Application of the Saddle Point Method for the Evaluation of Crosstalk Implications in an Arrayed-Waveguide Grating Interconnection

Thomas Kamalakis and Thomas Sphicopoulos, *Member, IEEE*

***Abstract*—The effect of in-band crosstalk can pose severe limitations in an optical network. In this paper, the implications of in-band crosstalk induced by an arrayed-waveguide grating (AWG) router in a passive $N \times N$ optical interconnection are analyzed with non-Gaussian statistics using a numerical model. The model is based on the saddle point approximation and takes into account fluctuations of the transfer function's sidelobes induced by the phase errors in the grating arms, phase noise, polarization variations, bit misalignment, and shot and thermal noise. The influence of these effects on the interconnection's bit error rate (BER) is analyzed. The validity of the Gaussian assumption for the crosstalk noise statistics is discussed. Finally, the model is used to examine the mean crosstalk requirements for various numbers of network nodes.**

***Index Terms*—Gratings, optical crosstalk, phased arrays, wavelength division multiplexing (WDM).**

I. Introduction

ARRAYED-WAVEGUIDE gratings (AWGs) [1], [2] are important components for the realization of modern optical wavelength division multiplexing (WDM) [3] networks on which they can serve as multiplexers, demultiplexers, and wavelength routers. Devices of this kind have been made commercially available, and there are techniques that allow polarization-insensitive operation [4]. However, due to fabrication imperfections, AWGs suffer from phase errors [5], [6], which tend to create sidelobes outside the main lobe of their transfer function, making the separation of the wavelength channels less than ideal.

In a network with wavelength reuse, this can lead to in-band crosstalk noise (which is on the same wavelength as that of the signal) as well as out-of-band crosstalk noise (which is on a different wavelength) [7]. Out-of-band crosstalk can be removed with additional filtering at the receiver side, but in-band crosstalk cannot be removed and can significantly increase the bit error rate (BER) of the received signal. In the network considered in this paper, in-band crosstalk originates from different laser sources at different network nodes. As a result, the noise components and the signal will be considered statistically independent of each other.

Many studies have been done to understand the implications of in-band crosstalk in an optical network (see [8]–[11] and references therein). In the case of many independent identically distributed (i.i.d.) interferers, the probability density function (pdf) of the crosstalk noise can be assumed approximately Gaussian because of the central limit theorem (CLT). The Gaussian approximation has also been used for the estimation of the performance of a $N \times N$ AWG interconnection [8] where the interferers are not of equal power. Although the CLT is still valid in this case, the Gaussian model may not provide a sufficiently accurate description. This is because the pdf of the crosstalk noise may not converge to a Gaussian shape as fast as in the case of i.i.d. interferers. Also, this study did not include the effects of shot noise and bit misalignment, which has been recently studied in [13] for a general optical add-drop multiplexer using the Gaussian model.

The saddle point approximation [12] can be used for the computation of the BER in the presence of noise with non-Gaussian statistics using its moment-generating function (MGF). This method has been applied [11] for the special case of i.i.d. interferers, taking into account phase noise, polarization variations, and shot and thermal noise. In this paper, the MGF of the in-band crosstalk noise which is present in a $N \times N$ AWG passive optical interconnection is evaluated, and the BER is calculated using the saddle point method. The strength of the interfering channels is estimated using the transfer function of the AWG, taking into account the phase errors on the grating arms. This allows an accurate description of the filtering characteristics of the AWG, including the chirp it induces both in the signal and in the noise components. The effects of phase noise, polarization variations, and shot and thermal noise are incorporated in the model. Bit misalignment is also taken into account in the computation of the MGFs. The validity of the Gaussian approximation is discussed both theoretically and numerically by comparing the results of the two methods. With the aid of some examples, the importance of the factors determining the BER is shown. Finally, the performance of the network for various values of the AWG average sidelobe level is presented in the form of diagrams.

II. Crosstalk Noise Model Description

A. General Considerations

A typical $N \times N$ interconnection employing an AWG wavelength router is depicted in Fig. 1(a). There are N nodes in the network that can communicate with each other resulting in N^2 simultaneous connections. The AWG is designed so that its free spectral range (FSR) is equal to N multiplied by the channel

Manuscript received February 11, 2002; revised May 2, 2002.

The authors are with the Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, GR-157 84, Athens, Greece (e-mail: thkam@di.uoa.gr; tsfikop@di.uoa.gr).

Digital Object Identifier 10.1109/JLT.2002.800807

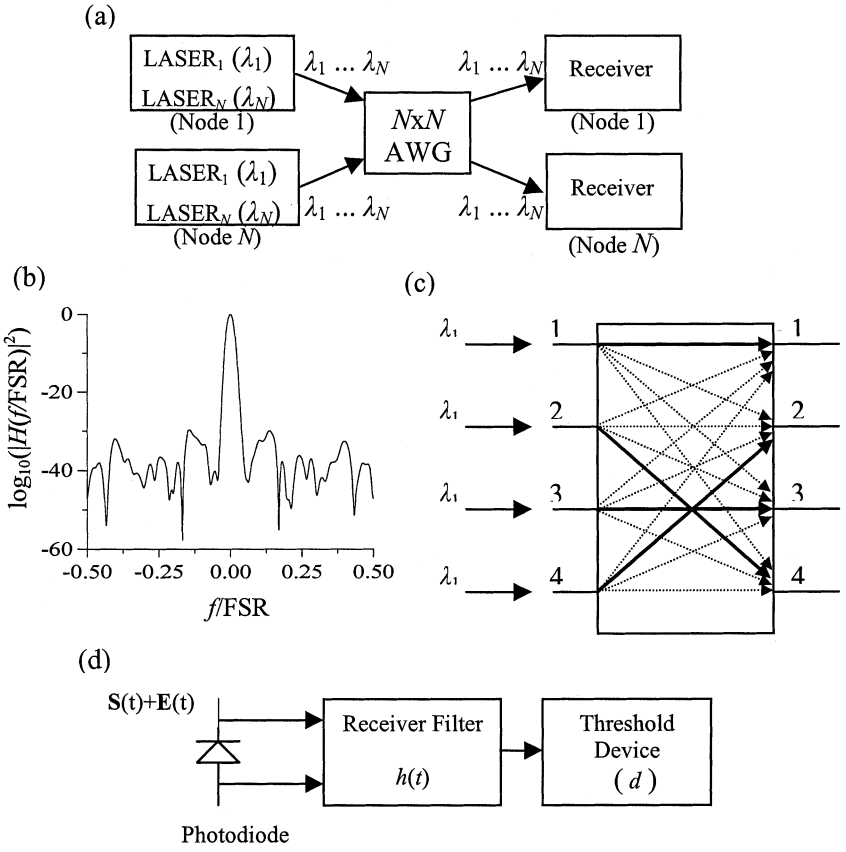

Fig. 1. (a) A $N \times N$ optical interconnection employing a AWG wavelength router. (b) The transmittance $|H(f/\mathrm{FSR})|^2$ of a 16×16 AWG with $M = 65$ waveguides and phase errors δ_m with standard deviation equal to $\pi/10$. (c) Routing of wavelength λ_1 between the input and output ports of the AWG. Solid arrows represent intended directions, and dashed arrows represent leaked components. (d) A typical receiver diagram.

spacing, and as a result, the total number of wavelengths required is N [14].

In Fig. 1(b) the transmittance $|H(f/\mathrm{FSR})|^2$ of the central input and output ports is plotted. The frequency f is measured around the central frequency of the AWG and is normalized by dividing it by FSR as usual [5]. Because of the cyclic properties of the AWG [14], the transfer function between input port p and output port q is $e_{pq}H(f-(p+q)\Delta f\mathrm{ch})$ where e_{pq} is the efficiency and Δf_{ch} is the channel spacing of the AWG. Each node is assumed to have N laser sources, each transmitting on a different wavelength $(\lambda_1, \ldots, \lambda_N)$. The same wavelengths are used in the other network nodes as well. Because of this wavelength reuse, the signal at a wavelength λ_1 from a network node will be accompanied by crosstalk noise components originating from the other network nodes.

Fig. 1(c) shows how the AWG routes the signals at wavelength λ_1 from its input ports to its output ports. The solid arrows represent the desired wavelength routes, and the dashed arrows represent the leaked components. For example, although light at output port 2 should come only from input port 4 at wavelength λ_1, there are also light components at λ_1 originating from input ports 1 to 3. This is because the transfer function between output port 2 and each of the input ports 1 to 3 cannot completely suppress light at wavelength λ_1 originating from input ports 1 to 3. It is interesting to note that in the case of an AWG interconnection, the in-band crosstalk noise components come from laser sources located at different network nodes. As a result, at a given output port, the crosstalk noise components are independent of each other and of the signal.

Fig. 1(d) shows the diagram of a simple receiver circuit without optical preamplification. The optical field entering the receiver's photodiode is the sum $\mathbf{S}+\mathbf{E}$ of the signal $\mathbf{S}$ on a wavelength λ and the in-band crosstalk noise $\mathbf{E}$ on the same wavelength,

$$\mathbf{S} = \mathbf{x}_0 e^{j\phi_0} g_0(t) \tag{1}$$

and

$$\mathbf{E} = \sum_{i=1}^{N-1} \mathbf{x}_i e^{j\phi_i} g_i(t) \tag{2}$$

where $\mathbf{S}$ denotes the envelope of the signal, and $\mathbf{E}$ denotes the envelope of the crosstalk noise. The unit vectors $\mathbf{x}_i$ denote the polarization direction of the signal $(i = 0)$ and the crosstalk components $(1 \le i \le N-1)$, while the symbols ϕ_i denote the random phases induced by the phase noise of the lasers. The function $g_0(t)$ describes the time variation of the desired signal at the considered output port of the AWG router, and the functions $g_i(t)$ $(i \ge 1)$ describe the time variation of the accompanying crosstalk noise components. The photocurrent $i(t)$, measured in photoelectrons per second, which is induced in the photodetector, is

$$\begin{aligned} i(t) =& \frac{\eta}{2hf}|\mathbf{S}+\mathbf{E}|^2 \\ \cong& \frac{\eta}{2hf}|g_0(t)|^2 \\ &+ \frac{\eta}{hf}\mathrm{Re}\left\{\sum_i \cos\theta_i g_i^*(t) g_0(t) e^{j(\phi_0-\phi_i)}\right\} \end{aligned} \quad (3)$$

where η is the photodetector quantum efficiency, h is Planck's constant, f is the frequency corresponding to the wavelength λ. In this equation, $\langle \mathbf{x}_0, \mathbf{x}_i \rangle$ was replaced by $\cos\theta_i$, where θ_i is the angle between the polarization vector $\mathbf{x}_0$ of the signal and the polarization vector $\mathbf{x}_i$ of crosstalk component i.

As will be shown next, the value of the BER is independent of the value of ϕ_0. The crosstalk–crosstalk beating term $|\mathbf{E}|^2$, which should be small in systems hoping to achieve small BER, is neglected, and g^* denotes the complex conjugate of g. To avoid carrying the factor $\eta/2hf$ in further calculations, the optical field is normalized so that $\eta/2hf = 1$. At the output of the receiver filter, the decision variable D is defined by

$$D(T) = \int_0^T h(T-t) i(t) dt \quad (4)$$

where T denotes the bit duration and $h(t)$ is the impulse response of the filter. For simplicity, the filter will be assumed to be an integrator ($h(t) = 1$ for $t \in [0, T]$). The decision variable can be decomposed into two parts, one due to the signal D_s and one due to the crosstalk noise $D_n (D = D_n + D_s)$ as follows:

$$D_s(T) = \int_0^T |g_0(t)|^2 dt \quad (5)$$

$$D_n(T) = 2\mathrm{Re}\left\{\sum_i \cos\theta_i e^{-j(\phi_i-\phi_0)} \int_0^T g_i^*(t) g_0(t) dt\right\}. \quad (6)$$

In (6), it is assumed that the random phases ϕ_i and the polarization angles θ_i change slowly and can therefore be considered constant over several bit periods.

In order to make the model as general as possible, it is imperative to take into account the induced chirp [15] in both the signal and the crosstalk terms due to the phase errors of the AWG. As a result, the functions $g_i(t)\ 0 \le i \le N-1$ should be assumed complex functions

$$g_i(t) = |g_i(t)| e^{jw_i(t)} \quad (7)$$

where $w_i(t)$ is the phase of $g_i(t)$, $w_i(t) = \arg(g_i(t))$. By replacing (7) in (6), the following equation is obtained:

$$D(T) = G_0 + 2\sum_i |G_i| \cos(\phi_i - r_i - \phi_0)\cos\theta_i \quad (8)$$

where

$$G_i = \int_0^T g_0(t) g_i^*(t) dt \quad (9)$$

and

$$r_i = \arg(G_i). \quad (10)$$

As seen by Fig. 1(c), the crosstalk components accompanying the signal are originating from different input ports. Hence, the phases ϕ_i are due to the phase noise of different laser sources and are considered independent. The same is true for the polarization direction angles θ_i. Both ϕ_i and θ_i are assumed to be uniformly distributed in $[0, 2\pi]$, and consequently, the moments of $\cos(\phi_i - r_i - \phi_0)$ for each value of r_i are calculated using

$$E\left\{\cos^{2n+1}(\phi_i - r_i - \phi_0)|r_i, \phi_0\right\} = 0 \quad (11a)$$

$$\begin{aligned} &E\left\{\cos^{2n}(\phi_i - r_i - \phi_0)|r_i, \phi_0\right\} \\ &\quad = \frac{1}{2\pi}\int_0^{2\pi} \cos^{2n}(\phi_i - r_i - \phi_0) d\phi_i \\ &\quad = \frac{1}{2^{2n}}\frac{(2n)!}{(n!)^2}. \end{aligned} \quad (11b)$$

In the above equations, $E\{\cdot\}$ denotes the expected value of a random variable. Equations (11a) and (11b) can be derived by expanding $\cos^l \phi$ in terms of $\cos(k\phi)$ $(1 \le k \le l)$ and performing the integration. Equations (11a) and (11b) express the fact that the moments of $\cos(\phi_i - r_i - \phi_0)$ are independent of ϕ_0 and r_i. Consequently, ϕ_0 and r_i do not have any bearing on the statistical behavior of the decision variable and can be set equal to zero ($r_i = \phi_0 = 0$), and D can be written as follows:

$$D(T) = G_0 + 2\sum_i |G_i| \cos\phi_i \cos\theta_i. \quad (12)$$

B. Modeling of the Imperfect Filtering Characteristic

Considering the central output port, the functions $g_i(t)$ $(1 \le i \le N-1)$, which are due to the imperfect filter characteristic of the router, can be calculated using

$$g_i(t) = \int_{-\infty}^{+\infty} H_i(f) F_i(f) e^{j2\pi f t} df. \quad (13)$$

In (13), $H_i(f)$ is the transfer function of the AWG between the input port i and the central output, and $F_i(f)$ is the input spectrum of interfering signal i centered on the same wavelength as the signal. Assuming that the channel uniformity is equal to unity ($e_{pq} \cong 1$) and using the cyclic property of the AWG, the transfer function $H_i(f)$ satisfies [5], [14]:

$$\begin{aligned} H_i(f + i\Delta f_{\mathrm{ch}}) =& H(f) \\ =& \sum_{k=-P}^{P} C_k \exp\left(j\frac{2\pi k}{\mathrm{FSR}} f\right)\exp(j\delta_k) \end{aligned} \quad (14)$$

where $H(f)$ is the transfer function between the central input and output ports of the AWG, C_k is the power intercepted by the k-th grating waveguide normalized to the total power, f is the optical frequency, and δ_k represents the phase errors of the AWG. The total number of grating waveguides is $N_w = 2P+1$.

The phase errors of the AWG are due to the imperfections of the grating waveguides, which cause a deviation of their effective index from its nominal value. In the simulations performed in this paper, where a large number of AWGs with different crosstalk characteristics must be considered, the phase errors will be assumed zero-mean Gaussian random variables [17]. Since the phase errors occur in different waveguides they will also be assumed independent. In any case, the validity of the

method described in this paper does not depend on the particular statistical behavior of the phase errors. A different behavior of the phase errors will only change the values of the BER in the simulations, but the method itself will still be applicable. Furthermore, in a real $N \times N$ interconnection, the phase errors of the AWG can be measured experimentally, and the transfer function can be calculated with high precision [16] using (14). The MGF of the crosstalk components can then be estimated as discussed in Section II-D, and the BER can be evaluated as discussed in Section II-F.

One can establish a relationship between the mean value of $|H(f)|^2$, the power distribution of the grating arms C_k, and the standard deviation of the phase errors σ:

$$E\{|H(f)|^2\} = \sum_{k=-P}^{P} C_k^2 \left(1 - e^{-\sigma^2}\right) + e^{-\sigma^2} |H_{\text{ideal}}(f)|^2 \tag{15}$$

where $H_{\text{ideal}}(f)$ is the transfer function between the central input and output port in the absence of phase errors ($\delta_k = 0$). Equation (15) is derived using the fact that $E\{e^{\pm j\delta_k}\} = e^{-\sigma^2/2}$ for Gaussian-distributed δ_k and that $E\{e^{j(\delta_k - \delta_n)}\} = E\{e^{j\delta_k}\}E\{e^{j\delta_n}\} = e^{-\sigma^2}$ for $n \neq k$ since δ_k are assumed independent. Because the sidelobes of $H_{\text{ideal}}(f)$ are very low, it is deduced that the mean value of $|H(f)|^2$ outside the main lobe is constant and equal to

$$X = E\{|H(f)|^2\} \cong \left(1 - e^{-\sigma^2}\right) \sum_{k=-P}^{P} C_k^2. \tag{16}$$

This expression relates the mean sidelobe level X and the standard deviation of the phase errors σ given the power distribution C_k of the grating arms. The quantity X is directly related to the sidelobe level of the transfer function of the AWG and will therefore be used in the results presented next.

C. Inclusion of Random Bit Misalignment

To include the random bit misalignment, one should allow for the function $f_i(t)$, which corresponds to the interfering spectra $F_i(f)$, to be displaced by a random offset τ_i. Because of the linearity of (13), the functions $g_i(t)$ will also be displaced by τ_i. Therefore, the random variables G_i used to calculate the decision variable $D(T)$ will depend on τ_i as follows:

$$G_i(\tau_i) = \int_0^T g_0(t) g_i^*(t - \tau_i) dt \tag{17}$$

and $D(T)$ can be written as

$$D(T) = G_0 + 2 \sum_i |G_i(\tau_i)| \cos\phi_i \cos\theta_i. \tag{18}$$

For simplicity, the time offsets are measured with respect to the offset of the signal, i.e., $\tau_0 = 0$. The value of the random variables G_i will depend on the two bits a and b of channel i that overlap with the bit of the signal within its duration T as follows:

$$G_i^{ab}(\tau_i) = \int_0^T g_0(t)\left(a q_i^*(t - \tau_i) + b q_i^*(t + T - \tau_i)\right) dt. \tag{19}$$

In the relation in (19), $q_i(t)$ is the function $g_i(t)$ corresponding to a single "1" bit at the i interfering channel. The time offset τ_i is assumed uniformly distributed in $[0, T]$ [9]. If the two bits are different ($a \neq b$), only one pulse $q_i(t)$ inside the interval $[0, T]$ must be considered. Consequently, $G_i^{01} = G_i^{10}$, and the expected value of G_i^{ab} can be written as

$$E\left\{G_i^{ab}\right\} = \frac{1}{4} E\left\{G_i^{11}\right\} + \frac{1}{2} E\left\{G_i^{01}\right\}. \tag{20}$$

In (20), a perfect extinction ratio was assumed, which implies that $G_i^{00} = 0$.

D. Calculation of the MGF

As pointed out in Section I, the MGF must be evaluated in order to use the saddle point method to calculate the BER. The MGF $M_i(s)$ of each crosstalk term i will be given by

$$M_i(s) = E\left\{\exp\left(s\left|G_i(\tau_i)\right| \cos\theta_i \cos\phi_i\right)\right\}. \tag{21}$$

Using the fact that $E\{\exp(2\cos\phi_i \cos\theta_i)\} = I_0^2(s)$ [11], where $I_0(s)$ is the modified Bessel function of zero order, the MGF can be written

$$M_i^{\text{PPM}}(s) = \frac{1}{4T}\int_0^T I_0^2\Big(s\left|G_i^{11}(\tau)\right|\Big) d\tau + \frac{1}{2T}\int_0^T I_0^2\Big(s\left|G_i^{01}(\tau)\right|\Big) d\tau + \frac{1}{4}. \tag{22a}$$

The initials PPM stand for the inclusion of polarization variations, phase noise, and bit misalignment, respectively. If no misalignment is assumed ($\tau_i = 0$), then the signal's bit coincides exactly with one bit b_i of each interfering channel i. Consequently, the adjacent bits of b_i do not affect the MGF, which will be given by

$$M_i^{\text{PP}}(s) = \frac{1}{2} I_0^2\Big(s\left|G_i^{11}(0)\right|\Big) + \frac{1}{2} \tag{22b}$$

where, similarly to (22a), the existence of the initials PP in $M_i^{\text{PP}}(s)$ stands for the inclusion of polarization variations and phase noise.

In the cases in which either the signals are copolarized ($\theta_i = 0$) or the phase noise is neglected ($\phi_i = 0$), the MGF is written as

$$M_i^{\text{PM}}(s) = \frac{1}{4T}\int_0^T I_0\Big(2s\left|G_i^{11}(\tau)\right|\Big) d\tau + \frac{1}{2T}\int_0^T I_0\Big(2s\left|G_i^{01}(\tau)\right|\Big) d\tau + \frac{1}{4}. \tag{22c}$$

In (22c), the fact that $E\{\exp(2s\cos\theta_i)\} = E\{\exp(2s\cos\phi_i)\} = I_0(2s)$ [11] was used. Once again, if bit misalignment is not present, the above result becomes

$$M_i^P(s) = \frac{1}{2} I_0\Big(2s\left|G_i^{11}(0)\right|\Big) + \frac{1}{2}. \tag{22d}$$

If neither the phase noise nor the polarization variations are taken into account, there are no cosines involved in the crosstalk

terms. In this case, using (6) and (9), it is easy to show that the MGF of each individual crosstalk term will be given by

$$M_i^M(s) = \frac{1}{4T}\int_0^T \exp\left(2s\mathrm{Re}\left\{G_i^{11}(\tau)\right\}\right)d\tau + \frac{1}{2T}\int_0^T \exp\left(2s\mathrm{Re}\left\{G_i^{01}(\tau)\right\}\right)d\tau + \frac{1}{4}. \quad (22e)$$

Finally, if bit misalignment is ignored, (22e) reduces to

$$M_i^{\mathrm{NONE}}(s) = \frac{1}{2}\exp\left(2s\mathrm{Re}\left\{G_i^{11}(0)\right\}\right) + \frac{1}{2}. \quad (22f)$$

Equations (22a)–(22f) can be used to calculate the MGFs of the interfering terms for all combinations of the factors (i.e., phase noise, polarization variation, and bit misalignment) that affect the crosstalk noise. The MGF $M'_{D|1}(s)$ of the decision variable D given that the signal bit is "1" is

$$M'_{D|1}(s) = \exp(sG_0)\prod_{i=1}^{N-1} M_i(s). \quad (23)$$

In the above equation, $\exp(sG_0)$ is the MGF of the signal, which, given the values of δ_k, is deterministic (since $\tau_0 = 0$).

E. Inclusion of Shot and Thermal Noise

Finally, the effects of thermal noise and of shot noise, due to the random arrival of the photons at the receiver, should also be included. Since the photon arrival process is a Poisson process, the variable s must be replaced by $e^s - 1$ [12]. On the other hand, the thermal noise is represented by a Gaussian random variable D_{th}, which is added to D with zero-mean value and $E\{D_{\mathrm{th}}^2\} = \sigma_{\mathrm{th}}^2$. The thermal noise D_{th} has MGF equal to $\exp\left(\sigma_{\mathrm{th}}^2 s^2/2\right)$ and is independent of the crosstalk noise. Hence, the MGF of the decision variable given that the signal bit is "1", including the shot and thermal noise, is written

$$M_{D|1}(s) = \exp\left(G_0(e^s - 1) + \frac{\sigma_{\mathrm{th}}^2}{2}s^2\right)\prod_{i=1}^{N-1} M_i(e^s - 1). \quad (24)$$

F. Calculation of the BER with the Saddle Point Method

Given the MGF $M(s)$ of a random variable, the saddle point method can be used to calculate numerically the cumulative probability distribution function (cpdf) of a random variable. This method is based on the fact that the MGF of a random variable is the Laplace transform of its pdf. Consequently, the cpdf reduces to the integral of $s^{-1}M(s)e^{-sd}$ (d being the decision threshold used at the receiver) on a complex contour. This integral can be approximated by a simple formula involving the logarithm of the integrand and its second derivative both calculated at the saddle point of the integrand [12].

In the case examined, the MGF $M_{D|1}(s)$, which can be calculated numerically performing the simple integrations in (22a) for each product component of (24), can be used to evaluate the BER using the saddle point method. If we define

$$\psi_1(s) = \ln\left(M_{D|1}(s)\right) - sd - \ln|s| \quad (25)$$

where d is the decision threshold used in the receiver, then by locating the negative saddle point s_1 of the function $\psi_1(s)$ on the real axis, the following formula can be used for the calculation of the BER P_{e1}, given that the signal bit is "1" [12]:

$$P_{e1} = P\{D \le d|1\} = \int_{-\infty}^{d} p_1(u)du \cong \frac{\exp\left(\psi_1(s_1)\right)}{\sqrt{2\pi\psi_1''(s_1)}} \quad (26)$$

where p_1 is the pdf of the decision variable (including the thermal noise and the shot noise contributions).

The BER, given that the signal bit is "0", is calculated by

$$P_{e0} = \frac{1}{2}\mathrm{erfc}\left(\frac{d}{\sigma_{\mathrm{th}}\sqrt{2}}\right) \quad (27)$$

since, due to the perfect extinction ratio assumption, only the thermal noise is present in this case. In some systems, a finite extinction ratio might have to be taken into account, resulting in the presence of crosstalk terms in the case of the "0" signal bit as well. In this case, the saddle point can be used again to calculate P_{e0} with MGFs taking into account the finite extinction ratio. In any case, the average BER will be given by

$$P_e = \frac{1}{2}(P_{e1} + P_{e0}). \quad (28)$$

The optimum decision threshold d_{opt} is the value of d that minimizes P_e and can be found by computing numerically the derivative P'_e of P_e and solving the equation $P'_e(d_{\mathrm{opt}}) = 0$. In the following, the BER values presented are the minimum values obtained by optimizing the decision threshold ($d = d_{\mathrm{opt}}$).

G. Accuracy of the Model

The saddle point method provides an accurate and efficient way to compute the BER in the case of a noise with a cpdf not known in closed form [12]. The accuracy of the method was investigated in [11], where a comparison between the BER predicted by the saddle point method and the experimentally measured BER was carried out for the case of i.i.d. (i.e., equal powered) crosstalk noise components. The theory was shown to be in good agreement with the experimental results. In this paper, an AWG interconnection is addressed, and since particular attention is paid to the precise modeling of the AWG transfer function, it is expected that the calculated BER should be in agreement with the experimentally measured BER in this case as well.

III. Validity of the Gaussian Approximation

The BER of an $N \times N$ passive optical interconnection has been previously studied using the Gaussian model [8]. However, since the interfering signals are expected to have different powers, the Gaussian model might not provide an accurate description of the noise behavior. In this section, the results of the

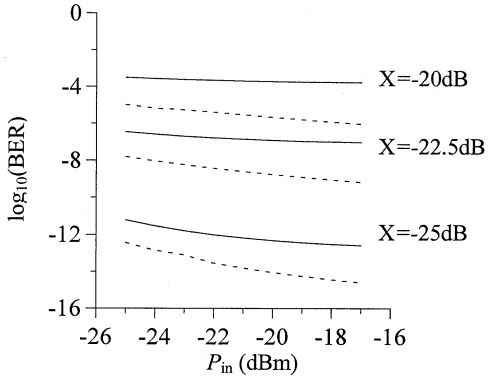

Fig. 2. Variation of the BER with respect to the power entering the input of a 16×16 AWG for various values of the mean sidelobe level (X). The dashed lines correspond to the results of the Gaussian model, and the solid lines correspond to the results of the saddle point method.

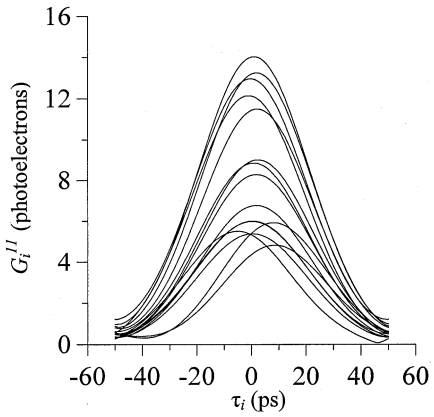

Fig. 3. A typical example of the functions $G_i^{11}(\tau_i)$ for the 15 interfering channels in the case of a 16×16 AWG.

saddle point method and of the Gaussian model will be compared.

In the Gaussian model, the noise, given that the signal bit is $b_s = 1$, is assumed to be a Gaussian random variable with noise power σ_n^2, and the BER is determined by

$$P_e = \frac{1}{4}\mathrm{erfc}\left(\frac{m_s - d}{\sqrt{2}\sigma_n}\right) + \frac{1}{4}\mathrm{erfc}\left(\frac{d}{\sqrt{2}\sigma_{\mathrm{th}}}\right). \tag{29}$$

In (29), m_s and σ_n are the mean value of the signal and the power of the crosstalk noise, respectively, in the case where $b_s = 1$. To calculate σ_n, (12) is used together with $E\{\cos^2\theta_i\} = E\{\cos^2\phi_i\} = 1/2$. After carrying out the computations, the following result is obtained:

$$\sigma_n^2 = \frac{1}{4}\sum_{i=1}^{N-1}\int_0^T d\tau_i \left(G_i^{11}(\tau_i)\right)^2 + \frac{1}{2}\sum_{i=1}^{N-1}\int_0^T d\tau_i \left(G_i^{01}(\tau_i)\right)^2 + G_0 + \sigma_{\mathrm{th}}^2. \tag{30}$$

In (30), the term G_0 is the power of the shot noise.

A comparison of the results obtained using the Gaussian approximation and the saddle point method is shown in Fig. 2 for the case of a 16×16 AWG router with three different mean sidelobe values X. The phase errors of each case were calculated using

$$\delta_k = \sigma(X)\delta_k^0 \tag{31}$$

where δ_k^0 is a Gaussian distribution with standard deviation equal to 1 (the same samples of δ_k^0 were used in the three cases of Fig. 2), and $\sigma(X)$ is calculated by solving (16) with respect to σ. The channel spacing is set to 200 GHz, and at the input of the AWG, 10 Gb/s Gaussian RZ pulses are assumed with full-width maximum equal to one-third of the bit duration, which is $T = 100$ ps. The thermal noise variance was set equal to $\sigma_{\mathrm{th}} = 5$. The input power P_{in} is the peak power of the "1" signal bit at the input of the AWG. These specifications will be used throughout this paper.

It is obvious that the Gaussian approximation does not predict an accurate value for the BER. In Fig. 3, the functions G_i^{11}, which determine the amplitude of each interferer, have been plotted for the case of an AWG with mean crosstalk $X = -22$ dB. As stated in Section I, the interferers are not of equal power, and hence, the noise pdf converges to a Gaussian shape slowly. This explains why the BER predicted by the Gaussian approximation does not agree with the BER predicted by the saddle point method. In fact, the Gaussian model seems to underestimate the BER.

To justify this statement theoretically, the higher moments of the crosstalk noise can be compared in the case of the more accurate, non-Gaussian statistics and in the case of Gaussian statistics. The $2n$th-order moment of a random variable R having pdf equal to $f_R(r)$ is given by $\int (r - \bar{R})^{2n} f_R(r)dr$ and is a measure of how fast the pdf of R tends to zero as $r \to \pm\infty$. For two random variables having the same standard deviation and mean value, the fourth-order moment can be used to determine which of the two has the pdf with the lower tails. This argument can be used when the two pdfs do not differ significantly in shape. As noted in Section I, the pdf of the crosstalk noise converges to a Gaussian shape as the number of interferers increase. Hence, by comparing the higher order moments of the actual noise distribution and its Gaussian approximation, it may be possible to conclude which one has the lower tails.

First, the higher order moments of the photocurrent y, with only a single interfering term, are considered. In this case, y is given by

$$y = 2\sqrt{2}\varepsilon\sigma\cos\theta\cos\phi \tag{32}$$

where ϕ and θ are uniformly distributed in $[0, 2\pi]$. To account for bit variations, ε is taken as a random variable assuming the values 0 and 1 with equal probability $(P(\varepsilon = 1) = P(\varepsilon = 0) = 1/2)$. Finally, σ^2 is the standard deviation of the crosstalk noise since

$$E\{y^2\} = 8\sigma^2 E\{\varepsilon^2\}E\{\cos^2\theta\}E\{\cos^2\phi\} = \sigma^2. \tag{33}$$

The fourth-order moment σ_4 of y is equal to

$$\sigma_4 = E\{y^4\} = 64\sigma^4 E\{\varepsilon^4\}E\{\cos^4\theta\}E\{\cos^4\phi\} = \frac{9}{2}\sigma^4 \tag{34}$$

where formula (11b) was used along with the fact that $E\{\varepsilon^4\} = 1/2$.

Let y_g be the Gaussian approximation of y, i.e., a Gaussian random variable with the same standard deviation as y, $E\{y_g^2\} = E\{y^2\} = \sigma^2$. The fourth-order moment of y_g is $E\{y_g^4\} = 3\sigma^4$, which is smaller than the fourth-order moment $E\{y^4\}$ of y given by (34). When there are more interferers present, the nth-order moment of the crosstalk noise can be calculated by applying successively the fact that, for two independent random variables r_1 and r_2,

$$E\{(r_1+r_2)^n\} = \sum_{k=0}^{n} \binom{n}{k} E\{r_1^k\} E\{r_2^{n-k}\}. \quad (35)$$

This equation can be applied regardless of the individual distribution of r_1 and r_2. Equation (35) can be used to show that if $E\{u_1^k\} \le E\{r_1^k\}$ and $E\{u_2^k\} \le E\{r_2^k\}$ for $k \le n$, then

$$E\{(u_2+u_1)^n\} \le E\{(r_2+r_1)^n\}. \quad (36)$$

Applying (36) successively, in the case of $n = 4$, it can be shown that when more interferers are present, the fourth-order moment σ_4 of the crosstalk noise will be again greater than that of a Gaussian random variable with the same standard deviation. Using the same reasoning, it can be ascertained that this holds for $n = 6, 8$, and 10. In other words, the higher order moments up to the tenth order corroborate the fact that the Gaussian approximation underestimates the BER.

In this analysis, the effect of shot noise was neglected. The shot noise adds one further random factor in the signal's statistics because the arrivals of the photons are assumed to obey a Poisson process. The MGF $M_c(s)$ of the photocurrent without the thermal and shot noises is given by

$$M_c(s) = \sum_{n=0}^{\infty} \frac{I_n}{n!} s^n \quad (37)$$

where $I_n = E\{(m+y)^n\}$ are the moments of the sum of the signal m and the crosstalk noise y. Using the binomial expansion of $(m+y)^n$ and the fact that $E\{y\} = E\{y^3\} = 0$, $E\{y^2\} = \sigma^2$, and $E\{y^4\} = \sigma_4$,

$$I_4 = m^4 + 6m^2\sigma^2 + \sigma_4 \quad (38a)$$
$$I_3 = m^3 + 3m\sigma^2 \quad (38b)$$
$$I_2 = m^2 + \sigma^2 \quad (38c)$$

and

$$I_1 = m. \quad (38d)$$

When the shot noise is taken into account, the MGF of the photocurrent becomes

$$M_{\rm csh}(s) = M_c(e^s - 1) = \sum_{n-0}^{\infty} \frac{I_n}{n!} (e^s - 1)^n. \quad (39)$$

The expectations $E\{Y^n\} = I'_n$ of the photocurrent Y in the presence of shot noise can be calculated using the fact that $I'_n = d^n M_{\rm csh}/ds^n|_{s=0}$. After carrying out the computations, the following result is obtained:

$$I'_4 = I_4 + 7I_3 + 6I_2 + I_1 \quad (40a)$$
$$I'_3 = I_3 + 3I_2 + I_1 \quad (40b)$$
$$I'_2 = I_2 + I_1 \quad (40c)$$
$$I'_1 = I_1 \quad (40d)$$

and

$$I'_1 = I_1. \quad (40d)$$

Using the binomial expansion of $(Y - I'_1)^4$, the fourth-order moment of Y is written as

$$E\left\{(Y - I'_1)^4\right\} = I'_4 - 4I'_3I'_1 + 6I'_2(I'_1)^2 - 3(I'_1)^4. \quad (41)$$

Using (38) and (40) within (41), the fourth-order moment of Y eventually takes the form

$$E\left\{\left(Y - E\{Y\}\right)^4\right\} = 3m^2 + 6m\sigma^2 + m + 7\sigma^2 + \sigma_4. \quad (42)$$

The fourth-order moment of a Gaussian distribution with standard deviation σ_g^2 that is equal to the sum of the power of the crosstalk and shot noises $\left(\sigma_g^2 = \sigma^2 + m\right)$ is $3\sigma_g^4$, or

$$3m^2 + 6m\sigma^2 + 3\sigma^4. \quad (43)$$

Comparing (42) with (43) and using the fact that $\sigma_4 \ge 3\sigma^4$, which was demonstrated previously, it is deduced that the fourth-order moment of the crosstalk noise is again larger than the fourth-order moment of its Gaussian approximation. In fact, the existence of m in (42) implies that, in this case, the shot noise has increased the difference between the fourth-order moment of the crosstalk noise and its Gaussian approximation.

This behavior is further illustrated in Fig. 4, where the logarithmic plot of the pdf of the crosstalk noise and its Gaussian approximation are presented for a 16×16 AWG with $X = -22$ dB and $P_{\rm in} = -24$ dBm. The pdfs are centered for convenience around $x = 0$. Shot and thermal noise are taken into account in both cases. The pdf of the crosstalk noise $f_{D|1}(x)$ was computed numerically using the fast Fourier transform (FFT) algorithm and the fact that $f_{D|1}(x)$ and $M_{D|1}(is)$ are related through a Fourier transform pair:

$$M_{D|1}(is) = E\{e^{isD}\} = \int_{-\infty}^{+\infty} f_{D|1}(x)\exp(isx)dx \quad (44)$$

and

$$f_{D|1}(x) = \int_{-\infty}^{+\infty} M_{D|1}(is)\exp(-isx)dx. \quad (45)$$

Since, as shown in Fig. 4, the tails of the Gaussian distribution are lower, it is evident that the Gaussian model underestimates the BER.

IV. Importance of the Different Effects Determining the Crosstalk Level

As can be seen from (12), for an AWG router with given phase errors in the grating arms δ_k, there are three random factors that determine the value of the interfering terms: the phase noise (ϕ_i), the polarization variations (θ_i), and the bit misalignment (which is related to the random time offset of each interferer τ_i). In Fig. 5, the BER resulting from taking into account different

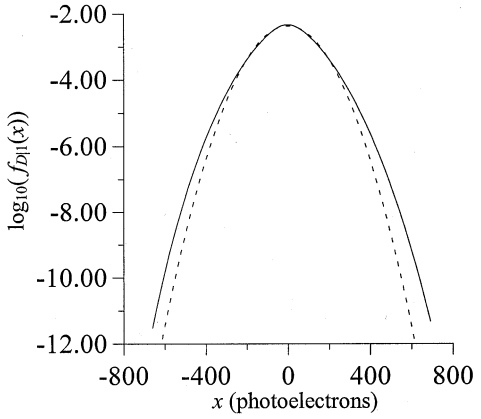

Fig. 4. Logarithmic plots of the exact (solid line) and approximate Gaussian (dashed line) pdfs $f_{D|1}(x)$, measured in (photoelectrons)$^{-1}$ and centered around $x = 0$. The mean sidelobe level of the AWG is $X = -22$ dB, and the input power is $P_{\rm in} = -24$ dBm.

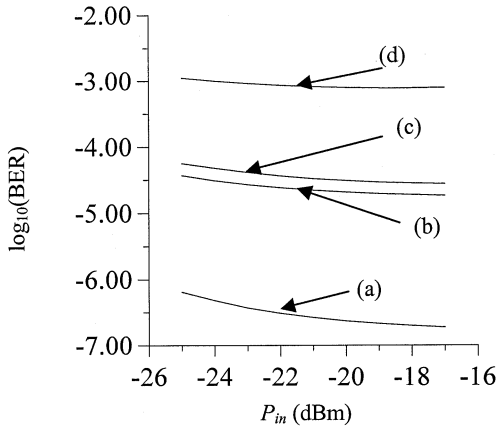

Fig. 5. The dependence of the BER as a function of the input powers $P_{\rm in}$ taking (a) polarization variations, phase noise, and misalignment, (b) polarization variations and phase noise only, (c) misalignment and either phase noise or polarization variations, and (d) either phase or polarization variations.

combinations of these factors is depicted for various values of $P_{\rm in}$ in the case of a 16×16 AWG with mean crosstalk $X = -22$ dB. From (12), it is evident that, since the phase noise ϕ_i and the polarization angle θ_i are identically distributed, they affect the statistical behavior of the interfering term i in the same way. Consequently, when only one of these random factors is taken into account in Fig. 5, it does not matter whether it is the phase or the polarization.

In Fig. 5, curve (a) corresponds to the existence of the combined effects of phase noise, polarization variations, and bit misalignments, and curve (b) corresponds to the existence of only the phase noise and the polarization variations in the statistics of the crosstalk noise. It is deduced that misalignment can decrease the BER to about two orders of magnitude and hence improve the performance of the network. This can be understood in terms of the noise power in these cases. When no misalignment is present $(\tau_i = 0)$, the power of the crosstalk contribution of the decision variable is given by

$$N_1 = \frac{1}{2} \sum_{i=1}^{N-1} \left| G_i^{11}(0) \right|^2 \tag{46}$$

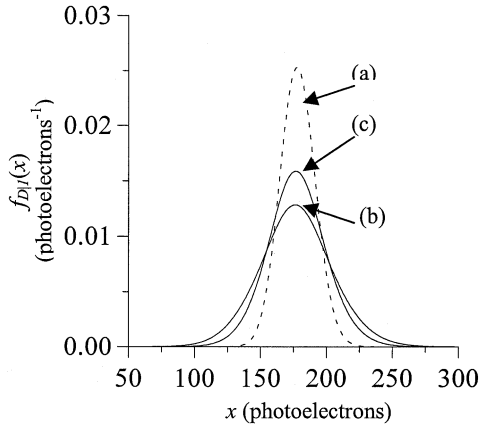

Fig. 6. The pdf of the decision variable D assuming that (a) no crosstalk is present, (b) there is crosstalk with no bit misalignment, and (c) there is crosstalk with bit misalignment. In all cases, shot noise is taken into account, but no thermal noise is assumed.

whereas when bit misalignment is present, the noise power is given by

$$N_2 = \frac{1}{4T} \sum_{i=1}^{N-1} \int_0^T \left| G_i^{11}(\tau_i) \right|^2 d\tau_i + \frac{1}{2T} \sum_{i=1}^{N-1} \int_0^T \left| G_i^{01}(\tau_i) \right|^2 d\tau_i. \tag{47}$$

These results can be derived in a way similar to the one used to derive σ_n^2 in (30). Alternatively, the noise powers can be calculated using the second-order derivative of the moment functions (22a) and (22b) at $s = 0$.

From Fig. 3, it is deduced that the majority of these functions take their maximum values around $\tau_i = 0$. The average of $\left| G_i^{ab}(\tau_i) \right|^2$ in (46) will be much smaller than $\left| G_i^{11}(0) \right|^2$ in (47) for the case of the RZ format. As a result, $N_1 < N_2$, and this explains why the BER is less when bit misalignment is present. The same behavior can be seen in Fig. 6, where the pdf of $D(T)$ has been plotted for $P_{\rm in} = -30$ dBm in three different cases. Pdf (a) corresponds to the case where only shot noise is present. Pdf (b) corresponds to the case of the crosstalk noise that includes bit misalignments, phase noise, polarization variations, and shot noise. Finally, pdf (c) corresponds to the same conditions as in (b) but without bit misalignment. Pdfs (b) and (c) are in agreement with the behavior of curves (a) and (b) of Fig. 5. It should be noted, however, that if a different input pulse format is used (ex. NRZ format), it is not guaranteed that bit misalignment will lead to a decrease in the BER.

In Fig. 5, curve (d) corresponds to the case where either phase noise or polarization variation is taken into account. By comparing curve (b) with curve (d), it is deduced that, when both of these random factors are present, the BER is reduced by at least one order of magnitude. This is because the standard deviation of the two cosines product appearing in (12) is equal to one-fourth, whereas if a single cosine term is present in (12) (when either phase or polarization variation is neglected), its standard deviation becomes one-half. In other words, the noise power is doubled, and the BER deteriorates.

In the case when both phase noise and polarization variation are neglected, however, the BER turns out to be much smaller

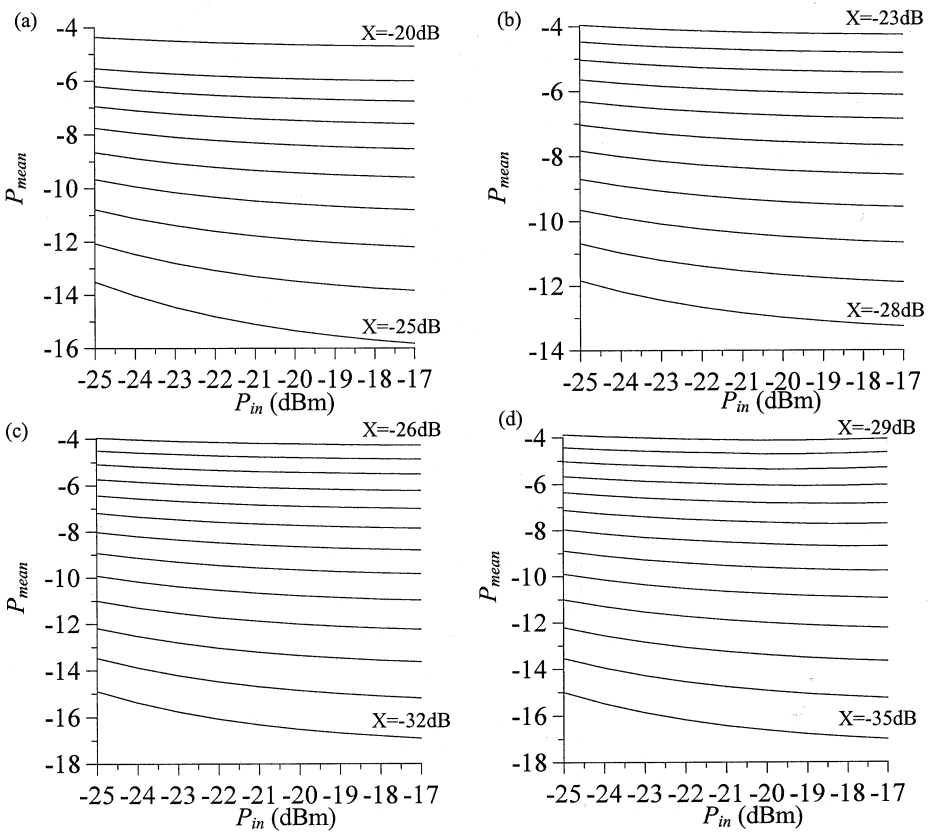

Fig. 7. The values of P_{mean} when (a) $N = 16$, (b) $N = 32$, (c) $N = 64$, (d) $N = 128$ for various values of the input power P_{in} in the AWG and the mean crosstalk level X. The difference in X for two adjacent curves in the same graph is 0.5 dB.

than in the previous cases. In this case, the statistics of the crosstalk noise is determined by the randomness of the interfering bits. By numerically calculating the pdf of the crosstalk noise under these conditions, it can be shown that the tails of the pdf are much lower than in the case where phase noise and polarization variations are taken into account. In addition, the amplitudes of the interferers are smaller in this case because only the real part of the functions G_i is present in (22e) and (22f).

V. BER Dependence on the Number of Channels

In order to assess the implications of crosstalk noise in the network's BER, a series of computer simulations have been performed. Various values for the expected sidelobe level X, for the input power P_{in}, and for the number of channels N have been considered. The MGFs used in the BER calculations included the contribution of polarization variations, phase noise, bit misalignment, and shot and thermal noise. The results are illustrated in Fig. 7(a)–(d) for $N = 16$, $N = 32$, $N = 64$, and $N = 128$, respectively. In each case, 25 different AWG transfer functions were created using Gaussian-distributed phase errors δ_k with the same standard deviation σ, which is related to the mean AWG crosstalk X through (16). The mean logarithmic BER, P_{mean} was computed from the BER's $P_{e,i}$ corresponding to each transfer function i, $1 \leq i \leq 25$ using

$$P_{\text{mean}} = \frac{1}{25} \sum_{i=1}^{25} \log_{10}(P_{e,i}). \tag{48}$$

As indicated by Fig. 7, a mean crosstalk level of -25 dB achieves very low BER when the number of channels is 16. However, as the number of channels increases to 128, mean crosstalk level equal to -35 dB is required to achieve the same BER. It is therefore imperative to use an AWG with lower crosstalk levels as the number of channels increases.

Another point of interest in these diagrams is that an increase in the input power (P_{in}) does not, in most cases, result in a drastic improvement of the BER. This is because the crosstalk contribution D_n to the decision variable D given by (6) depends linearly on the products $g_i^*(t)g_0(t)$. An increase in the input power by a factor of p is equivalent to the multiplication of each $g_i(t)$ by a factor of $\sqrt{p}$; therefore, the power of the signal (which is the expected value of D_s^2) and of crosstalk noise (which is the expected value of D_n^2) will both be multiplied by a factor of p^2. It is therefore expected not to have a drastic improvement in the BER.

VI. Conclusion

The implications of non-Gaussian AWG crosstalk noise on the performance of a passive $N \times N$ interconnection were examined using the saddle point method. A model for the calculation of the BER including the phase noise, polarization variations, and bit misalignments has been presented, and the importance of these factors has been evaluated. The thermal noise and the shot noise have also been incorporated into the model. The validity of the Gaussian approximation was examined, and it was shown that there are cases in which the Gaussian approximation does

not provide an accurate value for the BER. It was deduced that bit misalignments can lead to a decrease in the BER by approximately two orders of magnitude in such a network in the case of the RZ format. Finally, the performance of the network has been evaluated for various values of node number, input power, and mean AWG crosstalk. It was deduced that, as the number of channels increase, the mean AWG crosstalk must decrease significantly to achieve the same order of magnitude BER.

REFERENCES

[1] M. Smit and C. van Dam, "PHASAR-based WDM-devices: Principles, design and applications," *IEEE J. Select. Topics Quantum Electron.*, vol. 2, pp. 236–250, June 1996.

[2] C. Dragone, C. A. Edwards, and R. C. Kistler, "Intergrated optics $N \times N$ multiplexer on silicon," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 896–899, Oct. 1991.

[3] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *J. Select. Areas Commun.*, vol. 8, pp. 948–964, 1990.

[4] H. Takashi, Y. Hibino, and I. Nishi, "Polarization-insensitive arrayed-waveguide grating wavelength multiplexer on silicon," *Opt. Lett.*, vol. 17, pp. 499–501, 1992.

[5] C. Dragone, "Crosstalk caused by fabrication errors in a generalized Mach Zehnder interferometer," *Electron. Lett.*, vol. 33, no. 15, pp. 1326–1328, July 1997.

[6] K. Takada, H. Yamada, and Y. Inoue, "Origin of channel crosstalk in 100 GHz spaced silica-based arrayed waveguide grating multiplexer," *Electron. Lett.*, vol. 31, no. 14, pp. 1176–1178, July 1995.

[7] R. Ramaswami and K. N. Sivarajan, *Optical Networks a Practical Perspective*. San Mateo, CA: Morgan Kaufmann, 1998.

[8] H. Takahashi, K. Oda, and H. Toba, "Impact of crosstalk in an arrayed-waveguide multiplexer on $N \times N$ optical interconnection," *J. Lightwave Technol.*, vol. 14, pp. 1097–1105, June 1996.

[9] E. Iannone, R. Sabella, M. Avattaneo, and G. de Paolis, "Modeling of in-band crosstalk in WDM optical networks," *J. Lightwave Technol.*, vol. 17, pp. 1135–1141, July 1999.

[10] T. Gyselings, G. Morthier, and R. Baets, "Crosstalk analysis of multi-wavelegth optical cross connects," *J. Lightwave Technol.*, vol. 17, pp. 1273–1283, Aug. 1999.

[11] I. T. Monroy and E. Tangliongga, "Perfomance evaluation of optical crossconnects by saddlepoint approximation," *J. Lightwave Technol.*, vol. 16, March 1998.

[12] G. Einarsson, *Principles of Lightwave Communications*. New York: Wiley.

[13] R. Khosravani, M. I. Hayee, B. Hoanca, and A. E. Willner, "Reduction of coherent crosstalk in WDM add/drop multiplexing nodes by bit pattern misalignment," *IEEE Photon. Technol. Lett.*, vol. 11, no. 1, pp. 134–136, Jan 1999.

[14] H. Takahashi, K. Oda, H. Toba, and Y. Innoue, "Transmission characteristics of arrayed-waveguide $N \times N$ wavelength multiplexer," *J. Lightwave Technol.*, vol. 13, pp. 447–455, Mar. 1995.

[15] M. E. V. Segatto, G. D. Maxwell, R. Kashyap, and J. R. Taylor, "High-speed transmission and dispersion characteristics of an arrayed waveguide grating," *Opt. Commun.*, vol. 195, pp. 151–157, 2001.

[16] H. Yamada, K. Takada, Y. Inoue, and M. Horiguchi, "Measurement of the phase error distribution in silica based arrayed waveguide grating multiplexers by using Fourier transform spectroscopy," *Electron. Lett.*, vol. 31, no. 14, pp. 360–361, 1995.

[17] T. Goh, S. Suzuki, and A. Sugita, "Estimation of waveguide phase error in silica-based waveguides," *J. Lightwave Technol.*, vol. 15, pp. 2107–2213, November 1997.

Thomas Kamalakis, photograph and biography not available at time of publication.

Thomas Sphicopoulos (M'87), photograph and biography not available at time of publication.