An Estimation of Performance Degradation Due to Fabrication Errors in AWGs

Thomas Kamalakis, Thomas Sphicopoulos, *Member, IEEE*, and Dimitris Syvridis

***Abstract*—Arrayed waveguide gratings (AWGs) are important components for the realization of wavelength-division multiplexing optical networks. Their filtering performance is limited by the existence of phase errors in the grating waveguides due to fabrication imperfections. In this paper, the statistical properties of the phase errors are related to the waveguide imperfections using a variation of the effective index method. The filtering quality of the AWG is then investigated by considering the behavior of its transfer function in the presence of random phase errors. The probability density function of the transfer function's sidelobes is evaluated numerically, and the results are justified using theoretical considerations. Finally, the behavior of the maximum sidelobe level is also analyzed numerically, and universal diagrams are presented that allow the estimation of its mean value, standard deviation, and cumulative distribution function for every specific AWG.**

***Index Terms*—Crosstalk, gratings, integrated optics, optical filters, tolerance analysis, waveguide filters, wavelength-division multiplexing.**

I. Introduction

ARRAYED waveguide gratings (AWGs) [1], [2] are important components for the realization of modern optical communication networks employing wavelength-division multiplexing (WDM) [3] on which they can serve as wavelength multiplexers, demultiplexers, and routers. These devices have been made commercially available, and there are techniques for achieving polarization-insensitive operation [4]. There are also techniques that allow passband flattening [1]. The transfer function of the conventional AWG (without passband flattening) is of Gaussian shape at the passband and ideally has very low sidelobes, around -60 dB. Such a low sidelobe level would mean that networks employing AWGs would not suffer from the accumulation of in-band and out-band crosstalk noise at the receiver. Unfortunately practical AWGs exhibit a much higher sidelobe level, posing limitations in the achieved bit error rate (BER) of the receiver due to nonideal wavelength isolation.

Several studies have been made in order to understand the origin of crosstalk in AWGs [5]–[8]. It has been noted that it is mainly due to the phase errors between the different arms of the grating [6]. Although these studies focus on small phase errors, these can be quite significant in the case of an AWG with narrow channel spacing (such as 10 GHz). In addition, the statistics of the phase errors vary from arm to arm since the waveguide lengths are different. This is especially true for AWGs with small free spectral range (FSR) because the difference between the lengths of two adjacent grating waveguides ΔL is inversely proportional to the FSR.

In this paper, the behavior of the sidelobe level (directly related to the crosstalk) of a conventional AWG with Gaussian passband, and its relation to various design parameters and fabrication tolerances is described, even in the presence of larger phase errors with varying statistics from arm to arm. Waveguide fabrication errors, like nonvertical sidewalls [9], are taken into account with the effective refractive index method (ERIM). Next, the behavior of the transfer function of the AWG in the presence of the induced phase errors is analyzed. The statistical behavior of the sidelobes of the transfer function is shown, both by computer simulations and theoretical considerations, to be approximately that of an exponential random variable. Finally, using computer simulations, the behavior of the maximum sidelobe level is analyzed, and universal diagrams are given that enable the estimation, from the fabrication tolerances, of the mean value, the standard deviation, and the cumulative distribution function (CDF) of the maximum sidelobe level for most types of AWGs.

II. Relation Between Fabrication Imperfections and Phase Errors

As stated in [7], waveguide fabrication errors can produce a variation Δn_{eff} in the effective index n_{eff} of the fundamental TE and TM mode of a waveguide, thus causing phase errors in the grating arms of the AWG. Because of the etching process, the waveguide, while originally specified to be rectangular, is actually a trapezium [Fig. 1(a)], and there can be some deviation in the values of its width w and its height h. The refractive index of each region can also be different from its specified value due to errors in the gap wavelength λ_g. Fig. 1(b) shows the variation of n_{eff}, calculated using the ERIM for the fundamental TE mode of a waveguide which was specified to have core size 7 μm $\times$ 7 μm, core refractive index $n_1 = 1.5, n_2 = n_0$, and index contrast $\Delta = 0.75\%$, when only one of the seven parameters $p\,(=a, b, w, h, n_0, n_1, n_2)$ is modified. The ERIM is applied by segmenting, as shown in the Fig. 1(a), the waveguide in horizontal layers (typically 20 to 30). The width of each layer of the trapezium is assumed constant. The effective indexes of the various layers are found, and a vertical multilayered structure is obtained, from which the waveguide n_{eff}, both for the TE and the TM modes, is calculated as usual. Fig. 1(b) indicates that if the deviations Δp of the waveguide parameters p are small

Manuscript received March 8, 2002; revised June 7, 2002.

The authors are with the Department of Informatics and Telecommunications, University of Athens, Athens GR-15784, Greece.

Digital Object Identifier 10.1109/JLT.2002.802231

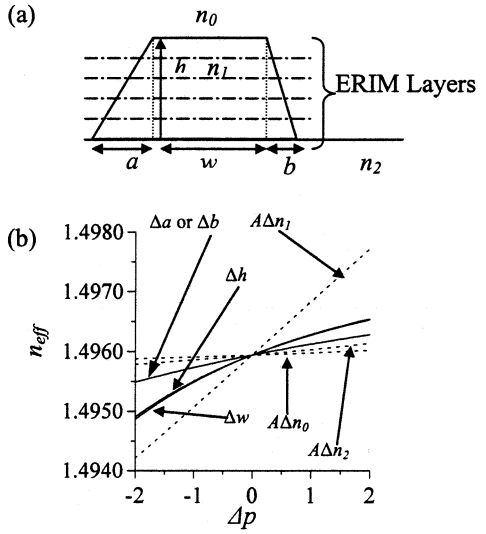

Fig. 1. (a) Geometry of a fabricated dielectric waveguide. (b) Variation of n_{eff} with respect to $\Delta p = (\Delta a, \Delta b, \Delta h, \Delta w, A\Delta n_0, A\Delta n_1, A\Delta n_2)$ with $A = 10^3$.

($\sim$0.1 μm for $w, a, b,$ and h and $\sim 10^{-4}$ for n_i), then Δn_{eff} is linear with respect to Δp; that is

$$\Delta n_{\text{eff}} = \frac{\partial n_{\text{eff}}}{\partial w}\Delta w + \frac{\partial n_{\text{eff}}}{\partial a}\Delta a + \frac{\partial n_{\text{eff}}}{\partial b}\Delta b + \frac{\partial n_{\text{eff}}}{\partial h}\Delta h + \frac{\partial n_{\text{eff}}}{\partial n_1}\Delta n_1 + \frac{\partial n_{\text{eff}}}{\partial n_2}\Delta n_2 + \frac{\partial n_{\text{eff}}}{\partial n_0}\Delta n_0. \quad (1)$$

As a result, the variance of ${\sigma_{n_{\text{eff}}}}^2$ of Δn_{eff} can be written as

$$\left(\sigma_{n_{\text{eff}}}^{\text{TE,TM}}\right)^2 = \sum_p \left(\partial n_{\text{eff}}^{\text{TE,TM}} \Big/ \partial p\right)^2 \sigma_{\Delta p}^2 \quad (2)$$

where $\sigma_{\Delta p}$ is the standard deviation of Δp. The values of the derivatives of (1) and (2) can be calculated using the above method. The deviations of the waveguide parameters Δp can be assumed Gaussian random variables with zero mean value ($\langle \Delta p \rangle = 0$) for $p \neq a, b$ and standard deviations $\sigma_{\Delta p}$ expressing the fabrication tolerances that can be roughly estimated directly or indirectly [7]. Because of the linearity mentioned previously and because a linear combination of independent Gaussian random variables is also a Gaussian random variable, Δn_{eff} will be Gaussian with standard deviation given by (2). The mean value of Δn_{eff}, which is due only to Δa and Δb, is small and does not play any significant role in the results. Hence, $\langle \Delta n_{\text{eff}} \rangle$ will be assumed equal to zero both in the theoretical calculations and the numerical simulations in this paper.

In most AWGs, the bending radii of the grating waveguides are large and do not affect the value of the derivatives in (2). However, if the bending radii become small, more elaborate two-dimensional (2-D) mode solvers [10], along with conformal mapping, must be used to determine these derivatives.

III. Effect of Waveguide Coupling

The above method treats each waveguide separately without including the presence of its adjacent waveguides. One might expect that, since adjacent grating arms are coupled, the phase errors in one waveguide could affect the phase error on its neighboring waveguides, or in other words, that coupling in the grating arms could correlate the phase errors.

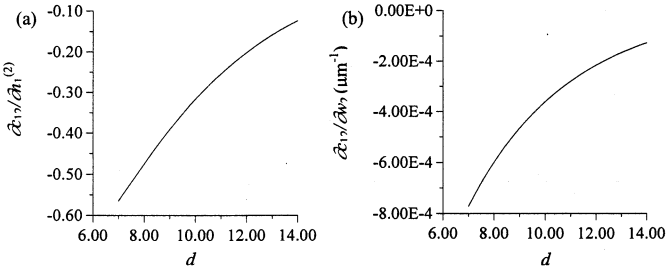

Fig. 2. Variation of the coupling coefficient's derivatives with respect to the center-to-center distance d of two waveguides.

The effect of waveguide coupling can be estimated using coupled-mode theory [11]. Assuming the effects of only adjacent waveguides, the complex amplitude $A_1(z)$ of the signal in a waveguide obeys

$$\frac{d}{dz}A_1(z) = -j\beta A_1(z) - jc_{10}A_0(z) - jc_{12}A_2(z) \quad (3)$$

where $A_1(z)$ is the complex amplitude of the considered waveguide, $A_0(z)$ and $A_2(z)$ are the complex amplitudes of its two adjacent waveguides, c_{10} and c_{12} are the corresponding coupling coefficients, while $\beta = 2\pi n_{\text{eff}}/\lambda$ is the propagation constant of the waveguide in the uncoupled case. Dividing by $A_1(z)$ and integrating with respect to z, the following result is obtained:

$$A_1(z) = A_1(0)\exp\left(-j\beta z - j\int_0^z c_{10}(z')\frac{A_0(z')}{A_1(z')}\,dz' - j\int_0^z c_{12}(z')\frac{A_2(z')}{A_1(z')}dz'\right). \quad (4)$$

From (4), it is deduced that the phase of the optical signal in a grating waveguide depends on the complex signal amplitudes of the adjacent waveguides and the coupling coefficients. A change in the parameters of the adjacent waveguides will affect the phase of the optical signal through the coupling coefficients c_{10} and c_{12}. These coefficients are functions of z since in AWGs, the grating waveguides start moving apart in the region after the first star coupler and move close to each other at the input of the second star coupler.

The derivatives of the coupling coefficients with respect to the waveguide parameters can be used to measure the change in the coupling coefficients and, hence, the phase-error correlation. If the width w_2 of waveguide 2 changes by a small amount Δw_2, then the change of the coupling coefficient c_{12} will be linear and will be given by $\Delta w_2(\partial c_{12}/\partial w_2)$. The induced phase change $\Delta\varphi_{12}$ in the signal of waveguide 1 will be

$$\Delta\phi_{12} = \int_0^z \frac{\partial c_{12}}{\partial w_2}\frac{A_2(z)}{A_1(z)}\Delta w_2\,dz. \quad (5)$$

Fig. 2 depicts the variation of the derivative $\partial c_{12}/\partial n_1^{(2)}$ and $\partial c_{12}/\partial w_2$ of the coupling coefficient c_{12} of two waveguides having core widths $w_1 = w_2 = 4\mu$m, core heights

$h_1 = h_2 = 4\mu$m, and core indexes $n_1^{(1)} = n_1^{(2)} = 1.5$ (where $n_1^{(i)}$ is the core index of the ith waveguide), calculated using the effective index method. The relative difference between the core indexes and the surrounding index is $\Delta = 0.33\%$ $(n_2 = 1.495)$, and the distance between the centers of the waveguides is initially $d = 7\,\mu$m (resulting in a gap of $3\,\mu$m between the cores). The corresponding values of the derivatives of the propagation constant β of the waveguides are $\partial\beta/\partial n_1^{(1)} = 1.29$, $\partial\beta/\partial w_1 = 8 \times 10^{-4}\ \mu\text{m}^{-1}$ and will be used to compare the change in the phase introduced by the change in β and by the change in the coupling coefficients. From Fig. 2, it is deduced that although the derivatives of the coupling coefficients are initially of the same order as the derivatives of the propagation constant, they gradually diminish as the waveguides move further apart. In addition, since the waveguides typically remain uncoupled in the curved section of the grating, it is expected that the phase error in waveguide 1, caused by the fabrication imperfections of the adjacent waveguide 2, will be small compared with the phase error caused by the fabrication imperfections of waveguide 1. It should also be noted that, in this example, the waveguides are placed close together, and the index contrast Δ is low, which implies that the coupling is somewhat strong. In practical AWGs, where a sharper index contrast Δ should be preferred [12] in order to reduce the chip size and where the grating waveguides often have larger initial spacing, the coupling-induced phase-error correlations will be even lower. As a result, the phase errors will be assumed uncorrelated throughout this paper.

IV. Statistics of the Phase Errors

Once $\sigma_{n_{\text{eff}}}$ is estimated by the fabrication tolerances $\sigma_{\Delta p}$, the standard deviation σ_k of the phase error δ_k of the kth waveguide can be calculated by

$$\sigma_k = 2\pi\sigma_{n_{\text{eff}}} L_k/\lambda \tag{6a}$$

where λ is the central wavelength and L_k the length of the kth arm $(0 \le k \le M-1)$. By replacing $L_k = L_0 + k\Delta L$ in (6a), where L_0 is the length of the shortest arm, the following equation is obtained for the standard deviations of the grating arms:

$$\sigma_k = \sigma_0 + k\Delta\sigma/(M-1) \tag{6b}$$

where $\Delta\sigma = 2\pi(M-1)\sigma_{n_{\text{eff}}}\Delta L/\lambda, \Delta L = c/(n_{\text{eff},o}\text{FSR})$ is the difference in the lengths of two consecutive arms, FSR the free spectral region, c the velocity of light in vacuum, $n_{\text{eff},o}$ the nominal effective refractive index of the fundamental waveguide mode, and M the number of grating arms. It is deduced by (6a) and (6b) that the phase error in each arm can have different standard deviation if ΔL is not negligible compared with L_0, as in the case of narrow channel spacing (e.g., 10 GHz).

Equation (6) does not account for the phase error induced by the imperfect photomask resolution. The effect of the photomask resolution can be incorporated by adding to the lengths of the waveguides a small random length Δl_m with mean value equal to zero and standard deviation equal to the resolution of the photomask σ_{ph}. The equation giving the standard deviation will change and will become

$$\sigma_k = \frac{2\pi}{\lambda}\sqrt{\sigma_{n_{\text{eff}}}^2 L_k^2 + n_{\text{eff},o}^2\sigma_{\text{ph}}^2}\,. \tag{7}$$

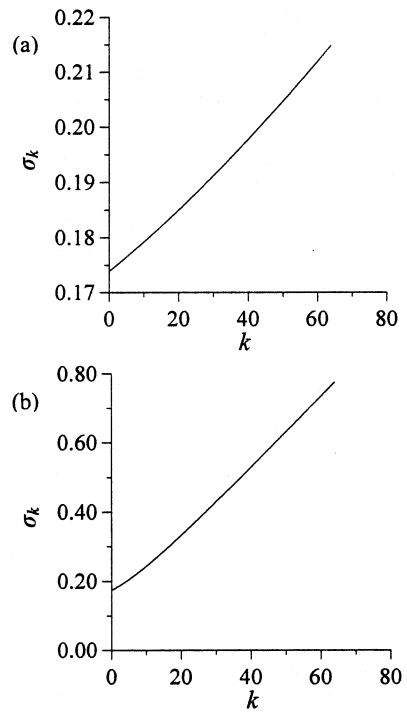

Fig. 3. The variation of σ_k with respect to k for the cases of a 16×16 AWG with (a) 100-GHz spacing and (b) 10-GHz spacing assuming $\sigma_{n_{\text{eff}}} = 2.1 \times 10^{-6}$, $\sigma_{\text{ph}} = 0.025\,\mu$m, $L_0 = 10^4\,\mu$m, and $M = 65$.

This equation exhibits a linear behavior with respect to k for practical values of its parameters. Using as an example the values of the fabrication tolerances measured in [7], which are $\sigma_{n_1} = \sigma_{n_2} = 2 \times 10^{-6}$ for the variation of the material refractive index and $\sigma_w = \sigma_h = 2 \times 10^{-3}\ \mu$m for the variations in the waveguide core dimensions, the value of $\sigma_{n_{\text{eff}}}$, calculated using the ERIM, is equal to 2.1×10^{-6}. The values of σ_k obtained by (7) are plotted in Fig. 3(a), for $L_0 = 10^4\,\mu$m, which is a typical value for the shortest grating arm, $\Delta L = 125\,\mu$m, which corresponds to a 16×16 AWG with channel spacing equal to 100 GHz, and photomask resolution $\sigma_{\text{ph}} = 0.025\,\mu$m [8]. The AWG is assumed to have 65 waveguides, while $n_{\text{eff},o}$ is taken approximately equal to 1.5. In Fig. 3(b), σ_k is also plotted for a 10-GHz spacing 16×16 AWG with the same number of waveguides and $\Delta L = 1250\,\mu$m. It is seen that in both cases, σ_k is approximately linear and can be written in the form of (6b) using curve fitting. As a result, the standard deviations σ_k of the phase errors in the grating arms will obey (6b) and depend on three parameters: the initial phase error σ_0, the difference $\Delta\sigma$ between the largest and the smallest value of σ_k, and the number of arms M.

V. Statistical Behavior of the Transfer Function

The transfer function between the central input and output ports of an AWG in the presence of the phase errors δ_k is given by [5], as follows:

$$H(v) = \sum_{k=0}^{M-1} C_k \exp(j2\pi kv)\exp(j\delta_k) \tag{8}$$

where C_k is the optical power of the kth grating waveguide normalized to the total power and $v = (f - f_0)/\text{FSR}$ is the normal-

ized optical frequency, f being the actual optical frequency and f_0 the central frequency of the AWG. In a conventional AWG, the coefficients C_k obey a Gaussian law, and in practical AWGs, the ratio $R_t = \min(C_k)/\max(C_k)$ should be very small to ensure that the sidelobes of the AWG in the absence of phase errors ($\delta_k = 0$) would be very low. Throughout the numerical simulations in this paper, the ratio R_t is set to be 2% in which case the sidelobes of the AWG, in the absence of phase errors, are about -60 dB. Different values of this ratio do not alter the sidelobe-level behavior noticeably. By varying R_t from 1% to 5% for $M = 50$, only a 0.5-dB variation in the maximum sidelobe level was observed considering a variety of different values for σ_0 and $\Delta\sigma$. This fact is also demonstrated in [8]. From the above remarks and (6a) and (6b), together with the frequency normalization by FSR, it is deduced that the statistical behavior of $H(v)$ of conventional AWGs will practically depend only on $\Delta\sigma, \sigma_0$ and M. In the case of flattened AWGs the distribution of C_k is no longer Gaussian, and as a result, the statistical behavior of the sidelobe level may be different from that of conventional AWGs.

The mean value of the transmittance $T(v) = |H(v)|^2$ between the central input port and the central output port of the AWG, is given by

$$\langle T(v)\rangle = \sum_{k,l=0}^{M-1} C_k C_l \exp(j2\pi(k-l)v)\langle\exp(j(\delta_k - \delta_l)\rangle. \tag{9a}$$

Since the phase errors are independent random variables, $\langle\exp(j2\pi(\delta_m - \delta_n))\rangle = \langle\exp(j2\pi\delta_m)\rangle\langle\exp(-j2\pi\delta_n)\rangle$ if $m \neq n$, and $\langle\exp(j2\pi(\delta_m - \delta_n))\rangle = 1$ if $m = n$. The phase errors δ_k are Gaussian random variables with $\langle\delta_k^2\rangle = \sigma_k^2$, and using the fact that $\langle\exp(j\delta_k)\rangle = \exp(-\sigma_\kappa^2/2)$, then $\langle T(v)\rangle$ can be written as

$$\begin{aligned}\langle T(v)\rangle &= \sum_{k\neq l} C_k C_l \exp(j2\pi(k-l)v) \\ &\quad\times \exp\big(-1/2t(\sigma_k^2+\sigma_l^2)\big) + \sum_{k=0}^{M-1} C_k^2 \\ &= \left|\sum_{k=0}^{M-1} C_k \exp\big(-\sigma_k^2/2\big)\exp(j2\pi kv)\right|^2 \\ &\quad + \sum_{k=0}^{M-1} C_k^2\big(1-\exp(-\sigma_k^2)\big).\end{aligned} \tag{9b}$$

The third part of (9b) is derived from the second part by adding $\sum C_k^2\exp(-\sigma_k^2)$ from the first sum and subtracting it from the second. As seen by the third part of (9b), the mean value of $T(v)$ can be decomposed into two parts: one that is frequency-dependent, and one that is constant. If the phase errors have all the same standard deviation $\sigma_k = \sigma$, then (9b) is reduced to

$$\langle T(v)\rangle = \exp(-\sigma^2)T_{\text{ideal}}(v) + \big(1-\exp(-\sigma^2)\big)\sum_{k=0}^{M-1} C_k^2 \tag{10}$$

where $T_{\text{ideal}}(v)$ is the transmittance of the AWG with no phase errors ($\delta_k = 0$). Since the sidelobes of the ideal transmittance $T_{\text{ideal}}(v)$ of the AWG are practically negligible, it is deduced that outside the main lobe of the transfer function the mean value of $T(v)$ is independent of the frequency v, as follows:

$$\langle T(v)\rangle \cong \big(1-\exp(-\sigma^2)\big)\sum_{k=0}^{M-1} C_k^2. \tag{11a}$$

This means that the sidelobes of $T(v)$ have constant mean value. In the case where the phase errors do not have the same standard deviation, the mean value of the sidelobes is written

$$\langle T(v)\rangle \cong \sum_{k=0}^{M-1} C_k^2\big(1-\exp(-\sigma_k^2)\big) \tag{11b}$$

since the first sum in (9b) can again be ignored. This is because it is equal to the transmittance $T_i(v)$ of an ideal AWG (with no phase errors) with power distribution equal to $C_k\exp(-\sigma_k^2/2)$. Since the sidelobes of the ideal transfer function are caused by the discontinuities of the power distribution, and $C_k\exp(-\sigma_k^2/2)$ is a smooth function of k, it is deduced that the sidelobes of $T_i(v)$ are very low.

At the central frequency ($v = 0$) of the AWG, the mean value of the transmittance, assuming $\sigma_k = \sigma$, is given by

$$\langle T(0)\rangle \cong \exp(-\sigma^2) \tag{12}$$

since, in this case, the first term in (10) dominates, and $T_{\text{ideal}}(0) = 1$. Equation (12) implies that the phase errors can cause some additional insertion losses. For $\sigma = \pi/10$, these insertion losses are -0.42 dB, but for $\sigma = \pi/4$, they become -2.67 dB. Since, theoretically, the fiber-to-fiber losses of an ideal router can be kept below -0.5 dB with proper waveguide design [13], the insertion losses due to phase errors could have an important contribution to the total insertion losses of the device.

After considerable mathematical manipulation, a very lengthy expression for the standard deviation $\text{std}(T(v))$ of the transmittance can also be derived (see Appendix). In Fig. 4(a),

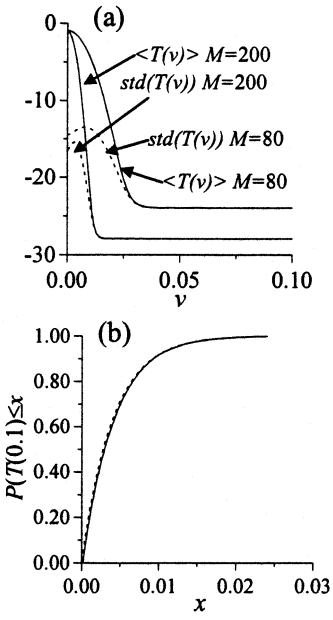

Fig. 4. (a) The mean value and the standard deviation of the transmittance of two AWGs having $M = 80$ and $M = 200$. In both cases, $\sigma_0 = \pi/10$ and $\Delta\sigma = \pi/10$. (b) The CDF of $T(0.1)$ for $M = 80$, $\sigma_0 = \pi/10$, and $\Delta\sigma = \pi/10$ (solid line) and the CDF of an exponential distribution with mean value equal to $\langle T(0.1)\rangle$ (dashed line).

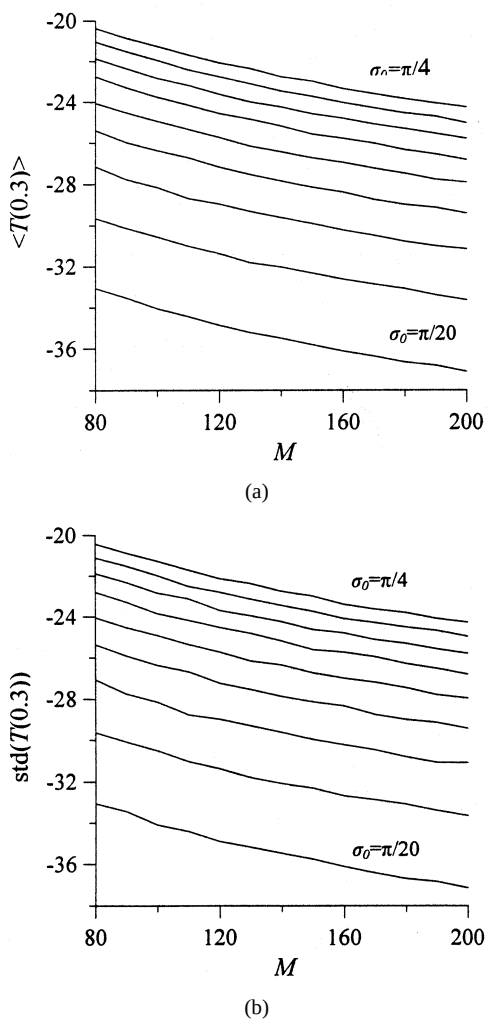

Fig. 5. (a) The mean value of the transfer function $T(v)$ at $v = 0.3$, and (b) the standard deviation of $T(v)$ at $v = 0.3$ for various values of the phase error σ_0 and the number of waveguides M.

the mean value and the standard deviation of $T(v)$ with $M = 80$ and $M = 200$ are plotted, assuming that $\sigma_0 = \pi/10$ and $\Delta\sigma = \pi/10$. It is obvious that the standard deviation and the mean value of $T(v)$ outside the main lobe are equal with a high degree of accuracy. This is a general result and suggests that the sidelobes are random variables with exponential distribution. In Fig. 4(b), the CDF, i.e., $P(T(v) \leq x)$, of the transmittance $T(v)$ of an AWG at $v = 0.1$ (solid line), calculated by (8) using computer simulations and the CDF of an exponential distribution given by $1-\exp(-x/\langle T(v)\rangle)$, with mean value $\langle T(0.1)\rangle$ given by (9) (dashed line), are plotted for the case of $M = 80, \sigma_0 = \pi/10$, and $\Delta\sigma = \pi/10$. From the excellent agreement between the two curves, it is deduced that the CDF of the sidelobes of $T(v)$ at a given frequency v can be calculated using an exponential approximation. This result seems to be general, and it is further illustrated in Fig. 5(a) and (b), where the mean value and the standard deviation of $T(v)$ at $v = 0.3$ have been plotted for various values of M and σ_0, assuming that $\Delta\sigma = 0$. These quantities have been calculated using computer simulations by generating 1000 samples of the transfer function in each case with Gaussian-distributed phase errors. By comparing Fig. 5(a) and (b), it is confirmed that the mean value and the standard deviation of $T(0.3)$ can be considered approximately equal, implying that the distribution of $T(0.3)$ will resemble an exponential distribution.

This behavior can also be theoretically justified. For small phase errors, the terms $\exp(j\delta_m)$ can be approximated by $1 + j\delta_m$, and the transfer function can be written as

$$H(v) = \sum_{m=0}^{M-1} C_m \exp(j2\pi mv) + j\sum_{m=0}^{M-1} C_m\delta_m \exp(j2\pi mv). \tag{13}$$

The first sum is the transfer function of the AWG without phase errors, which has negligible sidelobes. Consequently, the sidelobes of the transmittance $T(v)$ will be caused by the second sum

$$T(v) = |H(v)|^2 \cong \left|\sum_{m=0}^{M-1} C_m\delta_m \exp(j2\pi mv)\right|^2. \tag{14}$$

The real part $R(v)$ and the imaginary part $Y(v)$ of $H(v)$ are approximately given by

$$R(v) = \sum_{m=0}^{M-1} C_m\delta_m \sin(2\pi mv) \tag{15}$$

and

$$Y(v) = \sum_{m=0}^{M-1} C_m\delta_m \cos(2\pi mv). \tag{16}$$

Since $Y(v)$ and $R(v)$ are linear combinations of the Gaussian random variables δ_m, their joint distribution will also be Gaussian. The variances of $R(v)$ and $Y(v)$ are given by

$$\begin{aligned}\langle R^2(v)\rangle &= \sum_{m=0}^{M-1} C_m^2\sigma_m^2 \sin^2(2\pi mv)\\ &= \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 - \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 \cos(4\pi mv)\end{aligned} \tag{17a}$$

and

$$\begin{aligned}\langle Y^2(v)\rangle &= \sum_{m=0}^{M-1} C_m^2\sigma_m^2 \cos^2(2\pi mv)\\ &= \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 + \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 \cos(4\pi mv)\end{aligned} \tag{17b}$$

while their correlation is given by

$$\begin{aligned}\langle R(v)Y(v)\rangle &= \sum_{m=0}^{M-1} C_m^2\sigma_m^2 \cos(2\pi mv)\sin(2\pi mv)\\ &= \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 \sin(4\pi mv).\end{aligned} \tag{18}$$

By inspecting (17), it is deduced that the variances of $R(v)$ and $Y(v)$ consist of a constant term, which is the first sum in the right-hand side (RHS) of the equations and a transfer function $H_d(2v)$ with a smooth power distribution equal to $C_m^2\sigma_m^2$. Un-

less v is near $1/2$ or 0, this transfer function can be neglected compared with the constant term, and thus

$$\langle R^2(v)\rangle = \langle Y^2(v)\rangle = \frac{1}{2}\sum_{m=0}^{M-1} C_m^2\sigma_m^2 = \frac{\langle T(v)\rangle}{2}. \quad (19)$$

Using the same reasoning, it can be deduced that

$$\langle R(v)Y(v)\rangle \ll \langle R^2(v)\rangle = \langle Y^2(v)\rangle \quad (20)$$

and, consequently, that if v is not near $1/2$ or 0, $R(v)$ and $Y(v)$ can be regarded as independent Gaussian random variables with equal standard deviations. By assuming the transformation

$$R(v) = T(v)\cos W(v) \quad (21a)$$

$$Y(v) = T(v)\sin W(v) \quad (21b)$$

with $W(v) \in [-\pi\ \pi]$ and applying the theorem of transformation of random variables [14], the combined probability density function (pdf) $f_{T,W}$ of $T(v)$ and $W(v)$ will be given by

$$f_{T,W}(t,w) = f_{R,Y}(r,y)|J|. \quad (22)$$

In the above equation, $f_{R,Y}$ is the combined pdf of $R(v)$ and $Y(v)$, and J is the determinant of the Jacobian matrix of transformation (21). By carrying out the mathematical calculations it can be shown that, since R and Y are approximately independent Gaussian random variables with equal standard deviations, the pdf f_T of T is given by

$$f_T(t) = \frac{1}{\langle T(v)\rangle}\exp\left(-\frac{t}{\langle T(v)\rangle}\right) \quad (23)$$

and, as a result, the sidelobes of $T(v)$ exhibit the behavior of an exponential distribution.

VI. Statistical Behavior of the Maximum Sidelobe Level

Another characteristic measure of crosstalk degradation is the maximum sidelobe level $T_{\max}$ of the transmittance. To analyze its behavior for various values of $\sigma_0, \Delta\sigma$, and M, a large number of sample transfer functions with random Gaussian phase errors have been generated for each case, and $T_{\max}$ was measured for each sample. The mean value μ, the standard deviation s, and the CDF $P(T_{\max} \le x)$ of $T_{\max}$ have been evaluated based on the results. An interesting result observed by the simulations is that μ and s (in decibels) can be approximately written as the sum of two functions, one depending on σ_0 and $\Delta\sigma$, and the other depending on M, that is

$$\mu(M,\sigma_0,\Delta\sigma) \cong \mu_0(M) + \mu_1(\sigma_0,\Delta\sigma) \quad (\text{in dB}) \quad (24a)$$

and

$$s(M,\sigma_0,\Delta\sigma) \cong s_0(M) + s_1(\sigma_0,\Delta\sigma) \quad (\text{in dB}). \quad (24b)$$

Equations (24a) and (24b) are derived by investigating the results of the simulations. The functions μ_0 and s_0 are calculated by averaging out the functions μ and s for $\Delta\sigma$ and σ_0, respectively. In Fig. 6(a), μ_0 and s_0 are plotted, while in Fig. 6(b) and (c), the same is done for the functions μ_1 and s_1. Using

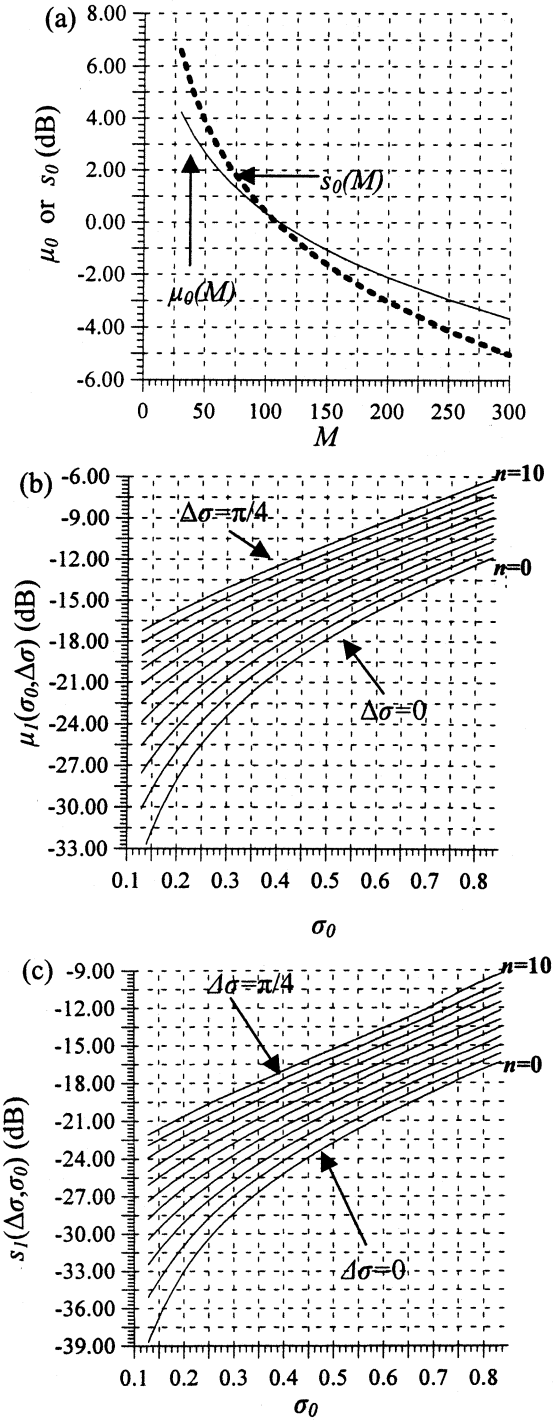

Fig. 6. (a) The functions $\mu_0 = \mu_0(M)$ (solid line) and $s_0 = s_0(M)$ (dashed line) (in dB), (b) the function $\mu_1 = \mu_1(\sigma_0, \Delta\sigma)$ (in dB), and (c) the function $s_1 = s_1(\sigma_0, \Delta\sigma)$ (in dB), for $\Delta\sigma = \Delta\sigma_n = n\pi/40$ where n designates the curves from the lower $(n = 0)$ to the higher $(n = 10)$.

the diagrams in Fig. 6, the values of μ and s can be calculated by applying (24a) and (24b) for every AWG once M, σ_0, and $\Delta\sigma$ are determined. In order to illustrate the accuracy of (24a) and (24b), the exact and approximate values of μ and s have been plotted in Fig. 7 for $\sigma_0 = \pi/8, \Delta\sigma = \pi/8$ and $\sigma_0 = \pi/4, \Delta\sigma = \pi/4$. It is deduced that the accuracy is satisfactory in all cases and improves considerably when the mean crosstalk is not too large.

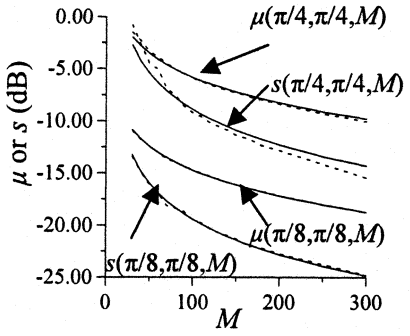

Fig. 7. The values of μ and s calculated with numerical simulations (dashed line) and their values calculated using the functions μ_1, μ_0 and s_1, s_0 using approximation (24) (solid line) for $\sigma_0 = \pi/8, \Delta\sigma = \pi/8$ and $\sigma_0 = \pi/4$, $\Delta\sigma = \pi/4$.

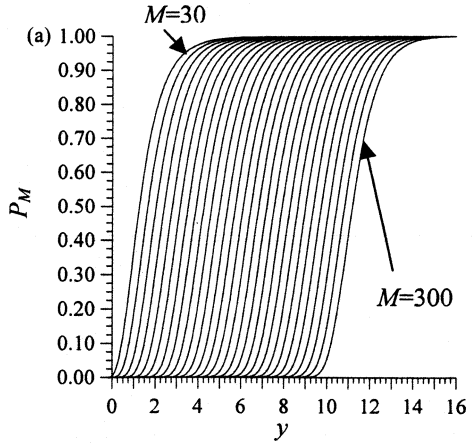

Fig. 8. The functions P_M used in the calculation of the CDF of $T_{\max}$. The left-most curve is for $M = 30$, the second-left curve for $M = 40$, and so on, up until the right-most, which is for $M = 300$.

By investigating the result of the simulations performed, it was observed that the shape of the CDF of $T_{\max}$, i.e., $P(T_{\max} \leq x)$, was primarily dependent on the number of arms M. In particular, by using the transformation $y = (x - \mu)/s$, the transformed CDF was observed mainly to depend on M and was not strongly influenced by σ_0 and $\Delta\sigma$, that is

$$P(T_{\max} \leq x) \cong P_M(y). \tag{25}$$

The functions $P_M(y)$, depending only on M and y, are plotted in Fig. 8, each displaced artificially on the horizontal axis by tM. The value of t is equal to the one tenth of the average value of the ratio between the mean value of the maximum sidelobe level and its standard deviation $\langle \mu/s \rangle/10 = 0.036$ and is chosen so that the graphs of Fig. 8 become distinguishable. The exact (dashed line) and approximate (solid line) CDFs have been plotted in Fig. 9(a) and (b) for $\sigma_0 = \pi/8, \Delta\sigma = \pi/8$, and $M = 80$ and $\sigma_0 = \pi/4, \Delta\sigma = \pi/4$, and $M = 50$, respectively. The mean values μ and standard deviation s were calculated using expressions (24a) and (24b) and the diagrams of Fig. 6. Then, the approximate CDFs were computed using the functions P_{50} and P_{80} of Fig. 8 by setting $x = sy + \mu - Mts$. As before, it is observed that the approximation is quite accurate when the mean crosstalk is not too large. Even in the presence of higher crosstalk, the approximation remains satisfactory.

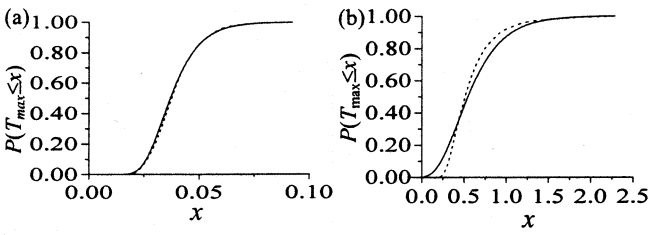

Fig. 9. The numerically calculated CDFs (dashed lines) of $T_{\max}$ in the case of (a) $\sigma_0 = \pi/8, \Delta\sigma = \pi/8$, and $M = 80$ and (b) $\sigma_0 = \pi/4, \Delta\sigma = \pi/4$, and $M = 50$ and their approximations (solid line) using the functions P_M using (24).

VII. AWG Crosstalk Evaluation Procedure

The diagrams of Figs. 6(a), (b), and 8 can be used to estimate the crosstalk performance for any AWG given the design parameters and the waveguide fabrication tolerances. In particular, these fabrication tolerances, expressed in terms of $\sigma_{\Delta p}$, can be used in order to determine $\sigma_{n_{\text{eff}}}$ by applying (2) and the ERIM. For each AWG, the parameters σ_0 and $\Delta\sigma = \sigma_{M-1} - \sigma_0$ can be calculated by applying (6). The CDF of the sidelobes will be that of an exponentially distributed random variable with mean value given by (9). The mean value and standard deviation of $T_{\max}$ can also be found using Fig. 6, (24a), and (24b). Finally, the CDF of $T_{\max}$ can be found from (25) and Fig. 8. This procedure can be applied for the TE and TM polarization separately, for every type of AWG with given geometrical characteristics ($L_0, \Delta L$, and M) by using the same diagrams obtained in this paper. For example, in Section IV, the values of the fabrication tolerances reported in [7] were used to calculate the standard deviation of the effective index for the TE mode, which was found to be $\sigma_{n_{\text{eff}}} = 2.1 \times 10^{-6}\,\mu\text{m}^{-1}$. Using $\sigma_{\text{ph}} = 0.025\,\mu\text{m}$ as in [8], $n_{\text{eff},o} \cong 1.5$, which is an approximate value for the effective index of the fundamental TE mode, and applying (7) for $k = 0$, we find $\sigma_0 = 0.17$ rad. For a 16×16 AWG with $M = 65$ waveguides and $\Delta L = 125\,\mu\text{m}$, applying (7) for $k = M - 1 = 64$, we have $\sigma_{M-1} = \sigma_{64} = 0.22$ rad and $\Delta\sigma = \sigma_{M-1} - \sigma_0 = 0.05$ rad. The values of $\Delta\sigma$ and σ_0 calculated in this manner are very close to the ones obtained by fitting σ_k of Fig. 3(a) with a linear curve. The expected maximum sidelobe level can now be calculated from the functions μ_0 and μ_1. From Fig. 6(a), one finds $\mu_0(M) = \mu_0(65) \cong 2.5\,\text{dB}$. The integer n for which $n(\pi/40)$ is closest to $\Delta\sigma$ is $n = 1$. Using the curve of Fig. 6(b) corresponding to $n = 1$, we obtain $\mu_1(\sigma, \Delta\sigma) = \mu_1(0.17, 0.05) \cong \mu_1(0.17, \pi/40) \cong -25\,\text{dB}$. Consequently, the expected maximum sidelobe level is $\mu = \mu_0 + \mu_1 \cong -22.5\,\text{dB}$. With a similar procedure, the standard deviation of the maximum sidelobe level turns out to be approximately $s \cong -27.5\,\text{dB}$. Finally, the CDFs of the maximum sidelobe level can be used to further appreciate the impact of the fabrication tolerances. For example, since $P_{65}(y) \cong P_{70}(y) = 80\,\%$ for $y \cong 3.8$ and $x = sy + \mu - Mts = 0.0082 \cong -20\,\text{dB}$, we expect that 80% of the fabricated devices will have maximum sidelobe level below $-20\,\text{dB}$.

In this paper, we have concentrated on the direct problem, which is the performance estimation of the fabricated AWGs given their fabrication tolerances. The results presented in this

work, can also be of use to a designer who wants to fabricate AWGs with a specified performance and has to choose from a number of fabrication processes with known tolerances ($\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$). For each fabrication process, the parameters σ_0 and $\Delta\sigma$ can be computed from $\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$ using (7). Having calculated these parameters, Figs. 6 and 8 can be used to estimate the performance of each fabrication process and determine which of them yields devices that fulfill a specified performance criterion.

It should be noted, however, that the diagrams presented in this paper can not be used directly to solve the inverse problem, that is the determination of the fabrication tolerances ($\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$) required in order for the fabricated AWGs to have a specified performance. For example, suppose that a designer wants to determine the fabrication tolerances required in order to ensure that the expected maximum sidelobe level μ is below -25 dB. In order to determine the maximum allowable values of $\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$, equation $\mu(\sigma_0, \Delta\sigma, M) = -25\,{\rm dB}$ must first be solved numerically in order to determine σ_0 and $\Delta\sigma$. The obtained values of σ_0 and $\Delta\sigma$ can then be used to determine $\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$ using (7). Since, however, $\mu(\sigma_0, \Delta\sigma, M) = -25\,{\rm dB}$ can have many solutions, there will be many values $\sigma_{n_{\rm eff}}$ and $\sigma_{\rm ph}$ that fulfill this performance criterion. It is, therefore, deduced that the inverse problem is complicated and requires further analysis and elaboration, which is outside the scope of the present work.

VIII. Conclusion

In this paper, the relation of the phase errors in the grating arms of the AWG, induced by the fabrication imperfections, and the sidelobe level of its transfer function, was investigated. A simple numerical tool was proposed in order to relate the fabrication tolerances of individual waveguides with the induced phase errors based on the effective index method. Waveguide coupling was studied, and it was deduced that it does not cause noticeable correlation between the phase errors in the different waveguides. The sidelobes of the AWG were shown to be exponential random variables both by using numerical simulations and theoretical considerations. The losses introduced on the peak of the transfer function due to the phase errors were also studied. Finally, the behavior of the maximum sidelobe level was analyzed using numerical simulations, and universal diagrams were presented that allow the estimation of its mean value, standard deviation, and CDF from the fabrication tolerances and the parameters of the AWG design.

Appendix
Calculation of the Standard Deviation of the Sidelobes

The variance of $T(v)$ is given by $\langle T^2(v)\rangle - \langle T(v)\rangle^2$. The expectation $\langle T(v)\rangle$ has been computed in (9). The computation of $\langle T^2(v)\rangle$ is much more involved since

$$\langle T^2(v)\rangle = \sum_{m,n,k,l} D_m D_n D_k^* D_l^* \langle \exp(j(\delta_m + \delta_n - \delta_k - \delta_l))\rangle \tag{A1}$$

where we have defined $D_m = C_m \exp(j2\pi m v)$, and the asterisk denotes complex conjugate. The expectation $\langle \exp(j(\delta_m + \delta_n - \delta_k - \delta_l))\rangle$ can be computed using the fact that

$$\langle \exp(j(\delta_m + \delta_n - \delta_k - \delta_l))\rangle = \langle \exp(j\delta_m)\rangle \langle \exp(j(\delta_n - \delta_k - \delta_l))\rangle \tag{A2a}$$

in the case where $m \neq n, m \neq k$, and $m \neq l$. If the aforementioned condition does not hold, the RHS of (A2a) will be different. If, for instance, $m = n, m \neq k$, and $m \neq l$

$$\langle \exp(j(\delta_m + \delta_n - \delta_k - \delta_l))\rangle = \langle \exp(2j\delta_m)\rangle \langle \exp(j(-\delta_l - \delta_k))\rangle. \tag{A2b}$$

Continuing this way, one can write $\langle T^2(v)\rangle$ as a sum of sums containing terms of the form $\langle \exp(j(q_m\delta_m + q_n\delta_n + q_k\delta_k))\rangle$, where the q_m, q_n, and q_k are integers.

The terms $\langle \exp(j(q_m\delta_m + q_n\delta_n + q_k\delta_k))\rangle$ can be expressed in a way similar to (A2) and so on, until $\langle T^2(v)\rangle$ is expressed as a sum of sums containing terms of the form $\langle \exp(jq_m\delta_m)\rangle = \exp(-q_m^2\sigma_m^2/2)$. The final expression is very complicated and will not be given here. It should be noted that, as shown theoretically in Section V, when v is in the vicinity of the sidelobes, the distribution of $T(v)$ is exponential with a high degree of accuracy. Therefore, for these values of v, the standard deviation of $T(v)$ is approximately equal to the square of its mean value, that is $\langle T^2(v)\rangle - \langle T(v)\rangle^2 \cong \langle T(v)\rangle^2$.

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, 1991.

[3] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 948–964, Aug. 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] T. Goh, S. Suzuki, and A. Sugita, "Estimation of waveguide phase error in silica-based waveguides," *J. Lightwave Technol.*, vol. 15, pp. 2107–2213, Nov. 1997.

[8] C. D. Lee, W. Chen, Q. Wang, Y. J. Chen, W. T. Beard, D. Stone, R. F. Smith, R. Mincher, and I. R. Stewart, "The role of photomask resolution on the perfomance of arrayed-waveguide grating devices," *J. Lightwave Technol.*, vol. 19, pp. 1726–1733, Nov. 2001.

[9] R. Mestric, M. Renaud, M. Bachmann, B. Martin, and F. Gaborit, "Design and fabrication of 1.31–1.55 μm phased array duplexer on InP," *IEEE J. Select. Topics Quantum Electron.*, vol. 2, pp. 251–256, June 1996.

[10] R. Scarmozino, A. Gopinath, R. Pregla, and S. Helfert, "Numerical techniques for modeling guided-wave photonic devices," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 150–161, Jan./Feb. 2000.

[11] H. A. Haus and W. P. Huang, "Mode coupling in tapered structures," *J. Lightwave Technol.*, vol. 7, pp. 729–730, Apr. 1989.

[12] S. Suzuki, M. Yanagisawa, Y. Hibino, and K. Oda, "High-density integrated planar lightwave circuits using SiO_2-GeO_2 waveguides with a high refractive index difference," *J. Lightwave Technology*, vol. 12, no. 5, pp. 790–796, May 1994.
[13] J. C. Chen and C. Dragone, "A study of fiber-to-fiber losses in waveguide grating routers," *J. Lightwave Technol.*, vol. 15, pp. 1895–1899, Oct. 1997.
[14] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1991.

Thomas Kamalakis was born in Athens in 1975. He received the degree in informatics and telecommunications and the M.Sc. degree in telecommunications from the University of Athens, Greece, in 1997 and 1999, respectively. Currently, he is working toward the Ph.D. degree on the design and modeling of arrayed-waveguide grating (AWG) devices with the same university.

Thomas Sphicopoulos (M'87) was born in Athens in 1953. He received the physics degree from Athens University in 1976, the D.E.A. and Ph.D. degrees in electronics from the University of Paris VI in 1977 and 1980, respectively, and the Doctorat Es Science from the Ecole Polytechnique Federale de Lausanne, Switzerland, in 1986.

From 1976 to 1977, he worked in Thomson CSF Central Research Laboratories on microwave oscillators. From 1977 to 1980, he was an Associate Researcher in Thomson's Aeronautics Infrastructure Division. In 1980, he joined the Electromagnetism Laboratory of the Ecole Polytechnique Federal de Lausanne, where he carried out research on applied electromagnetism. Since 1987, he has been with Athens University, where he has been engaged in research on broadband communications systems. Currently, he is a Professor, and since February 2000, head of the Department of Informatics and Telecommunications. His main areas of scientific interest are microwave and optical communication systems and networks. He has published more than 70 articles in scientific journals and conference proceedings.

Dimitris Syvridis, photograph and biography not available at the time of publication.