

# On the Error Probability Evaluation in Lightwave Systems With Optical Amplification

Enrico Forestieri, *Member, IEEE*, and Marco Secondini

**Abstract**—We review the time domain, frequency domain, and Fourier series Karhunen–Loève series expansion (KLSE) methods for exact BER evaluation in intensity- and phase-modulated direct-detection optically amplified systems. We compare their complexity and computational efficiency, and discuss the most relevant implementation issues. We show that the method based on a Fourier series expansion has the simplest implementation and requires the minimum number of eigenvalues to converge to the exact BER value for various kind of optical filters. For this method, we also introduce an equivalent form of the moment generating function, that avoids the singularity for eigenvalues equal to zero, and derive an alternative expansion where signal and noise are expanded on the same orthonormal basis.

**Index Terms**—Amplifier noise, direct detection, error probability evaluation, Karhunen–Loève expansion.

## I. INTRODUCTION

SINCE the advent of optical amplifiers in lightwave communications, the evaluation of the bit error ratio (BER)<sup>1</sup> in optically preamplified direct-detection systems has received much attention [1]–[17]. Due to the presence of amplified spontaneous emission (ASE) noise and of a square law detector, the distribution of the detected signal is non-Gaussian. Although the use of a Gaussian approximation has been shown to give fairly accurate results in OOK systems with optimum detection threshold [6], [17], the problem in general has to be modeled by using exact detection statistics rather than Gaussian approximations ( $Q$  factor approach). In particular, the exact BER can be evaluated from the moment generating function (MGF) of the output sample, which is obtained through a Karhunen–Loève series expansion (KLSE) of the noise. To this aim, different approaches have been reported in the literature and are commonly deployed. Though, to our knowledge a comparison among all those methods, taking into accounts the implementation issues and their computational efficiency, has not been done. Moreover, several works are still being published in which the authors resort to approximated approaches for BER evaluation, showing that the subject of exact BER evaluation has not yet been entirely

assimilated by the scientific community, and that some concerns on complexity and computational cost still exist.

In this paper, we review and compare the most known methods for exact BER evaluation, showing the differences in terms of implementation and efficiency. Among the many methods devised to this aim (see, for example, [8], [13]–[16], [18], and [19], to cite a few of them), we consider only those reported in [18] (time domain expansion), [8] (frequency domain expansion), and [15] (Fourier series expansion), as all others either are based upon them, or are their extensions, or are limited to particular combinations of optical and electrical filters. We show that the Fourier expansion method [15] is the most efficient in almost any practical case in terms of the minimum number of eigenvalues that are needed to obtain the exact BER value. For this method, we give an equivalent form of the MGF, that avoids the singularity which arises in the expression reported in [15] when one or more eigenvalues are equal to zero, and propose a simple formula to automatically determine the number of eigenvalues required for the convergence. We also derive an alternative expansion, where signal and noise are expanded on the same orthonormal basis, and discuss its advantages and disadvantages compared to the expansion on different bases that is adopted in [15].

The paper is organized as follows. Section II introduces the general formulation of the KLSE method, relating the exact BER evaluation to the evaluation of the MGF of the output sample. Sections III, IV, and V describe the time domain, frequency domain, and Fourier series expansions, respectively. Section VI compares the computational cost of the three methods for different optical filters. Finally, Section VII investigates the extension to DPSK systems.

## II. THE KLSE METHOD

The BER at the output of the decision gate can be evaluated by determining the probability density function (pdf) associated with each detected bit of a pseudorandom bit sequence of length  $2^n - 1$ , augmented by a 0 in the longest run of 0's such as to obtain a de Bruijn sequence  $\{a_k\}_{k=1}^N$  of length  $N = 2^n$  [20]. This sequence is such that all  $n$ -bit patterns occur exactly once in a single period and thus it is possible to accurately account for the intersymbol interference (ISI) due to  $n - 1$  bits. In the following, we describe the method for binary OOK detection, and only briefly outline the extension to DPSK as accounting for phase-modulated formats is quite straightforward, as explained in [21] and [22].

Due to the ISI and the nonstationarity of the signal-ASE beat noise term, the pdf of the sample  $y(t_k)$  is different for each sampling time  $t_k = t_s + kT_b$ ,  $T_b$  being the inverse of the bit

Manuscript received January 17, 2008; revised May 5, 2008. Current version published April 17, 2009. This work was supported in part by Ericsson and by the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) under the PRIN project STORiCo.

The authors are with Scuola Superiore Sant'Anna, Centro di Eccellenza per l'Ingegneria, dell'Informazione e della Comunicazione, I-56124 Pisa, Italy (e-mail: forestieri@sssup.it; marco.secondini@sssup.it).

Digital Object Identifier 10.1109/JLT.2008.927770

<sup>1</sup>As commonly done, we use the term BER instead of the more correct bit error probability (BEP).

rate  $R_b$ , and  $t_s$  a proper time offset. The average BER can be written as

$$\text{BER} = \frac{1}{N} \sum_{k=1}^N P(\mathcal{E}_k) \quad (1)$$

where  $\mathcal{E}_k$  is the error event associated with the  $k$ th bit  $a_k$ , whose probability, given the decision threshold  $y_{\text{th}}$ , can be evaluated through the pdf  $p_{y_k}(y)$  of the sample  $y_k \triangleq y(t_k)$  as

$$P(\mathcal{E}_k) = \begin{cases} P\{y_k < y_{\text{th}}\} = \int_{-\infty}^{y_{\text{th}}} p_{y_k}(y) dy & \text{if } a_k = 1 \\ P\{y_k > y_{\text{th}}\} = \int_{y_{\text{th}}}^{\infty} p_{y_k}(y) dy & \text{if } a_k = 0. \end{cases} \quad (2)$$

The minimum average BER is obtained by jointly minimizing (2) over  $y_{\text{th}}$  and  $t_s$ . However, most often the time offset  $t_s$  can be estimated more cheaply by choosing it such that the sampling times occur at the maximum eye opening. It can be shown that the probabilities that appear in (2) may be approximated with very high accuracy by the so-called *saddlepoint approximation* [23]

$$P(\mathcal{E}_k) \simeq \begin{cases} \frac{\exp[\Phi_{y_k}(s_0^-)]}{\sqrt{2\pi\Phi''_{y_k}(s_0^-)}} & \text{if } a_k = 1 \\ \frac{\exp[\Phi_{y_k}(s_0^+)]}{\sqrt{2\pi\Phi''_{y_k}(s_0^+)}} & \text{if } a_k = 0 \end{cases} \quad (3)$$

where  $\Phi''_{y_k}$  denotes the second-order derivative of the phase function  $\Phi_{y_k}(s)$  defined as

$$\Phi_{y_k}(s) \triangleq \ln \left[ \Psi_{y_k}(s) \frac{\exp(-y_{\text{th}}s)}{|s|} \right], \quad s \in \mathbb{R} \quad (4)$$

$\Psi_{y_k}(s) = E\{\exp(y_k s)\}$  being the MGF of  $y_k$ . The values  $s_0^+$  and  $s_0^-$  are the positive and negative, respectively, saddle points on the real  $s$  axis of  $\exp[\Phi_{y_k}(s)]$ , and may be evaluated as the roots of the equation  $\Phi'_{y_k}(s) = 0$ . So, to evaluate the BER<sup>2</sup>, one simply needs to know the MGF of the sample  $y_k$  at the output of the decision gate. This MGF can be analytically evaluated only for a few combinations of optical and electrical filters, such as Lorentzian shaped optical filter and integrate-and-dump electrical filter [14], [24]. Hence, all the works dealing with exact BER evaluation should really be regarded as methods for evaluating the MGF of the output sample.

In the following, we denote by

$$e(t) = x(t) + w(t) \quad (5)$$

the input field *before* the optical filter  $H_o(f)$  at the receiver end,  $x(t)$  being the signal and  $w(t)$  the ASE noise in the same polarization of the signal [modeled as additive white Gaussian noise (AWGN)]. Fourier transformed pairs will be denoted by lower and upper case letters, such that, as an example, the impulse response of the optical filter  $H_o(f)$  is  $h_o(t)$ . The MGF of the random variable  $z$  is defined as

$$\Psi_z(s) = E\{e^{zs}\} \quad (6)$$

where  $E\{\cdot\}$  denotes the statistical expectation operator.

<sup>2</sup>Note that the saddlepoint approximation could be used to compute the pdfs  $p_{y_k}(y)$ , and then the BER could be obtained by their numerical integration. However, it is more effective to directly evaluate the BER.

### III. EXPANSION IN THE TIME DOMAIN

We start describing the first method (to our knowledge) reported in the literature, which is based on a KLSE in the time domain [18]. In so doing, we extend this method to the case in which the electrical filter impulse response  $h_e(t)$  may also assume negative values, as in the literature it is described for nonnegative impulse responses only [11], [16], [18], [19]. Denoting by  $e_o(t) = s(t) + n(t)$  the receiver input field *after* the optical filter  $H_o(f)$ , where  $s(t) = x(t) \otimes h_o(t)$  and  $n(t) = w(t) \otimes h_o(t)$ , the electrical signal  $y(t)$  after the postdetection filter  $H_e(f)$  can be written as

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} |s(t-\tau) + n(t-\tau)|^2 h_e(\tau) d\tau \\ &\equiv \int_{-\infty}^{\infty} |s(t-\tau) + n(\tau)|^2 h_e(\tau) d\tau \end{aligned} \quad (7)$$

where the last equation should be interpreted in the sense that both sides are statistically equivalent, due to the stationarity of the filtered ASE  $n(t)$ . Expanding both signal and noise onto the functions  $\{\varphi_n(t)\}_{n=1}^{\infty}$  orthonormal with respect to the weight function  $h_e(t)$ , i.e., such that<sup>3</sup>

$$\int_{-\infty}^{\infty} \varphi_i(t) \varphi_j^*(t) h_e(t) dt = \gamma_i \delta_{i,j}, \quad \gamma_i = \pm 1 \quad (8)$$

yields

$$s(t-\tau) + n(\tau) = \sum_{k=1}^{\infty} [s_k(t) + n_k] \varphi_k(\tau) \quad (9)$$

where

$$s_k(t) = \gamma_k \int_{-\infty}^{\infty} s(t-\tau) \varphi_k^*(\tau) h_e(\tau) d\tau \quad (10)$$

$$n_k = \gamma_k \int_{-\infty}^{\infty} n(\tau) \varphi_k^*(\tau) h_e(\tau) d\tau. \quad (11)$$

Inserting (9) in (7) and taking into account the orthonormality condition (8), we obtain

$$y(t) = \sum_{k=1}^{\infty} \gamma_k |s_k(t) + n_k|^2. \quad (12)$$

As  $n(t)$  is Gaussian and zero mean, the coefficients  $n_k$  are complex Gaussian random variables and zero mean, too. Requiring that they are statistically uncorrelated (and, thus, independent, being they Gaussian), (11) implies that  $\{\varphi_n(t)\}$  should be such that

$$\begin{aligned} E\{n_i n_j^*\} &= N_0 \gamma_i \gamma_j \int \int_{-\infty}^{\infty} R(\tau_1 - \tau_2) \varphi_i(\tau_1) h_e(\tau_1) \\ &\quad \times \varphi_j^*(\tau_2) h_e(\tau_2) d\tau_1 d\tau_2 = 0 \\ &\quad \forall i \neq j \end{aligned} \quad (13)$$

where we accounted for the fact that  $h_e(t)$  is real and

$$R(\tau) = \frac{1}{N_0} E\{n(t+\tau) n^*(t)\} = \mathcal{F}^{-1}\{|H_o(f)|^2\} \quad (14)$$

<sup>3</sup>The orthonormality condition should be regarded here in a broader sense, as if  $h_e(t)$  is not always positive, we will see that for some  $i = j$  the integral (8) may be equal to  $-1$ .

is the autocorrelation function of  $n(t)$  normalized to the ASE spectral density per polarization mode  $N_0$ . Taking into account (8), it is apparent that (13) is satisfied if

$$\int_{-\infty}^{\infty} R(\tau_1 - \tau_2) \varphi_i(\tau_1) h_e(\tau_1) d\tau_1 = \lambda_i \varphi_i(\tau_2). \quad (15)$$

Notice that, when (15) holds, for  $i = j$ , (13) becomes

$$\begin{aligned} E\{|n_i|^2\} &= \lambda_i N_0 \int_{-\infty}^{\infty} |\varphi_i(t)|^2 h_e(t) dt \\ &= \gamma_i \lambda_i N_0 \end{aligned} \quad (16)$$

so it is clear that it should be  $\gamma_i = \text{sign}(\lambda_i)$ , and (9) is a Karhunen-Loève expansion, as the coefficients  $n_k$  in (11) are statistically uncorrelated and circularly symmetric complex random variables with zero mean and variance equal to  $|\lambda_k| N_0$ . Thus,  $y(t)$  in (12) is a sum/difference of noncentral chi-square independent random variables and its MGF is given by the product of the individual MGFs

$$\begin{aligned} \Psi_y(s) &= \prod_{k=1}^{\infty} \frac{1}{(1 - \lambda_k N_0 s)^P} \\ &\quad \times \exp\left(\frac{\text{sign}(\lambda_k) |s_k(t)|^2 s}{1 - \lambda_k N_0 s}\right) \end{aligned} \quad (17)$$

where  $P = 1$ . Had we accounted for the ASE in the orthogonal polarization, we would have obtained for the signal  $y(t)$

$$y(t) = \sum_{k=1}^{\infty} \text{sign}(\lambda_k) (|s_k(t) + n_k|^2 + |z_k|^2) \quad (18)$$

where the  $z_k$ s are distributed as the  $n_k$ s and independent of them, such that the MGF would have been as in (17) but now with  $P = 2$ . Had we considered the signal and noise components onto two general orthogonal polarizations, we would have obtained for  $y(t)$

$$\begin{aligned} y(t) &= \sum_{k=1}^{\infty} \text{sign}(\lambda_k) (|s_{k,1}(t) + n_{k,1}|^2 \\ &\quad + |s_{k,2}(t) + n_{k,2}|^2) \end{aligned} \quad (19)$$

where  $s_{k,p}(t)$  and  $n_{k,p}$  are relative to polarization  $p$ , such that the MGF would have been

$$\begin{aligned} \Psi_y(s) &= \prod_{k=1}^{\infty} \frac{1}{(1 - \lambda_k N_0 s)^2} \\ &\quad \times \exp\left(\frac{\text{sign}(\lambda_k) (|s_{k,1}(t)|^2 + |s_{k,2}(t)|^2) s}{1 - \lambda_k N_0 s}\right). \end{aligned} \quad (20)$$

Notice that, in the absence of PMD, the fiber Jones matrix is frequency independent and hence  $|s_{k,1}(t)|^2 + |s_{k,2}(t)|^2 = |s_k(t)|^2$ , whatever the two reference orthogonal polarizations. So, only in the absence of PMD, (17) and (20) would be equal.

#### A. Evaluating $\lambda_k$ and $s_k(t)$

Performing the integration in (15) over a limited time interval  $(T_1, T_2)$ , chosen such that outside it the impulse response  $h_e(t)$  of the electrical filter is negligible

$$\int_{-T_1}^{T_2} R(\tau_1 - \tau_2) \varphi_i(\tau_1) h_e(\tau_1) d\tau_1 \simeq \lambda_i \varphi_i(\tau_2) \quad (21)$$

and following the Nyström method [25], [26], we can apply a Gauss-Legendre quadrature rule [27] with appropriate weights  $w_k$  and nodes  $t_k$

$$\int_a^b g(t) dt \simeq \sum_{k=1}^{M_e} w_k g(t_k) \quad (22)$$

to evaluate (21) as

$$\sum_{k=1}^{M_e} w_k R(t_k - \tau_2) \varphi_i(t_k) h_e(t_k) \simeq \lambda_i \varphi_i(\tau_2). \quad (23)$$

Now, letting  $\tau_2$  assume the same values as the nodes  $t_n$ , we obtain the following  $M_e$  equations:

$$\begin{aligned} \sum_{k=1}^{M_e} w_k R(t_k - t_n) \varphi_i(t_k) h_e(t_k) &\simeq \lambda_i \varphi_i(t_n), \\ n &= 1, \dots, M_e. \end{aligned} \quad (24)$$

These equations can be written in matrix notation as

$$\mathcal{R}W\Phi = \Phi\Lambda \quad (25)$$

where  $\mathcal{R}$ ,  $W$ ,  $\Phi$ , and  $\Lambda$  are matrices whose elements are, respectively

$$\begin{aligned} r_{i,j} &= R(t_j - t_i) \\ w_{i,j} &= w_i h_e(t_i) \delta_{i,j} \\ \varphi_{i,j} &= \varphi_j(t_i) \\ \lambda_{i,j} &= \lambda_i \delta_{i,j} \end{aligned} \quad i, j = 1, \dots, M_e. \quad (26)$$

As evident from (25), the columns of  $\Phi$  are the eigenvectors of the matrix  $\mathcal{R}W$ , and the diagonal elements of  $\Lambda$  are its eigenvalues. The elements of the  $n$ th column of the matrix  $\Phi$  are the samples of the eigenfunction  $\varphi_n(t)$  at the times  $t_k$ ,  $k = 1, \dots, M_e$ . Notice that the matrix  $\mathcal{R}W$  is real but not symmetric, and  $W$  is not positive definite if  $h_e(t_k)$  is negative or zero for some  $t_k$ . However, as the correlation matrix  $\mathcal{R}$  is always symmetric and positive definite, we can factorize it by a Cholesky decomposition as <sup>4</sup>

$$\mathcal{R} = LL^T \quad (27)$$

such that multiplying (25) on the left by  $L^{-1}$  it can be written as

$$AB = B\Lambda \quad (28)$$

where  $A = L^T W L$  is symmetric and  $B = L^{-1} \Phi$ . So, instead of (25), we could solve (28) and then evaluate  $\Phi$  as  $\Phi = LB$ . Either methods should be taken into account, as one of them can succeed where the other fails. Indeed, finding eigenvalues and eigenvectors of a real nonsymmetric matrix may be numerically problematic, whereas a symmetric matrix gives no problems. But, to obtain the symmetric matrix  $A$  in (28), one needs the Cholesky decomposition (27) which, because of roundoff errors, in some cases might not be evaluated. We observed that the

<sup>4</sup>The Cholesky decomposition is straightforward to obtain, but the procedure may fail due to roundoff errors which may make the matrix nonpositive definite, even if theoretically it is so.

two methods never fail both at the same time, so a good strategy is trying to solve (25) and, if it fails, switching to (28). Note that a numerical failure with both methods could also be circumvented by slightly perturbing some parameter, but we find the mentioned procedure more effective.

Anyway, when solving either the eigenvalue problem (25) or (28) using standard packages, the returned eigenvectors are normalized to unit norm, such that the elements of the  $n$ th column of matrix  $\Phi$  are not equal but instead proportional to the samples of the eigenfunction  $\varphi_n(t)$  at the times  $t_k, k = 1, \dots, M_e$ . The proportionality constant for the  $n$ th eigenfunction can be easily found by equating to  $\gamma_n = \text{sign}(\lambda_n)$  the integral (8), for  $i = j = n$ , performed using the same quadrature rule and using the elements of column  $n$  of  $\Phi$  as the samples  $\varphi_n(t_k)$ .

Once  $\varphi_n(t_k)$  are known for  $k = 1, \dots, M_e$ , one could use (10) to directly evaluate  $s_n(t)$  through the Gaussian quadrature rule. This would require 1 FFT to evaluate  $s(t)$  (the required values  $s(t - t_k)$  should be obtained through interpolation, as the nodes of the quadrature rule are not equispaced), and  $NM_e$  integrals to obtain  $s_n(t)$  for all the  $N$  sampling times and the  $M_e$  eigenfunctions. Alternatively, a more efficient method is to evaluate  $\varphi_n(t)$  and  $h_e(t)$  through interpolation for  $t \neq t_k$ , such that

$$H_n(f) = \gamma_n \int_{-\infty}^{\infty} \varphi_n^*(t) h_e(t) e^{-j2\pi ft} dt \quad (29)$$

can be obtained through FFT, and then  $s_n(t)$  as

$$s_n(t) = \int_{-\infty}^{\infty} S(f) H_n(f) e^{j2\pi ft} df. \quad (30)$$

In this way, only a total of  $2M_e$  FFTs are needed ( $S(f)$  is the Fourier transform of the signal  $s(t)$  at the output of the optical amplifier).

#### IV. EXPANSION IN THE FREQUENCY DOMAIN

This method was proposed in [8] and is based on a KLSE in the frequency domain. The electrical signal  $y(t)$  after the post-detection filter  $H_e(f)$  can be written as

$$\begin{aligned} y(t) &= |e(t) \otimes h_o(t)|^2 \otimes h_e(t) \\ &= \mathcal{F}^{-1} \{ [E(f)H_o(f) \otimes E^*(-f)H_o^*(-f)] H_e(f) \} \\ &= \int \int_{-\infty}^{\infty} E(f_1) K(f_1, f_2) E^*(f_2) e^{j2\pi(f_1 - f_2)t} df_1 df_2 \end{aligned} \quad (31)$$

where

$$K(f_1, f_2) = H_o(f_1) H_e(f_1 - f_2) H_o^*(f_2). \quad (32)$$

As known, the eigenfunctions  $\{\phi_n(f)\}_{n=1}^{\infty}$  of the Fredholm integral equation

$$\int_{-\infty}^{\infty} K(f_1, f_2) \phi(f_1) df_1 = \lambda \phi(f_2) \quad (33)$$

form a complete set of orthonormal basis functions, and the corresponding eigenvalues  $\lambda_n$  are real and ordered as  $\lambda_1 \geq \lambda_2 \geq$

$\lambda_3 \geq \dots$ . Due to the completeness of  $\{\phi_n(f)\}_{n=1}^{\infty}$ , we can use the following expansion:

$$E(f) e^{j2\pi ft} = \sum_{n=1}^{\infty} e_n(t) \phi_n(f) \quad (34)$$

where the expansion coefficients  $e_n(t) = x_n(t) + w_n(t)$  are given by

$$e_n(t) = \int_{-\infty}^{\infty} E(f) \phi_n^*(f) e^{j2\pi ft} df. \quad (35)$$

Inserting (34) in (31) and taking into account (33), we obtain

$$\begin{aligned} y(t) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} e_n(t) e_m^*(t) \\ &\quad \times \int \int_{-\infty}^{\infty} \phi_n(f_1) K(f_1, f_2) \phi_m^*(f_2) df_1 df_2 \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} e_n(t) e_m^*(t) \lambda_n \int_{-\infty}^{\infty} \phi_n(f_2) \phi_m^*(f_2) df_2 \\ &= \sum_{n=1}^{\infty} \lambda_n |e_n(t)|^2 \\ &= \sum_{n=1}^{\infty} \lambda_n |x_n(t) + w_n(t)|^2. \end{aligned} \quad (36)$$

As  $w(t)$  is AWGN, the coefficients  $w_n(t)$  are statistically independent and circularly symmetric complex random variables with zero mean and variance equal to the ASE spectral density per polarization mode  $N_0$ . Thus,  $y(t)$  in (36) is a weighted sum of noncentral chi-square independent random variables and its MGF is given by the product of the individual MGFs

$$\Psi_y(s) = \prod_{n=1}^{\infty} \frac{1}{(1 - \lambda_n N_0 s)^P} \exp\left(\frac{\lambda_n |x_n(t)|^2 s}{1 - \lambda_n N_0 s}\right), \quad (37)$$

where  $P = 1$ . Had we accounted for the ASE in the orthogonal polarization, we would have obtained for the signal  $y(t)$  the expression

$$y(t) = \sum_{n=1}^{\infty} \lambda_n (|x_n(t) + w_n(t)|^2 + |\omega_n(t)|^2) \quad (38)$$

where  $\omega_n(t)$  are distributed as  $w_n(t)$  and independent of them, such that the MGF would have been as in (37) but now with  $P = 2$ . Had we considered the signal and noise components onto two general orthogonal polarizations, we would have obtained for  $y(t)$  the expression

$$\begin{aligned} y(t) &= \sum_{n=1}^{\infty} \lambda_n (|x_{n,1}(t) + w_{n,1}(t)|^2 \\ &\quad + |x_{n,2}(t) + w_{n,2}(t)|^2) \end{aligned} \quad (39)$$

where  $x_{n,p}(t)$  and  $w_{n,p}$  are relative to polarization  $p$ , such that the MGF would have been

$$\begin{aligned} \Psi_y(s) &= \prod_{n=1}^{\infty} \frac{1}{(1 - \lambda_n N_0 s)^2} \\ &\quad \times \exp\left(\frac{\lambda_n [|x_{n,1}(t)|^2 + |x_{n,2}(t)|^2] s}{1 - \lambda_n N_0 s}\right). \end{aligned} \quad (40)$$

In the absence of PMD,  $|x_{n,1}(t)|^2 + |x_{n,2}(t)|^2 = |x_n(t)|^2$ , whatever the two reference orthogonal polarizations, and (37) and (40) would be equal.

#### A. Evaluating $\lambda_n$ and $x_n(t)$

Performing the integration in (33) over a limited frequency interval  $(-F, F)$ , chosen such that outside it the optical filter frequency response  $H_o(f)$  is negligible [and, anyway, not larger than the bandwidth of  $x(t)$ , as suggested by (35)]

$$\int_{-F}^F K(f_1, f_2) \phi_n(f_1) df_1 \simeq \lambda_n \phi_n(f_2) \quad (41)$$

following again the Nyström method [25], [26], we can apply a Gauss-Legendre quadrature rule [27] with appropriate weights  $w_m$  and nodes  $\nu_m$ , to evaluate (41) as

$$\sum_{m=1}^{M_e} w_m K(\nu_m, f_2) \phi_n(\nu_m) \simeq \lambda_n \phi_n(f_2) \quad (42)$$

and letting  $f_2$  assume the same values as the nodes  $\nu_m$ , we have the following  $M_e$  equations:

$$\sum_{m=1}^{M_e} w_m K(\nu_m, \nu_k) \phi_n(\nu_m) \simeq \lambda_n \phi_n(\nu_k), \quad k = 1, \dots, M_e. \quad (43)$$

These equations can be written in matrix notation as

$$\mathcal{K}W\Phi = \Phi\Lambda \quad (44)$$

where  $\mathcal{K}$ ,  $W$ ,  $\Phi$ , and  $\Lambda$  are matrices whose elements are, respectively

$$\begin{aligned} \kappa_{i,j} &= K(\nu_j, \nu_i) \\ w_{i,j} &= w_i \delta_{i,j} \\ \phi_{i,j} &= \phi_j(\nu_i) \\ \lambda_{i,j} &= \lambda_i \delta_{i,j} \end{aligned} \quad i, j = 1, \dots, M_e. \quad (45)$$

As evident from (44), the columns of  $\Phi$  are the eigenvectors of the matrix  $\mathcal{K}W$ , whereas the diagonal elements of  $\Lambda$  are its eigenvalues. The elements of the  $n$ th column of the matrix  $\Phi$  are the samples of the eigenfunction  $\phi_n(f)$  at the frequencies  $\nu_k$ ,  $k = 1, \dots, M_e$ . Notice that the matrix  $\mathcal{K}W$  is not Hermitian, but, as the weights  $w_k$  are always positive, premultiplying (44) by the real diagonal matrix  $W^{1/2} = \text{diag}\{\sqrt{w_1}, \sqrt{w_2}, \dots, \sqrt{w_{M_e}}\}$ , it becomes

$$AB = BA \quad (46)$$

where now  $A = W^{1/2}\mathcal{K}W^{1/2}$  is Hermitian and  $B = W^{1/2}\Phi$ . So, instead of (44), it is better to solve (46) and then evaluate  $\Phi$  as  $\Phi = W^{-1/2}B$ , where  $W^{-1/2} = \text{diag}\{1/\sqrt{w_1}, 1/\sqrt{w_2}, \dots, 1/\sqrt{w_{M_e}}\}$ . Notice that this time the eigenvectors need not be normalized.

Once the samples  $\phi_n(\nu_k)$  are known for  $k = 1, \dots, M_e$ ,  $\phi_n(f)$ , for  $f \neq \nu_k$ , can be evaluated through interpolation, and then  $x_n(t)$  can be obtained (through FFT) as

$$x_n(t) = \int_{-\infty}^{\infty} X(f) \phi_n^*(f) e^{j2\pi ft} df \quad (47)$$

where  $X(f)$  is the Fourier transform of the signal  $x(t)$  at the input of the optical amplifier.

## V. FOURIER SERIES EXPANSION

We now briefly review the method proposed in [15]. When the signal term  $x(t)$  in (5) corresponds to a periodic binary sequence with period equal to  $N$ ,  $x(t)$  is also periodic with period equal to  $NT_b$ , so it can be expanded in Fourier series as

$$x(t) = \sum_{n=-\infty}^{\infty} x_n e^{j2\pi nt/NT_b}. \quad (48)$$

The noise  $w(t)$  is not periodic, but if the receiver has a finite memory, say  $T_0$ , then the signal  $y(t)$  at time  $t_k$  is only determined by the values of  $e(t)$  in the time interval  $(t_k - T_0, t_k)$ , i.e., an exact description of the noise  $w(t)$  in this time interval is a sufficient statistic. So we may write a Karhunen-Loève expansion for  $w(t)$  in the interval  $t_k - T_0 < t < t_k$  only, and, as  $w(t)$  is AWGN, we can choose any orthonormal basis for the expansion, for example, the Fourier basis

$$\left\{ \varphi_n(t) = \frac{1}{\sqrt{T_0}} \exp[j2\pi n(t - t_k + T_0)/T_0] \right\}_{n=-\infty}^{\infty}. \quad (49)$$

So, the relevant noise process becomes  $w(t) = \sum_n \omega_n \varphi_n(t)$ ,  $t_k - T_0 < t < t_k$ , where  $\omega_n$  are statistically independent and circularly symmetric complex random variables with zero mean and variance equal to the ASE spectral density per polarization mode  $N_0$ . For convenience, we rewrite  $w(t)$  as<sup>5</sup>

$$w(t) = \sum_{n=-\infty}^{\infty} \omega_n e^{j2\pi n(t - t_k + T_0)/T_0} \quad t_k - T_0 < t < t_k \quad (50)$$

where  $w_n$  are as  $\omega_n$  but with variance equal to  $N_0/T_0$ . The series (48) can be limited to the number of harmonics  $L$  carrying at least 99.9% of the signal energy after the optical filter, whereas, as the dimensionality of the space of periodic signals with bandwidth  $B$  and period  $D$  is about  $2BD + 1$ , we limit the series (50) to  $M_e = 2M + 1$  terms ( $-M \leq n \leq M$ ), where  $M = B_0 T_0$ ,

<sup>5</sup>In (50), using  $\exp(j2\pi n(t - t_k + T_0)/T_0)$  instead of simply  $\exp(j2\pi nt/T_0)$  is expedient to have the phase term automatically disappearing at the sampling time  $t_k$ . Equivalently, we could throw away any phase term as  $w_n$  in (50) are statistically equivalent to  $w_n \exp(j\varphi_n)$  for any  $\varphi_n$  (any phase rotation does not alter the statistics of a Gaussian process).

and  $B_0$  is such that the optical filter frequency response is negligible outside the interval  $(-B_0, B_0)$ . Thus, the Fourier transforms of  $x(t)$  and  $w(t)$  are taken approximately equal to

$$X(f) \simeq \sum_{\ell=-L}^L x_\ell \delta\left(f - \frac{\ell}{NT_b}\right) \quad (51)$$

$$W(f) \simeq \sum_{m=-M}^M w_m e^{-j2\pi m t_k / T_0} \delta\left(f - \frac{m}{T_0}\right). \quad (52)$$

Substituting  $E(f) = X(f) + W(f)$  in (31), after some simple algebra we obtain for  $t = t_k$

$$\begin{aligned} y(t_k) = & \sum_{\ell=-M}^M \sum_{m=-M}^M w_\ell^* K\left(\frac{m}{T_0}, \frac{\ell}{T_0}\right) w_m \\ & + \sum_{\ell=-L}^L \sum_{m=-M}^M x_\ell^* e^{-j2\pi \ell t_k / NT_b} K\left(\frac{m}{T_0}, \frac{\ell}{NT_b}\right) w_m \\ & + \sum_{\ell=-L}^L \sum_{m=-M}^M w_\ell^* K\left(\frac{m}{NT_b}, \frac{\ell}{T_0}\right) x_m e^{j2\pi m t_k / NT_b} \\ & + \sum_{\ell=-L}^L \sum_{m=-L}^L x_\ell^* e^{-j2\pi \ell t_k / NT_b} K\left(\frac{m}{NT_b}, \frac{\ell}{NT_b}\right) \\ & \times x_m e^{j2\pi m t_k / NT_b} \end{aligned} \quad (53)$$

or, in compact matrix notation

$$y(t_k) = \mathbf{n}^\dagger \mathbf{A} \mathbf{n} + \mathbf{v}_k^\dagger \mathbf{n} + \mathbf{n}^\dagger \mathbf{v}_k + d_k \quad (54)$$

where  $\dagger$  means transpose and conjugate,  $\mathbf{n}$  is a column vector whose  $M_e = 2M + 1$  components are

$$n_i = w_{i-M-1} \quad i = 1, \dots, M_e \quad (55)$$

$A$  is an Hermitian  $M_e \times M_e$  matrix whose elements are

$$\begin{aligned} a_{i,j} = & K\left(\frac{j-M-1}{T_0}, \frac{i-M-1}{T_0}\right) \\ = & H_o\left(\frac{j-M-1}{T_0}\right) H_e\left(\frac{j-i}{T_0}\right) H_o^*\left(\frac{i-M-1}{T_0}\right) \\ & i, j = 1, \dots, M_e \end{aligned} \quad (56)$$

$\mathbf{v}_k$  is a column vector whose  $M_e$  components are

$$\begin{aligned} v_{i,k} = & \sum_{\ell=-L}^L K\left(\frac{\ell}{NT_b}, \frac{i-M-1}{T_0}\right) x_\ell e^{j2\pi \ell t_k / NT_b} \\ = & H_o^*\left(\frac{i-M-1}{T_0}\right) \sum_{\ell=-L}^L x_\ell H_o\left(\frac{\ell}{NT_b}\right) \\ & \times H_e\left(\frac{\ell}{NT_b} - \frac{i-M-1}{T_0}\right) e^{j2\pi \ell t_k / NT_b} \\ & i = 1, \dots, M_e, \end{aligned} \quad (57)$$

and, finally,  $d_k$  is the signal sample at time  $t_k$  in the absence of noise<sup>6</sup>

$$\begin{aligned} d_k = & \sum_{\ell=-L}^L \sum_{m=-L}^L x_m x_\ell^* H_o\left(\frac{m}{NT_b}\right) H_e\left(\frac{m-\ell}{NT_b}\right) \\ & \times H_o^*\left(\frac{\ell}{NT_b}\right) e^{j2\pi(m-\ell)t_k / NT_b}. \end{aligned} \quad (58)$$

Notice that  $\mathbf{n}$  is a complex Gaussian random vector with zero mean and diagonal covariance matrix

$$E\{\mathbf{n}^* \mathbf{n}^T\} = \sigma^2 I \quad (59)$$

where

$$\sigma^2 = \frac{N_0}{T_0} \quad (60)$$

and  $I$  is the identity matrix.

The matrix  $A$  can be diagonalized by the unitary matrix  $U$  formed by its normalized eigenvectors arranged in the same order of their corresponding eigenvalues  $\lambda_i$

$$A = U \Lambda U^\dagger \quad (61)$$

where

$$\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_{M_e}\}. \quad (62)$$

Letting

$$\mathbf{z} = U^\dagger \mathbf{n} \quad (63)$$

$$\mathbf{b}_k = U^\dagger \mathbf{v}_k \quad (64)$$

and substituting (61) in (54), we have

$$\begin{aligned} y(t_k) = & \mathbf{z}^\dagger \Lambda \mathbf{z} + \mathbf{b}_k^\dagger \mathbf{z} + \mathbf{z}^\dagger \mathbf{b}_k + d_k \\ = & (\mathbf{z} + \Lambda^{-1} \mathbf{b}_k)^\dagger \Lambda (\mathbf{z} + \Lambda^{-1} \mathbf{b}_k) - \mathbf{b}_k^\dagger \Lambda^{-1} \mathbf{b}_k + d_k \\ = & \sum_{i=1}^{M_e} \left\{ \lambda_i \left| z_i + \frac{b_{i,k}}{\lambda_i} \right|^2 - \frac{|b_{i,k}|^2}{\lambda_i} \right\} + d_k \end{aligned} \quad (65)$$

where  $b_{i,k}$  and  $i = 1, \dots, M_e$ , are the components of vector  $\mathbf{b}_k$ . As

$$E\{\mathbf{z}^* \mathbf{z}^T\} = U^T E\{\mathbf{n}^* \mathbf{n}^T\} U^* = U^T \sigma^2 I U^* = \sigma^2 I \quad (66)$$

the  $z_i$ s are still independent and identically distributed Gaussian complex random variables, thus  $y(t_k)$  in (65) is a weighted sum of noncentral chi-square independent random variables plus two constant terms. Thus, its MGF may be written as

$$\Psi_y(s) = e^{d_k s} \prod_{i=1}^{M_e} \frac{1}{(1 - \lambda_i \sigma^2 s)^P} \exp\left(\frac{\sigma^2 |b_{i,k}|^2 s^2}{1 - \lambda_i \sigma^2 s}\right) \quad (67)$$

<sup>6</sup>Note that both (57) and (58) are more efficiently evaluated through FFT.

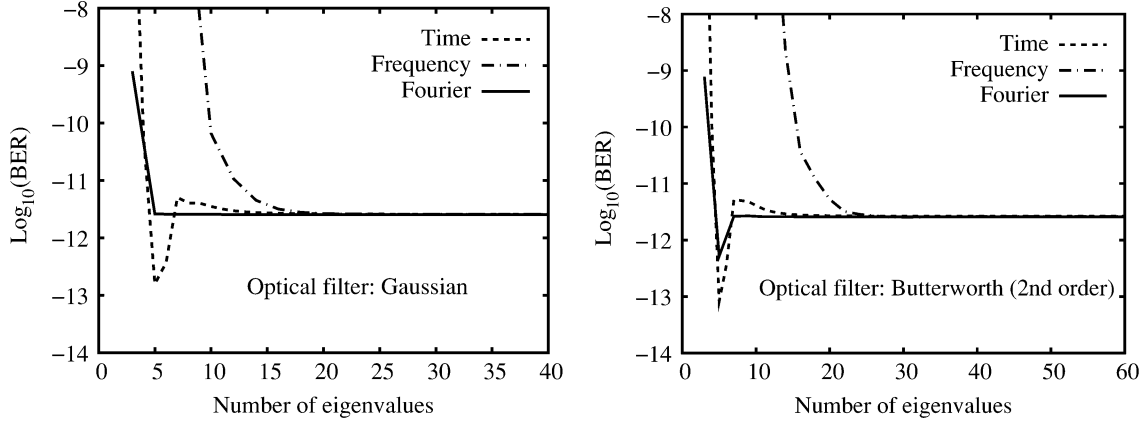


Fig. 1. Convergence of the evaluated BER with the number of eigenvalues: Gaussian and second-order Butterworth optical filters.

where  $\sigma^2 = N_0/T_0$  and  $P = 1$ . When taking into account the ASE noise in the orthogonal polarization, the additional term  $\sum_{i=1}^{M_e} \lambda_i |z'_i|^2$ , where the  $z'_i$ 's are independent random variables distributed as the  $z_i$ 's, should be added to the sample in (65), such that the MGF turns out to be as in (67), but with  $P = 2$ . When considering two general orthogonal polarizations, the MGF is

$$\Psi_y(s) = e^{d_k s} \prod_{i=1}^{M_e} \frac{1}{(1 - \lambda_i \sigma^2 s)^2} \times \exp\left(\frac{\sigma^2 (|b_{i,k}^{(1)}|^2 + |b_{i,k}^{(2)}|^2) s^2}{1 - \lambda_i \sigma^2 s}\right) \quad (68)$$

where  $|b_{i,k}^{(p)}|^2$  is relative to polarization  $p$ . Again, in the absence of PMD,  $|b_{i,k}^{(1)}|^2 + |b_{i,k}^{(2)}|^2 = |b_{i,k}|^2$ , such that (67) and (68) would be equal.

Note that the MGF in (67) is slightly different from that one reported in [15], where the contribution of the constant terms was separated. Although theoretically equivalent, the straight use of (67) avoids the singularity which arises in the expression reported in [15] when one or more eigenvalues are equal to zero.

#### A. Evaluating $\lambda_i$ and $b_{i,k}$

Differently to the previous methods, the quantities necessary to evaluate the MGF are very straightforward to obtain for the Fourier method, as it does not require Gauss-Legendre quadrature and interpolation. The matrix  $A$  depends only upon the optical and electrical filters and its elements are as in (56). Once evaluated eigenvalues and eigenvectors of  $A$ , the  $b_{i,k}$ 's are the components of the vector  $\mathbf{b}_k$  in (64) which is readily evaluated, as the components  $v_{i,k}$  of vector  $\mathbf{v}_k$  can be obtained through FFT from (57), requiring only the knowledge of the Fourier coefficients of the noise-free signal at the input of the optical filter (obtainable through FFT, too). The parameter  $T_0$  depends only upon the optical and electrical filters and may be chosen as

$$T_0 = \mu \left( \frac{1}{B_{o,N}} + \frac{1}{B_{e,N}} \right) \quad (69)$$

where  $B_{o,N}$  and  $B_{e,N}$  are the noise-equivalent bandwidths of the (lowpass equivalent) optical and electrical filters, respectively. Given a pair of filters, one chooses the smallest value of

$\mu$  which guarantees the convergence of the evaluated BER. For almost all practical filters a value of  $\mu = 2$  is sufficient when the parameter  $M_e$  is chosen as above, i.e.,  $M_e = 2B_0T_0 + 1$ , with  $B_0$  equal to the bandwidth in which about 99% of the lowpass equivalent optical filter impulse response energy is contained.

#### B. Alternative Expansion

In the Fourier expansion method described in this section, signal and noise are expanded on *different* orthonormal bases, such that the number of eigenvalues required is the minimum for the convergence of the noise term alone, and not of *both* signal and noise. Indeed, a (perhaps) more elegant formulation could be given to this method by expanding signal and noise on the *same* orthonormal basis, but sacrificing efficiency, as we will now see.

Expanding in Fourier series *both* signal and noise on a time interval  $T_0$  centered on  $t_k$ , (53) becomes

$$\begin{aligned} y(t_k) = & \sum_{\ell=-M}^M \sum_{m=-M}^M w_\ell^* K\left(\frac{m}{T_0}, \frac{\ell}{T_0}\right) w_m \\ & + \sum_{\ell=-M}^M \sum_{m=-M}^M x_\ell^* e^{-j2\pi\ell t_k/T_0} K\left(\frac{m}{T_0}, \frac{\ell}{T_0}\right) w_m \\ & + \sum_{\ell=-M}^M \sum_{m=-M}^M w_\ell^* K\left(\frac{m}{T_0}, \frac{\ell}{T_0}\right) x_m e^{j2\pi m t_k/T_0} \\ & + \sum_{\ell=-M}^M \sum_{m=-M}^M x_\ell^* e^{-j2\pi\ell t_k/T_0} K\left(\frac{m}{T_0}, \frac{\ell}{T_0}\right) \\ & \times x_m e^{j2\pi m t_k/T_0} \end{aligned} \quad (70)$$

where, however, the  $x_k$ 's are different from the ones in (48) as they are now the Fourier coefficients of the expansion of  $x(t)$  on the interval  $(t_k - T_0/2, t_k + T_0/2)$

$$x(t) = \sum_{n=-\infty}^{\infty} x_n e^{j2\pi n t/T_0} \quad t_k - \frac{T_0}{2} < t < t_k + \frac{T_0}{2}. \quad (71)$$

The sample  $y(t_k)$  in (70) can now be written in matrix notation as

$$y(t_k) = (\mathbf{x}_k + \mathbf{n})^\dagger A(\mathbf{x}_k + \mathbf{n}) \quad (72)$$

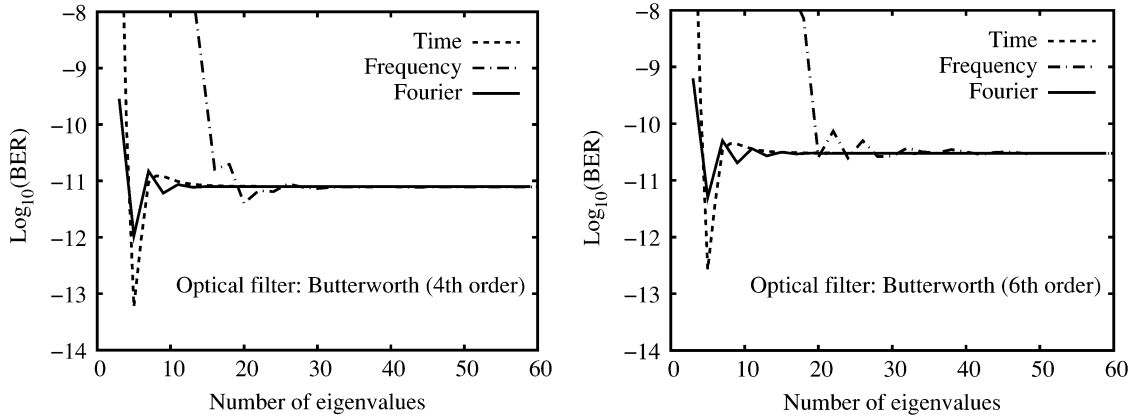


Fig. 2. Convergence of the evaluated BER with the number of eigenvalues: fourth- and sixth-order Butterworth optical filters.

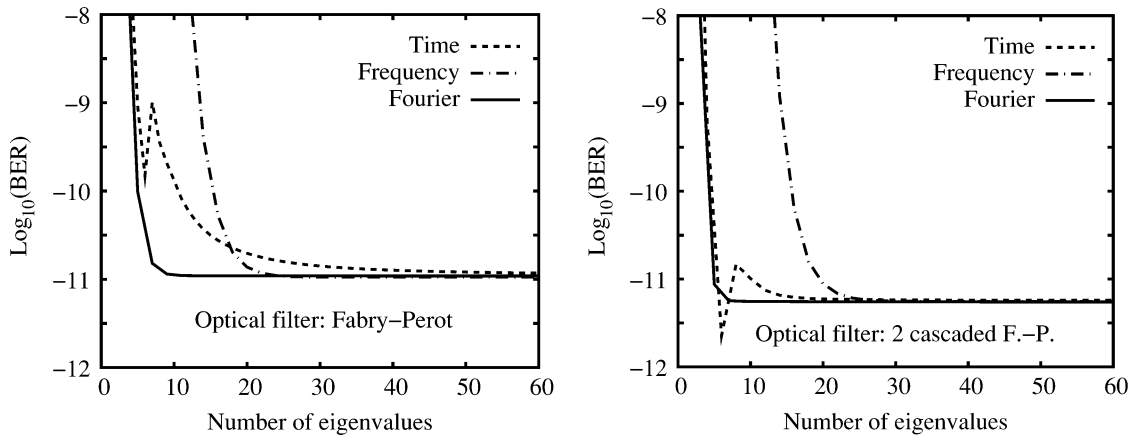


Fig. 3. Convergence of the evaluated BER value with the number of eigenvalues: Single and cascaded Fabry-Perot optical filters.

where  $\mathbf{n}$  and  $A$  are still the column vector and the matrix whose elements are as in (55) and (56), respectively, whereas  $\mathbf{x}_k$  is a column vector whose  $M_e = 2M + 1$  components are

$$\chi_{i,k} = x_{i-M-1} e^{j2\pi(i-M-1)t_k/T_0} \quad i = 1, \dots, M_e. \quad (73)$$

Letting now

$$\mathbf{c}_k = U^\dagger \mathbf{x}_k \quad (74)$$

and using (61) and (63), the sample  $y(t_k)$  in (72) may be written as

$$\begin{aligned} y(t_k) &= (\mathbf{c}_k + \mathbf{z})^\dagger \Lambda (\mathbf{c}_k + \mathbf{z}) \\ &= \sum_{i=1}^{M_e} \lambda_i |c_{i,k} + z_i|^2 \end{aligned} \quad (75)$$

which is formally equal to (36), so the MGF of  $y(t_k)$  is

$$\Psi_y(s) = \prod_{i=1}^{M_e} \frac{1}{(1 - \lambda_i \sigma^2 s)^P} \exp\left(\frac{\lambda_i |c_{i,k}|^2 s}{1 - \lambda_i \sigma^2 s}\right) \quad (76)$$

where, again,  $P = 1$  when neglecting the ASE noise in the orthogonal polarization, or  $P = 2$  when accounting for it.

Notice that this alternative method may be less efficient because  $N$  extra FFTs (one for each bit in the transmitted se-

quence) should be performed for evaluating  $\mathbf{x}_k$ . Moreover, a larger  $M_e$  is required, as the convergence of *both* signal and noise should be guaranteed. This was not the case with the previous method, as the signal was always evaluated with sufficient accuracy and only the convergence of the noise term was required. However, we only need to know the coefficients of the Fourier series expansion of the signal term at the *input* of the optical filter to be able to evaluate the MGF, as the elements of  $\mathbf{v}_k$  in (57) need not be evaluated.

Another alternative expansion could be also derived by expanding both signal and noise on a time interval  $NT_b$ , exactly as done in [28]. In this case, the convergence of the signal term and the  $N$  extra FFTs would not be required, but the size of the eigenvalue problem, that is proportional to the time-width of the expansion interval, would be fixed to a much larger value.

An alternative derivation of the Fourier series method using nonlinear Volterra theory is presented in [29], developing in detail the computational complexity of the method.

## VI. COMPARING THE VARIOUS METHODS

The computational cost of the methods described so far is determined by the length  $N = 2^n$  of the de Bruijn sequence  $\{a_k\}_{k=1}^N$ , and by the size  $M_e$  of the eigenvalue problem. When considering the same values of  $M_e$  and  $N$ , all the methods have approximately the same cost. However, while the ISI length  $n$  is



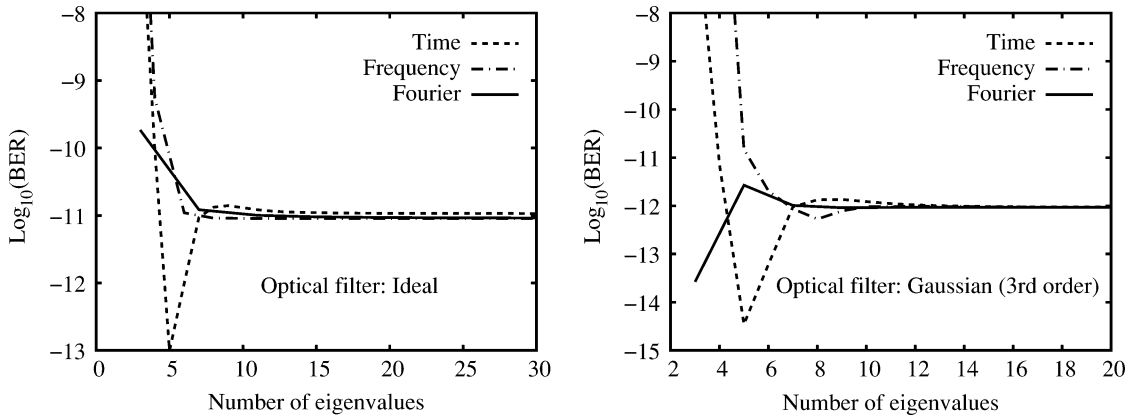


Fig. 4. Convergence of the evaluated BER value with the number of eigenvalues: Ideal rectangular and third-order Gaussian optical filters.

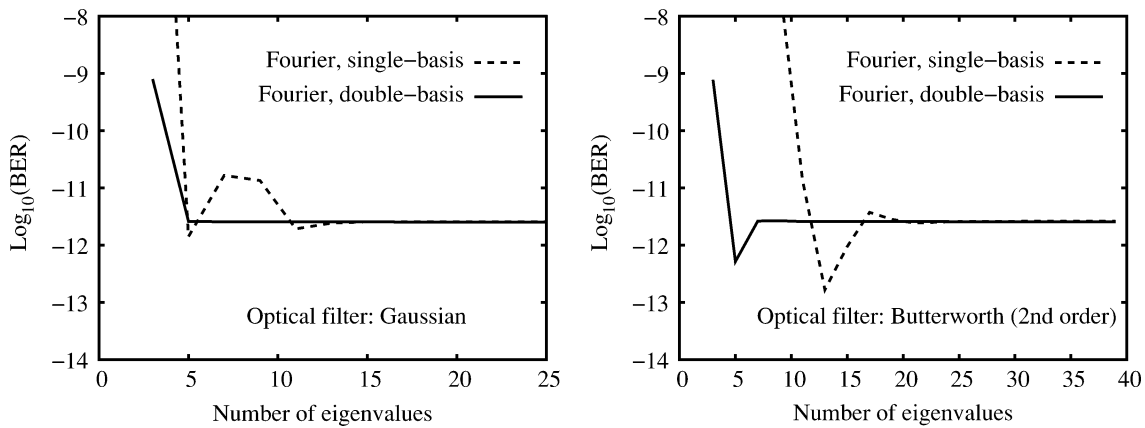


Fig. 5. Comparison between the Fourier methods using either the same basis or two different bases for the expansion of signal and noise: Gaussian and second-order Butterworth optical filters.

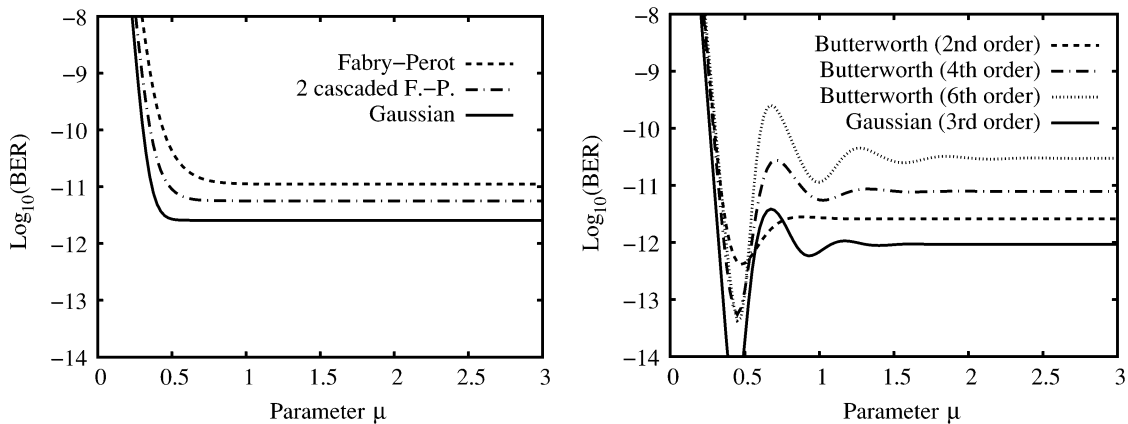


Fig. 6. Double-basis Fourier method: BER versus parameter  $\mu$  for various optical filters with 3-dB bandwidth equal to  $1.5R_b$ .

the same for all the methods, the value of  $M_e$ —corresponding either to the number of nodes of the quadrature rule for the time and frequency expansion, or to the number of coefficients for the Fourier expansion—required to achieve the convergence of the evaluated BER to its exact value may depend on which method is adopted. In the following, we compare the results obtainable using the methods illustrated in the previous sections in terms of the minimum number of eigenvalues  $M_e$  needed to obtain the exact value of the BER.

We take into account an OOK system with NRZ pulses in a back-to-back configuration. The electrical filter in the receiver is a fifth-order Bessel with 3-dB bandwidth equal to  $0.9R_b$ , whereas the optical filter has a 3-dB bandwidth equal to  $1.5R_b$  and is selected among various representative different types. The results obtained with a  $2^5$ -bit de Bruijn sequence for a fixed optical signal-to-noise ratio (OSNR) are summarized in Figs. 1 through 4. Similar results have been obtained when considering different electrical filters.

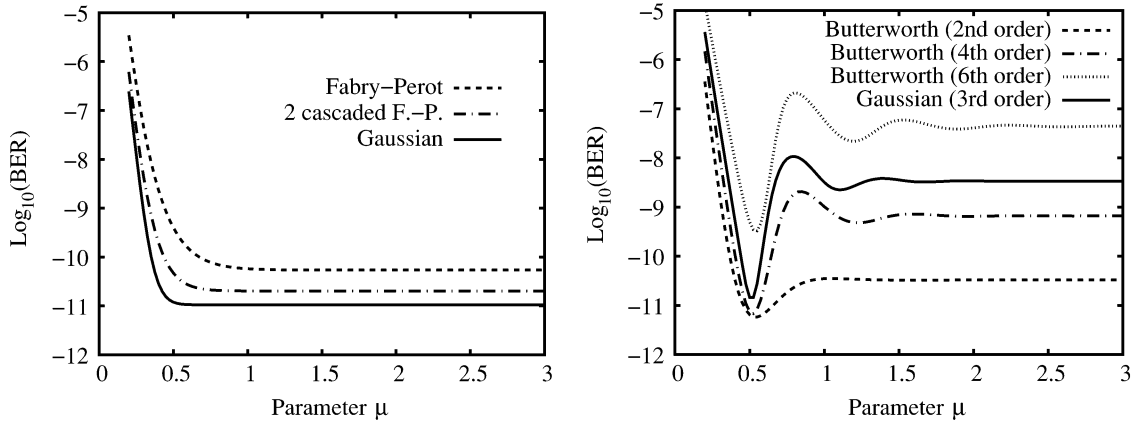


Fig. 7. Double-basis Fourier method: BER versus parameter  $\mu$  for various optical filters with 3-dB bandwidth equal to  $0.9R_b$ .

As can be seen, the number of eigenvalues required by the Fourier method is quite independent of the optical filter shape, as about 11 of them suffice in all cases taken into account. This is about the same number of eigenvalues required by the time domain method, except for the singly resonant type high-finesse Fabry-Perot optical filter, whereas the frequency domain method requires a doubled number of eigenvalues with respect to the other two methods, except when the optical filter has a very sharp rolloff, as shown in Fig. 4 for an ideal rectangular and a third-order Gaussian filters. The reason because the frequency domain method requires a higher number of eigenvalues is to be attributed to the fact that the eigenfunctions [whose bandwidth is of the same order as that of the optical filter, as can be inferred from (33)] should be known over a frequency range equal to the minimum between the signal bandwidth and the optical filter bandwidth, as can be seen from (47), where the optical filter bandwidth appears implicitly through the eigenfunctions. So, the smoother the optical filter rolloff, the larger the frequency range and so the larger the number of eigenfunction samples (i.e., of eigenvalues) to be known.

In order to avoid clutter, we have not reported the results about the alternative method described in Section V-B in Figs. 1–4, but it turns out that the minimum number of eigenvalues is practically the same as that required by the frequency domain method, as shown in Fig. 5 (to be compared with Fig. 1) for Gaussian and second-order Butterworth optical filters, confirming that the single-basis expansion, although more compact and elegant, is less efficient. As anticipated, this is due to the fact that the convergence of both signal and noise has to be guaranteed instead of that of the noise alone.

In general, when changing the bandwidth of the filters, a different number of eigenvalues is required. In the Fourier method, this number is automatically determined once the parameter  $\mu$  in (69), depending on the optical and electrical filters shapes (but not on their bandwidths), is determined. This could be done, once and for all, for a given pair of filters, otherwise one simply choose  $\mu = 2$ , a value which is safe with almost all practical filters, as shown in Figs. 6 and 7, relative to optical filters with 3-dB bandwidths of  $1.5R_b$  and  $0.9R_b$ , respectively.

As can be seen, filters whose impulse response is always positive (on the left in Figs. 6 and 7) monotonically converge to the exact BER as  $\mu$  increases. On the contrary, when the impulse response changes sign (on the right-hand side of Figs. 6 and 7), i.e., when it oscillates, the convergence is oscillatory, too. Filters with oscillatory impulse responses with duration longer

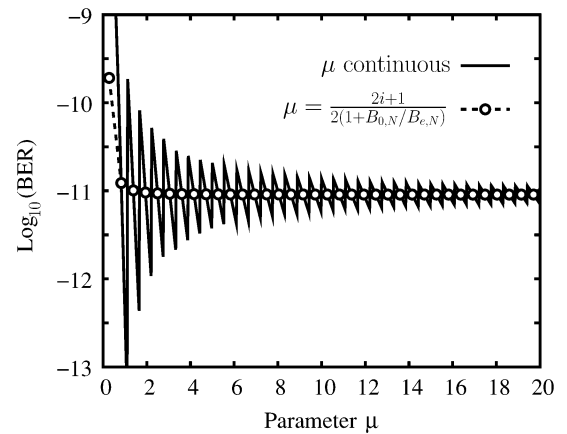


Fig. 8. Double-basis Fourier method: With an ideal filter, convergence is accelerated by using an appropriate sequence of values for  $\mu$ .

than  $2 \div 3$  times the inverse of the filter bandwidth, may require larger values of parameter  $\mu$  to converge to the exact BER, unless, to accelerate the convergence, appropriate values are used. For example, this is the case with an ideal rectangular filter, which deserves a discussion.

The impulse response of an ideal filter of bandwidth  $B$  is proportional to  $\text{sinc}(Bt) = \sin(\pi Bt)/(\pi Bt)$ , so it is very oscillatory in nature and decays very slowly, as its amplitude is always less than  $10^{-2}$  times its maximum value only for  $|t| > 30/B$ , meaning that a value of about  $\mu = 60$  would be required to converge to the exact BER. However, taking into account that the output signal is given by the convolution of the input signal with the filter impulse response, if we choose a sequence of values for  $\mu$  in (69) such that  $T_0$  falls in the middle of a lobe of the impulse response, i.e.,  $T_0 = (2i + 1)/(2B)$ ,  $i = 1, 2, \dots$ , then only the net area subtended by the impulse response between its local minima and maxima would play a role, and we would obtain a fast convergence to the exact BER, which is the value around which the BER obtained by continuously varying  $\mu$  oscillates. This is illustrated in Fig. 8, where the discrete values used for  $\mu$  are  $\mu_i = (2i + 1)/[2(1 + B_{o,N}/B_{e,N})]$ ,  $i = 0, 1, 2, \dots$ , such that  $T_0 = (2i + 1)/(2B_{o,N})$ .

## VII. DIFFERENTIAL PHASE-SHIFT KEYING (DPSK)

Differential direct detection is commonly employed in phase-modulated optical systems, like DPSK and DQPSK. The de-

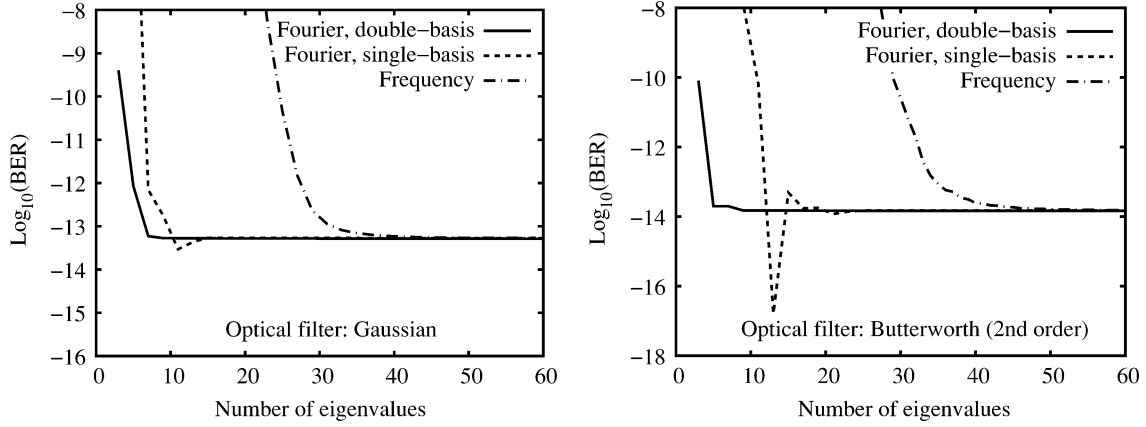


Fig. 9. Convergence of the evaluated BER with the number of eigenvalues: DPSK system, Gaussian, and second-order Butterworth optical filters.

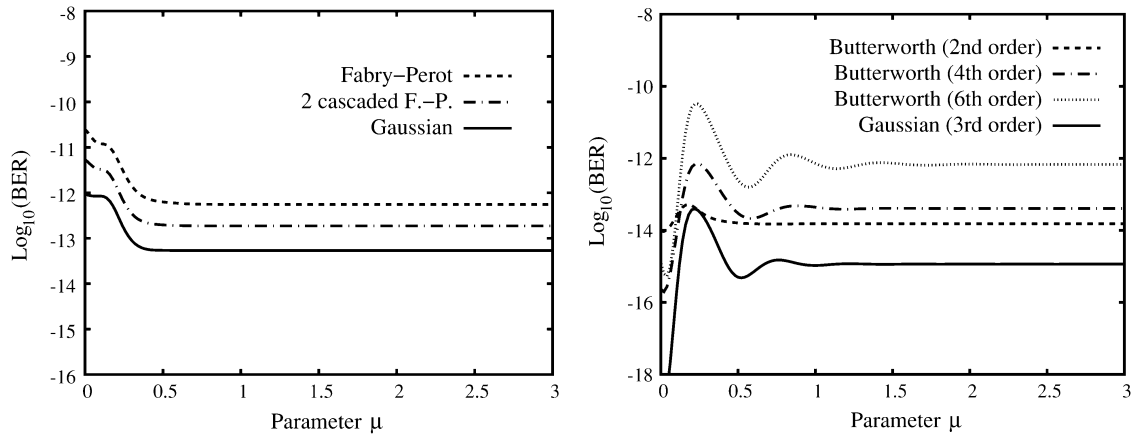


Fig. 10. Double-basis Fourier method: BER versus parameter  $\mu$  for a DPSK system and various optical filters with 3-dB bandwidth equal to  $1.5R_b$ .

tector consists of a Mach-Zehnder interferometer (MZI), with delay  $T_d$  and phase difference  $\Delta\theta$  between its output ports, and a couple of balanced photodetectors. The values of  $T_d$  and  $\Delta\theta$  depend on the modulation format, as well as on detector imperfections. In this case, it is easy to show [21], [22] that the signal  $y(t)$  at the output of the detector is still given by (31), but now with

$$K(f_1, f_2) = H_A(f_1)H_e(f_1 - f_2)H_A^*(f_2) - H_B(f_1)H_e(f_1 - f_2)H_B^*(f_2) \quad (77)$$

where

$$H_A(f) = H_o(f) \frac{\exp[-j(2\pi f T_d + \Delta\theta)] + 1}{2} \quad (78)$$

$$H_B(f) = H_o(f) \frac{\exp[-j(2\pi f T_d + \Delta\theta)] - 1}{2} \quad (79)$$

represent the (equivalent lowpass) transfer functions from the input of the optical filter to the sum (constructive) and difference (destructive) output ports of the MZI, respectively. Therefore, the BER can be evaluated by the same frequency or Fourier expansion methods illustrated in the previous sections, simply replacing (32) with (77). By contrast, the extension of the time domain expansion method to differential detection is not straightforward and will not be discussed in this paper.

In regard to the DPSK systems, the comparison of the various methods in terms of minimum number of eigenvalues required for the convergence to the exact BER value yields similar re-

sults to those presented in Section VI for OOK systems. As an example, Fig. 9 (to be compared with Figs. 1 and 5) shows the convergence of the frequency, double- and single-basis Fourier methods for the Gaussian and second-order Butterworth optical filters. The double-basis Fourier expansion is the most efficient, while the number of eigenvalues required by the frequency expansion method is even higher than in OOK, probably due to the fact that the evaluation of the integral (41) with the oscillating kernel function (77) requires a higher number of nodes  $M_e$  for the Gauss-Legendre quadrature rule. In regard to the convergence of the Fourier expansion for different filter shapes and bandwidths, this can be controlled as in OOK by choosing an appropriate value of the parameter  $\mu$  for the receiver memory that, including the delay  $T_d$  introduced by the MZI, can be expressed as

$$T_0 = T_d + \mu \left( \frac{1}{B_{o,N}} + \frac{1}{B_{e,N}} \right). \quad (80)$$

Fig. 10 shows the convergence of the double-basis Fourier expansion method to the exact BER value as  $\mu$  increases, for the same cases considered in Fig. 6. Again, depending on the filter shapes (but not on their bandwidths), an appropriate value for  $\mu$  can be determined once and for all for a given pair of filters. Otherwise,  $\mu = 2$  is a safe choice for all practical filters. Indeed, we note that the behavior for the OOK and DPSK is very similar, and that the same values of  $\mu$  can be used in both cases.

## VIII. CONCLUSION

We have reviewed the most known KLSE methods for exact BER evaluation in optically amplified direct-detection systems. Among the many methods devised to this aim, we have compared those based on a time domain, frequency domain, or Fourier series expansion. We have shown that the Fourier series expansion has the simplest implementation and requires the smallest number of eigenvalues to converge to the exact BER for various optical filters. We have further improved the Fourier series expansion by introducing an equivalent form of the MGF that avoids the singularity that arises when one or more eigenvalues are equal to zero, and we have proposed a simple formula to automatically determine the number of eigenvalues required for convergence. Finally, we have derived an alternative Fourier series expansion, where signal and noise are expanded on the same orthonormal basis, and discussed the implications of adopting the same basis or two different bases.

## REFERENCES

- [1] P. S. Henry, "Error-rate performance of optical amplifiers," in *Tech. Dig. OFC'89*, Feb. 1989, paper THK3..
- [2] N. A. Olsson, "Lightwave systems with optical amplifiers," *J. Lightw. Technol.*, vol. 7, no. 7, pp. 1071–1082, July 1989.
- [3] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightw. Technol.*, vol. 8, no. 12, pp. 1816–1823, Dec. 1990.
- [4] O. K. Tonguz and L. G. Kazovsky, "Theory of direct-detection lightwave receivers using optical amplifiers," *J. Lightw. Technol.*, vol. 9, no. 2, pp. 174–181, Feb. 1991.
- [5] D. Marcuse, "Calculation of bit-error probability for a lightwave system with optical amplifiers and post-detection Gaussian noise," *J. Lightw. Technol.*, vol. 9, no. 4, pp. 505–513, Apr. 1991.
- [6] P. A. Humblet and M. Azizoglu, "On the bit error rate of lightwave systems with optical amplifiers," *J. Lightw. Technol.*, vol. 9, no. 11, pp. 1576–1582, Nov. 1991.
- [7] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier system," *IEEE Photon. Technol. Lett.*, vol. 5, no. 3, pp. 304–306, Mar. 1993.
- [8] J.-S. Lee and C.-S. Shim, "Bit-error-rate analysis of optically preamplified receivers using an eigenfunction expansion method in optical frequency domain," *J. Lightw. Technol.*, vol. 12, no. 7, pp. 1224–1229, Jul. 1994.
- [9] N. G. Jensen, E. Bodtker, G. Jacobsen, and J. Strandberg, "Performance of preamplified direct detection systems under influence of receiver noise," *IEEE Photon. Technol. Lett.*, vol. 6, no. 12, pp. 1488–1490, Dec. 1994.
- [10] L. F. B. Ribeiro, J. R. F. Da Rocha, and J. L. Pinto, "Performance evaluation of EDFA preamplified receivers taking into account intersymbol interference," *J. Lightw. Technol.*, vol. 13, no. 2, pp. 225–232, Feb. 1995.
- [11] D. Ben-Eli, Y. E. Dallah, and S. Shamai, "Performance bounds and cutoff rates of quantum limited OOK with optical amplification," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 3, pp. 510–530, Apr. 1995.
- [12] S. L. Danielsen, B. Mikkelesen, T. Durhuus, C. Joergensen, and K. E. Stubkjaer, "Detailed noise statistics for an optically preamplified direct detection receiver," *J. Lightw. Technol.*, vol. 13, no. 5, pp. 977–981, May 1995.
- [13] C. Lawetz and J. C. Cartledge, "Performance of optically preamplified receivers with Fabry-Perot optical filters," *J. Lightw. Technol.*, vol. 14, no. 11, pp. 2467–2474, Nov. 1996.
- [14] I. T. Monroy and G. Einarsson, "Bit error evaluation of optically preamplified direct detection receivers with Fabry-Perot optical filters," *J. Lightw. Technol.*, vol. 15, no. 8, pp. 1546–1553, Aug. 1997.
- [15] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and post-detection filtering," *J. Lightw. Technol.*, vol. 18, no. 11, pp. 1493–1503, Nov. 2000.
- [16] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "A novel analytical approach to the evaluation of the impact of fiber parametric gain on the bit error rate," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2154–2163, Dec. 2001.
- [17] P. J. Winzer, M. Pfennigbauer, M. M. Strasser, and W. R. Leeb, "Optimum filter bandwidths for optically preamplified NRZ and RZ receivers," *J. Lightw. Technol.*, vol. 19, no. 9, pp. 1263–1273, Sept. 2001.
- [18] M. Kac and A. J. F. Siegert, "On the theory of noise in radio receivers with square law detectors," *J. Appl. Phys.*, vol. 18, no. 4, pp. 383–397, Apr. 1947.
- [19] C. W. Helstrom, "Distribution of the filtered output of a quadratic rectifier computed by numerical contour integration," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 4, pp. 450–463, Jul. 1986.
- [20] S. W. Golomb, *Shift Register Sequences*. San Francisco, CA: Holden-Day, 1967.
- [21] J. Wang and J. M. Kahn, "Impact of chromatic and polarization-mode dispersions on dpsk systems using interferometric demodulation and direct detection," *J. Lightw. Technol.*, vol. 22, no. 2, pp. 362–371, 2004.
- [22] A. H. Gnauck and P. J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 115–130, Jan. 2005.
- [23] C. W. Helstrom, "Approximate evaluation of detection probabilities in radar and optical communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 14, pp. 630–640, Jul. 1978.
- [24] H. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.
- [25] E. J. Nyström, "Über die praktische auflösung von integralgleichungen mit anwendungen auf randwertaufgaben," *Acta Mathematica*, vol. 54, pp. 185–204, 1930.
- [26] F. G. Tricomi, *Integral Equations*. New York: Dover, 1985.
- [27] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.
- [28] R. Holzlohner, V. S. Grigoryan, C. R. Menyuk, and A. L. Kath, "Accurate calculation of eye diagrams and bit error rates in optical transmission systems using linearization," *J. Lightw. Technol.*, vol. 20, no. 3, pp. 389–400, Mar. 2002.
- [29] M. Nazarathy, B. Livshitz, Y. Atzmon, M. Secondini, and E. Forestieri, "Optically amplified direct detection with pre- and post-filtering: A Volterra series approach," *J. Lightw. Technol.*, vol. 26, no. 22, pp. 3677–3693, Nov. 2008.

**Enrico Forestieri** (S'91-M'92) was born in Milazzo, Italy, in 1960. He received the Dr. Ing. degree in electronics engineering from the University of Pisa, Pisa, Italy, in 1988.

From 1989 to 1991, he has been a Postdoctoral scholar with the University of Parma, Italy, working on optical communication systems. From 1991 to 2000, he was a Research Scientist and faculty member of the University of Parma. Since 2001, he has been with Scuola Superiore Sant'Anna di Studi Universitari e di Perfezionamento, Pisa, where currently he is a Professor of Telecommunications. His research interests are in the general area of digital communication theory and optical communication systems, with special attention to adaptive optical and electronic equalization, channel coding, and advanced modulation formats for optical systems. He is the leader of the "Optical Transmission Theory and Techniques" area at the Integrated Research Center for Photonic Networks and Technologies (IRCPhoNeT), Pisa. His research activity has led to numerous scientific publications in leading international journals and conference proceedings, as well as a few patents.

Dr. Forestieri was General Chairman of the Tyrrhenian International Workshop on Digital Communications in 2004.



**Marco Secondini** was born in Rome, Italy, in 1975. He received the Dr. Ing. degree in electronics engineering from the University of Roma Tre, Rome, in 2000, and the Ph.D. degree from Scuola Superiore Sant'Anna, Pisa, Italy, in 2006.

During 2001, he was with QPlus Networks Inc., developing fiber optic systems for ultralong-haul communications. Since 2002, he has been with Scuola Superiore Sant'Anna, Pisa, where he is currently an Assistant Professor. He also collaborates with the National Photonic Networks Laboratory of the CNIT, Pisa. During 2005, he was a Visiting Faculty Research Assistant with the Photonics Group, University of Maryland Baltimore County. His current research interests are in the area of optical communication theory and include: fiber nonlinearities, chromatic and polarization mode dispersion, optical equalization techniques, and Monte Carlo methods for simulations.