Channel capacity and modeling of optical fiber communications — **Source link** ⤤

Salam Elahmadi, Mandyam D. Srinath, Dinesh Rajan, Richard Haberrman

Institutions: Southern Methodist University

Published on: 28 Apr 2009 - Wireless and Optical Communications Networks

Topics: Channel capacity, MIMO, Precoding, Intersymbol interference and Fading

Related papers:

- Channel capacity and outage probability analysis of sub carrier intensity modulated BPSK system over M-distribution free space optical channel

- OFDM channel estimation with dispersive fading channels

- The capacity of the cascaded fading channel in the low power regime

- Dispersion of the coherent MIMO block-fading channel

- Bounds on the Capacity of OFDM Underspread Frequency Selective Fading Channels

Channel Capacity And Modeling Of Optical Fiber Communications

Salam Elahmadi, Mandyam D. Srinath, Dinesh Rajan
Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275
Richard Haberrman
Department of Mathematics, Southern Methodist University, Dallas, TX 75275

Abstract—We use the method of multiple scales borrowed from perturbation theory to derive a new time-domain transfer function of the nonlinear fiber-optic wave-division multiplexing (WDM) communications channel. The obtained channel response, derived from the nonlinear Schrodinger equation is shown to be equivalent to the multi-path fading frequency selective channel encountered in wireless links. In the linear regime, the channel response is shown to be equivalent to a standard intersymbol interference (ISI) channel and is used to derive new bounds on the capacity of the dispersive optical fiber channel.

I. INTRODUCTION

Over the past few decades, fiber optic networks have experienced a remarkable evolution. Their bandwidth and reach have increased nearly a thousand-fold. Earlier networks used multi-mode lasers, multi-mode fibers with losses exceeding 2.5 dB/km, and direct modulation techniques. They carried data streams at modest rates (a few Mb/s) over modest distances (a few kilometers). Today's optical networks are capable of carrying traffic at hundreds of Gb/s over a few thousand kilometers. Many factors have contributed to this marked increase in performance. Chief among them are deployment of single-mode fibers with losses below 0.2 dB/km and use of distributed feedback (DFB) lasers, external modulation techniques, dispersion compensating fibers, forward error correction (FEC), and photonic amplification. The migration from single-channel to multi-channel systems has also contributed greatly to boosting the capacity of today's optical networks. In wave-division multiplexing (WDM), a widely deployed technology, multiple wavelengths are used to simultaneously carry data from multiple channels. The basic architecture of these systems is depicted in figure 1. Light from all channels is amplified at several sites along the line. The transmitted data is recovered by demultiplexing and converting the received light back into the electrical domain.

The focus of this paper is to answer the following basic question: What is the maximal error-free data rate that can be expected from a fiber optic medium? While Shannon's results address this question in a general sense, with the exception of the additive white Gaussian noise (AWGN) model, computing the theoretical information capacity has proved to be mathematically challenging for many channels. The fiber nonlinearities in WDM networks allow for coupling amongst channels, giving rise to undesired effects such as stimulated Raman scattering (SRS), cross-phase modulation (XPM), and four-wave mixing (FWM). Any attempt to derive the information capacity of the highly nonlinear WDM channel based on the AWGN assumption will evidently yield inaccurate results. The right approach is to obtain a model that takes into account the physical properties of the fiber optic WDM channel. In [1-4]], solutions to this problem were proposed but effects such as Raman scattering , optical-to-electrical (OE) regeneration, or dispersion were either ignored or treated in isolation. In sections II and III, we use the generalized nonlinear Schrodinger equation (NLSE) governing pulse propagation in an optical fiber to derive new channel models. We show that the channel response can be modeled as a fading, multipath, frequency-selective channel in the nonlinear
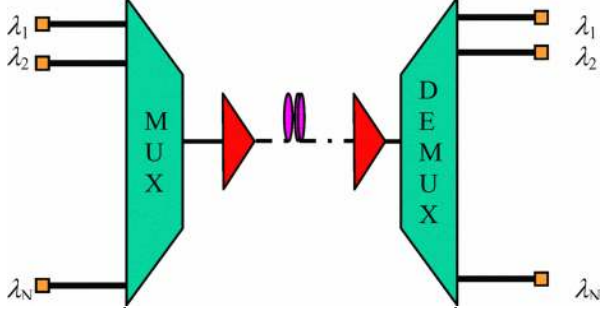
Fig. 1. Basic architecture of a WDM system.

regime and as an ISI channel in the linear regime. In sections IV and V, we make use of the rich literature treating the Shannon capacity of such a channel to derive new capacity bounds for the linear dispersive optical channel. Capacity bounds based on our solution to the nonlinear channel will be presented in a future publication.

II. PROPAGATION MODEL

Under the slowly-varying-envelope approximation which is valid for pulse widths greater than 1ps, and ignoring higher order dispersion coefficients, pulse propagation in the jth channel of a N-channel WDM system is governed by the following equation

$$\frac{\partial A_j}{\partial z} + \beta_1 \frac{\partial A_j}{\partial t} + \frac{i}{2}\beta_{2j}\frac{\partial^2 A_j}{\partial t^2} + \frac{\alpha_j}{2}A_j =$$
$$i\gamma_j\left(\left|A_j\right|^2 + 2\sum_{n=0,n\neq j}^{N-1}\left|A_n\right|^2\right)A_j - \frac{1}{2}\sum_{\substack{n=0\\n\neq j}}^{N-1}g_{jn}\left|A_n\right|^2 A_j \,, \tag{1}$$

where

A_j:	Envelope of the jth optical channel
β_{1j}:	Group delay of the jth channel
β_{2j}:	Group velocity dispersion (GVD) parameter of the jth channel
α_j:	Loss coefficient of the jth channel
γ_j:	Nonlinearity parameter for the jth channel
g_{kj}:	Raman coupling coefficient between channels j and n.

Equation (1) is known in the literature as the generalized nonlinear Schrodinger equation [5] (NLSE). The third and fourth terms on the left of side of equation (1) account for the linear effects: dispersion and fiber loss. The first term on the right-hand side is responsible for both self-phase modulation (SPM) through the pulse magnitude in the jth channel, and cross-phase modulation XPM through the pulse magnitudes from other channels, while the last term accounts for SRS crosstalk. In order to compute the information capacity of the channel described by (1), not only must a closed-form solution of equation (1) be found

but it must also include a mapping between the initial pulse $A(0,t)$ and the pulse $A(z,t)$ after it has propagated some distance z. Unfortunately, the NLSE equation has no known closed-form solution. To proceed further, the following assumptions are in order. The XPM term in (1) is twice as important as the SPM term. Moreover, as the number of channels in a WDM system increases, the XPM term will grow even larger and will in fact come to dominate so that SPM can be neglected. The second term on the left-hand side of (1) indicates that different wavelengths travel at different speeds. After a certain distance called pulse walk-off, pulses on different channels walk away from each other and the crosstalk caused by their interaction diminishes. In the work that follows, we will neglect the group delay term in (1) thereby assuming a worst-case scenario. . With these assumptions, equation (1) simplifies to

$$\frac{\partial A_j}{\partial z} + \frac{i}{2}\beta_{2j}\frac{\partial^2 A_j}{\partial t^2} = GA_j \,, \tag{2}$$

where

$$G = -\frac{\alpha_j}{2} + \sum_{n\neq j}(-\frac{g_{nj}}{2} + 2\gamma_n i)\left|A_n\right|^2 \,. \tag{3}$$

Neglecting SPM has allowed us to convert the nonlinear partial-differential equation (pde) in (1) into a variable-coefficient linear pde (2). Equation (2) is still a complex mathematical entity with no trivial solution, nevertheless its solution will be shown to be mathematically tractable.

III. LARGE SPACE-TIME SOLUTION TO THE LINEARIZED SCHRODINGER EQUATION

Equation (2) is a linearized version of the nonlinear Schrodinger equation and will be referred to, from hereon, as the linear Schrodinger equation (LSE). Although it is a linear equation, it admits no trivial solution because the damping function $G(z,t)$ defined in (3) is time-dependent. Nevertheless, the method of multiple scales borrowed from perturbation theory [6] can be used to find an asymptotic solution. Accordingly, we first proceed to solve the unperturbed problem by neglecting the damping function. That is we start by solving the following linear problem (the subscript j is dropped for simplicity)

$$\frac{\partial A}{\partial z} + \frac{i}{2}\beta_2\frac{\partial^2 A}{\partial t^2} = 0 \,. \tag{4}$$

Using Fourier analysis, the general solution of equation (4) is of the form

$$A(z,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(0,\omega) \exp\left(\frac{i}{2}\beta_2\omega^2 z - i\omega t\right) d\omega, \qquad (5)$$

where $\tilde{A}(0,\omega)$ is the Fourier transform of the initial pulse $A(0,t)$ (launched at $z=0$), defined as

$$\tilde{A}(0,\omega) = \int_{-\infty}^{\infty} A(0,t) \exp(i\omega t) dt. \qquad (6)$$

If the input pulse $A(0,t)$ travels a sufficiently large distance and is allowed to disperse, equation (5) will simplify, under the stationary phase approximation, to

$$A(z,t) \approx \sqrt{\frac{2\pi}{k''(\omega_0)z}} \tilde{A}(0,\omega_0) \exp(ikz - i\omega_0 t - i\phi)$$

$$\phi = \mathrm{sgn}(k''(\omega_0))\frac{\pi}{4} \qquad (7)$$

where k satisfies the dispersion relation

$$ik - \frac{i}{2}\beta_2\omega^2 = 0, \qquad (8)$$

$$k''(\omega_0) = \frac{d^2 k}{d\omega^2}\bigg|_{\omega_0}, \text{ and } \omega_0 \text{ is the solution to the following}$$

equation

$$\frac{dk}{d\omega}z - t = 0, \qquad (9)$$

where $d\omega/dk$ is the envelope group velocity. Solving for k in (8) yields

$$k = \frac{\beta_2}{2}\omega^2. \qquad (10)$$

Hence, the first and second derivatives of k are

$$\frac{dk}{d\omega} = \beta_2\omega, \qquad (11)$$

$$\frac{d^2 k}{d\omega^2} = \beta_2. \qquad (12)$$

Substituting the result from (11) in (9) and solving for ω_0, we obtain

$$\omega_0 = \frac{t}{\beta_2 z}. \qquad (13)$$

Finally, using (10), (12), and (13) in (7) and ignoring the constant phase, ϕ, we get

$$A(z,t) \approx \frac{\tilde{A}(0, \frac{t}{\beta_2 z})}{\sqrt{2\pi\beta_2 z}} \exp\left(-i\frac{t^2}{2\beta_2 z}\right). \qquad (14)$$

Equation (14) shows that the initial pulse will decay as the inverse square root of the dispersion-distance product but its shape will remain unchanged.

As a simple validity check of equation (14), consider the case of a Gaussian pulse for which $A(0,t)$ is given as

$$A(0,t) = \exp\left(-\frac{t^2}{2T_0^2}\right), \qquad (15)$$

where T_0 is the pulse width at the $1/e$-intensity point. The Fourier transform of this pulse is given as

$$\tilde{A}(0,\omega) = \sqrt{2\pi}T_0 \exp\left(-\frac{\omega^2 T_0^2}{2}\right). \qquad (16)$$

By substituting (16) in (5) and carrying out the integration, the exact complex envelope at any distance z is found to be

$$A(z,t) = \frac{T_0}{\sqrt{(T_0^2 - i\beta_2 z)}} \exp\left(-\frac{t^2}{2(T_0^2 - i\beta_2 z)}\right). \qquad (17)$$
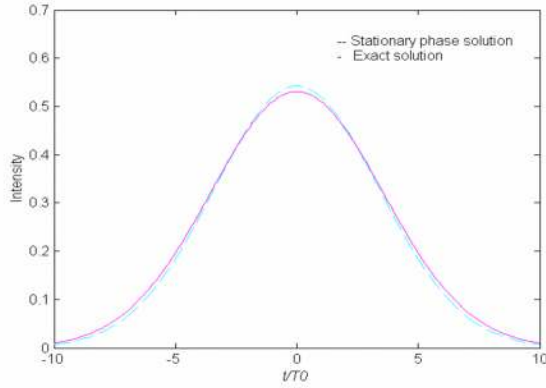
To obtain the stationary phase solution for a Gaussian input pulse, we evaluate (16) at ω_0 given in (13) and substitute the result in (14) to get

$$A(z,t) = \frac{T_0}{\sqrt{\beta_2 z}} \exp\left(-\frac{t^2 T_0^2}{2(\beta_2 z)^2}\right) \exp\left(-i\frac{t^2}{2\beta_2 z}\right). \qquad (18)$$

It can easily be shown that when

$$T_0^2 \ll \beta_2 z, \qquad (19)$$

the exact solution given by (17) will converge to the expression in (18). Figure 2.1 shows magnitude plots of equations (18) and (17) for different fiber lengths assuming $T_0 = 10$ ps and $\beta_2 = 17$ ps^2/Km. It is clear from the plots that when condition (19) is satisfied, as is the case for a
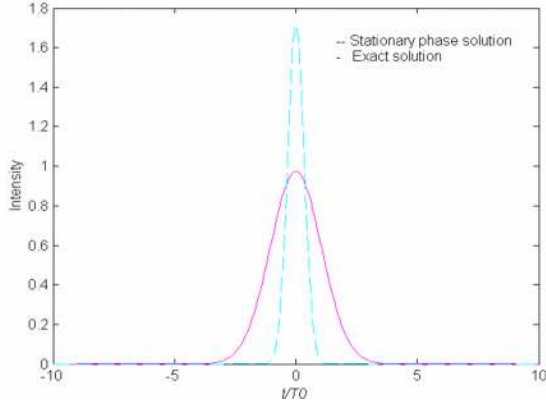
(a)

Fig. 2. Exact and stationary phase solutions of equation (4) when relation (19) is (a) satisfied and (b) (19) violated.

fiber length $z = 20$ Km (Figure 2.b), the stationary phase solution coincides nearly with the exact solution given by (17). For the smaller length $z = 2$ Km, (Figure 2.a), the stationary phase approximation does not match the exact solution as condition (19) no longer applies. Having solved the linear problem (4), we now turn our attention back to the LSE problem defined by equations (2-3). Let us assume that a solution to equation (2) exists and is of the form

$$A(z,t) = R(z,t)e^{i\theta(z,t)} \ , \tag{20}$$

where $R(z,t)$ is an unknown complex function and $\theta(z,t)$ is the phase obtained from the stationary phase solution, i.e.,

$$\theta(z,t) = -\frac{t^2}{2\beta_2 z} \ . \tag{21}$$

Let us denote the first and second partial derivatives of a function F with respect to x as F_x and F_{xx}, respectively. Then,

$$\begin{aligned}
A_z &= \left(i\theta_z R + R_z\right)e^{i\theta} \\
A_{tt} &= \left(-\theta_t^2 R + i(2\theta_t R_t + \theta_{tt}R) + R_{tt}\right)e^{i\theta} ,
\end{aligned} \tag{22}$$

where

$$\begin{aligned}
\theta_t &= -\frac{t}{\beta_2 z}, \\
\theta_{tt} &= -\frac{1}{\beta_2 z}, \\
\theta_z &= \frac{t^2}{2\beta_2 z^2} .
\end{aligned} \tag{23}$$

Substituting (22) in (2), we get

$$i\theta_z R + R_z + i\frac{\beta_2}{2}\left(R_{tt} - \theta_t^2 R\right) - \frac{\beta_2}{2}\left(2\theta_t R_t + \theta_{tt}R\right) = GR . \tag{24}$$

From the look of it, one might think, and rightly so, that equation (24) is mathematically more difficult to deal with than equation (4). Because of the terms θ_t^2 and R_{tt}, equation (24) admits no trivial solution. However, the dispersion relation in (8) implies that

$$\theta_t^2 = \frac{2}{\beta_2}\theta_z . \tag{25}$$

Moreover, after a sufficiently large length z, the input pulse will have dispersed and the term R_{tt} can be neglected without loss of generality. Hence, equation (24) simplifies to

$$R_z - \beta_2\theta_t R_t = R(G + \frac{\beta_2}{2}\theta_{tt}) . \tag{26}$$

We now call on the method of characteristics to solve for R in (26). Recall that R is a function of both space (z) and time (t) and so its spatial derivative, for an observer moving with the envelope, can be expressed as

$$\frac{dR}{dz} = R_z + \frac{\partial t}{\partial z}R_t . \tag{27}$$

By analogy with (27), solving equation (26) can be reduced to solving the following ordinary differential equation (ode)

$$\frac{dR}{dz} = R(G + \frac{\beta_2}{2}\theta_{tt}) ,$$ (28)

along the characteristic equation

$$\frac{dt}{dz} = -\beta_2\theta_t .$$ (29)

Replacing θ_t in (29) with the expression in (23) then integrating yields the following space-time relation

$$t = bz ,$$ (30)

where b is constant along the characteristic path described by (29). The o.d.e. in (27) can now be solved for any space-time point satisfying relation (30). Replacing θ_{tt} in (28) with the expression in (23) then integrating yields

$$R(z,t) = \frac{\tilde{A}(0,\frac{t}{\beta_2 z})}{\sqrt{\beta_2 z}}\exp\left(\int_0^z G(x,bx)dx\right),$$ (31)

where $\tilde{A}(0,\frac{t}{\beta_2 z})$ is the Fourier transform of the initial pulse, evaluated at $\omega = \frac{t}{\beta_2 z}$ and $b = t/z$. Finally, using expressions (31) and (21) in equation (20), we obtain

$$A(z,t) = \frac{\tilde{A}(0,\frac{t}{\beta_2 z})}{\sqrt{\beta_2 z}}\exp\left(\int_0^z G(x,\frac{t}{z}x)dx\right)\exp\left(-i\frac{t^2}{2\beta_2 z}\right)$$ (32)

Equation (32) is a large-space solution to the LSE equation. It shows that as the initial pulse propagates along the fiber, both its shape and phase (recall that G is a complex function) will be altered. Its magnitude will decay as the inverse square root of the dispersion-distance product and will be modulated, along a given characteristic path, by the fiber nonlinearities. It also provides a closed-form mapping between the initial pulse $A(0,t)$ and the output pulse $A(z,t)$ at any sufficiently large distance z. It thus constitutes a large-space (slightly different paths result from removing the large-space condition dictated by the stationary phase assumption) time-domain transfer function for the nonlinear WDM fiber optic channel. This result can be exploited to formulate the theoretical capacity problem of the nonlinear WDM communication system in an information theoretic framework.

IV. CHANNEL MODEL

Consider a WDM system as depicted in Fig. 1. The combined signal travels over a dispersive medium and is amplified before being regenerated at the receiver. Optical amplifiers, like any amplifier, not only amplify signals but they also corrupt them with additive noise. In the case of an erbium-doped amplifier (EDFA), a quasi-essential device in today's WDM systems, the added noise is attributed to a physical process known as amplified spontaneous emission (ASE). ASE noise is commonly modeled as an additive white Gaussian (AWG) stochastic process with a variance

$$\sigma_{ASE}^2 = n_{sp}h\upsilon_c(G-1)B_o ,$$ (33)

where n_{sp} is the spontaneous emission factor, h is Planck's constant, υ_c is the optical frequency, G is the gain, and B_o is the optical bandwidth. Let x_k represent the symbols transmitted over a WDM optical channel and let $h(t,L)$ be the combined transmitter and channel response. Furthermore, let the received signal, $y(t)$, be sampled at $t=nT$ where T is the symbol period. Then

$$y_k = \sum_n h_n x_{k-n} + w_k$$ (34)

where w_k represents the ASE noise generated by the optical amplifier. In the nonlinear regime, the combined sampled channel response given as

$$h_n = \tilde{A}(0,\frac{nT}{\beta_2 L})\rho(L,nT),$$ (35)

where

$$\rho(L,nT) = \frac{\exp\left(\int_0^L G_j(x,\frac{nT}{L}x)dx\right)}{\sqrt{\beta_{2j}L}}$$ (36)

is the nonlinear channel distortion, The term, $\tilde{A}(0,\frac{nT}{\beta_2 z})$

causes ISI and $\rho(L,nT)$, a random multiplicative noise coefficient, contributes to channel fading. A nonlinear WDM channel can thus be modeled as a multi-path, fading, frequency selective channel. This simple yet significant observation can be exploited to derive new bounds on the capacity of the nonlinear WDM optical fiber channel. In the linear regime, the combined response $h(nT, L)$ is given by (14) and the channel in (34) reduces to the classical discrete-time intersymbol interference channel.

V. CAPACITY BOUNDS

In this work, we limit our discussion to estimating the capacity of the linear optical channel defined by (14) and

(34). The capacity of the nonlinear channel defined by (34) and (35) will be presented in a future publication..

The channel capacity is defined as

$$C(Y;X) = \max_{p(A)} I(Y;X) , \qquad (37)$$

where $I(Y;X)$ is the mutual information between Y and X, defined as

$$I(Y;X) = E\left\{ \log\left(\frac{p(Y/X)}{E\{p(Y/X)\}} \right) \right\} . \qquad (38)$$

The maximization in (37) is carried out over the probability densities of the transmitted signal x_j subject to the power constraint

$$\sum_j \left| x_j \right|^2 = P , \qquad (39)$$

where P is the total available power.

In the linear regime where the classical discrete-time intersymbol interference channel model holds, the channel capacity is given as [6]

$$C_G = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left(1 + \frac{P\left|R_h(e^{-j\theta})\right|^2}{\sigma_w^2}\right) d\theta , \qquad (40)$$

where R_h is the discrete Fourier transform of h_n and P the launch power. The channel capacity in (40) is achieved only when the input symbols x_k are i.i.d. Gaussian random variables. In many communications systems this condition may not hold. Practical lower bounds however exist to evaluate the capacity of these systems [7]. One such bound, suitable for on-off keying (OOK) WDM systems is given as

$$C_{LWMF} = C_b\left(\frac{P}{\sigma_w^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left|R_h(e^{-j\theta})\right|^2 d\theta\right) \qquad (41)$$

where C_b is the capacity of the binary-input Gaussian memoryless channel [7]. Equations (40) and (41) set the capacity limits on the strictly dispersive channel (i.e., linear) . In Fig. 3, various bounds, including the standard AWG bound, C_{awg}, are plotted as a function of the channel signal-to-noise ratio (SNR) for a 100 Km SMF fiber link. From the plots, it is clear that rates above 2 bit/symbol are possible at a channel SNR of 10dB. Today, achievable rates for a similar channel SNR are around 0.2 bit/symbol.
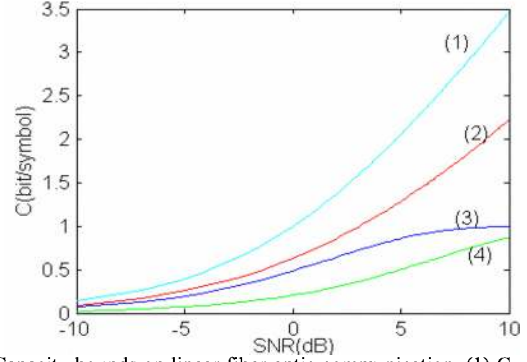

Fig. 3. Capacity bounds on linear fiber optic communication. (1) C_{awg}, (2) C_G, (3) C_b, (4)C_{LWMF}.

VI. CONCLUSION

The treatment of a fiber optic channel as a time-varying frequency selective channel has, to our knowledge, never been proposed. It bridges a long-standing gap between the physics of fiber optics and the theory of communication and opens fiber optic communication to powerful techniques already used in wireless systems. Estimating the capacity of a fiber optics system is a fundamental step to understanding how its performance may be improved. We have used the method of multiple scales to solve the nonlinear Schrodinger equation. The proposed solution is valid only after the initial pulse has traveled a sufficiently large distance and is allowed to disperse. In practice this condition is usually met (as is the case for single mode fibers with a loss coefficient of 0.22 dB/Km and a GVD around 17 ps/Km) after the pulse has propagated tens of kilometers. We then used the linear solution to derive new bounds on the capacity of linear but dispersive optical fiber channel. The capacity of the nonlinear channel based on our solution (35) will be presented in a future publication.

REFERENCES

[1]   P. P. Mitra and J. B. Stark, "Non-linear limits to the information capacity of optical fiber communications", Nature, 411, pp. 1027-1030, June 2001

[2]   E.E. Narimov et al,"The Channel Capacity of a Fiber Optics Communication System: Perturbation Theory" J. Light. Tech vol 20, no. 3, pp.530-537, March 2002.

[3]   J. Tang,'The Shannon channel capacity of dispersion-free nonlinear optical fiber transmission," J. Lightwave Technol., 19:1104–1109, 2001.

[4]   l. Djordjevic, B. Vasic, M. lvkovic, and I. Gabitov. Achievable information rates for high-speed long-haul optical transmission. J. Lightwave Technol., 23:3755–3763, 2005.

[5]   G. P. Agrawal, Nonlinear Fiber Optics, 2nd edition. Academic Press, San Diego, CA,1995.

[6]   R. Haberman, Elementary Applied Partial Differential Equations. Prentice-Hall, 1998.

[7]   S. Shamai et al," The intersymbol interference channel: lower bounds on capacity and channel precoding loss" Inform Theory, vol 42, no. 5, pp. 1388- 1404, Sept. 1996.