THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 55, No. 6, July–August 1976
Printed in U.S.A.

Statistical Evaluation of the Error Rate of the Fiberguide Repeater Using Importance Sampling

By P. BALABAN

(Manuscript received December 12, 1975)

Simulation of repeaters for an optical-fiberguide digital-communication system requires the calculation of the statistical error rate of the signal. The calculations of these error rates are difficult because of the non-gaussian nature of the noise in the optical detector. In this paper, statistical techniques are described which are useful for simulating tails of distributions. In particular, the importance-sampling procedure is used to modify the probability densities of the input values in a way that makes simulation possible. Application of this procedure gives more accurate results in reasonable computer times. The method is applied to the calculation of the error rate of a fiberguide repeater. Realistic examples are simulated. Results compare favorably with experimental measurements. The number of samples needed for simulation was reduced by five to six orders of magnitude. An alternative numerical solution to the problem is also developed.

I. INTRODUCTION

Traditional computation of error rates of digital repeaters is often done by assuming that the noise is gaussian in nature. For a linear equalizer, the error rate at a given sampling time can then be easily computed if the signal and standard deviation of the noise are given.[1]

The fiberguide repeater has essentially three kinds of noise sources. Two noise sources are located in the detector, while the third is a thermal gaussian noise generated by the input stage of the preamplifier (see Fig. 1).

The detector, which is an avalanche photodiode is governed by two intrinsic noise processes: the primary photo current, which is a Poisson process, and the avalanche multiplication process which produces the gain in the diode. Neither of these processes is gaussian.

The probability density functions (PDF) for the output current of the diode were derived by Personick[2,3] and McIntyre.[4] The expressions

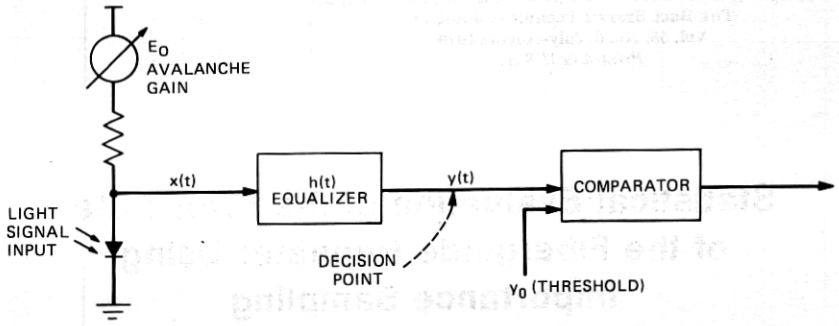


Fig. 1—Optical repeater.

describing these distributions are quite complex, and excessively time consuming for computer simulation of the repeater. An approximation to these probability densities was obtained by Webb, McIntyre, and Conradi (WMC).[5] The approximate PDF (referred to below as the WMC PDF) is, according to the authors, fairly accurate in the practical operating range, and is easily implemented on the computer.

The WMC PDF for the current at the output of the diode is:

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma\left(1 + \dfrac{x - M}{\sigma\lambda}\right)^{\frac{1}{2}}} \exp\left[\frac{-(x - M)^2}{2\sigma^2\left(1 + \dfrac{x - M}{\sigma\lambda}\right)}\right], \quad (1)$$

where

$M = (n_e q/T)\mathbf{G}$ = mean output current
q = charge of an electron
T = time interval
$\mathbf{G}$ = average avalanche gain
$n_e = (\eta T/\hbar\Omega)\mathbf{P}_{opt}$ = number of primary electrons
η = optical conversion efficiency
$\mathbf{P}_{opt}$ = average optical power in time T
$\hbar\Omega$ = energy of a photon
$\sigma^2 = n_e\mathbf{G}^2F_e(q/T)^2$ = variance of the diode current
$\lambda = \sqrt{n_eF_e}/(F_e - 1)$
$F_e = k_{eff}\mathbf{G} + [2 - (1/\mathbf{G})](1 - k_{eff})$ = excess noise factor
k_{eff} = effective ionization ratio (assumed constant over range of $\mathbf{G}$ considered)

k_{eff} and η are photodiode parameters.

Equation (1) is, more precisely, the PDF of the mean diode current over the time interval T for a given average optical input power,

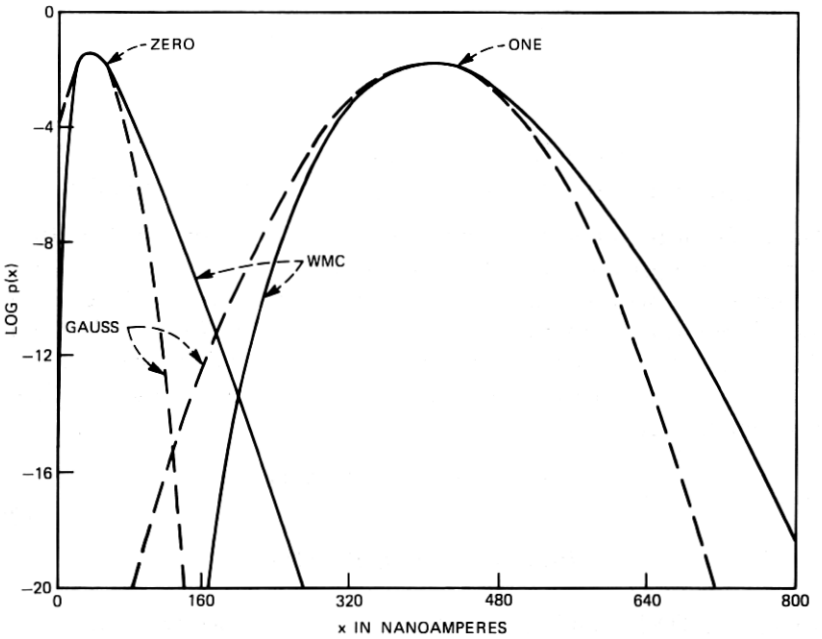


Fig. 2—Probability density function for ONE and ZERO signals using WMC and gaussian approximation.

$\mathbf{P}_{opt}$, and a given average avalanche gain **G**. It corresponds to the sum over all primary electrons of the product of the primary electron PDF with the avalanche gain PDF. This equation, therefore, combines the effects of both detector noise sources.

In Fig. 2 the solid lines represent current PDFs for two signals: a ONE where the optical power is $\mathbf{P}_{opt} = -51$ dBm and a ZERO where $\mathbf{P}_{opt} = -61$ dBm. The other parameters are $\mathbf{G} = 80$, $\eta = 0.667$, $k_{eff} = 0.04$, $T = 22$ ns. The broken lines represent error PDFs of the same signals and same diode parameters where a normal approximation to the distribution is assumed. The asymmetry of the WMC distributions is clearly seen.

The noise generated by the first stage of the amplifier has to be added to this detector-generated noise.

II. POSSIBLE SOLUTION TO THE PROBLEM

To compute the error rate of the repeater, one has to know the signal probability distribution at the decision-making point, i.e., at the output of the equalizer in Fig. 1.

The output of the equalizer with impulse response $h(t)$ is

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau, \tag{2}$$

where

$x(t) = x_s(t) + x_{TH}(t)$
$x_s(t)$ = signal output of the avalanche diode
$x_{TH}(t)$ = gaussian thermal noise generated by the amplifier.

The problem is to find the distribution of $y(t)$ for given distributions of $x_s(t)$ and $x_{TH}(t)$ and impulse response $h(t)$.

Figure 3 shows the shapes of two distributions: $P_1(y)$ is the cumulative distribution of y for a ONE signal transmitted, $P_0(y)$ is the probability that the output current exceeds y for a ZERO signal transmitted. y_0 is the decision threshold of the comparator (Fig. 1). These distributions are

$$P_1(y) = \int_{-\infty}^{y} p(y/\text{ONE})dy$$

$$P_0(y) = \int_{y}^{\infty} p(y/\text{ZERO})dy, \tag{3}$$

where $p(y/\text{ONE})$ and $p(y/\text{ZERO})$ are the respective conditional PDFs. It should be noted that, for practical laser sources, the ZERO input signal could be on the order of 10 percent of the ONE signal,

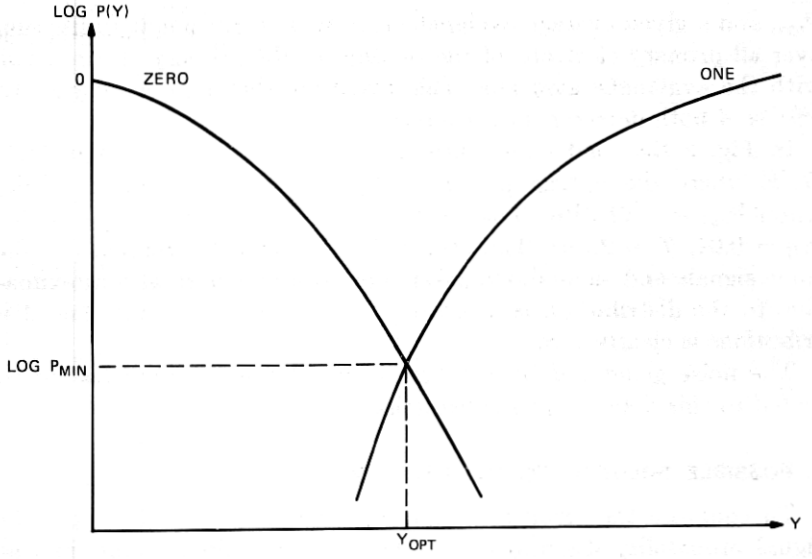


Fig. 3—Distributions for ZERO and ONE signals.

since the laser is not turned off completely. In addition, there is the dark current of the detector diode present. The error rate of a digital repeater is defined as

$$P_e = aP_1(y_0) + bP_0(y_0). \tag{4}$$

(The factors a and b are determined by the *a priori* distribution of the data. For unbiased data, $a = b = 0.5$.)

One of the features that has to be determined is the optimum placement of the decision threshold y_{opt}, that is, where P_e is minimum.

A general analytical solution to this problem is not possible; therefore, the problem has to be approximated numerically or simulated with the aid of a computer. Analysis of the system using a gaussian approximation to the detector noise sources was considered by S. D. Personick.[6] This analysis, however, predicts optimum decision thresholds considerably different from actual measured values.[7] Another solution to this problem is to replace the exact solution by a simpler solvable formulation that gives an upper-bound solution. This was done by Personick[3,8] using the Chernoff bounds. Finding upper bounds is a very effective technique since it gives a degree of confidence. Nevertheless, a method that gives more accurate results is needed to estimate how close the bounds are to the real solutions for this class of problems.

The problem can be solved either by the numerical solution of eq. (3) or by statistical Monte Carlo simulation using importance sampling. Both solutions were attempted and are described below. Particular emphasis was placed on the statistical simulation, since numerical solutions require excessive amounts of computer time.

III. DISCRETE STATISTICAL ANALYSIS (DIRECT COMPUTATION OF OUTPUT PDF)

The output of equalizer $y(t)$ is given in eq. (2). Since we are interested in the output at the sampling time t_0, it is not necessary to compute the convolution for every t. The value of t_0 is selected to maximize $y(t_0)$:

$$y(t_0) = \int_{-\infty}^{\infty} x(\tau)h(t_0 - \tau)d\tau \tag{5}$$

since

$$\begin{aligned} x(t) &= x_s(t) + x_{TH}(t) \\ y(t_0) &= \int_{-\infty}^{\infty} x_s(\tau)h(t_0 - \tau)d\tau + \int_{-\infty}^{\infty} x_{TH}(\tau)h(t_0 - \tau)d\tau \\ &= y_s(t_0) + y_{TH}(t_0). \end{aligned} \tag{6}$$

The problem is to find the probability distribution of $y(t_0)$.

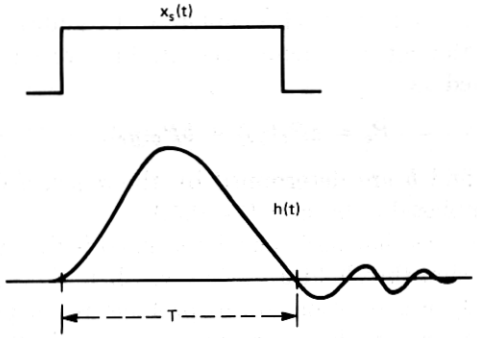


Fig. 4—Input signal and impulse response of the equalizer.

The probability distribution of $y_{TH}(t_0)$ is gaussian since x_{TH} is a thermally generated noise. The standard deviation of y_{TH} is

$$\sigma_{TH} = \sqrt{\frac{4kT\,B_w}{R_{TH}}}, \tag{7}$$

where B_w is the effective noise bandwidth of $h(t)$ and R_{TH} is the thermal resistance.

The effective noise bandwidth is defined as

$$B_w = \frac{1}{|H_0|^2}\int_0^\infty |H(f)|^2 df, \tag{8}$$

where $H(f)$ is the one-sided Fourier transform of $h(t)$ and $|H_0|$ is the maximum absolute value of $H(f)$. If the impulse response is approximated by a raised cosine of duration τ, $B_w = 3/4\tau$.

Typical waveforms for $x_s(t)$ and $h(t)$ are shown in Fig. 4. The discrete approximation of the signal $y_s(t_0)$ is:

$$y_s = \frac{T}{N}\sum_1^N x_n h_{N-n}, \tag{9}$$

where T is the duration of the signal and N is the number of samples used in the approximation. The samples x_n are independent and have a distribution of the form of eq. (1) for a time interval $\Delta T = T/N$.

3.1 Numerical computation of p(y)

If the random variables x_n are independent, then the density of their sum equals the convolution of their respective densities

$$p(y_s) = \frac{1}{\Delta T^N \cdot \prod_n |h_n|}\left[p\left(\frac{x_1}{\Delta T \cdot h_{N-1}}\right) * p\left(\frac{x_2}{\Delta T \cdot h_{N-2}}\right) * \cdots * p\left(\frac{x_N}{\Delta T \cdot h_0}\right)\right]. \tag{10}$$

Since the thermal noise is also additive and independent, the total density of the signal is

$$p(y) = p(y_s) * p(y_{TH}). \tag{11}$$

3.2 Fourier transform approach

The simplest and fastest way to compute $p(y)$ is to transform all partial densities using the FFT, multiply them all, and take an inverse FFT:

$$p(y) = \frac{1}{|(\Delta T)^N \prod_n (h_n)|} F^{-1}\left(\prod_n \left\{ F\left[p\left(\frac{x_n}{\Delta T \cdot h_{N-n+1}} \right) \right] \right\} \cdot F[p(y_{TH})] \right). \tag{12}$$

Unfortunately, the Fourier transform is computed as a sum of positive and negative numbers and, therefore, suffers from large numerical errors. Since we are interested in the tails of the distribution (error rate of 10^{-9}), we conclude that the numerical errors will be excessive.

3.3 Direct convolution

Direct convolution produces less truncation error, because densities are positive functions. But direct convolution requires an excessive amount of arithmetic operations, e.g., if we have n distributions each having M points, the number of operations (multiplication and addition) will be

$$N = \frac{n!M^2}{2^n}. \tag{13}$$

For $M = 100$ and $n = 2$, $N = 10^4$ which is a manageable number; but for $M = 100$ and $n = 10$, $N = 3.5 \times 10^7$ and the number of operations becomes excessive. A program using direct convolution was written for $n = 2$. A density for the average power of the input pulses was convolved with the density of the thermal noise. This can be regarded as a response of an integrate-and-dump equalizer (capacitive input).

The results of this convolution are shown in Fig. 5. The solid lines are the distributions for a ONE and ZERO being transmitted. The average power of the ONE in this case is -51 dBm and of the ZERO -61 dBm, with $T = 22$ ns, $R_{TH} = 2.3$ kilohms, and $B_w = 31$ MHz, which corresponds to a $\sigma_{TH} = 14.68$ nA.

The decision threshold θ is usually expressed in percent of the mean range of the output current as:

$$\theta = \frac{Y_0 - \bar{Y}_0}{\bar{Y}_1 - \bar{Y}_0} 100 \text{ [percent]}, \tag{14}$$

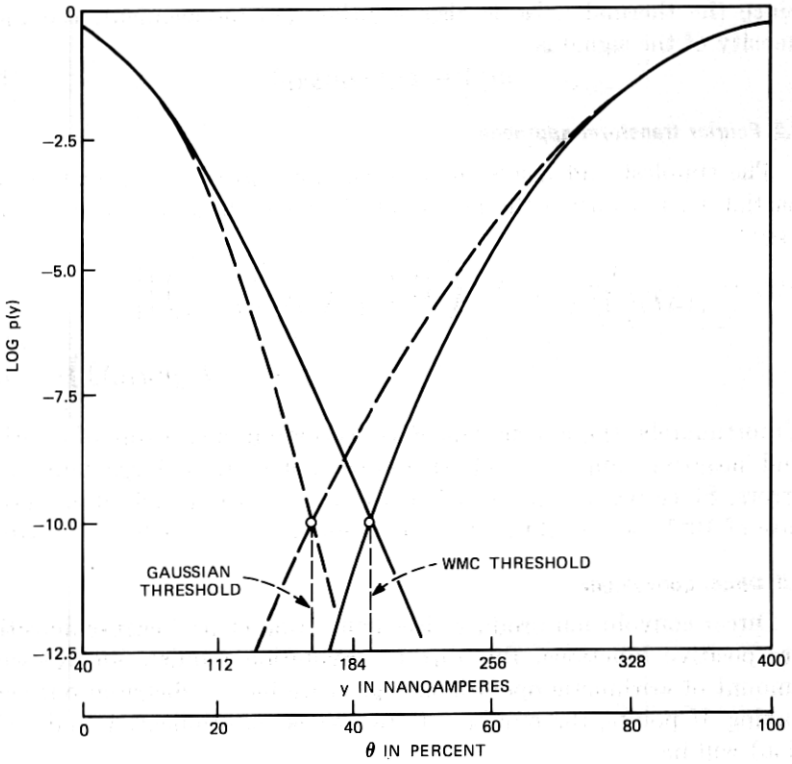


Fig. 5—Distributions for ONE and ZERO signals with thermal noise; WMC and gaussian approximations.

where

$\bar{Y}_1$ = mean output current for ONE
$\bar{Y}_0$ = mean output current for ZERO
Y_0 = threshold current.

As seen from the graph, the optimum decision threshold θ for this case is 42.3 percent and the minimum error rate is 9.86×10^{-11}. These results were also observed in an experimental setup.[7] The broken lines in Fig. 5 represent the same input but with the assumption that the probability density of the detector current is normal. As seen from the graph, the resultant threshold is considerably different from the one obtained with the WMC distributions.

IV. STATISTICAL (MONTE CARLO) ANALYSIS

Monte Carlo (MC) analysis has proven over the years to be a very effective method for analysis of systems whose parameters or inputs have known statistical distributions.[9]

However, straightforward application of MC analysis for calculation of error rates is impractical. The reason is that the error rates of interest are in the order of 10^{-6} to 10^{-9}. To have some confidence in the results, 10 to 100 samples have to be in that region of low probability. This will require on the average 10^{11} samples, which is impractical.

To overcome this shortcoming, we introduced a modified sampling method called importance sampling.[10] Conceptually, the idea is simple. If the regions which contribute to the result are known, modify the distributions in such a way that more samples are taken from the important regions.

Let us consider the situation where the problem is to estimate the value of

$$\xi = \int_{-\infty}^{\infty} g(x)p(x)dx, \tag{15}$$

where $p(x)$ is a PDF. The importance-sampling procedure consists in introducing another PDF $p^*(x)$, which is preferable for sampling purposes. Then

$$\begin{aligned}\xi &= \int_{-\infty}^{\infty} g(x)\,\frac{p(x)}{p^*(x)}\,p^*(x)dx \\ &= \int_{-\infty}^{\infty} g^*(x)p^*(x)dx.\end{aligned}$$

ξ is then estimated from

$$\xi = \frac{1}{N}\sum_{1}^{N} g^*(x_i),$$

where x_i is picked from the PDF $p^*(x_i)$.

4.1 Defining the important regions for error-rate computation

The problem is to define the regions where the errors come from. In Fig. 6a is shown the input waveform for a ONE, while Fig. 6b depicts the impulse response of the equalizer and 6c is the PDF of the diode current for one time interval ΔT.

Since the input signal is always positive, it is obvious that for positive h_n most of the errors, that is, outputs that are below the threshold θ, will come from region A in the PDF of Fig. 6c. Conversely, if h_n is negative, most of the errors will come from region C. Exactly the reverse is true for transmitted ZERO signals, since the signal is considered erroneous if it exceeds the threshold θ.

A very similar argument can be constructed for thermal noise, where the "important" regions are the positive tail for a ZERO transmitted and the negative tail for a ONE transmitted.

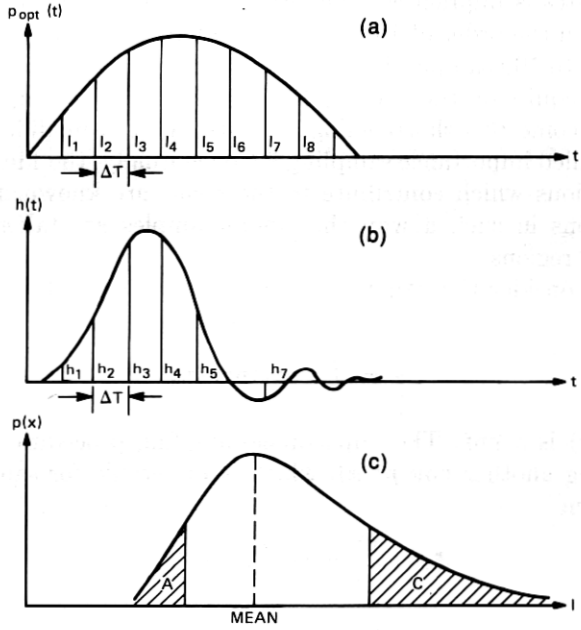


Fig. 6—(a) Optical input signal. (b) Impulse response of the equalizer. (c) PDF of average detector current for ΔT.

4.2 Biasing of the PDF

To implement the importance-sampling technique, the distributions have to be modified so that the important regions are emphasized. The PDFs in the random-number generator that we use are stored in tabular form, as shown in Fig. 7a. The region in question can be biased by a constant multiplier as in Fig. 7b, where $p^* = pB$ for $a \leqq x \leqq b$ and where the bias B is constant. In Fig. 7c, the PDF is biased such that it is quasi-uniform in the region prescribed. The bias B_i is constant in intervals between two adjacent points in the table. For each interval (x_{i-1}, x_i), this bias is

$$B_i = \frac{\alpha p_{\max}}{p_i}, \tag{16}$$

where p_i is the value of the PDF at x_i and α is a coefficient specified by the user. The biasing and normalization is done automatically by the program. The output of the random-number generator consists of two values: a random number RV and its associated bias BV. BV is the bias of the interval that RV was picked from.

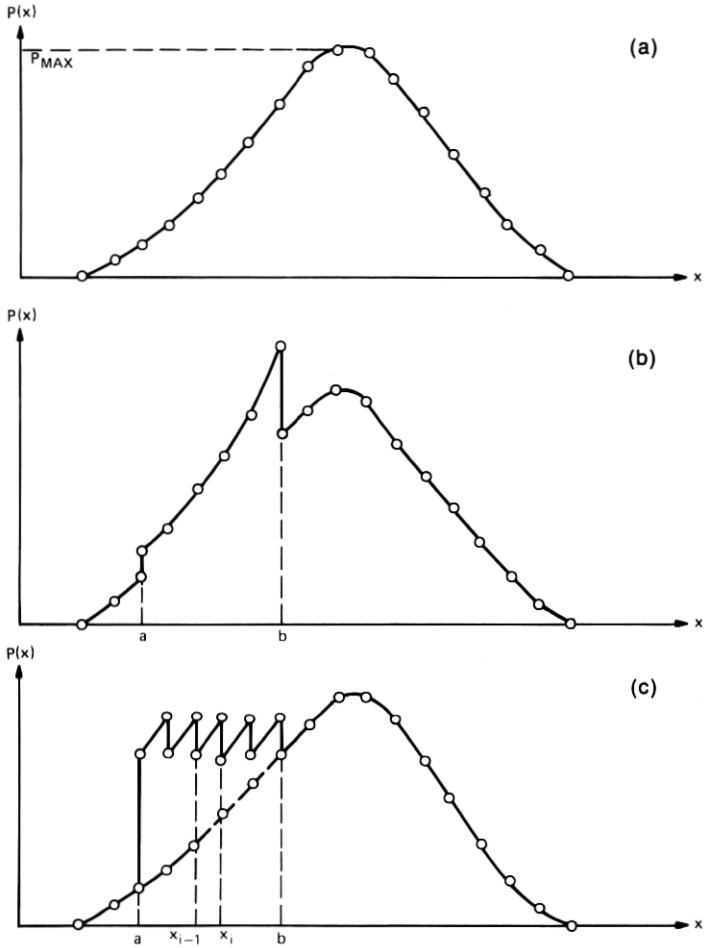


Fig. 7—Probability density functions. (a) Unbiased. (b) Constant bias. (c) Bias for quasi-uniform PDF.

4.3 Computing the histogram

4.3.1 The usual histogram

The standard output of a Monte Carlo analysis is the histogram. Usually the histogram is constructed by subdividing the output variable into M bins, Fig. 8.

The statistical analysis program counts the number of outputs which fall into each bin. For a large number of samples ($N \rightarrow \infty$), the histogram approximates the PDF. The probability p_i that the output y will

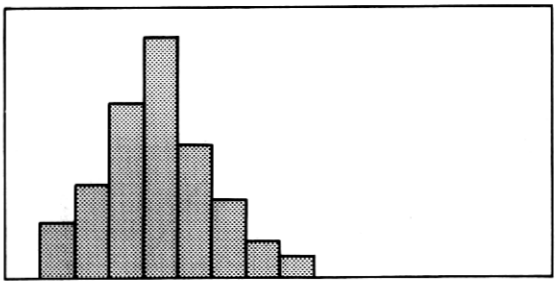


Fig. 8—Histogram.

be in the interval $y_{i-1} \leqq y \leqq y_i$ is

$$p_i = \lim_{N \to \infty} \frac{n_i}{N}, \tag{17}$$

where

n_i = number of samples in bin i
N = total number of samples.

The accuracy of estimation of p_i can be established by[11]

$$\epsilon = 2\frac{\sigma_i}{p} = 2\sqrt{\frac{(1-p)}{Np}}, \tag{18}$$

where $\pm\epsilon$ is the 95.4-percent confidence limit of p_i, p is the desired probability, and σ_i is the standard deviation of p_i.

4.3.2 Histogram for importance sampling

Let us assume we have M independent random variables $x_1, \cdots, x_m, \cdots, x_M$,

$$y = f(x_1, \cdots, x_m, \cdots, x_M),$$

$p^*(x_m)$ is the biased PDF of the random variable x_m. How do we compute the unbiased histogram for y.

If a sample x_{mi} is selected from $p^*(x_m)$, its probability to be selected is increased by the bias B_{mi}. Therefore, the weight of the sample x_{mi} is

$$w_{mi} = \frac{1}{B_{mi}}.$$

Since y_i is determined by M independent variables x_m the probability of y_i, to be selected from these regions, is increased by the product of all the biases of x_m. Therefore, the weight of y_i is $w_i = 1/(\prod_m B_{mi})$. The average weight of all the n_j samples that fell into bin j is given by

$$w_{avj} = \frac{1}{n_j}\sum_{i=1}^{n_j} w_{ij} = \frac{1}{n_j}\sum_{i=1}^{n_j}\left(\frac{1}{\prod_m B_{mi}}\right)_j. \tag{19}$$

The probability that the unbiased output will be $y_{j-1} \leqq y \leqq y_j$ is

$$p_j = w_{\mathrm{av}j} \frac{n_j}{N} = w_{\mathrm{av}j} p_j^*. \tag{20}$$

Since this method increases the probability p in eq. (18), the error ϵ is reduced considerably.

$$\epsilon = 2\sqrt{\frac{(1 - p^*)}{Np^*}} = 2\sqrt{\frac{w_{\mathrm{av}} - p}{Np}}. \tag{21}$$

For error-rate computation $p \ll 1$,

$$\epsilon \approx \frac{2}{\sqrt{Np^*}} = 2\sqrt{\frac{w_{\mathrm{av}}}{Np}}, \tag{22}$$

and since p_j^* is increased in the interesting areas by $1/w_{\mathrm{av}j}$, the confidence limit is decreased on the average by

$$\gamma = \frac{\epsilon_{\mathrm{IS}}}{\epsilon_{\mathrm{C}}} = \sqrt{w_{\mathrm{av}j}}, \tag{23}$$

where

ϵ_{IS} = Monte Carlo analysis with importance sampling
ϵ_{C} = conventional Monte Carlo analysis.

4.4 Example

A simple example to illustrate the principle of importance sampling is the problem of calculating the probability of obtaining a total of three when one tosses two dice. Each die has six faces labeled from one to six and each face has the probability $p = 1/6$ of being on top.

The problem, of course, can be solved analytically. Any particular combination of the dice has the probability of $1/6 \cdot 1/6 = 1/36$ of occurring. Since there are two combinations which make three (one-two and two-one), the probability of getting a three in a random toss of the dice is $2/36 = 1/18$.

If the probabilities of a one and a two are biased so that they occur twice as often as usual, that is, $p = 1/3$ rather than $1/6$, then the probability of getting a three will be four times as great as in the unbiased case. That is $p = 2/9$ instead of $1/18$. The weight for this example is

$$w_3 = 1/(B_1 B_3) = 1/4$$

and the probability of obtaining a 3 is being approximated by

$$\hat{p} = \frac{1}{4} \frac{n_3}{N}.$$

The error in eq. (18) is reduced by slightly more than a factor of two, since for

$$p = 1/18, \quad \epsilon = \frac{8.24}{\sqrt{N}}$$

and for

$$p = 2/9, \quad \epsilon = \frac{3.74}{\sqrt{N}}.$$

V. ERROR-RATE ANALYSIS

The importance-sampling method described above was applied to the error-rate analysis of the fiberguide repeater with an avalanche diode detector.

To demonstrate the analysis procedure, we will analyze one specific case. Let us assume that the input to the photodiode is as shown in Fig. 9. The ONE is a square pulse of −51 dBm power of 22 ns duration. ZERO is 10 percent of the ONE power or −61 dBm. The avalanche diode data are

$$\begin{aligned} &\text{average avalanche gain} = 80.0 \\ &\text{ionization ratio} \qquad k = 0.04 \\ &\text{conversion efficiency} \quad \eta = 0.667. \end{aligned}$$

The equalizer impulse response $h(t)$ is as in Fig. 10 and the thermal input resistance is $R_{TH} = 2.3$ kilohms.

The pulses (ONE and ZERO) are subdivided into nine intervals [eq. (9)], since this resolution is considered adequate for the impulse response in Fig. 10. Nine distributions* of the form in eq. (1) are computed for the ONE and ZERO pulses.

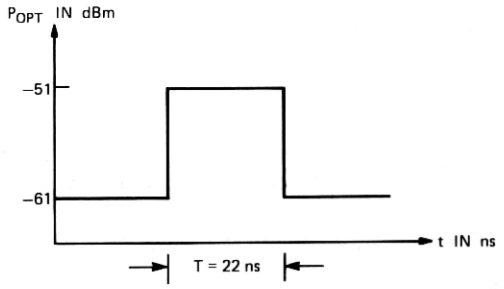


Fig. 9—Optical input.

* In this case the distributions are identical since the pulse is rectangular.

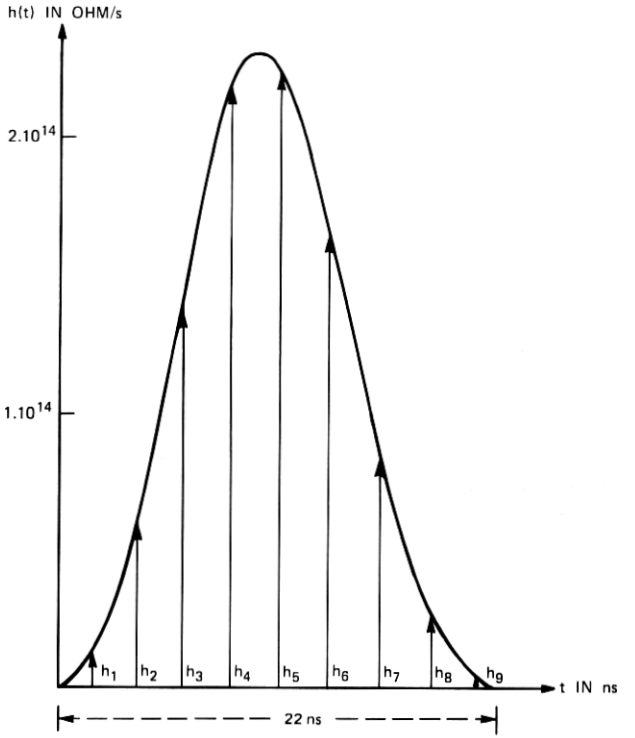


Fig. 10—Impulse response.

The mean for ONE is	$M_1 = 400.3$ nA
The standard deviation for ONE is	$\sigma_1 = 102.66$ nA
The mean for ZERO is	$M_0 = 40.03$ nA
The standard deviation for ZERO is	$\sigma_0 = 32.46$ nA

The effective noise bandwidth [eq. (8)] for the thermal noise is $B_w = 31$ MHz. The standard deviation is computed from eq. (7):

$$\sigma_{TH} = 14.68 \text{ nA}.$$

From the above considerations, the output for N intervals is

$$y = \frac{T}{N} \sum_{n=1}^{N} x_n h_{N-n} + \frac{T}{N} \sum_{n=1}^{N} h_n x_{TH}, \tag{24}$$

where the x_n and x_{TH} are samples from the respective PDFs and h's are discrete values obtained from $h(t)$. The impulse response is normal-

ized to produce unity gain

$$y^* = \frac{T}{N} \sum_{n=1}^{N} h_n \cdot y$$

$$h_n^* = \frac{h_{N-n}}{\sum_{n=1}^{N} h_n} \tag{25}$$

$$y^* = \sum_{n=1}^{N} h_n^* x_n + x_{\text{TH}}.$$

In addition, the above case has been analyzed for $N = 1$. This corresponds to the integrate-and-dump equalizer which was analyzed numerically in Section 3.3.

5.1 Computation of distributions and biasing

The calculations of distributions for the diode output are straightforward using the WMC distributions of eq. (1). The biasing is a more complex procedure and sometimes requires a few trials. The idea is to get a reasonable amount of samples in the "important" region of the output. The PDF for ONE was computed for $x_n = 0$ to 800 nA, and was biased according to eq. (16),

$$B_i = \alpha \frac{p_{\max}}{p_i} \text{ for 0 to 400 nA with } \alpha = 1.$$

The PDF for ZERO was computed from 0 to 400 nA and was biased from 40 nA to 400 nA with an $\alpha = 0.1$. The gaussian distribution for the thermal noise was tabulated for $\pm 7\sigma$. The biased regions were -7σ to 0 for the ONE signal and 0 to 7σ for the ZERO signal with an $\alpha = 0.5$. The PDFs were stored in tabular form for 100 points. Since the ZERO signal has a PDF which requires more detail for small currents (Fig. 2), the independent variable was spaced logarithmically. All others were linearly spaced.

The distributions were computed by a discrete form of eq. (3):

$$\text{Cum1}(n) = \sum_{i=1}^{n} \text{Hist1}(i)$$

$$\text{Cum0}(n) = \sum_{i=n}^{M} \text{Hist0}(i), \tag{26}$$

where Hist1(i) and Hist0(i) are histograms of the outputs.

To calculate the minimum error rate of eq. (4), Cum1(n) and Cum0(n) were smoothed and approximated by asymptotic functions. The minimum of the sum of these functions is the minimum error

rate, and the current where this minimum occurs is the optimum threshold.

5.2 Smoothing of the cumulative distributions

It was shown by S. B. Weinstein[12] that for the class of "exponential-type" distributions the error-probability distribution has an asymptote in the tails of the form

$$P_e(x) \sim \exp\left[-\left(\frac{x}{\sigma}\right)^{\nu}\right]. \tag{27}$$

For $\nu = 2$ the distribution is gaussian.

To obtain better accuracy, this asymptotic function was expanded to

$$P_e(x) = \exp\{-\exp[a(\log x)^2 + b \log x + c]\}$$

or

$$\log(-\log P_e) = a(\log x)^2 + b \log x + c; \tag{28}$$

i.e., the log log of P_e is a second-order polynomial of $\log x$. This equation reduces to Weinstein's form, eq. (27), for $a = 0$.

The second-order polynomials were then least square fit to both Cum1 and Cum0.

5.3 Minimum of error and optimum threshold

The probability of error, eq. (4), is defined as

$$\begin{aligned} P_e(I) &= \tfrac{1}{2}[P_{e_1}(I) + P_{e_0}(I)] \\ &= \tfrac{1}{2}(\exp\{-\exp[a_1(\log I)^2 + b_1(\log I) + c_1]\} \\ &\qquad + \exp\{-\exp[a_0(\log I)^2 + b_0(\log I) + c_0]\}), \end{aligned} \tag{29}$$

where I is the decision threshold current.

To find the minimum of $P_e(I)$, we compute I_{opt} from

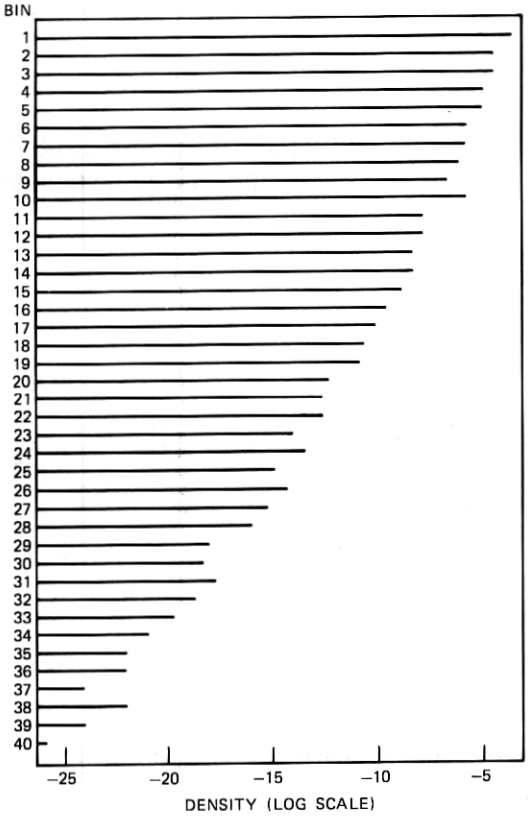
$$\left.\frac{dP_e(I)}{dI}\right|_{I=I_{opt}} = 0. \tag{30}$$

$$P_e(I)_{min} = P_e(I_{opt}).$$

Equation (30) is solved using the Newton-Raphson method. The threshold θ is usually expressed in percent as in eq. (14).

VI. RESULTS

Figure 11 shows a histogram for the ZERO signal for 10,000 samples. The abscissa is scaled logarithmically. In the tables on the right side are indicated the ordinate values for the bins, the number of samples in each bin, and the distribution for the bin. Below the histogram are the



I(nA)	NO. OF SAMPLES	CUM. 0 DISTRIB.
5.000E+01	85	2.232E−03
6.000E+01	111	6.067E−04
7.000E+01	165	4.128E−04
8.000E+01	168	1.667E−04
9.000E+01	216	9.279E−05
1.000E+02	277	2.061E−05
1.100E+02	293	1.573E−05
1.200E+02	331	9.328E−05
1.300E+02	339	6.318E−06
1.400E+02	383	5.809E−05
1.500E+02	427	7.079E−08
1.600E+02	390	4.162E−08
1.700E+02	456	1.349E−08
1.800E+02	454	7.077E−09
1.900E+02	458	1.984E−09
2.000E+02	455	3.315E−10
2.100E+02	440	1.114E−10
2.200E+02	433	2.208E−11
2.300E+02	444	5.407E−12
2.400E+02	426	3.579E−13
2.500E+02	375	1.461E−13
2.600E+02	380	6.224E−14
2.700E+02	319	1.003E−14
2.800E+02	307	8.280E−15
2.900E+02	252	9.154E−16
3.000E+02	208	7.867E−16
3.100E+02	178	9.662E−17
3.200E+02	194	1.057E−17
3.300E+02	156	1.557E−19
3.400E+02	142	1.030E−19
3.500E+02	113	8.583E−20
3.600E+02	98	7.897E−21
3.700E+02	76	4.744E−22
3.800E+02	49	2.739E−23
3.900E+02	54	2.839E−24
4.000E+02	42	1.640E−24
4.100E+02	29	8.339E−25
4.200E+02	18	8.271E−25
4.300E+02	20	5.535E−27
4.400E+02	13	2.754E−29

COEF. OF LEAST SQ. FIT $A_0 = 0.9954$ $A_1 = -1.2324$ $A_2 = 0.53019$
NUMBER OF SAMPLES = 10,000

Fig. 11—Histogram for ZERO signal.

coefficients of the fitted polynomial $R(I) = a_0 + a_1 \log I + a_2(\log I)^2$, where $R(I) = \log_{10}[-\log_{10} P_{eo}(I)]$.

Figure 12 shows the smoothed distributions for the above example for one section ($N = 1$) and 1000 samples. The stars represent numerical solutions for the same input parameters, as described in Section III. As seen from the data, if biasing is chosen rightly, a comparatively small sample size is adequate to simulate the extreme tails of the PDF.

Figure 13 shows the smoothed distributions when $h(t)$ is represented by nine sections ($N = 9$). The broken lines represent the distributions for $N = 1$. The sample size was 10,000 in both cases. As seen from the graph in this example, an integrate-and-dump equalizer analysis can be used as a first-order approximation to an equalizer design with an impulse response similar to the $h(t)$ shown in Fig. 10.

The minimum error-rate and optimum threshold were

$$N = 9 \quad P_{e\min} = 3.03 \times 10^{-10} \quad \theta = 43.61 \text{ percent}$$
$$N = 1 \quad P_{e\min} = 6.6 \times 10^{-11} \quad \theta = 42.68 \text{ percent.}$$

Figure 14 again shows the distribution of the above example ($N = 9$). The circled points are results obtained using Chernoff bound approximations as obtained by J. H. Bobsin[8] using Personick's[3] method. As seen from the graph, the optimal threshold predicted by the Chernoff approximation is identical with the one obtained by the statistical analysis. The minimum error rate shown by the Chernoff bound is about two orders of magnitude larger from the statistically simulated value. However, the error rate obtained by the Chernoff approximation is an upper bound. Therefore, if the result falls within the acceptable design specifications, no further analysis is necessary. If, on the other hand, the predicted error rate is marginal, a more accurate analysis technique such as the one described in this paper is advisable.

Figure 15 depicts curves for the minimum error rate $P_{e\min}$ and optimum threshold as a function of the standard deviation of the

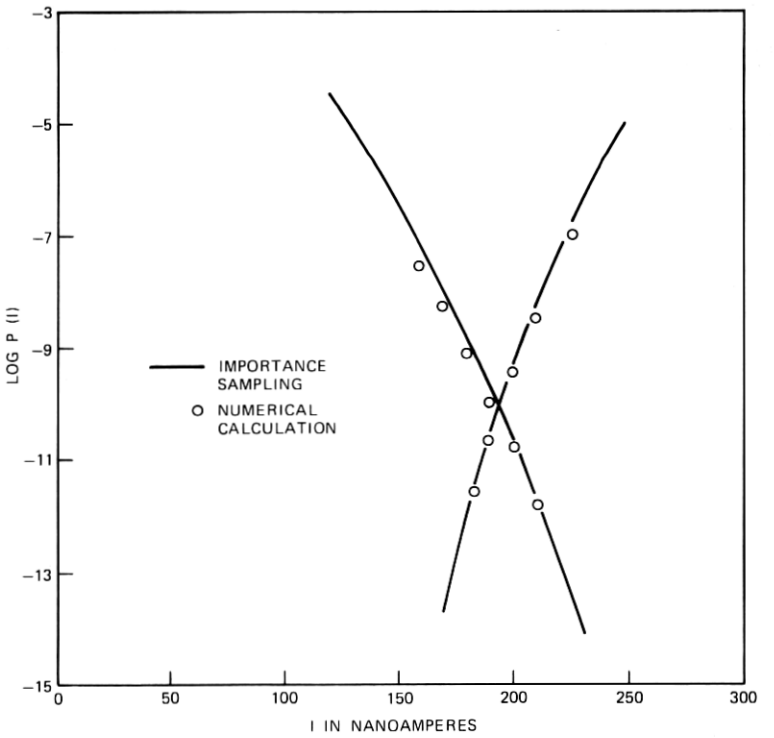


Fig. 12—Distributions for ONE and ZERO signals.

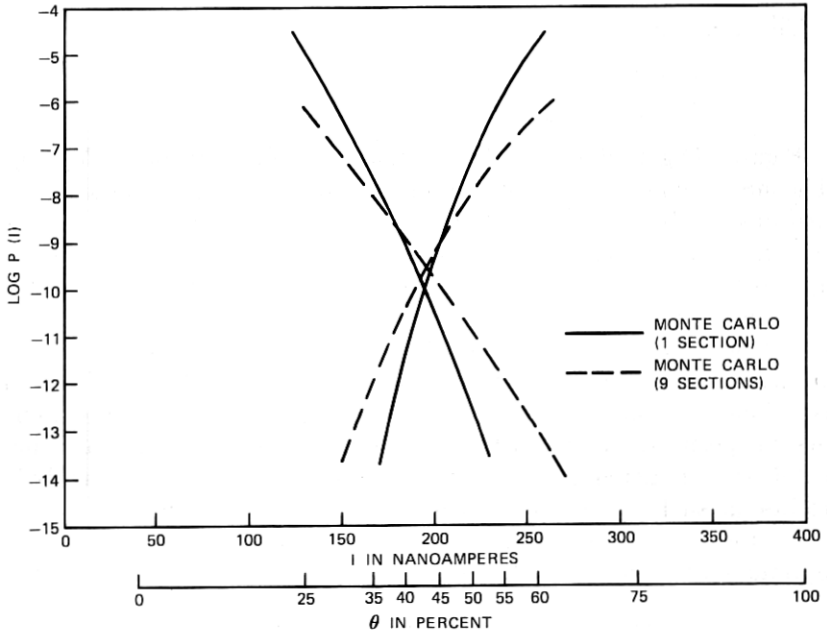


Fig. 13—Distributions for ONE and ZERO signals using Monte Carlo method for 1 and 9 sections.

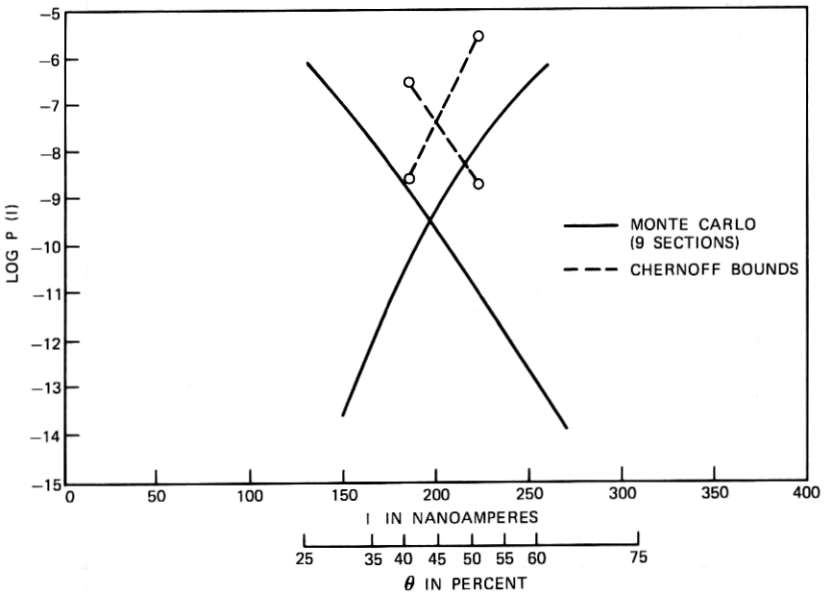


Fig. 14—Distributions for ONE and ZERO signals using Monte Carlo method and Chernoff bounds.

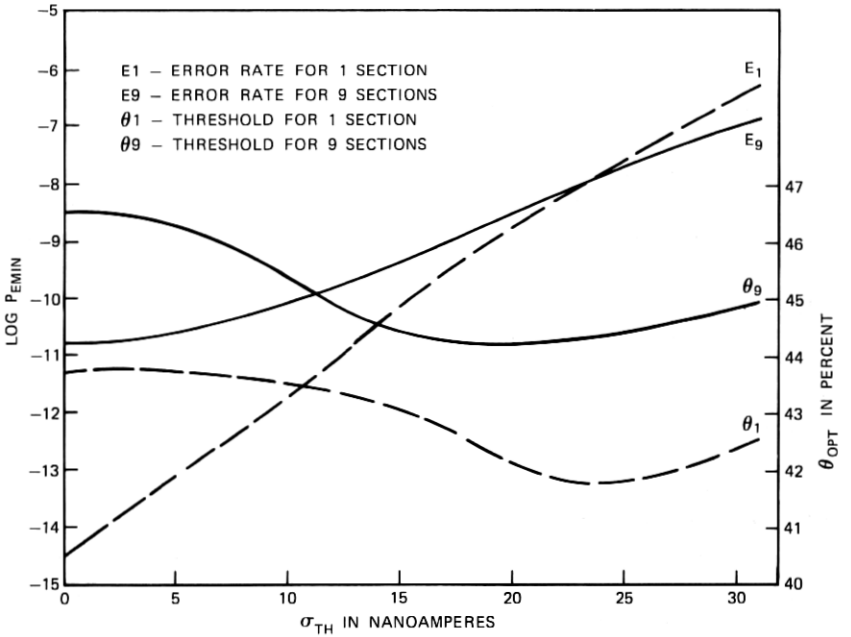


Fig. 15—Minimum error rate and optimum threshold for 1 and 9 sections as a function of thermal noise.

thermal noise σ_{TH}. For this example,

$$\sigma_{TH} = \frac{23.32}{\sqrt{R_{TH}}} [\text{nA}],$$

where R_{TH} is the input impedance of the equalizer in kilohms.

The solid lines represent $P_{e\,min}$ and θ_{opt} for $N = 9$. The broken lines depict the same case analyzed with $N = 1$.

As seen from the graphs, the optimum threshold for both types of receivers is between 42 and 46 percent. Interestingly, the threshold has a minimum at $\sigma = 25$ nA ($R_{TH} = 0.87$ kilohm).

The integrate-and-dump analysis has a lower error rate then the nine-section equalizer analysis in the normal operating range of the repeater.

Throughout this analysis, the avalanche gain was assumed to be constant. In a more realistic design, the avalanche gain can be adjusted to optimize the error rate. This will result in a higher avalanche gain for higher thermal noise.

In this example, the input signals and the equalizer impulse response were chosen so that there is no intersymbol interference. This was done to simplify the example. However, the analysis procedure is not limited to this case.

VII. CONCLUSIONS

The statistical importance-sampling technique was used to calculate error rates of optical repeaters. The application of this technique substantially reduced the number of samples needed for simulation of the repeater and thus made computer simulation practical.

The importance-sampling technique is very effective and has general applicability in areas where tails of distributions have to be computed or particular sensitive areas of the statistical distribution have to be emphasized.

This method cannot, however, be used automatically; a good prior knowledge of the behavior of the system is necessary. To get an accurate statistical simulation of a section of an output, all the important regions of input PDFs which may contribute to this section have to be emphasized.

VIII. ACKNOWLEDGMENTS

The author is grateful to D. D. Sell and S. D. Personick for introducing him to problems related to optical communication and for helpful discussion on the subject, C. L. Semmelman for his comments and suggestions in the area of statistical sampling techniques, and R. B. Hawkins for guidance and encouragement.

REFERENCES

1. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1968.
2. S. D. Personick, "New Results on Avalanche Multiplication Statistics with Applications to Optical Detection," B.S.T.J., *50*, No. 1 (January 1971), pp. 167–189.
3. S. D. Personick, "Statistics of a General Class of Avalanche Detectors With Applications to Optical Communications," B.S.T.J., *50*, No. 10 (December 1971), pp. 3075–3095.
4. R. J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory," IEEE Trans. Electron Dev., *ED-19* (June 1972), pp. 703–713.
5. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," RCA Rev., *35* (June 1974), pp. 234–276.
6. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems, I and II," B.S.T.J., *52*, No. 6 (July–August 1973), pp. 843–886.
7. D. D. Sell, unpublished material.
8. S. D. Personick, P. Balaban, and J. H. Bobsin, unpublished material.
9. Special issue on "Statistical Circuit Design," B.S.T.J., *50*, No. 4 (April 1971).
10. H. Kahn, "Use of Different Monte Carlo Sampling Techniques," Symposium on Monte Carlo Methods, edited by H. A. Mayer, New York: John Wiley, 1956, pp. 146–190.
11. R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics*," New York: MacMillan, 1970.
12. S. B. Weinstein, "Estimation of Small Probabilities by Linearization of the Tail of a Probability Distribution Function," IEEE Trans. Commun. Technol., *19*, No. 6 (December 1971), pp. 1149–1155.