DETECTION OF SCATTERED OPTICAL FIELDS WITH FOCAL PLANE RING DETECTORS

Victor A. Vilnrotter
Space and Communications Group
Hughes Aircraft Company.

ABSTRACT

It is demonstrated that when communicating through scattering channels order-of-magnitude improvement in system performance is possible by employing focal-plane processing techniques to recover some of the scattered radiation. Maximum a-posteriori and suboptimal system structures are derived and their performance evaluated. System sensitivity to the errors in the optimal weighting functions is discussed.

I. INTRODUCTION

It is well known that optical fields transmitted through randomly scattering media appear, at the receiver, to have originated from many different directions in space. This phenomenon may be viewed as a mapping by the medium of a point-source transmitter into an extended "equivalent source". Receivers designed to operate with point-source transmitters will suffer severe performance degradation when substantial power is scattered out of the main beam. Some of the scattered power can be recovered, however, by matching the receiving system to the equivalent source.

The receiver design is perhaps most easily described in terms of system spatial modes. Strictly speaking, a spatial mode is defined as an eigenfunction of the spatial signal-field coherence-function. When the equivalent source subtends a sufficiently large solid angle at the receiver, the system spatial modes are well approximated by diffraction-limited fields-of-view. Background effects can usually be attributed to a spatially white apparent source, which contributes the same average power to each receiver mode. (In addition to random signal and background components, a deterministic signal field may also be present at the receiver, representing an attenuated version of the transmitted field. However, analysis is simplified by considering its effects negligible. If necessary, the model can easily be altered to accomodate significant deterministic components).

System analysis is readily carried out in the focal-plane of the receiver. A d-cm lens with focal length f-cm gives rise, in the focal plane, to a diffraction-limited area $A_{dl} \simeq (\lambda f/d)^2$

cm^2: hence an R-cm radius circular region contains $N = \pi R^2/A_{dl}$ spatial modes. It is convenient to assume, therefore, that the focal-plane is partitioned into N disjoint detector-elements representing the spatial modes, each with modal area A_{dl}.

We can use the above models to analyze a communications example.

An optical transmitter attempts to transfer information through a random medium by sending T-second pulses. Conforming to well-established communications theoretic notation, transmission of a pulse will be assigned to hypothesis H_l, and no transmission to hypothesis H_o. Pulse duration is assumed sufficiently long to render delay-dispersion effects negligible. Defining B_o as the system optical bandwidth, it is further postulated that $T >> 1/2\ B_o$, implying that each counting interval contains a great many time modes. Because of this condition, the random electron-count k generated by a detector-element in response to an irradiating field can be legitimately modelled as a Poisson-distributed random variable. Assuming that only Gaussian fields are received, it can be shown that counts from different modes are independent.

II. FOCAL-PLANE PROCESSING

A. MAP Detection

A maximum a-posteriori (or MAP) detector, upon observing counts from all N spatial modes over time interval T, forms the count-vector $\underline{k} = (k_1, k_2, \ldots, k_N)$ and chooses that hypothesis for which the a-posteriori probability of k is greatest: in other words, a choice is made according to the rule $P(\underline{k} \mid H_1) \gtrless_{H_0}^{H_1} P(\underline{k} \mid H_0)$ Because modal counts are independent, the log-likelihood test can be expressed as

$$\sum_{i=1}^{N} k_i \ln\left(\frac{\mu_{si} + \mu_b}{\mu_b}\right) \gtrless_{H_0}^{H_1} \sum_{i=1}^{N} \mu_{si} \triangleq \eta$$

where k_i is the observed count from spatial mode i over a properly synchronized T-second time-interval, and μ_s and μ_b are the signal and background count-energies. The log-likelihood test defines the structure of the optimal receiver: it indicates that counts from each spatial mode should be weighted according to the natural log of the modal energies under the two hypotheses, summed and compared to the total collected signal energy. Furthermore, if disjoint regions exist in the focal plane over which signal average energies are essentially constant, then an additional simplification is possible: consider the focal-plane partitioned into K disjoint constant-energy regions $R_1, R_2, \ldots, R_K$. Within a given region, weighting factors are the same for all modes, and can be taken outside the partial summation:

$$\ln\left(\frac{\mu_{sR_1}+\mu_b}{\mu_b}\right)\sum_{i\in R_1} k_i + \cdots + \ln\left(\frac{\mu_{sR_K}+\mu_b}{\mu_b}\right)\sum_{i\in R_K} k_i \underset{H_0}{\overset{H_1}{\gtrless}} \eta \,.$$

System complexity may be greatly reduced, because in typical applications K is many orders of magnitude smaller than N.

When the signal equivalent source (sometimes called “irradiance function”) is circularly symmetric, the constant-energy regions are concentric rings $\lambda f/d$ in width (as long as receiver and irradiance function are correctly aligned).

Much broader rings may yield essentially the same performance if the irradiance function can be well approximated by simple functions (or steps) defined on a sequence of “wide” concentric rings. The exact number of rings needed for a good approximation depends on the particular irradiance function under consideration. In block diagram form, a spatially processing receiver which takes advantage of circular symmetries can be represented as in Fig. la.

We may view the likelihood test in terms of the threshold comparator input z, in which case the test reduces to $z \underset{H_0}{\overset{H_1}{\gtrless}} \eta$. If the conditional probability density of z were known, then error probabilities could be obtained by integrating over appropriate regions. Unfortunately, z is the sum of weighted Poisson varriates, whose density does not appear to be expressible in closed form. However, numerical techniques have been employed to derive error probabilities exactly.

The performance of the spatially weighted receiver should be compared to that of an unweighted receiver which collects energy from the same equivalent source, but weights only according to the total signal and noise energies collected. The unweighted receiver also performs spatial processing by not observing counts from all available modes, but it does not make use of the previously defined weighting function. Its performance therefore can never be better than that of the weighted receiver. When observing N modes, this type of receiver performs the well-known test

$$k \ln\left(\frac{\sum_i^N(\mu_{si}+\mu_b)}{N\mu_b}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

in this case the resulting system structure is very simple (shown in Fig. 1b) and error probabilities can be computed with relative ease.

B. Suboptimal Systems

A meaningful indicator of performance in suboptimal systems is the signal-to-noise power ratio, usually termed SNR_p. We shall determine the extent to which focal plane processing can improve signal-to-noise ratios.

Receiver mode j is assumed to contain average signal power component P_{sj} and background noise component σ_b^2. It has been shown (1) that when observing N modes, SNR_p is maximized if the contribution from the various spatial modes are weighted by the factor $q_j = \frac{\alpha}{2B_m}\left(\frac{P_{sj}}{P_{sj}+\sigma_b^2}\right)$ (where $\alpha = \frac{q.e.}{h\nu}$ and B_m is the electrical bandwidth) in which case SNR_p is expressible as a sum of weighted signal power components:

$$SNR_p = \sum_{j=1}^{N} q_j P_{sj}$$

If power levels are essentially constant over each of K disjoint regions, then we can again regroup terms to obtain

$$SNR_p = q_1 \sum_{j \in R_1} P_{sj} + \cdots + q_K \sum_{j \in R_K} P_{sj} \ .$$

Applying this result to the previous example, and defining each ring to be roughly a mode diameter in width, the variation of SNR_p with increasing number of rings can be investigated, for a given irradiance function.

Analogous to the unweighted MAP receiver, a simpler power-collecting system results if we set each modal weighting factor equal to a constant.
Defining

$$SNR_p^* = SNR_p \Big|_{q_j = \alpha/2B_m}$$

and substituting for q_j, we arrive at the familiar shot-noise limited signal-to-noise ratio associated with point detectors:

$$SNR_p^* = \frac{\alpha}{2B_m}\left(\sum_j^N P_{sj}\right) \Big/ \left(1 + \frac{N\sigma_b^2}{\sum_j^N P_{sj}}\right)$$

For comparison, both SNR_p and SNR_p* shall be evaluated in the example which follows.

III. NUMERICAL RESULTS

A. Error Probabilities

When a signal pulse is transmitted, the spatially weighted MAP receiver commits an error only if the value of the random variable z falls below the threshold : conversely, when a pulse is not transmitted during some T-second counting interval, an error is made if z exceeds the threshold. Mathematically, these conditional error probabilities are expressed as

$$P(E|H_1) = Pr[z < \eta | H_1]$$
$$P(E|H_0) = 1 - Pr[z < \eta | H_0]$$

where z is the weighted sum of Poisson distributed counts from the various disjoint rings in the detector-plane: $z = \sum_{m=1}^{K} a_m k_{R_m}$ Because counts from different rings are independent, the probability density of z can be obtained by convolving the densities of the individual weighted counts. Numerical techniques were invoked to perform the convolution and obtain the conditional density of z under the two hypotheses. Assuming the hypotheses equi-likely, the average error probability was computed according to the rule

$$P(E) = \frac{1}{2}[P(E|H_1) + P(E|H_0)]$$

It should be emphasized that values so obtained are exact error probabilities, and not error bounds. When the computation of very small error rates is desired, our ability to evaluate P (E) may be constrained by computer storage capacity. For the case of the unweighted receiver, count probabilities can be easily calculated and summed to obtain P (E).

For our numerical example, we have modelled the mode-energy variation in the focal-plane as a circularly symmetric Gaussian function with amplitude Q and variance 1/2. Mode energies due to background are assumed constant with level P. In terms of the spatial variable r, (denoting distance from the origin in the focal plane) the communications example can be modelled as follows: when H_1 is true, the count intensity over detector ring R_m can be expressed as

$$\sum_{R_m} (\mu_{si} + \mu_b) = 2\pi Q \int_{R_m} r e^{-r^2} dr + 2\pi P \int_{R_m} r\, dr$$

If H_o is true, then

$$\sum_{R_m} \mu_b = 2\pi P \int_{R_m} r\, dr$$

It should be noted that if the rings are much wider than a mode-diameter, and ring energies are defined as above, the same improvement in system performance will be realized, even though the resulting processing scheme may not be optimal.

With the above defined modal energies, the constant-energy rings were taken to be concentric rings of width $\Delta R = 0.5$ (arbitrary units), and P (E) was computed as the number of rings were increased. (It was observed that partitioning the detector into narrower rings did not result in noticable changes in either error probabilities or signal-to-noise ratios, therefore, we can consider figures 2 and 3 essentially correct even in the limit as $\Delta R \rightarrow 0$.)

Figure 2a is a graph of error probabilities as a function of increasing number of rings for various values of Q, with fixed noise level P = 0.1. Both weighted and unweighted receivers are represented. It can be seen that as the number of rings is increased, the error-probabilities of the weighted receiver decrease monotomically, and asymptotically approach a value dictated by Q. As expected, the unweighted detector (with the same radius as the weighted detector) never outperforms its weighted counterpart. In fact, the unweighted error probability reaches a minimum value: increasing the detector radius beyond this point results in poorer performance. With a very large detector, the error probability actually approaches the value P (E) = 1/2. The reason for this phenomenon is that in our model the total available signal energy is finite, but the total available noise energy is not. A large detector collects so much noise that the receiver cannot easily distinguish between the two hypotheses, and with high probability makes an error whenever H_o is true. The weighted receiver combats this problem by essentially ignoring the contribution from those modes in which average noise energy is much greater than average signal energy. The unweighted receiver should therefore be designed to allow P (E) to reach its minimum value. It is noteworthy that by weighting ring counts optimally, a 3-dB improvement in error probability can be achieved over the best unweighted receiver, with as few as five detector rings (see, for example, Fig. 2a, Q = 4 and Q = 6). Fig. 2b confirms an anticipated result: when noise levels are increased, minimum achievable error probabilities increase as well.

The performance of diffraction-limited receivers (that is, receivers which do not attempt to recover any of the scattered radiation) is well represented in Fig. 2a as we let the detector radius approach zero. Diffraction-limited error probabilities are clearly many orders of magnitude worse than the minimum achievable values.

B. Signal-To-Noise Ratios

As with MAP receivers, great improvements in signal-to-noise ratio is possible over diffraction limited receivers by matching system parameters to the signal equivalent source. Fig. 3 indicates that SNR_p is a monotonically increasing function of the number of detector rings used, and that SNR_p always upper bounds SNR_p*. However, both functions attain roughly the same maximum values. Due to its greater simplicity, the unweighted receiver is a better choice when maximum signal-to-noise ratio is the design criterion.

IV. CONCLUSIONS

We have shown that significant improvements can be achieved in optical communications systems when some of the signal energy (or power) scattered out of the main beam during transmission is recovered in an optimal manner. Analyses of performance curves indicate that by using ring-detector arrays, spatially processing MAP receivers can attain much lower error probabilities than diffraction-limited MAP receivers. It was also shown that performance is not a strong function of weighting errors, so that signal and noise energies need not be known with great accuracy. In many cases, system complexity can be reduced if a two-fold increase in error probability can be tolerated.

Analyses of suboptimal power-collecting systems also indicate that signal-to-noise ratios associated with diffraction limited receivers can be improved upon by properly collecting scattered radiation. Again, it was found that maximum achievable signal-to-noise ratios are not very sensitive to small errors in the weighting coefficients, implying that system complexity can be greatly reduced, without substantial penalties in performance.

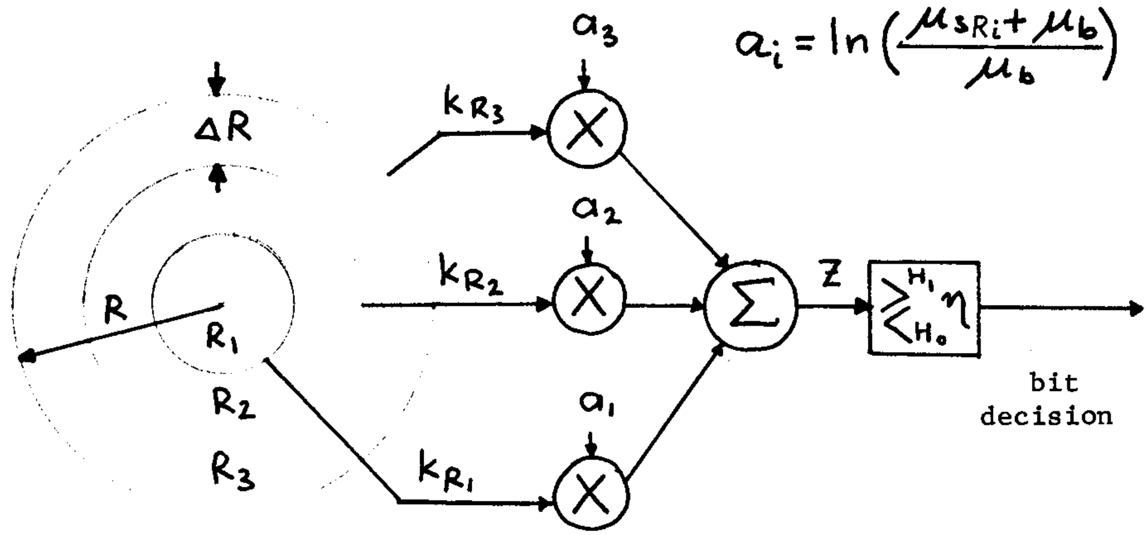
REFERENCES

1) Gagliardi and Karp, Optical Communications, Wiley, N. Y., (1976)

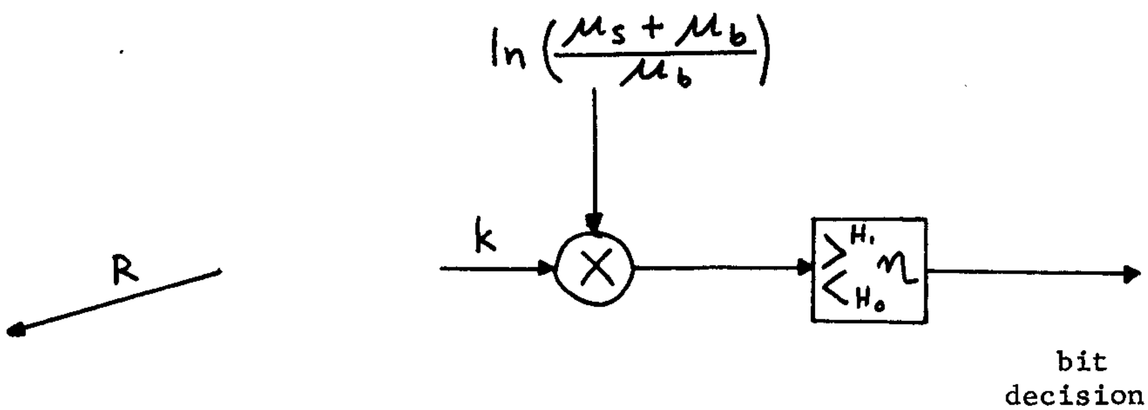
2) Papoulis, Probability, Random Variables and Stochastic Processes, McGraw Hill, N. Y., (1965)

3) Herrmann, G. "Optimum versus Suboptimum Detection Under the Poisson Regieme", IEEE Com. Trans., Vol. 21, (July, 1973)

4) Goodman, Fourier Optics, McGraw Hill, N. Y., (1968)

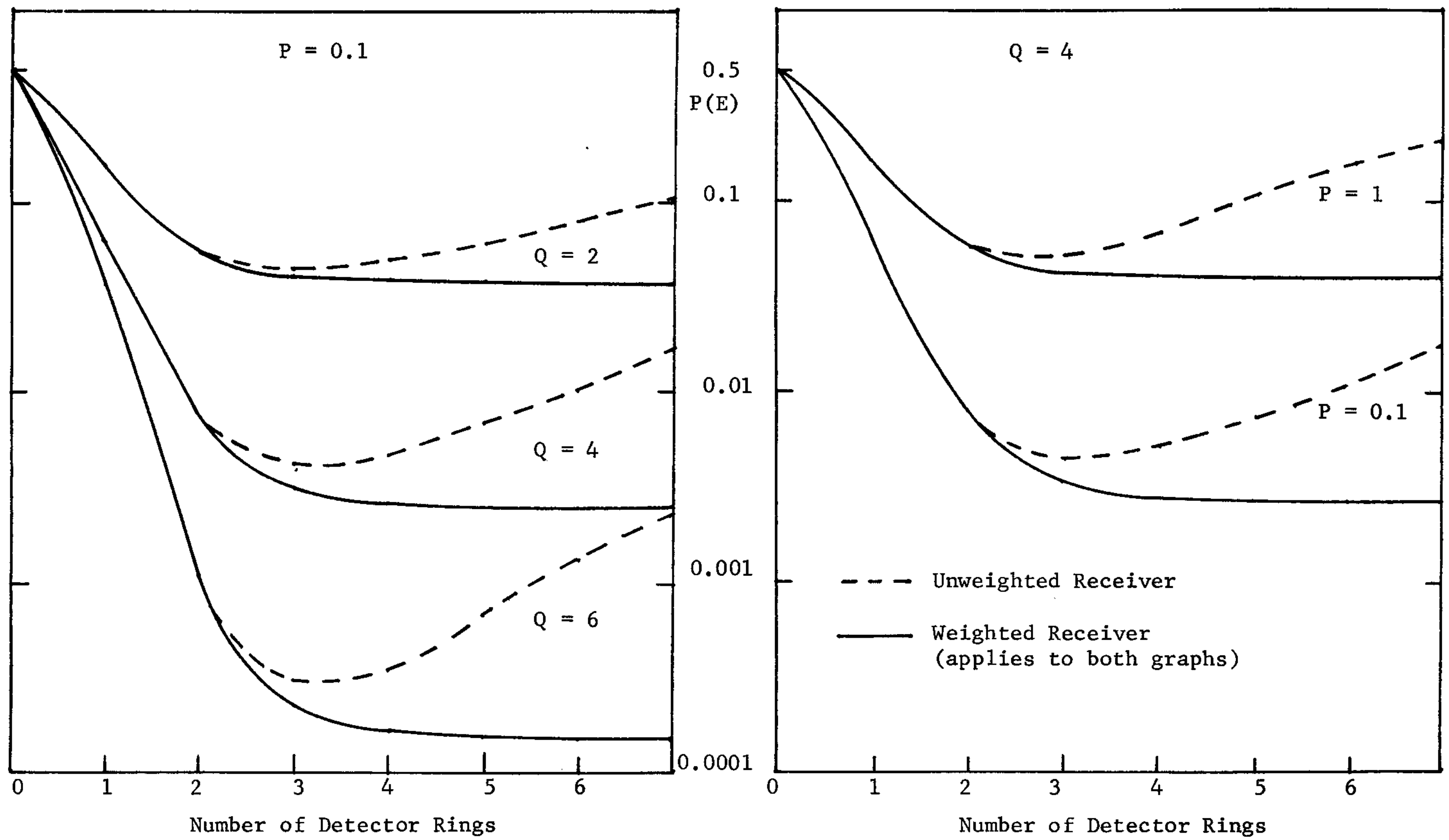


(a) WEIGHTED RECEIVER STRUCTURE



(b) UNWEIGHTED RECEIVER STRUCTURE

MAP RECEIVERS
Figure 1



(a) P(E) as a function of signal energy

(b) P(E) as a function of noise energy

Figure 2

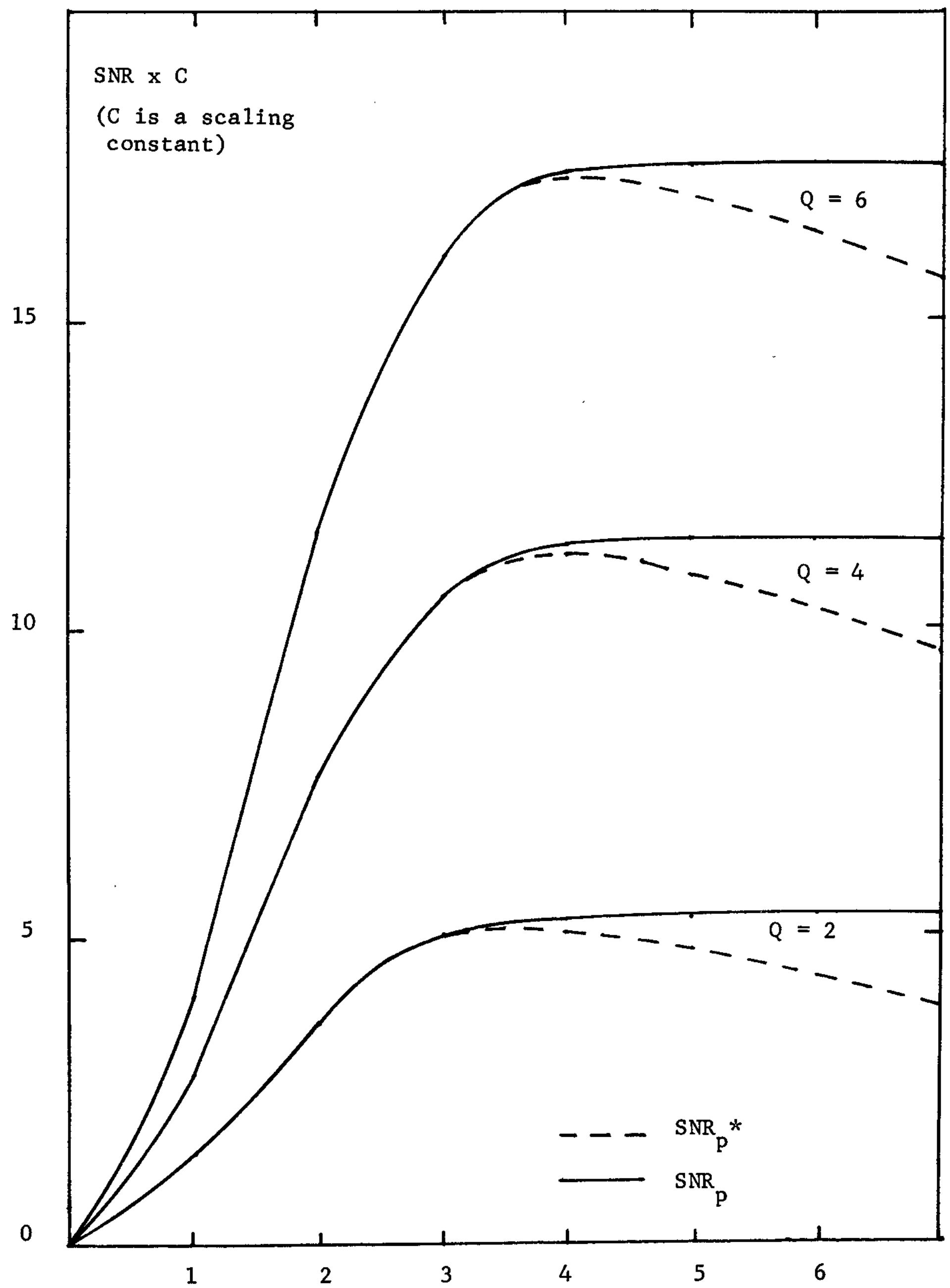


Figure 3. Signal-to-Noise Ratios