon and approximations to the capacity of the present, non-Gaussian system.

The ideas of information theory propounded by Shannon are very much akin to the ideas of statistical mechanics. For a signal made up of a sequence of letters (or field values), chosen from some alphabet (or range of values) one may define a probability distribution, $p(x)$, for the probability that a particular letter takes the value x. The amount of information per letter in such a sequence is given by the entropy of this distribution; $H(X)\equiv H[p(x)] = -\Sigma_x p(x)\log p(x)$. If the logarithm is taken in base 2, this gives the amount of information per letter in bits. One way to understand this is to think of compressing a signal. Compression involves removing redundancy in the signal (or alternatively correlations between letters) and so increasing the amount of information per letter. In this way, a signal with higher entropy will carry a higher density of information.

When one attempts to communicate via a channel, each letter, x, of the input signal will be received as some letter, y, at the output. Allowing for the uncertainty in the transmission process, the output for an input, x, will be given by some conditional distribution $p(y|x)$. For an information theorist, this defines the channel of communication. For each input letter, the output may take a range of values. The entropy of this additional spreading at the output is given by the conditional entropy; $H(Y|X)=\Sigma_x p(x)H(Y|X=x) = -\Sigma_x p(x)[\Sigma_y p(y|x)\log p(y|x)]$.

When a sequence of letters is transmitted, the amount of information in common between the input and output, otherwise known as their mutual information, is given by the total entropy of the output signal minus the additional entropy due to signal distortion;

$$I(X;Y)=H(Y)-H(Y|X). \tag{23}$$

When using a particular channel, the conditional distribution $p(y|X)$ is fixed and the remaining freedom is in the choice of input distribution $p(x)$. The idea is that if certain letters are more distorted than others, it pays one to use them less frequently, even though this means a reduction in the amount of information that the input signal can carry. The maximum transmitted information density that can be achieved is given by functionally optimizing the mutual information over the choice of input signal distribution;

$$C=\max_{p(x)} I(X;Y). \tag{24}$$

Shannon formulated these ideas rigorously and showed that this is not only an upper bound upon the rate of transmission of information, but it is also an achievable bound. A summary of these basics of information theory may be found in Ref. [2]. The functional optimization in Eq. (24) is carried out under the various system constraints, such as the average signal power. In fact, the only case for which there exists an explicit analytical solution is when the joint distributions of the input and output signals are Gaussian. In this case, the capacity is given by

$$C=\log\left[\frac{\langle x^*x\rangle\langle y^*y\rangle}{\begin{vmatrix}\langle x^*x\rangle & \langle x^*y\rangle \\ \langle y^*x\rangle & \langle y^*y\rangle\end{vmatrix}}\right]. \tag{25}$$

When the signal is band limited, the corresponding expressions for the capacity per unit bandwidth, or the spectral efficiency, is given by integrating over frequency;

$$c=\frac{1}{2\pi\Delta}\int d\omega \log\left[\frac{\langle x_\omega^* x_{-\omega}\rangle\langle y_\omega^* y_{-\omega}\rangle}{\begin{vmatrix}\langle x_\omega^* x_{-\omega}\rangle & \langle x_\omega^* y_{-\omega}\rangle\\ \langle y_\omega^* x_{-\omega}\rangle & \langle y_\omega^* y_{-\omega}\rangle\end{vmatrix}}\right]. \quad (26)$$

The factor of $1/2\pi$ arises because ω is the angular frequency conjugate to t.

We are now in a position to use these results from information theory, together with our knowledge of the quadratic and quartic statistics of the WDM system to obtain a lower bound and estimate of the capacity of the WDM optical communications system. A rigorous lower bound on the capacity of a channel defined by a conditional distribution, $p(y|x)$, may be found as follows: The quadratic input-output statistics are calculated for a Gaussian input distribution. A channel whose statistics are Gaussian and whose quadratic statistics correspond to those of the actual channel, will have capacity given by Eq. (26). This gives a lower bound on the capacity of the actual channel. A proof of this is given in Ref. [10]. The basic idea behind it is that deviations from Gaussian statistics correspond to extra correlations in the noise that the channel adds to the signal. These extra correlations allow one to make progress in determining which part of the received signal is noise and which part is signal. Ignoring these correlations— as one does in the Gaussian approximation— leads one to underestimate the extent to which noise may be removed from the signal and so to an underestimate of the system capacity. Notice that since a Gaussian distribution maximizes the entropy for a fixed variance, any deviation from Gaussianity will increase correlations.

We are interested in two ways of using the WDM optical fiber system. In the first case, both the amplitude and phase of the electrical signal are used to communicate. We call this coherent communication. In the second situation, only the amplitude of the electrical field is used to communicate. This latter case is more representative of current optical communication systems. Coherent communication was considered by Mitra and Stark in Ref. [10]. The quadratic statistics given by Eq. (4) are combined with Eq. (26) to give the following expression for the spectral efficiency in terms of the single-particle Green's function of the channel:

$$c\geqslant\int\frac{d\omega}{2\pi\Delta}\log[1+P_{eff}(\omega)/N_{eff}(\omega)], \quad (27)$$

with the effective signal and noise powers given by

$$P_{eff}(\omega)=P|\langle\mathcal{G}_d(\omega)\rangle|^2,$$

$$N_{eff}(\omega)=P(1-|\langle\mathcal{G}_d(\omega)\rangle|^2)+N, \quad (28)$$

respectively. This result is a modification of the famous Shannon [1] result for the capacity of a linear channel with additive white noise, allowing for the conversion of signal power to noise power by scattering from signals in the other subbands. The single-particle Green's function calculated in Sec.IV, $\langle\mathcal{G}_d(L,\omega)\propto\exp[-\text{const}\times P^2L]$ is substituted into Eq. (28) to obtain the functional dependence of the capacity upon system parameters. The result is sketched in Fig. 5. There are two main features worthy of comment. The first is the peak

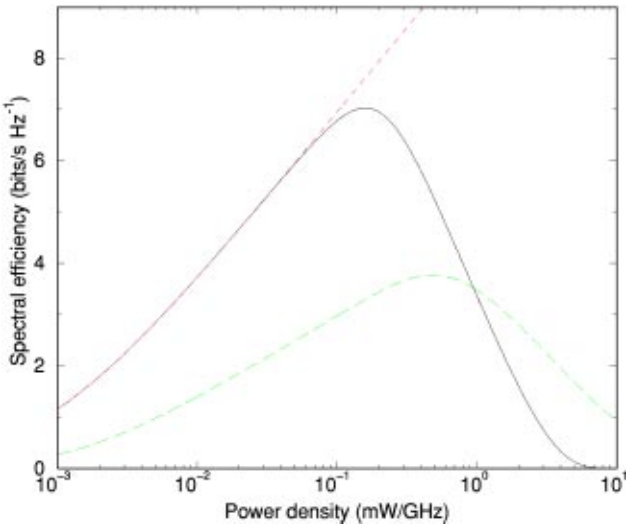

FIG. 5. Spectral efficiency vs input power: Plots of (a) Eq. (26), solid line; (b) Shannon formula, $c=\log[1+P/N]$, short-dashed line; (c) Eq. (29), long-dashed line.

in the capacity for a particular power; increasing the signal power relative to the noise power is productive up to a point, but when the signal power is very large, nonlinear interactions between signals in the different subbands become the dominant mechanism of signal distortion. The appearance of an optimal power is a common feature of the effects of nonlinearity. It occurs in the Gordon-Haus jitter of solitons [4] and in the nonlinear (or Gordon-Mollenauer) phase noise [5] (see, for example, Ref. [12]). The second feature worthy of comment is the dependence of the capacity upon length at large distances. This is important, since it is for very long distance communication that propagation nonlinearities have the most important effect. The Gaussian approximation to coherent communication gives a lower bound on capacity that decays exponentially with length at large distances. We do not expect this exponential decay to correctly reproduce the dependence of the system capacity, since it is in this limit that non-Gaussian corrections to the input-output statistics are likely to be most important. We will see below how taking account of these non-Gaussian corrections is likely to change the dependence of capacity at large distances.

Unfortunately, there exist no rigorous bounds on the capacity for incoherent communication in which the phase of the electric field is ignored. Nevertheless, expressions similar to Eq. (28) can provide useful estimates of the incoherent capacity. One such estimate is obtained by fitting an approximate Gaussian distribution of intensity about its mean values [given by Eq. (4)] to the quadratic statistics of intensity given by Eq. (5). The capacity of the channel with this Gaussian distribution is an approximation to the capacity of the actual channel. It is given by

$$c\approx\frac{1}{2}\int\frac{d\omega}{2\pi\Delta}\times\log\left(1+\frac{P^2 2\pi|\mathcal{G}_d^{\mathrm{II}}(\omega)|^2/\Delta}{P^2(1-2\pi|\mathcal{G}_d^{\mathrm{II}}(\omega)|^2/\Delta)+2PN+N^2}\right). \quad (29)$$

The prefactor of 1/2 arises because the intensity is a real field. In the limit of strong nonlinearity and large signal to noise ratio, Eq. (29) reduces to $C \approx \sqrt{\pi/2}\Delta_{eff} L i_{3/2}[P^2/(P+N)^2]$. This has an appealing interpretation. The effects of nonlinearity are contained entirely within a prefactor and provide an effective bandwidth $\Delta_{eff} = 6\Delta\sqrt{L_D^2/\eta L_{tot}^3}$. Our analysis of jitter showed that XPM mainly introduces timing errors and not amplitude error. The effective bandwidth enters because signal pulses must be separated by more than the timing error in order to be distinguishable. The information carried by each pulse is determined by the additive noise power. Notice that in the limit of strong nonlinearity, the coherent capacity has an exponential decay with length, whereas Eq. (29) has a power-law decay. The actual coherent capacity is greater than the incoherent capacity, since a greater number of degrees of freedom are used in coherent communication. The behavior of Eqs. (28) and (29) at large powers points to limitations of the Gaussian approximation. A comparison of the coherent and incoherent capacity is shown in Fig. 5. The incoherent capacity shows a peak at some optimal power. This peak is at slightly higher power than that for the coherent capacity, because phase or timing fluctuations are more sensitive to XPM than amplitude fluctuations.

It seems plausible that it should be possible to express the capacity of a channel purely in terms of its multiple-order Green's functions. Unfortunately, we do not have such an expression. Indeed, only for a channel that is entirely determined by its two-point (or single particle) Green's function is it possible to write down such an expression. In this case the channel is Gaussian and the result is Eq. (28) as obtained by Shannon [1]. Obtaining such expressions would be a significant contribution to the understanding of communication in nonlinear media. The bound on coherent communication and the estimate of capacity for incoherent communication given above give an indication of the importance of including higher-order Green's functions. Vertex corrections to higher-order Green's functions determine the extent to which the channel has non-Gaussian statistics. Including these corrections can dramatically modify the functional dependence of ones estimates of system performance upon system parameters.

VII. CONCLUSIONS

In conclusion, we have considered the impediments to communication caused by nonlinear interaction between the subbands of a WDM optical fiber system. These interactions between subbands lead to non-Gaussian input-output statistics. We have provided a simple characterization of these non-Gaussian statistics in terms of vertex corrections to higher-order Green's functions of the channel. Propagation of light in the WDM system under the action of XPM is described by a Schrödinger equation with a spatially and temporally random potential. This potential encodes the effects of XPM. We have calculated the Green's function for propagation in this channel using Feynman path-integral and diagrammatic techniques. In the case where the signal correlations are approximated by δ functions in time, these Green's functions may be obtained in closed form. The result shows a superdiffusive spreading of a signal pulse and of its arrival time. We have interpreted these results in terms of the amplitude and timing jitter of received pulses; both of which are standard measures of optical system performance. We find that the received signal pulse shows a superdiffusive spreading of its distribution of arrival times; $\langle t^2 \rangle \propto L^3$. Finally, we have used the results to obtain lower bounds and estimates of the capacity of the WDM system. These are based upon the Shannon result for the capacity of a Gaussian channel. In the case of coherent communication, they lead to a strict lower bound upon the capacity. In the case of incoherent communication, where only the signal amplitude and not its phase is used for communication, we obtain an estimate of the system capacity. The different dependences of our estimates of capacity for coherent and incoherent communication at large distances, exponential and power-law, respectively, show the importance of accounting for the non-Gaussian nature of the channel.

APPENDIX: CALCULATION OF AMPLITUDE FLUCTUATIONS

The results contained in Eqs. (22) may be derived using rules for the composition

$$\int dt' \mathcal{G}(x_1,x';t_1,t')\mathcal{G}(x',x_2;t',t_2) = \mathcal{G}(x_1,x_2;t_1,t_2), \tag{A1}$$

unitarity,

$$\mathcal{G}(x=0;t_1,t_2) = \delta(t_1-t_2), \tag{A2}$$

and time reversal

$$\mathcal{G}^*(x_1,x_2;t_1,t_2) = \mathcal{G}(x_2,x_1;t_2,t_1) \tag{A3}$$

of the Green's function for a particular realization of the potential, $V(x,t)$. These properties of the Green's function may be confirmed by considering the path-integral formulation, Eq. (6).

Using these relations, the first of Eqs. (22) may be deduced by the following series of manipulations:

$$\begin{aligned}\langle E_{out}(t)\rangle &= \left\langle A\int dt' \mathcal{G}(t,t') + n(t)\right\rangle \\ &= A\left\langle \int dt' \mathcal{G}(t,t')\right\rangle \\ &= A\langle \mathcal{G}(\omega=0)\rangle.\end{aligned}$$

This is nothing more than the average of Eq. (3). The second of Eqs. (22) requires a few more manipulations;

$$\langle |E_{out}(t)|^2\rangle = A^2\left\langle \int dt_1 dt_2 \mathcal{G}(t,t_1)\mathcal{G}^*(t,t_2)\right\rangle + \langle |n(t)|^2\rangle + A\left\langle n(t)\int dt_1 \mathcal{G}^*(t,t_1)\right\rangle + A\left\langle n(t)^*\int dt_1 \mathcal{G}(t,t_1)\right\rangle$$

$$= A^2\int dt dt_1 dt_2 \mathcal{G}(t,t_1)\mathcal{G}^*(t,t_2)/\delta(0) + N\delta(0)$$

$$= A^2\int dt dt_1 dt_2 \mathcal{G}(0,0;t_2,t_1)/\delta(0) + N\delta(0)$$

$$= A^2\int dt dt_1 dt_2 \delta(t_1-t_2)/\delta(0) + N\delta(0)$$

$$= (A^2+N)\delta(0).$$

We have used Eq. (A3) in moving from the first to second line, followed by Eqs. (A1) and (A2) moving between lines two and three and three and four, respectively. We have also used the fact that the averages over the noise and over the potential are temporally invariant. the third of Eqs. (22c) may be deduced as follows:

$$\langle \delta|E_{out}(t)|^2\,\delta|E_{out}(t)|^2\rangle = \langle |E_{out}(t)|^2|E_{out}(t)|^2\rangle - \langle |E_{out}(t)|^2\rangle^2$$

$$= A^4\int dt_1 dt_2 dt_3 dt_4 \langle \mathcal{G}^*(L;t,t_1)\mathcal{G}^*(L;t,t_2)\mathcal{G}(L;t,t_3)\mathcal{G}(L;t,t_4)\rangle$$

$$+ 4NA^2\int dt_1 dt_2 \langle \mathcal{G}^*(L;t,t_1)\mathcal{G}(L;t,t_2)\rangle + [2N^2-(A^2+N)^2]\delta(0)^2$$

$$\approx A^4\int dt_1 dt_2 dt_3 dt_4 \langle \mathcal{G}^*(L;t,t_1)\mathcal{G}(L;t,t_3)\rangle\langle \mathcal{G}^*(L;t,t_2)\mathcal{G}(L;t,t_4)\rangle + [N(2A^2+N)-A^4]\delta(0)^2$$

$$\approx A^4\left[\int dt dt_1 dt_2 \langle \mathcal{G}^*(L;t,t_1)\mathcal{G}(L;t,t_3)\rangle/\delta(0)\right]^2 + [N(2A^2+N)-A^4]\delta(0)^2 \approx N(2A^2+N)\delta(0)^2.$$

The third line is simply a rearrangement of the various terms arising after substitution of Eq. (3). A complete calculation of the first term in this expression requires evaluation of the four-particle Green's function. We do not carry out this calculation here. Instead we make the approximation discussed in the text and ignore vertex corrections. This approximation is embodied in the step between the third and fourth lines. In moving to the fifth line, we have used the fact that the average over the noise and random potential leads to temporally invariant expressions. We may then use Eqs. (A1)–(A3) in the same way as before in order to derive our final result.

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

[2] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[3] G.P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, San Diego, 1995).

[4] J.P. Gordon and H.A. Haus, Opt. Lett. **11**, 665 (1986).

[5] J.P. Gordon and L.F. Mollenauer, Opt. Lett. **15**, 1351 (1990).

[6] L. F. Mollenauer, S. G. Evangelides, and J. P. Gordon, J. Lightwave Technol. **9**, 362 (1991).

[7] D. Marcuse, J. Lightwave Technol. **8**, 1816 (1990).

[8] P. A. Humblet and M. Azizoglu, J. Lightwave Technol. **9**, 1576 (1991).

[9] R. Hui, M. O'Sullivan, A. Robinson, and M. Taylor, J. Lightwave Technol. **15**, 1071 (1997).

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

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

[12] P.P. Mitra, J.B. Stark, and A.G. Green, Opt. Photon. News **13**, 522 (2002).

[13] G.E. Falkovich, I. Kolokolov, V. Lebedev, and S.K. Turisyn, Phys. Rev. E **63**, 025601(R) (2001).

[14] R.P. Feynman and A.R. Hibbs, *Path Integrals in Quantum Mechanics* (McGraw-Hill, New York, 1965).

[15] K. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, 1999).

[16] L.G.L. Wegenger, M. Povinelli, A.G. Green, P.B. Littlewood, and P. P. Mitra (unpublished).

[17] M. Kardar, Phys. Rev. Lett. **55**, 2235 (1985).

[18] M. Kardar, G. Parisi, and Yi-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).

[19] M. Kardar and Yi-C. Zhang, Phys. Rev. Lett. **58**, 2087 (1987).