Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2012 October 1; 29(10): 2204–2216.

Asymptotic ideal observers and surrogate figures of merit for signal detection with list-mode data

Eric Clarkson

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

Eric Clarkson: clarkson@radiology.arizona.edu

Abstract

The asymptotic form for the likelihood ratio is derived for list-mode data generated by an imaging system viewing a possible signal in a randomly generated background. This calculation provides an approximation to the likelihood ratio that is valid in the limit of large number of list entries, i.e., a large number of photons. These results are then used to derive surrogate figures of merit, quantities that are correlated with ideal-observer performance on detection tasks, as measured by the area under the receiver operating characteristic curve, but are easier to compute. A key component of these derivations is the determination of asymptotic forms for the Fisher information for the signal amplitude in the limit of a large number of counts or a long exposure time. This quantity is useful in its own right as a figure of merit (FOM) for the task of estimating the signal amplitude. The use of the Fisher information in detection tasks is based on the fact that it provides an approximation for ideal-observer detectability when the signal is weak. For both the fixed-count and fixed-time cases, four surrogate figures of merit are derived. Two are based on maximum likelihood reconstructions; one uses the characteristic functional of the random backgrounds. The fourth surrogate FOM is identical in the two cases and involves an integral over attribute space for each of a randomly generated sequence of backgrounds.

1. INTRODUCTION

For signal-detection problems, there is an optimal observer known as the ideal observer. Optimality in this case can be defined in several ways, one of which is the area under the receiver operating characteristic (ROC) curve. For a given statistical model of the data in the signal-absent and signal-present populations, the ideal observer maximizes the area under the ROC curve, also known as the area under the curve (AUC). The AUC is a figure of merit (FOM) for detection tasks that can be directly related to observer performance on two-alternative forced choice tests, where the observer is presented with two data vectors and must decide which one was drawn from the signal-present population. The ideal-observer AUC therefore can be used as a FOM quantifying the usefulness of the data produced by an imaging system for the particular detection task at hand.

The ideal observer is a mathematical observer that makes a decision for a given data vector by computing a likelihood ratio from the data and comparing it to a threshold [1]. Another quantity, the ideal-observer detectability, is monotonically related to the ideal-observer AUC and therefore provides an equivalent FOM for signal detection. The detectability for the ideal observer can be difficult to compute when there is randomness in the background of

the signal as well as noise in the imaging system itself. It is therefore desirable to have a surrogate FOM (SFOM), a quantity that is easier to compute than the ideal-observer detectability but that correlates with it over a wide range of parameters specifying the background model, the signal model, and the imaging system. We have developed some SFOMs for imaging systems with binned data [2], and in this work we want to apply the same methods to systems that produce list-mode data [3,4]. A key component of this method is the Fisher information matrix (FIM).

The FIM is usually thought of as an important component of statistical estimation theory. The FIM is defined for a parameterized family of probability distribution functions (PDFs), where the parameter is a finite dimensional vector. In estimation theory, the Cramer–Rao bound is derived from the FIM and provides a lower bound on the variances of unbiased estimators of the components of the vector parameter [5]. One problem with the Cramer–Rao lower bound is that the minimum variance is only achieved for exponential families of PDFs, a very restricted class. In imaging, if we want to include object variability as well as system noise, then, as we will see below, the PDFs are mixture distributions, which do not form exponential families. Another problem with the Cramer–Rao bound is that it depends on the true parameter value, which is unknown and often a random quantity itself. A second property of the FIM is that the inverse of this matrix is asymptotically equal to the covariance matrix for the maximum likelihood estimator of the vector parameter [5]. Other asymptotic results relate the FIM to ideal-observer detection performance [6]. The difficulty with using these results in the imaging context is that we do not typically operate in an asymptotic regime, where you have many independent samples from a given member of the family of PDFs. In the imaging context we usually only have one image to work with and need to estimate a vector parameter and/or detect a signal without generating more images from the given object.

In previous publications [2,7] we have shown that there is a connection between the FIM and the ideal-observer detectability. The context in which this connection arises is the problem of detecting a small change in the vector parameter of a parameterized family of PDFs for the data. For example, if a signal is weak, then we may introduce a signal amplitude parameter and reduce the signal-detection problem to the detection of a small change in this parameter from zero. If we plot the detectability of the ideal observer versus signal amplitude, then the slope of this curve at the origin is the Fisher information for the amplitude parameter. One useful aspect of this relationship is that there are no special requirements on the form of the family of PDFs. This is also not an asymptotic result; it applies to the problem of detecting a signal from one data vector generated by the imaging system. In the sections below, we will use the FIM approximation to ideal-observer detectability to derive SFOMs for list-mode data. The essential strategy is to compute the FIM for the signal amplitude and then find approximations to this quantity that are valid when the number of list entries, the number of detected photons, is a large number. We will consider the fixed-counts case first, and then move on to the fixed-time case.

2. FIXED COUNTS

We do fixed counts first because it is easier in some ways and results from this case can be modified for use in the fixed-time case. We assume therefore that the number of counts is fixed ahead of time. We do not assume that the time it takes to collect these counts is measured, although in practice this could be useful information especially for certain estimation tasks. Thus, time is a random variable, and its distribution would have to be included in the analysis if it were being measured. The discussion in this section could be modified to take this into account.

A. Null Functions

We need to understand null functions for the system in order to make sense of asymptotic ideal observers in what follows. First we will review the basics of list-mode data for fixed counts. Then we will define the concept of a null function in this context and provide the conditions that null functions have to satisfy.

The conditional probability density function for a list entry $\hat{\mathbf{A}}$ in a list-mode imaging system conditioned on the object function $\mathbf{f}$ is given by

$$\text{pr}(\hat{\mathbf{A}}|\mathbf{f}) = \frac{\int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r})s(\mathbf{r})f(\mathbf{r})d\mathbf{r}}{\int_S s(\mathbf{r})f(\mathbf{r})d\mathbf{r}} = \frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}}{\mathbf{s}^\dagger \mathbf{f}}, \quad (1)$$

where S is the support region for all object functions, $s(\mathbf{r})$ is the sensitivity function (which gives the probability that a photon emitted from location $\mathbf{r}$ will be detected), and $\text{pr}(\hat{\mathbf{A}}|\mathbf{r})$ is the conditional probability that a detected photon emitted from location $\mathbf{r}$ will give the list entry $\hat{\mathbf{A}}$. The space of list entries will be denoted by L . We will switch between the function notation and the vector notation at will. The vector notation may refer either to vectors in an infinite dimensional Hilbert space for theoretical purposes or to vectors in a finite dimensional space if the functions involved are discretized for computational purposes. In either case the space of objects will be denoted by U . Note that the vector $\mathbf{p}(\hat{\mathbf{A}})$ represents the function $\text{pr}(\hat{\mathbf{A}}|\mathbf{r})s(\mathbf{r})$. This fact gives us the relation

$$\int_L \mathbf{p}(\hat{\mathbf{A}})d\hat{\mathbf{A}} = \mathbf{s}, \quad (2)$$

which will be used below. Also note that the adjoint dagger symbol notation is used to denote the Hilbert space inner products represented by the integrals in Eq. (1).

Without loss of generality, we can assume that $s(\mathbf{r}) > 0$ everywhere in the support region S . If the sensitivity function vanishes in some subregion of S , we might as well remove that subregion from consideration since photons being emitted from there will never contribute to the data. The quantity

$$\frac{s(\mathbf{r})f(\mathbf{r})}{\int_S s(\mathbf{r})f(\mathbf{r})d\mathbf{r}} \quad (3)$$

is the conditional probability that a detected photon was emitted from location $\mathbf{r}$ given that the object function is $f(\mathbf{r})$. Note that this fact implies that $f(\mathbf{r}) \geq 0$ everywhere in S and

$$\int_S s(\mathbf{r})f(\mathbf{r})d\mathbf{r} > 0. \quad (4)$$

We will assume that all object functions satisfy these two positivity constraints.

The conditional probability for entire list is given by the product

$$\text{pr}(\mathcal{A}|\mathbf{f}) = \prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}), \quad (5)$$

where J is the number of entries in the list. In this section J is assumed to be fixed and known. In Section 3, the exposure time will be fixed and known, and J will become a random number. If the time to reach J counts is measured, then this PDF would be multiplied by the PDF for this time, which would be a gamma distribution.

In the following $\mathbf{f}_a$ will usually represent the actual static activity distribution, the object that produced the given data list. A null function at $\mathbf{f}_a$ is a function $\mathbf{f}_n$ that satisfies the condition

$$\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a+\mathbf{f}_n)=\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a) \quad (6)$$

for all arbitrary list entries $\hat{\mathbf{A}}$. Our first task is to see what conditions define the null functions at $\mathbf{f}_a$. To this end, let $\mathbf{f} = \mathbf{f}_a + \mathbf{f}_n$ and note that this function must satisfy the two positivity constraints. Define the normalization integrals for each object by

$$A = \int_S s(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} > 0, \quad (7)$$

$$A_a = \int_S s(\mathbf{r}) f_a(\mathbf{r}) d\mathbf{r} > 0. \quad (8)$$

We next define the scaled objects by $\mathbf{f}'_a = A_a^{-1} \mathbf{f}_a$ and $\mathbf{f}' = A^{-1} \mathbf{f}$. The conditional probability for a list entry is invariant under scaling, so we have $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}') = \text{pr}(\hat{\mathbf{A}}|\mathbf{f}'_a)$. The scaled object functions therefore satisfy

$$\int_S s(\mathbf{r}) f'(\mathbf{r}) d\mathbf{r} = \int_S s(\mathbf{r}) f'_a(\mathbf{r}) d\mathbf{r} = 1, \quad (9)$$

and, for all $\hat{\mathbf{A}}$,

$$\int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) s(\mathbf{r}) f'(\mathbf{r}) d\mathbf{r} = \int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) s(\mathbf{r}) f'_a(\mathbf{r}) d\mathbf{r}. \quad (10)$$

We may write $\mathbf{f}' = \mathbf{f}'_a + \mathbf{f}'_n$ and derive the equivalent conditions for $\mathbf{f}'_n$:

$$\int_S s(\mathbf{r}) f'_n(\mathbf{r}) d\mathbf{r} = 0, \quad (11)$$

and, for all $\hat{\mathbf{A}}$,

$$\int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) s(\mathbf{r}) f'_n(\mathbf{r}) d\mathbf{r} = 0. \quad (12)$$

If we integrate the second of these equations over $\hat{\mathbf{A}}$ we arrive at the first equation. Thus, the second equation for $f'_n(\mathbf{r})$ is all that we need. Now reverting to the original unscaled objects, we find that

$$\mathbf{f} = \left(\frac{A}{A_a} \right) \mathbf{f}_a + A \mathbf{f}'_n = \mathbf{f}_a + \left[\left(\frac{A}{A_a} \right) - 1 \right] \mathbf{f}_a + A \mathbf{f}'_n. \quad (13)$$

Therefore, we may say that the null functions at $\mathbf{f}_a$ all have the form $\mathbf{f}_n = C \mathbf{f}_a + \tilde{\mathbf{f}}$, where, for all $\hat{\mathbf{A}}$,

$$\int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r})s(\mathbf{r})\tilde{f}(\mathbf{r})d\mathbf{r}=0 \quad (14)$$

and $C > -1$. Of the two positivity constraints, the second one is automatically satisfied at this point. The first one gives us $\tilde{f}(\mathbf{r}) = -(1 + C)f_a(\mathbf{r})$ for all $\mathbf{r}$ in S .

An equivalent description can be formulated in terms of the space of constant conditional probability at $\mathbf{f}_a$, i.e., the set of objects $\mathbf{f}$ satisfying the two positivity constraints and $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}) = \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$. These functions have the form $\mathbf{f} = B\mathbf{f}_a + \tilde{\mathbf{f}}$ with $B > 0$ and $\tilde{\mathbf{f}}$ as above. Again the second positivity constraint is automatically satisfied, and the first positivity constraint reduces to $\tilde{f}(\mathbf{r}) = -Bf_a(\mathbf{r})$ for all $\mathbf{r}$ in S . We will call this space $U(f_a)$, and it will make an appearance below when we discuss the asymptotic values of the likelihood ratio. It is straightforward to show that $U(f_a)$ is a convex cone, although this will not play a role in the discussions below.

B. Ideal Observer for Large Counts

The ideal observer for a detection task computes the likelihood ratio and compares this number to a threshold to decide the presence or absence of a signal. We will show that, in the asymptotic limit of a large number of counts, the likelihood ratio has a limiting value that depends only on the conditional probability $\text{pr}(\hat{\mathbf{A}}|\mathbf{f})$. We start by looking at asymptotic values for the conditional probability of the list.

The conditional PDF for the list $\mathcal{A} = \{\hat{\mathbf{A}}_1, \dots, \hat{\mathbf{A}}_J\}$ is given by the product

$$\text{pr}(\mathcal{A}|\mathbf{f}) = \prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}). \quad (15)$$

The unconditional PDF for the list is the integral

$$\text{pr}(\mathcal{A}) = \int_U \text{pr}(\mathcal{A}|\mathbf{f})\text{pr}(\mathbf{f})d\mathbf{f}. \quad (16)$$

Both of these PDFs are defined on the space of lists $\mathcal{L} = L^J$. We assume that the support for the prior PDF $\text{pr}(\mathbf{f})$ is contained within the set of objects satisfying the two positivity constraints. In the case where the object space U is infinite dimensional, the concept of a prior PDF is problematic since stochastic processes, in general, do not have PDFs. In this case we would replace this last equation with the expectation

$$\text{pr}(\mathcal{A}) = \langle \text{pr}(\mathcal{A}|\mathbf{f}) \rangle_{\mathbf{f}}, \quad (17)$$

which has a well-defined meaning for any stochastic process. We will use the integral notation throughout, but all expressions may be easily converted to the expectation notation as needed.

As a first step toward the asymptotic ideal observer for signal detection, we rewrite the conditional probability in the form

$$\text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) = \exp \left\{ \sum_{j=1}^J \ln[\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f})] \right\}. \quad (18)$$

For a large number of counts, we will get, with high probability, a good approximation by replacing the sum with an expectation value:

$$\text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) \cong \exp \left\{ J \int_L \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a) \ln[\text{pr}(\hat{\mathbf{A}}|\mathbf{f})] d\hat{\mathbf{A}} \right\}. \quad (19)$$

The object $\mathbf{f}_a$ is the actual object that produced the data. This object is often specified in simulations used to evaluate imaging approaches but is not directly known in real clinical scans. This type of approximation will be used often below and has a rigorous meaning from information theory. That is, we can show that, for any $\varepsilon > 0$, the set $\mathcal{L}(\varepsilon)$ of lists $\mathcal{A}$ that satisfy

$$\left| \frac{1}{J} \ln[\text{pr}(\widehat{\mathcal{A}}|\mathbf{f})] - \int_L \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a) \ln[\text{pr}(\hat{\mathbf{A}}|\mathbf{f})] d\hat{\mathbf{A}} \right| < \varepsilon, \quad (20)$$

has probability

$$\int_{\mathcal{L}(\varepsilon)} \text{pr}(\widehat{\mathcal{A}}|\mathbf{f}_a) d\widehat{\mathcal{A}} > 1 - \varepsilon \quad (21)$$

if J is sufficiently large. This type of convergence is called convergence in probability. Note that the quantity on the right-hand side in the approximation for $\text{pr}(\widehat{\mathcal{A}}|\mathbf{f})$ does not depend on the actual list itself but only on the object function that generated it. This means that, for large J , most lists, in the probabilistic sense, are approximately equally probable. In information theory this fact is called the asymptotic equipartition property [8]. The integral in the exponent in the approximation to $\text{pr}(\widehat{\mathcal{A}}|\mathbf{f})$ is the cross entropy, and it can be expressed in terms of the Kullback–Leibler distance,

$$D[\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a), \text{pr}(\hat{\mathbf{A}}|\mathbf{f})] = \int_L \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a) \ln \left[\frac{\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)}{\text{pr}(\hat{\mathbf{A}}|\mathbf{f})} \right] d\hat{\mathbf{A}}, \quad (22)$$

and the differential entropy,

$$H[\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)] = - \int_L \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a) \ln[\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)] d\hat{\mathbf{A}}. \quad (23)$$

Notice that each of these two expressions involves an average over attribute vector $\hat{\mathbf{A}}$. We will use a more concise notation $D_L(\mathbf{f}_a, \mathbf{f})$ for the Kullback–Leibler distance and $H_L(\mathbf{f}_a)$ for the differential entropy. The subscript L is used to emphasize that these quantities are calculated from the conditional PDFs in list space corresponding to the objects. In terms of these two quantities, the approximate value for $\text{pr}(\widehat{\mathcal{A}}|\mathbf{f})$ is given by

$$\text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) \cong \exp[-JD_L(\mathbf{f}_a, \mathbf{f})] \exp[-JH_L(\mathbf{f}_a)]. \quad (24)$$

The Kullback–Leibler distance is always nonnegative and equals zero only if the two conditional PDFs $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$ and $\text{pr}(\hat{\mathbf{A}}|\mathbf{f})$ are the same. The differential entropy can be positive or negative.

The likelihood ratio is the test statistic for the ideal observer and is given by

$$\Lambda(\widehat{\mathcal{A}}) = \frac{\int_U \text{pr}(\widehat{\mathcal{A}}|\mathbf{f} + \Delta\mathbf{f}) \text{pr}(\mathbf{f}) d\mathbf{f}}{\int_U \text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (25)$$

We assume the signal $\Delta\mathbf{f}$ is a nonnegative function. This is needed to assure that $\text{pr}(\widehat{\mathcal{A}}|\mathbf{f} + \Delta\mathbf{f})$ is always well defined for objects $\mathbf{f}$ that satisfy the two positivity constraints. An alternative form for the likelihood is given by

$$\Lambda(\widehat{\mathcal{A}}) = \frac{\int_U \text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) \text{pr}(\mathbf{f} - \Delta\mathbf{f}) d\mathbf{f}}{\int_U \text{pr}(\widehat{\mathcal{A}}|\mathbf{f}) \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (26)$$

Using the cross-entropy approximation, we may write

$$\Lambda(\widehat{\mathcal{A}}) \cong \frac{\int_U \exp[-JD_L(\mathbf{f}_a, \mathbf{f})] \text{pr}(\mathbf{f} - \Delta\mathbf{f}) d\mathbf{f}}{\int_U \exp[-JD_L(\mathbf{f}_a, \mathbf{f})] \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (27)$$

Note the quantity on the right-hand side in this expression is not itself a test statistic since it does not depend on the actual data. What this expression tells us is that, if the actual object is $\mathbf{f}_a$ and we generate many data lists from this object, then with high probability the value for the corresponding likelihood ratio will be close to the quantity on the right-hand side.

We expect that, for a large number of counts, with high probability the objects that give a small value for $D_L(\mathbf{f}_a, \mathbf{f})$ will give the overwhelming contribution to the integrals in this last expression. Expanding the Kullback–Leibler distance around the actual object gives, to the lowest order, a quadratic function, and we then arrive at the approximation

$$\Lambda(\widehat{\mathcal{A}}) \cong \frac{\int_U \exp \left\{ -\frac{J}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) \right\} \text{pr}(\mathbf{f} - \Delta\mathbf{f}) d\mathbf{f}}{\int_U \exp \left\{ -\frac{J}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) \right\} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (28)$$

The matrix $\mathbf{F}(\mathbf{f}_a)$ is the FIM for the conditional probability $\text{pr}(\hat{\mathbf{A}}|\mathbf{f})$ evaluated at the actual object. For an infinite dimensional object space, the FIM is an operator. We discuss this last approximation further in Subsection 2.C.

C. FIM and the Null Space

If we examine the approximation in Eq. (27), we can see that the quantity in the exponent is zero if $\mathbf{f}$ and $\mathbf{f}_a$ differ by a null function at $\mathbf{f}_a$. We want to see whether the approximation in Eq. (28) has the same property. This will help us to see what happens to the likelihood ratio as $J \rightarrow \infty$.

We can compute the FIM by using the definition of the conditional probability $\text{pr}(\hat{\mathbf{A}}|\mathbf{f})$ from above. Computing the score is the first step, and this vector is defined by $\text{sc}(\hat{\mathbf{A}}|\mathbf{f}) = \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}|\mathbf{f})]$. A short calculation leads to

$$\text{sc}(\hat{\mathbf{A}}|\mathbf{f}) = \frac{\mathbf{p}(\hat{\mathbf{A}})}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}} - \frac{\mathbf{s}}{\mathbf{s}^\dagger \mathbf{f}}. \quad (29)$$

The two positivity constraints on $\mathbf{f}$ ensure that the components of this vector are always finite numbers. The FIM is the covariance of the score $\mathbf{F}(\mathbf{f}) = \langle \text{sc}(\hat{\mathbf{A}}|\mathbf{f}) \text{sc}^\dagger(\hat{\mathbf{A}}|\mathbf{f}) \rangle_{\hat{\mathbf{A}}|\mathbf{f}}$ and can be written as

$$\mathbf{F}(\mathbf{f}) = \left\langle \frac{\mathbf{p}(\hat{\mathbf{A}})\mathbf{p}(\hat{\mathbf{A}})^\dagger}{[\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}]^2} \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}} - \frac{\mathbf{s}\mathbf{s}^\dagger}{(\mathbf{s}^\dagger \mathbf{f})^2}. \quad (30)$$

In practice this matrix can be estimated by drawing samples from the conditional PDF $\text{pr}(\hat{\mathbf{A}}|\mathbf{f})$.

Now we want to relate the FIM to the null functions described above. Let us define the difference function $f_d = f - f_a$. Then the quadratic form in the exponent in the FIM approximation to the likelihood ratio in Eq. (28) is given by

$$\mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d = \left\langle \left[\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_d}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} - \left(\frac{\mathbf{s}^\dagger \mathbf{f}_d}{\mathbf{s}^\dagger \mathbf{f}_a} \right)^2. \quad (31)$$

This quantity vanishes if f_d is a null function at f_a . On the other hand, we may write the quadratic form as

$$\mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d = \left\langle \left[\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_d}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} - \frac{\mathbf{s}^\dagger \mathbf{f}_d}{\mathbf{s}^\dagger \mathbf{f}_a} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}}. \quad (32)$$

If this quantity vanishes, then, for almost all $\hat{\mathbf{A}}$, we have the equality

$$\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_d}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} = \frac{\mathbf{s}^\dagger \mathbf{f}_d}{\mathbf{s}^\dagger \mathbf{f}_a}. \quad (33)$$

If the numerator on the right in this equation is zero, the numerator on the left is also, and f_d is a null function. If the numerator on the right is nonzero, we have the equality

$$\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_d}{\mathbf{s}^\dagger \mathbf{f}_d} = \frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a}{\mathbf{s}^\dagger \mathbf{f}_a}. \quad (34)$$

As above this gives us $\mathbf{f}_d = B\mathbf{f}_a + \tilde{\mathbf{f}}$ with $\mathbf{p}(\hat{\mathbf{A}})^\dagger \tilde{\mathbf{f}} = 0$ for all $\hat{\mathbf{A}}$. Therefore, f_d is also a null function at f_a in this case. Note that the $B = 0$ situation is covered in the first case, where the numerators are zero. On the other hand, the exponent in Eq. (28) will be zero when $B = -1$ even though f_d is not a null function at f_a strictly speaking for this value of B . Other than that caveat, which is not important for any of the integrals involved, we can say that the

quadratic form in the exponent in Eq. (28) is zero if and only if f_d is a null function at f_a . This means that the quadratic expansion used in Subsection 2.B for the Kullback–Leibler distance is an expansion for f near the space of constant conditional probability $U(f_a)$ within the object space. In particular f does not have to be near f_a for the quadratic expansion to be a valid approximation to the KL distance. All that is required is that $f - f_a$ is near a null function at f_a . Now we will return to the likelihood ratio and see what happens as $J \rightarrow \infty$.

D. Asymptotic Form for the Ideal Observer

Now we return to the FIM approximation of the likelihood ratio in Eq. (28) and rewrite the denominator as

$$\int_U \exp \left[-\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d \right] \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d. \quad (35)$$

Define $U_0 = U_0(f_a)$ as the null space of $\mathbf{F}(\mathbf{f}_a)$ and $U_1 = U_1(f_a)$ as the orthogonal complement to U_0 . Also let $\mathbf{f}_d = \mathbf{f}_{d0} + \mathbf{f}_{d1}$ be the corresponding orthogonal decomposition for the difference vector. We have shown above that U_0 is the linear span of the functions in $U(f_a)$, and that these functions are essentially all of the nonnegative functions in U_0 . Define the eigenvalues and eigenvectors of the FIM by $\mathbf{F}(\mathbf{f}_a) \mathbf{u}_n = \lambda_n \mathbf{u}_n$ and consider the eigenvector decomposition of the difference vector:

$$\mathbf{f}_d = \sum_{n=1}^N f_{dn} \mathbf{u}_n. \quad (36)$$

The quadratic form in the exponent of the FIM approximation to Λ can be written as

$$\mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d = \sum_{n=1}^R \lambda_n |f_{dn}|^2 = \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1}, \quad (37)$$

where R is the rank of $\mathbf{F}(\mathbf{f}_a)$. Now the denominator in Eq. (28) can be written as

$$\int_{U_1} \exp \left[-\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right] \int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}. \quad (38)$$

By normalizing the Gaussian function that appears in both the numerator and denominator of the FIM approximation to Λ in Eq. (28), we can see that, as $J \rightarrow \infty$, the value of the likelihood ratio asymptotically approaches a limit:

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a - \Delta \mathbf{f}) d\mathbf{f}_{d0}}{\int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a) d\mathbf{f}_{d0}}. \quad (39)$$

Since the vector $\mathbf{f}_a$ is in the subspace U_0 , we may write this limit as

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U_0} \text{pr}(\mathbf{f}_{d0} - \Delta \mathbf{f}) d\mathbf{f}_{d0}}{\int_{U_0} \text{pr}(\mathbf{f}_{d0}) d\mathbf{f}_{d0}}. \quad (40)$$

Finally, we use the fact that we have a positive signal and that the prior PDF is supported on the set of functions satisfying the two positivity constraints to write

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U(f_a)} \text{pr}(\mathbf{f} - \Delta \mathbf{f}) d\mathbf{f}}{\int_{U(f_a)} \text{pr}(\mathbf{f}) d\mathbf{f}}, \quad (41)$$

as $J \rightarrow \infty$. The ratio in this last equation is the likelihood ratio for the observer that is directly observing the subspace of objects whose conditional probabilities are identical with the conditional probability of the actual object. The performance of this observer defines the maximum performance for the original ideal observer. Another way of thinking about this observer is that it is observing the conditional PDF $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$ directly instead of J samples from this PDF.

The performance of this asymptotic ideal observer provides an upper limit for the performance of the ideal list-mode observer for the given signal-detection task. It is possible that the AUC of the asymptotic ideal observer is unity, in which case the detectability for the list-mode ideal observer will increase without bound as $J \rightarrow \infty$. To see how this is possible, suppose first that the signal is absent. Then $\mathbf{f}_a$ is in the support of the prior distribution. If $\mathbf{f} - \Delta \mathbf{f}$ is not within this support for any $\mathbf{f}$ in $U(f_a)$, then the numerator in the asymptotic likelihood ratio in Eq. (41) will be zero. If this property of the signal holds for any signal-absent $\mathbf{f}_a$, then the asymptotic likelihood ratio is always zero when the signal is absent. Now suppose that the signal is present. Then $\mathbf{f}_a - \Delta \mathbf{f}$ is in the support of the prior distribution. If $\mathbf{f}$ is not in this support for any $\mathbf{f}$ in $U(f_a)$, then the denominator in the asymptotic likelihood ratio in Eq. (41) is zero. If this property holds for any signal-present $\mathbf{f}_a$, then the asymptotic likelihood ratio is always infinite when the signal is present. If the signal satisfies both of these properties, then the AUC for the asymptotic ideal observer will be unity.

As an example of how the scenario described in the previous paragraph can happen, suppose that we have a planar optical imaging system and the actual signal-free objects are flat with unknown amplitude. The support of the prior is then the one-dimensional space of flat objects. Suppose that the signal has frequency components within the passband of the imaging system. For any signal-absent object $\mathbf{f}_a$, the space $U(f_a)$ consists of sums of flat objects and objects whose frequency components are all higher than the band limit of the imaging system. Then $\mathbf{f} - \Delta \mathbf{f}$ is not flat for any $\mathbf{f}$ in $U(f_a)$. On the other hand, for any signal-present object $\mathbf{f}_a$, the space $U(f_a)$ consists of sums of flat objects, the signal, and objects whose frequency components are all higher than the band limit of the imaging system. Then $\mathbf{f}$ is not flat for any $\mathbf{f}$ in $U(f_a)$. This is an extreme example, but it illustrates that an AUC of unity can occur for the asymptotic likelihood ratio if the prior is supported on a relatively low-dimensional parameterized space, the signal is not a null object, and the signal is not in the support of the prior.

E. SFOMs

We could use the AUC or detectability of the ideal list-mode observer as an FOM for the imaging system on the detection task, but the likelihood ratio is difficult to compute accurately. We have developed Markov chain Monte Carlo (MCMC) methods to do this computation, but they are time consuming. An SFOM is a quantity that can be computed more easily and will correlate with the ideal-observer detectability when system or signal parameters are varied. We will examine some SFOMs that arise from a relation between ideal-observer detectability and Fisher information that we have derived elsewhere.

To derive an SFOM, we first introduce an amplitude parameter α for the signal $\Delta \mathbf{f}$ and write

$$\text{pr}(\widehat{\mathcal{A}}|\alpha) = \int_U \text{pr}(\widehat{\mathcal{A}}|f + \alpha \Delta f) \text{pr}(f) df. \quad (42)$$

For convenience of notation we will define $\text{pr}(\widehat{\mathcal{A}}) = \text{pr}(\widehat{\mathcal{A}}|0)$. The SFOM will arise from the relation between the detectability for the ideal observer and the Fisher information for the amplitude parameter α . Therefore, the first step is to compute the score for α . The Fisher information is then the expectation of the square of the score evaluated at $\alpha = 0$.

The score for the amplitude parameter at $\alpha = 0$ is defined in the usual way as

$$\text{sc}(\widehat{\mathcal{A}}) = \frac{d}{d\alpha} \ln[\text{pr}(\widehat{\mathcal{A}}|\alpha)] \Big|_{\alpha=0}. \quad (43)$$

A short calculation leads to an integral expression

$$\text{sc}(\widehat{\mathcal{A}}) = \int_U \Delta f^\dagger \text{sc}(\widehat{\mathcal{A}}|f) \text{pr}(f|\widehat{\mathcal{A}}) df, \quad (44)$$

where the local score is defined by $\text{sc}(\widehat{\mathcal{A}}|f) = \nabla_f \ln[\text{pr}(\widehat{\mathcal{A}}|f)]$ and the posterior PDF $\text{pr}(f|\widehat{\mathcal{A}})$ satisfies the equation $\text{pr}(f|\widehat{\mathcal{A}})\text{pr}(\widehat{\mathcal{A}}) = \text{pr}(\widehat{\mathcal{A}}|f)\text{pr}(f)$. Thus, the (global) score $\text{sc}(\widehat{\mathcal{A}})$ is the inner product of the signal with the posterior mean of the local score. The word “local” here refers to the fact that the local score is conditioned on a given object f .

Now we can calculate the local score and get an expression for the score. The result is

$$\text{sc}(\widehat{\mathcal{A}}) = \int_U \Delta f^\dagger \left[\sum_{j=1}^J \text{sc}(\widehat{\mathbf{A}}_j|f) \right] \text{pr}(f|\widehat{\mathcal{A}}) df, \quad (45)$$

where $\text{sc}(\widehat{\mathbf{A}}_j|f)$ is given by Eq. (29). The FIM approximation to ideal-observer detectability as a function of amplitude follows from the Taylor series expansion of the square of the ideal-observer detectability as a function of α :

$$d^2(\alpha) = \alpha^2 \langle [\text{sc}(\widehat{\mathcal{A}})]^2 \rangle_{\widehat{\mathcal{A}}} + \dots \quad (46)$$

The next term in the expansion is cubic in the amplitude parameter and involves the derivative of the FIM as a function of α . The fourth-order term is not defined unless the kurtosis of the score is zero. We will use the first nonzero term, i.e., the quadratic term shown above, as our approximation for the detectability:

$$d^2(\alpha) \cong \alpha^2 \langle \langle [\text{sc}(\widehat{\mathcal{A}})]^2 \rangle_{\widehat{\mathcal{A}}|f_a} \rangle_{f_a}. \quad (47)$$

In this expression we have written out the expectation as a two-stage process, where we average over lists given an object and then average over the objects. This approximation will be good for weak signals. Note that the ideal-observer detectability is related to the ideal-observer AUC by the relation $2[\text{AUC}(\alpha)] = 1 + \text{erf}[1/2d(\alpha)]$.

We could stop here, but our expression for the score involves a posterior mean, which will require MCMC techniques to evaluate. MCMC methods tend to be computationally

intensive, so we will examine approximations to the FIM approximation and hope they lead us to something easier to compute. The ultimate test of any resulting SFOM is to simulate a list-mode system where the ideal-observer detectability can be computed (as noted earlier, we have to use MCMC methods for this) and see if the proposed SFOM correlates with this quantity over a range of signals and/or system parameters.

1. Fixed-Counts SFOM 1—For the first SFOM, we use the law of large numbers to approximate the sum of local scores with an expectation:

$$sc(\widehat{\mathcal{A}}) \cong J \int_U \Delta \mathbf{f}^\dagger \langle \mathbf{sc}(\widehat{\mathbf{A}}|\mathbf{f}) \rangle_{\widehat{\mathbf{A}}|\mathbf{f}_a} \text{pr}(\mathbf{f}|\widehat{\mathcal{A}}) d\mathbf{f}. \quad (48)$$

Ideal-observer detectability can then be approximated by

$$d^2(\alpha) \cong \alpha^2 J^2 \left\langle \left\langle \left[\int_U \Delta \mathbf{f}^\dagger \langle \mathbf{sc}(\widehat{\mathbf{A}}|\mathbf{f}) \rangle_{\widehat{\mathbf{A}}|\mathbf{f}_a} \text{pr}(\mathbf{f}|\widehat{\mathcal{A}}) d\mathbf{f} \right]^2 \right\rangle_{\widehat{\mathcal{A}}|\mathbf{f}_a} \right\rangle_{\mathbf{f}_a}. \quad (49)$$

So far this does not help much as we will still need MCMC methods to compute the integral in square brackets. However, if we interchange the expectation and the integration within the square brackets, we arrive at

$$d^2(\alpha) \cong \alpha^2 J^2 \left\langle \left\langle \left[\int_U \Delta \mathbf{f}^\dagger \mathbf{sc}(\widehat{\mathbf{A}}|\mathbf{f}) \text{pr}(\mathbf{f}|\widehat{\mathcal{A}}) d\mathbf{f} \right]_{\widehat{\mathbf{A}}|\mathbf{f}_a}^2 \right\rangle_{\widehat{\mathcal{A}}|\mathbf{f}_a} \right\rangle_{\mathbf{f}_a}. \quad (50)$$

The integral is the posterior mean (PM) estimate of the quantity $\Delta \mathbf{f}^\dagger \mathbf{sc}(\widehat{\mathbf{A}}|\mathbf{f})$ from the data list. One option at this point to eliminate the need for MCMC methods is to replace this PM estimate with a maximum likelihood (ML) estimate. We get an ML estimate of the local score by producing an ML estimate of the object and then inserting it into the local score. This replacement would then result in the approximation

$$d^2(\alpha) \cong \alpha^2 J^2 \left\langle \left\langle \left[\Delta \mathbf{f}^\dagger \mathbf{sc}[\widehat{\mathbf{A}}|\widehat{\mathbf{f}}_{\text{ML}}(\widehat{\mathcal{A}})] \right]_{\widehat{\mathbf{A}}|\mathbf{f}_a}^2 \right\rangle_{\widehat{\mathcal{A}}|\mathbf{f}_a} \right\rangle_{\mathbf{f}_a}. \quad (51)$$

This approximation requires computation of an ML estimate of the object from the list $\widehat{\mathcal{A}}$ and then ordinary Monte Carlo sampling to compute the expectation of $\Delta \mathbf{f}^\dagger \mathbf{sc}[\widehat{\mathbf{A}}|\widehat{\mathbf{f}}_{\text{ML}}(\widehat{\mathcal{A}})]$ over $\widehat{\mathbf{A}}$ given $\mathbf{f}_a$. Then we square and use Monte Carlo sampling to average over $\widehat{\mathcal{A}}$ given $\mathbf{f}_a$, and finally average over $\mathbf{f}_a$. This is a lot of Monte Carlo averaging, but at least there are no MCMC computations.

2. Fixed-Counts SFOM 2—For this SFOM, we start with the same expectation approximation for the score as in SFOM 1, but now we write the expectation in terms of cross entropy:

$$\Delta \mathbf{f}^\dagger \langle \mathbf{sc}(\widehat{\mathbf{A}}|\mathbf{f}) \rangle_{\widehat{\mathbf{A}}|\mathbf{f}_a} = \Delta \mathbf{f}^\dagger \nabla_{\mathbf{f}} \int_L \ln[\text{pr}(\widehat{\mathbf{A}}|\mathbf{f})] \text{pr}(\widehat{\mathbf{A}}|\mathbf{f}_a) d\widehat{\mathbf{A}}. \quad (52)$$

For large values of J , the posterior distribution is localized near the space $U(f_a)$, so we will use the quadratic approximation for the KL distance:

$$\Delta \mathbf{f}^\dagger \langle \mathbf{sc}(\hat{\mathbf{A}}|\mathbf{f}) \rangle_{\hat{\mathbf{A}}|f_a} \cong -\frac{1}{2} \Delta \mathbf{f}^\dagger \nabla_{\mathbf{f}} [(\mathbf{f}-\mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a)(\mathbf{f}-\mathbf{f}_a)]. \quad (53)$$

Now we take the gradient

$$\Delta \mathbf{f}^\dagger \langle \mathbf{sc}(\hat{\mathbf{A}}|\mathbf{f}) \rangle_{\hat{\mathbf{A}}|f_a} \cong -\Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f}-\mathbf{f}_a). \quad (54)$$

This gives an approximation for the score in terms of the PM estimate of the actual object: $\mathbf{sc}(\hat{\mathbf{A}}) \cong J \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) [\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathbf{A}}) - \mathbf{f}_a]$.

This score approximation gives an approximation for the detectability via the Fisher information. The approximate detectability is expressed as a nested pair of expectations:

$$d^2(\alpha) \cong \alpha^2 J^2 \langle \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) \langle [\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a] [\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a]^\dagger \rangle_{\hat{\mathcal{A}}|f_a} \mathbf{F}(\mathbf{f}_a) \Delta \mathbf{f} \rangle_{f_a}. \quad (55)$$

MCMC methods are still needed here to compute the PM object estimate. To get a more tractable SFOM, we replace the PM estimate with an ML estimate and write

$$d^2(\alpha) \cong \alpha^2 J^2 \langle \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) \langle [\hat{\mathbf{f}}_{\text{ML}}(\hat{\mathcal{A}}) - \mathbf{f}_a] [\hat{\mathbf{f}}_{\text{ML}}(\hat{\mathcal{A}}) - \mathbf{f}_a]^\dagger \rangle_{\hat{\mathcal{A}}|f_a} \mathbf{F}(\mathbf{f}_a) \Delta \mathbf{f} \rangle_{f_a}. \quad (56)$$

We will see below that, in the asymptotic limit, there is actually no difference in this approximation when we replace the PM estimate with an ML estimate. Note that we will still have three expectations to compute in this approximation since the matrix $\mathbf{F}(\mathbf{f}_a)$ requires an expectation using samples from $\text{pr}(\hat{\mathbf{A}}|f_a)$ to evaluate.

3. Fixed-Counts SFOM 3—For this SFOM, we first find the asymptotic form of the vector that appears in SFOM 2:

$$\mathbf{F}(\mathbf{f}_a) [\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a] \cong \mathbf{F}(\mathbf{f}_a) \frac{\int_U (\mathbf{f}-\mathbf{f}_a) \exp \left[-\frac{J}{2} (\mathbf{f}-\mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f}-\mathbf{f}_a) \right] \text{pr}(\mathbf{f}) d\mathbf{f}}{\int_U \exp \left[-\frac{J}{2} (\mathbf{f}-\mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f}-\mathbf{f}_a) \right] \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (57)$$

With the difference vector $\mathbf{f}_d$ as above, we can rewrite these integrals and insert them into the Fisher information detectability approximation. The result is

$$d^2(\alpha) \cong \alpha^2 J^2 \left\langle \left\{ \frac{\int_U \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d \exp \left[-\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d \right] \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d}{\int_U \exp \left[-\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d \right] \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d} \right\}^2 \right\rangle_{f_a}. \quad (58)$$

Both integrals in the curly brackets are inner products between a shifted prior PDF and a function whose Fourier transform is known. This inner product could be performed in Fourier space if the characteristic function for the prior PDF is known, which is often the case. One complication of this approach is that the matrix $\mathbf{F}(\mathbf{f}_a)$ is singular due to the

presence of null functions at $\mathbf{f}_a$, but this does not present insurmountable difficulties if the pseudoinverse of $\mathbf{F}(\mathbf{f}_a)$ can be computed. If the Fourier space approach is viable, then there will only be two Monte Carlo calculations involved in the computation of the SFOM: one for computing $\mathbf{F}(\mathbf{f}_a)$ and the other for the expectation over $\mathbf{f}_a$. This approximation to the ideal-observer detectability should be good for large J .

What happens to this SFOM in the limit as $J \rightarrow \infty$? To answer this last question, we decompose object space using the singular value decomposition of $\mathbf{F}(\mathbf{f}_a)$ as we did for the likelihood ratio. The end result is

$$\mathbf{F}(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a] \cong \mathbf{F}(\mathbf{f}_a) \frac{\int_{U_1} \mathbf{f}_{d1} \exp \left\{ -\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right\} \int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}{\int_{U_1} \exp \left\{ -\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right\} \int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}. \quad (59)$$

By normalizing the Gaussian in the numerator and denominator and letting $J \rightarrow \infty$, we find that $\mathbf{F}(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a] \rightarrow 0$ independently of the actual list $\hat{\mathcal{A}}$. This shows that the quantity in the angle brackets in our third approximation to detectability is approaching zero as $J \rightarrow \infty$. We know that the detectability is bounded by the detectability of the asymptotic ideal observer discussed above. Therefore, the decreasing value of the expectation must cancel the J^2 dependence when the asymptotic ideal detectability is finite.

We cannot say that the PM estimator is asymptotically unbiased due to the existence of null functions. It does however follow from this discussion that, as $J \rightarrow \infty$, we have $\text{pr}(\hat{\mathbf{A}}|\hat{\mathbf{f}}_{\text{PM}}) \rightarrow \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$. This relation is as close as we can get to an asymptotic unbiased condition. Now we will find the asymptotic form for the PM estimator itself. For large J we have the approximation

$$\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a \cong \frac{\int_{U_1} \exp \left\{ -\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right\} \int_{U_0} (\mathbf{f}_{d0} + \mathbf{f}_{d1}) \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}{\int_{U_1} \exp \left\{ -\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right\} \int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}. \quad (60)$$

Normalizing the Gaussian in the usual way, we find that, in the limit as $J \rightarrow \infty$, we get

$$\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a \rightarrow \frac{\int_{U_0} \mathbf{f}_{d0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a) d\mathbf{f}_{d0}}{\int_{U_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a) d\mathbf{f}_{d0}}. \quad (61)$$

This shows that the PM estimator is biased asymptotically. The bias is given by the null function on the right in this relation. Note that the bias depends on the actual object $\mathbf{f}_a$. We may also write this relation in the form

$$\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) \rightarrow \frac{\int_{U(\mathbf{f}_a)} \mathbf{f} \text{pr}(\mathbf{f}) d\mathbf{f}}{\int_{U(\mathbf{f}_a)} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (62)$$

This shows that the PM estimator asymptotically depends only on the prior PDF $\text{pr}(\mathbf{f})$ and the conditional PDF $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$, which is what we expect.

From the asymptotic form of the conditional probability as $J \rightarrow \infty$,

$$\text{pr}(\hat{\mathcal{A}}|\mathbf{f}) \rightarrow \exp \left\{ -\frac{J}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) \right\}, \quad (63)$$

we can see that any ML estimate will have the asymptotic form $\hat{\mathbf{f}}_{\text{ML}} \rightarrow \mathbf{f}_a + \mathbf{f}_n$, where $\mathbf{f}_n$ is a null function at $\mathbf{f}_a$. Since the PM estimate has the same form asymptotically, the replacement of $\mathbf{F}(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) - \mathbf{f}_a]$ with $\mathbf{F}(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{ML}}(\hat{\mathcal{A}}) - \mathbf{f}_a]$ in Approximation 2 results in no penalty in the limit as $J \rightarrow \infty$. If the PM of the local score was asymptotically the same as $\text{sc}(\hat{\mathbf{A}}|\hat{\mathbf{f}}_{\text{PM}})$, we could say something similar about Approximation 1, but these two estimates of the local score are probably not the same.

4. Fixed-Counts SFOM 4—For this last SFOM for fixed counts, we go back to SFOM 2 in Eq. (56) and use asymptotic properties of ML estimators to simplify the expression for approximate detectability. An ML estimator is a solution to the equation

$$\sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j|\hat{\mathbf{f}}_{\text{ML}})] = 0. \quad (64)$$

Using Taylor's theorem, we can expand to the first order about $\mathbf{f}_a$ and write this as

$$\sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}_a)] + \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}_\beta)] (\hat{\mathbf{f}}_{\text{ML}} - \mathbf{f}_a) = 0, \quad (65)$$

where $\mathbf{f}_\beta = \beta \hat{\mathbf{f}}_{\text{ML}} + (1 - \beta)\mathbf{f}_a$ and $0 \leq \beta \leq 1$. These conditions on $\mathbf{f}_\beta$ place it on the line segment connecting $\mathbf{f}_a$ and $\hat{\mathbf{f}}_{\text{ML}}$.

Now define a random vector $\mathbf{y}$ by

$$\mathbf{y} = \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}_a)]. \quad (66)$$

Each component in the sum for the vector $\mathbf{y}$ is the local score evaluated at a list entry. These are independent identically distributed random vectors with mean zero and covariance $\mathbf{F}(\mathbf{f}_a)$. The covariance of $\mathbf{y}$ is therefore $J\mathbf{F}(\mathbf{f}_a)$. Next, using the law of large numbers, we get the first approximate equality in

$$-\frac{1}{J} \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}_\beta)] \cong \mathbf{F}(\mathbf{f}_\beta) \cong \mathbf{F}(\mathbf{f}_a) \quad (67)$$

for large J . The second approximate equality here follows from the fact noted above that, as $J \rightarrow \infty$, $\hat{\mathbf{f}}_{\text{ML}} \rightarrow \mathbf{f}_a + \mathbf{f}_n$, where $\mathbf{f}_n$ is a null function at $\mathbf{f}_a$. Since $\mathbf{f}_\beta$ is between $\mathbf{f}_a$ and $\hat{\mathbf{f}}_{\text{ML}}$, it will have the same asymptotic limit.

Putting these facts together, we have, for large J , the approximation $\mathbf{y} \cong J\mathbf{F}(\mathbf{f}_a)(\hat{\mathbf{f}}_{\text{ML}} - \mathbf{f}_a)$, and the covariance of the vector on the right in this approximate equality is therefore approximately $J\mathbf{F}(\mathbf{f}_a)$. Together with Eq. (56), this gives us a simpler SFOM:

$$d^2(\alpha) \cong \alpha^2 J \langle \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) \Delta \mathbf{f} \rangle_{\mathbf{f}_a}. \quad (68)$$

In explicit form this SFOM can be written as

$$d^2(\alpha) \cong \alpha^2 J \left\{ \left\langle \left\langle \left[\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} \right\rangle_{\mathbf{f}_a} - \left\langle \left(\frac{\mathbf{s}^\dagger \Delta \mathbf{f}}{\mathbf{s}^\dagger \mathbf{f}_a} \right)^2 \right\rangle_{\mathbf{f}_a} \right\}. \quad (69)$$

In terms of computation, this SFOM requires the least amount of Monte Carlo sampling. The dependence on J is somewhat troubling since we know that the detectability cannot increase indefinitely with J when the asymptotic ideal observer has finite detectability. At present we have no convincing solution to this difficulty. However, this may not detract from the use of this approximation as an SFOM since we only need something correlated with ideal-observer performance for this application.

Using Eq. (1), we may write SFOM 4 as a more analytic expression:

$$d^2(\alpha) \cong \alpha^2 J \left\{ \left\langle \frac{1}{\mathbf{s}^\dagger \mathbf{f}_a} \int_L \frac{[\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}]^2}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} d\hat{\mathbf{A}} \right\rangle_{\mathbf{f}_a} - \left\langle \left(\frac{\mathbf{s}^\dagger \Delta \mathbf{f}}{\mathbf{s}^\dagger \mathbf{f}_a} \right)^2 \right\rangle_{\mathbf{f}_a} \right\}. \quad (70)$$

If the integral over attribute space can be performed, then we only need to sample from the prior to calculate this quantity.

3. FIXED TIME

Now we fix the exposure time and let the number of counts become a Poisson random variable. The number of counts is now part of the data and gives us a little more information about the object than we have in the fixed-counts case. We will be concerned with the same issues in this section as in the previous one. We want to know the asymptotic form of the likelihood ratio, and we want to use the asymptotic analysis to develop SFOMs for ideal-observer performance.

A. Null Functions

As with the fixed-counts case, we need to know what the null functions are in order to proceed. The null space is less complicated for the fixed-time case and does not depend on the actual object.

The mean number of counts involves the exposure time τ :

$$\bar{J}(\mathbf{f}) = \tau \int_{\mathcal{S}} s(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} = \tau \mathbf{s}^\dagger \mathbf{f}. \quad (71)$$

The data are the list $\mathcal{A}$ as in Section 2, plus the total count J . We write $G = (\mathcal{A}, J)$ for the data. The conditional probability for the data given the object is given by

$$\text{pr}(G|\mathbf{f}) = \frac{[\bar{J}(\mathbf{f})]^J}{J!} \exp[-\bar{J}(\mathbf{f})] \prod_{j=1}^J \text{pr}(\hat{\mathbf{A}}_j|\mathbf{f}) = \text{Pr}(J|\mathbf{f}) \text{pr}(\hat{\mathcal{A}}|\mathbf{f}, J). \quad (72)$$

The conditions for a null function $\mathbf{f}_n$ at $\mathbf{f}_a$ are $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a + \mathbf{f}_n) = \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$ for all attribute vectors $\hat{\mathbf{A}}$, and $J(\mathbf{f}_a + \mathbf{f}_n) = J(\mathbf{f}_a)$. These conditions are satisfied if and only if $\mathbf{f}_n$ satisfies Eq. (12) for all $\hat{\mathbf{A}}$. The denominator in the conditional probability $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$ is fixed by the mean count and the time, so there is no scale invariance as there was for the fixed-count case. The null space is therefore independent of $\mathbf{f}_a$ for fixed time. This null space will be denoted by $\tilde{U}_0$ and its orthogonal complement by $\tilde{U}_1$.

B. Ideal Observer for Long Time Interval

The PDF for the data is given by the integral

$$\text{pr}(G) = \int_U \text{pr}(G|\mathbf{f}) \text{pr}(\mathbf{f}) d\mathbf{f}. \quad (73)$$

As with the fixed-count case, an approximation to the value of the likelihood ratio that is accurate with high probability is given by

$$\Lambda(G) \cong \frac{\int_U \text{Pr}(J|\mathbf{f}) \exp\{-JD_{\text{KL}}[\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a), \text{pr}(\hat{\mathbf{A}}|\mathbf{f})]\} \text{pr}(\mathbf{f} - \Delta\mathbf{f}) d\mathbf{f}}{\int_U \text{Pr}(J|\mathbf{f}) \exp\{-JD_{\text{KL}}[\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a), \text{pr}(\hat{\mathbf{A}}|\mathbf{f})]\} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (74)$$

Note that, for long exposure times, the mean $J(\bar{\mathbf{f}}_a)$ of the Poisson distribution for the count J is very large, and this means that, with very high probability, J itself is very large since the standard deviation of a Poisson distribution is the square root of the mean. This fact justifies the use of this approximation for the value of the likelihood ratio. As with the fixed-counts case, the quantity on the left in this relation is not itself a test statistic but represents with very high probability an approximate value for the likelihood ratio.

Using the same FIM $\mathbf{F}(\mathbf{f}_a)$ as before, we may also write an approximation to the approximation:

$$\Lambda(G) \cong \frac{\int_U \text{Pr}(J|\mathbf{f}) \exp\left\{-\frac{J}{2}(\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a)\right\} \text{pr}(\mathbf{f} - \Delta\mathbf{f}) d\mathbf{f}}{\int_U \text{Pr}(J|\mathbf{f}) \exp\left\{-\frac{J}{2}(\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a)\right\} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (75)$$

We will make use of this expression to examine the asymptotic form of the likelihood ratio.

C. Limiting Forms for Ideal Observer

Introducing the difference vector $\mathbf{f}_d$ as before, the denominator in the approximation to Λ in Eq. (75) can be written as

$$\int_U \text{Pr}(J|\mathbf{f}_d + \mathbf{f}_a) \exp\left[-\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_d\right] \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d. \quad (76)$$

By decomposing the object space U into the null space U_0 of $\mathbf{F}(\mathbf{f}_a)$ and its orthogonal complement U_1 , the denominator can be written

$$\int_{U_1} \exp \left[-\frac{J}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}(\mathbf{f}_a) \mathbf{f}_{d1} \right] \int_{U_0} \Pr(J|\mathbf{f}_{d0}+\mathbf{f}_{d1}+\mathbf{f}_a) \times \Pr(\mathbf{f}_{d0}+\mathbf{f}_{d1}+\mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}. \quad (77)$$

Note that U_0 is not the null space discussed in Subsection 3.A; it is the same U_0 as we used in the fixed-counts case.

By normalizing the Gaussian function in the numerator and the denominator, we can see that, as $J \rightarrow \infty$, with high probability, the value of the likelihood ratio converges to a limit:

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U_0} \Pr(J|\mathbf{f}_{d0}+\mathbf{f}_a-\Delta\mathbf{f}) \Pr(\mathbf{f}_{d0}+\mathbf{f}_a-\Delta\mathbf{f}) d\mathbf{f}_{d0}}{\int_{U_0} \Pr(J|\mathbf{f}_{d0}+\mathbf{f}_a) \Pr(\mathbf{f}_{d0}+\mathbf{f}_a) d\mathbf{f}_{d0}}. \quad (78)$$

The invariance of the subspace U_0 under translation by $\mathbf{f}_a$ reduces this expression to

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U_0} \Pr(J|\mathbf{f}_{d0}-\Delta\mathbf{f}) \Pr(\mathbf{f}_{d0}-\Delta\mathbf{f}) d\mathbf{f}_{d0}}{\int_{U_0} \Pr(J|\mathbf{f}_{d0}) \Pr(\mathbf{f}_{d0}) d\mathbf{f}_{d0}}. \quad (79)$$

Now we can again use the fact that we have a positive signal to write:

$$\Lambda(\widehat{\mathcal{A}}) \rightarrow \frac{\int_{U(\mathbf{f}_a)} \Pr(J|\mathbf{f}-\Delta\mathbf{f}) \Pr(\mathbf{f}-\Delta\mathbf{f}) d\mathbf{f}}{\int_{U(\mathbf{f}_a)} \Pr(J|\mathbf{f}) \Pr(\mathbf{f}) d\mathbf{f}}. \quad (80)$$

In the numerator and denominator of this last relation, we are again integrating over the space of object functions that satisfy $\Pr(\hat{\mathbf{A}}|\mathbf{f}) = \Pr(\hat{\mathbf{A}}|\mathbf{f}_a)$, but, in contrast to the fixed-counts case, the prior probability at each point in this space is now multiplied by the corresponding probability for the observed value of J . This is the likelihood ratio for the observer who has direct access to the normalized mean of the Poisson point process in attribute space that generates the list and to the number of entries in that list.

We can take this a step further and expand the probability for the number of counts to the second order about the actual object:

$$\Pr(J|\mathbf{f}) \cong \frac{1}{J!} \exp \left\{ -\bar{J}(\mathbf{f}_a) + J \ln[\bar{J}(\mathbf{f}_a)] + \left[\frac{J - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f} - \mathbf{f}_a) - \frac{J}{2\bar{J}^2(\mathbf{f}_a)} [\bar{J}(\mathbf{f} - \mathbf{f}_a)]^2 \right\}. \quad (81)$$

Define a modified Fisher matrix by the equation

$$\mathbf{F}_s(\mathbf{f}_a) = \mathbf{F}(\mathbf{f}_a) + \frac{1}{(\mathbf{s}^\dagger \mathbf{f}_a)^2} \mathbf{s} \mathbf{s}^\dagger = \left\langle \frac{\mathbf{p}(\hat{\mathbf{A}}) \mathbf{p}(\hat{\mathbf{A}})^\dagger}{[\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a]^2} \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a}. \quad (82)$$

After canceling some terms, we may write an approximation for the likelihood ratio:

$$\Lambda(G) \cong \frac{\int_U \exp \left\{ -\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_d + \left[\frac{J - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_d) \right\} \text{pr}(\mathbf{f}_d + \mathbf{f}_a - \Delta \mathbf{f}) d\mathbf{f}_d}{\int_U \exp \left\{ -\frac{J}{2} \mathbf{f}_d^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_d + \left[\frac{J - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_d) \right\} \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d}. \quad (83)$$

Decompose object space using the SVD of $\mathbf{F}_s(\mathbf{f}_a)$. A null vector $\mathbf{f}_n$ for this matrix satisfies

$$\left\langle \left[\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_n}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a^2} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} = 0, \quad (84)$$

which is true if and only if $\mathbf{f}_n$ is in $\tilde{U}_0$. Also note that $J(\mathbf{f}_n) = 0$ in this case. Let $\mathbf{f}_d = \tilde{\mathbf{f}}_{d0} + \tilde{\mathbf{f}}_{d1}$ be the corresponding decomposition for the difference vector. We may rewrite the denominator in this expression for the approximate likelihood ratio as

$$\int_{\tilde{U}_1} \exp \left\{ -\frac{J}{2} \tilde{\mathbf{f}}_{d1}^\dagger \mathbf{F}_s(\mathbf{f}_a) \tilde{\mathbf{f}}_{d1} + \left[\frac{J - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\tilde{\mathbf{f}}_{d1}) \right\} \int_{\tilde{U}_0} \text{pr}(\tilde{\mathbf{f}}_{d0} + \tilde{\mathbf{f}}_{d1} + \mathbf{f}_a) d\tilde{\mathbf{f}}_{d0} d\tilde{\mathbf{f}}_{d1}. \quad (85)$$

Note that, for a given $\tilde{\mathbf{f}}_{d1}$, the linear term in the exponent is of the order of $\tau^{1/2}$ as $\tau \rightarrow \infty$ since J is a Poisson random variable with mean $J(\mathbf{f}_a)$. The quadratic term, on the other hand, will be of the order of τ . Thus, the quadratic term will dominate in the limit and produce a delta function at zero when the Gaussian is normalized. By treating the numerator the same way, we find that

$$\Lambda(\hat{\mathcal{A}}) \rightarrow \frac{\int_{\tilde{U}_0} \text{pr}(\tilde{\mathbf{f}}_{d0} + \mathbf{f}_a - \Delta \mathbf{f}) d\tilde{\mathbf{f}}_{d0}}{\int_{\tilde{U}_0} \text{pr}(\tilde{\mathbf{f}}_{d0} + \mathbf{f}_a) d\tilde{\mathbf{f}}_{d0}} \quad (86)$$

as $\tau \rightarrow \infty$. We define the space $\tilde{U}(\mathbf{f}_a)$ to be those objects $\mathbf{f}$ that satisfy the two positivity constraints, $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}) = \text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$ for all $\hat{\mathbf{A}}$ and $J(\mathbf{f}) = J(\mathbf{f}_a)$. Now we can again use the fact that we have a positive signal to write:

$$\Lambda(\hat{\mathcal{A}}) \rightarrow \frac{\int_{\tilde{U}(\mathbf{f}_a)} \text{pr}(\mathbf{f} - \Delta \mathbf{f}) d\mathbf{f}}{\int_{\tilde{U}(\mathbf{f}_a)} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (87)$$

This is the likelihood ratio for the observer who has access to the normalized mean to the Poisson process for the data list and to the mean number of counts. By multiplying this function by this number, this observer has access to the mean

$$\bar{u}(\hat{\mathbf{A}}) = \int_S \text{pr}(\hat{\mathbf{A}}|\mathbf{r}) s(\mathbf{r}) f_a(\mathbf{r}) d\mathbf{r} \quad (88)$$

of the Poisson process that generates the data list. This function in attribute space contains all of the possible information that the imaging system can deliver about the object in question.

If we write the integral relationship in Eq. (88) as $\bar{u} = \mathcal{L}f_a$, we can formulate the condition under which the asymptotic ideal observer will have a finite value for detectability. In order

for this to occur, there must be a set $C \subset U$ of signal-absent objects such that the prior probability assigned to C is nonzero and for each object $f(\mathbf{r})$ in C there is another signal-free object $f'(\mathbf{r})$ such that $\mathcal{L}f = \mathcal{L}f' + \mathcal{L}\Delta f$. If the function $\mathcal{L}f$ is observed by the asymptotic ideal observer, both $f_a = f$ and $f_a = f' + \Delta f$ are possibilities for the actual object. In this case, the numerator and denominator of the likelihood ratio in Eq. (87) are nonzero numbers, and the decision of asymptotic ideal observer will depend on the threshold. Since the set C has finite prior probability, there is a finite probability that the asymptotic ideal observer will make the wrong decision, at least for thresholds in a certain range of values. Thus, the detectability for this observer will be finite. If the set of confusion C is empty, then the asymptotic ideal observer will always make the correct decision, and the detectability will be infinite. In terms of the likelihood ratio in Eq. (87), this means that the numerator will be zero when the signal is absent and the denominator zero when the signal is present. The condition for finite detectability can also be stated in object space: there is a set of signal-free objects $C \subset U$ with finite prior probability such that, for any $f \in C$, there is a signal-free object f' and a null function f_n satisfying $f + f_n = f' + \Delta f$.

D. SFOMs

The SFOMs will again be based on the Fisher information approximation to the ideal-observer detectability. As in the fixed-counts case, we introduce amplitude parameter α ,

$$\text{pr}(G|\alpha) = \int_U \text{pr}(G|\mathbf{f} + \alpha\Delta\mathbf{f}) \text{pr}(\mathbf{f}) d\mathbf{f}, \quad (89)$$

and define $\text{pr}(G) = \text{pr}(G|0)$. The score for the amplitude parameter at $\alpha = 0$ is defined as

$$\text{sc}(G) = \frac{d}{d\alpha} \ln[\text{pr}(G|\alpha)] \Big|_{\alpha=0}, \quad (90)$$

and a short calculation leads to the integral

$$\text{sc}(G) = \int_U \Delta\mathbf{f}^\dagger \text{sc}(G|\mathbf{f}) \text{pr}(\mathbf{f}|G) d\mathbf{f}. \quad (91)$$

As before, this integral involves the local score, $\text{sc}(G|\mathbf{f}) = \nabla_{\mathbf{f}} \ln[\text{pr}(G|\mathbf{f})]$, and the posterior PDF $\text{pr}(\mathbf{f}|G)$. The global score is again the PM of the local score.

Now we can calculate the local score and get an expression for the score:

$$\text{sc}(G) = \int_U \Delta\mathbf{f}^\dagger \left[\sum_{j=1}^J \text{sc}(\hat{\mathbf{A}}_j|\mathbf{f}) + \frac{J\tau}{J(\mathbf{f})} \mathbf{s} - \tau \mathbf{s} \right] \text{pr}(\mathbf{f}|G) d\mathbf{f}. \quad (92)$$

By simplifying this expression, we arrive at

$$\text{sc}(G) = \int_U \Delta\mathbf{f}^\dagger \left[\sum_{j=1}^J \frac{\mathbf{p}(\hat{\mathbf{A}}_j)}{p(\hat{\mathbf{A}}_j)^\dagger \mathbf{f}} \right] \text{pr}(\mathbf{f}|G) d\mathbf{f} - \tau \Delta\mathbf{f}^\dagger \mathbf{s}. \quad (93)$$

The FIM approximation to ideal-observer detectability as a function of amplitude is then given by $d^2(a) \approx a^2 \langle [\text{sc}(G)]^2 \rangle_G$. We can also write this SFOM in terms of nested expectations as

$$d^2(\alpha) \cong \alpha^2 \langle [\text{sc}(G)]^2 \rangle_{G|\mathbf{f}_a} \cdot \quad (94)$$

As we did for the fixed-counts case, we will seek to find asymptotic approximations to the expectation on the right in this relation.

1. Fixed-Time SFOM 1—For this approximation, we replace the sum in the score with an expectation:

$$\text{sc}(G) \cong J \int_U \left\langle \frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}} \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} \text{pr}(\mathbf{f}|G) d\mathbf{f} - \tau \Delta \mathbf{f}^\dagger \mathbf{s}. \quad (95)$$

Interchanging the expectation with the integral gives us

$$\text{sc}(G) \cong J \left\langle \int_U \left[\frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}} \right] \text{pr}(\mathbf{f}|G) d\mathbf{f} \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} - \tau \Delta \mathbf{f}^\dagger \mathbf{s}. \quad (96)$$

Once again the integral represents a PM estimate of the quantity in the square brackets. To avoid MCMC computations, we could replace the PM estimate with an ML estimate and use

$$\text{sc}(G) \cong J \left\langle \frac{\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_{\text{ML}}(G)} \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} - \tau \Delta \mathbf{f}^\dagger \mathbf{s} \quad (97)$$

in the expression for the approximate ideal-observer detectability. Computation of this approximation would then require only Monte Carlo evaluations of the expectations and an algorithm for finding an ML estimate of the object. The covariance of the score then gives the Fisher information.

2. Fixed-Time SFOM 2—For this approximation, we use the unsimplified form for the score and replace the sum with an expectation:

$$\text{sc}(G) \cong J \int_U \Delta \mathbf{f}^\dagger \left\langle \text{sc}(\hat{\mathbf{A}}|\mathbf{f}) \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} \text{pr}(\mathbf{f}|G) d\mathbf{f} + \left[J \int_U \frac{\text{pr}(\mathbf{f}|G)}{\mathbf{s}^\dagger \mathbf{f}} d\mathbf{f} - \tau \right] \mathbf{s}^\dagger \Delta \mathbf{f}. \quad (98)$$

Now we follow the same steps that we used in the fixed-counts case to approximate the expectation in the first term of this expression with a quadratic function. The end result is that the score can be written in terms of the PM estimate of the object and what can be viewed as a PM estimate of the exposure time:

$$\text{sc}(G) \cong -J \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) [\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] + [\hat{\tau}_{\text{PM}}(G) - \tau] \Delta \mathbf{f}^\dagger \mathbf{s}. \quad (99)$$

The PM estimate of the exposure time is the integral

$$\hat{\tau}_{\text{PM}}(G) = \int_U \frac{J}{\mathbf{s}^\dagger \mathbf{f}} \text{Pr}(\mathbf{f}|G) d\mathbf{f}. \quad (100)$$

Of course we do not need to estimate the exposure time since we measure it, but since the actual exposure time is given by

$$\tau = \frac{\overline{J}(\mathbf{f}_a)}{\mathbf{s}^\dagger \mathbf{f}_a}, \quad (101)$$

this PM expression is not an unreasonable estimate of τ if it was not measured. We could approximate this estimate of the exposure time by

$$\hat{\tau}_{\text{PM}}(G) \cong \frac{J}{\mathbf{s}^\dagger \mathbf{f}_{\text{PM}}(G)}. \quad (102)$$

Since we want to avoid MCMC calculations, we will replace PM estimates with ML estimates to produce an approximate score to use in our SFOM:

$$\text{sc}(G) \cong -J \Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) [\hat{\mathbf{f}}_{\text{ML}}(G) - \mathbf{f}_a] + [\hat{\tau}_{\text{ML}}(G) - \tau] \Delta \mathbf{f}^\dagger \mathbf{s}, \quad (103)$$

where

$$\hat{\tau}_{\text{ML}}(G) = \frac{J}{\mathbf{s}^\dagger \mathbf{f}_{\text{ML}}(G)}. \quad (104)$$

The use of the ML estimate in the first term of the approximate score can actually be rigorously justified in the limit as $\tau \rightarrow \infty$ because, as we will see below, the PM estimate of the object and the ML estimate of the object differ by a null vector of $\mathbf{F}(\mathbf{f}_a)$ in this limit. As in the first approximation, the resulting approximation to the ideal-observer detectability gives an SFOM that requires only Monte Carlo evaluations of expectations and an ML reconstruction algorithm.

3. Fixed-Time SFOM 3—For this approximation, we expand the second term in the score approximation to the first order about the actual object:

$$[\hat{\tau}_{\text{PM}}(G) - \tau] \cong \frac{J}{\mathbf{s}^\dagger \mathbf{f}_{\text{PM}}(G)} - \tau \cong \frac{J}{\mathbf{s}^\dagger \mathbf{f}_a} - \tau - \frac{J \mathbf{s}^\dagger}{(\mathbf{s}^\dagger \mathbf{f}_a)^2} [\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a]. \quad (105)$$

Now the score is approximated by

$$\text{sc}(G) \cong -J \Delta \mathbf{f}^\dagger \mathbf{F}_s(\mathbf{f}_a) [\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] + \left(\frac{J}{\mathbf{s}^\dagger \mathbf{f}_a} - \tau \right) \Delta \mathbf{f}^\dagger \mathbf{s}. \quad (106)$$

To proceed further, we use the approximation

$$\mathbf{F}_s(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] \cong \mathbf{F}_s(\mathbf{f}_a) \frac{\int_U (\mathbf{f} - \mathbf{f}_a) \exp \left\{ -\frac{j}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}_s(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f} - \mathbf{f}_a) \right\} \text{pr}(\mathbf{f}) d\mathbf{f}}{\int_U \exp \left[-\frac{j}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}_s(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f} - \mathbf{f}_a) \right] \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (107)$$

With the difference vector $\mathbf{f}_d$ as above, we can write this as

$$\mathbf{F}_s(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] \cong \mathbf{F}_s(\mathbf{f}_a) \frac{\int_U \mathbf{f}_d \exp \left\{ -\frac{j}{2} \mathbf{f}_d^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_d + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_d) \right\} \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d}{\int_U \exp \left[-\frac{j}{2} \mathbf{f}_d^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_d + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_d) \right] \text{pr}(\mathbf{f}_d + \mathbf{f}_a) d\mathbf{f}_d}. \quad (108)$$

Again, both integrals in the curly brackets are inner products between a shifted prior PDF and a function whose Fourier transform is known. This inner product could be performed in Fourier space if the characteristic function for the prior PDF is known and the pseudoinverse of $\mathbf{F}_s(\mathbf{f}_a)$ can be computed. This approximation should be good for large τ . What happens in the limit as $\tau \rightarrow \infty$?

Using the decomposition of object space into $\tilde{U}_0$ and $\tilde{U}_1$, we can write

$$\mathbf{F}_s(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] \cong \mathbf{F}_s(\mathbf{f}_a) \frac{\int_{\tilde{U}_1} \mathbf{f}_{d1} \exp \left\{ -\frac{j}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_{d1} + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_{d1}) \right\} \int_{\tilde{U}_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}{\int_{\tilde{U}_1} \exp \left\{ -\frac{j}{2} \mathbf{f}_{d1}^\dagger \mathbf{F}_s(\mathbf{f}_a) \mathbf{f}_{d1} + \left[\frac{j - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f}_{d1}) \right\} \int_{\tilde{U}_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_{d1} + \mathbf{f}_a) d\mathbf{f}_{d0} d\mathbf{f}_{d1}}. \quad (109)$$

By normalizing the Gaussians and letting $\tau \rightarrow \infty$, we find that $\mathbf{F}_s(\mathbf{f}_a)[\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a] \rightarrow 0$ independently of G .

We cannot say that the PM estimator is asymptotically unbiased due to the existence of null functions. It does however follow from this discussion that, as $\tau \rightarrow \infty$, we have, for all $\hat{\mathbf{A}}$, the relation $\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_{\text{PM}} \rightarrow \mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a$. This last equation is as close as we can get to an asymptotic unbiased condition for the fixed-time case.

We find as in the fixed-counts case that the asymptotic form as $\tau \rightarrow \infty$ for the PM estimator itself is determined by

$$\hat{\mathbf{f}}_{\text{PM}}(G) - \mathbf{f}_a \rightarrow \frac{\int_{\tilde{U}_0} \mathbf{f}_{d0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a) d\mathbf{f}_{d0}}{\int_{\tilde{U}_0} \text{pr}(\mathbf{f}_{d0} + \mathbf{f}_a) d\mathbf{f}_{d0}}. \quad (110)$$

This shows that the PM estimator is biased asymptotically. The bias is given by the null function on the right in this relation. Note that the bias depends on the actual object $\mathbf{f}_a$. As with the fixed-count case, we may also write this relation in the form

$$\hat{\mathbf{f}}_{\text{PM}}(\hat{\mathcal{A}}) \rightarrow \frac{\int_{\tilde{U}(\mathbf{f}_a)} \mathbf{f} \text{pr}(\mathbf{f}) d\mathbf{f}}{\int_{\tilde{U}(\mathbf{f}_a)} \text{pr}(\mathbf{f}) d\mathbf{f}}. \quad (111)$$

This shows that the PM estimator asymptotically depends only on the prior PDF $\text{pr}(\mathbf{f})$, the conditional PDF $\text{pr}(\hat{\mathbf{A}}|\mathbf{f}_a)$, and the exposure time τ , which is again what we expect.

In the asymptotic limit, we are using the fact that

$$\text{pr}(G|\mathbf{f}) \rightarrow \frac{1}{J!} \exp \left\{ -\bar{J}(\mathbf{f}_a) + J \ln[\bar{J}(\mathbf{f}_a)] - \frac{J}{2} (\mathbf{f} - \mathbf{f}_a)^\dagger \mathbf{F}_s(\mathbf{f}_a) (\mathbf{f} - \mathbf{f}_a) + \left[\frac{J - \bar{J}(\mathbf{f}_a)}{\bar{J}(\mathbf{f}_a)} \right] \bar{J}(\mathbf{f} - \mathbf{f}_a) \right\}. \quad (112)$$

Therefore, any ML estimate must satisfy $\mathbf{f}_{\text{ML}} \rightarrow \mathbf{f}_a + \mathbf{f}_n$ as $\tau \rightarrow \infty$ where $\mathbf{f}_n$ is a null function. This shows that the PM estimate of the object and the ML estimate differ by a null function in this limit.

4. Fixed-Time SFOM 4—For this last SFOM for fixed counts, we go back to SFOM 2 in Eq. (103) and use asymptotic properties of ML estimators to simplify the expression for approximate detectability. An ML estimator is a solution to the equation

$$\left(\frac{J}{\mathbf{s}^\dagger \hat{\mathbf{f}}_{\text{ML}}} - \tau \right) \mathbf{s} + \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j | \hat{\mathbf{f}}_{\text{ML}})] = 0. \quad (113)$$

Using Taylor's theorem, we can expand to the first order about $\mathbf{f}_a$ and write this as

$$\left(\frac{J}{\mathbf{s}^\dagger \mathbf{f}_a} - \tau \right) \mathbf{s} + \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j | \mathbf{f}_a)] + \left\{ -\frac{J}{(\mathbf{s}^\dagger \mathbf{f}_\beta)^2} \mathbf{s} \mathbf{s}^\dagger + \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j | \mathbf{f}_\beta)] \right\} (\hat{\mathbf{f}}_{\text{ML}} - \mathbf{f}_a) \approx 0, \quad (114)$$

where $\mathbf{f}_\beta = \beta \hat{\mathbf{f}}_{\text{ML}} + (1 - \beta) \mathbf{f}_a$ and $0 < \beta < 1$ as before.

Define a random vector $\mathbf{y}$ as before via the equation

$$\mathbf{y} = \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j | \mathbf{f}_a)]. \quad (115)$$

For a fixed J , we know that the mean of this vector is $\mathbf{0}$, and therefore the mean remains $\mathbf{0}$ when the averaging over J is included. For fixed J the vector $\mathbf{y}$ has covariance $J\mathbf{F}(\mathbf{f}_a)$. After averaging over J , we find that the overall covariance of $\mathbf{y}$ is $J\mathbf{F}(\mathbf{f}_a)$. For large exposure time τ , the count J is also large with high probability. Using the law of large numbers, we get, with high probability,

$$\frac{1}{J} \left\{ \frac{J}{(\mathbf{s}^\dagger \mathbf{f}_a)^2} \mathbf{s} \mathbf{s}^\dagger - \sum_{j=1}^J \nabla_{\mathbf{f}} \ln[\text{pr}(\hat{\mathbf{A}}_j | \mathbf{f}_\beta)] \right\} \cong \mathbf{F}_s(\mathbf{f}_\beta) \cong \mathbf{F}_s(\mathbf{f}_a). \quad (116)$$

The second approximate equality here follows from the fact noted above that, as $J \rightarrow \infty$, $\hat{\mathbf{f}}_{\text{ML}} \rightarrow \mathbf{f}_a + \mathbf{f}_n$, where $\mathbf{f}_n$ is a null function. Since $\mathbf{f}_\beta$ is between $\mathbf{f}_a$ and $\hat{\mathbf{f}}_{\text{ML}}$, it will have the same asymptotic limit.

Putting these facts together, we have, for large τ , the approximation

$$\mathbf{y} \cong J\mathbf{F}_s(\mathbf{f}_a) (\hat{\mathbf{f}}_{\text{ML}} - \mathbf{f}_a) - \left(\frac{J}{\mathbf{s}^\dagger \mathbf{f}_a} - \tau \right) \mathbf{s}, \quad (117)$$

and the covariance of this vector is approximately $\bar{\mathbf{J}}\mathbf{F}(\mathbf{f}_a)$. By taking this relation together with Eq. (106) and the fact that the ML estimate of the object and the PM estimate differ by a null function in the asymptotic limit, we get

$$d^2(\alpha) \cong \alpha^2 \tau \langle (\mathbf{s}^\dagger \mathbf{f}_a) (\Delta \mathbf{f}^\dagger \mathbf{F}(\mathbf{f}_a) \Delta \mathbf{f}) \rangle_{\mathbf{f}_a}. \quad (118)$$

Note the similarity with SFOM 4 from the fixed-counts case. If we write this approximation out explicitly, we have a convenient expression for this SFOM:

$$d^2(\alpha) \cong \alpha^2 \tau \left\langle (\mathbf{s}^\dagger \mathbf{f}_a) \left\{ \left\langle \left[\frac{\mathbf{p}(\hat{\mathbf{A}}^\dagger \Delta \mathbf{f})}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} \right]^2 \right\rangle_{\hat{\mathbf{A}}|\mathbf{f}_a} - \left(\frac{\mathbf{s}^\dagger \Delta \mathbf{f}}{\mathbf{s}^\dagger \mathbf{f}_a} \right)^2 \right\} \right\rangle_{\mathbf{f}_a}. \quad (119)$$

Once again we have the problem that the detectability increases without bound as the exposure time increases, which we know cannot be the case if the asymptotic ideal observer has finite detectability. It must be the case that some or all of the approximations that lead to this final approximation to the detectability become worse as the exposure time increases, but we have not been able to determine where the error enters into the calculation.

If we use Eq. (1), we may rewrite this SFOM in the more analytic form,

$$d^2(\alpha) \cong \alpha^2 \tau \left\langle \int_L \frac{[\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}]^2}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} d\hat{\mathbf{A}} - \frac{(\mathbf{s}^\dagger \Delta \mathbf{f})^2}{\mathbf{s}^\dagger \mathbf{f}_a} \right\rangle_{\mathbf{f}_a}. \quad (120)$$

As with the fixed-counts SFOM 4, if the integral over attribute space can be performed, then we only need to sample from the prior to evaluate this quantity.

4. CONCLUSION

In view of the similarity between the result in Eq. (70) and the result in Eq. (120), we propose that the most useful SFOM to investigate at this time is the quantity

$$\text{SFOM4} = \left\langle \int_L \frac{[\mathbf{p}(\hat{\mathbf{A}})^\dagger \Delta \mathbf{f}]^2}{\mathbf{p}(\hat{\mathbf{A}})^\dagger \mathbf{f}_a} d\hat{\mathbf{A}} - \frac{(\mathbf{s}^\dagger \Delta \mathbf{f})^2}{\mathbf{s}^\dagger \mathbf{f}_a} \right\rangle_{\mathbf{f}_a}.$$

There are several advantages that this SFOM possesses over other SFOMs discussed in this paper and in the work on SFOMs for binned data in [2]. The first advantage is that the computation of SFOM 4 does not require any reconstructions. SFOMs 1 and 2 in this paper and all of the SFOMs in [2] require an object reconstruction from simulated data for each sample drawn from the object ensemble. The second advantage for SFOM 4 is that the integral is over the low-dimensional attribute space as opposed to SFOM 3 for fixed-count and fixed-time cases, which requires an integral over the large-dimensional object space. Another advantage for SFOM 4 is that it is nearly identical to an approximation to ideal-observer detectability derived in [3] by a completely different approach. In that paper this approximation was shown to correlate more closely with ideal-observer performance as calculated by MCMC methods [9] than the Hotelling signal-to-noise ratio. The Hotelling

observer is the ideal observer when all of the statistics are Gaussian, and the Hotelling SNR is equal to the ideal-observer detectability in that case.

Our next step is therefore to compute SFOM 4 in simulations where the ideal-observer detectability can be directly calculated. By varying background, signal, and system parameters, we can then observe the degree of correlation between SFOM 4 and ideal-observer detectability. If these studies have positive results, then we can start applying SFOM 4 to the optimization of list-mode imaging systems. For systems in the design stage, SFOM 4 could be used to rapidly assess various design options in terms of signal detectability. Since SFOM 4 is based on the FIM, these systems would be simultaneously optimized for estimating signal amplitude. For adaptive imaging systems [10], the initial or scout scan can be used to reduce the object ensemble to those objects compatible with the scout data. Then the average over objects in SFOM 4 can be performed over this reduced ensemble to select the optimum system configuration for the diagnostic scan.

Acknowledgments

The author wishes to acknowledge Harrison Barrett and Matthew Kupinski for their help with this work, and National Institutes of Health grants R37 EB000803, P41 EB002035, and RC1 EB010974 for financial support.

References

1. 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.
2. Clarkson E, Shen F. Fisher information and surrogate figures of merit for the task-based assessment of image quality. *J Opt Soc Am A*. 2010; 27:2313–2326.
3. Caucci L, Barrett HH. Objective assessment of image quality. V. Photon counting detectors and list-mode data. *J Opt Soc Am A*. 2012; 29:1003–1016.
4. Barrett HH, White T, Parra LC. List-mode likelihood. *J Opt Soc Am A*. 1997; 14:2914–2923.
5. Shao, J. *Mathematical Statistics*. Springer: 1999.
6. Kay, SM. *Fundamentals of Statistical Signal Processing II: Detection Theory*. Prentice Hall; 2008.
7. Shen F, Clarkson E. Using Fisher information to approximate ideal-observer performance on detection tasks for lumpy-background images. *J Opt Soc Am A*. 2006; 23:2406–2414.
8. Cover, TM.; Thomas, JA. *Elements of Information Theory*. Wiley; 1991.
9. 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.
10. Clarkson E, Kupinski MA, Barrett HH, Furenlid L. A task-based approach to adaptive and multimodality imaging. *Proc IEEE*. 2008; 96:500–511.