Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2012 June 1; 29(6): 1003–1016.

Objective assessment of image quality. V. Photon-counting detectors and list-mode data

Luca Caucci* and Harrison H. Barrett

College of Optical Sciences and Department of Radiology, University of Arizona, Tucson, Arizona 85724

Abstract

A theoretical framework for detection or discrimination tasks with list-mode data is developed. The object and imaging system are rigorously modeled via three random mechanisms: randomness of the object being imaged, randomness in the attribute vectors, and, finally, randomness in the attribute vector estimates due to noise in the detector outputs. By considering the list-mode data themselves, the theory developed in this paper yields a manageable expression for the likelihood of the list-mode data given the object being imaged. This, in turn, leads to an expression for the optimal Bayesian discriminant. Figures of merit for detection tasks via the ideal and optimal linear observers are derived. A concrete example discusses detection performance of the optimal linear observer for the case of a known signal buried in a random lumpy background.

1. Introduction

This paper is the fifth in a sequence of contributions to the theory of objective or task-based assessment of image quality [1–4], colloquially referred to as the OAIQs. The premise of these papers is that image quality for medical or scientific images must be defined in terms of the ability of some observer to extract useful information from the images. This task-based approach to image quality is expounded in the book *Foundations of Image Science* [5], and it is developed in detail in many additional papers in this journal [6–22] and elsewhere [23–26]. These papers consider a variety of detection and estimation tasks, and they treat human observers as well as mathematical model observers, including the ideal Bayesian observer for detection and discrimination tasks, and the ideal linear observer, known in the image-quality literature as the Hotelling observer [27]. Many of them also consider the channelized Hotelling observer [28], which is often a good predictor of the performance of a human observer on a detection task.

One thing that all of these papers have in common, however, is that they deal with images stored as arrays of pixel values or other forms of histograms. In the present paper we consider imaging systems in which a fundamentally different form of data storage, called list mode, is used. This storage mode is useful, and often unavoidable, when individual photons are detected and characterized by their position on the detector, time of detection, energy, or other attributes.

A wide variety of photon-counting detectors is used in such systems. For optical imaging, a common approach [29–32] is to use a microchannel plate or other image intensifier so that a single optical photon on the detector input can produce many secondary electrons on the intensifier output. The secondary electrons are then converted to electrical output signals by

an array of electrodes or by an output phosphor and an array of photodetectors. Alternatively, the optical photons can be incident on a large-area avalanche photodiode (APD), and the amplified charge can be collected by electrodes along the edges of the APD [33]. In all of these cases, one incident optical photon produces a set of electrical signals from which the two-dimensional (x, y) position of the photon can be estimated, and it is also straightforward to estimate its time of arrival. Thus detectors in this category estimate three attributes, x , y and t , for each detected optical photon. These estimated attributes are continuous variables, and there is no need to bin them into pixel arrays or into discrete time frames.

Optical or infrared imaging spectrometers [34] can include photon-counting detectors and estimation of spatial and spectral coordinates.

Similar methods are used for detecting x-ray or gamma-ray photons and estimating their attributes. A prime example is the Anger scintillation camera [35], which has been the standard imaging instrument in clinical nuclear medicine for over 50 years. Used to image the distribution of a gamma-ray emitting radiopharmaceutical in the patient's body, the Anger camera consists of a parallel-bore collimator or pinhole to form the image, a scintillation crystal to convert each absorbed gamma ray to a flash of light, a thick glass window to allow the light to spread, and an array of photomultipliers (PMTs) to collect the light and convert it to electrical signals. Because of its high energy, each gamma-ray photon produces thousands of optical photons, just as if an image intensifier were present, and each PMT detects some fraction of the light from each flash. All of the resulting signals are used to estimate the location within the scintillation crystal where the gamma photon was absorbed, and the final spatial resolution is much better than the size of a single PMT. As the crystal is necessarily thick in order to absorb the gamma rays efficiently, it is helpful to estimate the depth of interaction z as well as the lateral position, (x, y) . It is also possible to estimate the energy of the gamma rays from the PMT signals, which is useful in discriminating against photons that have undergone Compton scattering in the body. Finally, because the radiotracer in the patient can vary with time, it is useful to record the time of arrival of each gamma ray. Thus each gamma photon detected in an Anger camera can be characterized by five attributes: interaction position (x, y, z) , energy $\mathcal{E}$ and time of arrival t . Moreover, when the camera is used in single-photon emission computed tomography (SPECT), images must be acquired from multiple directions, and we can add an angular variable specifying the projection direction, for a total of six attributes per gamma-ray event.

Even more attributes can be measured in positron-emission tomography (PET) [36], where the radiopharmaceuticals are labeled with isotopes that decay by emission of positrons (antielectrons). A positron has a range of a few millimeters in tissue, and at the end of this range it annihilates with an electron and produces two simultaneous gamma-ray photons, each of energy 511 keV, moving in opposite directions. Typically a ring of scintillation detectors surrounding the patient captures the two coincident photons, and in the most sophisticated systems, x , y , z , $\mathcal{E}$, and t are estimated for each photon, making a total of ten estimated attributes per nuclear decay event.

Similar photon-counting systems are applicable in many other fields. X-ray computed tomography (CT) is now using photon-counting semiconductor detectors with energy resolution [37, 38], and advanced maximum-likelihood estimation methods can be used to get sub-pixel spatial resolution [39, 40]. Compton telescopes in gamma-ray astronomy [41–43] necessarily record multiple attributes of two or more interactions produced by a single incident gamma photon.

By analogy to PET, photon-counting detectors and attribute estimation also readily apply to other forms of two-photon emission, e.g., spontaneous parametric down-conversion [44] or

two-photon emission from semiconductors [45]. Even quantum-entangled two-photon imaging and interferometry [46, 47] could be performed with these methods.

When many attributes are estimated for each photon event, it may be impractical to store them in conventional binned arrays. If we measure P attributes and bin each to B bits, then it requires 2^{BP} bins to store the data. For example, if we measure six attributes and discretize each to 8 bits (256 bins per attribute), then we need $2^{48} \approx 3 \times 10^{14}$ bins. Because the total number of detected events is certain to be small compared to 10^{14} , almost all of these bins would contain no data, with a small fraction containing one event each.

The alternative to binning is a simple list of the attributes. If J events are detected and P attributes of each are estimated to 8-bit accuracy, we need PJ bytes to store the whole data set. For example, 10^9 events with six attributes per event requires a manageable 6 GBytes of storage in list-mode. List-mode data storage is now routine when photon-counting detectors are used in medical tomography [38], and it is increasingly being used in other imaging methods as suitable detectors are developed.

There is a large literature on tomographic reconstruction from list-mode data [48–56], but relatively little has been done on task-based assessment of image quality. Barrett *et al.* [57] developed an expression for detectability of a nonrandom signal on a nonrandom background from list-mode data, and Parra and Barrett [51] considered simple estimation and detection problems in time-of-flight PET. A number of papers have also treated detection performance on tomographic images reconstructed from list-mode data, using observer models that are standard for pixel-array images [49, 58, 59]. Figures of merit obtained this way thus apply to the joint quality of the raw list-mode data and the reconstruction algorithm.

The goal of this paper is to develop a rigorous statistical theory of image quality for detection or discrimination tasks with list-mode data. The theory includes the randomness of the objects being imaged; randomness in how each object generates data in photon-counting detectors; randomness in the resulting estimated attributes, and the effect of various mathematical observers, including the ideal observer and the Hotelling observer. In order to focus on the intrinsic quality of list-mode data sets, these observers act directly on the attribute lists, and image reconstruction is not considered. The theory is equally applicable to imaging of optical photons and x-ray or gamma-ray photons.

This paper is organized as follows. Section 2 introduces the mathematical notation and derives basic results concerning the ideal (Bayesian) observer, which are then applied in Sections 3 and 4 to two relevant cases: detection of a known signal in a known background, and detection of a known signal in a random background, respectively. Section 5 considers optimal linear detection with list-mode data, and concepts developed in this section are applied in Section 6 to a concrete detection problem. Finally, Section 7 concludes this paper by summarizing the main results.

2. General Theory

Much of the terminology and the basic statistics used in this paper are found in an earlier paper entitled “List-mode Likelihood” [57], and further discussion is given in [5].

2.A. Terminology and Notation

An object being imaged is assumed to be a scalar-valued function of spatial position, independent of time for simplicity. When we wish to emphasize its functional dependence, we denote the object as $f(\mathbf{r})$, where $\mathbf{r}$ is a two-dimensional (or three-dimensional) position vector. The units of $f(\mathbf{r})$ are emitted photons per unit area (or volume) per unit time.

We shall also use the notation $\mathbf{f}$ to mean a vector description of the object; in principle, $\mathbf{f}$ is the vector in Hilbert space that corresponds to the function $f(\mathbf{r})$, but in computer studies $\mathbf{f}$ may be a finite expansion in pixels, voxels or other basis functions [5].

The object is viewed by some photon-counting imaging system either for a specified time or until a specified number of detection events has been reached. In an earlier publication [57] these two modes were called preset-time and preset-counts, respectively, and it was shown that slightly different likelihood expressions were obtained in the two cases. In this paper we assume a preset acquisition time, denoted τ , so the number of counts J is a random variable, obeying Poisson statistics.

Each detection event has some attributes, such as position, energy and time of arrival. For the j^{th} event, these attributes are collected into a P -dimensional attribute vector $\mathbf{A}_j$, ($j = 1, \dots, J$), where P is the number of attributes per event and J is the number of events. We do not perfectly measure these attributes, but instead estimate them from detector signals. The estimated attribute vector for one event is denoted $\hat{\mathbf{A}}_j$, and the full list of measured attributes for J events [57], is

$$\hat{\mathcal{A}} \equiv \{\hat{\mathbf{A}}_j, j=1, \dots, J\},$$

where $\hat{\mathbf{A}}_j$ is referred to as a list entry, and the curly brackets denote a set. In preset-time mode, the data set $\mathbf{G}$ for inferences about the object consists of the J list entries and J itself, so we define $\mathbf{G} \equiv (\hat{\mathcal{A}}, J)$.

2.B. Statistical Models

For list-mode image reconstruction [48–52,56,60] or for assessment of image quality from list-mode data [51, 57], we need a likelihood, which is the probability of $\mathbf{G}$ conditional on some hypothesis H_k about the object. For a detection or classification task, H_k is the hypothesis that the object was drawn from class k . For an estimation task the hypothesis is that some parameter vector takes on a certain value from a continuous or discrete range of values.

By the usual rules of conditional probabilities, we can write the likelihood as

$$\text{pr}(\mathbf{G}|H_k) = \text{pr}(\hat{\mathcal{A}}, J|H_k) = \text{pr}(\hat{\mathcal{A}}|J, H_k) \text{Pr}(J|H_k).$$

The notation here needs some comment. We customarily [5] use $\text{pr}(\dots)$ for the probability density function (PDF) for a continuous-valued random variable or vector and $\text{Pr}(\dots)$ for the probability of a discrete-valued random variable or vector. Here, $\mathbf{G}$ is a random vector with both continuous and discrete components, so we use $\text{pr}(\dots)$ for this mixed statistical description. Moreover, subscripts on $\text{pr}(\dots)$ and $\text{Pr}(\dots)$ are omitted unless needed to avoid ambiguity, and the random quantity is assumed to be the argument of the function. Otherwise the notation is conventional, with $\text{pr}(\hat{\mathcal{A}}|J, H_k)$ being read as the PDF of $\hat{\mathcal{A}}$ conditional on J and H_k , and $\text{Pr}(J|H_k)$ indicating the probability of J conditional on H_k only.

If the objects being imaged are random, we can express the likelihood as an expectation with respect to the random objects:

$$\text{pr}(\mathbf{G}|H_k) = \int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f}) \text{pr}(\mathbf{f}|H_k) = \int d\mathbf{f} \text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) \text{Pr}(J|\mathbf{f}) \text{pr}(\mathbf{f}|H_k) \equiv \langle \text{pr}(\mathbf{G}|\mathbf{f}) \rangle_{\mathbf{f}|H_k} \quad (1)$$

where $\langle \dots \rangle$ denotes ensemble average. In principle the integral above is over all components of the object vector, though in practice the expectation might be performed by Monte Carlo sampling. Unless otherwise noted, all integrals in this paper are assumed to run over the full, usually infinite, range of the variables of integration.

For a given object, the events are statistically independent, so

$$\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) = \text{pr}(\{\hat{\mathbf{A}}_j|J, \mathbf{f}\}) = \prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}),$$

and hence

$$\text{pr}(\hat{\mathcal{A}}|J, H_k) = \langle \text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) \rangle_{\mathbf{f}|H_k} = \int d\mathbf{f} \left[\prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) \right] \text{pr}(\mathbf{f}|H_k),$$

Further applying familiar rules of probability calculus, we find

$$\text{pr}(\hat{\mathcal{A}}|J, H_k) = \int d\mathbf{f} \left[\prod_{j=1}^J \int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \text{pr}(\mathbf{A}_j|\mathbf{f}) \right] \text{pr}(\mathbf{f}|H_k) = \int d\mathbf{f} \left\{ \prod_{j=1}^J \int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \left[\int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) \text{pr}(\mathbf{r}|\mathbf{f}) \right] \right\} \text{pr}(\mathbf{f}|H_k),$$

where $\text{pr}(\mathbf{A}_j|\mathbf{r})$ is the probability density function of the true attribute vector conditioned on the event having originated at a point $\mathbf{r}$ in the object.

Now, with

$$\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) = \prod_{j=1}^J \int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \left[\int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) \text{pr}(\mathbf{r}|\mathbf{f}) \right]$$

and (1), the likelihood is given by

$$\text{pr}(\mathbf{G}|H_k) = \int d\mathbf{f} \left\{ \prod_{j=1}^J \int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \left[\int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) \text{pr}(\mathbf{r}|\mathbf{f}) \right] \right\} \text{Pr}(J|\mathbf{f}) \text{pr}(\mathbf{f}|H_k).$$

From earlier work [5, 57], we know that

$$\text{pr}(\mathbf{r}|\mathbf{f}) = \frac{f(\mathbf{r})s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}')s(\mathbf{r}')}, \quad (2)$$

where $s(\mathbf{r})$ is the system sensitivity, defined as the probability that a photon emitted at point $\mathbf{r}$ is recorded in the list. Thus $\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f})$ can be written in either of two equivalent forms:

$$\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) = \frac{\int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) f(\mathbf{r})s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}')s(\mathbf{r}')} = \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r})s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}')s(\mathbf{r}')}. \quad (3)$$

With the first of these forms, the overall probability density of the list-mode data set is

$$\text{pr}(\mathbf{G}|\mathbf{H}_k) = \int d\mathbf{f} \left\{ \prod_{j=1}^J \frac{\int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}') s(\mathbf{r}')} \right\} \text{Pr}(J|\mathbf{f}) \text{pr}(\mathbf{f}|\mathbf{H}_k), \quad (4)$$

Note that, conditioned on $\mathbf{f}$, J is a Poisson random variable, so $\text{Pr}(J|\mathbf{f})$ can be written as

$$\text{Pr}(J|\mathbf{f}) = \exp[-\bar{J}(\mathbf{f})] \frac{[\bar{J}(\mathbf{f})]^J}{J!},$$

where

$$\bar{J}(\mathbf{f}) = \tau \int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r}). \quad (5)$$

Thus a factor of $[\int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r})]^J$ in (4) cancels, and we find

$$\text{pr}(\mathbf{G}|\mathbf{H}_k) = \frac{\tau^J}{J!} \int d\mathbf{f} \left\{ \prod_{j=1}^J \int d\mathbf{A}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j) \left[\int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right] \right\} \exp[-\bar{J}(\mathbf{f})] \text{pr}(\mathbf{f}|\mathbf{H}_k).$$

In this expression, the task is defined by $\text{pr}(\mathbf{f}|\mathbf{H}_k)$; the modality is defined by $\text{pr}(\mathbf{A}_j|\mathbf{f})$ or equivalently, $\int d\mathbf{r} \text{pr}(\mathbf{A}_j|\mathbf{r}) \text{pr}(\mathbf{r}|\mathbf{f})$; and the properties of the detector are characterized by $\text{pr}(\hat{\mathbf{A}}_j|\mathbf{A}_j)$.

With the second form in (3), which does not show the role of the detector explicitly, we have

$$\text{pr}(\mathbf{G}|\mathbf{H}_k) = \int d\mathbf{f} \left\{ \prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}') s(\mathbf{r}')} \right\} \text{Pr}(J|\mathbf{f}) \text{pr}(\mathbf{f}|\mathbf{H}_k) = \left\langle \left\{ \prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r}' f(\mathbf{r}') s(\mathbf{r}')} \right\} \text{Pr}(J|\mathbf{f}) \right\rangle_{\mathbf{f}|\mathbf{H}_k},$$

where, we recall, the angle brackets $\langle \dots \rangle_{\mathbf{f}|\mathbf{H}_k}$ denote an average over objects in class k .

Other equivalent forms are

$$\begin{aligned} \text{pr}(\mathbf{G}|\mathbf{H}_k) &= \frac{\tau^J}{J!} \int d\mathbf{f} \left\{ \prod_{j=1}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right\} \exp[-\bar{J}(\mathbf{f})] \text{pr}(\mathbf{f}|\mathbf{H}_k) = \int d\mathbf{f} \left\{ \prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) \right\}_{(7)} \exp[\\ &\quad - \bar{J}(\mathbf{f})] \frac{[\bar{J}(\mathbf{f})]^J}{J!} \text{pr}(\mathbf{f}|\mathbf{H}_k). \end{aligned}$$

3. Detection of a Known Signal on a Known Background

3.A. Likelihood Ratio

For an SKE/BKE (signal known exactly, background known exactly [5]) detection task, H_k is specified by the known object, which we can denote as $\mathbf{f}_0$ in the signal-absent case and $\mathbf{f}_1 \equiv \mathbf{f}_0 + \Delta\mathbf{f}$ in the signal-present case. Thus [cf. (7)],

$$\text{pr}(\mathbf{G}|H_k) = \text{pr}(\mathbf{G}|\mathbf{f}_k) = \frac{\tau^J}{J!} \left\{ \prod_{j=1}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_k(\mathbf{r}) s(\mathbf{r}) \right\} \exp[-\bar{J}(\mathbf{f}_k)].$$

The expression above is obtained from (7) by assuming $\text{pr}(\mathbf{f}|H_k) = \delta(\mathbf{f} - \mathbf{f}_k)$, where $\delta(\dots)$ denotes the delta function. The log-likelihood is given by

$$\log[\text{pr}(\mathbf{G}|\mathbf{f}_k)] = J \log \tau - \log J! - \bar{J}(\mathbf{f}_k) + \sum_{j=1}^J \left\{ \log \left[\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_k(\mathbf{r}) s(\mathbf{r}) \right] \right\}.$$

The ideal observer uses the likelihood ratio $\Lambda(\mathbf{G}|\mathbf{f}_0) = \text{pr}(\mathbf{G}|H_1)/\text{pr}(\mathbf{G}|H_0)$ or its logarithm $\lambda(\mathbf{G}|\mathbf{f}_0) = \log \Lambda(\mathbf{G}|\mathbf{f}_0)$ as the test statistic. For the present SKE/BKE problem, the likelihood ratio is given by

$$\Lambda(\mathbf{G}|\mathbf{f}_0) = \frac{\text{pr}(\mathbf{G}|\mathbf{f}_0 + \Delta\mathbf{f})}{\text{pr}(\mathbf{G}|\mathbf{f}_0)} = \exp[-\bar{J}(\mathbf{f}_0 + \Delta\mathbf{f}) + \bar{J}(\mathbf{f}_0)] \prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) [f_0(\mathbf{r}) + \Delta f(\mathbf{r})] s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \quad (8)$$

and its logarithm becomes

$$\begin{aligned} \lambda(\mathbf{G}|\mathbf{f}_0) &= \log \left[\frac{\text{pr}(\mathbf{G}|\mathbf{f}_0 + \Delta\mathbf{f})}{\text{pr}(\mathbf{G}|\mathbf{f}_0)} \right] \\ &= -\bar{J}(\mathbf{f}_0 + \Delta\mathbf{f}) + \bar{J}(\mathbf{f}_0) + \log \prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) [f_0(\mathbf{r}) + \Delta f(\mathbf{r})] s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \\ &= -\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) + \sum_{j=1}^J \log \left[1 + \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right]. \end{aligned} \quad (9)$$

If the signal is weak, the logarithms can be expanded to yield:

$$\lambda(\mathbf{G}|\mathbf{f}_0) = -\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) + \sum_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} - \frac{1}{2} \sum_{j=1}^J \left[\frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right]^2 \dots \quad (10)$$

3.B. Figures of Merit for SKE/BKE Detection by an Ideal Observer

We can use the log of the likelihood ratio from (10) to compute an ideal-observer detectability by noting that the terms in the sum in (9) are independent and identically distributed, so the log-likelihood is normally distributed by the central-limit theorem. Normal log-likelihoods have some powerful properties, discussed in detail in Sec. 4C of [3]. In particular, the means and variances of a normal log-likelihood under the two hypotheses must be related by:

$$\langle \lambda \rangle_1 = -\langle \lambda \rangle_0, \quad \text{var}_1(\lambda) = \text{var}_0(\lambda) = 2\langle \lambda \rangle_1, \quad (11)$$

where the angle brackets denote expectations with respect to the list-mode data under the subscripted hypotheses; thus $\langle \lambda \rangle_k$ is a shorthand for $\langle \lambda(\mathbf{G}) \rangle_{\mathbf{G}|\mathbf{f}_k} \equiv \int d\mathbf{G} \lambda(\mathbf{G}) \text{pr}(\mathbf{G}|\mathbf{f}_k)$.

The ideal-observer SNR in this problem is given by [3]

$$\text{SNR}^2 \equiv \frac{[\langle \lambda \rangle_1 - \langle \lambda \rangle_0]^2}{\frac{1}{2}[\text{var}_0(\lambda) + \text{var}_1(\lambda)]} = 2\langle \lambda \rangle_1 - 2\langle \lambda \rangle_0.$$

Moreover, for any normally distributed test statistic, the area under the ROC curve (AUC) is given by [5] $\text{AUC} = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{1}{2} \text{SNR}\right)$, where $\text{erf}(\dots)$ denotes the error function. Thus, because the log-likelihood is normal, the ideal-observer performance, specified by either SNR or AUC, is fully determined by $\langle \lambda \rangle_0$.

Let us now compute $\langle \lambda \rangle_0$ for the expression in (10), term by term. The first term, $-\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r})$, is not random in this SKE/BKE problem, so it is unchanged by the averaging. The average of the second term is given by

$$\begin{aligned} & \left\langle \sum_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\text{GIH}_0} \\ &= \sum_{J=0}^{\infty} \text{Pr}(J|f_0) \sum_{j=1}^J \int d\hat{\mathbf{A}}_j \text{pr}(\hat{\mathbf{A}}_j|f_0) \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \\ &= \sum_{J=0}^{\infty} \text{Pr}(J|f_0) \sum_{j=1}^J \int d\hat{\mathbf{A}}_j \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f_0(\mathbf{r}) s(\mathbf{r})} \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \\ &= \sum_{J=0}^{\infty} \text{Pr}(J|f_0) \sum_{j=1}^J \frac{\int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \int d\hat{\mathbf{A}}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r})}{\int d\mathbf{r} f_0(\mathbf{r}) s(\mathbf{r})} \\ &= \sum_{J=0}^{\infty} \text{Pr}(J|f_0) \sum_{j=1}^J \frac{\int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f_0(\mathbf{r}) s(\mathbf{r})}, \end{aligned}$$

where, in the second line we have used (4) and in the fourth line we have used $\int d\hat{\mathbf{A}}_j \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) = 1$. Finally, we note that all J terms in the sum yield the same result, so, with (5),

$$\left\langle \sum_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\text{GIH}_0} = \frac{\int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f_0(\mathbf{r}) s(\mathbf{r})} \sum_{J=0}^{\infty} \text{Pr}(J|f_0) J = \tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}),$$

which just cancels the nonrandom first term in (10).

If we truncate (10) after the terms quadratic in Δf and perform manipulations similar to the above, we obtain

$$\langle \lambda \rangle_0 \approx -\frac{1}{2} \sum_{J=0}^{\infty} \text{Pr}(J|f_0) \sum_{j=1}^J \int d\hat{\mathbf{A}}_j \text{pr}(\hat{\mathbf{A}}_j|f_0) \left[\frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right]^2 = -\frac{1}{2} \tau \int d\hat{\mathbf{A}} \frac{\left[\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})}. \quad (12)$$

Thus the ideal-observer SNR^2 for a weak-signal SKE/BKE problem is given by

$$\text{SNR}^2 \approx \tau \int d\hat{\mathbf{A}} \frac{\left[\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})}. \quad (13)$$

Note that SNR^2 increases linearly with the acquisition time and hence with the mean number of photons in the data set.

An equivalent form is

$$\text{SNR}^2 \approx \bar{J}(f_0) \int d\hat{\mathbf{A}} \text{pr}(\hat{\mathbf{A}}|f_0) \left[\frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right]^2 = \bar{J}(f_0) \left\langle \left[\frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) f_0(\mathbf{r}) s(\mathbf{r})} \right]^2 \right\rangle_{\hat{\mathbf{A}}|f_0}.$$

This form is useful if we wish to evaluate SNR^2 by drawing Monte Carlo samples of $\hat{\mathbf{A}}$.

4. SKE/BKS Detection

The assumption of a known background in Sec. 3 is obviously unrealistic, and we should treat the background as stochastic. In this section we modify the results of Sec. 3 to include the effect of background randomness. The key difference is that the individual $\hat{\mathbf{A}}_j$ are statistically independent for a given object, but the randomness of the object induces statistical dependence. Indeed:

$$\begin{aligned} \text{pr}(\hat{\mathbf{A}}_j, \hat{\mathbf{A}}_{j'}|H_k) \\ &= \int d\mathbf{f} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{f}) \text{pr}(\mathbf{f}|H_k) \neq \left[\int d\mathbf{f} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) \text{pr}(\mathbf{f}|H_k) \right] \left[\int d\mathbf{f} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{f}) \text{pr}(\mathbf{f}|H_k) \right] \\ &= \text{pr}(\hat{\mathbf{A}}_j|H_k) \text{pr}(\hat{\mathbf{A}}_{j'}|H_k). \end{aligned}$$

4.A. Likelihood Ratio

Consider the detection of a known signal that, when it is present, is simply added to a random background $b(\mathbf{r})$. In the literature, this task is referred to as SKE/BKS (signal known exactly, background known statistically). For this task, the object under the signal-absent hypothesis is given by $f(\mathbf{r}) = b(\mathbf{r})$, and under the signal-present hypothesis it is given by $f(\mathbf{r}) = b(\mathbf{r}) + \Delta f(\mathbf{r})$. In vector notation, the two hypotheses are $H_0 : \mathbf{f} = \mathbf{b}$ and $H_1 : \mathbf{f} = \mathbf{b} + \Delta \mathbf{f}$, where now the background $\mathbf{b}$ is the only random quantity in the problem; its probability density function is denoted $\text{pr}_{\mathbf{b}}(\mathbf{b})$. Thus averaging over $\mathbf{f}$ conditional on H_0 is the same as averaging over $\mathbf{b}$.

From (1),

$$\begin{aligned} \text{pr}(\mathbf{G}|H_0) &= \int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f}) \text{pr}_{\mathbf{b}}(\mathbf{f}) = \int d\mathbf{f} \text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) \text{Pr}(J|\mathbf{f}) \text{pr}_{\mathbf{b}}(\mathbf{f}), \\ \text{pr}(\mathbf{G}|H_1) &= \int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f} + \Delta \mathbf{f}) \text{pr}_{\mathbf{b}}(\mathbf{f}) = \int d\mathbf{f} \text{pr}(\hat{\mathcal{A}}|J, \mathbf{f} + \Delta \mathbf{f}) \text{Pr}(J|\mathbf{f} + \Delta \mathbf{f}) \text{pr}_{\mathbf{b}}(\mathbf{f}). \end{aligned}$$

The likelihood ratio is

$$\Lambda(\mathbf{G}) \equiv \frac{\text{pr}(\mathbf{G}|H_1)}{\text{pr}(\mathbf{G}|H_0)} = \frac{\int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f}+\Delta\mathbf{f}) \text{pr}_b(\mathbf{f})}{\int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f}) \text{pr}_b(\mathbf{f})} = \frac{\langle \text{pr}(\mathbf{G}|\mathbf{f}+\Delta\mathbf{f}) \rangle_{\mathbf{f}|H_0}}{\langle \text{pr}(\mathbf{G}|\mathbf{f}) \rangle_{\mathbf{f}|H_0}}.$$

Because $\Delta\mathbf{f}$ is nonrandom, we can take it out of the expectation integrals and then expand $\Lambda(\mathbf{G})$ in a Taylor series as

$$\Lambda(\mathbf{G}) = 1 + \frac{\Delta\mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} + \frac{1}{2} \frac{\Delta\mathbf{f}^\dagger \mathbf{B} \Delta\mathbf{f}}{\text{pr}(\mathbf{G}|H_0)} + \dots, \quad (14)$$

where

$$\mathbf{a} = \int d\mathbf{f} \left[\frac{\partial \text{pr}(\mathbf{G}|\mathbf{f})}{\partial \mathbf{f}} \right] \text{pr}_b(\mathbf{f}) = \left\langle \frac{\partial \text{pr}(\mathbf{G}|\mathbf{f})}{\partial \mathbf{f}} \right\rangle_{\mathbf{f}|H_0}$$

and

$$\mathbf{B} = \int d\mathbf{f} \left[\frac{\partial^2 \text{pr}(\mathbf{G}|\mathbf{f})}{\partial \mathbf{f} \partial \mathbf{f}^\dagger} \right] \text{pr}_b(\mathbf{f}) = \left\langle \frac{\partial^2 \text{pr}(\mathbf{G}|\mathbf{f})}{\partial \mathbf{f} \partial \mathbf{f}^\dagger} \right\rangle_{\mathbf{f}|H_0}.$$

The notation used here is discussed in detail in Sec. A.10.2 of [5]. In brief, $\partial/\partial \mathbf{f}$ is the gradient with respect to $\mathbf{f}$; that is, if $\mathbf{f}$ is represented as an $N \times 1$ column vector with components f_n , then $\partial \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f}$ is a column vector with components $\partial \text{pr}(\mathbf{G}|\mathbf{f})/\partial f_n$. With this notation, $\Delta\mathbf{f}^\dagger [\partial \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f}]$ indicates the scalar product of the two vectors $\Delta\mathbf{f}$ and $\partial \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f}$. Similarly, $\partial^2 \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f} \partial \mathbf{f}^\dagger$ is an $N \times N$ matrix of second derivatives, and the quadratic form $\Delta\mathbf{f}^\dagger [\partial^2 \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f} \partial \mathbf{f}^\dagger] \Delta\mathbf{f}$ is to be interpreted as the scalar product of $\Delta\mathbf{f}$ with $[\partial^2 \text{pr}(\mathbf{G}|\mathbf{f})/\partial \mathbf{f} \partial \mathbf{f}^\dagger] \Delta\mathbf{f}$. If $\mathbf{f}$ is the infinite-dimensional Hilbert-space vector corresponding to the function $f(\mathbf{r})$, then $\partial/\partial \mathbf{f}$ is interpreted as a functional or Fréchet derivative [5].

The notation does not show it, but $\mathbf{a}$ and $\mathbf{B}$ are random functions of the entire list-mode data set $\mathbf{G}$.

From (14), the logarithm of the likelihood ratio is

$$\lambda(\mathbf{G}) = \frac{\Delta\mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} + \frac{1}{2} \frac{\Delta\mathbf{f}^\dagger \mathbf{B} \Delta\mathbf{f}}{\text{pr}(\mathbf{G}|H_0)} - \frac{1}{2} \left[\frac{\Delta\mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} + \frac{1}{2} \frac{\Delta\mathbf{f}^\dagger \mathbf{B} \Delta\mathbf{f}}{\text{pr}(\mathbf{G}|H_0)} \right]^2 + \dots$$

Through terms quadratic in $\Delta\mathbf{f}$,

$$\lambda(\mathbf{G}) \approx \frac{\Delta\mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} + \frac{1}{2} \frac{\Delta\mathbf{f}^\dagger \mathbf{B} \Delta\mathbf{f}}{\text{pr}(\mathbf{G}|H_0)} - \frac{1}{2} \left[\frac{\Delta\mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} \right]^2. \quad (15)$$

This expression is the generalization of (9) for detection of weak nonrandom signals on random backgrounds (SKE/BKS tasks).

4.B. Figures of Merit

The first term in (15) is given explicitly by

$$\frac{\Delta \mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} = \frac{\Delta \mathbf{f}^\dagger \int d\mathbf{f} \left[\frac{\partial}{\partial \mathbf{f}} \text{pr}(\mathbf{G}|\mathbf{f}) \right] \text{pr}_b(\mathbf{f})}{\int d\mathbf{f} \text{pr}(\mathbf{G}|\mathbf{f}) \text{pr}_b(\mathbf{f})}.$$

From (7), we know that

$$\text{pr}(\mathbf{G}|\mathbf{f}) = \frac{\tau^J}{J!} \exp[-\bar{J}(\mathbf{f})] \prod_{j=1}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}).$$

With the chain rule, we can show that

$$\begin{aligned} \Delta \mathbf{f}^\dagger \frac{\partial}{\partial \mathbf{f}} \text{pr}(\mathbf{G}|\mathbf{f}) &= -\Delta \mathbf{f}^\dagger \frac{\partial \bar{J}(\mathbf{f})}{\partial \mathbf{f}} \text{pr}(\mathbf{G}|\mathbf{f}) \\ &+ \frac{\tau^J}{J!} \exp[-\bar{J}(\mathbf{f})] \sum_{j=1}^J \left\{ \left[\Delta \mathbf{f}^\dagger \frac{\partial}{\partial \mathbf{f}} \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right] \prod_{\substack{j'=1 \\ j' \neq j}}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right\} \\ &= -\tau \text{pr}(\mathbf{G}|\mathbf{f}) \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \\ &+ \frac{\tau^J}{J!} \exp[-\bar{J}(\mathbf{f})] \sum_{j=1}^J \left\{ \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r}) \prod_{\substack{j'=1 \\ j \neq j'}}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right\}, \end{aligned}$$

where the product over j' now has $J-1$ factors, with $j'=j$ excluded.

Taking the average with respect to $\text{pr}_b(\mathbf{f})$ and doing a little algebra, we find

$$\Delta \mathbf{f}^\dagger \mathbf{a} = -\tau \text{pr}(\mathbf{G}|H_0) \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) + \frac{\tau^J}{J!} \left\langle \exp[-\bar{J}(\mathbf{f})] \sum_{j=1}^J \left\{ \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})} \prod_{j'=1}^J \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}) \right\} \right\rangle_{\mathbf{f}|H_0}.$$

With (5),

$$\begin{aligned} \Delta \mathbf{f}^\dagger \mathbf{a} &= -\tau \text{pr}(\mathbf{G}|H_0) \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \\ &+ \frac{1}{J!} \left\langle \exp[-\bar{J}(\mathbf{f})] [\bar{J}(\mathbf{f})]^J \sum_{j=1}^J \left\{ \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})} \prod_{j'=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r})} \right\} \right\rangle_{\mathbf{f}|H_0}. \end{aligned}$$

Thus, with (7),

$$\frac{\Delta \mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} = -\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) + \frac{\left\langle \exp[-\bar{J}(\mathbf{f})][\bar{J}(\mathbf{f})]^J \sum_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})} \prod_{j'=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_{j'}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\mathbf{f}|H_0}}{\left\langle \exp[-\bar{J}(\mathbf{f})][\bar{J}(\mathbf{f})]^J \prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\mathbf{f}|H_0}}. \quad (16)$$

To simplify this complicated expression, we note that

$$\prod_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} f(\mathbf{r}) s(\mathbf{r})} = \text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}).$$

If we interpret $\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f})$ as a function of $\mathbf{f}$ for fixed $\hat{\mathcal{A}}$ and J , then $\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f})$ is the list-mode likelihood for estimation of $\mathbf{f}$. By definition, $\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f})$ peaks at the maximum-likelihood estimate of $\mathbf{f}$, which, under H_0 , we denote as $\hat{\mathbf{b}}_{\text{ML}}$. If this peak is sharp compared to the width of $\text{pr}_b(\mathbf{f})$, we can say that

$$\text{pr}(\hat{\mathcal{A}}|J, \mathbf{f}) \approx C \delta(\mathbf{f} - \hat{\mathbf{b}}_{\text{ML}}). \quad (17)$$

Of course, $\hat{\mathbf{b}}_{\text{ML}}$ depends on the list-mode data $\hat{\mathcal{A}}$, but we suppress this dependence for notational convenience.

When we use the delta function to perform the two averages over $\mathbf{f}$ in (16), many cancellations occur and we find

$$\frac{\Delta \mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} = -\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) + \sum_{j=1}^J \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \hat{\mathbf{b}}_{\text{ML}}(\mathbf{r}) s(\mathbf{r})}.$$

For notational clarity, we define

$$\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \equiv \frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \hat{\mathbf{b}}_{\text{ML}}(\mathbf{r}) s(\mathbf{r})}. \quad (18)$$

Notice that

$$\left\langle \frac{\Delta \mathbf{f}^\dagger \mathbf{a}}{\text{pr}(\mathbf{G}|H_0)} \right\rangle_{\mathbf{G}|H_0} = \Delta \mathbf{f}^\dagger \left\{ \int d\mathbf{G} \int d\mathbf{f} \left[\frac{\partial}{\partial \mathbf{f}} \text{pr}(\mathbf{G}|\mathbf{f}) \right] \text{pr}_b(\mathbf{f}) \right\} = \Delta \mathbf{f}^\dagger \left\{ \int d\mathbf{f} \left[\frac{\partial}{\partial \mathbf{f}} \int d\mathbf{G} \text{pr}(\mathbf{G}|\mathbf{f}) \right] \text{pr}_b(\mathbf{f}) \right\} = 0,$$

and, similarly

$$\left\langle \frac{\Delta \mathbf{f}^\dagger \mathbf{B} \Delta \mathbf{f}}{\text{pr}(\mathbf{G}|H_0)} \right\rangle_{\mathbf{G}|H_0} = \Delta \mathbf{f}^\dagger \left\{ \int d\mathbf{G} \int d\mathbf{f} \left[\frac{\partial^2}{\partial \mathbf{f} \partial \mathbf{f}^\dagger} \text{pr}(\mathbf{G}|\mathbf{f}) \right] \text{pr}_b(\mathbf{f}) \right\} \Delta \mathbf{f} = \Delta \mathbf{f}^\dagger \left\{ \int d\mathbf{f} \left[\frac{\partial^2}{\partial \mathbf{f} \partial \mathbf{f}^\dagger} \int d\mathbf{G} \text{pr}(\mathbf{G}|\mathbf{f}) \right] \text{pr}_b(\mathbf{f}) \right\} = 0,$$

where we used the fact that $\int d\mathbf{G} \text{pr}(\mathbf{G}|\mathbf{f}) = 1$. Using the results above and with the help of (11), we have

$$\text{SNR}^2 \approx \left\langle \sum_{j=1}^J \sum_{j'=1}^J \Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}}) \right\rangle_{\mathbf{G}|\mathbf{H}_0} - \left[\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2.$$

In the double sum over j and j' , there are J terms with $j = j'$ and $J^2 - J$ terms with $j \neq j'$. For each case, the terms are identically distributed, so we can write

$$\text{SNR}^2 \approx \left\langle J [\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}})]^2 \right\rangle_{\mathbf{G}|\mathbf{H}_0} + \left\langle (J^2 - J) \Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}}) \right\rangle_{\mathbf{G}|\mathbf{H}_0} - \left[\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2 \quad (9)$$

with the understanding that $j \neq j'$.

In more detail, the first term in (19) is

$$\begin{aligned} & \left\langle J [\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}})]^2 \right\rangle_{\mathbf{G}|\mathbf{H}_0} \\ &= \left\langle \left\langle J [\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}})]^2 \right\rangle_{j|\hat{\mathbf{b}}_{\text{ML}}} \right\rangle_{\hat{\mathbf{A}}_j|\hat{\mathbf{b}}_{\text{ML}}, \hat{\mathbf{b}}_{\text{ML}}|\mathbf{H}_0} \\ &= \left\langle \bar{J}(\hat{\mathbf{b}}_{\text{ML}}) \left\langle [\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}})]^2 \right\rangle_{\hat{\mathbf{A}}_j|\hat{\mathbf{b}}_{\text{ML}}} \right\rangle_{\hat{\mathbf{b}}_{\text{ML}}|\mathbf{H}_0} \\ &= \tau \left\langle \int d\hat{\mathbf{A}}_j \frac{\left[\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}_j|\mathbf{r}) \hat{\mathbf{b}}_{\text{ML}}(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\hat{\mathbf{b}}_{\text{ML}}|\mathbf{H}_0}, \end{aligned}$$

where we have used (18), (5) and (1). Similarly, the second term in (19) is

$$\left\langle (J^2 - J) \Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}}) \right\rangle_{\mathbf{G}|\mathbf{H}_0} = \left\langle \left\langle J^2 - J \right\rangle_{j|\hat{\mathbf{b}}_{\text{ML}}} \left\langle \Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}}) \right\rangle_{\hat{\mathbf{A}}_j, \hat{\mathbf{A}}_{j'}|\hat{\mathbf{b}}_{\text{ML}}} \right\rangle_{\hat{\mathbf{b}}_{\text{ML}}|\mathbf{H}_0}.$$

Because J conditional on its mean is a Poisson random variable, and because $\langle J^2 - J \rangle = \langle J \rangle^2$ for a Poisson random variable,

$$\langle J^2 - J \rangle_{j|\hat{\mathbf{b}}_{\text{ML}}} = [\bar{J}(\hat{\mathbf{b}}_{\text{ML}})]^2.$$

Moreover, conditional on $\hat{\mathbf{b}}_{\text{ML}}$, $\Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}})$ and $\Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}})$ are statistically independent for $j \neq j'$, and the averages can be performed with (5) and (1). After some algebra, we find

$$\left\langle (J^2 - J) \Phi(\hat{\mathbf{A}}_j, \hat{\mathbf{b}}_{\text{ML}}) \Phi(\hat{\mathbf{A}}_{j'}, \hat{\mathbf{b}}_{\text{ML}}) \right\rangle_{\mathbf{G}|\mathbf{H}_0} = \left[\tau \int d\mathbf{r} \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2.$$

Thus the second and third terms in (19) cancel, and we are left with [cf. (12)]

$$\text{SNR}^2 \approx \tau \left\langle \int d\hat{\mathbf{A}} \frac{\left[\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r}) \right]^2}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \hat{b}_{\text{ML}}(\mathbf{r}) s(\mathbf{r})} \right\rangle_{\hat{b}_{\text{ML}}|H_0}. \quad (20)$$

This expression gives the performance of an ideal observer on a weak-signal SKE/BKS task, provided the approximation in (17) is valid.

In a form that is amenable to Monte Carlo evaluation [cf. (13)], we can also write

$$\text{SNR}^2 \approx \left\langle \bar{J}(\hat{b}_{\text{ML}}) \left\langle \left[\frac{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \Delta f(\mathbf{r}) s(\mathbf{r})}{\int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \hat{b}_{\text{ML}}(\mathbf{r}) s(\mathbf{r})} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\hat{b}_{\text{ML}}} \right\rangle_{\hat{b}_{\text{ML}}|H_0}. \quad (21)$$

If the approximations above are not applicable (and that would happen, for example, when the signal is not weak so that the approximation in (14) is not valid) or when computation requirements to evaluate (21) are prohibitive, other computational methods can be used. As an example, Kupinski *et al.* [15] show the application of Markov chain Monte Carlo (MCMC) techniques [61] to the problem of estimating ideal-observer detection performance by calculating $\Lambda(\mathbf{G})$ for randomly generated datasets $\mathbf{G}$. Although [15] assumed pixelated data, the same scheme can be adapted to the case of list-mode data $\mathbf{G}$

$$\Lambda(\mathbf{G}) \approx \frac{1}{N} \sum_{n=1}^N \Lambda(\mathbf{G}|\mathbf{f}_{\Phi_n}), \quad (22)$$

where $\Phi_1, \dots, \Phi_N$ are parameterizations of the background object $\mathbf{f} = \mathbf{f}_{\Phi_n}$ MCMC-generated from list-mode data $\mathbf{G}$, and $\Lambda(\mathbf{G}|\mathbf{f}_{\Phi_n})$ is the SKE/BKE likelihood ratio

$$\Lambda(\mathbf{G}|\mathbf{f}_{\Phi_n}) = \frac{\text{pr}(\mathbf{G}|\mathbf{f}_{\Phi_n} + \Delta \mathbf{f})}{\text{pr}(\mathbf{G}|\mathbf{f}_{\Phi_n})},$$

defined in (8). We refer the interested reader to [61, 62] for a detailed theoretical treatment of MCMC methods. Details of our MCMC calculation will be provided in a separate paper currently under preparation. Theoretical results derived by Barrett *et al.* [3] show that for the case of the log-likelihood ratio, $\text{SNR}^2 = -8 \log \langle \Lambda^{1/2} \rangle_0$ in which the ensemble average $\langle \Lambda^{1/2} \rangle_0$ is estimated as the sample average of the $\Lambda^{1/2}(\mathbf{G}_1), \dots, \Lambda^{1/2}(\mathbf{G}_M)$, where each $\Lambda(\mathbf{G}_m)$ is calculated via MCMC sampling of (22).

5. List-mode Hotelling Observer

5.A. Poisson Point Process in Attribute Space

For the conventional Hotelling observer $t(\mathbf{g}) = \mathbf{w}^T \mathbf{g}$ operating on binned data $\mathbf{g}$ [5, 28], the signal s to be detected is the difference $\langle \mathbf{g} \rangle_1 - \langle \mathbf{g} \rangle_0$ in the mean data under the two hypotheses

and $\mathbf{w} = \bar{\mathbf{K}}_g^{-1} \mathbf{s}$, in which $\bar{\mathbf{K}}_g = \frac{1}{2} \mathbf{K}_{g|H_0} + \frac{1}{2} \mathbf{K}_{g|H_1}$ is the mean covariance matrix of $\mathbf{g}$. It is easy to see, however, that the difference in means of two list-mode data sets is not a useful discriminant. Consider, for example, a simple photon-counting imaging system where the attributes are the 2D interaction positions on a detector plane. The mean position is of little use if both the signal-

absent and signal-present objects are symmetrically distributed around some origin; the mean of every attribute is zero under both hypotheses in that case.

We can, however, salvage the Hotelling observer and other concepts from linear-systems analysis by using the attribute list to construct a Poisson point process in attribute space [57, 63], defined as

$$u(\hat{\mathbf{A}}) \equiv \sum_{j=1}^J \delta(\hat{\mathbf{A}} - \hat{\mathbf{A}}_j).$$

The intuitive meaning of this point process can be seen by continuing with the example where the attributes are 2D positions, in which case we construct $u(\hat{\mathbf{A}})$ by placing a delta-function point at each of the J locations.

The point process is specified by $J + 1$ random variables: the J attribute vectors in $\hat{\mathbf{A}}$ and J itself. Because the terms in the sum over j are independent and identically distributed, the mean of $u(\hat{\mathbf{A}})$, conditional on $\mathbf{f}$, is given by

$$\bar{u}(\hat{\mathbf{A}}|\mathbf{f}) = \langle u(\hat{\mathbf{A}}|\mathbf{f}) \rangle_{\mathbf{g}|\mathbf{f}} = \sum_{J=0}^{\infty} \Pr(J|\mathbf{f}) J \int d\hat{\mathbf{A}}_j \Pr(\hat{\mathbf{A}}_j|\mathbf{f}) \delta(\hat{\mathbf{A}} - \hat{\mathbf{A}}_j) = \bar{J}(\mathbf{f}) \Pr(\hat{\mathbf{A}}|\mathbf{f}).$$

With (2) and (5), we now find

$$\bar{u}(\hat{\mathbf{A}}|\mathbf{f}) = \tau \int d\mathbf{r} \Pr(\hat{\mathbf{A}}|\mathbf{r}) f(\mathbf{r}) s(\mathbf{r}).$$

This equation defines a linear operator $\mathcal{L}$ such that

$$\bar{u}(\hat{\mathbf{A}}|\mathbf{f}) = \tau [\mathcal{L}\mathbf{f}](\hat{\mathbf{A}}),$$

or, more abstractly,

$$\bar{\mathbf{u}} = \tau \mathcal{L}\mathbf{f}.$$

This is the list-mode counterpart of the familiar relation [5] for binned data, $\bar{\mathbf{g}} = \mathcal{H}\mathbf{f}$. The key difference is that $\mathcal{H}$ is a mapping from an infinite-dimensional Hilbert space ($\mathbb{L}_2(\mathbb{R}^3)$ if $\mathbf{r}$ is a three-dimensional vector) to an M -dimensional Euclidean space, where M is the total number of bins in $\mathbf{g}$. The operator $\mathcal{L}$, on the other hand, maps $\mathbb{L}_2(\mathbb{R}^3)$ to $\mathbb{L}_2(\mathbb{R}^P)$, where P is the number of attributes per event [63].

To compute the Hotelling test statistic or the Hotelling SNR², we need not only the mean of $u(\hat{\mathbf{A}})$ but also its covariance function, which we denote $K_{u|\mathbf{f}}(\hat{\mathbf{A}}, \hat{\mathbf{A}}')$ for a fixed object or $K_{u|H_k}(\hat{\mathbf{A}}, \hat{\mathbf{A}}')$ when we consider an ensemble of objects appropriate to hypothesis H_k . The definitions are

$$\begin{aligned} K_{u|f}(\hat{A}, \hat{A}') &\equiv \left\langle \left[u(\hat{A}) - \bar{u}(\hat{A}|f) \right] \left[u(\hat{A}') - \bar{u}(\hat{A}'|f) \right] \right\rangle_{u|f}; \\ K_{u|H_k}(\hat{A}, \hat{A}') &\equiv \left\langle \left[u(\hat{A}) - \bar{u}_k(\hat{A}) \right] \left[u(\hat{A}') - \bar{u}_k(\hat{A}') \right] \right\rangle_{u|f|H_k}, \end{aligned}$$

where

$$\bar{u}_k(\hat{A}) \equiv \left\langle \bar{u}(\hat{A}|f) \right\rangle_{f|H_k} = \tau \int d\mathbf{r} \, \text{pr}(\hat{A}|\mathbf{r}) \bar{f}_k(\mathbf{r}) s(\mathbf{r}), \quad (23)$$

with $\bar{f}_k(\mathbf{r})$ being the mean of the object ensemble for H_k .

Methods for computing these covariance functions are detailed in Barrett and Myers [5] (see especially Secs. 8.5.3, 11.3.3 and 11.3.7). The results are

$$\begin{aligned} K_{u|f}(\hat{A}, \hat{A}') &= \bar{u}(\hat{A}|f) \delta(\hat{A} - \hat{A}'); \\ K_{u|H_k}(\hat{A}, \hat{A}') &= \bar{u}_k(\hat{A}) \delta(\hat{A} - \hat{A}') + \tau^2 [\mathcal{L} \mathcal{K}_{f|H_k} \mathcal{L}^\dagger](\hat{A}, \hat{A}'), \end{aligned} \quad (24)$$

where $\mathcal{L}^\dagger$ is the adjoint [5] of $\mathcal{L}$ and

$$[\mathcal{L} \mathcal{K}_{f|H_k} \mathcal{L}^\dagger](\hat{A}, \hat{A}') = \int d\mathbf{r} \int d\mathbf{r}' \, \text{pr}(\hat{A}|\mathbf{r}) s(\mathbf{r}) K_{f|H_k}(\mathbf{r}, \mathbf{r}') \text{pr}(\hat{A}'|\mathbf{r}') s(\mathbf{r}'), \quad (25)$$

with $K_{f|H_k}(\mathbf{r}, \mathbf{r}')$ being the covariance function (the kernel of the operator $\mathcal{K}_{f|H_k}$) of the random objects under H_k . Operator notations for (23) and (24) are

$$\bar{u}_k = \tau \mathcal{L} \bar{\mathbf{f}}_k, \quad \mathcal{K}_{u|H_k} = \tau (\mathcal{L} \bar{\mathbf{f}}_k) \mathcal{I}_{\hat{A}} + \tau^2 \mathcal{L} \mathcal{K}_{f|H_k} \mathcal{L}^\dagger,$$

where $\mathcal{I}_{\hat{A}}$ is the unit operator in attribute space, with kernel $\delta(\hat{A} - \hat{A}')$.

5.B. Hotelling Test Statistic and Figure of Merit

For a binary discrimination task, a linear observer computes a scalar product between some template $\mathbf{w}$ and the data to get a test statistic t , and then compares t to a threshold to make the decision. For list-mode data, the test statistic is a scalar product in attribute space, with the form

$$t(\mathbf{G}) = \mathbf{w}^\dagger \mathbf{u} \equiv \int d\hat{A} \, \mathbf{w}(\hat{A}) u(\hat{A}).$$

From (23), the mean of $t(\mathbf{G})$ under H_k is

$$\bar{t}_k = \mathbf{w}^\dagger \bar{\mathbf{u}}_k = \tau \int d\hat{A} \, \mathbf{w}(\hat{A}) \int d\mathbf{r} \, \text{pr}(\hat{A}|\mathbf{r}) \bar{f}_k(\mathbf{r}) s(\mathbf{r}).$$

The average of the variance of $t(\mathbf{G})$ under the two hypotheses is

$$\overline{\sigma_t^2} \equiv \frac{1}{2} \text{var}\{t(\mathbf{G})|H_1\} + \frac{1}{2} \text{var}\{t(\mathbf{G})|H_2\} = \int d\hat{\mathbf{A}} \int d\hat{\mathbf{A}}' w(\hat{\mathbf{A}}) \overline{\mathbf{K}}_u(\hat{\mathbf{A}}, \hat{\mathbf{A}}') w(\hat{\mathbf{A}}') = \mathbf{w}^\dagger \overline{\mathcal{K}}_u \mathbf{w},$$

where

$$\overline{\mathbf{K}}_u(\hat{\mathbf{A}}, \hat{\mathbf{A}}') = \frac{1}{2} K_{u|H_0}(\hat{\mathbf{A}}, \hat{\mathbf{A}}') + \frac{1}{2} K_{u|H_1}(\hat{\mathbf{A}}, \hat{\mathbf{A}}'),$$

or, in operator form,

$$\overline{\mathcal{K}}_u = \frac{1}{2} \mathcal{K}_{u|H_1} + \frac{1}{2} \mathcal{K}_{u|H_2}.$$

With these results, the figure of merit for this general linear discriminant in attribute space is [3]

$$\text{SNR}_w^2 = \frac{(\mathbf{w}^\dagger \Delta \overline{\mathbf{u}})^2}{\mathbf{w}^\dagger \overline{\mathcal{K}}_u \mathbf{w}},$$

where

$$\Delta \overline{\mathbf{u}} = \overline{\mathbf{u}}_1 - \overline{\mathbf{u}}_0 = \tau(\mathcal{L} \Delta \mathbf{f}).$$

It can be shown (Barrett and Myers [5], Sec. 13.2.12) that this figure of merit is maximized when

$$\overline{\mathcal{K}}_u \mathbf{w} = \Delta \overline{\mathbf{u}} \quad \text{or} \quad \mathbf{w} = \overline{\mathcal{K}}_u^{-1} \Delta \overline{\mathbf{u}}.$$

This choice of $\mathbf{w}$ defines the list-mode Hotelling observer, first given by Lehovich [63], for which the figure of merit is

$$\text{SNR}_{\text{Hot}}^2 = \Delta \overline{\mathbf{u}}^\dagger \overline{\mathcal{K}}_u^{-1} \Delta \overline{\mathbf{u}}, \quad (26)$$

Explicitly, from (23), (24), and (25), the operator $\overline{\mathcal{K}}_u$ has the kernel given by

$$\overline{K}_u(\hat{\mathbf{A}}, \hat{\mathbf{A}}') = \left[\tau \int d\mathbf{r} \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) \overline{\bar{f}}(\mathbf{r}) s(\mathbf{r}) \right] \delta(\hat{\mathbf{A}} - \hat{\mathbf{A}}') + \tau^2 \int d\mathbf{r} \int d\mathbf{r}' \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) s(\mathbf{r}) \overline{K}_f(\mathbf{r}, \mathbf{r}') \text{pr}(\hat{\mathbf{A}}'|\mathbf{r}') s(\mathbf{r}'),$$

where $\overline{\bar{f}}(\mathbf{r}) = \frac{1}{2} [\bar{f}_0(\mathbf{r}) + \bar{f}_1(\mathbf{r})]$ and, similarly, $\overline{K}_f(\mathbf{r}, \mathbf{r}') = \frac{1}{2} [K_{f|H_0}(\mathbf{r}, \mathbf{r}') + K_{f|H_1}(\mathbf{r}, \mathbf{r}')]$. Because covariance operators are positive-semidefinite, the “diagonal” term on the first line of (27) guarantees that the inverse in (26) exists, provided only that the integral in square brackets in (27) does not vanish for any $\hat{\mathbf{A}}$.

Various methods for obtaining or approximating $\text{SNR}_{\text{Hot}}^2$ without actually performing the operator inverse are discussed in Barrett and Myers [5], Sec. 14.3.2; the results there are specifically for covariance matrices, but they all generalize readily to operators. Some insights might be gained by manipulating the expression for $\overline{\mathcal{K}}_u^{-1}$. For example,

$$\overline{\mathcal{K}}_u^{-1} = \{\tau \overline{\mathcal{B}} + \tau^2 \mathcal{L} \overline{\mathcal{K}}_f \mathcal{L}^\dagger\}^{-1} = \{\tau [\overline{\mathcal{B}} + \tau \mathcal{L} \overline{\mathcal{K}}_f \mathcal{L}^\dagger]\}^{-1} = \left\{ \tau \overline{\mathcal{B}} \left[\mathcal{I}_{\mathbf{A}} + \tau \overline{\mathcal{B}}^{-1} \mathcal{L} \overline{\mathcal{K}}_f \mathcal{L}^\dagger \right] \right\}^{-1} = \frac{1}{\tau} \left[\mathcal{I}_{\mathbf{A}} + \tau \overline{\mathcal{B}}^{-1} \mathcal{L} \overline{\mathcal{K}}_f \mathcal{L}^\dagger \right]^{-1} \overline{\mathcal{B}}^{-1},$$

where we introduced

$$\overline{\mathcal{B}} = (\mathcal{L} \overline{\mathbf{f}}) \mathcal{I}_{\mathbf{A}}, \quad \text{and} \quad \overline{\mathcal{K}}_f = \frac{1}{2} \mathcal{K}_{f|H_0} + \frac{1}{2} \mathcal{K}_{f|H_1}.$$

Notice that the operator $\overline{\mathcal{B}}$ is diagonal; so its inverse is easy to calculate.

6. Example

The model system for the example presented here is a planar (2D) optical source imaged with a lens onto a photon-counting imaging detector such as a multianode position-sensitive photomultiplier (MAPMT) [64]. The same mathematics applies to a planar gamma-ray source imaged with a pinhole or parallel-hole collimator onto a scintillation camera [5] which uses a photomultiplier array to sense the light from a scintillation event produced by a gamma ray. In both cases, we can think of the j^{th} photon (optical or gamma-ray) as emerging from the object at a point $\mathbf{r}_j = (x_j, y_j)$ and impinging on the detector at $\mathbf{R}_j = (X_j, Y_j)$. Then the signals from the PMTs are used to estimate the position $\mathbf{R}_j$, with the estimate being denoted by $\hat{\mathbf{R}}_j = (\hat{X}_j, \hat{Y}_j)$. Thus the true attribute of the j^{th} detected photon is $\mathbf{A}_j = \mathbf{R}_j$, and the list-mode data set is

$$\mathcal{A} = \{\hat{\mathbf{R}}_j, j=1, \dots, J\}.$$

We assume that the magnification of the imaging system is unity and that the sensitivity $s(\mathbf{r})$ is a constant s_0 , independent of position but proportional to the square of the numerical aperture of the optics (or the square of the diameter of the pinhole in the gamma-ray case). The optics is assumed to be linear and shift invariant; and its point spread function (PSF) is denoted by $p_{\text{opt}}(\mathbf{R})$, which is normalized such that $\int d\mathbf{R} p_{\text{opt}}(\mathbf{R}) = 1$. Thus

$$\text{pr}(\mathbf{R}|\mathbf{r}) = p_{\text{opt}}(\mathbf{R} - \mathbf{r}).$$

The detector is also assumed to be linear and shift invariant, and its PSF is denoted by $p_{\text{det}}(\mathbf{R})$, so that

$$\text{pr}(\hat{\mathbf{R}}|\mathbf{R}) = p_{\text{det}}(\hat{\mathbf{R}} - \mathbf{R}).$$

Thus both PSFs, normally interpreted deterministically, play the role of conditional PDFs on photon position in this case [5]. In this paper we assume, for simplicity, that both PSFs are Gaussian:

$$p_{\text{opt}}(\mathbf{R} - \mathbf{r}) = \frac{1}{2\pi\sigma_{\text{opt}}^2} \exp \left[-\frac{|\mathbf{R} - \mathbf{r}|^2}{2\sigma_{\text{opt}}^2} \right],$$

$$p_{\text{det}}(\hat{\mathbf{R}} - \mathbf{R}) = \frac{1}{2\pi\sigma_{\text{det}}^2} \exp \left[-\frac{|\hat{\mathbf{R}} - \mathbf{R}|^2}{2\sigma_{\text{det}}^2} \right].$$

As in previous sections, the object being imaged is divided into signal and background components. The signal, when present, is taken to be Gaussian:

$$\Delta f(\mathbf{r}) = \frac{\Delta f_0}{2\pi r_s^2} \exp \left[-\frac{|\mathbf{r}|^2}{2r_s^2} \right],$$

where $|\mathbf{r}| = \sqrt{x^2 + y^2}$ and r_s is a positive number characterizing the width of the signal.

To specify a random background we use the stationary lumpy-background model [65], for which

$$f(\mathbf{r}) = f_0 \sum_{k=1}^K \ell(\mathbf{r} - \mathbf{r}_k),$$

where the lump function is also assumed to be a Gaussian

$$\ell(\mathbf{r} - \mathbf{r}_k) = \frac{1}{2\pi r_b^2} \exp \left[-\frac{|\mathbf{r} - \mathbf{r}_k|^2}{2r_b^2} \right].$$

Here the number of lumps, K , is a Poisson random variable with mean $\bar{K}$ and the lump centers $\mathbf{r}_k$ are random 2D vectors uniformly distributed over a square area of size $w \times w$, and the detector is also $w \times w$. With these assumptions, $f(\mathbf{r})$ is a filtered, stationary, Poisson random process. For further details, see Rolland [65] or Barrett and Myers [5]. Because the signal is known, the covariance $K_{f|H_k}(\mathbf{r}, \mathbf{r}')$ is the same under the two hypotheses and given by [65]

$$K_{f|H_0}(\mathbf{r}, \mathbf{r}') = K_{f|H_1}(\mathbf{r}, \mathbf{r}') = \bar{K} f(\mathbf{r}, \mathbf{r}') = \frac{p f_0^2}{4\pi r_b^2} \exp \left[-\frac{|\mathbf{r} - \mathbf{r}'|^2}{4r_b^2} \right].$$

in which we defined $p = \bar{K}/w^2$. We will assume $r_b \gg w$ and $r_s \gg w$, so that the density $\text{pr}(\hat{\mathbf{R}}|\mathbf{r})$ can be approximated with the convolution $(p_{\text{opt}} * p_{\text{det}})(\hat{\mathbf{R}} - \mathbf{r})$ between p_{opt} and p_{det} evaluated at $\hat{\mathbf{R}} - \mathbf{r}$

$$\text{pr}(\hat{\mathbf{R}}|\mathbf{r}) = \frac{1}{2\pi\sigma^2} \exp \left[-\frac{|\hat{\mathbf{R}} - \mathbf{r}|^2}{2\sigma^2} \right],$$

where $\sigma^2 = \sigma_{\text{opt}}^2 + \sigma_{\text{det}}^2$. With this result, the quantity $[\mathcal{E}_f \mathcal{E}_f^\dagger](\hat{\mathbf{R}}, \hat{\mathbf{R}}')$ assumes the form

$$[\mathcal{L}\overline{\mathcal{H}}_{\mathbf{f}}\mathcal{L}^{\dagger}](\hat{\mathbf{R}},\hat{\mathbf{R}}')=\frac{pf_0^2}{4\pi(r_b^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}-\hat{\mathbf{R}}'|^2}{4(r_b^2+\sigma^2)}\right].$$

In a similar way,

$$\begin{aligned} [\mathcal{L}\overline{\mathbf{f}}_0](\hat{\mathbf{R}}) &= pf_0, \\ [\mathcal{L}\overline{\mathbf{f}}_1](\hat{\mathbf{R}}) &= pf_0 + \frac{\Delta f_0}{2\pi(r_s^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}|^2}{2(r_s^2+\sigma^2)}\right]. \end{aligned}$$

For the example considered here, the kernel of the operator $\overline{\mathbf{u}}$ is

$$\overline{K}_u(\hat{\mathbf{R}},\hat{\mathbf{R}}')=\tau pf_0\left\{1+\frac{\Delta f_0}{4\pi pf_0(r_s^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}|^2}{2(r_s^2+\sigma^2)}\right]\right\}\delta(\hat{\mathbf{R}}-\hat{\mathbf{R}}')+\tau^2\frac{pf_0^2}{4\pi(r_b^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}-\hat{\mathbf{R}}'|^2}{4(r_b^2+\sigma^2)}\right],$$

and if $\Delta f_0 \ll pf_0(r_s^2+\sigma^2)$, we can also write

$$\overline{K}_u(\hat{\mathbf{R}},\hat{\mathbf{R}}')\approx\tau pf_0\left\{\delta(\hat{\mathbf{R}}-\hat{\mathbf{R}}')+\frac{\tau f_0}{4\pi(r_s^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}-\hat{\mathbf{R}}'|^2}{4(r_b^2+\sigma^2)}\right]\right\}.$$

By the same token, $\overline{\mathcal{B}}^{-1}\approx(pf_0)^{-1}\mathfrak{Q}_{\hat{\mathbf{R}}}$, so that

$$[\overline{\mathcal{B}}^{-1}\mathcal{L}\overline{\mathcal{H}}_{\mathbf{f}}\mathcal{L}^{\dagger}](\hat{\mathbf{R}},\hat{\mathbf{R}}')\approx\frac{f_0}{4\pi(r_b^2+\sigma^2)}\exp\left[-\frac{|\hat{\mathbf{R}}-\hat{\mathbf{R}}'|^2}{4(r_b^2+\sigma^2)}\right].$$

Notice that $\overline{\mathcal{B}}^{-1}\mathfrak{L}_{\mathbf{f}}\mathfrak{L}^{\dagger}$ is shift invariant, this allows us to write

$$\overline{\mathcal{B}}^{-1}\mathcal{L}\overline{\mathcal{H}}_{\mathbf{f}}\mathcal{L}^{\dagger}=\mathcal{U}\mathcal{D}\mathcal{U}^{\dagger}, \quad (29)$$

provided that

$$\begin{aligned} [\mathcal{U}](\hat{\mathbf{R}},\rho) &= e^{i2\pi\hat{\mathbf{R}}\rho}, \\ [\mathcal{D}](\rho,\rho') &= f_0\delta(\rho-\rho')e^{-4\pi^2(r_b^2+\sigma^2)|\rho|^2}. \end{aligned}$$

The form above can be verified by substituting and performing the integrals. Then from (28)

$$\left\{\mathcal{I}_{\hat{\mathbf{R}}}+\tau\overline{\mathcal{B}}^{-1}\mathcal{L}\overline{\mathcal{H}}_{\mathbf{f}}\mathcal{L}^{\dagger}\right\}^{-1}=\{\mathcal{I}_{\hat{\mathbf{R}}}+\tau\mathcal{U}\mathcal{D}\mathcal{U}^{\dagger}\}^{-1}=\{\mathcal{U}\mathcal{U}^{\dagger}+\tau\mathcal{U}\mathcal{D}\mathcal{U}^{\dagger}\}^{-1}=\{\mathcal{U}[\mathcal{I}_{\rho}+\tau\mathcal{D}]\mathcal{U}^{\dagger}\}^{-1}=\mathcal{U}[\mathcal{I}_{\rho}+\tau\mathcal{D}]^{-1}\mathcal{U}^{\dagger},$$

in which we have used the fact that $\mathcal{U}$ is unitary ($\mathcal{U}^{\dagger}\mathcal{U}=\mathcal{U}\mathcal{U}^{\dagger}=\mathfrak{Q}$). Notice that $\mathfrak{Q}_{\rho}+\tau\mathcal{D}$ is diagonal, thus its inverse is readily available. Using (26) and (28) we can write

$$\text{SNR}_{\text{Hot}}^2 \approx \tau(\mathcal{L}\Delta\mathbf{f})^\dagger \mathcal{U} \{\mathcal{I}_\rho + \tau\mathcal{D}\}^{-1} \mathcal{U}^\dagger \overline{\mathcal{B}}^{-1} (\mathcal{L}\Delta\mathbf{f}),$$

where

$$\begin{aligned} ((\mathcal{L}\Delta\mathbf{f})^\dagger \mathcal{U})(\rho) &= \Delta f_0 e^{-2\pi^2(r_s^2 + \sigma^2)|\rho|^2}, \\ (\mathcal{U}^\dagger \overline{\mathcal{B}}^{-1} \mathcal{L}\Delta\mathbf{f})(\rho) &= \frac{\Delta f_0}{pf_0} e^{-2\pi^2(r_s^2 + \sigma^2)|\rho|^2}. \end{aligned}$$

If we carry out the calculation, we find

$$\text{SNR}_{\text{Hot}}^2 \approx \frac{(\Delta f_0)^2}{pf_0^2} \int_{\mathbb{R}^2} d^2\rho \frac{f_0 \tau e^{-4\pi^2(r_s^2 + \sigma^2)|\rho|^2}}{1 + f_0 \tau e^{-4\pi^2(r_b^2 + \sigma^2)|\rho|^2}}. \quad (30)$$

To simplify the notation, let us introduce $\alpha = f_0\tau$, $\beta = (\Delta f_0)^2/(pf_0)$, $a = 4\pi^2(r_s^2 + \sigma^2)$, $b = 4\pi^2(r_b^2 + \sigma^2)$, and $\gamma = a/b$. With a change of variables [66, 67], we obtain

$$\begin{aligned} \text{SNR}_{\text{Hot}}^2 &\approx 2\pi\beta \int_0^\infty d\rho \frac{\alpha \rho e^{-a\rho^2}}{1 + \alpha e^{-b\rho^2}} \\ &= \frac{\pi\beta}{b} \int_0^1 dy \frac{y^{\gamma-1}}{1/\alpha + y} = \frac{\pi\beta}{a} {}_2F_1(I, \gamma; I + \gamma - \alpha) = \frac{\pi\beta}{a} \frac{\alpha}{1 + \alpha} {}_2F_1\left(I, I; I + \gamma; \frac{\alpha}{1 + \alpha}\right), \end{aligned} \quad (31)$$

where the function ${}_2F_1(a, b; c; z)$ denotes the hypergeometric series [66]:

$${}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$$

where $(x)_n = x(x+1) \dots (x+n-1)$.

We are now interested in studying the behavior of $\text{SNR}_{\text{Hot}}^2$ as the exposure time τ increases. We will consider three different cases: $r_s < r_b$, $r_s = r_b$, and $r_s > r_b$ and we will use

$\text{SNR}_{\text{Hot}}^2(\tau, r_s < r_b)$, $\text{SNR}_{\text{Hot}}^2(\tau, r_s = r_b)$, and $\text{SNR}_{\text{Hot}}^2(\tau, r_s > r_b)$ to denote the values of SNR^2 for each case.

The $r_s = r_b$ case is simple to handle. Calculation of the integral yields

$$\text{SNR}_{\text{Hot}}^2(\tau, r_s = r_b) = \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 + \sigma^2)} \log(1 + f_0\tau).$$

As expected, $\text{SNR}_{\text{Hot}}^2(\tau, r_s = r_b)$ increases as the exposure time τ increases. The same is true for the $r_s < r_b$ cases. Indeed, by expanding the series in (31) and noticing that $\alpha/(1 + \alpha)$ is positive and monotonic increasing with $\alpha = f_0\tau$, we see that $\text{SNR}_{\text{Hot}}^2(\tau, r_s \neq r_b)$ must increase as τ increases.

For the $r_s < r_b$ cases, consider first

$$\text{SNR}_{\text{Hot}}^2 = \frac{\pi\beta}{a} \sum_{n=0}^{\infty} \frac{(1)_n}{(1+\gamma)_n} \left(\frac{\alpha}{1+\alpha} \right)^{n+1},$$

where

$$\frac{(1)_n}{(1+\gamma)_n} = \frac{1 \times 2 \times \cdots \times n}{(1+\gamma) \times (2+\gamma) \times \cdots \times (n+\gamma)}.$$

If $r_s < r_b$, then $\gamma < 1$ and $(1)_n/(1+\gamma)_n > 1/(n+1)$ provided that $n \geq 0$. Therefore, after substitutions

$$\text{SNR}_{\text{Hot}}^2(\tau, r_s < r_b) > \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 + \sigma^2)} \sum_{n=1}^{\infty} \frac{1}{n+1} \left(\frac{\alpha}{1+\alpha} \right)^{n+1} = \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 + \sigma^2)} \log(1 + f_0\tau).$$

The result we just obtained shows that, for the $r_s < r_b$ case and as τ increases, $\text{SNR}_{\text{Hot}}^2$ grows faster than it grew in the $r_s = r_b$ case. With a Taylor expansion for $1/\tau \approx 0$, we can make this concept more precise

$$\text{SNR}_{\text{Hot}}^2(\tau, r_s < r_b) \approx \frac{(\Delta f_0)^2}{pf_0^2} \text{CSC} \left(\pi \frac{r_s^2 + \sigma^2}{r_b^2 + \sigma^2} \right) \frac{1}{4(r_b^2 + \sigma^2)} (f_0\tau)^{1 - \frac{r_s^2 + \sigma^2}{r_b^2 + \sigma^2}}, \quad \tau \gg 1.$$

On the other hand, if $r_s > r_b$

$$\text{SNR}_{\text{Hot}}^2(\tau, r_s > r_b) < \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 + \sigma^2)} \sum_{n=1}^{\infty} \frac{1}{n+1} \left(\frac{\alpha}{1+\alpha} \right)^{n+1} = \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 + \sigma^2)} \log(1 + f_0\tau),$$

and

$$\lim_{\tau \rightarrow \infty} \text{SNR}_{\text{Hot}}^2(\tau, r_s > r_b) = \lim_{\tau \rightarrow \infty} \frac{\pi\beta}{a} \frac{\alpha}{1+\alpha} {}_2F_1 \left(1, 1; 1+\gamma; \frac{\alpha}{1+\alpha} \right) = \frac{\pi\beta}{a} \left[\frac{\Gamma(\gamma+1)}{\Gamma(\gamma)} \right] \left[\frac{\Gamma(\gamma)}{\Gamma(\gamma-1)} \right]^{-1} = \frac{\pi\beta}{a} \frac{\gamma}{\gamma-1},$$

in which $\Gamma(\gamma) = \int_0^{\infty} dt \, t^{\gamma-1} e^{-t}$ and we have used $\Gamma(\gamma+1) = \gamma\Gamma(\gamma)$ [68]. Upon substitution

$$\text{SNR}_{\text{Hot}}^2(\tau, r_s > r_b) \approx \frac{(\Delta f_0)^2}{pf_0^2} \frac{1}{4\pi(r_s^2 - r_b^2)}, \quad \tau \gg 1.$$

Fig. 1 reports plots of $\text{SNR}_{\text{Hot}}^2$ for three different cases as functions of total mean count number $\bar{J}$, which is related to the exposure time as $\bar{J} = \tau pf_0 w^2$, where w is the detector side length. For comparison, Fig. 1 also includes plots of SNR^2 for the ideal observer calculated via MCMC methods, which we briefly alluded to at the end of Sec. 4.B. The ideal observer's SNR^2 is asymptotically linear with the total mean count number. This asymptotic behavior confirms the result of (20), which was made possible by the assumption in (17). On the other hand, for

low counts, the approximation in (17) breaks down and, as shown in Fig. 1, the ideal observer's SNR^2 is no longer proportional to the total mean count number (or, equivalently, to the exposure time).

The results derived above show that the $\text{SNR}_{\text{Hot}}^2$ keeps growing as the exposure time increases only for the $r_s = r_b$ case. Otherwise, $\text{SNR}_{\text{Hot}}^2$ is upper limited. Furthermore, $\text{SNR}_{\text{Hot}}^2$ grows as the signal intensity grows, and it decreases if the background intensity and/or the lump density p increases. The peculiar behavior we have in the $r_s > r_b$ case can be intuitively explained by noting that the signal has a Gaussian shape, so do the lumps in the background. If the signal is wider than the lumps, it is possible to have a false positive when a few lumps cluster at the signal location in such a way that they actually look like the signal we are looking for. On the other hand, when the signal is narrower than the lumps ($r_s < r_b$), false positives are reduced because no lump arrangement will look like the signal we are to detect.

The $r_s = r_b$ case deserves a different discussion. One could argue that because $r_s = r_b$, a background lump placed at $\mathbf{r} = (0, 0)$ could mistakenly be interpreted as the signal Δf we want to detect, which in our case is indeed centered at $(0, 0)$ and it is a scaled version of the lump function $\ell(\mathbf{r})$. Thus, the $\text{SNR}_{\text{Hot}}^2$ should be bounded when $r_s = r_b$. However, if we recall that the lump centers $\mathbf{r}_1, \dots, \mathbf{r}_K$ are 2D points uniformly distributed over a square of size $w \times w$, we immediately realize that the probability of having a lump centered at $(0, 0)$ is zero. Thus, with probability 1, a lump centered at $(0, 0)$ must be understood as the signal Δf .

7. Summary and Conclusions

In this paper, we have considered the problem of signal detection with list-mode data. Two observers have been considered: the ideal observer and the optimal linear observer, known in the literature as the Hotelling observer. Three distinct processes are involved in the development of these observers: randomness in the object being imaged, randomness in the attribute vectors, and, finally, randomness in their estimates from noisy detector outputs. The detection problems we have considered are detection of a known signal in a known background and detection of a known signal in a random background. Associated figures of merit have been calculated as well. In particular, the ideal observer's SNR^2 is found to be asymptotically proportional to the exposure time. The optimal linear observer is applied to list-mode data in an example at the end of the paper. A followup paper on the Markov chain Monte Carlo calculation of SNR^2 for the ideal observer is currently under preparation.

Acknowledgments

The authors would like to acknowledge Eric Clarkson and Matthew Kupinski for stimulating discussions and NIH grants R37 EB000803 and P41 EB002035 for financial support.

References

1. Barrett HH. Objective assessment of image quality: Effects of quantum noise and object variability. *J. Opt. Soc. Am. A.* 1990; 7:1266–1278. [PubMed: 2370589]
2. Barrett HH, Denny JL, Wagner RF, Myers KJ. Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. *J. Opt. Soc. Am. A.* 1995; 12:834–852.
3. Barrett HH, Abbey CK, Clarkson E. Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions. *J. Opt. Soc. Am. A.* 1998; 15:1520–1535.
4. Barrett HH, Myers KJ, Devaney N, Dainty C. Objective assessment of image quality. IV. Application to adaptive optics. *J. Opt. Soc. Am. A.* 2006; 23:3080–3105.

5. Barrett, HH.; Myers, KJ. Foundations of Image Science. Hoboken, NJ: Wiley-Interscience; 2004.
6. Myers KJ, Barrett HH, Borgstrom MC, Patton DD, Seeley GW. Effect of noise correlation on detectability of disk signals in medical imaging. *J. Opt. Soc. Am. A.* 1985; 2:1752–1759. [PubMed: 4056949]
7. Smith WE, Barrett HH. Hotelling trace criterion as a figure of merit for the optimization of imaging systems. *J. Opt. Soc. Am. A.* 1986; 3:717–725.
8. Fiete RD, Barrett HH, Smith WE, Myers KJ. Hotelling trace criterion and its correlation with human-observer performance. *J. Opt. Soc. Am. A.* 1987; 4:945–953. [PubMed: 3598746]
9. Hanson KM. Method of evaluating image-recovery algorithms based on task performance. *J. Opt. Soc. Am. A.* 1990; 7:1294–1304.
10. Burgess AE. Statistically defined backgrounds: performance of a modified nonprewhitening observer model. *J. Opt. Soc. Am. A.* 1994; 11:1237–1242.
11. Burgess AE, Li X, Abbey CK. Visual signal detectability with two noise components: anomalous masking effects. *J. Opt. Soc. Am. A.* 1997; 14:2420–2442.
12. Bochud FO, Abbey CK, Eckstein MP. Visual signal detection in structured backgrounds. III. Calculation of figures of merit for model observers in statistically non-stationary backgrounds. *J. Opt. Soc. Am. A.* 2000; 17:193–205.
13. Abbey CK, Eckstein MP. Derivation of a detectability index for correlated responses in multiple-alternative forced-choice experiments. *J. Opt. Soc. Am. A.* 2000; 17:2101–2104.
14. Eckstein MP, Abbey CK, Bochud FO. Visual signal detection in structured backgrounds. IV. Figures of merit for model performance in multiple-alternative forced-choice detection tasks with correlated responses. *J. Opt. Soc. Am. A.* 2000; 17:206–217.
15. Kupinski MA, Hoppin JW, Clarkson E, Barrett HH. Ideal-observer computation in medical imaging with use of Markov-chain Monte Carlo techniques. *J. Opt. Soc. Am. A.* 2003; 20:430–438.
16. Chen L, Barrett HH. Task-based lens design with application to digital mammography. *J. Opt. Soc. Am. A.* 2005; 22:148–167.
17. Rolland J, O'Daniel J, Akcay C, DeLemos T, Lee K-I, Cheong KS, Clarkson E, Chakrabarti R, Ferris R. Task-based optimization and performance assessment in optical coherence imaging. *J. Opt. Soc. Am. A.* 2005; 22:1132–1142.
18. Caucci L, Barrett HH, Devaney N, Rodríguez JJ. Application of the Hotelling and ideal observers to detection and localization of exoplanets. *J. Opt. Soc. Am. A.* 2007; 24:B13–B24.
19. Burgess AE, Judy PF. Signal detection in power-law noise: effect of spectrum exponents. *J. Opt. Soc. Am. A.* 2007; 24:B52–B60.
20. Abbey CK, Eckstein MP. Classification images for simple detection and discrimination tasks in correlated noise. *J. Opt. Soc. Am. A.* 2007; 24:B110–B124.
21. Dixon TD, Canga EF, Nikolov SG, Troscianko T, Noyes JM, Canagarajah CN, Bull DR. Selection of image fusion quality measures: objective, subjective, and metric assessment. *J. Opt. Soc. Am. A.* 2007; 24:B125–B135.
22. Platiša L, Goossens B, Vansteenkiste E, Park S, Gallas BD, Badano A, Philips W. Channelized Hotelling observers for the assessment of volumetric imaging data sets. *J. Opt. Soc. Am. A.* 2011; 28:1145–1163.
23. Wagner RF, Brown DG. Unified SNR analysis of medical imaging systems. *Phys. Med. Biol.* 1985; 30:489–518.
24. Frey EC, Gilland KL, Tsui BMW. Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE Trans. Med. Imag.* 2002; 21:1040–1050.
25. He X, Frey EC, Links JM, Gilland KL, Segars WP, Tsui BMW. A mathematical observer study for the evaluation and optimization of compensation methods for myocardial SPECT using a phantom population that realistically models patient variability. *IEEE Trans. Nucl. Sci.* 2004; 51:218–224.
26. Gang GJ, Lee J, Prince JL, Stayman JW, Tward DJ, Zbijewski W, Siewerdsen JH. Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance. *Med. Phys.* 2011; 38:1754–1768. [PubMed: 21626910]
27. Hotelling H. The generalization of Student's ratio. *Ann. of Math. Stat.* 1931; 2:360–378.

28. Myers KJ, Barrett HH. Addition of a channel mechanism to the ideal-observer model. *J. Opt. Soc. Am. A*. 1987; 4:2447–2457. [PubMed: 3430229]
29. Papaliolios, C.; Mertz, L. New two-dimensional photon camera. In: Crawford, DL., editor. *Instrumentation in Astronomy IV*. Vol. vol. 331. Proc. SPIE; 1982. p. 360-369.vol. 331
30. Papaliolios C, Nisenson P, Ebstein S. Speckle imaging with the PAPA detector. *Appl. Optics*. 1985; 24:287–292.
31. Timothy, JG.; Morgan, JS.; Slater, DC.; Kasle, DB.; Bybee, RL. MAMA detector systems: A status report. In: Huffman, RE., editor. *Ultraviolet technology III*. Vol. vol. 1158. Proc. SPIE; 1989. p. 104-117.vol. 1158
32. Timothy JG. Optical detectors for spectroscopy. *Publications of the Astronomical Society of the Pacific*. 1983; 95:810–834.
33. Bruyndonckx P, Lemaître C, Schaart D, Maas M, (jan) van der Laan DJ, Krieguer M, Devroede O, Tavernier S. Towards a continuous crystal APD-based PET detector design. *Nucl. Instrum. and Meth. A*. 2007; 571:182–186.
34. Bostroem, KA.; Aloisi, A.; Bohlin, R.; Diaz, R.; Dixon, V.; Goudfrooij, P.; Hodge, P.; Lennon, D.; Long, C.; Niemi, S.; Osten, R.; Proffitt, C.; Walborn, N.; Wheeler, T.; Wolfe, M.; York, B.; Zheng, W. *STIS Instrument Handbook, Version 10.0*. Baltimore, MD: Space Telescope Science Institute; 2010.
35. Anger HO. Scintillation camera. *Rev. Sci. Instrum*. 1958; 29:27–33.
36. Wernick, MN.; Aarsvold, JN. *Emission Tomography: The Fundamentals of PET and SPECT*. San Diego, CA: Elsevier Academic Press; 2004.
37. Johns, PC.; Dubeau, J.; Gobbi, DG.; Li, M.; Dixit, MS. Opto-Canada: SPIE Regional Meeting on Optoelectronics, Photonics, and Imaging. Vol. vol. TD01. Proc. SPIE; 2002. Photon-counting detectors for digital radiography and X-ray computed tomography; p. 367-369.vol. TD01
38. Shikhaliev PM, Xu T, Molloy S. Photon counting computed tomography: Concept and initial results. *Med. Phys.* 2005; 32:427–436. [PubMed: 15789589]
39. Hunter WCJ, Barrett HH, Furenlid LR. Calibration method for ML estimation of 3D interaction position in a thick gamma-ray detector. *IEEE Trans. Nucl. Sci.* 2009; 56:189–196. [PubMed: 20191099]
40. Hesterman JY, Caucci L, Kupinski MA, Barrett HH, Furenlid LR. Maximum-likelihood estimation with a contracting-grid search algorithm. *IEEE Trans. Nucl. Sci.* 2010; 57:1077–1084. [PubMed: 20824155]
41. Takahashi, T.; Nakazawa, K.; Kamae, T.; Tajima, H.; Fukazawa, Y.; Nomachi, M.; Kokubun, M. High resolution CdTe detectors for the next-generation multi-Compton gamma-ray telescope. In: Truemper, JE.; Tananbaum, HD., editors. *X-Ray and Gamma-Ray Telescopes and Instruments for Astronomy*. Vol. vol. 4851. Proc. SPIE; 2002. p. 1228-1235.vol. 4851
42. Bellm EC, Boggs SE, Bandstra MS, Bowen JD, Perez-Becker D, Wunderer CB, Zoglauer A, Amman M, Luke H-K, Chang J-L, Chiu J-S, Liang Y-H, Chang Z-K, Liu W-C, Hung C-H, Lin PN, Huang MA, Jean P. Overview of the nuclear Compton telescope. *IEEE Trans. Nucl. Sci.* 2009; 56:1250–1256.
43. Zoglauer A, Andritschke R, Kanbach G. Data analysis for the MEGA prototype. *New Astron. Rev.* 2004; 48:231–235.
44. Kitaeva GK, Penin AN. Spontaneous parametric down-conversion. *JETP Lett.* 2005; 82:330–335.
45. Hayat A, Ginzburg P, Orenstein M. Observation of two-photon emission from semiconductors. *Nature Photon.* 2008; 2:238–241.
46. Nasr MB, Minaeva O, Goltsman GN, Sergienko AV, A. Saleh BE, Teich MC. Submicron axial resolution in an ultrabroadband two-photon interferometer using superconducting single-photon detect. *Opt. Express*. 2008; 16:15104–15108. [PubMed: 18795048]
47. Abouraddy AF, Stone PR, Sergienko AV, Saleh BEA, Teich MC. Entangled-photon imaging of a pure phase object. *Phys. Rev. Lett.* 2004; 93 213903.
48. Reader AJ, Ally S, Bakatselos F, Manavaki R, Walledge RJ, Jeavons AP, Julyan PJ, Zhao S, Hastings DL, Zweit J. One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays. *IEEE Trans. Nucl. Sci.* 2002; 49:693–699.

49. Khurd I-T, Hsiao P, Rangarajan A, Gindi G. A globally convergent regularized ordered-subset EM algorithm for list-mode reconstruction. *IEEE Trans. Nucl. Sci.* 2004; 51:719–725.
50. Reader AJ, Erlandsson K, Ott RJ, Flower MA. Attenuation and scatter correction of list-mode data driven iterative and analytic image reconstruction algorithms for rotating 3D PET systems. *IEEE Trans. Nucl. Sci.* 1999; 46:2218–2226.
51. Parra L, Barrett HH. List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET. *IEEE Trans. Med. Imag.* 1998; 17:228–235.
52. Snyder DL, Polite DG. Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements. *IEEE Trans. Nucl. Sci.* 1983; 30:1843–1849.
53. Reader AJ, Erlandsson K, Flower MA, Ott RJ. Fast accurate iterative 2D reconstruction for low-statistics positron volume imaging. *Phys. Med. Biol.* 1998; 43:835–846. [PubMed: 9572508]
54. Byrne C. Likelihood maximization for list-mode emission tomographic image reconstruction. *IEEE Trans. Med. Imag.* 2001; 20:1084–1092.
55. Huesman RH, Klein GJ, Moses WW, Qi J, Reutter BW, Virador PRG. List-mode maximum-likelihood reconstruction applied to positron emission mammography (PEM) with irregular sampling. *IEEE Trans. Med. Imag.* 2000; 19:532–537.
56. Levkovitz R, Falikman D, Zibulevsky M, Ben-Tal A, Nemirovski A. The design and implementation of COSEN, an iterative algorithm for fully 3-D listmode data. *IEEE Trans. Med. Imag.* 2001; 20:633–642.
57. Barrett HH, White T, Parra LC. List-mode likelihood. *J. Opt. Soc. Am. A.* 1997; 14:2914–2923.
58. Surti, S.; Karp, JS.; Popescu, LM.; Daube-Witherspoon, ME.; Werner, M. Investigation of image quality and NEC in a TOF-capable PET scanner. *IEEE Nuclear Science Symposium Conference Record*; Rome; Italy. 2004. p. 4032-4037.
59. Guérin B, Fakhri GE. Novel scatter compensation of list-mode PET data using spatial and energy dependent corrections. *IEEE Trans. Med. Imag.* 2011; 30:759–773.
60. Snyder DL. Parameter estimation for dynamic studies in emission-tomography systems having list-mode data. *IEEE Trans. Nucl. Sci.* 1984; 31:925–931.
61. Gilks, WR.; Richardson, S.; Spiegelhalter, DJ. Introducing Markov chain Monte Carlo. In: Gilks, WR.; Richardson, S.; Spiegelhalter, DJ., editors. *Markov chain Monte Carlo in Practice*. Vol. chap. 1. Boca Raton, FL: Chapman & Hall/CRC; 1996.
62. Neal, RM. Tech. Rep. CRG-TR-1. Department of Computer Science, University of Toronto; 1993. Probabilistic inference using Markov chain Monte Carlo methods.
63. Lehovich, A. Ph.D. thesis. Tucson, AZ: University of Arizona; 2005. List-mode SPECT reconstruction using empirical likelihood.
64. Korpar S, Križan P, Pestotnik R, Gorišek A, Stanovnik A, Stari M, Škrk D. Multinode photomultipliers as position-sensitive detectors of single photons. *Nucl. Instrum. Methods Phys. Res., A.* 2000; 442:316–321.
65. Rolland JP, Barrett HH. Effect of random background inhomogeneity on observer detection performance. *J. Opt. Soc. Am. A.* 1992; 9:649–658. [PubMed: 1588452]
66. Wang, ZX.; Guo, DR. *Special functions*. Singapore: World Scientific; 1989.
67. Apelblat, A. *Table of Definite and Infinite Integrals*. Amsterdam, Netherlands: Elsevier Scientific Pub. Co.; 1983.
68. Andrews, GE.; Askey, R.; Roy, R. *Special functions*. Cambridge, England: Cambridge University Press; 2000.

Fig. 1.

Plots of SNR^2 for three different cases. For all plots, $\Delta f_0 = 100 \text{ s}^{-1}$, $f_0 = 1000 \text{ s}^{-1}$, $p = 20000 \text{ m}^{-2}$, $\sigma = 0.001 \text{ m}$, and $r_b = 0.005 \text{ m}$. For the solid curves, $r_s = 0.0045 \text{ m}$; for the dashed curves, $r_s = r_b = 0.0050 \text{ m}$; and, finally, for the dash-dotted curves $r_s = 0.0055 \text{ m}$. Thick curves corresponds to the SNR^2 for the Hotelling observer, while thin curves to the MCMC-calculated SNR^2 for the ideal observer