Biol. Cybern. 43, 87–96 (1982)

Biological Cybernetics
© Springer-Verlag 1982

Multiplication Noise in the Human Visual System at Threshold: 2. Probit Estimation of Parameters*

Paul R. Prucnal and Malvin Carl Teich

Columbia University, New York, NY, USA

Abstract. A mathematical technique is described that relates detection model parameters to stimulus magnitude and experimental probability of detection. The normalizing transform is used to make the response statistics approximately Gaussian. Conventional probit analysis is then applied. From measurements at M stimulus levels, a system of M equations is solved and estimates of M unknown parameters of the detection model are obtained. The technique is applied to a threshold vision model based on additive and multiplicative Poisson noise. Results are obtained for the parameter estimates for individual subjects, and for the standard deviation of the estimates, for various values of the stimulus energy and number of trials. A frequency-of-seeing experiment is performed using a point-source stimulus that randomly assumes 3 energy levels with 200 trials per level. With a central efficiency of 50%, the estimated ocular quantum efficiency for our four subjects lies between 12% and 23%, the average dark count at the retina lies between 8 and 36 counts, and the threshold count for our (low false-report rate) data lies between 11 and 32. The theoretical results reduce to those obtained by Barlow (J. Physiol. London **160**, 155–168, 1962), in the absence of dark light and multiplication noise.

1. Introduction

Probit analysis provides a convenient mathematical device for converting a normal sigmoid response curve into a straight line. This enables quantitative data derived from certain binary-response experiments, e.g., frequency-of-seeing data, to be readily analyzed. Though the technique appears to have originated with Fechner (1860) in his discussion of the relationship of the difference between two weights to the proportion of trials in which a subject correctly judges which was the heavier, it has been most extensively developed in connection with biological assays (Gaddum, 1933; Bliss, 1934, 1935; Finney, 1952). It was Bliss (1934) who coined the term "probit", an abbreviation for *prob*ability un*it*.

In 1962, Barlow used the method to estimate the overall quantum efficiency of visual discriminations. Following the model set down by Hecht et al. in 1942, he assumed that Poisson photon fluctuations governed the variability of subject response in vision experiments at low light levels. In this model the frequency of seeing, as a function of the average number of photons delivered to the cornea in a brief flash of light, is determined by the threshold count number at the retina, t, and by the quantum efficiency η. Additive noise and loss in the optic tract are not considered. Barlow demonstrated that, in place of the array of experimental stimulus energies generally used to trace out such frequency-of-seeing curves, data at two stimulus energies sufficed to provide estimates (and sampling errors of the estimates) for the two model parameters t and η. The two-point method promised, and delivered, greater accuracy than the full-curve method because of the optimal use of the limited number of trials available with human subjects. It offered further advantages in the simplicity of the experimental procedure and calculation.

Inasmuch as the probit method is based on a normally distributed response variable, the Poisson fluctuations of the number of effective photons absorbed at the retina had first to be converted to a normal random variable. Barlow used the square-root normalizing transform for this purpose, though slightly more accurate normalizing transforms are available (Bartlett, 1947; Prucnal and Teich, 1980).

* This work was supported by the National Science Foundation

0340-1200/82/0043/0087/$02.00

In a similar way, we use probit analysis to estimate various threshold parameters in human vision. But whereas Barlow emphasized the estimation of quantum efficiency in the context of the Hecht-Shlaer-Pirenne model, we approach the problem first from a general point of view, and then in the context of a recently developed neural-counting model for threshold vision that incorporates stimulus fluctuations, dark light, and multiplication noise (Teich et al., 1982a). In our treatment, the underlying neural counting distribution is the Neyman Type-A (Neyman, 1939; McGill, 1967, 1971) rather than the Poisson. In the absence of dark light and multiplication noise, our result reduces to that obtained by Barlow.

In Sect. 2 we provide a general theoretical treatment of parameter estimation for the threshold detection problem. We consider an example in which a Poisson process drives the mean of a second stochastic process. A special case of such "multiplied processes" is the shot-noise-driven doubly stochastic Poisson point process (Saleh and Teich, 1982a), which we have previously related to the maintained dark discharge in the retinal ganglion cell (Teich and Saleh, 1981). In Sect. 3, we apply the results of Sect. 2 to vision at threshold, and the accuracy of the method is determined. In Sect. 4, the probit method is used in conjunction with the frequency-of-seeing data recently collected by Teich et al. (1982a), to provide estimates for the threshold parameters defined by their model. The conclusion is presented in Sect. 5. In Part 3 of this set of papers (Teich et al., 1982b) we deal with quantum fluctuations for non-Poisson stimuli.

2. Theory

2.1. General Detection Problem

A general class of detection problems is defined as follows. The response x to a stimulus E is represented by a random variable X, with conditional probability density function (pdf) $p_X(x:E)$. The pdf $p_X(x:E)$ is expressed in terms of its statistical parameters $q_1, q_2, \ldots$. If a mathematical model with parameters $a_1, a_2, \ldots$ is proposed to describe the relationship of response to stimulus, then the q's may be defined in terms of the a's. For example, a simple mathematical model might specify that the mean response $q_1 = \langle x \rangle$ is proportional, with constant a_1, to the stimulus E, viz.

$$\langle x \rangle = a_1 E. \tag{1}$$

More complex models with additional parameters can easily be imagined. Based upon measurements of the response x it is desired to obtain values for the parameters $a_1, a_2, \ldots$ of the stimulus-response model. Often the response cannot be measured directly however, but rather it is convenient to simply determine whether the response exceeds a threshold t, that is, whether the response is detected. The probability of detection is then (van Trees, 1968)

$$P_D(E, t, a_1, a_2, \ldots) = \int_t^\infty p_X(x:E)\, dx. \tag{2}$$

An experimental estimate based upon N_i measurements of P_D corresponding to a stimulus level E_i will be denoted p_i. Thus, based upon experimental estimates p_i, it is desired to obtain values for the parameters $t, a_1, a_2, \ldots$ of the stimulus-response model, as described in the following analysis.

Assume that a transformation of random variables $\xi = h(x)$ is performed, and that the transformed random variable is Gaussian with mean $\langle h(x:E) \rangle$ and unity variance. The form of such a "normalizing transformation" has received considerable attention for many years in the statistics literature (Bartlett, 1936; Cochran, 1940; Curtiss, 1943; Anscombe, 1948; Freeman and Tukey, 1950) and more recently has been used to study increment threshold laws for arbitrary stimulus statistics in psychophysics (Prucnal and Teich, 1980) and to characterize and evaluate the performance of optical detection systems having non-Gaussian statistics (Prucnal, 1980a, b). The transform $\xi = h(x)$ can be chosen to be monotonic (Prucnal and Teich, 1980), which implies that the inverse transform $x = h^{-1}(\xi)$ exists. The exact form of $h(x)$ is known for lognormal (Prucnal and Teich, 1980), chi-squared and non-central chi-squared random variables (Saleh, 1978). A number of approximate normalizing transforms have been obtained for specific cases such as the Poisson, binomial, and negative binomial (Bartlett, 1936; Cochran, 1940; Curtiss, 1943; Anscombe, 1948; Freeman and Tukey, 1950). In particular the square-root normalizing transform is commonly used for Poisson random variables (Barlow, 1962; McGill, 1971). Furthermore, if the variance σ_x^2 can be expressed as a function of the mean $\langle x \rangle$

$$\sigma_x^2 = D^2(\langle x \rangle), \tag{3}$$

then an approximate normalizing transform is given by

$$h(x) \simeq \int \frac{d\langle x \rangle}{D(\langle x \rangle)}, \tag{4}$$

(Kendall and Stuart, 1966; Prucnal and Teich, 1980; Prucnal, 1980b) where the integral is evaluated at $\langle x \rangle = x$, and where it is understood that $\langle x \rangle$ may be a function of E. It is apparent that (4) leads to the above-described square-root transform for Poisson random variables. Equation (4) has also been used in cases where the function $D(\langle x \rangle)$ is more complex, and the

accuracy of the approximation has been discussed (Prucnal, 1980b).

For cases where a normalizing transform is known, (2) can be expressed in terms of the transformed random variable ξ as

$$\psi[\langle h(x:E)\rangle - h(t)] = P_D(E,t), \tag{5}$$

where t is the threshold, E is the stimulus, and ψ is the standard normal distribution function

$$\psi[z] = \int_{-\infty}^{z} (2\pi)^{-1/2} \exp(-v^2/2)\, dv. \tag{6}$$

Equation (6) expresses the probability ψ in terms of an integral which is cumbersome to deal with analytically. Instead, it is more convenient to employ the probit function Y (Finney, 1952) that relates the probability ψ to the upper limit z of the integral in (6) through the simple relation

$$Y(\psi) = z + 5. \tag{7}$$

Substituting (5) in (7) we obtain

$$Y(P_D) = 5 + \langle h(x:E)\rangle - h(t). \tag{8}$$

2.2. Single-Parameter Fit

Let the probability of detection p_1 be measured for a single stimulus level E_1. Equation (8) can be solved explicitly for the threshold t, yielding

$$t = h^{-1}[\langle h(x:E_1)\rangle + 5 - Y(p_1)], \tag{9}$$

where $h^{-1}(\cdot)$ is the inverse of the normalizing transform. Considering that p_1 is a measured quantity subject to binomial sampling error, it is a random variable. The variance of t can be expressed in terms of the variance of p_1 as

$$\sigma_t^2 \simeq \left(\frac{dt}{dY}\right)^2 \left(\frac{dY}{dP_D}\right)^2 \sigma_{p_1}^2, \tag{10}$$

where an approximate result for the variance of a function of a random variable has been utilized (Papoulis, 1965, p. 152). If the estimate of the binomial random variable p_1 is based upon N_1 measurements of the response to a stimulus E_1, (10) becomes

$$\sigma_t^2 \simeq \left(\frac{dt}{dY}\right)^2 \left[\left(\frac{dY}{dP_D}\right)^2 p_1(1-p_1)\right] \frac{1}{N_1}. \tag{11}$$

The term in square brackets in (11) is the inverse of the "weighting coefficient" w_1, and has been tabulated as a function of Y (Finney, 1952, p. 250). Using (9) to evaluate the derivative of t with respect to Y, (11) becomes

$$\sigma_t^2 \simeq \{h^{-1}[\langle h(x:E_1)\rangle + 5 - Y(p_1)]'\}^2 \frac{1}{N_1 w_1} \tag{12}$$

where the prime denotes the derivative of h^{-1}.

2.3. Multiple Parameter Fits

We address the two parameter case first. To obtain the parameters t and a_1, the probabilities of detection p_1 and p_2 are measured at two stimulus levels E_1 and E_2. Evaluating (8) using E_1, E_2, p_1, and p_2 and subtracting yields

$$\langle h(x:E_1)\rangle - \langle h(x:E_2)\rangle = d, \tag{13}$$

where

$$d = Y(p_1) - Y(p_2). \tag{14}$$

Note that (13) corresponds to the generalized increment threshold law derived by Prucnal and Teich (1980) and that the dependence on the threshold t has been eliminated. For a two parameter model, (13) can first be solved for a_1 as a function of E_1, E_2, p_1, and p_2. This value of a_1 can then be used to solve for t using (9). In a similar manner, to solve for $M+1$ parameters, the probabilities of detection $p_1, p_2, \ldots, p_{M+1}$ are estimated for stimuli $E_1, E_2, \ldots, E_{M+1}$. Using (8) yields a system of $M+1$ equations which can then be solved for the $M+1$ unknowns $t, a_1, \ldots, a_M$.

2.4. Application to Multiplied Processes

As an example of the single and multiple parameter fitting procedure, consider a model in which the pdf $p_X(x:E)$ corresponds to a multiplied process (Saleh and Teich, 1982a; Matsuo et al., 1982) in which

$$\sigma_x^2 = (\varrho+1)\langle x:E\rangle, \tag{15}$$

where $\varrho = a_1$. Here it is assumed that the expected value of x is conditioned on E. Utilizing (3) and (4), the normalizing transform is

$$h(x:E) \simeq 2(x)^{1/2}(\varrho+1)^{-1/2}. \tag{16}$$

The single-parameter fit of the threshold t yields

$$t \simeq \frac{(\varrho+1)}{4}\left[2\left(\frac{\langle x:E_1\rangle}{\varrho+1}\right)^{1/2} + 5 - Y(p_1)\right], \tag{17}$$

with variance

$$\sigma_t^2 \simeq (\varrho+1)t/N_1 w_1, \tag{18}$$

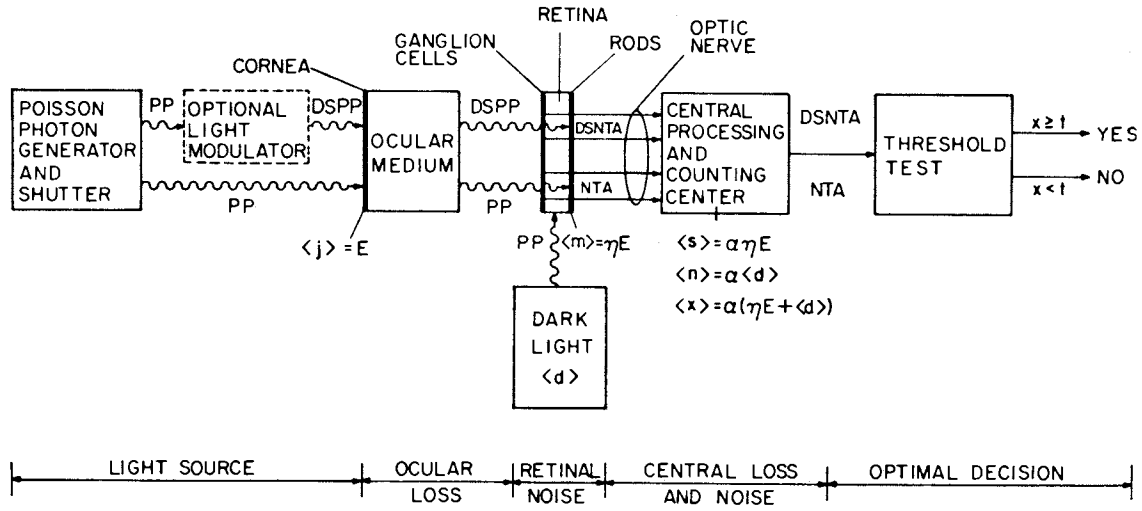

Fig. 1. Block diagram of model for visual system processing at threshold. Only the lower light path, which excludes the optional light modulator, is considered here. For $\alpha \to 0$, with $\langle s \rangle$, $\langle n \rangle$, and $\langle x \rangle$ fixed, the model becomes identical to that proposed by Barlow (1956). For $\alpha \to 0$ and $\langle c \rangle = \langle n \rangle = \langle d \rangle \to 0$, with $\langle s \rangle$ and $\langle x \rangle$ fixed, the model reduces to that used by Hecht et al. (1942) and by Barlow (1962). [After Teich et al. (1982a)]

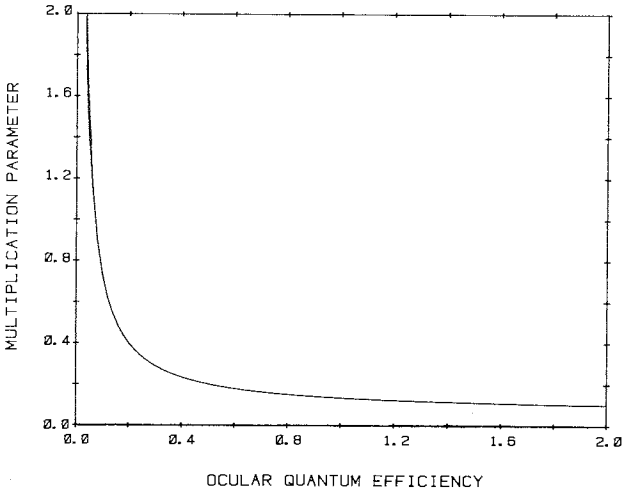

Fig. 2. Multiplication parameter α vs. ocular quantum efficiency η normalized so that $\eta = 0.2$ for $\varrho = \alpha = \frac{1}{2}$

from (9) and (12). For a two parameter fit, using (13), ϱ can first be obtained as solution to the equation

$$\langle x : E_1 \rangle^{1/2} - \langle x : E_2 \rangle^{1/2} \simeq \frac{(\varrho+1)^{1/2}}{2} d , \tag{19}$$

and using this value of ϱ, t can then be obtained from (17).

3. Application to Vision at Threshold

The above fitting procedure for multiplied processes can be applied to the threshold vision model proposed by Teich et al. (1982a). In this model (see Fig. 1) a stimulus energy E is transformed into a discrete sequence of neural impulses which ultimately constitutes a flow of information in the sensory channel. Noise impulses, corresponding to dark light, and indistinguishable from the stimulus impulses, are present in the channel. The sequence of impulses is monitored at the counting center. The number of impulses x counted in a fixed interval T_c is compared with a threshold t. This form of processing constitutes an optimal decision (Prucnal and Teich, 1978, 1979).

The transformation from stimulus energy to impulses is governed, within the context of this model, by a multiplied process with mean

$$\langle x : E \rangle = \alpha\eta(E + \langle c \rangle) \tag{20}$$

where $\alpha = a_2$ and with variance given by (15). The ocular quantum efficiency η is defined as the fraction of photons incident on the cornea that contribute to activation of the rods. The parameter $\langle n \rangle$ represents the average number of noise impulses counted in the interval T_c, whereas $\langle d \rangle = \langle n \rangle / \alpha$ and $\langle c \rangle = \langle n \rangle / \alpha\eta$ are the corresponding quantities at the retina and cornea, respectively.

This vision model relates the stimulus E to the response x through a multiplied Poisson process with parameters t, $a_1 = \varrho$, $a_2 = \alpha$, $a_3 = \eta$, and $a_4 = \langle d \rangle$. It is desired to obtain values for these parameters based upon experimental estimates of the probability of detection. Substituting (20) in (17), the threshold is obtained as a function of the measurements E_1 and p_1 and the remaining parameters ϱ, α, η and $\langle c \rangle$,

$$t \simeq \frac{(\varrho+1)}{4}\left[2\left(\frac{\alpha\eta(E+\langle c \rangle)}{\varrho+1}\right)^{1/2} + 5 - Y(p_1)\right]^2 . \tag{21}$$

Substituting (20) in (19), the quantum efficiency

$$\eta=\frac{\varrho+1}{4\alpha}\left[\frac{Y(p_2)-Y(p_1)}{\sqrt{E_2+\langle c\rangle}-\sqrt{E_1+\langle c\rangle}}\right]^2 \tag{22}$$

is obtained as a function of the measurements E_1, E_2, p_1, and p_2 and the remaining parameters ϱ, α and $\langle c\rangle$. Evaluating (22) at the stimulus levels E_1 and $E_3 \neq E_2$, and equating the result to (22) obtains the effective average number of noise impulses at the cornea as a function of the measurements E_1, E_2, E_3, p_1, p_2, and p_3,

$$\langle c\rangle=\frac{(y^2E_3-z^2E_2)^2+2E_1(y^2E_3+z^2E_2)-E_1^2}{4yz(zE_2-yE_3+E_1)}, \tag{23}$$

where

$y=[Y(p_1)-Y(p_2)]/[Y(p_3)-Y(p_2)]$

and

$z=y-1$.

The values of the remaining parameters ϱ and α can be obtained provided the additional measurements E_4, E_5 and p_4, p_5 are performed. However, the resulting set of five equations in five unknowns (19) does not lend itself to convenient analytic solution, and therefore must be solved numerically. This slight complication can be avoided by assigning the value $\frac{1}{2}$ to ϱ and α, in accordance with the value proposed by Teich et al. (1982a). It is apparent from (23) that $\langle c\rangle$ is invariant to the particular choice of ϱ and α. On the other hand, η and t are functions of ϱ and α as seen in (22), and (21). For η, the dependence on ϱ and α enters through a proportionality factor $(\varrho+1)/\alpha$. For example, if the choice $\varrho=\alpha=\frac{1}{2}$ and the above-described measurements lead to the estimate $\eta=0.20$, then the estimate of η for other choices of $\alpha=\varrho$ is shown in Fig. 2. In a similar way, the dependence of the estimate of t on the choice of ϱ and α can also be determined.

The vision model, then, has three unknown parameters which can be estimated from the response to three stimulus levels. From the above solution, $\langle c\rangle$ is determined from these three responses [Eq. (23)], η is determined from two responses and $\langle c\rangle$ [Eq. (22)], $\langle d\rangle$ is determined from $\langle d\rangle=\eta\langle c\rangle$, and t is determined from one response, $\langle c\rangle$ and η (21).

The above solution is approximate since the normalizing transformation (16) used at the outset was approximate. The accuracy of the solution is tested as follows. Known parameters are used to generate theoretical probabilities of detection p_i, $i=1,2,3$. Substituting the values of p_i into (21)–(23) yields estimates of the parameters t, η, and $\langle c\rangle$. The estimates are then compared with the known parameter values.

As an example, let the known parameter values be $\alpha=\frac{1}{2}, t=28, \eta=0.20$, and $\langle c\rangle=150$, and let the counting distribution associated with the multiplied Poisson process be the Neyman Type-A, for which $\varrho=\alpha$ (see Teich, 1981). From (2) and (20), with $p_X(x:E_i)$ given by the Neyman Type-A, detection probabilities p_i can be generated, and the corresponding probits $Y(p_i)$ can be determined from a table (Finney, 1952). For stimulus energies $E_1=0, E_2=23.8$, and $E_3=226.4$ photons at the cornea, the corresponding probits are $Y(p_1)=2.670$, $Y(p_2)=3.158$, and $Y(p_3)=6.392$. Substitution of E_i and $Y(p_i)$, $i=1, 2, 3$, into (21)–(23) yields the parameter estimates $t=28.1, \eta=0.202$, and $\langle c\rangle=148.6$, which are very near the known parameter values. The agreement is equally good for other examples.

Since p_1, p_2, and p_3 are measured quantities subject to binomial sampling error, they are random variables. The variance of t can be expressed in terms of the variance of p_1, η, and $\langle c\rangle$ as

$$\sigma_t^2\simeq\left(\frac{\partial t}{\partial Y(p_1)}\right)^2\left(\frac{dY(p_1)}{dP_D}\right)^2\sigma_{p_1}^2 + \left(\frac{\partial t}{\partial\eta}\right)^2\sigma_\eta^2+\left(\frac{\partial t}{\partial\langle c\rangle}\right)^2\sigma_{\langle c\rangle}^2. \tag{24}$$

Evaluating the derivatives and using the definition of the weighting coefficient given after (11), yields

$$\sigma_t^2\simeq t\left(\frac{\varrho+1}{N_1w_1}+\frac{\alpha(E_1+\langle c\rangle)}{\eta}\sigma_\eta^2+\frac{\alpha\eta}{E_1+\langle c\rangle}\sigma_{\langle c\rangle}^2\right). \tag{25}$$

The variance of η can be expressed in terms of the variance of p_1, p_2, and $\langle c\rangle$ as

$$\sigma_\eta^2\simeq\left(\frac{\partial\eta}{\partial(Y(p_1)-Y(p_2))}\right)^2\left[\left(\frac{dY(p_1)}{dP_D}\right)^2\sigma_{p_1}^2 + \left(\frac{dY(p_2)}{dP_D}\right)^2\sigma_{p_2}^2\right]+\left(\frac{\partial\eta}{\partial\langle c\rangle}\right)^2\sigma_{\langle c\rangle}^2. \tag{26}$$

Evaluating the derivatives and substituting the weighting coefficients yields

$$\sigma_\eta^2\simeq\eta^2\left[\left(\frac{1}{N_1w_1}+\frac{1}{N_2w_2}\right)\frac{1}{Y(p_2)-Y(p_1)} + \frac{1}{(E_1+\langle c\rangle)(E_2+\langle c\rangle)}\sigma_{\langle c\rangle}^2\right]. \tag{27}$$

The variance of $\langle c\rangle$ can be expressed in terms of the variance of p_1, p_2, and p_3 as

$$\sigma_{\langle c\rangle}^2\simeq\left(\frac{d\langle c\rangle}{dy}\right)^2\left\{\left(\frac{\partial y}{\partial(Y(p_1)-Y(p_2))}\right)^2\cdot\left[\left(\frac{dY(p_1)}{dP_D}\right)^2\sigma_{p_1}^2+\left(\frac{dY(p_2)}{dP_D}\right)^2\sigma_{p_2}^2\right] + \left(\frac{\partial y}{\partial(Y(p_2)-Y(p_3))}\right)^2\cdot\left[\left(\frac{dY(p_2)}{dP_D}\right)^2\sigma_{p_2}^2+\left(\frac{dY(p_3)}{dP_D}\right)^2\sigma_{p_3}^2\right]\right\}. \tag{28}$$

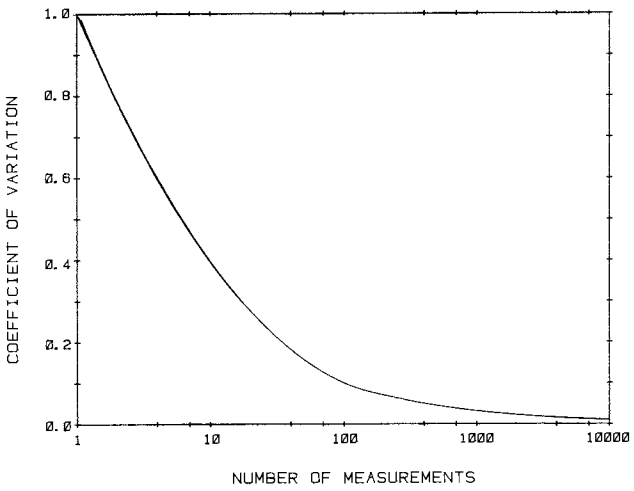

Fig. 3. Relative coefficient of variation c_v vs. number of measurements N_i at each stimulus level E_1, E_2, and E_3

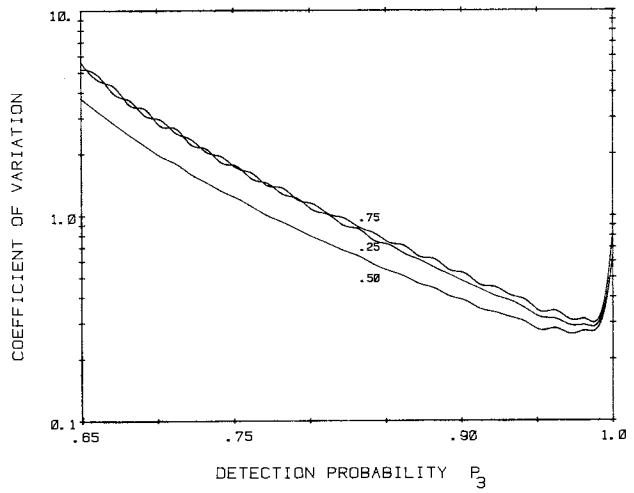

Fig. 4. Coefficient of variation c_v vs. detection probability p_3, where $p_1=1-p_3$. The probability p_2 falls at a specified fraction Δ of the distance between p_1 and p_3. The cases $\Delta=0.25$, 0.50, and 0.75 are shown. For all cases, $\varrho=\alpha=\frac{1}{2}$, $t=28$, $\eta=0.20$, and $\langle c\rangle=150$

Evaluating the derivatives and substituting the weighting coefficients yields

$$\sigma^2_{\langle c\rangle} \simeq \langle c\rangle^2 \left\{ \frac{1-2y}{yz} + \frac{E_3-E_2}{E_1+zE_2-yE_3} + \frac{4(z^2E_2-y^2E_3)(zE_2-yE_3)}{(z^2E_2-y^2E_3)^2-2E_1(z^2E_2+y^2E_3)+E_1^2} \right\} \cdot \left\{ \frac{1}{[Y(p_2)-Y(p_1)]^2}\left(\frac{1}{N_1w_1}+\frac{1}{N_2w_2}\right) + \frac{[Y(p_1)-Y(p_2)]^2}{[Y(p_3)-Y(p_2)]^4}\left(\frac{1}{N_3w_3}+\frac{1}{N_2w_2}\right)\right\}, \quad (29)$$

where y and z are defined in (23). Thus, the variance of t, η, and $\langle c\rangle$ is given in (25), (27), and (29), respectively, as a function of E_1, E_2, E_3, $Y(p_1)$, $Y(p_2)$, $Y(p_3)$, N_1, N_2, N_3, w_1, w_2, w_3 and parameter magnitudes.

A useful measure of the overall uncertainty in the estimate of these three parameters is the coefficient of variation

$$c_v=\left(\frac{\sigma_t^2}{\langle t\rangle^2}+\frac{\sigma_\eta^2}{\langle\eta\rangle^2}+\frac{\sigma^2_{\langle c\rangle}}{\langle c\rangle^2}\right)^{1/2}. \quad (30)$$

Examination of (25), (27), (29), and (30) reveals that c_v is proportional to $1/\sqrt{N_i}$. For the case $N_1=N_2=N_3$, Fig. 3 represents the relative coefficient of variation, normalized to its value at $N_i=1$, plotted against N_i.

Examination of (25), (27), (29), and (30) also reveals that c_v depends on the separation of E_1, E_2, and E_3, as well as the separation of $Y(p_1)$, $Y(p_2)$, and $Y(p_3)$. This implies that a minimum c_v can be obtained from some optimal choice of the three stimulus levels. A good deal of insight into the location of this optimum is obtained from the following example. The coefficient of variation is calculated for the case $\varrho=\alpha=\frac{1}{2}$, $t=28$, $\eta=0.20$, $\langle c\rangle=150$, and $N_i=10^4$. The energies E_1, E_2, and E_3 are selected in such a way that the theoretical detection probabilities p_1 and p_3 fall symmetrically about 50%, and the probability p_2 falls some specified fraction Δ of the distance between p_1 and p_3. The resulting detection probabilities are $p_1=1-p_3$, $p_2=1+p_3(2\Delta-1)-\Delta$, and p_3. The coefficient of variation for this case is shown in Fig. 4 for $\Delta=0.25$, 0.50, and 0.75. The minimum c_v is obtained for the case $p_1=2\%$, $p_2=50\%$, $p_3=98\%$. The behavior of c_v for $\Delta=0.25$ and $\Delta=0.75$ is practically identical. The coefficient of variation is always least for $\Delta=0.5$. The undulations in the curves are artifacts introduced by the discrete nature of the search algorithm used to find the nearest E_i producing the desired p_i. The location of the minimum is consistent with that determined by Barlow (1962) for the two-point case. A heuristic interpretation of the result is that a maximum separation between p_1, p_2, and p_3 is desirable, except that for p_3 too close to 100%, the binomial sampling error becomes very large. As a rule-of-thumb, in performing an experiment, the stimulus levels should be selected in such a way that the probabilities p_1, p_2, and p_3 fall in the vicinity of 2%, 50%, and 98%.

4. Experiment

The parameter estimation procedure described in the last section will now be applied to data collected by Teich et al. (1982a). Based upon the experimental frequency of seeing at three stimulus levels, estimates of the parameters t, η, and $\langle c\rangle$ will be calculated. These estimates are within the context of the threshold vision

model discussed in Sect. 3. As described earlier, to maintain a low value of c_v for the estimates, p_1 must fall in the vicinity of 2%. For this reason, only the low false-positive-rate data ($p_1 \leqq 2.2\%$) of Teich et al. will be utilized. Since the experimental procedure and results were previously reported in detail in Part 1 of this set of papers, only the essentials are repeated below.

The subject viewed a small dim red fixation target produced by a Maxwellian system and received a 5′ disk-shaped plane-polarized light stimulus in the left eye at $17\frac{1}{2}^\circ$ horizontal eccentricity on the temporal retina. The stimulus was generated by a feedback-stabilized Spectra-Physics Model 162 Ar^+ ion laser that oscillated at 514.5 nm. The flash energy was controlled by an acousto-optic light modulator and the light was attenuated by neutral-density filters. An electronic shutter was opened just before, and closed just after, the presentation of the stimulus, to minimize the transmission of stray light to the subject. Three subject's switches were used to START the trial, and to indicate YES and NO responses. Absolute photometric calibrations were made using an EG&G radiometer with a silicon photodiode at the front end, substituted in place of the subject's eye.

Subjects were four males ranging in age from 23 to 38 years. All had normal vision. Alignment in the apparatus was maintained by the use of a dental-impression mouthbite. A trial consisted of the presentation of a 1-ms flash of light or of a blank. The flash energy was chosen quasi-randomly from one of 10 mean levels, separated by 0.115 log units; this yielded a total range of about 1 log unit. A block of trials consisted of 5 trials at each mean energy level plus 10 blanks, for a total of 60 trials per block. A session began with 35 min of total dark adaptation, interrupted only by 2 to 4 blocks of preliminary experimental trials, to allow the subject to re-acquaint with the task. An experimental session was typically comprised of 6 blocks of trials and lasted from 1 to $1\frac{1}{2}$ h.

Frequency-of-seeing data were generated by this method, and these were fit by theoretical probability-of-detection curves associated with the model illustrated in Fig. 1. The criterion used for "best fit" was the minimum sum-of-squares error between the data and the corresponding theoretical probabilities of detection. The first fitting procedure used was, perhaps, the simplest that could be envisioned: the parameters $\varrho = \alpha$, η, and $\langle d \rangle$ were fixed at constant values while the one remaining parameter, the threshold count t, was permitted to vary parametrically. Suitable estimates for the collective group of subjects were obtained essentially by trial and error. It was found that the theoretical family $\{\varrho = \alpha = 0.5, \eta = 20\%, \langle d \rangle = 30\}$ collectively fit all of the frequency-of-seeing data obtained with Poisson light. For the 4 sets of low false-positive rate (FPR) data, the best-fitting threshold values ranging between $t = 26$ and $t = 28$ (see Fig. 2, Teich et al., 1982a). The minimum sum-of-squares ranged between 0.0122 and 0.0546 for these four cases. Alternative one-parameter individual fits to the same data were also obtained with parameters in the following ranges: $\{\varrho = \alpha = 0.5, \eta = 20\%, 16 \leqq \langle d \rangle \leqq 38, 18 \leqq t \leqq 32\}$. The sum-of-squares error using this second procedure ranged between 0.0176 and 0.0498.

In the fitting procedure used by Teich et al. (1982a), for individual subjects, one parameter (t) was estimated from measurements at 11 stimulus energies. Because the total number measurements that could be reliably obtained during an experimental session was small (approximately 600), only about 60 measurements were obtained at each stimulus level. The small magnitude of N_i resulted in rather unreliable estimates of the parameter t.

In the present analysis, the number of stimulus energies required equals the number of parameters to be estimated. In this way, the limited number of measurements available are optimally distributed over the fewest possible number of stimulus levels. For example, assuming that all of the other pertinent parameters were known *a priori*, the single parameter t could be estimated by concentrating all 600 measurements at one stimulus energy, rather than distributing these measurements over 11 stimulus energies.

Carrying this argument further, three parameters can be estimated by distributing the 600 measurements over three stimulus energies. In this way, using the same number of total measurements (approximately 600), more parameters can be estimated (3 rather than 1) with greater reliability (200 trials per point rather than 60), than with the procedure used in Part 1.

We would not expect to gain an improvement in accuracy by re-processing the (60 trial per point) 11-point data of Teich et al. in a manner that utilizes only three of these data points. In fact, one would expect to loose accuracy by using only 180 measurements compared to approximately 600 measurements. Nevertheless, we will proceed with this re-interpretation for two reasons: it provides an illustration of the new technique; it allows the estimation of 3 parameters instead of just one. Clearly, it would be preferable to perform a suitably designed experiment with the measurements concentrated at three carefully chosen stimulus levels. Such a preliminary experiment, with 200 trials at three levels, was performed and will be discussed subsequently.

For each of the 4 sets of low FPR data, then, 3 stimulus energies are chosen, together with the associated experimental values of the frequency of seeing. Using (21)–(23), (25), (27), (29), and (30), estimates of t, η,

Table 1. Results of the 3-point parameter estimation technique applied to the frequency-of-seeing data collected by Teich et al. (1982a). Bottom row represents data collected at three energies with 200 trials/point

Subject/ Condition	Average stimulus energy at cornea E (number of photons)	Experimental frequency of seeing (%)	Associated Probit Y	Number of trials N	Assumed multiplication parameters $\alpha=\varrho$	Estimated ocular quantum efficiency η (%)	Estimated average dark noise at retina $\langle d\rangle$ (counts)	Estimated threshold at counting center t (counts)	Coefficient of variation c_v
PRP	0	1	2.67	120	0.5	23 ± 36	36 ± 88	32 ± 67	3.59
(low FPR)	97.3	32	4.53	60					
	170.9	76	5.71	60					
MEB	0	2.2	2.99	120	0.5	18 ± 28	22 ± 35	20 ± 36	2.90
(low FPR)	97.3	44	4.85	60					
	226.4	96	6.75	60					
GV	31.6	5	3.36	60	0.5	15 ± 25	19 ± 14	16 ± 28	2.48
(low FPR)	97.3	55	5.13	60					
	226.4	97	6.88	60					
MCT	0	1.5	2.83	120	0.5	12 ± 25	8 ± 13	11 ± 26	3.49
(low FPR)	129	52	5.05	60					
	300	98	7.05	60					
PRP	0	4.5	3.30	200	0.5	17 ± 15	10 ± 14	11 ± 12	1.95
(intermediate	31.5	20.5	4.18	200					
FPR)	129	88.5	6.20	200					

and $\langle d\rangle$ are calculated, together with the corresponding standard deviation and overall c_v.

The results of our analysis are presented in the top four rows of Table 1. [The bottom row of Table 1, labeled PRP (intermediate FPR), will be discussed subsequently.] Inspection of the entries reveals that, when $\varrho=\alpha$ is taken to be 0.5, the ranges of estimated parameters for our four subjects are $12\%\leqq\eta\leqq23\%$, $8\leqq\langle d\rangle\leqq36$, and $11\leqq t\leqq32$. The coefficient of variation c_v is specified in the last column of Table 1. Though these uncertainties are large, as is evident from Table 1, they are rather confined in terms of the possible ranges of these parameters. The parameters all appear to have sensible values and, indeed, they fall in the same ranges as those obtained for the one-parameter fitting procedures discussed above where η and $\langle d\rangle$ were obtained for the collective group of subjects by trial and error. This three-parameter estimation technique permits the construction of individual sensitivity-reliability curves for each subject (see Fig. 6, Teich et al., 1982a).

Theoretical probability-of-detection curves can be constructed using the probit estimates, and when we compare these with the full 11-point experimental frequency-of-seeing curves, we find that the sum-of-squares errors are just about a factor of 2 larger than those obtained using the 1-parameter fitting procedure (which had the benefit of three times as many data points). Therefore, although only 3 energy values are used (for a total of 240 trials), the fits are comparable with those obtained by means of the 1-parameter procedure (with a total of 720 trials).

The uncertainties can be reduced, of course, by increasing the number of trials, as evidenced by the $1/\sqrt{N}$ behavior that emerges from the accuracy analysis (see Fig. 3). The uncertainty can also be reduced by selecting the appropriate energy values which minimize the coefficient of variation (see Fig. 4). An example is provided in the bottom row of Table 1 where 200 trials were conducted at each of 3 energy values. The coefficient of variation in this case is significantly reduced. It is expected that relocation of the middle stimulus energy nearer the optimal value would have further reduced c_v. These estimates have been used to generate the probability-of-detection curve shown in Fig. 5. The cross ($\times$) on the left ordinate represents the associated theoretical false-positive rate $P_F=0.039$. The experimental data are represented by the three open circles; the left-most data point is the experimental false-positive rate $\hat{P}_F=0.045\pm0.015$. The vertical bars are the $\pm1-\sigma$ brackets.

We note that the procedure used in our experiments is identical to "Procedure 3" used by Barlow (1962). Our estimated values for η are substantially higher than his, however; this is because we include the effects of dark light and multiplication noise. Thus, our measured η represents the ocular quantum efficiency. Note that for $\varrho=\alpha\to0$ and $\langle c\rangle=\langle n\rangle=\langle d\rangle\to0$, with $\langle s\rangle$ and $\langle x\rangle$ fixed, our analysis reduces identically

to that formulated by Barlow (1962), whereas for $\varrho=\alpha\to 0$, with $\langle s\rangle$, $\langle n\rangle$, and $\langle x\rangle$ fixed, our theory provides a probit analysis for Barlow's (1956) model for threshold vision, which incorporates additive Poisson retinal noise and Poisson stimulus fluctuations.

5. Conclusion

A mathematical technique has been described [see (8)] that relates detection model parameters to the stimulus magnitude and experimental probability of detection. The technique relies on the normalizing transform [see (4)] to stabilize the variance of the response. In this way, the statistics of the response can be made approximately Gaussian, and conventional probit analysis can be utilized.

Based upon measurements at M stimulus levels, a system of M equations is solved (see Sects. 2.2 and 2.3) yielding estimates of M unknown parameters of the model. The variance of the estimate was obtained [see (12)] in the single-parameter case.

The technique was shown to be applicable for extracting estimates of threshold parameters of the human visual system (see Sect. 3). For a model based on a multiplied Poisson process with multiplication parameters equal to one-half, the ranges of estimated parameters for four subjects were $12\%\leqq\eta\leqq 23\%$, $8\leqq\langle d\rangle\leqq 36$, $11\leqq t\leqq 36$ (see Table 1). The variance of these estimates was shown to be proportional to $1/\sqrt{N}$ (Fig. 3), and to be sensitive to the separation of stimulus energies and detection probabilities (Fig. 4). For the three-parameter estimate, a relatively small coefficient of variation was obtained with 200 trials per point. To estimate more than three parameters, it is expected that the coefficient of variation will be more sensitive to the separation of stimulus energies and a substantially larger number of trials per point will be necessary to provide satisfactory results.

The analysis was specifically carried out in terms of the Neyman Type-A ($\varrho=\alpha$) since this distribution plays an archetypical role in the neural-counting model for vision at threshold recently developed by Teich et al. (1982a). However the Neyman Type-A is a member of a broad class of distributions in which a Poisson random variable drives the mean of a second stochastic process (see Sect. 2.4), or cascade of such processes (Matsuo et al., 1982). These multiplied processes all share in common a count variance that is proportional to the count mean, and evidently they are closely related to the Poisson in this respect. The proportionality of variance to mean remains whether the process is stationary or nonstationary (Saleh and Teich, 1982b; Matsuo et al., 1982). Thus our analysis, which relies on the square-root transform, can be directly applied to any of these processes, with a simple substitution of constants [see (16)]. As in the case of the Poisson a slight improvement may be available by using a transform more complex than the square-root (Bartlett, 1947; Prucnal and Teich, 1980).

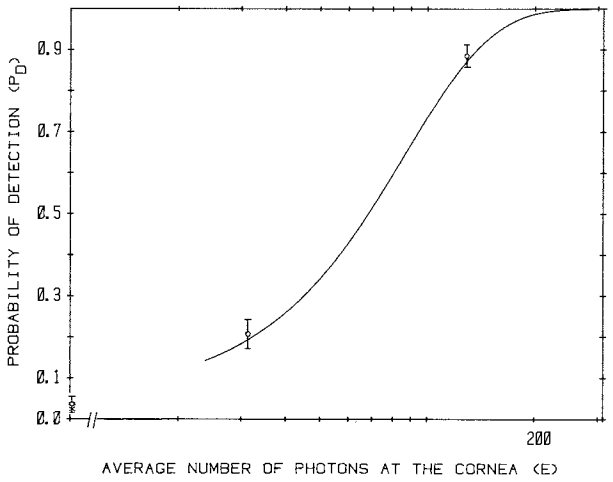

Fig. 5. Theoretical probability-of-detection curve generated by using the probit parameter estimates provided in the bottom row of Table 1 ($\varrho=\alpha=0.5$, $\eta=17\%$, $\langle d\rangle=10$, $t=11$). The cross ($\times$) represents the theoretical false-positive rate $P_F=0.039$. The experimental data are represented by the three open circles; the left-most data point is the experimental false-positive rate $\hat{P}_F=0.045\pm 0.015$. The vertical bars are the $\pm 1-\sigma$ brackets. The 3-point sum-of-squares error is 3.1×10^{-3}

The general formulation presented in Sect. 2 could also be used with signals of very different statistical character. Thus our analysis may be useful for studying the increment threshold behavior of the visual system at supra-threshold energy levels, where effects associated with refractoriness evidently set in (Teich and Diament, 1980; Teich and Saleh, 1981). In this case the associated neural counting distributions are quite different in character from the Neyman Type-A and the Poisson (e.g., the mean is less than the variance) so that a different normalizing transform is appropriate.

The formulation can also be adapted to other paradigms. Though the treatment presented here has been in terms of the non-orthogonal (yes-no) format, it can easily be extended to an orthogonal format (e.g., forced choice) by calculating the statistics of the appropriate random variable (Prucnal, 1981), and the associated normalizing transform. In the same manner, a rating experiment of the kind performed by Sakitt (1972, 1974) can by analyzed by obtaining a probit regression line together with a suitable normalizing transform. Indeed, Sakitt (1974) interpreted her results as suggesting a discrete Gaussian-like internal decision random variable with a variance greater than the mean. The family of multiplied Poisson counting

distributions (e.g., the Neyman Type-A distribution) has just such characteristics (Teich, 1981).

Though the emphasis in this work has been on the application of this model to vision at threshold, it clearly has far greater generality and will likely find use in areas such as biological assay and entomology.

References

Anscombe, F.J.: The transformation of Poisson, binomial and negative binomial data. Biometrika **35**, 246–254 (1948)

Barlow, H.B.: Retinal noise and absolute threshold. J. Opt. Soc. Am. **46**, 634–639 (1956)

Barlow, H.B.: A method of determining the overall quantum efficiency of visual discriminations. J. Physiol. (London) **160**, 155–168 (1962); Measurements of the quantum efficiency of discrimination in human scotopic vision. J. Physiol. (London) **160**, 169–188 (1962)

Bartlett, M.S.: The square-root transformation in the analysis of variance. J. R. Stat. Soc. Suppl. **3**, 68–78 (1936)

Bartlett, M.S.: The use of transformations. Biometrics **3**, 39–52 (1947)

Bliss, C.I.: The method of probits. Science **79**, 38–39 (1934); The method of probits – a correction. Science **79**, 409–410 (1934)

Bliss, C.I.: The calculation of the dosage-mortality curve. Ann. Appl. Biol. **22**, 307–333 (1935)

Cochran, W.G.: The analysis of variance when experimental errors follow the Poisson or binomial laws. Ann. Math. Stat. **9**, 335–347 (1940)

Curtiss, J.H.: On transformations used in the analysis of variance. Ann. Math. Stat. **14**, 107–122 (1943)

Fechner, G.T.: Elemente der Psychophysik. Leipzig: Breitopf und Härtel 1860

Finney, D.J.: Probit analysis (2nd edn.) Cambridge: Cambridge University Press 1952

Freeman, M.F., Tukey, J.W.: Transformations related to the angular and square root. Ann. Math. Stat. **21**, 607–611 (1950)

Gaddum, J.H.: Reports on biological standards. III. Methods of biological assay depending on a quantal response. Spec. Rep. Ser. Med. Res. Coun. (London) **183** (1933)

Hecht, S., Shlaer, S., Pirenne, M.H.: Energy, quanta, and vision. J. Gen. Physiol. **25**, 819–840 (1942)

Kendall, M.G., Stuart, A.: The advanced theory of statistics, Vol. 3. New York: Hafner 1966

Matsuo, K., Saleh, B.E.A., Teich, M.C.: Cascaded Poisson processes. (1982) (submitted for publication)

McGill, W.J.: Neural counting mechanisms and energy detection in audition. J. Math. Psychol. **4**, 351–376 (1967)

McGill, W.J.: Poisson counting and detection in sensory systems. In: Concepts of communications: interpersonal, intrapersonal, and mathematical. Beckenbach, E.F., Tompkins, C.B. (eds.). Chap. 9, pp. 257–281. New York: Wiley 1971

Neyman, J.: On a new class of "contagious" distributions, applicable in entomology and bacteriology. Ann. Math. Stat. **10**, 35–57 (1939)

Papoulis, A.: Probability, random variables, and stochastic processes. New York: McGraw-Hill 1965

Prucnal, P.R.: Generalized performance parameter for single-threshold detection systems. Appl. Opt. **19**, 3606–3610 (1980a)

Prucnal, P.R.: Receiver performance evaluation using photocounting cumulants with application to atmospheric turbulence. Appl. Opt. **19**, 3611–3616 (1980b)

Prucnal, P.R.: Single-threshold processing for orthogonal likelihood-ratio detection with application to FSK fiber-optic communications. IEEE Trans. Commun. Theory **29**, 743–749 (1981)

Prucnal, P.R., Teich, M.C.: Single-threshold detection of a random signal in noise with multiple independent observations. 1: discrete case with application to optical communications. Appl. Opt. **17**, 3576–3583 (1978)

Prucnal, P.R., Teich, M.C.: Single-threshold detection of a random signal in noise with multiple independent observations. Part 2: continuous case. IEEE Trans. Inf. Theory **25**, 213–218 (1979)

Prucnal, P.R., Teich, M.C.: An increment threshold law for stimuli of arbitrary statistics. J. Math. Psychol. **21**, 168–177 (1980)

Sakitt, B.: Counting every quantum. J. Physiol. (London) **223**, 131–150 (1972)

Sakitt, B.: Canonical ratings. Percept. Psychophys. **6**, 478–488 (1974)

Saleh, B.E.A.: Photoelectron statistics. New York: Springer-Verlag 1978

Saleh, B.E.A., Teich, M.C.: Multiplied-Poisson noise in pulse, particle, and photon detection. Proc. IEEE (1982a) (in press)

Saleh, B.E.A., Teich, M.C.: Statistical properties of a nonstationary Neyman-Scott cluster process. (1982b) (submitted for publication)

Teich, M.C.: Role of the doubly stochastic Neyman Type-A and Thomas counting distributions in photon detection. Appl. Opt. **20**, 2457–2467 (1981)

Teich, M.C., Diament, P.: Relative refractoriness in visual information processing. Biol. Cybern. **38**, 187–191 (1980)

Teich, M.C., Prucnal, P.R., Vannucci, G., Breton, M.E., McGill, W.J.: Multiplication noise in the human visual system at threshold: 1. Quantum fluctuations and the minimum detectable energy. J. Opt. Soc. Am. (1982a) (in press)

Teich, M.C., Prucnal, P.R., Vannucci, G., Breton, M.E., McGill, W.J.: Multiplication noise in the human visual system at threshold: 3. Quantum fluctuations for non-Poisson stimuli. (1982b) (submitted for publication)

Teich, M.C., Saleh, B.E.A.: Interevent-time statistics for shot-noise-driven self-exciting point processes in photon detection. J. Opt. Soc. Am. **71**, 771–776 (1981)

van Trees, H.L.: Detection, estimation and modulation theory, Part I. New York: Wiley 1968

Received: May 24, 1981

Prof. Dr. Malvin Carl Teich
Columbia University
Department of Electrical Engineering
Seeley W. Mudd Building
New York, NY 10027
USA