

Using Fisher information to approximate ideal-observer performance on detection tasks for lumpy-background images

Fangfang Shen and Eric Clarkson

Department of Radiology, University of Arizona, P.O. Box 245067, Tucson, Arizona 85724-5067

Received September 2, 2005; revised April 13, 2006; accepted May 11, 2006; posted May 17, 2006 (Doc. ID 64596)

When building an imaging system for detection tasks in medical imaging, we need to evaluate how well the system performs before we can optimize it. One way to do the evaluation is to calculate the performance of the Bayesian ideal observer. The ideal-observer performance is often computationally expensive, and it is very useful to have an approximation to it. We use a parameterized probability density function to represent the corresponding densities of data under the signal-absent and the signal-present hypotheses. We develop approximations to the ideal-observer detectability as a function of signal parameters involving the Fisher information matrix, which is normally used in parameter estimation problems. The accuracy of the approximation is illustrated in analytical examples and lumpy-background simulations. We are able to predict the slope of the detectability as a function of the signal parameter. This capability suggests that the Fisher information matrix itself evaluated at the null parameter value can be used as the figure of merit in imaging system evaluation. We are also able to provide a theoretical foundation for the connection between detection tasks and estimation tasks. © 2006 Optical Society of America

OCIS codes: 110.2960, 110.3000, 110.4280.

1. INTRODUCTION

In medical imaging, two of the most important tasks are signal classification and signal estimation. Signal detection, such as finding tumors and looking for bone fractures, belongs to the classification category. In this paper, our study focuses on signal detection because it is a very important part of medical imaging. When building an imaging system for detection tasks, we need to evaluate how well the system performs before the system can be optimized. One way to evaluate a system is to calculate the performance of the Bayesian ideal observer on a detection task. The Bayesian ideal observer uses all of the information available to make decisions and is optimal by several criteria,¹ including the area under the receiver operating characteristic (ROC) curve (AUC). Because the ideal observer uses the raw data, its performance is not affected by the reconstruction methods. Therefore it is very useful in optimizing imaging system hardware. The AUC of an ideal observer is often not easy to compute, and it is quite useful to have an approximation to it. There are two hypotheses to be considered in detection tasks, signal-absent and signal-present. In this paper, a single data probability density function with a variable scalar or vector parameter is used to represent the corresponding densities under the two hypotheses. For example, consider the probability density function of data vector \mathbf{g} with a parameter θ , where θ represents the intensity of the signal. When $\theta=0$, $\text{pr}(\mathbf{g}|\theta)$ represents the conditional probability density function of \mathbf{g} under the signal-absent hypothesis; when $\theta>0$, it represents the density function under the signal-present hypothesis. We derive approximations to the ideal-observer AUC by using the Fisher information matrix and illustrate their accuracy in some ex-

amples. In the lumpy object simulation, the Fisher information approach requires Markov chain Monte Carlo (MCMC) simulation only for images with one parameter value θ_0 ; then the detectability can be extrapolated, and a multitude of signal values can be extrapolated. On the other hand, the direct detectability estimation requires MCMC simulation at each signal value of interest. Barrett and Clarkson,^{2,3} Hanson,⁴ and Moore *et al.*⁵ have shown the close relationship between detection tasks and estimation tasks. Müller *et al.*⁶ have shown some important results on the maximum likelihood estimator. Since the Fisher information matrix is often useful in signal estimation tasks, our approximations provide a firmer foundation for the connection between the two tasks.

This study makes use of the concepts related to the ROC curve, the AUC, detectability, the likelihood-generating function, and Fisher information, which are introduced in Section 2. In Section 3, we show that the likelihood-generating function and the ideal-observer AUC can be made dependent on a signal parameter. This allows us to derive our primary result, an approximation of the ideal-observer detectability that is explicitly dependent on the Fisher information matrix. In Sections 4 and 5, the accuracy of our approximation is illustrated in analytical examples and numerical simulations. In Section 6, we present our conclusions.

2. BACKGROUND

We represent the data set as an $M \times 1$ column vector \mathbf{g} . In signal detection tasks, for each vector, we must decide between two hypotheses, signal-absent H_0 and signal-present H_1 , using a test statistic t . The test statistic maps

the data vector to a real number. When t is greater than a certain threshold x , the decision is positive, or signal-present. Otherwise, the decision is negative, or signal-absent. The true positive fraction (TPF),¹ which is the fraction of positive decisions while the signal is present in the data, is defined as

$$\text{TPF}(x) = \text{pr}(t \geq x|H_1) = \int_x^\infty \text{pr}(t|H_1)dt. \quad (2.1)$$

The false positive fraction (FPF) is defined similarly as

$$\text{FPF}(x) = \text{pr}(t \geq x|H_0) = \int_x^\infty \text{pr}(t|H_0)dt. \quad (2.2)$$

The trade-off between TPF and FPF is controlled by the threshold x and can be illustrated by the ROC curve, which is a plot of $\text{TPF}(x)$ versus $\text{FPF}(x)$. The curve starts at (1,1) and ends at (0,0) as the threshold increases from $-\infty$ to ∞ . The AUC (Ref. 7) is defined as

$$\text{AUC} = \int_0^1 \text{TPF} d(\text{FPF}) = - \int_{-\infty}^\infty dx \text{TPF}(x) \frac{d}{dx} \text{FPF}(x). \quad (2.3)$$

An ideal observer uses the likelihood ratio given below as the test statistic to maximize the AUC. The likelihood ratio¹ is defined as

$$\Lambda(\mathbf{g}) = \frac{\text{pr}(\mathbf{g}|H_1)}{\text{pr}(\mathbf{g}|H_0)}. \quad (2.4)$$

Barrett and Clarkson¹ derived an approximation of the ideal-observer AUC shown below:

$$\text{AUC}_\Lambda \approx \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{1}{2}\sqrt{2G(0)}\right), \quad (2.5)$$

where

$$G(\beta) = \frac{\ln\langle\Lambda^{\beta+1/2}\rangle_0}{(\beta+1/2)(\beta-1/2)} \quad (2.6)$$

is called the likelihood-generating function. Clarkson and Barrett⁸ derived an exact integral equation for the ideal-observer AUC using moments of the likelihood ratio:

$$\text{AUC}_\Lambda = 1 - \frac{1}{2\pi} \int_0^\infty \langle\Lambda^{1/2+i\alpha}\rangle_0 \langle\Lambda^{1/2-i\alpha}\rangle_0 \frac{d\alpha}{\alpha^2 + 1/4} \quad (2.7)$$

where $\langle\cdot\rangle_0$ represents the conditional mean under the signal-absent hypothesis. The conditional mean is calculated by averaging over the statistical noise realizations in an ensemble of signal-absent images, including the background variability. The detectability d_A of the observer can also be used as the figure of merit. It is related to the AUC by the formula

$$\text{AUC}_\Lambda = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{1}{2}d_A\right), \quad (2.8)$$

and it can be approximated by the likelihood-generating function as

$$d_A \approx \sqrt{2G(0)}. \quad (2.9)$$

Equation (2.7), which is used to derive our approximations of the ideal-observer performance, shows the connection between the ideal-observer AUC and the moments of Λ .

When θ is the parameter describing the signal, the Fisher information matrix $\mathbf{F}=(F_{ij})$ as a function of θ evaluated at θ_0 is used in our approximation, where θ_0 (the null parameter) denotes the absence of a signal. The entries of the matrix are defined as

$$F_{ij}(\theta_0) = \left\langle \frac{1}{\text{pr}^2(\mathbf{g}|\theta)} \frac{\partial \text{pr}(\mathbf{g}|\theta)}{\partial \theta_i} \frac{\partial \text{pr}(\mathbf{g}|\theta)}{\partial \theta_j} \right\rangle_{\theta_0}.$$

The conditional mean $\langle\cdot\rangle_{\theta_0}$ is computed by averaging over an ensemble of noisy signal-absent data, and for a generic function $f(\theta)$ such a conditional mean is defined as

$$\langle f(\theta) \rangle_{\theta_0} = \int f(\theta) \text{pr}(\mathbf{g}|\theta_0) d\mathbf{g}.$$

This definition is useful for deriving many equations in this paper. In estimation theory, the Fisher information matrix is used to calculate the lower bound of the variance of the uniformly minimum variance unbiased estimator (UMVUE). This bound is inversely proportional to the norm of the Fisher information matrix.

3. APPROXIMATIONS TO THE IDEAL-OBSERVER PERFORMANCE

A. Fisher Information Approximation

Consider the probability density function $\text{pr}(\mathbf{g}|\theta)$ with a variable parameter θ . For example, θ can represent the intensity or the size of the signal in the detection task. When $\theta=\theta_0$, $\text{pr}(\mathbf{g}|\theta_0)$ represents the probability density function under the signal-absent hypothesis. The likelihood ratio for two different values of θ is given by

$$\Lambda_\theta(\mathbf{g}) = \frac{\text{pr}(\mathbf{g}|\theta)}{\text{pr}(\mathbf{g}|\theta_0)}, \quad (3.1)$$

and the corresponding likelihood-generating function is given by

$$G_\theta(\beta) = \frac{\ln\langle\Lambda_\theta^{\beta+1/2}\rangle_{\theta_0}}{(\beta+1/2)(\beta-1/2)} \quad (3.2)$$

Therefore, by substituting Eq. (3.1) into Eq. (2.7), the ideal-observer AUC can be written as

$$\text{AUC}_\Lambda(\theta) = 1 - \frac{1}{2\pi} \int_0^\infty \langle\Lambda_\theta^{1/2+i\alpha}\rangle_{\theta_0} \langle\Lambda_\theta^{1/2-i\alpha}\rangle_{\theta_0} \frac{d\alpha}{\alpha^2 + 1/4}. \quad (3.3)$$

Clarkson and Barrett⁸ derived the following equation, which relates the AUC and the detectability:

$$\text{AUC}_\Lambda(\boldsymbol{\theta}) = 1 - \frac{1}{2\pi} \int_0^\infty \exp\left[-\left(\alpha^2 + \frac{1}{4}\right)d_\Lambda^2(\boldsymbol{\theta})\right] \frac{d\alpha}{\alpha^2 + 1/4}. \quad (3.4)$$

To simplify the notation, we make the following substitutions. Let $\gamma(\boldsymbol{\theta}) = d_\Lambda^2(\boldsymbol{\theta})$ and $\mu_\theta(\alpha) = \langle \Lambda_\theta^{1/2+i\alpha} \rangle_{\theta_0} \langle \Lambda_\theta^{1/2-i\alpha} \rangle_{\theta_0}$. By making these substitutions, the following integral equation is obtained from Eqs. (3.3) and (3.4):

$$\int_0^\infty \left\{ \exp\left[-\left(\alpha^2 + \frac{1}{4}\right)\gamma(\boldsymbol{\theta})\right] - \mu_\theta(\alpha) \right\} \frac{d\alpha}{\alpha^2 + 1/4} = 0. \quad (3.5)$$

Taking the first derivative of Eq. (3.5) gives

$$\frac{\partial}{\partial \theta_i} \gamma(\boldsymbol{\theta}) = - \frac{\int_0^\infty \frac{\partial}{\partial \theta_i} \mu_\theta(\alpha) \frac{d\alpha}{\alpha^2 + 1/4}}{\int_0^\infty \exp[-(\alpha^2 + 1/4)\gamma(\boldsymbol{\theta})] d\alpha}. \quad (3.6)$$

The derivative in the numerator of Eq. (3.6) can be expanded as

$$\begin{aligned} \frac{\partial}{\partial \theta_i} \mu_\theta(\alpha) &= \left(\frac{1}{2} + i\alpha\right) \\ &\times \left\langle \Lambda_\theta^{-\frac{1}{2}+i\alpha} \frac{1}{\text{pr}(\mathbf{g}|\boldsymbol{\theta}_0)} \left[\frac{\partial \text{pr}(\mathbf{g}|\boldsymbol{\theta})}{\partial \theta_i} \right] \right\rangle_{\theta_0} \langle \Lambda_\theta^{1/2-i\alpha} \rangle_{\theta_0} \\ &+ \text{c. c.}, \end{aligned} \quad (3.7)$$

where c.c. represents the corresponding complex conjugate term. When evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, it becomes

$$\begin{aligned} \left. \frac{\partial}{\partial \theta_i} \mu_\theta(\alpha) \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= \left(\frac{1}{2} + i\alpha\right) \\ &\times \left\langle \frac{1}{\text{pr}(\mathbf{g}|\boldsymbol{\theta}_0)} \left[\frac{\partial \text{pr}(\mathbf{g}|\boldsymbol{\theta})}{\partial \theta_i} \right] \right\rangle_{\theta_0} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \text{c. c.} \\ &= \left(\frac{1}{2} + i\alpha\right) \frac{\partial}{\partial \theta_i} \left\langle \frac{\text{pr}(\mathbf{g}|\boldsymbol{\theta})}{\text{pr}(\mathbf{g}|\boldsymbol{\theta}_0)} \right\rangle_{\theta_0} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \text{c. c.} \end{aligned} \quad (3.8)$$

Note that the conditional mean in the second line of Eq. (3.8) equals unity regardless of the value of $\boldsymbol{\theta}$. Therefore the numerator of Eq. (3.6) equals zero when evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. We also have $\gamma(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0$ from its definition. As $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_0$, the numerator of Eq. (3.6) approaches zero, while the denominator increases without bound. These facts imply that

$$\left. \frac{\partial}{\partial \theta_i} \gamma(\boldsymbol{\theta}) \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0. \quad (3.9)$$

We expect the first derivatives of $\gamma(\boldsymbol{\theta})$ to be zero when evaluated at $\boldsymbol{\theta}_0$ because $\gamma(\boldsymbol{\theta})$ is nonnegative, and it reaches its minimum value at $\boldsymbol{\theta}_0$. By taking two derivatives of the first and the second term in the integrand of

Eq. (3.5), respectively, and evaluating the results at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, the following two equations are derived:

$$\begin{aligned} &\left. \frac{\partial^2}{\partial \theta_i \partial \theta_j} \exp\left[-\left(\alpha^2 + \frac{1}{4}\right)\gamma(\boldsymbol{\theta})\right] \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ &= -\left(\alpha^2 + \frac{1}{4}\right) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \gamma(\boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \end{aligned} \quad (3.10)$$

$$\left. \frac{\partial^2}{\partial \theta_i \partial \theta_j} \mu_\theta(\alpha) \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -2\left(\alpha^2 + \frac{1}{4}\right) F_{ij}(\boldsymbol{\theta}_0), \quad (3.11)$$

where F_{ij} is the (i, j) element of the Fisher information matrix. The derivation of Eq. (3.11) is given in Appendix A. As $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_0$, the integrand of Eq. (3.5) approaches a constant as a function of α . Since the right-hand side of Eq. (3.5) is always zero, the constant must be zero as well. As a result, we also derive the second derivative of $\gamma(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$:

$$\left. \frac{\partial^2}{\partial \theta_i \partial \theta_j} \gamma(\boldsymbol{\theta}) \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 2F_{ij}(\boldsymbol{\theta}_0). \quad (3.12)$$

Using Taylor expansion, the previous results provide a second-order Fisher information approximation to the detectability squared in terms of $\boldsymbol{\theta} - \boldsymbol{\theta}_0$:

$$d_\Lambda(\boldsymbol{\theta}) = \sqrt{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathbf{F}(\boldsymbol{\theta}_0) (\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \dots}. \quad (3.13)$$

Using similar arguments, we are able to derive the third derivatives of $\gamma(\boldsymbol{\theta})$:

$$\left. \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} \gamma(\boldsymbol{\theta}) \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \left[\frac{\partial F_{jk}(\boldsymbol{\theta})}{\partial \theta_i} + \frac{\partial F_{ik}(\boldsymbol{\theta})}{\partial \theta_j} + \frac{\partial F_{ij}(\boldsymbol{\theta})}{\partial \theta_k} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (3.14)$$

from which the third-order approximations to the detectability squared are obtained. The derivation of Eq. (3.14) is given in Appendix A. Unfortunately, we have been unable to compute higher-order derivatives by using this approach. We also find that, by using an equivalent parameterization of the signal, the approximation will not be affected. If $\tau(\theta)$ is used to parameterize the probability density function instead of θ and its Jacobian matrix is nonsingular at θ_0 , our approximated detectability values derived from these two parameterizations at corresponding values θ' and τ' are the same. The proof is provided in Appendix B.

B. Approximating the $G(0)$ Approximation

When computing a higher-order approximation to the AUC, we note that $G(0)$ takes the form of

$$G_\theta(0) = -4 \ln(\langle \Lambda_\theta^{1/2} \rangle_{\theta_0}). \quad (3.15)$$

By definition, $G_\theta(0)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0$. Since $G_\theta(0)$ has the analytical form of Eq. (3.15), we are able to compute its derivatives with respect to $\boldsymbol{\theta}$. With $\Lambda_\theta(\mathbf{g})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 1$, the first derivative of $G_\theta(0)$ is derived as follows:

$$\begin{aligned}
 \left. \frac{\partial}{\partial \theta_i} G_\theta(0) \right|_{\theta=\theta_0} &= -2 \left. \frac{\left\langle \Lambda_\theta^{-1/2} \frac{\partial \Lambda_\theta}{\partial \theta_i} \right\rangle_{\theta_0}}{\langle \Lambda_\theta^{1/2} \rangle_{\theta_0}} \right|_{\theta=\theta_0} \\
 &= -2 \left. \left\langle \frac{1}{\text{pr}(\mathbf{g}|\theta_0)} \left[\frac{\partial \text{pr}(\mathbf{g}|\theta)}{\partial \theta_i} \right] \right\rangle_{\theta_0} \right|_{\theta=\theta_0} \\
 &= -2 \left. \frac{\partial}{\partial \theta_i} \left\langle \frac{\text{pr}(\mathbf{g}|\theta)}{\text{pr}(\mathbf{g}|\theta_0)} \right\rangle_{\theta_0} \right|_{\theta=\theta_0} = 0. \quad (3.16)
 \end{aligned}$$

Similarly, we are able to compute the higher-order derivatives of $G_\theta(0)$, and these derivatives provide the terms in a Taylor expansion approximation of $G_\theta(0)$. Substituting these approximations into relation (2.9), we realize that the approximation obtained from $G_\theta(0)$ is the same as the approximation obtained by using the Fisher information matrix up to the third order. The derivations of the second- and third-order approximations are given in Appendix C. We are able to compute higher-order derivatives of $G_\theta(0)$ to approximate the ideal-observer AUC.

C. Example

From the probability density functions under the two hypotheses, H_0 and H_1 , we can define a parameterized density function with a scalar parameter ρ :

$$\text{pr}(\mathbf{g}|\rho) = \frac{\text{pr}^\rho(\mathbf{g}|H_1)\text{pr}^{1-\rho}(\mathbf{g}|H_0)}{\langle \Lambda^\rho \rangle_0} = \frac{\Lambda^\rho(\mathbf{g})\text{pr}(\mathbf{g}|H_0)}{\langle \Lambda^\rho \rangle_0},$$

where $\text{pr}(\mathbf{g}|\rho) = \text{pr}(\mathbf{g}|H_1)$ when $\rho=1$ and $\text{pr}(\mathbf{g}|\rho) = \text{pr}(\mathbf{g}|H_0)$ when $\rho=0$. This family of densities appears in information theory.^{9,10} The advantage of using this density is that it can interpolate between any two densities. It does not require that $\text{pr}(\mathbf{g}|H_0)$ and $\text{pr}(\mathbf{g}|H_1)$ have the same functional form with a varying parameter. The Fisher information matrix approximation method can be applied to this parameterized density. The Fisher information matrix can be written as

$$\begin{aligned}
 F(0) &= \left\langle \frac{\partial}{\partial \rho} [\ln \text{pr}(\mathbf{g}|\rho)] \frac{\partial}{\partial \rho} [\ln \text{pr}(\mathbf{g}|\rho)] \right\rangle_{\rho=0}, \\
 \frac{\partial}{\partial \rho} [\ln \text{pr}(\mathbf{g}|\rho)] &= \ln \Lambda(\mathbf{g}) - \frac{1}{\langle \Lambda^\rho \rangle_0} \left\langle \frac{\partial}{\partial \rho} e^{(\ln \Lambda)\rho} \right\rangle_0 \\
 &= \ln \Lambda(\mathbf{g}) - \frac{1}{\langle \Lambda^\rho \rangle_0} \langle \Lambda^\rho \ln \Lambda \rangle_0 \\
 &= \ln \Lambda(\mathbf{g}) - \int \frac{\Lambda^\rho}{\langle \Lambda^\rho \rangle_0} \text{pr}(\mathbf{g}|H_0) \ln \Lambda \, d\mathbf{g} \\
 &= \lambda(\mathbf{g}) - \langle \lambda \rangle_\rho \\
 &= s(\mathbf{g}).
 \end{aligned}$$

The notation $\langle \cdot \rangle_\rho$ represents the conditional mean over the density $\text{pr}(\mathbf{g}|\rho)$, λ is the log-likelihood ratio given by λ

$= \ln \Lambda$, and $s(\mathbf{g})$ is the score. The Fisher information matrix of the parameterized density function becomes

$$F(0) = \langle s^2 \rangle_0 = \text{var}_0 \lambda.$$

This gives us a second-order approximation of detectability:

$$d_A = \sqrt{\text{var}_0 \lambda + \dots} \quad (3.17)$$

The third-order approximation can be computed similarly, and its formulation is given by

$$d_A = \sqrt{\text{var}_0 \lambda + \frac{1}{2} \langle s^3 \rangle_0 + \dots}, \quad (3.18)$$

A detailed derivation is shown in Appendix D.

4. ANALYTICAL EXAMPLE

Here we are able to compare analytically the approximations using methods derived in this paper with the exact

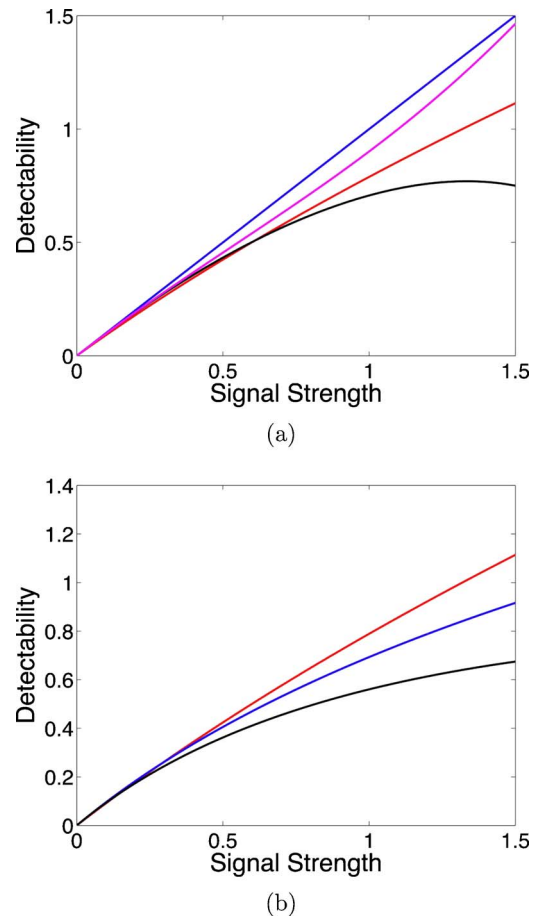


Fig. 1. (Color online) Exact detectability and approximations to detectability for Poisson noise. The horizontal axis is the signal strength Ms , and the vertical axis is the detectability. When the signal strength is 1 in (a), from top to bottom, we have the Fisher information approximation given by Eq. (3.13), the fourth-order $G_\theta(0)$ approximation, the exact value of the detectability, and the third-order approximation using Eqs. (3.13) and (3.14). When the signal strength is 1 in (b), from top to bottom, we have the exact value of the detectability, the second-order $\text{pr}(\mathbf{g}|\rho)$ approximation given by Eq. (3.17), and the third-order $\text{pr}(\mathbf{g}|\rho)$ approximation given by Eq. (3.18).

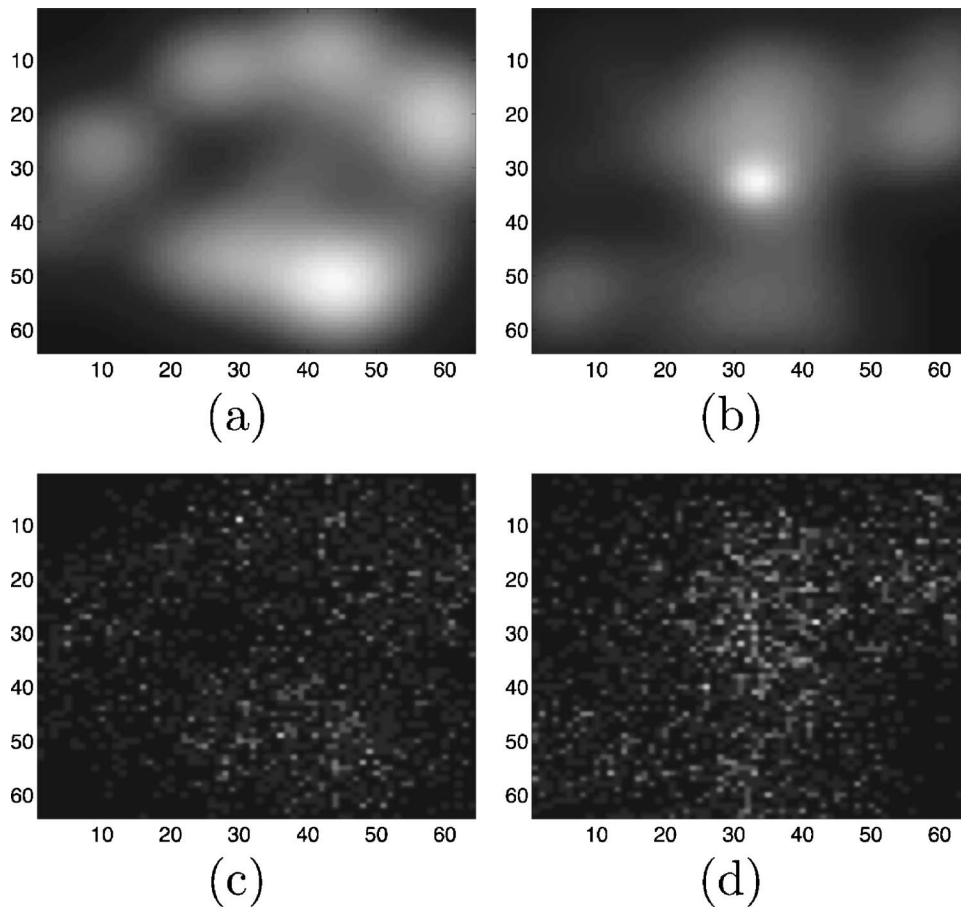


Fig. 2. Lumpy model objects and their noisy image realizations. A signal-absent lumpy object and a signal-present lumpy object are shown in (a) and (b), respectively; their noisy realizations are shown in (c) and (d), respectively. The number and location of lumps in these objects are random.

AUC values for some distributions. One example is Poisson statistics, which are often used when the number of photons collected by the imaging system is relatively small. When there is no signal present, the conditional probability of the data is given by

$$\text{pr}(\mathbf{g}|H_0) = \prod_{m=1}^M \frac{\exp[g_m \log(b_m) - b_m]}{g_m!}. \quad (4.1)$$

The signal changes the mean data vector from \mathbf{b} to $\mathbf{b} + \mathbf{s}$. We assume that both \mathbf{b} and \mathbf{s} are positive vectors and that the values of both vectors are known. In the special case where $b_1 = b_2 = \dots = b_M = b$, $s_1 = s_2 = \dots = s_M = s$, and the background intensity Mb is fixed at $Mb = 1$, we apply the approximation methods using Ms as the signal parameter. We compare the approximations with the exact value of AUC given by Clarkson and Barrett.⁸ The result is shown in Fig. 1, and other analytical results have been presented by Shen and Clarkson.¹¹ As shown in this example, the second-order approximations correctly capture the slope of the detectability at $\theta = 0$ and the third-order approximations follow the curvature of the detectability more closely. The fourth-order $G_\theta(0)$ approximations do not provide more accurate results when compared with lower-order approximations because they do not approximate the true value of the detectability directly. Instead,

they are approximations of the right-hand side of relation (2.9).

5. SIMULATIONS

It is very important to be able to approximate the performance of the ideal observer for more realistic images. Therefore, in our simulation study, we adopt the lumpy object model to test our methods. A lumpy background has N Gaussian lumps randomly distributed in the field of view, where N is a Poisson-distributed number with mean \bar{N} . We use 64×64 images with $\bar{N} = 25$ in our simulations. Various types of noise can be introduced. In this paper, we choose Poisson noise. The imaging model can be expressed by using the following formula:

$$\bar{g}_m = \sum_{n=1}^N \int_{-\infty}^{\infty} h_m(\mathbf{r}) L(\mathbf{r} - \mathbf{c}_n) d\mathbf{r}. \quad (5.1)$$

The function $h_m(\mathbf{r})$, which is a Gaussian function, represents the point-spread function of a linear imaging system for the m th pixel in the image. Within one imaging system, the height and width of $h_m(\mathbf{r})$ can vary among different pixels to characterize an imaging system's noise level and resolution at different locations. Two different sets of $h_m(\mathbf{r})$ can represent two imaging systems, respec-

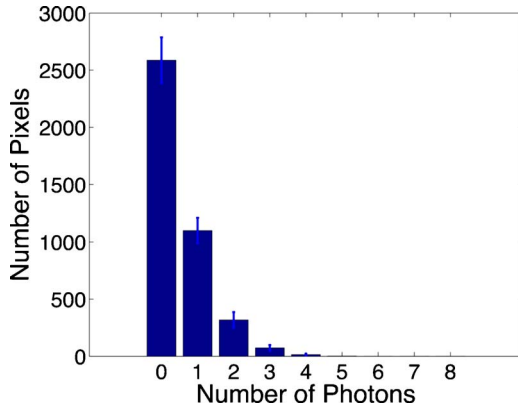


Fig. 3. (Color online) Single point histogram of the lumpy objects.

tively. In our simulation, the shapes of the Gaussian functions are the same for all pixels. The function $L(\mathbf{r})$ is also Gaussian, whose width and height describe the size and brightness of the lumps. The vector \mathbf{c}_n is the position of the n th lump in the background. To generate signal-present images, we add a signal lump at the center of the lumpy background. The lumpy object model is a background-known-statistically (BKS) and signal-known-exactly model. Figure 2 shows a signal-absent lumpy object (upper left) and a signal-present object (upper right) with a relatively small lump representing the signal. Two corresponding noisy realizations of the image data of the objects are shown in the lower row. The bright signal is used for illustration purposes only. In our simulations, the signal is much weaker, and it would be extremely difficult for human observers to detect the signal. Figure 3 shows the single point histogram distribution from a lumpy object model, which is not able to be fitted by a Gaussian model.

Kupinski *et al.* have developed Markov chain Monte Carlo (MCMC) algorithms to compute the ideal-observer detectability¹² and to compute the Fisher information.¹³ We simulate the likelihood ratio for each image by generating a long chain of objects for which we know the background parameters, such as the locations of the background lumps. We compute the BKE likelihoods (the likelihoods of obtaining the image data from each of the objects in the chain). The only randomness in the BKE likelihoods is the Poisson noise in our simulation. The likelihood of each object’s image data is computed by averaging the BKE likelihoods over the chain. The Fisher information is simulated in a similar manner. Note that, for the approximation, we need to simulate the Fisher information matrix only at $\theta = \theta_0$ to obtain the approximation for any θ . First, we consider a signal with fixed size but varying magnitude represented by θ , where $\theta_0 = 0$. We also consider the case where the magnitude of the signal is fixed. The varying signal parameter (the relative size) is defined by

$$\theta = \frac{\text{signal lumps full width at half-magnitude}}{\text{background lumps full width at half-magnitude}},$$

and $\theta_0 = 1$. In both cases, since the parameter is a scalar, the Fisher information also becomes a scalar.

The ideal-observer detectabilities and their Fisher information approximations provided by Eq. (3.13) are shown in Figs. 4 and 5. In Fig. 4, the detectability increases as the signal amplitude is increased. In Fig. 5, the signal is not distinguishable when the signal size equals the background lump size. As the signal size is reduced, it gradually differs from the background lumps. Therefore the detectability increases as the relative size of the signal decreases. Note that the vertical scales in Figs. 4 and 5 are different. We also studied the case where the signal size is greater than the background lump size (not shown). As expected, both the detectability and the value of its Fisher information approximation increased as the relative size of the signal increased. As shown in these figures, the second-order Fisher information approximation lies close to the detectability when the signal is weak. These are the situations when the human observers and the ideal observers perform poorly, and we need to improve the imaging system to have better performance from both observers. When $\theta \gg \theta_0$, which means that the signal is strong, human observers perform very well, and the ideal observer’s performance is nearly perfect (AUC

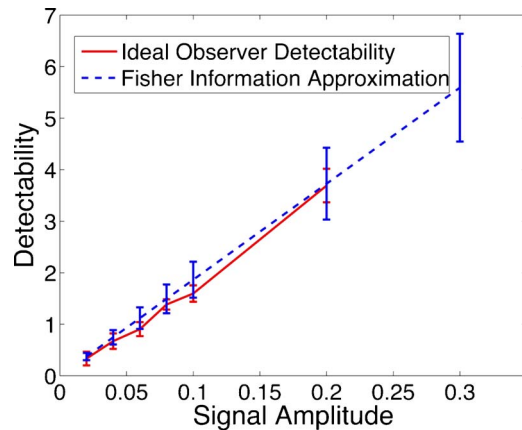


Fig. 4. (Color online) Illustration of the second-order Fisher information approximation of the detectability for lumpy-background images. The parameter considered here is the amplitude of the signal. The simulated ideal-observer detectability is infinite when the signal amplitude equals 0.3 and therefore is not shown in the figure.

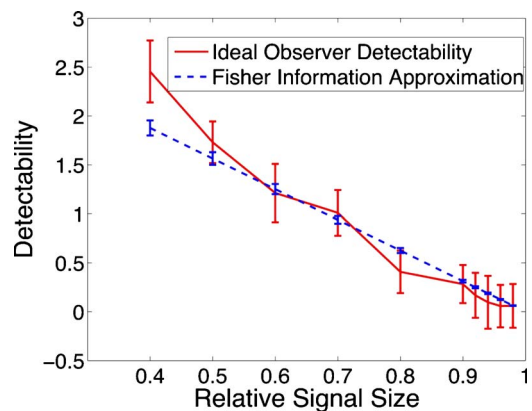


Fig. 5. (Color online) Illustration of the second-order Fisher information approximation of the detectability for lumpy-background images. The parameter considered here is the size of the signal.

≈ 1). Even without an accurate approximation, we are still able to predict the ideal-observer performance. The point at which our approximation starts to diverge from the detectability depends on the magnitude of the third-order term, for which we do not yet have a bound.

6. CONCLUSIONS

Through both the analytical examples and the numerical simulations, we can see that the second-order Fisher information approximation correctly captures the slope of the ideal-observer detectability. When the third-order approximations have reasonable values (for example, always positive), they follow the curvature of the detectability more closely than do the second-order approximations of the same method. The fourth-order approximations from $G_{\theta}(0)$ do not necessarily give a better approximation, even though they are of higher order. For the examples considered so far, we can see that the approximations from the Fisher information matrix generally agree most closely with the exact calculation of detection performance. To obtain a better approximation, one may consider computing the higher-order derivatives of the function $\gamma(\theta)$.

Since we can approximate the ideal-observer detectability at any signal strength by using the Fisher information matrix evaluated at the null value θ_0 , we suggest that the Fisher information matrix at this specific value can be used as a figure of merit of the imaging system. The advantage of this figure of merit is that it is independent of the strength of the signal. Our approximation also provides theoretical proof of the connection between the detection tasks and the corresponding estimation tasks. The Fisher information appears in the lower bound of the variance of the uniformly minimum variance unbiased estimator (UMVUE). For this particular estimator, as the norm of the Fisher information matrix increases, the lower bound of the variance decreases, and the mean square error becomes smaller. Meanwhile, in signal-detection tasks, how fast the detectability grows as a function of the signal strength depends on the norm of the Fisher information matrix. When we improve an imaging system by obtaining a Fisher information matrix with greater norm, we increase the ideal-observer performance for a detection task and, at the same time, we also decrease the variance of the UMVUE for the corresponding estimation task. We can optimize an imaging system for both tasks by constructing it in such a way that the greatest amount of Fisher information is obtained.

APPENDIX A: DERIVATION OF THE SECOND AND THIRD DERIVATIVES OF $\Gamma(\Theta)$

The second derivative of the $\mu_{\theta}(\alpha)$ is derived as

$$\begin{aligned} \left. \frac{\partial^2}{\partial \theta_i \partial \theta_j} \mu_{\theta}(\alpha) \right|_{\theta=\theta_0} &= - \left(\alpha^2 + \frac{1}{4} \right) \left\langle \Lambda_{\theta}^{-3/2+i\alpha} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \right\rangle_0 \left\langle \Lambda_{\theta}^{1/2-i\alpha} \right\rangle_0 \Big|_{\theta=\theta_0} + \left(\frac{1}{2} + i\alpha \right) \left\langle \Lambda_{\theta}^{-1/2+i\alpha} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} \right\rangle_0 \left\langle \Lambda_{\theta}^{1/2-i\alpha} \right\rangle_0 \Big|_{\theta=\theta_0} \\ &+ \left(\alpha^2 + \frac{1}{4} \right) \left\langle \Lambda_{\theta}^{-1/2+i\alpha} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \right\rangle_0 \left\langle \Lambda_{\theta}^{-1/2-i\alpha} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \right\rangle_0 \Big|_{\theta=\theta_0} + \text{c. c.} = -2 \left(\alpha^2 + \frac{1}{4} \right) \left\langle \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \right\rangle_0 \\ &= -2 \left(\alpha^2 + \frac{1}{4} \right) F_{ij}(\theta_0). \end{aligned}$$

Some of the conditional means vanish if we follow an argument similar to that shown in Eq. (3.8). Similarly, the only nonzero terms in the third derivative of μ_{θ} evaluated at $\theta=\theta_0$ are

$$\begin{aligned} \left. \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} \mu_{\theta}(\alpha) \right|_{\theta=\theta_0} &= - \left(\alpha^2 + \frac{1}{4} \right) \left\langle \Lambda_{\theta}^{-3/2+i\alpha} \left(\frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_k} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} + \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_j \partial \theta_k} + \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \right) \right\rangle_0 \left\langle \Lambda_{\theta}^{1/2-i\alpha} \right\rangle_0 \Big|_{\theta=\theta_0} \\ &- \left(\alpha^2 + \frac{1}{4} \right) \left(-\frac{3}{2} + i\alpha \right) \left\langle \Lambda_{\theta}^{-5/2+i\alpha} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \right\rangle_0 \left\langle \Lambda_{\theta}^{1/2-i\alpha} \right\rangle_0 \Big|_{\theta=\theta_0} + \text{c. c.} \\ &= \left(\alpha^2 + \frac{1}{4} \right) \left\langle 3 \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial \Lambda_{\theta}}{\partial \theta_k} - 2 \left(\frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_j \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} \right) \right\rangle_0 \Big|_{\theta=\theta_0} = - \left(\alpha^2 + \frac{1}{4} \right) \\ &\times \left[\frac{\partial}{\partial \theta_i} F_{jk}(\theta) + \frac{\partial}{\partial \theta_j} F_{ik}(\theta) + \frac{\partial}{\partial \theta_k} F_{ij}(\theta) \right]_{\theta=\theta_0}. \end{aligned}$$

The third derivative of $\gamma(\theta)$ is

$$\frac{\partial^3 \gamma(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} = \left[\frac{\partial}{\partial \theta_i} F_{jk}(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_j} F_{ik}(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_k} F_{ij}(\boldsymbol{\theta}) \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}.$$

APPENDIX B: FISHER INFORMATION APPROXIMATION USING DIFFERENT REPRESENTATIONS OF THE SIGNAL PARAMETER

The Fisher information approximation can be written as

$$d\boldsymbol{\theta}^\dagger F(\boldsymbol{\theta}_0) d\boldsymbol{\theta} = d\boldsymbol{\theta}^\dagger \langle \nabla_{\boldsymbol{\theta}} \ln \text{pr}(\mathbf{g}|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} \ln \text{pr}(\mathbf{g}|\boldsymbol{\theta})^\dagger \rangle_{\boldsymbol{\theta}_0} d\boldsymbol{\theta}.$$

Consider an equivalent signal parameterization $\boldsymbol{\tau}$, which is a function of $\boldsymbol{\theta}$ with a nonsingular Jacobian matrix \mathbf{J} at $\boldsymbol{\theta}_0$, where

$$J_{ij} = \partial \tau_i / \partial \theta_j.$$

Following the Jacobian matrix, we are able to obtain the following transformations:

$$\begin{aligned}
 \mathbf{J} d\boldsymbol{\theta} &= \begin{bmatrix} \frac{\partial \tau_1}{\partial \theta_1} & \frac{\partial \tau_1}{\partial \theta_2} & \dots & \frac{\partial \tau_1}{\partial \theta_n} \\ \frac{\partial \tau_2}{\partial \theta_1} & \frac{\partial \tau_2}{\partial \theta_2} & \dots & \frac{\partial \tau_2}{\partial \theta_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_n}{\partial \theta_1} & \frac{\partial \tau_n}{\partial \theta_2} & \dots & \frac{\partial \tau_n}{\partial \theta_n} \end{bmatrix} \begin{pmatrix} d\theta_1 \\ d\theta_2 \\ \vdots \\ d\theta_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \frac{\partial \tau_1}{\partial \theta_i} d\theta_i \\ \sum_{i=1}^n \frac{\partial \tau_2}{\partial \theta_i} d\theta_i \\ \vdots \\ \sum_{i=1}^n \frac{\partial \tau_n}{\partial \theta_i} d\theta_i \end{pmatrix} \\
 &= \begin{pmatrix} d\tau_1 \\ d\tau_2 \\ \vdots \\ d\tau_n \end{pmatrix} = d\boldsymbol{\tau},
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{J}^\dagger \nabla_{\boldsymbol{\tau}} &= \begin{bmatrix} \frac{\partial \tau_1}{\partial \theta_1} & \frac{\partial \tau_2}{\partial \theta_1} & \dots & \frac{\partial \tau_n}{\partial \theta_1} \\ \frac{\partial \tau_1}{\partial \theta_2} & \frac{\partial \tau_2}{\partial \theta_2} & \dots & \frac{\partial \tau_n}{\partial \theta_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_1}{\partial \theta_n} & \frac{\partial \tau_2}{\partial \theta_n} & \dots & \frac{\partial \tau_n}{\partial \theta_n} \end{bmatrix} \begin{pmatrix} \frac{\partial}{\partial \tau_1} \\ \frac{\partial}{\partial \tau_2} \\ \vdots \\ \frac{\partial}{\partial \tau_n} \end{pmatrix} \\
 &= \begin{pmatrix} \sum_{i=1}^n \frac{\partial}{\partial \tau_i} \frac{\partial \tau_i}{\partial \theta_1} \\ \sum_{i=1}^n \frac{\partial}{\partial \tau_i} \frac{\partial \tau_i}{\partial \theta_2} \\ \vdots \\ \sum_{i=1}^n \frac{\partial}{\partial \tau_i} \frac{\partial \tau_i}{\partial \theta_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \theta_1} \\ \frac{\partial}{\partial \theta_2} \\ \vdots \\ \frac{\partial}{\partial \theta_n} \end{pmatrix} = \nabla_{\boldsymbol{\theta}}.
 \end{aligned}$$

These two transformations are sufficient to prove the following equation:

$$\begin{aligned}
 d\boldsymbol{\tau}^\dagger \nabla_{\boldsymbol{\tau}} \ln \text{pr}(\mathbf{g}|\boldsymbol{\tau}(\boldsymbol{\theta})) &= (\mathbf{J} d\boldsymbol{\theta})^\dagger (\mathbf{J}^\dagger)^{-1} \nabla_{\boldsymbol{\theta}} \ln \text{pr}(\mathbf{g}|\boldsymbol{\theta}) \\
 &= d\boldsymbol{\theta}^\dagger \nabla_{\boldsymbol{\theta}} \ln \text{pr}(\mathbf{g}|\boldsymbol{\theta}),
 \end{aligned}$$

which leads to the equivalence of the Fisher information approximation using different parameterizations:

$$d\boldsymbol{\theta}^\dagger F(\boldsymbol{\theta}_0) d\boldsymbol{\theta} = d\boldsymbol{\tau}^\dagger F(\boldsymbol{\tau}_0) d\boldsymbol{\tau}.$$

APPENDIX C: DERIVATIVES OF $G_{\boldsymbol{\theta}}(0)$

The second derivative of $G_{\boldsymbol{\theta}}(0)$ is derived as

$$\begin{aligned}
 \frac{\partial^2}{\partial \theta_i \partial \theta_j} G_{\boldsymbol{\theta}}(0) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= \left(\frac{\left\langle \Lambda_{\boldsymbol{\theta}}^{-3/2} \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_i} \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_j} \right\rangle_0}{\langle \Lambda_{\boldsymbol{\theta}}^{1/2} \rangle_0} - 2 \frac{\left\langle \Lambda_{\boldsymbol{\theta}}^{-1/2} \frac{\partial^2 \Lambda_{\boldsymbol{\theta}}}{\partial \theta_i \partial \theta_j} \right\rangle_0}{\langle \Lambda_{\boldsymbol{\theta}}^{1/2} \rangle_0} + \frac{\left\langle \Lambda_{\boldsymbol{\theta}}^{-1/2} \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_i} \right\rangle_0 \left\langle \Lambda_{\boldsymbol{\theta}}^{-1/2} \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_j} \right\rangle_0}{\langle \Lambda_{\boldsymbol{\theta}}^{1/2} \rangle_0^2} \right) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\
 &= \left\langle \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_i} \frac{\partial \Lambda_{\boldsymbol{\theta}}}{\partial \theta_j} \right\rangle_0 \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = F_{ij}(\boldsymbol{\theta}_0).
 \end{aligned}$$

The third derivative of $G_{\boldsymbol{\theta}}(0)$ is derived as

$$\begin{aligned}
 \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} G_{\theta}(0) \Big|_{\theta=\theta_0} &= \left[\frac{\partial}{\partial \theta_k} \left(\frac{\left\langle \Lambda_{\theta}^{-3/2} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \right\rangle_0}{\left\langle \Lambda_{\theta}^{1/2} \right\rangle_0} - 2 \frac{\left\langle \Lambda_{\theta}^{-1/2} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} \right\rangle_0}{\left\langle \Lambda_{\theta}^{1/2} \right\rangle_0} + \frac{\left\langle \Lambda_{\theta}^{-1/2} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \right\rangle_0 \left\langle \Lambda_{\theta}^{-1/2} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \right\rangle_0}{\left\langle \Lambda_{\theta}^{1/2} \right\rangle_0^2} \right) \Big|_{\theta=\theta_0} \right] \\
 &= \left[\frac{\left\langle \Lambda_{\theta}^{-3/2} \left(\frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_j \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} \right) \right\rangle_0}{\left\langle \Lambda_{\theta}^{1/2} \right\rangle_0} - \frac{3 \left\langle \nabla \Lambda_{\theta}^{-5/2} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \right\rangle_0}{2 \left\langle \Lambda_{\theta}^{1/2} \right\rangle_0^2} \right] \Big|_{\theta=\theta_0} \\
 &= \left\langle \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_j \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_k} + \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \frac{\partial^2 \Lambda_{\theta}}{\partial \theta_i \partial \theta_j} - \frac{3}{2} \frac{\partial \Lambda_{\theta}}{\partial \theta_i} \frac{\partial \Lambda_{\theta}}{\partial \theta_j} \frac{\partial \Lambda_{\theta}}{\partial \theta_k} \right\rangle_0 \Big|_{\theta=\theta_0} \\
 &= \frac{1}{2} \left[\frac{\partial}{\partial \theta_i} F_{jk}(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_j} F_{ik}(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_k} F_{ij}(\boldsymbol{\theta}) \right] \Big|_{\theta=\theta_0}.
 \end{aligned}$$

APPENDIX D: DERIVATIVES OF THE THIRD-ORDER APPROXIMATION FOR A PARAMETERIZED DENSITY FUNCTION PR(G|P)

Following Eq. (3.14), the third derivative of the detectability squared as a function of ρ can be written as

$$\begin{aligned}
 \frac{\partial^3 \gamma(\rho)}{\partial \rho^3} &= 3 \frac{\partial^3}{\partial \rho} F(\rho) \\
 &= \left\langle 6 \left[\frac{\partial}{\partial \rho} \ln \text{pr}(\mathbf{g}|\rho) \right] \left[\frac{\partial^2}{\partial \rho^2} \ln \text{pr}(\mathbf{g}|\rho) \right] - 3 \left[\frac{\partial}{\partial \rho} \ln \text{pr}(\mathbf{g}|\rho) \right]^3 \right\rangle_0 \\
 &= \langle 6(\lambda - \langle \lambda \rangle_{\rho}) (\langle \lambda^2 \rangle_{\rho} - \langle \lambda \rangle_{\rho}^2) - 3(\lambda - \langle \lambda \rangle_{\rho})^3 \rangle_0.
 \end{aligned}$$

When evaluated at $\rho=0$, it becomes

$$\frac{\partial^3 \gamma(\rho)}{\partial \rho^3} = -3 \langle (\lambda - \langle \lambda \rangle_0)^3 \rangle_0 = -3 \langle s^3 \rangle_0.$$

Fangfang Shen and Eric Clarkson may be reached by e-mail at, respectively, fang@math.arizona.edu and clarkson@radiology.arizona.edu.

REFERENCES

1. H. H. Barrett and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," *J. Opt. Soc. Am. A* **15**, 1520–1535 (1998).

2. H. H. Barret "Objective assessment of image quality: effects of quantum noise arid object variability," *J. Opt. Soc. Am. A* **7**, 1266–1278 (1990).
3. H. H. Barrett and K. Myers, *Foundations of Image Science* (Wiley, 2004).
4. K. M. Hanson, "Variations in task and the ideal observer," in *Proc. SPIE* **419**, 60–67 (1983).
5. S. C. Moore, D. J. de Vries, B. Nandram, M. F. Kijewski, and S. P. Müller, "Collimator optimization for lesion detection incorporating prior information about lesion size," *Med. Phys.* **22**, 703–713 (1995).
6. S. P. Müller, C. K. Abbey, F. J. Rybicki, S. C. Moore, and M. F. Kijewski, "Measures of performance in nonlinear estimation tasks: prediction of estimation performance at low signal-to-noise ratio," *Phys. Med. Biol.* **50**, 3697–3715 (2005).
7. H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *J. Opt. Soc. Am. A* **12**, 834–852 (1995).
8. E. Clarkson and H. H. Barrett, "Approximations to ideal-observer performance on signal-detection tasks," *Appl. Opt.* **39**, 1783–1793 (2000).
9. S. Kullback, *Information Theory and Statistics* (Dover, 1997).
10. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, 1991).
11. F. Shen and E. Clarkson, "Using Fisher information to compute ideal-observer performance on detection tasks," in *Proc. SPIE* **5372**, 22–30 (2004).
12. M. A. Kupinski, J. Hoppin, E. Clarkson, and H. Barrett, "Ideal observer computation in medical imaging with use of Markov chain Monte Carlo," *J. Opt. Soc. Am. A* **20**, 430–438 (2003).
13. M. A. Kupinski, E. Clarkson, K. Gross, and J. W. Hoppin, "Optimizing imaging hardware for estimation tasks," in *Proc. SPIE* **5043**, 309–313 (2003).