1 May 1998

ELSEVIER

Optics Communications 150 (1998) 305–318

OPTICS COMMUNICATIONS

Full length article

On timing Jitter in wavelength-division multiplexed soliton systems

Mark J. Ablowitz [a], Gino Biondini [a], Sarbarish Chakravarty [b], Rudy L. Horne [a]

[a] *Department of Applied Mathematics, University of Colorado at Boulder, Campus Box 526, Boulder, CO 80309-0526, USA*
[b] *School of Mathematics, University of New South Wales, Sydney 2052, Australia*

Received 29 October 1997; revised 15 January 1998; accepted 20 January 1998

Abstract

Collision-induced timing shifts in a wavelength-division multiplexed soliton system are computed when damping, amplification, filtering and positive dispersion management following the loss profile are included. A statistical analysis is presented which takes into account the resulting effect of the large number of collisions occurring in the fiber. Analytic expressions are derived for the root mean square timing jitter and the maximum length of error-free transmission with an arbitrary number of channels. An extensive analysis of system performance corresponding to situations with and without filters and/or dispersion management is carried out. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 42.97.S; 42.65.Tg; 42.81.Dp
Keywords: Optical communications; Fiber optics; Solitons; Wavelength-division multiplexing

1. Introduction and framework

Considerable progress has been achieved in recent experiments on fiber-optic long-distance soliton data transmission [1,2] with wavelength-division multiplexing (WDM). When compared to conventional single-channel soliton systems, WDM offers the potential for a large increase in the total capacity of soliton communication devices. However, the use of wavelength-division multiplexing raises a number of issues of theoretical and practical importance. For example, due to the periodic distribution of amplifiers, a resonant instability created by the nonlinear terms (four-wave mixing interactions) can seriously degrade the signal. This issue has been the subject of recent papers [3,4]. In Ref. [3] it was shown how proper use of dispersion management following the loss profile can alleviate the negative effects of four-wave mixing. Another serious problem that arises in soliton systems is caused by the frequency shifts and the associated displacements in pulse arrival times created by interaction of the solitons with amplifier noise, an effect which was was first discovered in the well known study by Gordon and Haus [5]. As shown in Refs. [6,7], this type of jitter can be substantially reduced by the introduction of guiding filters.

In WDM soliton systems there can also be serious timing displacement effects due to nonuniform soliton collisions in the presence of amplifiers, which induce permanent frequency and velocity shifts of the solitons [8]. Mecozzi and Haus [9] showed that the presence of filters allows the soliton parameters to relax to their original, unperturbed value. More recently, in a two-channel system, the average effect of a single collision was considered [10], and the mean square timing jitter due to soliton collisions was computed through numerical simulations [11,12]. However, to our knowledge no comprehensive statistical analysis has been published which takes into account the large number of collisions that occur in the fiber, and no analytical expressions are known for the collision-induced timing jitter. Similarly, no theory exists which includes the combined effects of filters, dispersion management, and more than two channels. In a previous Letter [13] the problem of collision-induced timing jitter

PII S0030-4018(98)00045-5

was discussed when amplifiers are present, and an analytical method to find the mean square time displacements was introduced. In this paper the method is enhanced in order to include the effects of filters and/or positive dispersion management following the loss profile. Given typical values of the system parameters, we calculate the total root mean square timing jitter and the expected bit error rate of the system. These analytical results are then extended in order to include multi-channel systems and dispersion management. In the case of two channels and constant dispersion, our results are consistent with the numerical simulations presented in Refs. [11,12]. Important system implications, such as the maximum number of channels compatible with a given system length, are explored.

The framework of this paper is the following: In Section 2, via soliton perturbation techniques, we derive the fundamental equations which allow us to compute the frequency deviation (see Eq. (2.10)) and the corresponding timing shift (see Eqs. (2.12)) resulting from a single collision between two wavelength-division multiplexed solitons in a non-ideal fiber. In Section 3 we describe a method which allows us to perform a statistical analysis of the cumulative effect of the large number of collisions that can take place in a two-channel transmission line. We give explicit formulae for the root mean square timing jitter (see Eqs. (3.3)), and we use these formulae to analyze the bit error rate of the system. In Section 4 we generalize the previous results in order to investigate multi-channel systems. We obtain estimates for the maximum number of channels compatible with a desired system length (see Eqs. (4.2)) Finally, in Section 5 we introduce dispersion management and obtain modified equations for the timing shifts (see Eqs. (5.4)). We show that a suitable choice of dispersion map allows a significant increase in the maximum number of channels. The notion of ''optimal'' dispersion management is introduced. Broadly speaking, the paper contains a methodology which allows a detailed study of timing jitter in multi-channel WDM soliton systems with a variety of perturbative effects included, such as amplification, filtering and dispersion management.

We note that very recently a similar problem was independently studied by Mecozzi [14] from a different perspective. While the topics studied are similar, the two investigations differ in the analytic approach. In particular, our results are derived by following the conventional practice of approximating the filter action with a continuous equivalent distributed along the line, while Mecozzi's work deals explicitly with the lumped nature of the filter's response. In the appendix we present a brief comparison between the two approaches, namely the lumped and distributed model of filters. We find that the results obtained with the lumped model represent an asymptotic perturbation of the distributed case for the range of parameter values considered.

2. Collision-induced timing shifts

In this section we derive equations for the timing shift resulting from a single collision between two solitons in a non-ideal fiber, both with and without the presence of filters. The fundamental propagation equation for the dimensionless field amplitude q is the perturbed nonlinear Schrödinger equation (NLS) with damping, amplification, and filtering terms:

$$iq_z + (1/2)D(z)q_{tt} + |q|^2 q = iP[z]q + iF[t]q, \quad (2.1)$$

where the operator $P[z]$ describes the periodic damping/amplification and $F[t]$ represents the filter action, averaged over an amplification period. The function $D(z)$ describes the particular choice of dispersion. In this section we consider the case of constant dispersion, i.e. we set $D(z) \equiv 1$. We will deal with dispersion-managed fibers (i.e. $D(z)$ not constant) in Section 5. Expanding the filter response in Taylor series up to third order $F[t]$ can be written as

$$F[t] = \eta_0 - \eta_2(i\partial_t - \omega_f)^2 - i\eta_3(i\partial_t - \omega_f)^3, \quad (2.2)$$

where ω_f is the filter peak frequency, $\eta_{2,3}$ are the distributed filter parameters and η_0 is the additional gain required to overcome the energy loss of the solitons due to the presence of filters. The variables z and t are the usual nondimensional space and retarded time, normalized to the dispersion length z_* and the characteristic time t_*: $z_* = 2\pi c t_*^2/(\lambda^2 \overline{D})$ and $t_* = \tau/1.763$, respectively, where $\lambda = 1.550$ μm is the central wavelength, $\overline{D}$ is the average dispersion parameter, τ is the full width at half maximum of the pulse intensity and c is the speed of light in vacuum. If the filtering is accomplished by Fabry-Perot étalons, the corresponding dimensionless parameters are

$$\eta_2 = \left[2R/(1-R)^2\right]\left(d^2/c^2 t_*^2 z_a\right),$$

$$\eta_3 = \left[(1+R)/(1-R)\right](2d/3ct_*)\eta_2,$$

where d is the mirror spacing, R is the mirror reflectivity (cf. Ref. [15]) and $z_a = l_a/z_*$ is the amplifier spacing in dimensionless units (see below). The value of d determines the frequency separation between adjacent maxima of the filter function, and is chosen so that this separation coincides with the frequency/wavelength separation between channels, $d = \pi c/\Delta\omega = \lambda^2/(2\Delta\lambda)$. In turn the frequency separation usually depends on the pulse width τ in the time domain (see Section 4). For $\Delta\lambda = 0.63$ nm the corresponding value of d is $d = 1.90$ mm. The quantity R is essentially a free parameter which determines the amount of filtering applied to the pulses. For $\tau = 20$ ps and and amplifier spacing $l_a = 25$ km, choosing $R = 0.045$ and $d = 1.90$ mm yields $\eta_2 = 0.25$ and $\eta_3 = 0.41\,\eta_2$. If $R = 0.083$ is used, the corresponding values are $\eta_2 = 0.50$ and $\eta_3 = 0.44\,\eta_2$. (Note that $\overline{D} = 0.5$ ps/(nm km) and $z_* = 201.94$ km is used throughout this paper.)

The effects of damping and amplification are described by taking $P[z]$ to be [16]

$$P[z] = -\Gamma + (e^{\Gamma z_a} - 1) \sum_{n=-\infty}^{\infty} \delta(z - nz_a), \tag{2.3}$$

where $\Gamma = \gamma z_*$ is the dimensionless damping coefficient and $z_a = l_a/z_*$ is the dimensionless amplifier spacing, while $\delta(z)$ is the Dirac delta function. Typical experimental values for γ and l_a are $2\gamma = 0.20$ dB/km $=$ 0.046 km^{-1} and $l_a = 25$ km. For $\tau = 20$ ps these values yield $\Gamma = 4.62$ and $z_a = 0.12$. As usual we rescale the field amplitude as $q(z,t) = [g(z)]^{1/2} u(z,t)$, where the function $g(z)$ denotes the periodic energy gain/loss cycle and varies on the length scale of the amplifier distance – which is small compared to the dispersion distance, since $z_a \ll 1$. That is, $g(z)$ is the periodic function

$$g(z) = a_0^2 \exp[-2\Gamma(z - nz_a)],$$
$$nz_a \le z < (n+1) z_a. \tag{2.4}$$

The quantity a_0^2 (which represents the ratio between the pulse power after an amplifier and the average power) is chosen so that the average of $g(z)$ is unity over an amplification cycle; i.e.,

$$a_0^2 = 2\Gamma z_a / [1 - \exp(-2\Gamma z_a)].$$

With the substitution $q(z,t) = [g(z)]^{1/2} u(z,t)$ Eq. (2.1) becomes

$$iu_z + (1/2) D(z) u_{tt} + g(z)|u|^2 u = iF[t]\, u. \tag{2.5}$$

In the ideal case $g(z) = 1$ and $\eta_{0,2,3} = 0$.

We are interested in the physically significant case when the solitons are widely separated in frequency space. We first deal with two solitons; the multi-soliton case can be discussed in a similar way, and will be the subject of Section 4. To leading order we decompose u as

$$u(z,t) \sim u_1(z,t) + u_2(z,t) + O(\epsilon) \tag{2.6}$$

(cf. Refs. [17,18]), where $u_j = A_j \operatorname{sech} S_j \exp(i\chi_j)$, $S_j = A_j[t - \Omega_j(z - z_o)]$, and $\chi_j = \Omega_j t - (\Omega_j^2 - A_j^2) z/2$, with S_j chosen so that two solitons collide at $z = z_o$. We assume A_j and Ω_j to be slowly varying functions of z with respect to the characteristic amplification period z_a. Also, we take A_j to be normalized to unity and we set initially $\Omega_2 = -\Omega_1 =: \Omega > 0$, so that the dimensionless frequency separation is 2Ω and $\epsilon = 1/\Omega \ll 1$ (cf. Refs. [17,18]). The corresponding physical wavelength separation between channels is $\Delta\lambda = (\lambda^2/\pi c t_*)\Omega$. Substituting u into (2.5) and separating out frequency channels – i.e. expanding in the neighborhood of the frequencies Ω_1 and Ω_2 – yields two coupled equations for the soliton channels. Explicitly, u_1 satisfies [8]

$$iu_{1,z} + (1/2) u_{1,tt} + g(z)\left(|u_1|^2 + 2|u_2|^2\right) u_1 = iF[t] u_1. \tag{2.7}$$

The equation for u_2 is simply obtained by interchanging u_1 and u_2 in Eq. (2.7).

If there are no perturbations, the NLS equation admits an infinite number of conserved quantities. The first two of them correspond to the total energy W and mean frequency M. For each pulse these two quantities are defined as [8]

$$W_j = \int |u_j|^2 \, dt, \tag{2.8a}$$

$$M_j = \operatorname{Im}\left[\int (\partial u_j/\partial t) u_j^* \, dt\right] / W_j, \tag{2.8b}$$

where the asterisk denotes complex conjugation (hereafter all integrals are from $-\infty$ to ∞ unless explicit integration limits are given). In particular, when the explicit expression for u_1 and u_2 are used, $W_j = 2A_j$ and $M_j = \Omega_j$.

Soliton perturbation theory allows us to obtain the variation (slow with respect to the amplifier period) of the quantities A_j and Ω_j by taking the z derivative of Eqs. (2.8) and using the evolution equations for $u_{1,2}$ (cf. Ref. [8]). Then, using the explicit expression for $u_{1,2}$ and requiring that the W_j be constant determines the extra gain: $\eta_0 = \eta_2[(\Omega - \omega_f)^2 + 1/3]$. Similarly, by forming the equation for M_2 and substituting the expressions for u_1, u_2 yields the following differential equation for Ω (see also Refs. [10,12]):

$$\frac{d\Omega}{dz} = g(z) \frac{d}{dz} f(z - z_o) - (4/3)\eta_2(\Omega - \omega_f), \tag{2.9}$$

where $f(z) = (2/\Omega)[2\Omega z \cosh(2\Omega z) - \sinh(2\Omega z)] \operatorname{csch}^3(2\Omega z)$ expresses the frequency shift in an ideal collision. Although Eq. (2.9) is well known in the literature, our analysis differs significantly from previous studies on timing jitter.

We recall that the multiple peaks of the Fabry-Perot étalon filters are taken to coincide with the unperturbed soliton frequency. Therefore, since we are studying the evolution of solitons in channel 2, we set the filter frequency ω_f to the unperturbed value of Ω_2, i.e. $\omega_f = \Omega(-\infty)$, where for convenience we take the left end of the fiber to be $-\infty$. (Of course, had we wanted to study solitons in channel 1 we would put $\omega_f = -\Omega(-\infty)$.) Treating the first term in the right hand side of Eq. (2.9) as a forcing we obtain a first order ordinary differential equation which can be integrated exactly to obtain:

$$\Delta\Omega(z) = \int_{-\infty}^{z} \exp[-(4/3)\eta_2(z - z')] g(z') \times \frac{d}{dz'} f(z' - z_o) \, dz', \tag{2.10}$$

where $\Delta\Omega(z) = \Omega(z) - \Omega(-\infty)$. The residual frequency shift is defined as $\Delta\Omega(\infty)$. When filters are present, we see immediately that $\Delta\Omega(\infty) = 0$. Without filters we have instead $\Delta\Omega(\infty) = \int g(z) f'(z - z_o) \, dz$.

The frequency shift $\Delta\Omega(z)$ induces a shift in the arrival time, through the direct coupling between the soli-

ton carrier frequency and its group velocity. That is, the timing shift at a given position L in the fiber is given by

$$\delta t = -\int_{-\infty}^{L} \Delta\Omega(z)\,\mathrm{d}z$$
$$= -\int_{-\infty}^{L} \mathrm{d}z\, \mathrm{e}^{-(4/3)\eta_2 z} \int_{-\infty}^{z} \mathrm{d}z'\, \mathrm{e}^{(4/3)\eta_2 z'} g(z') \times \frac{\mathrm{d}}{\mathrm{d}z'} f(z'-z_\mathrm{o}). \tag{2.11}$$

After an integration by parts Eq. (2.10) yields the following convenient formula in the limit $L \to \infty$:

$$\delta t_\mathrm{f}(z_\mathrm{o}) = (3/4\eta_2)\int f(z-z_\mathrm{o})\, g'(z)\mathrm{d}z \tag{2.12a}$$

(Note that the integral that yields the timing shift when filters are present is the same that yields the residual frequency shift in the case without filters.) In the limit when the filter strength η_2 vanishes Eq. (2.10) yields instead, for large L:

$$\delta t_\mathrm{nf}(z_\mathrm{o}) = (L-z_\mathrm{o})\int f(z-z_\mathrm{o})\, g'(z)\mathrm{d}z - \int f(z-z_\mathrm{o})\mathrm{d}z - \int f(z-z_\mathrm{o})(g(z)-1)\mathrm{d}z - \int (z-z_\mathrm{o})\, g'(z) f(z-z_\mathrm{o})\mathrm{d}z, \tag{2.12b}$$

where exponentially small terms are dropped. The first term is the dominant contribution, while the second term is the timing shift in the ''ideal'' ($g=1$) case. (Note that in Ref. [13] only the dominant term was included.) It is interesting to observe that, by considering the dominant contribution in (2.12b) (i.e. by keeping only the first term), for large L we obtain the simple relation

$$\delta t_\mathrm{f}/\delta t_\mathrm{nf} = 3/\left[4\eta_2(L-z_\mathrm{o})\right], \tag{2.13}$$

which allows a direct comparison of the timing shift resulting from collisions with and without the presence of filters.

The integrals in Eqs. (2.12) can be readily calculated. To a good approximation we can take Ω to be constant in the integrals. Since $g(z)$ is periodic, all integrals involving $g(z)$ can be written in terms of Fourier series. The Fourier coefficients are

$$g_m = \frac{a_0^2}{z_a}\int_0^{z_a} \mathrm{e}^{-2\Gamma z + 2m\pi i z/z_a}\,\mathrm{d}z = \frac{\Gamma z_a}{\Gamma z_a - m\pi i}, \quad m = 0, \pm 1, \pm 2, \ldots \tag{2.14}$$

Hence we substitute (2.4) into (2.12a) and use Eq. (2.14), together with the integral

$$\int \mathrm{e}^{2n\pi i z/z_a} f(z)\,\mathrm{d}z = \left(n^2\pi^4/4\Omega^4 z_a^2\right)\mathrm{csch}^2\left(n\pi^2/2\Omega z_a\right)$$

(cf. Ref [8]). In this way, after rearranging the terms in the summation, we can express $\delta t_\mathrm{f}(z_\mathrm{o})$ as a sine Fourier series:

$$\delta t_\mathrm{f}(z_\mathrm{o}) = -(3/4\eta_2)\sum_{m=1}^{\infty} a_m \sin\left[2m\pi z_\mathrm{o}/z_a - \phi_m\right], \tag{2.15a}$$

where

$$a_m = c_m|g_m|, \quad \phi_m = \arg[g_m],$$
$$c_m = \left(\pi^5/\Omega^4 z_a^3\right) m^3 \,\mathrm{csch}^2\left(m\pi^2/2\Omega z_a\right),$$

and where

$$|g_m| = \Gamma z_a/\sqrt{(\Gamma z_a)^2 + (m\pi)^2},$$
$$\arg[g_m] = \arctan\left[m\pi/(\Gamma z_a)\right].$$

(Note that, in the previous formulae, Ω refers to the leading order contribution, i.e. the unperturbed value.) In the case without filters, substituting (2.4) into (2.12b) and using $\int f(z)\,\mathrm{d}z = 1/\Omega^2$ yields instead

$$\delta t_\mathrm{nf}(z_\mathrm{o}) = -(L-z_\mathrm{o})\sum_{m=1}^{\infty} a_m \sin\left[2m\pi z_\mathrm{o}/z_a - \phi_m\right] - \sum_{m=1}^{\infty} b_m \cos\left[2m\pi z_\mathrm{o}/z_a - \phi_m\right] - 1/\Omega^2, \tag{2.15b}$$

where

$$b_m/a_m = (3z_a/2\pi m)\, h\left(\pi^2 m/2\Omega z_a\right)$$

and $h(x) = (2/3)x\coth(x) - 1$. In both cases the timing shift depends on the position of the collision point z_o relative to the amplifiers, a fact that is important when computing the statistical averages.

Eqs. (2.15), together with (2.13), are the fundamental results of this section. In the following parts of this paper we extend these formulae in order to derive a general theory which allows us to obtain the timing jitter and the bit error rate of a WDM soliton system in various physically interesting situations.

3. Statistical analysis of the timing shifts

Next we turn our attention to the statistical analysis of the timing shifts in a WDM soliton transmission line. We want to compute the relative shift in the arrival time of two adjacent solitons in a given channel. For the sake of concreteness, consider two solitons in channel 2 at locations n and $n+1$ along the data stream. We assume that when the nth soliton in channel 2 is interacting with a soliton in channel 1 (at position z_n), the $(n+1)$th soliton $n+1$ in channel 2 also interacts with the adjacent soliton in channel 1 (at position z_{n+1}). Since adjacent solitons in

any given channel are separated by a distance which is much smaller than the characteristic amplifier length, the value of the collision point z_o is effectively the same for such adjacent collisions. Therefore the relative timing shift between the nth and the $(n+1)$th solitons in channel 2 as an effect of the interaction with solitons in channel 1, is given by (cf. Refs. [13,19])

$$\delta t_{n,n+1} = (b_{n+1} - b_n)\delta t(z_n), \tag{3.1}$$

where b_n can be 0 or 1 depending on the arbitrary encoding of data in channel one: Here $b_n = 1$ indicates the effective presence of a soliton in channel 1, while $b_n = 0$ implies that no soliton is present. The total relative timing shift between two adjacent solitons in channel 2 is given by the sum over all the relative shifts due to collisions occurring over the total length of the system:

$$\Delta t = \sum_{n=1}^{N} \delta b_n \delta t(z_n), \tag{3.2}$$

where $\delta b_n = b_{n+1} - b_n$, and N is the total number of collisions that can take place within a fiber of total length L (in units of z_*). A convenient normalized distance is the minimum distance that a soliton in channel 2 would propagate between each successive collision, $z_s = T/(2\Omega)$, where T is the dimensionless bit period (also called time window). As usual we assume a separation of 5 pulse widths between adjacent pulses in the same channel; that is, $T = 5\tau/t_*$. Given z_s, the total number of collisions in the fiber is obtained as $N = L/z_s$. It is also customary to introduce the ''collision length'' z_c of the two channels as the distance between the points of the fiber where the initial half-power point of the soliton in one channel overlaps in time with the final half-power point of the soliton in the other channel (see Ref. [8]). Then $z_c = \tau/\Omega t_*$, i.e. z_s is related to the collision length by the relation $z_s = (5/2)z_c$. We also note that the $L = N_a z_a$, where N_a is the total number of amplifiers present in the system. Hence the total number of collisions can also be expressed as $N = (2/5)(z_a/z_c)N_a$, where, in terms of dimensional quantities, the dimensionless ratio z_c/z_a is $z_c/z_a = 2\tau/(\bar{D} l_a \Delta\lambda)$.

Since the total number of collisions along the fiber is large (usually more than a hundred), in Eq. (3.2) we treat b_n and $\delta t(z_n)$ as independent random variables. For b_n the randomness comes from the arbitrary encoding of data in channel 1; for $\delta t(z_n)$ it comes from the arbitrariness of the location of the collision point with respect to the location of the amplifiers. That is, we assume that each collision is characterized by a random value of z_n. As a result, the values of $\delta t(z_n)$ will also be random (except of course for the contribution coming from the ideal case, which is always constant and equal to $1/\Omega^2$). Assuming a random sequence of bits b_n, the difference δb_n takes the value 0 with probability 1/2 and the values ± 1 with probability 1/4. Thus we have [13,20] $\langle \delta b_n \delta t(z_n)\rangle = \langle \delta b_n\rangle\langle \delta t(z_n)\rangle = 0$; that is, the relative average timing shift is exactly zero. Also, $\langle(\delta b_n \delta t(z_n))^2\rangle = \langle(\delta b_n)^2\rangle\langle \delta t(z_n)^2\rangle = \langle \delta t(z_n)^2\rangle/2$. Finally, since Δt is the sum of a large number of small independent random variables, its statistics is described by a Gaussian distribution. The variance of the total timing shift is the sum of the individual variances. Since $\delta t(z_o)$ is periodic in z_o, the averages on z_n are taken over an amplifier period. Then, using Eq. (2.15a) we have $\langle \delta t(z_n)\rangle_f = 0$ and $\langle \delta t(z_n)^2\rangle_f = (3/4\eta_2)^2\|a\|^2/2$, where we define

$$\|a\|^2 = \sum_{m=1}^{\infty} a_m^2 .$$

That is, $\langle \delta t(z_n)^2\rangle$ is independent of n, unlike the case without filters [13]. Hence, using $\langle(\delta b_n)^2\rangle = 1/2$ we obtain the variance of the total timing shift simply as

$$\langle(\Delta t)^2\rangle_f = \sum_{n=1}^{N} \langle \delta t(z_n)^2\rangle_f/2 = (3/8\eta_2)^2\|a\|^2(L/z_s). \tag{3.3a}$$

A similar analysis can be done to obtain the improved estimate when no filters are present. In this case we use Eq. (2.15b). Though more involved because of the presence of cross products, the calculations proceed much in the same way as before. Due to the phase difference between terms in the two Fourier series in (2.15b), all the cross-products average to zero, and the final result is

$$\langle(\Delta t)^2\rangle_{nf} = (1/12)\|a\|^2 L z_s(L/z_s + 1)(L/z_s + 1/2) + \|b\|^2(L/4z_s) + 1/(2\Omega^4), \tag{3.3b}$$

where the relation $\sum_{n=1}^{N} n^2 = N(N+1)(2N+1)/6$ was used. The first term in (3.3b) is the dominant contribution, while the third one (which is independent of the system length L) represents the collision-induced timing jitter in an ideal fiber and the second term in (3.3b) results from the second term in (2.15b) – i.e. from the combined effects of the last two integrals in (2.12b). In particular, by taking the dominant contribution, we have

$$\langle(\Delta t)^2\rangle_{nf} = (1/12)\|a\|^2 L^3/z_s, \tag{3.3c}$$

which is the result appearing in Ref. [13]. From the ratio of (3.3a) and (3.3c) we get

$$\sqrt{\langle(\Delta t)^2\rangle_f}\Big/\sqrt{\langle(\Delta t)^2\rangle_{nf}} = 3\sqrt{3}/(4\eta_2 L), \tag{3.4}$$

which allows a direct comparison of the root mean square (rms) timing jitter in systems with and without filters.

In Fig. 1 we plot the dimensional rms timing jitter $\sqrt{\langle(\Delta t)^2\rangle}$ (in ps) versus the total length of fiber for a two-channel system: a) without filters, keeping all terms in Eq. (3.3b); b) with filters included, Eq. (3.3a). In both cases we plot the resulting curves with $z_c/z_a = 2$, 1.4, 0.7 for $\tau = 20$ ps, $l_a = 25$ km and $\eta_2 = 0.25$. Note that here and in rest of the paper, we convert the dimensionless

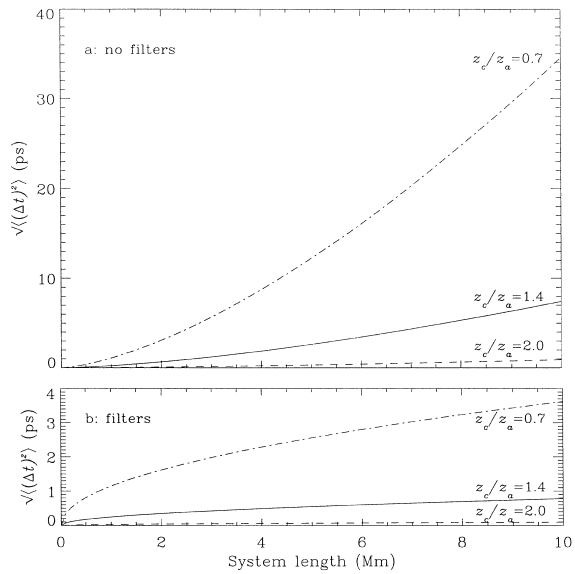

Fig. 1. The root mean square timing shift in a two-channel WDM system for $\tau = 20$ ps, $\overline{D} = 0.5$ ps/(nm km), $l_a = 25$ km and $\eta_2 = 0.25$: (a) without filters; (b) with filters. Dashed lines: $z_c/z_a = 2.0$; solid lines: $z_c/z_a = 1.4$; dotted lines: $z_c/z_a = 0.7$.

quantities into physical units by multiplying them with appropriate scale factors. In particular, the rms timing jitter is scaled by $t_* = 11.34$ ps and the fiber length by $z_* = 201.94$ km in order to obtain the corresponding dimensional quantities. It is evident from these figures that there is an order of magnitude decrease in the collision-induced timing jitter when filters are included. This analytical result is in qualitative agreement with the numerical simulations reported in Refs. [11,12]. The value of $\sqrt{\langle(\Delta t)^2\rangle}$ in ps at 10000 km as a function of z_c/z_a is displayed in Fig. 2. Note the maximum for $z_c/z_a \simeq 0.6$. The characteristic behavior of $\sqrt{\langle(\Delta t)^2\rangle}$ can be explained by noting that, when the collision is distributed over many amplifiers, i.e. for large values of z_c/z_a, the contributions of the amplifiers begin to cancel each other. Conversely, for very small values of z_c/z_a the collision is too rapid in order for the amplifiers to have a significant effect. (See also Ref. [8] in the case without filters.)

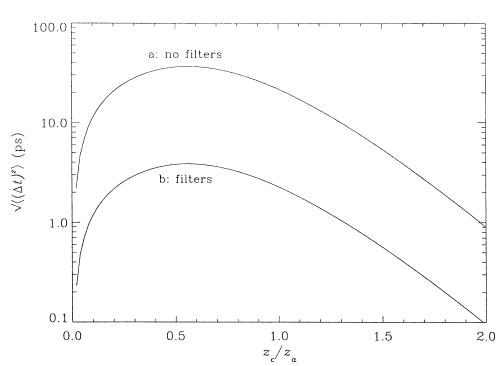

Fig. 2. The root mean square timing jitter at 10000 km as a function of collision length: (a) without filters; (b) with filters. The values of τ, $\overline{D}$, l_a and η_2 are the same as in Fig. 1.

Since the values of Δt follow a Gaussian distribution, the expected bit error rate (BER) of a WDM soliton system is given by (see e.g. Ref. [21])

$$\mathrm{BER} = \mathrm{erfc}\left[rT/\sqrt{2\langle(\Delta t)^2\rangle} \right], \tag{3.5}$$

where $\mathrm{erfc}(x)$ is the complementary error function and r is a parameter which measures the sensitivity of the receiver, which is defined by assuming that the maximum time displacement tolerated by the receiver is rT. Like the timing jitter, the bit error rate has a peak for $z_c/z_a \simeq 0.6$. We also recall that in Ref. [8] the value $z_c/z_a = 2$ was regarded as the limit of the ''safe region'' for data transmission. This bound on z_c would seriously limit the maximum allowable frequency separation between channels, and hence the maximum number of channels in a WDM system (see below). However, in the presence of filters the bit error rate for a two-channel system is always less than 10^{-10}, even for values of r as low as 0.2, and is always less than 10^{-40} if $r = 0.4$. From this discussion we expect that with filters the previous limit can be significantly relaxed, and a larger number of ''error-free'' channels in a long-distance WDM soliton system will be possible. A detailed analysis of these results, the multi-soliton case, and the system implications of our analysis will be the subject of the next section.

4. Timing jitter in multi-channel soliton systems

The results derived in the previous section can be generalized to the case when more than two channels are present. It is well-known that in the ideal case soliton interactions are pairwise (see e.g. Refs. [16–18]). We assume that the same is true to leading order when perturbation terms are included. Then the total variance in the relative arrival times of adjacent solitons in the jth channel is simply the sum of the variances resulting from the interactions with the other $J-1$ channels, J being the total number of channels. That is, the total mean square timing jitter each channel j is

$$\langle(\Delta t)^2\rangle_j = \sum_{\substack{k=1\\k\neq j}}^{J} \langle(\Delta t)^2\rangle_{jk}, \tag{4.1}$$

with $\langle(\Delta t)^2\rangle_{jk}$ given by either (3.3a) or (3.3b), where Ω is replaced by $\Omega_{jk} = |\Omega_j - \Omega_k|/2$. Note that, as a consequence, z_c becomes $z_c^{jk} = \tau/(\Omega_{jk} t_*)$. In dimensional variables, $\Delta\lambda$ is replaced by $\Delta\lambda_{jk} = |\lambda_j - \lambda_k|$, which yields $\Omega_{jk} = (\pi c t_*/\lambda^2)\Delta\lambda_{jk}$.

The channel wavelengths are taken to be $\lambda_j = \lambda + \Delta\lambda_{\min}[j-(J+1)/2]$, where $\lambda = 1.550$ μm is the central wavelength. The wavelength separation then becomes $\Delta\lambda_{jk} = (j-k)\Delta\lambda_{\min}$, where $\Delta\lambda_{\min}$ is determined by requiring (as is standard) that adjacent channels be separated by five spectral widths. For sech-shaped pulses, $\Delta\lambda_{\min} = 1.575\,\lambda^2/c\tau$. In particular, for $\tau = 20$ ps, $\Delta\lambda_{\min} = 0.63$ nm. The dimensionless frequency separation is then $2\Omega_{jk} = (2\pi c t_*/\lambda^2)\Delta\lambda_{jk} = (j-k)\Delta\Omega_{\min}$, where $\Delta\Omega_{\min} = 5.6$. In turn, for $l_a = 25$ km this value yields $z_c^{jk}/z_a = \tau/\Omega_{jk} t_* z_a = 5.08/(j-k)$, which means that adjacent channels are characterized by a relatively large value of z_c/z_a, and this ratio decreases with increasing frequency separation.

As the number of channels increases, the bit error rate of the system grows not only because of the increased number of collisions, but also because the outermost channels undergo rapid collisions, which are characterized by values of z_c/z_a closer to the peak value 0.6 . Another way of looking at this effect is to note that the coefficients a_m defined in the previous section decay exponentially for very small values of Ω and algebraically for large values of Ω. Therefore, as the frequency separation Ω between the two channels increases, we expect that a_m will get closer to their maximum value.

This effect is exemplified in Table 1, where, for a typical 8-channel WDM soliton system, we report the total timing jitter experienced by solitons in a given channel as a result of collisions with the solitons in all the other channels, as obtained by using (4.1) with either (3.3a) or (3.3c). (No appreciable differences are found in the case without filters by using the more accurate estimate given by (3.3b).) As it can be seen, the outermost channels offer the worst performance.

Using the formulae derived in the previous section for the total timing jitter we can write an estimate for the maximum distance of error-free transmission attainable with any given number of channels, or, equivalently, for the maximum number of channels compatible with a desired system length. In fact, since $\mathrm{erfc}(x) < 10^{-9}$ for $x \geq 4.3201$, if we require the bit error rate of the system to be less than 10^{-9} Eq. (3.5) yields a condition on the root mean square timing shift: $(\Delta t)_{\mathrm{rms}} \leq 0.1637\,rT$. For example, for $r = 0.4$ and $\tau = 20$ ps the maximum rms timing jitter in dimensional units tolerated by the system would be 6.55 ps. More in general, in order to have a total bit error rate less than 10^{-9} in a multi-channel system, it is necessary that the rms timing jitter in each of the individual channels be less than $0.1637\,rT$. Or, in dimensional units, the maximum rms timing jitter needs to be less than $0.818\,r\tau$. The previous conditions, in turn, pose an upper limit to the system length. With filters we have, using Eq. (3.3a),

$$\max_{j=1,\ldots,J}\left[\sum_{\substack{k=1\\k\neq j}}^{J} \|a\|_{jk}^2/z_c^{jk}\right] L \leq 37.0\,\eta_2^2 r^2 .$$

Substituting $z_c^{jk} = \tau/(\Omega_{jk} t_*)$ this yields the dimensionless maximum length

$$L_{\max}^{\mathrm{f}} = 65.2\,\eta_2^2 r^2 \Big/ \max_{j=1,\ldots,J} C_j , \tag{4.2a}$$

where $C_j = \sum_{k=1;\,k\neq j}^{J} \Omega_{jk}\|a\|_{jk}^2$. (Of course, the corresponding limit in dimensional units is obtained by multiplying (4.2) by the unit length z_*.) Without filters, the dominant contribution (3.3c) yields, with the same arguments,

$$L_{\max}^{\mathrm{nf}} = 4.79\,r^{2/3} \Big/ \max_{j=1,\ldots,J} C_j^{1/3} . \tag{4.2b}$$

In both cases the maximum length is limited by the channel with the largest overall timing jitter. By increasing the number of channels the corresponding C_j increase both because the sum in k is taken over a larger number of channels and because the channels added are the outermost ones in the frequency domain, and interact more strongly with all the others.

The resulting values for the system length in physical units are plotted in Fig. 3 versus the total number of channels, for $\tau = 20$ ps, $l_a = 25$ km, $\eta_2 = 0.25$ and $r = 0.4$. As expected, the maximum length of error-free transmission decreases rapidly as the number of channels is increased. Note also that $L_{\max}$ depends in a very simple way on the receiver sensitivity r and the filter strength η_2, which means that the results corresponding to different values of these parameters can be obtained by simple rescaling.

For any fixed length L, Eqs. (4.2) can also be used to obtain a limit on the allowed number of channels. By taking 10000 km as the limit distance for transoceanic communications we see that, for $r = 0.4$, only a four-channel system offers a satisfactory performance in the case without filters. Even with guiding filters (with $\eta_2 = 0.25$) the maximum number of channels is limited to 9. (For

Table 1
The total timing jitter and bit error rate in each of the eight channels of a system with $\tau = 20$ ps, $\overline{D} = 0.5$ ps/(nm km), $l_a = 25$ km and $\eta_2 = 0.25$

Channel	λ (nm)	Without filters		With filters	
		Δt_{rms} (ps)	BER	Δt_{rms} (ps)	BER
1	1547.79	50.33	0.43	5.28	4×10^{-14}
2	1548.42	37.39	0.28	3.92	1×10^{-57}
3	1549.05	23.80	9×10^{-2}	2.50	$<10^{-60}$
4	1549.68	11.68	6×10^{-4}	1.23	$<10^{-60}$
5	1550.32	11.68	6×10^{-4}	1.23	$<10^{-60}$
6	1550.95	23.80	9×10^{-2}	2.50	$<10^{-60}$
7	1551.58	37.39	0.28	3.92	1×10^{-57}
8	1552.21	50.33	0.43	5.28	4×10^{-14}

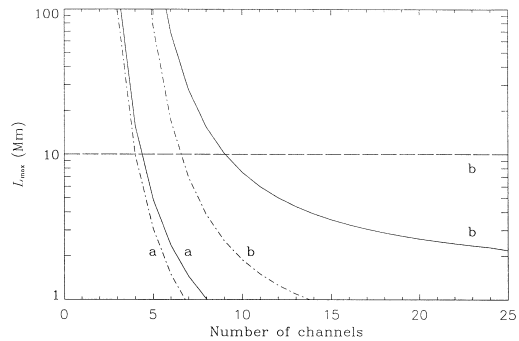

Fig. 3. The maximum length of error-free transmission for a given number of channels: a) without filters; b) with filters. Solid lines: $r = 0.4$; dot-dashed lines: $r = 0.2$. System parameters are the same as in Fig. 1.

$r = 0.2$ this last number further decreases to 7.) A better result can be obtained if stronger filtering is employed: for $\eta_2 = 0.5$ the previous limit raises to 21 channels. As we are going to see in the next section, all these estimates are going to be dramatically improved by the addition of dispersion management.

5. Timing jitter in dispersion-managed soliton systems

The reduction of collision-induced timing jitter in two-channel systems with dispersion management but without filters was studied theoretically and numerically in Refs. [23–25]. Dispersion-management was also shown to have a beneficial effect on Gordon-Haus jitter in a recent experimental paper [22]. Conveniently, the analytical methods presented in the previous sections can be readily adapted to include dispersion management. We go back to Eq. (2.5), where now we allow the dispersion D to be a function of z. As is standard (cf. Refs. [23,25]), we perform the rescaling $\zeta(z) = \int_0^z D(z)\,\mathrm{d}z$, where $D(z)$ is normalized such that the average of $D(z)$ is one over an amplification period, i.e. $(1/z_a)\int_0^{z_a} D(z)\,\mathrm{d}z = 1$. This choice allows an immediate comparison with the constant dispersion case and ensures that $\zeta(z_a) = z_a$. Note however that the transformation from z to ζ is invertible only if $D(z)$ is always of one sign; here we consider the case of $D(z)$ being positive. The evolution equation for u becomes now

$$iu_\zeta + (1/2)u_{tt} + G(z(\zeta))|u|^2 u = iF[t]u/D(z(\zeta)), \tag{5.1}$$

where $G(z) = g(z)/D(z)$. To leading order we write the solution of (5.1) as $u \sim u_1 + u_2$, where the solitons $u_{1,2}$ are now given by $u_j = A_j \operatorname{sech} A_j[t - \Omega_j(\zeta - \zeta_o)] \exp i[\Omega_j t - (\Omega_j^2 - A_j^2)\zeta/2]$, with $\zeta_o = \zeta(z_o)$. That is, solitons now evolve in ζ, not z. This seemingly formal change has important consequences in the following.

We apply the same methods described in the previous sections for the case of constant dispersion. Namely, we construct differential equations for the pulse energy and average frequency following the same procedures used in Section 2. (As it might be expected, the same condition holds for the extra gain.) The evolution of the average frequency is now given by

$$\frac{\mathrm{d}(\Delta\Omega)}{\mathrm{d}\zeta} = G(z(\zeta))\frac{\mathrm{d}}{\mathrm{d}\zeta} f(\zeta - \zeta_o) - (4/3)\eta_2 \Delta\Omega / D(z(\zeta)).$$

In order to solve this equation it is convenient to rewrite it in terms of the variable z. This is easily accomplished by noting that $\mathrm{d}\zeta = D(z)\,\mathrm{d}z$:

$$\frac{\mathrm{d}(\Delta\Omega)}{\mathrm{d}z} = G(z)\frac{\mathrm{d}}{\mathrm{d}z} f(\zeta(z) - \zeta_o) - (4/3)\eta_2 \Delta\Omega, \tag{5.2}$$

We note that the evolution equation for $\Delta\Omega(z)$ is very similar to the case of constant dispersion. More precisely, Eq. (5.2) is identical to Eq. (2.9), *except that*: (i) $g(z)$ is replaced by $G(z)$, and (ii) the solitons are now functions of ζ instead of z, which is reflected in the argument of f. Like we did for (2.9), we solve (5.2) exactly (using the simplicity of the integrating factor when the equation is written in terms of the variable z), and we obtain:

$$\Delta\Omega(z) = \int_{-\infty}^{z} \exp[-(4/3)\eta_2(z - z')] G(z') \times \frac{\mathrm{d}}{\mathrm{d}z'} f(\zeta(z') - \zeta_o)\,\mathrm{d}z'. \tag{5.3}$$

Then (as in (2.11–2.12) we integrate $\Delta\Omega(z)$ to get (after an integration by parts)

$$\delta t_{\mathrm{f}}(z_o) = (3/4\eta_2)\int f(\zeta - \zeta_o)\frac{\mathrm{d}}{\mathrm{d}\zeta} G(z(\zeta))\,\mathrm{d}\zeta, \tag{5.4a}$$

where we have performed the inverse substitution (i.e. from z to ζ) in order to write the result in a more convenient form. In the limit $\eta_2 = 0$ we have instead, for the dominant contribution (see also (2.12b)):

$$\delta t_{\mathrm{nf}}(z_o) = (L - z_o)\int f(\zeta - \zeta_o)\frac{\mathrm{d}}{\mathrm{d}\zeta} G(z(\zeta))\,\mathrm{d}\zeta. \tag{5.4b}$$

Comparing Eqs. (5.4) and (2.12) we see that the only change from the case of constant dispersion is that z_o is substituted by ζ_o, and the Fourier coefficients g_m of $g(z)$ are replaced by the Fourier coefficients G_m of $G(z(\zeta))$ (see (5.6) below). Of course, the ideal choice of dispersion would be an exponential tapering that exactly matches the shape of $g(z)$ (i.e. $D(z) = g(z)$), in which case we would have $G(z) \equiv 1$, and all the Fourier coefficients except the

first one would be zero. However, reproducing the ideal exponential shape is not practical, and one resorts to a stepwise approximation. In this scheme the amplifier distance is divided into S spans identified by the end points $z_0, z_1, \ldots, z_S$ (with $z_0 = 0$ and $z_S = z_a$), and the dispersion assumes the constant value D_s in each of the sub-intervals $z_{s-1} \le z < z_s$. Given a sequence of intermediate points z_s, a convenient choice for the values D_s is obtained by requiring that the average of $G(z)$ is one in each of the sub-intervals, i.e. $[1/(z_s - z_{s-1})]\int_{z_{s-1}}^{z_s} G(z)\,\mathrm{d}z = 1$:

$$D_s = \frac{a_0^2}{z_s - z_{s-1}} \int_{z_{s-1}}^{z_s} \mathrm{e}^{-2\Gamma z}\,\mathrm{d}z = \frac{z_a}{z_s - z_{s-1}} \frac{\mathrm{e}^{-2\Gamma z_{s-1}} - \mathrm{e}^{-2\Gamma z_s}}{1 - \mathrm{e}^{-2\Gamma z_a}} \tag{5.5}$$

(cf. Ref. [25]). Note that this choice ensures that the average of $D(z)$ is one over one amplification period, i.e. $(1/z_a)\int_0^{z_a} D(z)\,\mathrm{d}z = 1$. The Fourier coefficients of $G(z(\zeta))$ are then computed by breaking the resulting integral into S sub-intervals:

$$G_m = \frac{1}{z_a}\int_0^{z_a} G(z(\zeta))\,\mathrm{e}^{2m\pi i\zeta/z_a}\,\mathrm{d}\zeta = a_0^2 \sum_{s=1}^{S} \frac{q_{m,s}}{2\Gamma z_a - 2m\pi i D_s}, \tag{5.6}$$

where the relation $G(z(\zeta))\,\mathrm{d}\zeta = g(z)\,\mathrm{d}z$ was used; the coefficients $q_{m,s}$ are defined as follows:

$$q_{m,s} = \exp\left[-(2\Gamma z_{s-1} - im\kappa_{s-1})\right] - \exp\left[-(2\Gamma z_s - im\kappa_s)\right],$$

with $\kappa_s = (2\pi/z_a)\int_0^{z_s} D(z)\,\mathrm{d}z$, while a_0 was defined in Section 2. Note that $G_0 = 1$, because the average of $G(z(\zeta))$ is unity over an amplification cycle, i.e. $(1/z_a)\int_0^{z_a} G(z(\zeta))\,\mathrm{d}\zeta = 1$.

Once the change from g_m to G_m is made, *all the procedures presented in the previous sections can still be used in order to (i) compute the root mean square timing jitter, (ii) the bit error rate and (iii) to estimate the maximum system length.* In other words, Eqs. (3.3)–(3.5) and (4.1), (4.2) remain valid provided that the a_m are replaced by $a_m H_m$, where the quantities H_m express the reduction factor of the Fourier coefficients, and are given by

$$H_m = \left|\frac{G_m}{g_m}\right| = \frac{\sqrt{(2\Gamma z_a)^2 + (2m\pi)^2}}{1 - \mathrm{e}^{-2\Gamma z_a}} \times \sqrt{\left(\sum_{s=1}^{S} \frac{\alpha_{m,s}}{d_{m,s}}\right)^2 + \left(\sum_{s=1}^{S} \frac{\beta_{m,s}}{d_{m,s}}\right)^2}, \tag{5.7}$$

where

$$\alpha_{m,s} = 2\Gamma z_a \mathrm{Re}\left[q_{m,s}\right] - 2m\pi D_s \mathrm{Im}\left[q_{m,s}\right],$$
$$\beta_{m,s} = 2\Gamma z_a \mathrm{Im}\left[q_{m,s}\right] + 2m\pi D_s \mathrm{Re}\left[q_{m,s}\right],$$
$$d_{m,s} = (2\Gamma z_a)^2 + (2m\pi D_s)^2.$$

In the special case of a two-step approximation, the only free parameter of the dispersion management is the ratio θ of the intermediate length z_1 to the amplification period z_a, that is $\theta = z_1/z_a$. The values of the dispersion are then

$$D_1 = \exp\left[(1-\theta)\Gamma z_a\right]\sinh\left[\theta\Gamma z_a\right]/\left(\theta\sinh\left[\Gamma z_a\right]\right),$$

$$D_2 = \exp\left[-\theta\Gamma z_a\right]\sinh\left[(1-\theta)\Gamma z_a\right] / \left((1-\theta)\sinh\left[\Gamma z_a\right]\right).$$

In Fig. 4 we plot the *first* reduction factor H_1 as a function of $\theta = z_1/z_a$ for a two-step approximation, for $l_a = 25$, 50, 75 km. The resulting curves are remarkably similar in some cases to the behavior of the timing shift as a function of θ when filters are present; or, equivalently, to the behavior of the residual frequency shift in the case without filters (cf. Ref. [25]. Also note that the same reduction factor applies, independently of whether or not filters are present, cf. Eq. (3.4)). In fact, the minimum of the curve $H_1(\theta)$ for $l_a = 50$ km occurs for $\theta = 0.3418$, which corresponds exactly to the value $l_{\min} = 17.09$ km appearing in Ref. [25]. This striking result is an indication that, in the sums contained in Eqs. (3.2) and (3.3), the first term alone ($m = 1$) is often sufficient to obtain a very good approximation to the correct result for the timing jitter. In particular, this happens when the frequency separation between channels is not too large. However, for very large values of Ω (i.e. for small values of the ratio z_c/z_a) more terms in the Fourier series come into play, and the optimal value of θ cannot be found on the simple basis of the previous argument.

In Fig. 5 we compare the resulting timing jitter in ps for a two-channel system with a two-step approximation, and with $\tau = 20$ ps $l_a = 25$ km, $\eta_2 = 0.25$ at a typical value $z_c/z_a = 1.4$: (a) without filters; (b) with filters. In both

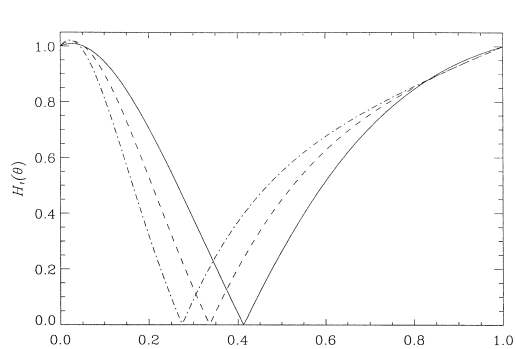

Fig. 4. The behavior of the coefficient H_1 defined in (5.7) as a function of θ. Solid line: $l_a = 25$ km; dashed line: $l_a = 50$ km; dot-dashed line: $l_a = 75$ km. The minima occur respectively for $\theta = 0.4163$, $\theta = 0.3148$ and $\theta = 0.2814$.

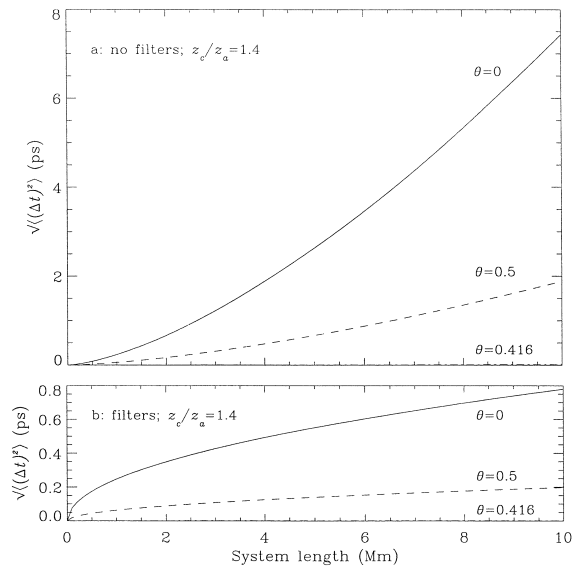

Fig. 5. The root mean square timing jitter in a two-channel WDM system with $z_c/z_a = 1.4$: (a) without filters; (b) with filters. Solid lines: no dispersion management; dashed lines: $\theta = 0.5$; dot-dashed lines: $\theta = 0.416$. The values of τ, $\overline{D}$, l_a and η_2 are the same as in Fig. 1. For $\theta = 0.416$ the value of $\Delta t_{\rm rms}$ at 10000 km is 1.416×10^{-2} ps or 1.484×10^{-3} ps for the cases without and with filters, respectively.

cases we display the results in the following three situations: (i) no dispersion management; (ii) fiber spans of equal length in real units (i.e. $\theta = 0.5$); (iii) the value of θ which yields the minimum of H_1 ($\theta = 0.416$). The reduction in timing jitter in the cases presented ranges from a factor of 4 (with $\theta = 0.5$) to almost 100 (with $\theta = 0.416$)!

Fig. 6 displays the rms timing jitter at 10000 km as a function of z_c/z_a for systems with and without filters in the following three cases: (i) no dispersion management (i.e. $\theta = 0$); (ii) $\theta = 0.5$; (iii) the value of θ that yields the maximum reduction, as determined numerically for each value of z_c/z_a. It should be noted that, while a very large reduction factor applies for $z_c/z_a \geq 1.5$, for smaller values of z_c/z_a (and in particular near the peak, which is now located at $z_c/z_a \simeq 0.3$) the reduction factor becomes relatively small. The reason for this behavior is that, as we said earlier, for small values of the ratio z_c/z_a many terms in the Fourier series contribute to determine the final value of the timing jitter. Therefore, even if a numerical minimization of $\Delta t_{\rm rms}$ as a function θ is performed, the resulting reduction factor can not be expected to be very large, because a simple two-step dispersion map does not have enough free parameters to minimize enough of the relevant Fourier coefficients.

A large reduction in timing jitter is also found when dealing with the multi-channel case. As an example, in Table 2 we present for comparison the results relative to the same eigth-channel system presented in Table 1, where in this case we add a two-step approximation with $\theta = 0.5$. In Fig. 7 we show the estimates of the maximum system length in dimensional units versus the number of channels (for $\tau = 20$ ps $l_a = 25$ km, $\eta_2 = 0.25$ and $r = 0.4$), for the same three cases described in Fig. 5. The changes are dramatic. In particular, we see that a system with no filters and suitable dispersion management offers performances which in some cases are comparable to those of a system with filters but without dispersion management. However, the presence of filters is still needed if a large number of channels is desired.

Using the same method described in Section 4, we can also estimate the maximum number of channels for error-free transmission up to 10000 km. With a two-step dispersion map with $\theta = 0.5$, and for $r = 0.4$, the maximum number of channels goes from 4 to 9 in the case without filters; with filters and dispersion management this number increases from 9 to 39! If $\theta = 0.416$ is used, the maximum

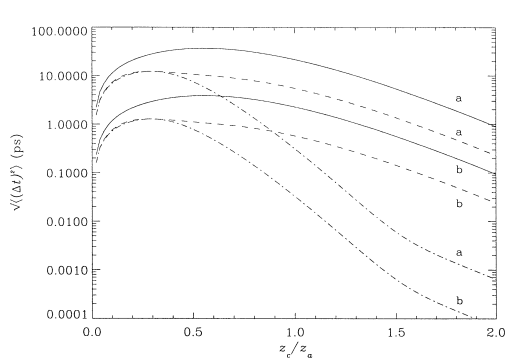

Fig. 6. The root mean square timing jitter at 10000 km as a function of collision length: (a) without filters; (b) with filters. Solid lines: no dispersion management; dashed lines: $\theta = 0.5$; dot-dashed lines: the value of θ that yields the minimum for $\Delta t_{\rm rms}$, as determined numerically for each value of z_c/z_a. The values of τ, $\overline{D}$, l_a and η_2 are the same as in Fig. 2.

Table 2
The total timing jitter and bit error rate in each of the eight channels of a system with dispersion management. A two-step dispersion map with $\theta = 0.5$ is used. System parameters have the same value as in Table 1

Channel	λ (nm)	Without filters		With filters	
		$\Delta t_{\rm rms}$ (ps)	BER	$\Delta t_{\rm rms}$ (ps)	BER
1	1547.79	12.87	2×10^{-3}	1.35	$<10^{-60}$
2	1548.42	9.51	3×10^{-5}	1.00	$<10^{-60}$
3	1549.05	6.04	3×10^{-11}	0.63	$<10^{-60}$
4	1549.68	2.96	2×10^{-41}	0.31	$<10^{-60}$
5	1550.32	2.96	2×10^{-41}	0.31	$<10^{-60}$
6	1550.95	6.04	3×10^{-11}	0.63	$<10^{-60}$
7	1551.58	9.51	3×10^{-5}	1.00	$<10^{-60}$
8	1552.21	12.87	2×10^{-3}	1.35	$<10^{-60}$

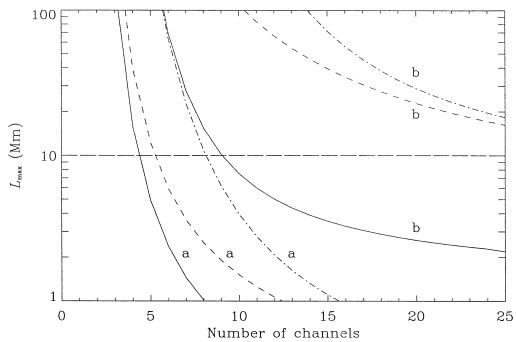

Fig. 7. The maximum length of error-free transmission for a given number of channels in a system with and without dispersion management: (a) without filters; (b) with filters. Solid lines: no dispersion management; dashed lines: $\theta = 0.5$; dot-dashed lines: $\theta = 0.416$. The values of τ, $\overline{D}$, l_a and η_2 are the same as in Fig. 3, while $r = 0.4$.

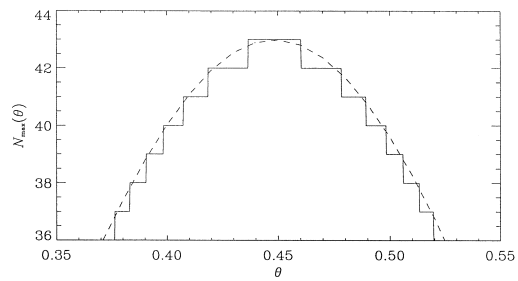

Fig. 8. The maximum number of channels for error-free transmission over 10000 km as a function of the parameter θ in a two-step dispersion map. The values of τ, $\overline{D}$, l_a, η_2 and r are the same as in Fig. 7. The dashed line is a level curve of a two-dimensional surface plot of L_{max} as a function of both N and θ, taken at $L_{max} = 10000$ km.

number of channels further increases to 9 and 41 for systems without and with filters, respectively. (For $r = 0.2$ and $\theta = 0.416$ the corresponding numbers are 7 and 17.) These estimates are summarized in Table 3.

Yet another consequence of these results is that dispersion management could permit the relaxation of some system constraints. For example, with $l_a = 50$ km and a two-step approximation with $\theta = 0.315$ (which is the value that minimizes $H_1(\theta)$ when $l_a = 50$ km), error-free transmission is possible with 4 and 11 channels for the cases without and with filters, respectively. (Without dispersion management these numbers would be respectively 2 and 3). Some other choices of parameters are presented in Table 3.

We conclude this section by noting that, when timing jitter is considered, an ''optimal'' dispersion map can be found as follows. As we have seen, to any given distance there corresponds a maximum number of channels N_{max} compatible with error-free transmission. In general, for a two-step dispersion map, this number will depend on the particular value chosen for θ.

Table 3
The maximum number of channels for error-free transmission, with $\overline{D} = 0.5$ ps/(nm km) and various choices of the remaining system parameters. ''No DM'' refers to constant dispersion; ''DM'' to a two-step dispersion map with $\theta = 0.416$

System parameters				Maximum number of channels			
τ (ps)	l_a (km)	η_2	r	Without filters		With filters	
				No DM	DM	No DM	DM
20	25	0.25	0.4	4	9	9	41
20	25	0.25	0.2	4	8	7	18
20	25	0.5	0.4	4	9	21	> 50
20	50	0.25	0.4	2	4	3	9
20	50	0.5	0.4	2	4	4	32
10	25	0.25	0.4	1	2	2	4
10	25	0.5	0.4	1	2	2	10

In Fig. 8 we plot the maximum number of channels N_{max} as a function of θ by requiring error-free transmission over 10000 km (employing the standard values used before for the system parameters, i.e. $\tau = 20$ ps, $l_a = 25$ km, $\eta_2 = 0.25$ and $r = 0.4$). Fig. 8 clearly shows an absolute maximum number of channels $N_M = 43$ for $\theta = \theta_M = 0.450$. We also note that $\theta_M \neq 0.416$, the latter corresponding to minimizing the reduction factor for the first Fourier coefficient, i.e. $H_1(\theta)$ (cf. Fig. 4). The reason for this difference is that, when the number of channels in the system becomes large, the outer channels are characterized by values of z_c/z_a which yield the maximum contribution to the timing jitter (cf. Fig. 6). As shown previously, for these channels many Fourier coefficients are significant in determining the resulting value of the timing jitter. In fact, we estimate that, for $N \simeq 40$, the first five Fourier coefficients all contribute significantly to the timing jitter. We also note that, in general, the function $H_m(\theta)$ has m minima, and the location of these minima is different for each of these functions. Therefore, reducing only the first Fourier coefficient is not sufficient anymore, and one can expect that a higher number of steps will be required in order to achieve an efficient reduction of collision-induced timing jitter. In this situation the behavior of a multi-channel system (and therefore the optimal choice of the dispersion map) can differ significantly from the corresponding behavior of a two-channel system.

6. Conclusions

In this paper we have developed and used an analytical method to study collision-induced timing jitter in a WDM soliton system. The main features of our investigation are the following: (i) Analytical formulae for the timing shift

with/without filters are obtained. (ii) The filter action is approximated using the conventional approach of converting to a continuous equivalent. (iii) A stochastic analysis is developed to obtain the rms timing shift which takes into account a large number of collisions. (iv) Employing analytical formulae for the bit error rate, the concept of ''error-free'' transmission is introduced and discussed for a two-channel WDM soliton system. (v) Multi-channel WDM soliton systems are studied by appropriate extensions of the two-channel approach. (vi) The concept of dispersion management is introduced in the context of collision-induced timing jitter. Both two-channel and multi-channel systems are investigated for typical values of system parameters, comparing the results of a two-step dispersion map to the case of constant dispersion. vii) Minimizing the first reduction factor $H_1(\theta)$ of the Fourier coefficients g_m demonstrates excellent agreement with previous results of two-channel systems [25]. Upon consideration of a multi-channel system, the notion of ''optimal'' dispersion map is discussed. (viii) Using conventional values of system parameters ($\tau = 20$ ps, $\overline{D} = 0.5$ ps/(nm km), $l_a = 25$ km and $\eta_2 = 0.25$) we find that ''error-free'' transmission over 10000 km is possible in the following cases: (a) 4 channels without filters and without dispersion management; (b) 9 channels without filters but with dispersion management; (c) 9 channels with filters but without dispersion management; (d) 43 channels with both filters and ''optimal'' dispersion management.

Naturally, when such a large number of channels is considered, many factors that we have neglected in our discussion can be expected to be relevant. Amongst the others: third order dispersion, shock terms, Raman scattering, amplifier noise, and the effect of sliding the filter frequency. Similarly, dispersion management will need to be more fully analyzed in order to find efficient ways of reducing resonant four-wave mixing instabilities. We plan to investigate some of these issues in future papers.

Acknowledgements

We are pleased to thank A. Mecozzi for useful discussions and for sending us a preprint of his recent work. We also thank R.B. Jenkins and J.R. Sauer; indeed, Jenkins' Ph.D. Thesis was invaluable during the realization of this work. This effort was partially sponsored by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number F49620-97-1-0017. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government.

Appendix A. A comparison between lumped and distributed filter models

In this appendix we compare the expression for the timing shift derived in this paper to that obtained recently by Mecozzi [14]. A major difference between the two investigations is that in our study we approximate the filter transfer function by a continuous equivalent distributed along the fiber, whereas in Ref. [14] the lumped nature of the filter action is dealt with explicitly. This is equivalent to replacing the constant η_2 in $F(t)$ appearing in Eq. (2.5) by

$$\eta(z) = \eta_2 z_a \sum_{n=-\infty}^{\infty} \delta(z - nz_a), \tag{A.1}$$

which takes into account the fact that filters are located at discrete positions in the fiber. Note in (A.1) the presence of the factor z_a which is necessary to ensure a proper normalization of the filter action (cf. the definition of η_2 in Section 2). For simplicity we consider the filters to be positioned at amplifier locations. While this model represents a closer approximation to the physical system, it also introduces some complications in the analytical treatment of the problem. Here we show that the correction on the timing jitter obtained with the lumped model can be viewed asymptotically as a perturbation of the simpler situation in which the continuous approximation is used. In fact, for all the cases discussed in this paper the results obtained with the continuous approach are a close approximation to those resulting from the lumped model. To our knowledge, this constitutes the first explicit discussion of the validity of the distributed filter model in a study of filtering effects on timing displacements.

To estimate the timing jitter with the lumped filter model we insert Eq. (A.1) into Eq. (2.5). The evolution of the average frequency Ω can be found using the same procedures described earlier in this paper, and is governed by

$$\frac{\mathrm{d}(\Delta\Omega)}{\mathrm{d}z} = g(z)\frac{\mathrm{d}}{\mathrm{d}z}f(z - z_o) - (4/3)\eta(z)\Delta\Omega, \tag{A.2}$$

which replaces Eq. (2.9). For simplicity we neglect dispersion management, for the moment. Its effects can be introduced in a straightforward way, and will be discussed later. Like Eq. (2.9), this differential equation can be integrated exactly to obtain

$$\Delta\Omega(z) = \int_{-\infty}^{z} \exp[-(N(z) - N(z'))]g(z') \times \frac{\mathrm{d}}{\mathrm{d}z'}f(z' - z_o)\,\mathrm{d}z', \tag{A.3}$$

where

$$N(z) = (4/3)\int_{0+}^{z} \eta(z')\mathrm{d}z' = (4/3)\eta_2 z_a \lfloor z/z_a \rfloor ,$$

and the floor function is defined as $\lfloor x \rfloor = n$ for $n \le x < n + 1$. By introducing the ''sawtooth'' function $s(x)$ as

$$s(x) = x - (n + 1/2), \quad n \le x < n + 1, \tag{A.4}$$

we have $\lfloor x \rfloor = x - s(x) - 1/2$. Therefore we can rewrite the function $N(z)$ as

$$N(z) = (4/3)\eta_2 z_a \left[z/z_a - s(z/z_a) - 1/2 \right].$$

The reason for introducing $s(x)$ is to separate explicitly in $N(z)$ the part which contains the linear growth $z/z_a - 1/2$ from the periodic perturbation $s(z/z_a)$. That is, unlike $N(z)$ the function $s(z/z_a)$ is periodic with period z_a and zero mean. This allows us to make a convenient comparison with the results coming from the distributed model. As before, we obtain the timing shift by integrating Eq. (A.3):

$$\delta t = -\int_{-\infty}^{\infty} \mathrm{d}z\, \mathrm{e}^{-(4/3)\eta_2 z + (4/3)\eta_2 z_a s(z/z_a)} \times \int_{-\infty}^{z} \mathrm{d}z'\, \mathrm{e}^{(4/3)\eta_2 z' - (4/3)\eta_2 z_a s(z'/z_a)} g(z') f'(z' - z_o), \tag{A.5}$$

Note that $|s(x)| \le 1/2$, while $|\eta_2| \ll 1$ and $|z_a| \ll 1$. Therefore we expand the exponentials in Eq. (A.5) as $\exp[\pm(4/3)\eta_2 z_a s(z/z_a)] \simeq 1 \pm (4/3)\eta_2 z_a s(z/z_a)$. In this way we can write the timing shift as

$$\delta t = \delta t_0 + \delta t_1 + O\left(\eta_2^2 z_a^2\right), \tag{A.6}$$

where δt_0 is the result coming from the continuous approximation (namely, Eq. (2.12a)), and the leading order correction δt_1 is given by (after an appropriate integration by parts)

$$\delta t_1 = z_a \int \mathrm{d}z\, s(z/z_a) g(z) f'(z - z_o) - (4/3)\eta_2 z_a \int \mathrm{d}z\, s(z/z_a) \Delta\Omega(z). \tag{A.7}$$

(Note that in the previous formula $\Delta\Omega(z)$ refers to the distributed approximation, given by (2.10).) The integrals in (A.7) can be calculated by expanding $s(z/z_a)$ and $g(z)$ in Fourier series. The Fourier coefficients of $g(z)$ are given in (2.14), while the Fourier coefficients of $s(z/z_a)$ are

$$s_m = \frac{1}{z_a}\int_0^{z_a} s_a(z/z_a)\, \mathrm{e}^{2m\pi i z/z_a} \mathrm{d}z = \frac{1}{2m\pi i}, \quad m = \pm 1, \pm 2, \ldots, \tag{A.8}$$

with $s_0 = 0$. Once these expansions are made, both integrals in (A.7) break up into a double series, whose terms can be calculated by repeating the same steps used in Section 2 to obtain (2.15a) from (2.12a). Explicitly, the first integral in (A.7) is given by

$$\delta t_1 = -(i/2) z_a \sum_{m,n=-\infty}^{\infty} s_m g_n c_{m+n}\, \mathrm{e}^{2(m+n)\pi i z_o/z_a},$$

where the coefficients g_m and c_m were defined in Section 2, and where higher order contributions, including the second integral in (A.7), have been neglected. After rearranging the terms in the double series, the resulting expression for δt_1 is

$$\delta t_1 = (z_a/2\pi) \sum_{m=1}^{\infty} c_m |\xi_m| \times \sin\left[2m\pi z_o/z_a + \arg[\xi_m] + \pi/2\right], \tag{A.9}$$

where

$$\xi_m = \sum_{n=1}^{\infty} (g_{m-n} - g_{m+n})/n.$$

Therefore the correction δt_1 can be written as a sine Fourier series which is formally identical to the one that yields the continuous approximation (i.e. Eq. (2.15a)). The only differences are: i) the Fourier coefficients g_m in (2.15a) are substituted by the ξ_m in (A.9); ii) the factor appearing in front of the series is $z_a/2\pi$ in (A.9) instead of $3/4\eta_2$ in (2.15a). In fact, a first comparison of the two results can be done by simply taking the ratio of these two quantities. For $\tau = 20$ ps, $\overline{D} = 0.5$ ps/(nm km), $l_a = 25$ km and $\eta_2 = 0.25$, this ratio is $2\eta_2 z_a/3\pi = 6.56 \times 10^{-3}$, suggesting that the term δt_1 is only a small correction to the continuous approximation.

A more accurate comparison of the two results can be done by taking the ratio of the L^2 norms of the Fourier series: $\|\delta t_1\|/\|\delta t_0\|$, where $\|\delta t_1\|^2 = (z_a/2\pi)^2 \times \sum_{m=1}^{\infty} c_m^2 |\xi_m|^2$, and $\|\delta t_0\|$ is defined in a similar way. For the previous values of the system parameters, this ratio is almost independent of the frequency separation Ω, and is never larger than 0.05. That is, $\|\delta t_1\|/\|\delta t_0\| < 5\%$, independently of Ω. Also, since the ξ_m depend linearly on the g_m, the same method holds if dispersion management is taken into account. Namely, all the steps we used to obtain Eq. (A.9) can be reproduced exactly when dispersion-managed systems following the loss profile are considered. The only change will be the substitution of the Fourier coefficients g_m with G_m, which parallels the corresponding change in Eq. (2.15a); the asymptotic analysis is the same. We compared the lumped and distributed cases for 2-, 3- and 4-step dispersion maps for numerous values of frequency separation between two channels. While the contribution in the lumped case increases, nevertheless we find that for 2-step dispersion map with equal step sizes, $\|\delta t_1\|/\|\delta t_0\| < 15\%$ throughout the range of channel spacings considered. Upon varying the step sizes, we find that the previous estimate also holds whenever $\|\delta t_0\|$ (expressed in dimensional units) is larger than 0.1 ps. We note that values of $\|\delta t_0\|$ larger than 0.1 ps are the

only ones which could result in a BER of 10^{-9} at 10000 km, even for $r = 0.2$. For 3- and 4-step dispersion maps with equal step sizes, our initial numerical calculations indicate that, while the values of $\|\delta t_0\|$ decrease with increasing number of steps, the corresponding values of $\|\delta t_1\|$ remain relatively unaffected. Therefore the difference on timing displacements between lumped and distributed filtering can be more significant than the 2-step case. We will address this issue in more detail in a separate publication.

References

[1] L.F. Mollenauer, P.V. Mamyshev, M.J. Neubelt, Demonstration of soliton WDM transmission at up to 8×10 Gbit/s, error-free over transoceanic distances, OFC1996, San Jose, USA, February 1996, Paper PD22.
[2] M. Nakazawa, K. Suzuki, H. Kubota, A. Sahara, E. Yamada, Electron. Lett. 33 (1997) 1233.
[3] P.V. Mamyshev, L.F. Mollenauer, Optics Lett. 21 (1996) 396.
[4] M.J. Ablowitz, G. Biondini, S. Chakravarty, R.B. Jenkins, J.R. Sauer, Optics Lett. 21 (1996) 1646.
[5] J.P. Gordon, H.A. Haus, Optics Lett. 11 (1986) 665.
[6] A. Mecozzi, J.D. Moores, H.A. Haus, Y. Lai, Optics Lett. 16 (1991) 1841.
[7] Y. Kodama, A. Hasegawa, Optics Lett. 17 (1992) 31.
[8] L.F. Mollenauer, S.G. Evangelides, J.P. Gordon, J. Lightwave Technol. 9 (1991) 362.
[9] A. Mecozzi, H.A. Haus, Optics Lett. 17 (1992) 988.
[10] M. Midrio, P. Franco, F. Matera, M. Romagnoli, M. Settembre, Optics Comm. 112 (1994) 283.
[11] F. Matera, M. Romagnoli, M. Settembre, S. Wabnitz, Y. Kodama, Ultimate capacity of WDM soliton transmission with sliding filters, in: Nonlinear Guided Waves and Their Applications OSA Technical Digest Series, vol. 6, Optical Society of America, Washington, DC, 1995.
[12] Y. Kodama, S. Wabnitz, Optics Comm. 113 (1995) 395.
[13] R.B. Jenkins, J.R. Sauer, S. Chakravarty, M.J. Ablowitz, Optics Lett. 20 (1995) 1964.
[14] A. Mecozzi, J. Opt. Soc. Am. B 13 (1998) 152.
[15] E.A. Golovchenko, A.N. Pilipetskii, C.R. Menyuk, J.P. Gordon, L.F. Mollenauer, Optics Lett. 20 (1995) 539.
[16] A. Hasegawa, Y. Kodama, Phys. Rev. Lett. 66 (1991) 161.
[17] S. Chakravarty, M.J. Ablowitz, J.R. Sauer, R.B. Jenkins, Optics Lett. 20 (1995) 136.
[18] M.J. Ablowitz, G. Biondini, S. Chakravarty, R.B. Jenkins, J.R. Sauer, J. Opt. Soc. Am. B 14 (1997) 1788.
[19] R.B. Jenkins, Ph.D. thesis, University of Colorado, Boulder, Colorado, 1995.
[20] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.
[21] J.P. Powers, An introduction to Fiber Optics Systems, Aksen Associate Series in Electrical and Computer Engineering, 1993.
[22] G.M. Carter, J.M. Jacob, C.R. Menyuk, E.A. Golovchenko, A.N. Pilipetskii, Optics Lett. 22 (1997) 513.
[23] A. Hasegawa, S. Kumar, Y. Kodama, Optics Lett. 21 (1996) 39.
[24] H. Sugahara, H. Kato, Y. Kodama, Electron. Lett. 33 (1997) 1065.
[25] W. Forysiak, J.F.L. Devaney, N.J. Smith, N.J. Doran, Optics Lett. 22 (1997) 600.