10-4-78
247

DR. 553

MASTER

UCRL-52490

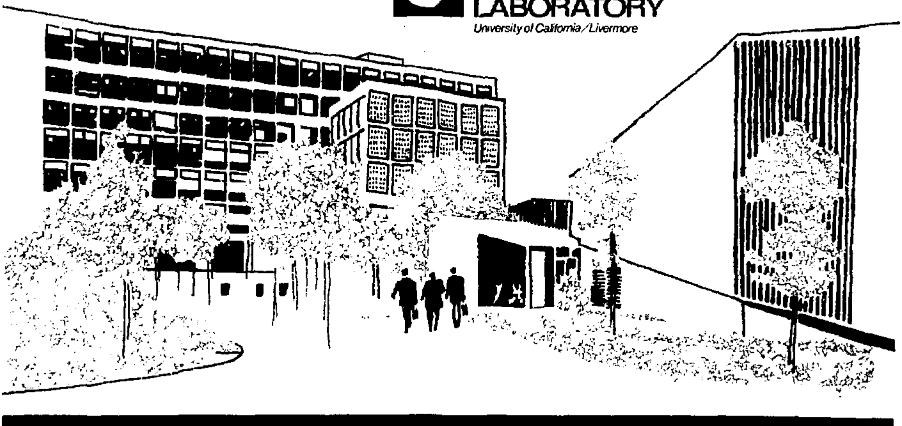
THE APPLICATION OF RANDOM-POINT PROCESSES TO THE DETECTION OF RADIATION SOURCES

J. W. Woods

June 28, 1978

Work performed under the auspices of the U.S. Department of Energy by the UCLLL under contract number W-7405-ENG-48.

LAWRENCE
LIVERMORE
LABORATORY
University of California/Livermore



Distribution Category
UC-41

LAWRENCE LIVERMORE LABORATORY
University of California Livermore, California 94550

UCRL-52490

THE APPLICATION OF RANDOM-POINT PROCESSES TO THE DETECTION OF RADIATION SOURCES

J. W. Woods

MS. date: June 28, 1978

CONTENTS

THE APPLICATION OF RANDOM-POINT PROCESSES TO THE DETECTION OF RADIATION SOURCES

ABSTRACT

In this report we review the mathematical theory of random-point processes and show how use of the theory can obtain optimal solutions to the problem of detecting radiation sources. As noted, the theory also applies to image processing in low-light-level or low-count-rate situations. Paralleling Snyder's work, we extend the theory to the multichannel case of a continuous, two-dimensional (2-D), energy-time space. This extension essentially involves showing that the data are doubly stochastic Poisson (DSP) point processes in energy as well as time. Further, we present a new 2-D recursive formulation for the radiation-detection problem with large computational savings over nonrecursive techniques when the number of channels is large ($\geqslant 30$). Finally, we present and discuss some adaptive strategies for on-line "learning" of unknown, time-varying signal and background-intensity parameters and statistics. These adaptive procedures apply when a complete statistical description is not available *a priori*.

INTRODUCTION

The spreading use of nuclear materials with the consequent possibility of diversion has increased concern over the problem of detecting radiation sources. [1] This detection* problem is made difficult by both shielding and by the inverse—r^2 loss of power on a sensor of fixed size, which leads to low measured signal power that manifests itself as a low count rate at the sensor. Because of the relatively large background count rate, due to naturally occurring gamma rays, the resulting signal-in-noise detection problem is very difficult.

In addition to the detection problem (i.e., deciding whether the source is present or not), we are often interested in additional information, e.g., source strength, energy spectrum, number of sources, and spatial distribution. In this respect, imaging sensors should be useful. They could sense composite sources made up of simple sources and display their various spatial and intensity relationships. The theory presented below also applies to such estimation problems in that the optimal detector will be given as a functional of the optimal causal estimate and the noisy data.

One method widely used to detect signals in noise is match-filtering the noisy data with a mask that represents the temporal (spatial) structure of the signal. This is the classic linear filter method for a signal in Gaussian noise. However, because of the counting or random-point-process character of the radiation signal and noise, significant further improvement can be gained by using the nonlinear detectors and estimators developed below. The improvement can be especially significant in cases of low count rate, or when a low false alarm rate is required, or both, which would in turn be implied by the necessity of short decision intervals.

Other related problems that can be solved by the theory and methods of random-point processes include restoration of the low-photon-count images provided in flash x-ray imaging, [2] x-ray astronomy, [3] and low-dosages biomedical imaging. [4] (One current method for enhancing low-count-rate images is use of coarse-grained film, but at the cost of lost resolution.) The methods presented below provide an optimal tradeoff of noise level vs resolution.

*The words "detection" and "detector" are used in the sense of estimation/detection theory. Thus, they refer to a decision device and not to a sensor.

Past approaches to solving the radiation-detection problem have mainly been *ad hoc*. Peeling is a common technique used to assess elemental contributions to a measured spectrum. [5,6] A more sophisticated technique is unweighted least squares, which fits a linear combination of known elemental spectra to a measured sample spectrum. Often some type of weighting is required to produce acceptable results; however, the proper weighting for any given problem is not known and must be determined empirically. [6,7] An even more serious drawback for both weighted and unweighted least squares is that they are appropriate only for Gaussian statistics, not for Poisson statistics. The two techniques result in differences that are most pronounced in the low-count-rate case—which is our major concern.

Another example of past approaches addresses the problem of time-varying background intensity. If a constant threshold is chosen for a single-channel system with a varying background count rate, a less-than-optimal detection results when the background intensity is low. A current solution to this problem is to estimate the background intensity from data in neighboring spectral windows and then subtract the estimate from the data on the main channel. [8] This approach has little statistical justification and doubles the photon noise on the main channel; it often leads to negative estimates for the signal intensity—a physical impossibility! In *Adaptive Strategies*, below, we show how to avoid this problem (p. 13).

For these reasons, a theoretical approach to the problem of detecting radiation sources is needed. This work should be based upon work done by Snyder [9,10] and others [11–13] on the theory of random point processes (RPP). Their work has resulted in a theory that is fully suited to the radiation-detection problem. It has been successfully applied in optical communications [14] and in biomedical signal processing. [7]

In this report we undertake four tasks:

- Applying RPP theory to the radiation-detection problem;
- Extending the theory to multichannel energy-time detection by noting that the count data constitute a two-dimensional (2-D) random point field (RPF);
- Deriving a novel 2-D recursive formulation that greatly reduces the computational burden in the case of many data channels; and
- Presenting various adaptive strategies to deal with an unknown time variation in background or signal intensity.

MATHEMATICAL MODELS AND SOLUTIONS

In this section we consider four generic cases of increasing difficulty and discuss the practical application of each. Each of the first three cases considers only a single energy channel. The first is the case of known intensity; the second, randomly varying intensity; and the third, parameterized intensity. The fourth case is the extension of the above concepts to a multichannel environment, where we consider a general, randomly varying intensity. Our analysis is based on the use of Poisson statistics. This model has been shown to be appropriate for radioactive decay by Evans [15] and for optical processes by Kelley and Kleiner. [16]

Single Channel, Known Intensities

We now consider detecting gamma and neutron sources whose output is modeled by the Poisson RPP. No spectral data are available, hence the detection is single-channel. We further assume that time-varying source and background intensities (or expected count rate per unit time are known). The model is expressed mathematically as a binary hypothesis problem in detection theory. Under hypothesis zero, H_0, the observed counts are due to the background alone. Under hypothesis one, H_1, the counts are due to the superposition of a source on the background. Under each hypothesis, the data are assumed to come from Poisson sources with $\lambda_s(t)$ = signal intensity and $\lambda_b(t)$ = background intensity. Thus, we have

$$H_0: N(t) = N_b(t); \lambda(t) = \lambda_b(t)$$

and

$$H_1: N(t) = N_s(t) + N_b(t); \lambda(t) = \lambda_s(t) + \lambda_b(t) ,$$

where $0 \leq t \leq T$. The optimal solution to this detection problem was derived by Reiffen and Sherman[17] in 1963. It consists of calculating the following likelihood ratio statistic:

$$\ell(T) = \int_0^T \ell n[(\lambda_s(t) + \lambda_b(t))/\lambda_b(t)]dN(t) - \int_0^T \lambda_s(t)dt \,. \tag{1}$$

The test consists of comparing this statistic to a threshold determined by the *a priori* probabilities of the individual hypotheses:

$$\ell(T) \underset{H_0}{\overset{H_1}{\gtrless}} \ell_0 \triangleq \ell n \left[\frac{Pr(H_0)}{Pr(H_1)}\right] \,.$$

Figure 1 shows a system diagram. The test can be derived in a simple but rather inelegant way by partitioning the time axis into segments Δt_i and using the Poisson probability law on each of the segments,

$$\exp[\ell(t)] = \lim_{\substack{M \to \infty \\ |\Delta t_i| \to 0}} \times \left\{ \prod_{i=1}^{M} \frac{[\lambda_s(t_i) + \lambda_b(t_i)]^{\Delta N(t_i)}}{[\lambda_b(t_i)]^{\Delta N(t_i)}} \exp[-\lambda_s(t_i)] \right\}$$

or

$$\ell(t) = \lim_{\substack{M \to \infty \\ |\Delta t_i| \to 0}} \times \left\{ \sum_{i=1}^{M} \ell n \left[\frac{\lambda_s(t_i) + \lambda_b(t_i)}{\lambda_b(t_i)}\right] \Delta N(t_i) - \sum_{i=1}^{M} \lambda_s(t_i)\, \Delta t_i \right\} , \tag{2}$$

where we see the defining sums for the integrals in Eq. (1) above. An interesting byproduct of this derivation is that the finite sum given above is the optimal statistic for binning the data in time. Further, we can see that if the weighting function $\ell n[1 + \lambda_s(t)/\lambda_b(t)]$ does not change much over a bin width, the binned data should produce almost optimal results. This can be used to judiciously choose the bin width for a practical system.

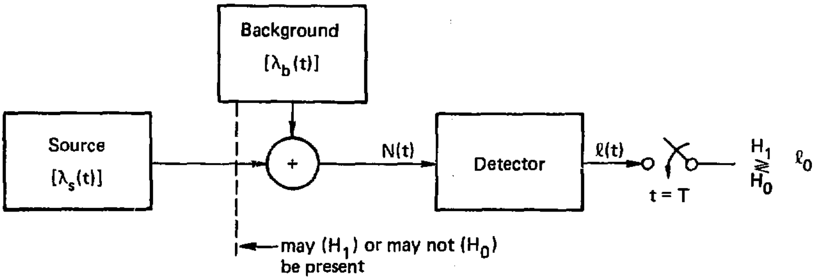


Fig. 1. Optimal detection of time-varying Poisson source.

For an application of this simple single-channel detector, we consider the case of a stationary doorway sensor or a sensor mounted at a fixed distance from a conveyor belt. We further assume that the possible source is of known intensity. We can then precalculate the source and background intensities as a function of time by knowing the object position and velocity, which in turn can be measured by an auxiliary device. Another application would be a moving sensor passing by a possible source in a known trajectory with respect to the source. However, the background intensity would have to be determined in previous passes through the same area. (This restriction is lifted in the random-intensity case discussed below.)

The above detector, Eqs. (1) and (2), is optimal for this radiation-detection problem because of the Poisson nature of the radiation statistics. Aside from the bias term $\int_0^t \lambda_s(t)dt$, the detector statistic is just the weighted sum of the counts in the elemental time bins dt. As such, it is relatively easy to calculate and its implementation compares favorably with the matched-filter detector, which is known to be optimal for the more familiar white Gaussian-noise case. Thus, there is no reason to prefer the linear matched filter to the optimal detector on practical grounds. However, it is well known that in the high-count case, the Poisson distribution becomes approximately Gaussian,[9] with mean and variance as determined by the Poisson distribution,

$$E\{N_s(t)\} = \int_0^t \lambda_s(\tau)d\tau = \mathrm{Var}\{N_s(t)\} .$$

Thus, it might be expected that, when the background count is large, the data should be approximately Gaussian and, hence, that the likelihood ratio, Eq. (1), should have approximately the same performance as the matched-filter detector. Closer examination, however, reveals two related flaws in this argument. First, the likelihood ratio is, by definition, the ratio of two probabilities, each of which is approximately Gaussian, due to the large background-count rate. However, in the case of a weak signal, the log-likelihood ratio using the Gaussian approximation may be significantly in error, because the difference of the logs of the probabilities will emphasize small errors in the approximation. Second, the Gaussian approximation does not hold out on the tails of the distribution.[18] Yet the requirement of a low false-alarm rate moves the evaluation point out to the non-Gaussian tails of the distribution. In any event, the simplicity of Eqs. (1) and (2) shows there is no reason not to implement the optimal detector.

Performance of the optimal detector is unfortunately not analytically tractable; however, bounds are available.[9] (See the Appendix.) Also, the performance in the Gaussian case, for which formulas are available, will set a lower bound. When we consider a Gaussian process into the optimal Gaussian detector, we find it to be the optimal linear detector for any process with the same second-order statistics. Therefore, the performance would he the same for the Poisson process input to this detector, so long as the Poisson process has the same second order statistics. It is clear that the best nonlinear detector for this Poisson process would outperform the best linear detector.

Single Channel, Random Intensities

In the case of known intensities we assumed that the signal and background intensities were completely known. Thus, the detector derived above does not apply in the case where one or more of these intensities is unknown. However, it is a remarkable fact that when the unknown intensities are themselves modeled as random processes, the optimal detector is very similar to the one derived above. That detector need only be modified by the insertion of causal minimum-mean-square error (MMSE) estimates of the unknown intensities in place of the known intensities.[10] On defining $\eta(t) \triangleq [N(\tau), 0 \leq \tau < t]$ and

$$\hat{\lambda}_s^{(1)}(t) \triangleq E\{\lambda_s(t)/\eta(t), H_1\} ,$$

$$\hat{\lambda}_b^{(i)}(t) \triangleq E\{\lambda_b(t)/\eta(t), H_i\} , i = 0 \text{ or } 1 ,$$

we have,

$$\ell(T) = \int_0^T \ln\left[\frac{\hat{\lambda}_s^{(1)}(t) + \hat{\lambda}_b^{(1)} t}{\hat{\lambda}_b^{(0)}(t)}\right] dN(t)$$

$$- \int_0^T \left[\hat{\lambda}_s^{(1)}(t) + \hat{\lambda}_b^{(1)}(t) - \hat{\lambda}_b^{(0)}(t)\right] dt \; . \qquad (3)$$

Actually only two (causal) estimates need be calculated: the background intensity under H_0 and the total (signal plus background) intensity under H_1. (For a system diagram see Fig. 2.) This Eq. (3) representation then reduces the detection problem to one of causal estimation. In the case of finite-order dynamical models, the following equations have been obtained by Snyder.[10] He models the random intensity λ as a function of a vector-diffusion process x(t) given by the Itô equation:

$$d\underline{x}(t) = \underline{\underline{A}}(t)\underline{x}(t)dt + \underline{\underline{B}}(t)d\underline{w}(t) \, , \qquad \underline{x}(0) \qquad (4)$$

given where $\underline{w}(t)$ is the vector Wiener process with covariance

$$E[\underline{w}(t)\underline{w}^T(u)] = \underline{\underline{I}}\,\min(t,u) \, .$$

The approximate filtering equations are a special case of those in Ref. 9. We have

$$d\underline{x}^*(t) = \underline{\underline{A}}(t)\underline{x}^*(t)\,dt + \underline{\underline{\Sigma}}^*(t)\nabla\lambda[\underline{x}^*(t)]\lambda^{-1}[\underline{x}^*(t)][dN(t) - \lambda[\underline{x}^*(t)dt] \, , \qquad (5)$$

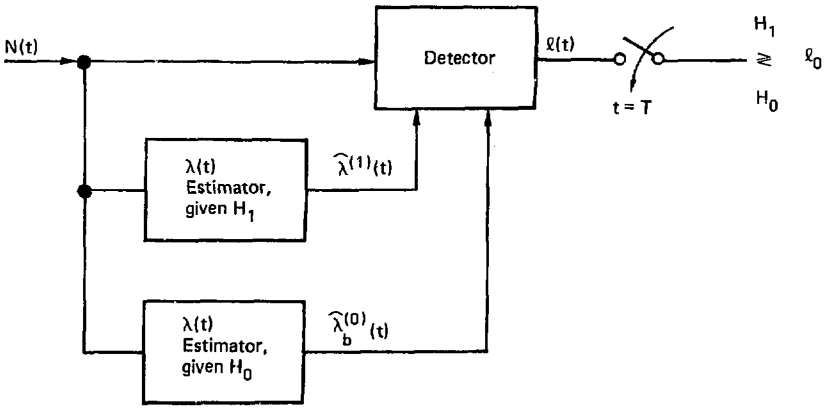


Fig. 2. Optimal estimator/detector for doubly stochastic Poisson (DSP) process.

where $\underline{x}^*(0) = E[\underline{x}(0)]$. The equation for the approximate error covariance $\underline{\underline{\Sigma}}^*(t)$ is

$$d\underline{\underline{\Sigma}}^*(t) = (\underline{\underline{A}}(t)\underline{\underline{\Sigma}}^*(t) + \underline{\underline{\Sigma}}^*(t)\underline{\underline{A}}^T(t))dt + \underline{\underline{B}}(t)\underline{\underline{B}}^T(t)dt$$
$$-\underline{\underline{\Sigma}}^*(t)\,\underline{\nabla}^2\lambda(\underline{x}^*(t))\underline{\underline{\Sigma}}^*(t) + (\underline{\underline{\Sigma}}^*(t)\,\underline{\nabla}^2\ell n\lambda(\underline{x}^*(t))\underline{\underline{\Sigma}}^*(t))dN(t)\,, \qquad (6)$$

with initial condition $\underline{\underline{\Sigma}}^*(0) = \underline{\underline{Cov}}(\underline{x}(0))$. In Eqs. (5) and (6)

$$\underline{\nabla}f(\underline{x}) \triangleq \left[\frac{\partial f}{\partial x_i}\right] \quad \text{and} \quad \underline{\underline{\nabla}}^2 f(\underline{x}) \triangleq \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}\,.$$

The filtering equation, Eq. (5), is a nonlinear equation of the prediction-update variety; it produces the suboptimal estimate $\underline{x}^*(t)$. The covariance Eq. (6) seems similar to the Ricatti equation that arises in Kalman filtering,[19] however, inspection of the last term in Eq. (6) reveals a dependence on the data, which makes $\underline{\underline{\Sigma}}^*(t)$ random, unlike the Kalman case, but like the extended Kalman case.[19]

The approximate Eqs. (5) and (6) were derived as a close approximation to the optimal nonlinear equations in Ref. 9 where a number of examples are presented to show excellent agreement between the approximate and the exact estimates. The approach is analogous to that used in the extended Kalman filter[18] and essentially consists of a Taylor series expansion of the new estimate about the old, to linearize the incremental equation.

Applications for this case abound in the various scenarios for radiation-source detection. The model of this section has added the realistic feature of uncertainty in the source and background intensities. They are characterized by their mean or expected value plus a random deviation determined by a Markov diffusion. The added realism allows the modeling of a moving sensor, e.g., in a van, where the random variation of the background with position is much more than the simple Poisson deviation.[6] Equations (5) and (6) are also modeling a source whose intensity is not exactly known. This can account for variations in sources, shielding and distance from source to sensor and hence model additional uncertainty that could not be treated in the model developed for the known-intensity case.

In comparison with linear detectors, we note that in the present case the distribution does not converge to normal if $\lambda(t) \geqslant 0$ is asymptotically nondeterministic in the sense that

$$\lim_{t\to\infty} \frac{V(T)}{M(T)} > 0$$

where

$$M(T) \triangleq E\left\{\Lambda(T) = \int_0^T \lambda(t)dt\right\} \qquad (7a)$$

and

$$V(T) \triangleq \mathrm{Var}\{\Lambda(T)\}. \qquad (7b)$$

This is verified by a corollary in Grandell.[20] Instead of converging to normal, the counting variable N(T) converges to the weighted sum of a normal random variable and a centered version of (T), which could not be normal because of its nonnegativeness. Thus, in the case of random intensities (also called compound Poisson and doubly stochastic Poisson), the "linear," or Gaussian, detectors would be definitely suboptimal in many ways.

Grandell[21] presents detailed examples of linear and nonlinear estimates, with $\lambda(t)$ a jump process having two levels. He finds the nonlinear estimate can be much more accurate than the best linear estimate, the differences becoming more pronounced as the distance between the two levels increases. No examples are given for the case of smooth $\lambda(t)$.

Single Channel, Parameterized Intensities

This section treats a special case of the random-intensity problem. In particular, we assume the unknown signal is known, but for a finite set of parameters that are to be estimated as part of the proposed optimal detection algorithm. We assume the following problem statement:

Observe $N(t)$ over $[0,T]$ and decide H_0 or H_1, where:

H_0: $N(t)$ is Poisson with intensity $\lambda_b(t)$,

H_1: $N(t)$ is conditionally Poisson with intensity

$\lambda_b(t) + \lambda_s(t,\underline{c})$, where the density $p_{\underline{c}}(\underline{C})$ is known.

For this problem, the optimal detector has the form of Eq. (3) specialized to deterministic $\lambda_b(t)$:

$$\ell(T) = \int_0^T \ell n \left[1 + \frac{\hat{\lambda}_s(t,\underline{c})}{\lambda_b(t)}\right] dN(t) - \int_0^T \hat{\lambda}_s(t,\underline{c})dt \ , \tag{8}$$

where

$$\hat{\lambda}_s(t,c) \triangleq E\{\lambda_s(t,\underline{c}) \,|\, \eta(t), H_1\} \ .$$

To complete the solution, we must solve for $\hat{\lambda}_s$. Following Ref. 9, we make the approximation $\hat{\lambda}_s(t,\underline{c}) \simeq \lambda_s(t,\hat{\underline{c}}(t))$, where $\hat{\underline{c}}(t) = E\{\underline{c} \,|\, \eta(t), H_1\}$. The resulting approximate equations are

$$d\hat{\underline{c}}(t) = -R(t)\underline{\nabla}\lambda_s(t,\hat{\underline{c}}(t))dt + \underline{R}(t)\underline{\nabla}\ell n\lambda_s(t,\hat{\underline{c}}(t))dN(t) \ , \tag{9}$$

with initial condition

$$\hat{c}(0) = E_0(c) \triangleq \int \underline{C}p_c(\underline{C})dC = \bar{\underline{c}} \ .$$

The error covariance $\underline{\underline{R}}(t)$ evolves according to

$$d\underline{\underline{R}}(t) = -\underline{\underline{R}}(t)\underline{\underline{J}}_1(t,\hat{\underline{c}}(t))\underline{\underline{R}}(t)dt + \underline{\underline{R}}(t)\underline{\underline{J}}_2(t,\hat{c}(t))\underline{\underline{R}}(t)dN(t) \ , \tag{10}$$

where

$$\underline{\underline{R}}(0) = E_0[(\underline{c} - \bar{\underline{c}})(\underline{c} - \bar{\underline{c}})^T] \ ,$$

and

$$(\underline{J}_1)_{ij} \triangleq \left. \frac{\partial^2 \lambda_s(t,\underline{c})}{\partial c_i \partial c_j} \right|_{\underline{c} = \hat{\underline{c}}(t)} , \qquad (\underline{J}_2)_{ij} \triangleq \left. \frac{\partial^2 \ln \lambda_s(t,\underline{c})}{\partial c_i \partial c_j} \right|_{\underline{c} = \hat{\underline{c}}(t)}$$

These equations are seen to be a special case of Eqs. (5) and (6) corresponding to a dynamical model with $d\underline{x}(t) \equiv 0$.

To illustrate with an example, let $c = c$, a scalar, and let $\lambda_s(t,c) = c\lambda_s(t)$. This would correspond to a detection problem where the signal intensity is known except for its overall level, a situation that often occurs in practice. In this case, Eqs. (9) and (10) become,

$$d\hat{c}(t) = -R(t)\lambda_s(t)dt + (R(t)/\hat{c}(t))dN(t) , \qquad \hat{c}(0) = \text{given}, \tag{11a}$$

and

$$dR(t) = -(R(t)/\hat{c}(t))^2 dN(t) . \qquad R(0) = \text{given}. \tag{11b}$$

These equations show that R(t), the error variance for the estimate of th unknown parameter c, decreases monotonically as more data are processed. Equation (11b) also shows that the decrease tends to zero as R(t) tends to zero, thus keeping R(t) nonnegative. For nondegenerate cases, $R(t) \searrow 0$ as $t \nearrow \infty$, indicating that $\hat{c}(t)$ will tend to its exact value. Basically, this requires that $\lambda_s(t)$ not go to zero for large t.

Equations (11) could easily be simulated to gain added understanding of the optimal solution to this problem. In this regard, it is interesting to note that Eq. (8) is recursive in T. As T increases to $T + \Delta T$, only the increment $\ell(T + \Delta T) - \ell(T)$ need be calculated, the integrals over (0,T) need not be repeated. This happens because the estimates in Eq. (8) [and also Eq. (3)] are causal and, hence, do not require future data for their evaluation.

For very long observation intervals, a theorem of Grandell [20] shows that N(t) tends to a Gaussian distribution. This corresponds to the case,

$$\lim_{T \to \infty} \frac{V(T)}{M(T)} = 0 \quad [\text{see Eq. (7)}]$$

and occurs because $\lambda_s(t)$ is deterministic given a finite set of parameters whose values are eventually learned as $T \nearrow \infty$. As a result of this observation, the linear estimator could be used for this case. However, as noted above, this would not be accurate for the case of small false alarm probability.

Multiple Channels, Random Intensities

In this section we examine the general, multichannel DSP process, with initial attention focused on the case of known intensities with no time variation considered.

Known Intensity, No Time Variation

If we argue that the resulting process is Poisson in the energy variable e, instead of the time variable t, it follows that the techniques for the three cases already presented apply to this case. Secondly, we add time variation to the energy variation to yield a 2-D DSP field, as studied by Fishman and Snyder [22] in a general context for space-time fields. Finally in this section, we develop a 2-D recursive formulation for energy-time estimation or detection.

To treat the multichannel case, we must first determine the probability distributions. To focus on the energy variable, we specialize to the case of no time variation, so that the data consist of the 1-D record of counts vs energy. (See Fig. 3.) We idealize the channel width to a differential de and consider a continuum of energies. Thus the data become

$$\eta(E) \triangleq [N(e), 0 \leq e < E] ,$$

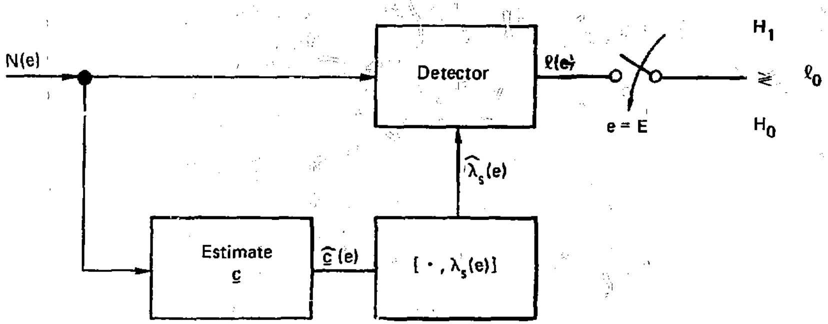


Fig. 3. Optimal estimator/detector for multichannel, time-invariant DSP process with parameterized signal intensity.

where $N(e) \stackrel{\Delta}{=}$ number of photons received with energy less than or equal to e. We can then obtain the practical cases of finite energy bins of various widths by quantizing the continuous case. We show next that the random process (N(e)) is, to a good approximation, a Poisson RPP in the parameter e. To do this, we first recall the set of necessary and sufficient conditions for an RPP to be Poisson.

By definition, a Poisson process is an RPP $\{N(e), 0 \leq e \leq E\}$ with the following three properties:

(1) $\Pr[N(0) = 0] = 1$

(2) For all $e_0 < e_1$, the increment forms $N(e_0, e_1) = N(e_1) - N(e_0)$. (It is Poisson distributed with parameter $\lambda(e_1) - \lambda(e_0)$, where $\lambda(e) \geq 0$ and non-decreasing on [0,E].)

(3) $[N(e), 0 \leq e \leq E]$ has independent increments.

Since N(e) is a counting process in energy, condition (1) is trivially satisfied. Condition (2) is implied by the counting observations of Evans [15] for radioactive decay and Kelley and Kleiner [16] for photon counting in various energy ranges. Condition (3) is less evident. It states that the total count in disjoint energy bands must be independent. This is a common assumption and appears to be true in practice. This point should probably be subjected to experimental hypothesis testing. Accepting condition (3), we thus arrive at the conclusion that N(e) is a Poisson RPP on the energy continuum [0,E].

We can now apply the results derived for the three earlier cases to the energy variable to derive optimal detectors and estimators for the multichannel, time-independent case. The key point in an application of these results is that the (average) intensity of the radiation λ(e) is not a function of time. As a particular example, we take the results of the third case—for parameterized intensity—and convert to the multichannel, time-invariant case, to obtain the detector,

$$\ell(E) = \int_0^E \ln\left[1 + \frac{\hat{\lambda}_s(e,\underline{c})}{\lambda_b(e)}\right] dN(e) - \int_0^E \hat{\lambda}_s(e,\underline{c})de \ , \tag{12}$$

and the estimator,

$$d\underline{\hat{c}}(e) = -\underline{\underline{R}}(e)\,\underline{\nabla}\lambda_s(e,\underline{\hat{c}}(e))de + \underline{\underline{R}}(e)\,\underline{\nabla}\ln s(e,\underline{\hat{c}}(e))dN(e)\ , \tag{13}$$

$$dR(e) = -\underline{\underline{R}}(e)\underline{\underline{J}}_1(e,\underline{\hat{c}}(e))\underline{\underline{R}}(e)de + \underline{\underline{R}}(e)\underline{\underline{J}}_2(t,\underline{\hat{c}}(e))\underline{\underline{R}}(e)dN(e)\ . \tag{14}$$

As an application, we consider the case of detecting a signal composed of varying amounts (c_i) of known elements ($\lambda_{si}(e)$) in a background flux ($\lambda_b(e)$):

H_0: N(e) is DSP with $\lambda_b(e)$

H_1: N(e) is DSP with $\lambda_b(e) + \sum_{i=1}^{M} c_i\lambda_{si}(e)$

where $p_{\underline{c}}(\underline{C})$ is a given *a priori* density. In this very practical case, Eqs. (13) and (14) simplify to:

$$d\hat{\underline{c}}(e) = -\underline{\underline{R}}(e)\underline{\lambda}_s(e)de + \underline{\underline{R}}(e)[\underline{\lambda}_s(e)/\underline{\lambda}_s(e,\hat{\underline{c}}(e))]\, dN(e)$$

or

$$d\hat{\underline{c}}(e) = \underline{\underline{R}}\,(e)\lambda_s(e)\left[\frac{dN(e)}{\lambda_s(e,\hat{\underline{c}}(e))} - de\right], \quad (15)$$

and

$$d\underline{\underline{R}}(e) = -[\underline{\underline{R}}(e)\underline{\lambda}_s(e)\underline{\lambda}_s^T(e)\underline{\underline{R}}(e)]\, dN(e)\,, \quad (16)$$

where

$$\underline{\lambda}_s^T(e) \triangleq [\lambda_{s1}(e), \lambda_{s2}(e), \ldots, \lambda_{sM}(e)]\,.$$

For a system diagram, see Fig. 3.

These pairs of equations, (13) and (14), and (15) and (16), provide recursive solutions for multichannel detection on an energy continuum. We expect them to be useful for the situation where a large number of channels are available, in that a recursive processor can be more efficient than a nonrecursive multichannel processor. If we partition the energy space into a finite number of bins, the optimal detection statistic is simply the sum of the log-likelihood ratios for the individual bins, because of the independent increment property of N(e). (See example 3.3.1 on multichannel analysis in Ref. 9.) For a small number of channels, this nonrecursive (in energy) solution would be computationally preferable; however, for a fairly dense set of channels, the numerical solution of Eqs. (15) and (16) would be computationally preferable.

Energy and Time Variation

We now turn to the 2-D problem that arises when we have both energy and time variation in an intensity $\lambda(e,t)$. We first present the nonrecursive formulation for multidimensional RPF's developed by Fishman and Snyder,[22] but specialized to the case of the 2-D DSP field. Then we present a new 2-D recursive procedure for the estimator that is based on a similar derivation of the reduced update Kalman filter[23] for linear 2-D signal estimation.

Vector-Recursive Formulation. The results in Ref. 22 cover a more general case than is needed here. Although the results are presented for space-time RPF's, they, of course, hold for any multidimensional parameter (e.g., $\underline{e},\underline{t}$). The resulting estimator structure is causal in $\underline{t}$ only; thus, the observed data at time t consist of the set,

$$\eta(t) \triangleq \{N(e,\tau);\ 0 \leq \tau < t,\ 0 \leq e \leq E\}\,. \text{ (See Fig. 4).}$$

The formulas presented by Fishman and Snyder are very similar to those of the 1-D case. For the problem

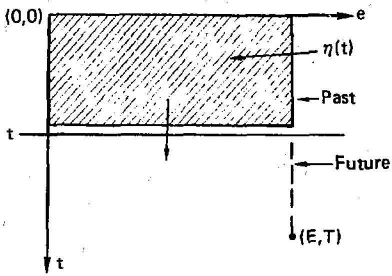


Fig. 4. Diagram on the energy-time plane showing casual structure of vector recursive formulation.

H_0: $N(e,t)$ is Poisson with $\lambda_b'(e,t)$

H_1: $N(e,t)$ is Poisson with $\lambda_s(e,t) + \lambda_b(e,t)$,

we obtain the log-likelihood ratio,

$$\ell(T) = \int_0^T \int_0^E \ell n \left[1 + \frac{\lambda_s(e,t)}{\lambda_b(e,t)}\right] dN(e,t) - \int_0^T \int_0^E \lambda_s(e,t)\, de\, dt \ . \tag{17}$$

For the case of $\lambda_s(e,t)$ random, we obtain Eq. (17), but with $\lambda_s(e,t)$ replaced by its causal (in t) estimate,

$$\hat{\lambda}_s(t) = E\{\lambda_s(e,t)/\eta(t), H_1\} = \hat{E}_t\{\lambda_s(e,t)\} \ .$$

For $\lambda_s(e,t)$ of the form $\lambda_s(e,t;\underline{x}(e))$ where $\underline{x}(e)$ is a Markov diffusion in energy, the following recursive representation is derived in Ref. 22:

$$d\hat{E}_t(\underline{x}(e)) = \int_0^E [\hat{E}_t(\underline{x}(e)\lambda_s) - \underline{x}(e)\lambda_s][dN(e,t) - \lambda_s^{-1}(e,t)\, de\, dt] \ . \tag{18}$$

We can obtain an approximate equation for the Eq. (18) estimate by using a procedure analogous to that used for Eq. (5), that is, expand in a Taylor series about the conditional mean.[21]

Computationally, Eq. (18) is seen to be recursive in $\underline{t}$ and nonrecursive in $\underline{e}$. Thus, if the dynamics are assumed to be much slower than the point arrivals, a numerical approximation to Eq. (18) would involve a weighted sum over the range [0,E] for the estimate $\hat{E}_t(\underline{x}(e))$ at each energy value. The computations might be excessive for a large number of energy channels. In the next section, we develop an estimator that is recursive in both e and t and, thus, is more efficient.

2-D Recursive Formulation. The simplest way to develop a 2-D recursive formulation for the energy-time detection is to replace the (e,t) continuum by a scanned representation (Fig. 5). If the scan lines are sufficiently dense (close together), the loss in detection efficiency due to this replacement can be made negligible. The advantage is that the detection problem for the scanned data is a 1-D detection problem and is solved above. We note that this 1-D process is Poisson, hence, we can immediately give the log-likelihood ratio. In the limit, as the scan line spacing goes to zero, we obtain formally

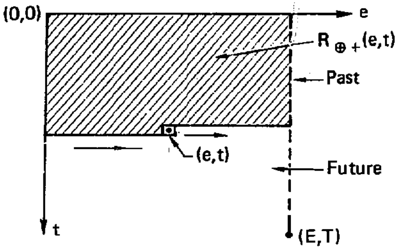


Fig. 5. Diagram on the energy-time plane showing casual structure of 2-D recursive formulation.

$$\ell(E,T) = \int_{R_{\oplus+}(E,T)} \ln \left[1 + \frac{\lambda_s(e,t)}{\lambda_b(e,t)}\right] dN(e,t) - \int_{R_{\oplus+}(E,T)} \lambda_s(e,t)\, de\, dt$$

where

$$\mathscr{R}_{\oplus+}(E,T) \triangleq \{e \leq E, t \leq t\} \cup \{e > E, t < T\} .$$

To consider a case of random variation in λ_s, as in *Single Channel, Random Intensities*, we first develop a diffusion equation for the random field $\underline{x}(e,t)$, specifying $\lambda_s(\underline{x}(e,t))$. Formally generalizing the 1-D case, in the limit as the scan line spacing decreases to zero, we obtain

$$d_e\underline{x}(e,t) = \underline{\underline{A}}(e,t)\underline{x}(e,t)\, de + \underline{\underline{B}}(e,t) d_e\underline{w}(e,t) \tag{19}$$

with boundary conditions $\underline{x}(0,t)$ and $\underline{x}(e,0)$ given. We note in Eq. (19) that d_e indicates an increment due to a differential increment in e. The equation for the d_t increment is similar, but is not needed for a recursion, as shown in Fig. 5. Using Eq. (19) to model the random field $\underline{x}(e,t)$, we obtain the following log-likelihood ratio formulation:

$$\ell(E,T) = \int_{R_{\oplus+}(E,T)} \ln \left[1 + \frac{\hat{\lambda}_s(e,t)}{\lambda_b(e,t)}\right] dN(e,t) - \int_{R_{\oplus+}(E,T)} \hat{\lambda}_s(e,t)\, de\, dt \ , \tag{20}$$

where

$$\hat{\lambda}_s(e,t) \triangleq E[\lambda_s(\underline{x}(e,t))/\eta(e,t), H_1]$$

where

$$\eta(e,t) \triangleq \{N \text{ on } \mathscr{R}_{\oplus+}(e,t)\} .$$

As before, we replace $\hat{\lambda}_s(e,t)$ by $\lambda_s[\underline{x}^*(e,t)]$, where $\underline{x}^*(e,t)$ is the approximate MMSE recursive estimate obtained from the linearized equation

$$d_e\underline{x}^*(e,t) = \underline{\underline{A}}(e,t)\underline{x}^*(e,t)de + \underline{\underline{\Sigma}}^*(e,t)\, \underline{\nabla}\lambda_s[\underline{x}^*(e,t)]\lambda_s^{-1}[x^*(e,t)]$$

$$[d_eN(e,t) - \lambda_s(\underline{x}^*(e,t))de] . \tag{21}$$

The error covariance $\underline{\underline{\Sigma}}^*$ is the solution to the 2-D version of Eq. (6), in the raster-scanning sense. The form of Eq. (21) allows us to obtain the estimate in a recursive manner. The form of Eq. (20) allows us to compute the detection statistic recursively along with the estimate as new values $\hat{\lambda}(e,t)$ become available. The discrete approximation to Eq. (19) would take the form of an NSHP recursive filter,[25] with state vector as shown in Fig. 6. The computations can be greatly reduced and made equivalent to having the state vector slightly larger than the model support, shown by the dotted lines in Fig. 6. This would be accomplished analogously to the reduced update Kalman filter[23] developed for Gaussian noise statistics. This is possible because of the Kalman-like prediction-plus-update form of Eq. (21). Even more efficient formulations are possible in the case of the parametrically unknown model discussed above. Finally, generalizations containing both random parameters and random fields in both the background and the signal are clearly possible. This would have particular application to the moving (e.g., in a van) multichannel detector.

If we let the two coordinates (e,t) become spatial coordinates (x,y), the above equations give a 2-D recursive solution for the estimator and detector on DSP random pictures such as would arise, for example, in low-photon images. In this case, because of the large number of points in each dimension (>500), the recursive formulation would be particularly efficient in comparison with nonrecursive (in space) formulations.[22]

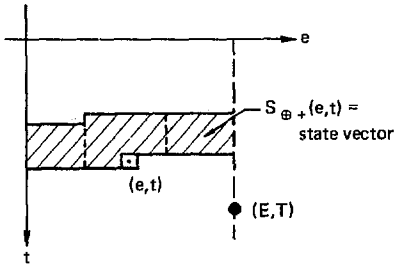


Fig. 6. Diagram on the energy-time plane showing the state vector for the 2-D recursive formulation.

ADAPTIVE STRATEGIES

In this section we introduce the concept of adaptive processing for the problems discussed above. We describe a processing method or solution as "adaptive" if some parameters in the solution are *unknown* and are *learned* in the course of the processing. The word "learned" is synonymous with the word "estimated," but it has additional implication of low estimation error. Accordingly, we assume that a complete *a priori* description of the statistics of the signal and noise is not available. Then we estimate the unknown parameters and use these learned parameters to make the signal estimate. These interpretations agree with the literature on adaptive systems.

A simple example that arises in radiation detection concerns the single-channel detector in a moving vehicle. As the detector moves, the background noise intensity $\lambda_b(t)$ varies in an unknown manner. Thus, the simple optimal test of the kind presented in the first case considered above cannot be performed since it uses $\lambda_b(t)$ as a "parameter." An adaptive solution to this problem would first estimate $\lambda_b(t)$ online, then use this estimate $\tilde{\lambda}_b(t)$ in place of $\lambda_b(t)$ in the detection equation corresponding to the known case in Eq. (1). While such a solution is superficially very close to the optimal solution presented for the random-intensity, single-channel case; a significant difference exists. In the earlier instance we assumed a complete statistical description of the unknown background (and signal) intensity so that it was possible to formulate an optimal detector. The remarkable conclusion of the random-intensity case is that the optimal detector, Eq. (3), is the same as the detector in the known intensity case, Eq. (1), but with the optimal causal estimates (under H_i) substituted for the unknown, random intensities. However, in this section, we assume that a detailed statistical description is now available, so that such optimal causal estimates are not possible. Rather, it is possible only to make various *ad hoc* estimates.

We can conclude two important points from the above discussion: First, the adaptive estimator/detector is not necessarily better and in fact may be much worse than the detectors presented in the preceding section. Second, for a case where either $\lambda(t)$ cannot be easily modeled to the required accuracy *or* a case where the resulting optimal detector is too complicated to implement, the adaptive method offers a useful alternative.

To return to the example, if $\lambda_b(t)$ is unknown, we can form a simple finite-memory, running average estimate over, say, τ seconds:

$$\tilde{\lambda}_b(t) = \frac{N(t) - N(t-\tau)}{\tau} \quad . \tag{22}$$

This estimate could then be used in Eq. (1) to get an adaptive detection statistic,

$$\ell(T) = \int_0^T \ell n \left[1 + \frac{\lambda_s(t)}{\tilde{\lambda}_b(t)}\right] dN(t) - \int_0^T \lambda_s(t)dt \quad .$$

An immediate problem with this strategy is that Eq. (22) yields a good estimate for $\lambda_b(t)$ only if τ is judiciously chosen as a good compromise between tracking the "dynamical" variations in $\lambda_b(t)$ and reducing the statistical variation of the right side of Eq. (22). However, it is even more important to note that Eq. (22) is useful as an estimate of $\lambda_b(t)$ only if the signal is not present in the interval $(t-\tau,t]$. This problem has been studied in estimation theory as "decision directed estimation."[23] This approach feeds back the detector output (i.e., H_0 or H_1) to control the estimate Eq. (22). For example, a simple strategy is to update $\tilde{\lambda}_b(t)$ only if no detections occur in $(t-\tau,t]$.

In a more sophisticated variant of the above procedure, we assume that $\lambda_b(t)$ is governed by a Markov diffusion [Eq. (4)], where one or more parameters in the system matrices $\underline{A}(t)$ and $\underline{B}(t)$ are unknown. In such a situation, these parameters could be estimated using "distribution-free" methods such as Eq. (22), followed by insertion of these "learned" parameters into the estimate equations (5) and (6). Such a procedure would have the advantage, with respect to the simpler method mentioned in the previous paragraph, of allowing a continuation of the dynamical trajectory in $\lambda_b(t)$ during intervals of signal detection, when no updating is permitted.

Adaptive techniques can also be applied in the multichannel detection problems discussed above. A fully adaptive implementation would allow the multichannel intensity models for both background and signal to acquire their unknown parameters online, using decision-directed estimation. Concurrently, causal estimates are made of the intensities. Also concurrently, these estimates are used in the recursive calculation of the detection statistic (E,T) (see Fig. 7).

As noted in the Introduction; a current adaptive solution for multichannel detection estimates the background intensity from data in neighboring windows. Then this estimate is subtracted from the total intensity in the signal window to estimate the signal intensity. Finally, the signal-intensity estimate thus obtained is used as the detection statistic. A very serious defect in this procedure often gives the estimated signal intensity a negative value. However, our study of optimal detection methods indicates that several improvements can be made to this current procedure. Beginning with the energy version of Eq. (3), we see that it is not necessary to estimate the signal intensity; rather it is necessary only to estimate the total received intensity under each hypothesis. Thus the subtraction that leads to negative values should not be performed. Then, the proper sufficient statistic is not $\hat{\lambda}_s(e)$, but rather the log-likelihood ratio given by Eq. (3). Thus, the following procedure is recommended as a simple improvement.

1. Use the data in neighboring windows to estimate $\lambda_b(e)$ in the signal window. Call it $\tilde{\lambda}_b^{(0)}(e)$.
2. Estimate the total intensity, $\lambda(e) = \lambda_b(e) + \lambda_s(e)$ in the signal window by smoothing N(e). Call the estimate $\tilde{\lambda}^{(1)}(e)$.
3. Form the statistic.

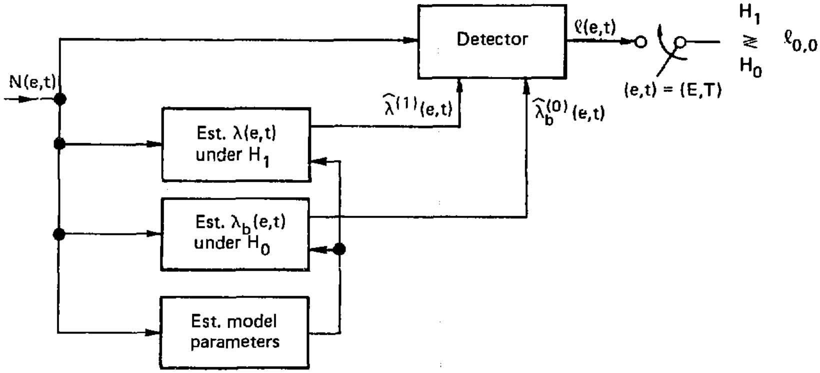


Fig. 7. Adaptive approximation to 2-D version of optimal detector in *Single Channel, Random Intensities.*

$$\ell(e) = \int_{e-\Delta}^{e+\Delta} \ell n \left[\frac{\tilde{\lambda}^{(1)}(\nu)}{\tilde{\lambda}_b(\nu)}\right] dN(\nu) - \int_{e-\Delta}^{e+\Delta} \left[\tilde{\lambda}^{(1)}(\nu) - \tilde{\lambda}_b^{(0)}(\nu)\right] d\nu .$$

If the signal window is very narrow, this reduces to

$$\ell(e) = \ell n[\tilde{\lambda}^{(1)}(e)/\tilde{\lambda}_b^{(0)}(e)]\, n(e) - 2\Delta[\tilde{\lambda}^{(1)}(e) - \hat{\lambda}_b^{(0)}(e)] \tag{23}$$

where $n(e) \triangleq N(e + \Delta) - N(e - \Delta)$ = number of counts in the signal window $(e - \Delta, e + \Delta]$. Equation (23) is no more difficult than the current procedure, but should give results much closer to optimal.

CONCLUSION

In this report we review the recently developed mathematical theory of random-point processes and show how it can obtain optimal solutions to the problem of detecting radiation sources. We also show how the theory can be applied to image processing for both imaging detectors and other low-photon imagers. These applications rest on the experimentally verified conclusion that statistical counting noise can be modeled by the Poisson random process and the doubly stochastic Poisson random process. We extend the theory extended to the multichannel energy-time continuum. Also we introduce new 2-D recursive formulation for energy-time detection with large computational savings for the case of a large number of channels ($\geqslant$30). We discuss adaptive strategies to extend the theory to the practical case of unknown parameters that must be learned on-line. Finally, we introduce a theoretically motivated improvement to a currently used adaptive strategy.

ACKNOWLEDGMENT

The computer programming and graphical output discussed in the Appendix result from the work of K. A. Dines of the Information Processing Section.

REFERENCES AND NOTES

1. Nuclear Science Symposium Issue, *IEEE Trans. Nuc. Sci.* (Jan. 1977).
2. See any treatise on x-ray imaging devices.
3. M. P. Ekstrom and J. W. Woods, "Application of Digital Image Restoration in X-Ray Astronomy," *EASCON Convention Record*, Wash., D.C. (Oct. 7-9, 1974).
4. M. M. Ter-Pogassian, *The Physical Aspects of Diagnostic Radiology* (Harper and Row, New York, 1967).
5. W. Simon, *Mathematical Techniques for Physiology and Medicine* (Academic Press, New York, 1972).
6. R. Gunnink and J. B. Niday, *Computerized Qualitative Analysis By Gamma-Ray Spectrometry*, Lawrence Livermore Laboratory, Livermore, CA, UCRL-51061, Vol. 1 (1972).
7. J. Markham *et al.*, "A Numerical Implementation of the Maximum-Likelihood Method of Parameter Estimation for Tracer-Kinetic Data," *Math. Biosci.* **28**, 275 (1976).
8. *Detector Workshop Notes*, EG&G Inc., Santa Barbara, CA (January 1976).
9. D. L. Snyder, *Random Point Processes* (John Wiley & Sons, New York, 1975).
10. D. L. Snyder, "Filtering & Detection for Doubly Stochastic Poisson Processes," *IEEE Trans. Inform. Theory* **IT-18**, 91 (Jan. 1972).
11. I. Rubin, "Regular Point Processes and Their Detection," *IEEE Trans. Inform. Theory* **IT-18**, 547 (Sept. 1972).
12. A. Segal, "Recursive Estimation from Discrete-Time Point Processes," *IEEE Trans. Inform. Theory* **IT-22**, 422 (July 1976).
13. J. H. VanSchuppen, "Filtering, Prediction, and Smoothing for Counting Process Observations, a Martingale Approach," *SIAM J. Appl. Math.* **32** (May 1977).
14. E. V. Hoversten *et al.*, "Direct-Detection Optical Communication Receivers," *IEEE Trans. Comm.* **COM-22**, 17 (Jan. 1974).
15. R. D. Evans, *The Atomic Nucleus* (McGraw-Hill, New York, 1965).
16. P. L. Kelley and W. H. Kleiner, "Theory of Electromagnetic Field Measurement and Photoelectron Counting," *Phys. Rev.* **136**, A316 (Oct. 1964).
17. B. Reiffen and H. Sherman, "An Optimum Demodulator for Poisson Processes: Photon Source Detectors," *Proc. IEEE* **51**, 1316 (Oct. 1963).
18. B. V. Gnedenko, *Theory of Probability* (Chelsea, New York, 1963).
19. A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications & Control* (McGraw-Hill, New York, 1971).
20. J. Grandell, "On Stochastic Processes Generated by a Stochastic Intensity Function," *Skand. Aktuar Tidskr.*, 204 (1971).
21. J. Grandell, *Doubly Stochastic Poisson Processes*, Vol. 529, in *Lecture Notes in Mathematics*, Springer-Verlag (1976).
22. P. M. Fishman and D. L. Snyder, "The Statistical Analysis of Space-Time Point Processes," *IEEE Trans. Inform. Theory* **IT-22**, 257 (May 1976).
23. J. W. Woods and C. H. Radewan, "Kalman Filtering in Two-Dimensions," *IEEE Trans. Inform. Theory* **IT-23**, 473 (July 1977).
24. D. B. Cooper, "On Some Convergence Properties of 'Learning with a Probabilistic Teacher' Algorithms," *IEEE Trans. Inform. Theory* **IT-21**, 699 (Nov. 1975).
25. J. W. Woods and M. P. Ekstrom, "Nonsymmetric Half-Plane Recursive Filters—Characterization, Stability Theorem and Test," *Proc. 1975 Int'l. Sympos. on Circuits & Systems*, Newton, Mass., 447 (April 1975).

JRS

APPENDIX

In this appendix we consider a simple problem consisting of a vector observation of two Poisson RPP's. We obtain the detection probabilities using the Gaussian approximation and also bound the probabilities using a Chernoff bound.

The problem can be simple stated as follows. For a T-second interval we observe $\underline{N}(t) = (N_1(t), N_2(t))^T$ and must decide between two alternatives.

$$H_1{:}N(t) \text{ is Poisson with rate } \underline{\lambda} = (\lambda_{01} + \lambda_{11}, \lambda_{02} + \lambda_{12})^T$$

$$H_0{:}\ N(t) \text{ is Poisson with rate } \underline{\lambda} = (\lambda_{01}, \lambda_{02})^T .$$

The two channels are assumed mutually independent. The values for the λ_{ij} are given constants, with channel 1 being stronger. Three strategies are considered:

- Using only the stronger channel.
- Adding the two channels.
- Computing the optimal log-likelihood ratio.

For strategy 3, due to the mutual independence of the channels and the constant rates, we obtain from Eq. (1) of the main text,

$$\ell(T) = a_1 N_1(T) + a_2 N_2(T) \underset{H_0}{\overset{H_1}{\gtrless}} \ell_0 \qquad (A1)$$

where

$$a_i \triangleq \ln\left(1 + \frac{\lambda_{1i}}{\lambda_{0i}}\right) \quad i = 1,2 \ .$$

The likelihood ratio for strategies 1 and 2 are just special cases of Eq. (A1).

To use the Gaussian approximation, via the Central Limit Theorem, [17] for calculation of the detection probabilities, we note: Under H_0,

$$E[\ell(T)] = \lambda_{01}a_1 + \lambda_{02}a_2 \triangleq m_0$$

and

$$\mathrm{Var}[\ell(T)] = \lambda_{01}a_1^2 + \lambda_{02}a_2^2 \triangleq \sigma_0^2 .$$

Similarly for H_1,

$$E[\ell(T)] = (\lambda_{01} + \lambda_{11})\, a_1 + (\lambda_{02} + \lambda_{12})\, a_2 \triangleq m_1$$

and

$$\mathrm{Var}[\ell(T)] = (\lambda_{01} + \lambda_{11})\, a_1^2 + (\lambda_{02} + \lambda_{12})\, a_2^2 = \sigma_1^2 .$$

Thus we may obtain

$$P_F = \Pr\left\{\ell(T) \geq \ell_o/H_o\right\} \simeq Q\left(\frac{\ell_o - m_o}{\sigma_o}\right) \tag{A2}$$

and

$$P_D = \Pr\left\{\ell(T) \geq \ell_o/H_1\right\} \simeq Q\left(\frac{\ell - m_1}{\sigma_1}\right) \tag{A3}$$

where

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-1/2\, x^2}\, dx.$$

These equations are accurate near the means.

To develop the Chernoff bounds on these probabilities (following Ref. 7), we find the moment-generating functions for $\ell(T)$ under H_1 and H_0, respectively:

$$H_0:\ M_{\ell(T)}^{(0)}(s) = \exp\left[\sum_{i=1}^{2} \lambda_{oi}(e^{sa_i} - 1)\right]$$

$$H_1:\ M_{\ell(T)}^{(1)}(s) = \exp\left[\sum_{i=1}^{2} (\lambda_{oi} + \lambda_{1i})(e^{sa_i} - 1)\right]$$

In terms of these functions, we obtain the following bounds on P_F and P_D:

$$P_F(\ell_o) \leq \min_{s \geq 0} M_{\ell(T)}^{(0)}(s)\, e^{-s\ell_o}$$

$$P_D(\ell_o) \geq 1 - \min_{s \leq 0} M_{\ell(T)}^{(1)}(s)\, e^{-s\ell_o}$$

In Figs. A1 and A2, we see that the Chernoff bound to the Gaussian approximation are fairly far apart. This is because the Chernoff bound is accurate in the sense of correct exponential dependence on the threshold. The σ_i values are about 0.6, thus, the range of the threshold on the plot extends out to three standard deviations from the mean. The detection performance is poor because the standard deviation is equal to twice the difference in the means under the two hypotheses. The bounds to Gaussian approximations are plotted in Figs. A3-A5 for various combinations of the parameters:

$$\lambda_{01} = 800\,, \lambda_{02} = 1000$$

$$0 \leq \lambda_{11} \leq 100\ ,\quad \lambda_{12}/\lambda_{11} = \frac{1}{2}, \frac{3}{5}, \frac{3}{4}\,.$$

Figures A3, A4, and A5 show plots of P_D vs the signal intensity on the strongest channel. They are plotted using the Gaussian approximation, which can be expected to be accurate near the mean. The threshold is set such that P_F is 10^{-2}. Since T = 1 s, the false-alarm rate would be 10^{-2}/second.

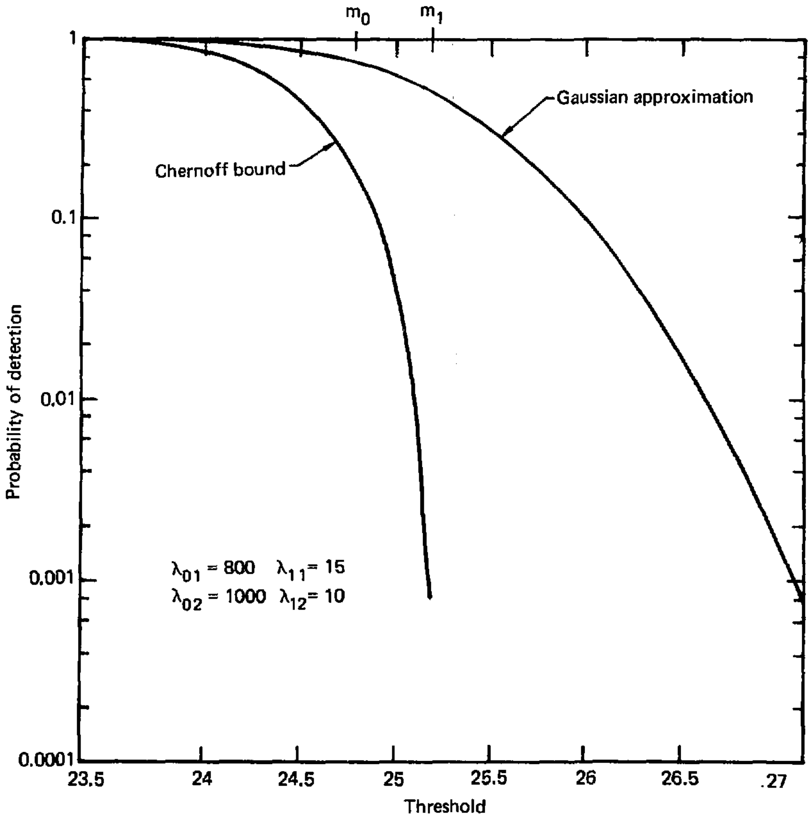


Fig. A1. Performance bounds on P_D for Poisson detection: $\lambda_{0\ell} + k\lambda_{1\ell}$ = total received signal rate in channel ℓ under Hypothesis k; m_k = mean of process under Hypothesis k (k = 0,1).

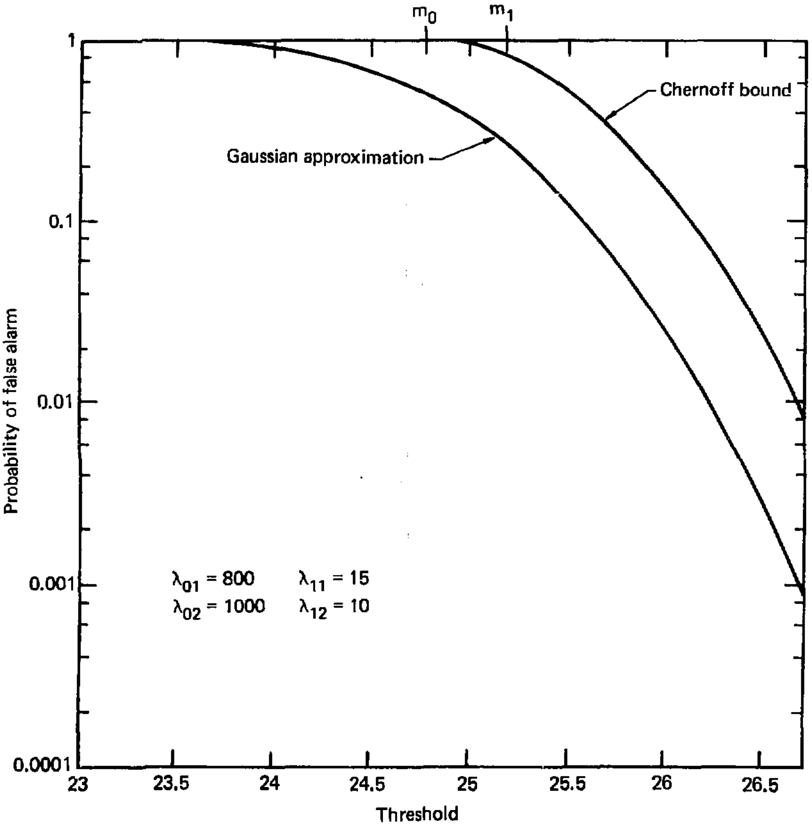


Fig. A2. Performance bounds on P_F for Poisson detection: $\lambda_{o\ell} + k\lambda_1$ = total received signal rate in channel ℓ under Hypothesis k; m_k = mean of process under Hypothesis k (k = 0,1).

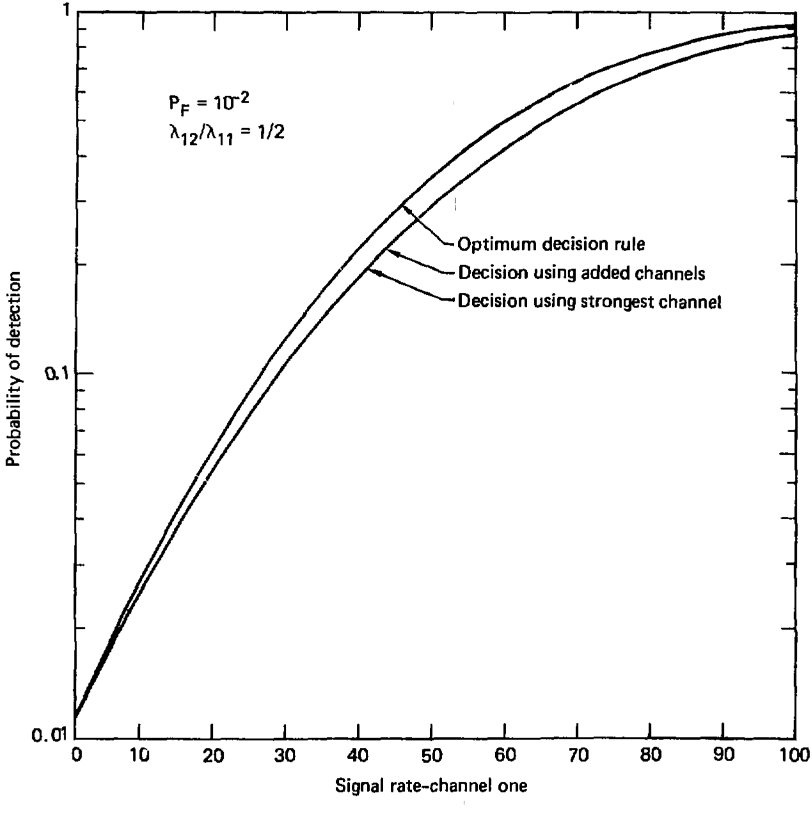


Fig. A3. P_D vs λ_{11} for three different decision rules ($\lambda_{12}/\lambda_{11} = 3/4$).

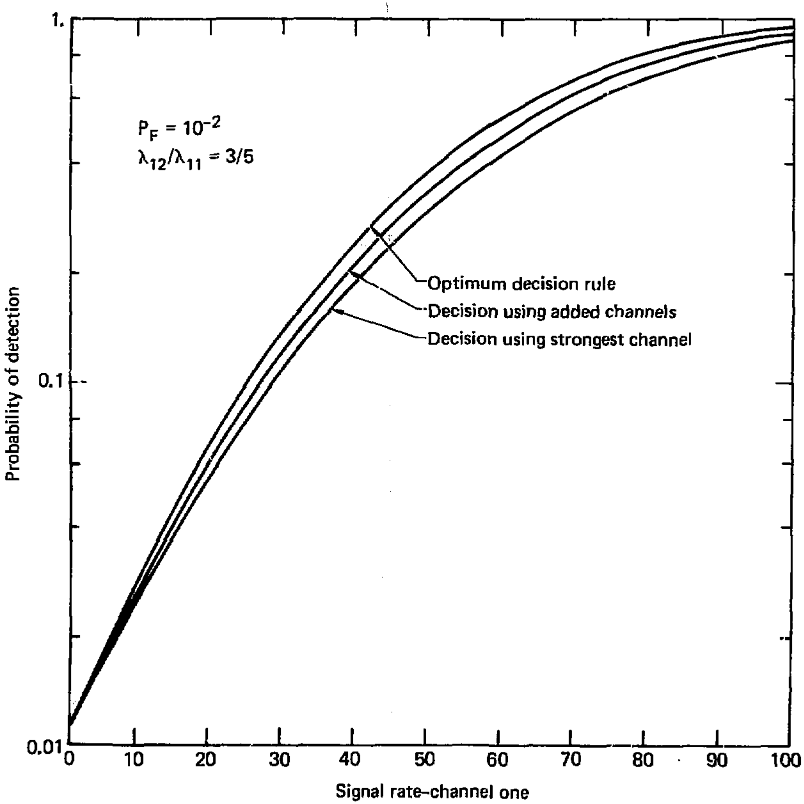


Fig. A4. P_D vs λ_{11} for three different decision rules ($\lambda_{12}/\lambda_{11} = 3/5$).

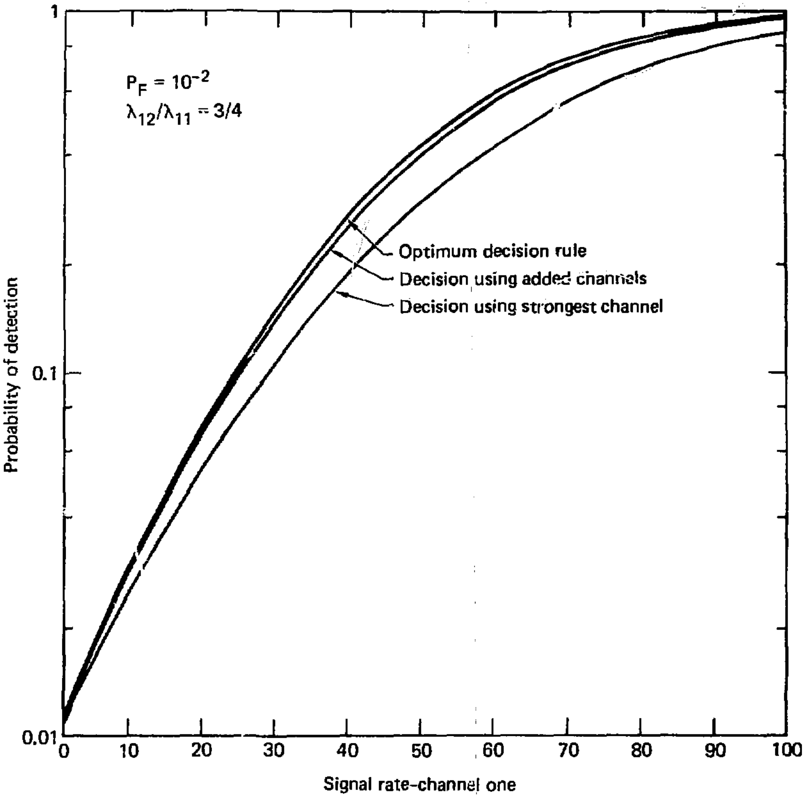


Fig. A5. P_D vs λ_{11} for three different decision rules ($\lambda_{12}/\lambda_{11} = 1/2$).