Some Statistical Remarks on the Derivation of BER in Amplified Optical Communication Systems

Felix Abramovich and Polina Bayvel

***Abstract*— We consider the signal detection problem in amplified optical transmission systems as a statistical hypothesis testing procedure, and we show that the detected signal has a well-known chi-squared distribution. In particular, this approach considerably simplifies the derivation of bit-error rate (BER). Finally, we discuss the accuracy of the Gaussian approximations to the exact distributions of the signal.**

I. Introduction

IN THIS LETTER we consider the analysis and derivation of the bit-error rate (BER) in an optical transmission system with all-optical amplifiers, such that the amplifiers add Gaussian amplified spontaneous emission noise, dominant over other possible noise sources. This is a well-known problem, studied by a number of authors (see, for example, Marcuse [4]). In this contribution, we consider this problem as a standard statistical hypothesis testing procedure that, in particular, automatically yields results obtained by Marcuse through rather complex calculus. The suggested approach provides valuable insight into the evaluation of performance in optical transmission systems. We also discuss the accuracy of Gaussian approximations for the exact signal distributions.

II. Optical Transmission System: Statistical Model

Consider a transmission system with optical amplifiers (Fig. 1) with an input noiseless signal $E(t)$ being either logical ZERO or logical ONE. The amplifiers generate spontaneous emission noise $e(t)$ added to $E(t)$, and the resulting *noisy* signal $z(t) = E(t) + e(t)$ is filtered by a bandpass optical filter and then incident on the detector. The detector produces an electrical current proportional to the absolute square value of the amplitude of the noisy signal, so the output signal from detector is $I(t) = K|z(t)|^2 = K|E(t) + e(t)|^2$. The output current $I(t)$ is averaged over one bit period T, and the observed datum y is

$$y = \frac{1}{T}\int_0^T I(t)\,dt = \frac{K}{T}\int_0^T |E(t) + e(t)|^2\,dt.$$

Various aspects of statistical analysis of detecting noisy signals are considered, for example, in Whalen [5]. From a statistical viewpoint, detecting a signal is essentially a hypothesis testing procedure. Given an observed noisy y, one tests whether $E(t)$ is a logical ZERO (null hypothesis H_0) or a logical ONE (alternative hypothesis H_1). We construct an appropriate test using the Fourier representation of signals:

$$E(t) = \sum_{\nu=0}^{\infty} C_\nu e^{2\pi\nu i t/T} \quad \text{and} \quad e(t) = \sum_{\nu=0}^{\infty} c_\nu e^{2\pi\nu i t/T} \tag{1}$$

where C_ν and c_ν are the Fourier coefficients of a noiseless signal $E(t)$ and noise $e(t)$, respectively. Since, prior to the square-law detector, the signal is passed through an optical bandpass filter (see Fig. 1) whose bandwidth is between $\nu = \nu_1$ and $\nu = \nu_2$, only these frequencies will remain in (1), so that

$$E(t) = \sum_{\nu=\nu_1}^{\nu_2} C_\nu e^{2\pi\nu i t/T} \quad \text{and} \quad e(t) = \sum_{\nu=\nu_1}^{\nu_2} c_\nu e^{2\pi\nu i t/T}.$$

The Fourier coefficients z_ν of the resulting noisy signal $z(t)$ are $z_\nu = C_\nu + c_\nu$, and the observed y is

$$y = K\sum_{\nu=\nu_1}^{\nu_2} |z_\nu^2| = K\sum_{\nu=\nu_1}^{\nu_2} (z_{r\nu}^2 + z_{i\nu}^2)$$

where $z_{r\nu}$ and $z_{i\nu}$ are real and imaginary parts of z_ν, respectively. Assuming that the noise $e(t)$ is Gaussian with a variance $\sigma^2, z_{r\nu} \sim \mathcal{N}(C_{r\nu}, \sigma^2), z_{i\nu} \sim \mathcal{N}(C_{i\nu}, \sigma^2)$, and all $z_{r\nu}$'s, $z_{i\nu}$'s are independent due to the orthogonality of the Fourier series.

Under the null hypothesis H_0 (logical ZERO), the average $\langle y \rangle = 0$, while under the alternative hypothesis H_1 (logical ONE), $\langle y \rangle = \Sigma_{\nu=\nu_1}^{\nu_2} (C_{r\nu}^2 + C_{i\nu}^2) > 0$. Performing any statistical test on noisy data, there are two types of possible errors: to reject erroneously H_0, known in statistics as a Type I error, and to accept erroneously H_0, a Type II error (e.g., [5]), corresponding to an incorrect detection of a logical ONE or a logical ZERO, respectively.

In the problem at hand, given y, the likelihood ratio test for testing H_0 versus H_1 is of the form: reject H_0 (and hence, accept H_1) if $y > I_d$, where I_d is a critical threshold value. The threshold current I_d is usually chosen to minimize the total probability of an error in detecting any signal, i.e., the bit-error rate $\mathrm{BER} = \frac{1}{2}[P_0(I_d) + P_1(I_d)]$, where $P_0(I_d)$ and $P_1(I_d)$ are probabilities of Type I and Type II errors.

III. Derivation of BER

To calculate $P_0(I_d)$ and $P_1(I_d)$ (and hence, BER), one needs to derive first the distributions of the test statistic y under both hypotheses.

We start with H_0, under which all $C_{r\nu} = C_{i\nu} = 0$, and hence, y is the sum of squares of $2M$ independent Gaussian

Paper approved by J. J. O'Reilly, the Editor for Optical Communications of the IEEE Communications Society. Manuscript received July 1, 1996; revised February 4, 1997.

F. Abramovich is with the Department of Statistics & Operations Research, The Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Israel.

P. Bayvel is with the Department of Electronic & Electrical Engineering, University College London, London WCIE 7JE, England.

Publisher Item Identifier S 0090-6778(97)06614-2.

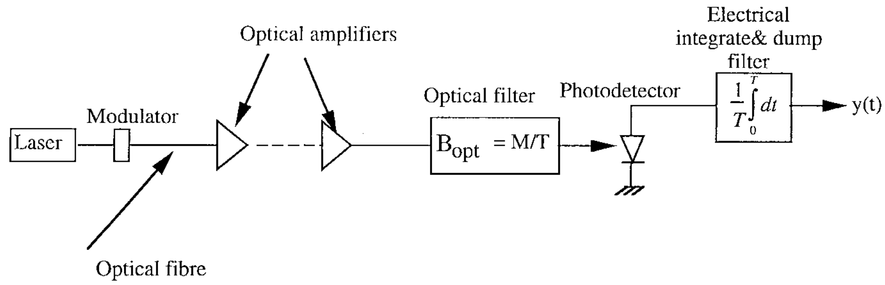

Fig. 1. Schematic diagram of an optical communication system.

random variables with zero means and variances $K\sigma^2$, where $M = \nu_2 - \nu_1 + 1$. Then, $y \sim K\sigma^2\chi^2_{2M}$, where χ^2_{2M} is a (central) chi-squared distribution with $2M$ degrees of freedom [5] whose probability density is

$$\chi^2_{2M}(x) = \frac{1}{2^M}\,\frac{x^{M-1}}{(M-1)!}\,e^{-x/2} \tag{2}$$

and thus, the probability density $W_0(y)$ of y under H_0 is

$$\begin{aligned} W_0(y) &= \frac{1}{K\sigma^2}\,\chi^2_{2M}(y/K\sigma^2) \\ &= \left(\frac{1}{2K\sigma^2}\right)^M \frac{y^{M-1}}{(M-1)!}\,e^{-y/2K\sigma^2}. \end{aligned}$$

Using the standard results for the chi-squared distribution, one immediately has

$$\begin{aligned} \overline{I}_0 &= \langle y\rangle = 2MK\sigma^2 \\ \sigma_0^2 &= \mathrm{var}(y) = 4MK^2\sigma^4 = \overline{I}_0^2/M. \end{aligned}$$

All of these directly obtained results are identical to [4, eqs. (13), (17), (18)], derived there after rather tedious calculus.

By repeated integration by parts, the probability of Type I error is

$$\begin{aligned} P_0(I_d) &= \int_{I_d}^{\infty} W_0(y)\,dy \\ &= \frac{1}{2K\sigma^2}\,e^{-(I_d/2K\sigma^2)} \sum_{j=0}^{M-1} \frac{(I_d/2K\sigma^2)^j}{j!}. \end{aligned} \tag{3}$$

The sum in (3) can be well approximated by the (last) largest two terms, leading to Marcuse's approximation (16) for $P_0(I_d)$:

$$\begin{aligned} P_0(I_d) \approx &\frac{1}{(M-1)!}\left(\frac{I_d}{2K\sigma^2}\right)^{M-1}\left(1 + \frac{2K\sigma^2(M-1)}{I_d}\right) \\ &\cdot e^{-(I_d/2K\sigma^2)}. \end{aligned}$$

Consider now the distribution of y in the case of a logical ONE signal. Under the alternative hypothesis H_1, y is the sum of squares of $2M$ independent Gaussian random variables with means $C_{r\nu}$'s and $C_{i\nu}$'s, and therefore, is distributed as $y \sim K\sigma^2\chi^2_{2M;|E_1|^2}$, where $\chi^2_{2M;|E_1|^2}$ is a *noncentral* chi-squared distribution with $2M$ degrees of freedom and the noncentrality parameter $|E_1|^2 = \Sigma_{\nu=\nu_1}^{\nu_2}(C^2_{r\nu} + C^2_{i\nu})$ [5]. The probability density of a noncentral chi-squared random variable is an infinite weighted sum of the usual (central) chi-squared densities, where the weights are successive Poisson probabilities (Johnson and Kotz, [3]:

$$\chi^2_{2M;|E_1|^2} = \sum_{j=0}^{\infty} \frac{e^{-|E_1|^2/2}(|E_1|^2/2)^j}{j!}\,\chi^2_{2M+2j}.$$

Thus, the exact probability density of y under H_1 is given by the infinite sum

$$W_1(y) = \frac{1}{K\sigma^2}\sum_{j=0}^{\infty}\frac{e^{-|E_1|^2/2}(|E_1|^2/2)^j}{j!}\,\chi^2_{2M+2j}(y/K\sigma^2)$$

where the (central) chi-squared distributions $\chi^2_{2M+2j}(\cdot)$ are given in (2). The average and the variance of y are

$$\begin{aligned} \langle y\rangle &= \sigma^2K(2M + |E_1|^2) = \overline{I}_0 + \overline{I}_1 \\ \sigma_1^2 &= 2\sigma^4K^2(2M + 2|E_1|^2) = \sigma_0^2 + \frac{2}{M}\,\overline{I}_0\overline{I}_1 \end{aligned} \tag{4}$$

where $\overline{I}_1 = \sigma^2K|E_1|^2$ [5]. (Note that σ^2 is erroneously missing in Marcuse's definition (23) of $\overline{I}_1$.) Again, (4) coincides with analogous results of Marcuse [4, eqs. (46), (47)].

The probability of Type II error $P_1(I_d) = \int_0^{I_d} W_1(y)\,dy$. However, to obtain a closed form for $P_1(I_d)$, one has to use various approximations for $W_1(y)$. Indeed, one of these was used by Marcuse [4, eq. (24)].

The critical threshold value I_d that minimizes BER $= \frac{1}{2}[P_0(I_d) + P_1(I_d)]$ is defined by solving the equation $W_0(I_d) = W_1(I_d)$. To derive an analytical approximation for $I_d, W_1(I_d)$ should be replaced by one of its approximations. The explicit expressions for I_d and for the resulting BER are given in Marcuse [4, Sect. III].

Summarizing, we can conclude that the statistical hypothesis testing approach introduced in this section allows a direct and straightforward derivation of BER in optical communication systems.

IV. Gaussian Approximation

In this section, we discuss the Gaussian approximations to $W_0(y)$ and $W_1(y)$. This problem was considered by Marcuse [4], who demonstrated that the Gaussian approximations are not appropriate for small M. We would like to make some statistical comments to clarify this issue.

As already mentioned, y is the sum of $2M$ independent random variables (each of them is a squared Gaussian), and therefore, according to the central limit theorem, for large M, the distribution of y is asymptotically Gaussian. Thus, *asymptotically* for large $M, W_0(y)$ is $\mathcal{N}(\overline{I}_0, \overline{I}_0^2/M)$ and $W_1(y)$ is $\mathcal{N}(\overline{I}_0+\overline{I}_1, \sigma_0^2+(2/M)\overline{I}_0\overline{I}_1)$ or, more precisely, the distributions of the standardized $y^* = (y/(K\sigma^2) - 2M)/\sqrt{4M}$ under H_0 and of $y^{**} = (y/(K\sigma^2) - 2M - |E_1|^2/\sqrt{4(M+|E_1|^2)}$ under H_1 converge to the standard normal distribution $\mathcal{N}(0,1)$ as M increases. However, substituting the corresponding cumulants of the chi-squared distribution [3] into the Edgeworth expansion of the exact cumulative distribution function $F_0(y^*)$ of y^* under H_0, we have

$$F_0(y^*) = \Phi(y^*) - \phi(y^*)\{H_2(y^*)/(3\sqrt{M}) + H_3(y^*)/(4M) + H_5(y^*)/(18M)\} + O(M^{-3/2}) \tag{5}$$

where $\Phi(\cdot), \phi(\cdot)$ are the cumulative distribution function and the density function of the standard normal distribution, respectively, and $H_j(\cdot)$ is the Hermite polynomial of degree $j, j = 2,3,5$ [2, ch. 4]. The expansion (5) indicates that the rate of convergence of $F_0(\cdot)$ to $\Phi(\cdot)$ is rather slow $[O(M^{-1/2})]$, and essentially explains why the Gaussian approximation to the chi-squared distribution performs satisfactorily when the number of degrees of freedom $(2M)$ is sufficiently large (in practice, usually is on the order of 30 or even more). Since, physically, M represents the optical filter bandwidth, practical low-noise optical receivers will restrict M to be less than 10. Hence, it is not surprising that the Gaussian approximations behave poorly, and the original chi-squared distributions should be used.

Alternatively, the accuracy of the Gaussian approximations can be improved even for small M by using the Wilson–Hilferty transformations of the original chi-squared variables: $y^*_{\rm WH} = ((y/(K\sigma^2)/2M))^{1/3}$ and $y^{**}_{\rm WH} = ((y/(K\sigma^2)/2M + |E_1|^2))^{1/3}$. Then, $y^*_{\rm WH}$ under H_0 is approximately Gaussian with mean $1 - 1/(9M)$ and variance $1/(9M)$ [6], while $y^{**}_{\rm WH}$ under H_1 is approximately Gaussian with mean $1 - (4(M+|E_1|^2)/9(2M+|E_1|^2))$ and variance $(4(M+|E_1|^2)/9(2M+|E_1|^2)^2$, respectively [1]. It is key to note that the rate of convergence of the standardized $y^*_{\rm WH}$ and $y^{**}_{\rm WH}$ to the standard normal distribution is now $O(M^{-3/2})$, which is faster than that in (5), and the corresponding Gaussian approximations work quite satisfactorily, even for small numbers of degrees of freedom [6], [1]. 1931.

V. Conclusions

It was shown that the statistical hypothesis testing approach allows a direct and straightforward derivation of the detected signal distribution, and hence, error probabilities in detecting a signal in an amplified optical transmission system. Specifically, it was shown that this distribution is a well-known chi-squared distribution. The statistical analysis, introduced here, gives a useful insight into the evaluation of performance in optical communication systems. Moreover, this approach can also be extended to analyze systems with more general noise sources, including noise generated by nonlinear transmission effects and in the electrical detector circuit. The accuracy of the Gaussian approximations to the original chi-squared distributions of the signal was discussed, and it was shown that it can be improved by using the Wilson–Hilferty transformation.

Acknowledgment

The authors thank P. Lane and L. Moura (UCL) for useful clarifications and comments.

References

[1] S. H. Abdel-Aty, "Approximate formulae for the percentage points and the probability integral of the noncentral χ^2-distribution," *Biometrika*, vol. 41, pp. 538–540, 1954.

[2] O. E. Barndorff-Nielsen and D. R. Cox, *Asymptotic Techniques for Use in Statistics*. London, U.K.: Chapman and Hall, 1989.

[3] N.I. Johnson and S. Kotz, *Continuous Univariate Distributions*. Boston, mA: Wiley, 1970.

[4] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightwave. Technol.*, vol. 8, pp. 1816–1823, 1990.

[5] A. D. Whalen, *Detection of Signals in Noise*. Orlando, FL: Academic, 1971.

[6] E. B. Wilson and M. M. Hilferty, "The distribution of chi-square," in *Proc. Nat. Acad. Sci.*, vol. 17, 1931, pp. 684–688.