Asymptotic Source Detection Performance of Gamma-Ray Imaging Systems Under Model Mismatch

Daniel J. Lingenfelter, *Student Member, IEEE*, Jeffrey A. Fessler, *Fellow, IEEE*, Clayton D. Scott, *Member, IEEE*, and Zhong He, *Senior Member, IEEE*

Abstract—Likelihood-based test statistics for the task of detecting a radioactive source in background using a gamma-ray imaging system often have intractable distributions. This complicates the tasks of predicting detection performance and setting thresholds that ensure desired false-alarm rates. Asymptotic distributions of test statistics can aid in predicting performance and in setting detection thresholds. However, in applications with complex sensors, like gamma-ray imaging, often only approximate statistical models for the measurements are available. Standard asymptotic approximations can yield inaccurate performance predictions when based on misspecified models. This paper considers asymptotic properties of detection tests based on maximum likelihood (ML) estimates under model mismatch, i.e., when the statistical model used for detection differs from the true distribution. We provide general expressions for the asymptotic distribution of likelihood-based test statistics when the number of measurements is Poisson, and expressions specific to gamma-ray source detection that one can evaluate using a modest amount of data from a real system or Monte Carlo simulation. Considering a simulated Compton imaging system, we show that the proposed expressions yield more accurate detection performance predictions than previous expressions that ignore model mismatch. These expressions require less data and computation than conventional empirical methods.

Index Terms—Asymptotics, Compton scatter camera, detection, hypothesis testing, misspecified models.

I. INTRODUCTION

MANY detection strategies use test statistics based on maximum-likelihood (ML) parameter estimates. However, in many practical detection problems, the true likelihood is not available. Estimators derived from approximate models may not enjoy the desirable properties of the ML estimator, such as asymptotic unbiasedness and asymptotic efficiency. Predicting asymptotic detection performance under the incorrect assumption that these properties are satisfied can lead to inaccurate, and sometimes overly optimistic results. This work focuses on gamma-ray source detection using position-sensitive detectors, an application where model mismatch is often significant due to system model approximation.

Position-sensitive gamma-ray detectors are useful for security, medicine, and nuclear nonproliferation. Position-sensitive detectors are advantageous because they allow image formation and spatial localization of sources. However, some types of position-sensitive detectors have unknown or intractable models for the system response, motivating the use of approximate models. Examples of approximations that result in model mismatch are substituting a non-Gaussian uncertainty with a Gaussian distribution, linearizing the system response, and neglecting physical processes, such as Doppler broadening and pair production. Examples of position-sensitive gamma-ray detectors and approximate models for the interaction event probabilities are given in [1]–[4].

Asymptotic approximations are useful for predicting detection performance with complex sensors, but conventional asymptotic methods for predicting detection performance assume correct models. We analyzed the asymptotic detection performance of gamma-ray imaging systems in [5], proving that position-sensitivity always improves detection performance when the detector sensitivity is uniform. The formulas for asymptotic detection performance in [5] can give inaccurate results in practical applications where model mismatch exists. Our work extends previous work on asymptotic detection performance of gamma-ray detectors by accounting for model mismatch.

Asymptotics of estimators and likelihood ratios have been studied previously in the case of a nonrandom number of measurements [6]. The results for a nonrandom number of measurements are not applicable to gamma-ray imaging systems, which record a Poisson number of measurements in a fixed-time scan [7, p. 470]. Estimators derived from mispecified models are known as quasi-maximum-likelihood estimators (QMLE) [6]. For a nonrandom number of independent and identically distributed (i.i.d.) measurements, a QMLE derived from a mispecified model is asymptotically normal provided that certain regularity conditions are satisfied [6]. A QMLE can also be viewed as an M-estimate. The asymptotic normality and convergence of

M-estimates is discussed in [8] and [9]. To our knowledge, previous analyses have not considered the case where the number of measurements is random, even in recent work [10].

To address these issues, we extend the theory on convergence of the QMLE for a fixed number of measurements in [6] to the case where the number of measurements is Poisson. We then specialize this theory to gamma-ray detection to characterize the performance of the source intensity test (SIT) [5] and generalized likelihood ratio test (GLRT) [11] under model mismatch. The specialized theory allows efficient computation of the distribution of the QMLE using Monte Carlo methods.

We report experiments that demonstrate the accuracy of our methods and their favorable performance compared to a naïve application of expressions in [5] when model mismatch is present. In these experiments, we evaluate the asymptotic performance of tests for detecting a point-source in a distributed background using a simulated parallel-plate Compton detector and the approximate system model from [4]. We illustrate that evaluating the asymptotic performance in terms of the receiver operating characteristic (ROC) using the method outlined in this paper requires less data than generating the ROC empirically. We use ROC analysis rather than risk minimization in our experiments because it does not require prior probabilities on the presence and absence of a source.

The contributions of this paper are twofold: the extension of the asymptotic convergence results of [6] from the case of a fixed number of measurements to the case of a Poisson number of measurements, and the novel application of the asymptotic convergence results to detection performance approximation for gamma-ray imaging systems.

This paper is organized as follows: Section II introduces the problem, Section III gives results about convergence and asymptotic normality of the QMLE when the number of measurements is Poisson, Section IV gives specialized expressions for the asymptotic distribution of the SIT and GLRT for gamma-ray imaging systems, and Section V gives the results of numerical experiments using the proposed asymptotic theory.

II. BACKGROUND

To characterize asymptotic detection performance, we must first define the true distribution of recorded events and reasonable models for it. The true distribution is governed by physics but its exact form is typically unknown or difficult to compute in practice. We also describe a model distribution for a list of recorded events, which is often an approximation of the true distribution. We describe the model in general terms to allow adaptation to different types of gamma-ray detectors and perhaps other applications such as list-mode positron emission tomography (PET) scans [12].

A. True Distribution of Recorded Events

We assume that the true distribution of recorded events follows the list-mode model of [13] for Poisson measurements. During a fixed-duration scan, a gamma-ray imaging system records attribute vectors, such as interaction locations within the detector and deposited energy, for each photon interaction event. Let the list of attribute vectors recorded by the system be $\boldsymbol{x} = (x_1, x_2, \ldots, x_J)$. The random number of measurements

is $J \sim \text{Poisson}(\bar{J})$, where $\bar{J} = \lambda_{\text{s}}\tau$, λ_{s} is the rate of recorded events in counts per unit time, and τ is the deterministic scan time chosen by the user. Each $x_j \in \mathcal{R}$ where $\mathcal{R}$ is the set of all possible event attribute vectors. For example, in an energy and position-sensitive detector, x_j is a vector of the coordinates and energies of all interactions of the jth photon.

Let $\mathsf{p}(x)$ denote the true density[1] of the recorded attributes for a single interacting photon. In general, we do not parameterize this true distribution because it represents the true physical process. Provided that the count rate is low enough to avoid dead-time effects [13], the true probability density of the list of event attributes $\boldsymbol{x}$ is given by

$$\mathsf{p}(\boldsymbol{x}) = e^{-\bar{J}} \bar{J}^J / J! \prod_{j=1}^{J} \mathsf{p}(x_j). \qquad (1)$$

B. Measurement Model

The true distribution of the observations in (1) is not always known exactly, so we now give an observation model. Since the mean number of emissions $\bar{J}$ is unknown, we model it by $\tilde{J}(\boldsymbol{\theta})$, where $\tilde{J}(\boldsymbol{\theta}) = \tilde{\lambda}(\boldsymbol{\theta})\tau$, $\tilde{\lambda}(\boldsymbol{\theta})$ is the modeled photon emission rate in counts per unit time, and $\boldsymbol{\theta}$ is a parameter vector that lies in the set $\Theta \subset \mathbb{R}^d$ for some integer d. Examples of parameters of interest in gamma-ray imaging are source intensity, source position, and source energy. Let $\tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ denote the modeled attribute density, which should approximate the true density $\mathsf{p}(x)$. Under the above assumptions, we model the likelihood of the observations as follows [5], [13]:

$$\tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}) \triangleq e^{-\tau\tilde{\lambda}(\boldsymbol{\theta})} \left[\tau\tilde{\lambda}(\boldsymbol{\theta})\right]^J / J! \prod_{j=1}^{J} \tilde{\mathsf{p}}(x_j; \boldsymbol{\theta}). \qquad (2)$$

Throughout, we use superscript "$\sim$" to denote functions or distributions that are part of the model and might differ from the true underlying functions or distributions that they represent.

We define model mismatch in Definition 1 [6].

Definition 1: Model mismatch exists if for all $\boldsymbol{\theta} \in \Theta$ such that $\tilde{\lambda}(\boldsymbol{\theta}) = \lambda_{\text{s}}$, there exists an $x \in \mathcal{R}$ such that $\tilde{\mathsf{p}}(x; \boldsymbol{\theta}) \neq \mathsf{p}(x)$.

This definition means that there is no parameter in the parameter space such that the modeled and true distributions and count rates are identical.

C. Estimator Definition

The quasi maximum likelihood (QML) estimate of the parameter vector $\boldsymbol{\theta}$ is given by

$$\tilde{\boldsymbol{\theta}}_\tau \triangleq \arg\max_{\boldsymbol{\theta} \in \Theta} \log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}), \qquad (3)$$

assuming a solution exists and is unique. The subscript τ is included to emphasize that the estimate is a function of a list of event attributes acquired during a scan of duration τ. In the next section, we examine the asymptotic properties of $\tilde{\boldsymbol{\theta}}_\tau$ as $\tau \to \infty$, i.e., as one records more events by increasing the scan time.

---

[1]Throughout, density refers to the probability density function if the attributes are continuous random variables, or the probability mass function if the attributes are discrete random variables. For more general or mixed distributions, density refers to the Radon-Nikodym density [14] with respect to an appropriate base measure.

Section V shows that the asymptotic approximation can accurately characterize the distribution of $\tilde{\boldsymbol{\theta}}_\tau$, even for a finite scan time.

III. PROPERTIES OF THE QMLE

We now describe the asymptotic behavior of the QMLE in (3). The results, stated in Theorems 1 and 2, are extensions of the results of [6] to the case where the number of measurements is Poisson. The results in [6] apply only to the case where the number of measurements is not random, and therefore are not directly applicable to gamma-ray detection systems. The regularity conditions required for Theorems 1 and 2 are similar to those required for the results of [6], except we add conditions on the modeled count rate $\tilde{\lambda}(\boldsymbol{\theta})$ to ensure that the modeled likelihood in (2) is well defined.

Let $\nabla_{\boldsymbol{\theta}}$ denote the $d \times 1$ column gradient with respect to $\boldsymbol{\theta}$, and let $\nabla_{\boldsymbol{\theta}}^2$ denote the $d \times d$ Hessian with respect to $\boldsymbol{\theta}$. By the definition in (3), assuming that the log-likelihood is differentiable and the maximizer is in the interior of the parameter space, the QMLE $\tilde{\boldsymbol{\theta}}_\tau$ satisfies

$$\nabla_{\boldsymbol{\theta}} \log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta})|_{\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_\tau} = \mathbf{0}. \tag{4}$$

We also define

$$\tilde{g}(\boldsymbol{\theta}) \triangleq \mathsf{E}\left[\log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta})\right] = \int_{\mathcal{R}} \log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}) \mathsf{p}(\boldsymbol{x}) d\boldsymbol{x} \tag{5}$$

where the expectation is with respect to the true distribution $\mathsf{p}(\boldsymbol{x})$. Even though the true distribution is unknown, we assume one can compute expectations by Monte Carlo integration with samples obtained from it.

A. Convergence

Let $\tilde{\boldsymbol{\mu}} \in \Theta$ be given by

$$\tilde{\boldsymbol{\mu}} \triangleq \arg\max_{\boldsymbol{\theta} \in \Theta} \tilde{g}(\boldsymbol{\theta}), \tag{6}$$

assuming that such a solution exists and is unique. Under suitable conditions, $\tilde{\boldsymbol{\mu}}$ is the limit of the sequence of estimates $\tilde{\boldsymbol{\theta}}_\tau$ as $\tau \to \infty$. Appendix A describes one sufficient set of regularity conditions that guarantees existence, uniqueness, and convergence. These conditions are similar to those in [6], with extensions to the case of a random number of measurements. The regularity conditions apply only to the model, with the exception of the assumption that the attributes of different photon interactions are i.i.d.. The regularity conditions ensure that the log-likelihood function exists and that the QMLE exists and is unique. We discuss verification of these conditions in Appendix C. The convergence of the QMLE is stated in Theorem 1, which extends [6, Theorem 2.2] to the case of a Poisson number of measurements.

Theorem 1: Under suitable regularity conditions, $\tilde{\boldsymbol{\theta}}_\tau \overset{\text{a.s.}}{\to} \tilde{\boldsymbol{\mu}}$ as $\tau \to \infty$.

A useful result for interpretation of $\tilde{\boldsymbol{\mu}}$ is given in Corollary 1. The additional regularity conditions required for Corollary 1 place restrictions on the true distribution. The proof and sufficient regularity conditions are given in Appendix A.

Corollary 1: Under suitable regularity conditions, $\tilde{\boldsymbol{\mu}}$ minimizes the Kullback-Leibler divergence [15] of the true distribution with respect to the modeled distribution.

This Theorem and Corollary are shown in [6] for the case of J nonrandom. In the absence of model mismatch, the true distribution (1) and modeled distributions (2) are equal for some $\boldsymbol{\theta} \in \Theta$ such that $\lambda_{\mathsf{s}} = \tilde{\lambda}(\boldsymbol{\theta}_{\text{true}})$. Let $\boldsymbol{\theta}_{\text{true}}$ be the parameter value, assuming a unique solution, under which this equality is achieved. Then, in the absence of model mismatch, $\tilde{\boldsymbol{\mu}} = \boldsymbol{\theta}_{\text{true}}$ by Corollary 1, and Theorem 1 reduces to the usual asymptotic consistency of ML estimates [16] in the absence of model mismatch.

B. Asymptotic Normality

A QMLE may also be asymptotically normal if certain conditions on the model and true distributions are met. Theorem 2 is an extension of [6, Th. 3.2] to the case of a Poisson number of measurements. The regularity conditions ensure that the expectations in Theorem 2 are well defined. The proof is given in Appendix B.

Theorem 2: Under suitable regularity conditions, asymptotically as $\tau \to \infty$

$$\sqrt{\tau}(\tilde{\boldsymbol{\theta}}_\tau - \tilde{\boldsymbol{\mu}}) \overset{d}{\to} \mathcal{N}\left(\mathbf{0}, \Sigma(\tilde{\boldsymbol{\mu}})\right) \tag{7}$$

where

$$\Sigma(\boldsymbol{\theta}) = \tilde{\mathsf{H}}^{-1}(\boldsymbol{\theta}) \tilde{\mathsf{G}}(\boldsymbol{\theta}) \tilde{\mathsf{H}}^{-1}(\boldsymbol{\theta}), \tag{8}$$

$$\tilde{\mathsf{G}}(\boldsymbol{\theta}) \triangleq \lambda_{\mathsf{s}} \mathsf{E}\left[\left(\nabla_{\boldsymbol{\theta}} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) + \nabla_{\boldsymbol{\theta}} \log \tilde{\lambda}(\boldsymbol{\theta})\right), \right.$$
$$\left. \times \left(\nabla_{\boldsymbol{\theta}} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) + \nabla_{\boldsymbol{\theta}} \log \tilde{\lambda}(\boldsymbol{\theta})\right)^T\right], \tag{9}$$

$$\tilde{\mathsf{H}}(\boldsymbol{\theta}) \triangleq -\lambda_{\mathsf{s}} \nabla_{\boldsymbol{\theta}}^2 \log \tilde{\lambda}(\boldsymbol{\theta}) + \nabla_{\boldsymbol{\theta}}^2 \tilde{\lambda}(\boldsymbol{\theta}) - \lambda_{\mathsf{s}} \mathsf{E}\left[\nabla_{\boldsymbol{\theta}}^2 \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta})\right], \tag{10}$$

and expectations are with respect to the true distribution.

This theorem differs from [6, Theorem 3.2] because convergence is shown as the scan time τ, rather than the number of measurements, goes to infinity. Our proof in Appendix B avoids dividing the scan into intervals and naturally shows convergence as scan time goes to infinity.

In the absence of model mismatch, $\tilde{\mathsf{G}}(\boldsymbol{\theta})$ and $\tilde{\mathsf{H}}(\boldsymbol{\theta})$ are equal to the time-normalized Fisher information matrix $\mathsf{F}(\boldsymbol{\theta})$ [6], where

$$\mathsf{F}(\boldsymbol{\theta}) \triangleq -\frac{1}{\tau} \mathsf{E}\left[\nabla_{\boldsymbol{\theta}}^2 \log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta})\right].$$

In this special case, the covariance in (8) simplifies to $\Sigma(\boldsymbol{\theta}) = \mathsf{F}^{-1}(\boldsymbol{\theta})$. Thus, Theorem 2 reduces to the classic asymptotic normality and asymptotic efficiency of ML estimation [16] in the absence of model mismatch.

C. Using Asymptotic Distributions for Approximation

It may be challenging to verify the regularity conditions for Theorems 1 and 2 listed in Appendices A and B. Our numerical results show that Theorems 1 and 2 are useful as practical approximations for gamma-ray detection problems. The conditions listed in the Appendix are sufficient, but not necessary. A discussion of how one could verify the regularity conditions is given in Appendix C.

The asymptotic mean $\tilde{\boldsymbol{\mu}}$ in (6) and covariance $\Sigma(\tilde{\boldsymbol{\mu}})$ in (7) both depend on the true attribute distribution $\mathsf{p}(x)$. In some gamma-ray imaging problems, the exact form of $\mathsf{p}(x)$ is computationally intractable, but it is relatively easy to compute expectations (9) and (10) with respect to $\mathsf{p}(x)$ by Monte Carlo methods [17]. The next section uses Theorems 1 and 2 to develop specialized expressions for the asymptotic distribution of test statistics applied to gamma-ray source detection problems.

IV. ASYMPTOTICS FOR GAMMA-RAY SOURCE DETECTION

The true and modeled distributions in Section II are more general than required for gamma-ray source detection problems. This section develops a model and a true distribution for the specific task of source detection, and uses the model to develop asymptotic distributions of useful test statistics for detecting a source in background.

A. True Distribution

In Section II, we made no assumption about the form of the true distribution of recorded attributes $\mathsf{p}(x)$. In gamma-ray source detection problems, the measurements are often generated by a mixture of source and background emission processes. Let α^{t} be the nonnegative true mean number of *emitted* source counts per unit time, λ_b^{t} be the nonnegative true mean number of background counts *recorded* per unit time, and s^{t} be the true sensitivity, or probability that a photon emitted from the source is recorded. We assume that the true distribution follows the mixture

$$\mathsf{p}(x) = \frac{\alpha^{\mathsf{t}} s^{\mathsf{t}} \mathsf{p}_{\mathsf{S}}(x) + \lambda_b^{\mathsf{t}} \mathsf{p}_{\mathsf{B}}(x)}{\alpha^{\mathsf{t}} s^{\mathsf{t}} + \lambda_b^{\mathsf{t}}} \qquad (11)$$

where $\mathsf{p}_{\mathsf{S}}(x)$ is the density of recorded attributes given that they originated from the source and $\mathsf{p}_{\mathsf{B}}(x)$ is the density of recorded attributes given that they originated from the background. We adopt a mixture model because the recorded attributes of events that originate from the source have a different distribution than the attributes of recorded events that originate from the background. One can derive (11) using the law of total probability. The nonnegative sensitivity s^{t} is a function of the spatial relationship between the source and detector. The quantities α^{t} and λ_b^{t} are usually not known in practice, but are needed for simulation. One can estimate s^{t} by simulation for a particular detector geometry, source position, and energy spectrum. The total mean number of received counts obeys the relation $\lambda_{\mathsf{s}} = \alpha^{\mathsf{t}} s^{\mathsf{t}} + \lambda_b^{\mathsf{t}}$. The density $\mathsf{p}(x)$ is a probability density function by the nonnegativity of α^{t}, s^{t}, and λ_b^{t}.

We formulate source detection as the hypothesis testing problem

$$H_1 : \alpha^{\mathsf{t}} > 0$$
$$H_0 : \alpha^{\mathsf{t}} = 0.$$

B. Model Distribution

We consider a modeled attribute distribution similar to that of [5] for detecting a source in background. The parameters characterizing the source are the intensity α with units of counts

emitted per unit time and source position[2] $\boldsymbol{\phi} \in \Phi$. In the 3D far-field with a known source energy, the set Φ could be $[0, 2\pi] \times [0, \pi]$, representing all possible azimuth and polar angles on a sphere. We parameterize the background intensity by the background count rate λ_b with units of gamma-ray counts *recorded* per unit time. Let $\boldsymbol{\theta}$ be the vector of all parameters, where $\boldsymbol{\theta}$ lies in the d-dimensional parameter space Θ. In what follows, we assume that $\boldsymbol{\theta}$ takes the form:

$$\boldsymbol{\theta} = (\alpha, \boldsymbol{\phi}). \qquad (12)$$

Let the modeled sensitivity $\tilde{s}(\boldsymbol{\phi})$ approximate the probability that a photon emitted from a source positioned at $\boldsymbol{\phi}$ is recorded. We model the rate of recorded photons by

$$\tilde{\lambda}(\boldsymbol{\theta}) \triangleq \lambda_b + \alpha\tilde{s}(\boldsymbol{\phi}). \qquad (13)$$

We adopt the model in (13) because the total rate of recorded photons is the sum of the recorded count rate due to background λ_b and the recorded count rate due to the source $\alpha\tilde{s}(\boldsymbol{\phi})$.

Let $\tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi})$ denote the modeled distribution of a recorded attribute vector x given it originated from a source at positon $\boldsymbol{\phi}$ and let $\tilde{\mathsf{p}}_{\mathsf{B}}(x)$ denote the modeled distribution of a recorded attribute vector x given that it originated from the background. Note that $\tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi})$ depends only on the source position, and $\tilde{\mathsf{p}}_{\mathsf{B}}(x)$ does not depend on any of the parameters in (12).

We model the overall distribution of recorded attributes as a mixture of $\tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi})$ and $\tilde{\mathsf{p}}_{\mathsf{B}}(x)$ given by

$$\tilde{\mathsf{p}}(x; \boldsymbol{\theta}) = \frac{\alpha\tilde{s}(\boldsymbol{\phi})\tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi}) + \lambda_b\tilde{\mathsf{p}}_{\mathsf{B}}(x)}{\alpha\tilde{s}(\boldsymbol{\phi}) + \lambda_b}. \qquad (14)$$

As in (11), $\tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ in (14) is a valid probability density function when α, $\tilde{s}(\boldsymbol{\phi})$, and λ_b are nonnegative.

C. Asymptotic Performance of Source Intensity Test (SIT) Under Model Mismatch

The SIT [5] for detecting the presence of a radiation source of unknown intensity α is given by

$$\tilde{\alpha}_\tau \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \qquad (15)$$

where $\tilde{\alpha}_\tau$ is the QMLE for α. By Theorem 2, assuming that the regularity conditions are satisfied,

$$\sqrt{\tau}\left(\tilde{\alpha}_\tau - \tilde{\mu}_{[1]}\right) \overset{d}{\to} \mathcal{N}\left(0, \Sigma(\tilde{\boldsymbol{\mu}})_{[1,1]}\right) \qquad (16)$$

as $\tau \to \infty$, where $\Sigma(\boldsymbol{\theta})$ is defined in (8), $[1, 1]$ denotes the $[1, 1]$ component of the matrix, and $\tilde{\mu}_{[1]}$ is the asymptotic mean of $\tilde{\alpha}_\tau$ defined by the first element of $\tilde{\boldsymbol{\mu}}$ in (6). Define $\tilde{\boldsymbol{\mu}}^{(1)}$ to be the solution to (6) under H_1 and $\tilde{\boldsymbol{\mu}}^{(0)}$ to be the solution to (6) under H_0. Thus, $\tilde{\alpha}_\tau$ is a consistent estimator of $\tilde{\mu}_{[1]}$ under H_1 and a consistent estimator of $\tilde{\mu}_{[1]}^{(0)}$ under H_0. Because of model mismatch, $\tilde{\mu}_{[1]}$ may not be equal to the true intensity and $\tilde{\mu}_{[1]}^{(0)}$ may not be zero.

We calculate $\tilde{\mu}_{[1]}^{(1)}$ and $\tilde{\mu}_{[1]}^{(0)}$ by simulating or recording a large number of observations from the true distribution under H_1 and H_0 and solving for the QMLE using (4) under both hypotheses. Although obtaining accurate estimates of $\tilde{\mu}_{[1]}^{(1)}$ and $\tilde{\mu}_{[1]}^{(0)}$ requires

---

[2]$\boldsymbol{\phi}$ could also denote a vector containing both spatial position and energy

many observations, we found that more are required to generate an ROC empirically with high accuracy. We calculate $\Sigma(\tilde{\boldsymbol{\mu}}^{(1)})$, and $\Sigma(\tilde{\boldsymbol{\mu}}^{(0)})$ using (8) and evaluate the expectations in (9) and (10) by Monte Carlo integration with data sampled from the true distribution. One can use these values with (16) to predict the corresponding asymptotic ROC curve.

D. Asymptotic Distribution of GLRT Under Model Mismatch

The GLRT is another common detection method given by [11]

$$\Lambda_{\text{GLRT}} \triangleq \frac{\tilde{\mathsf{p}}(\boldsymbol{x}; \tilde{\boldsymbol{\theta}}_\tau)}{\arg\max_{\boldsymbol{\theta}, \alpha=0} \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta})} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{17}$$

By a similar argument to that in [6] for a nonrandom number of measurements, asymptotically as $\tau \to \infty$

$$2\log\Lambda_{\text{GLRT}} \approx \tilde{\alpha}_\tau^2 \left(\tilde{\mathsf{H}}^{-1}(\tilde{\boldsymbol{\theta}}_\tau)_{[1,1]}\right)^{-1}. \tag{18}$$

The expression in (18) resembles the Wald test [11], which is a variance-normalized statistic of the ML estimates.

Combining Theorem 2 with (18) and performing algebraic manipulations, we have that under model mismatch for the GLRT with a scalar parameter and τ large,

$$2\log\Lambda_{\text{GLRT}} \sim \begin{cases} a\left(\tilde{\boldsymbol{\mu}}^{(1)}\right) \chi_1^2\left(\eta\left(\tilde{\boldsymbol{\mu}}^{(1)}\right)\right), & \text{under } H_1 \\ a\left(\tilde{\boldsymbol{\mu}}^{(0)}\right) \chi_1^2\left(\eta\left(\tilde{\boldsymbol{\mu}}^{(0)}\right)\right), & \text{under } H_0 \end{cases} \tag{19}$$

where $\chi_k^2(\eta)$ denotes a noncentral chi-square random variable with noncentrality parameter η and k degrees of freedom [11], and

$$\eta(\boldsymbol{\theta}) \triangleq \tau \alpha^2 \Sigma^{-1}(\boldsymbol{\theta})_{[1,1]} \tag{20}$$

$$a(\boldsymbol{\theta}) \triangleq \left(\Sigma^{-1}(\boldsymbol{\theta})_{[1,1]} \tilde{\mathsf{H}}^{-1}(\boldsymbol{\theta})_{[1,1]}\right)^{-1}. \tag{21}$$

The result in (19) generalizes the result from [11] that the GLRT is asymptotically noncentral chi-square under H_1 and central chi-square under H_0. The multiplicative factors $a(\tilde{\boldsymbol{\mu}}^{(1)})$ and $a(\tilde{\boldsymbol{\mu}}^{(0)})$ equal unity in the absence of model mismatch.

One can use (19) to compute the asymptotic area under the ROC curve (AUC) of the GLRT using Monte Carlo simulation to estimate $\tilde{\boldsymbol{\mu}}^{(1)}$, $\tilde{\boldsymbol{\mu}}^{(0)}$, $\Sigma(\tilde{\boldsymbol{\mu}}^{(1)})$, and $\Sigma(\tilde{\boldsymbol{\mu}}^{(0)})$. Note that $a(\boldsymbol{\theta})$ in (21) and $\eta(\boldsymbol{\theta})$ in (20) are functions of a generic parameter vector $\boldsymbol{\theta}$. When α appears on the right-hand side (RHS) of (20), it represents the first element of $\boldsymbol{\theta}$, which corresponds to the source intensity parameter α. One can use the distributions in (19) to compute the ROC and AUC.

E. Asymptotic Distributions for Gamma-Ray Imaging

The covariance matrix in (7) has a specific form when the recorded attributes are distributed according to the mixture model in (14) and the parameter vector has the form (12). In this section, we derive expressions for $\tilde{\mathsf{G}}(\boldsymbol{\theta})$ and $\tilde{\mathsf{H}}(\boldsymbol{\theta})$ in (8). Let

$$\tilde{\mathsf{F}}(\boldsymbol{\theta}) = \tilde{\mathsf{F}}_{\text{c}}(\boldsymbol{\theta}) \begin{bmatrix} K(\boldsymbol{\theta})_{[1,1]} & \frac{\alpha K(\boldsymbol{\theta})_{[2,1]}^T}{\tilde{s}(\boldsymbol{\phi})} \\ \frac{\alpha K(\boldsymbol{\theta})_{[2,1]}}{\tilde{s}(\boldsymbol{\phi})} & \frac{\alpha^2 K(\boldsymbol{\theta})_{[2,2]}}{\tilde{s}^2(\boldsymbol{\phi})} \end{bmatrix} \tag{22}$$

where

$$K(\boldsymbol{\theta})_{[i,j]} \triangleq \mathsf{E}\left[\frac{\tilde{q}_i(x; \boldsymbol{\theta})\tilde{q}_j^T(x; \boldsymbol{\theta})}{\tilde{\mathsf{p}}^2(x; \boldsymbol{\theta})}\right] \tag{23}$$

$$\tilde{q}_1(x; \boldsymbol{\theta}) \triangleq \tilde{\mathsf{p}}_{\text{S}}(x; \boldsymbol{\phi})$$

$$\tilde{q}_2(x; \boldsymbol{\theta}) \triangleq \nabla_{\boldsymbol{\phi}}\left(\tilde{s}(\boldsymbol{\phi})\tilde{\mathsf{p}}_{\text{S}}(x; \boldsymbol{\phi})\right)$$

$$\tilde{\mathsf{F}}_{\text{c}}(\boldsymbol{\theta}) \triangleq \frac{\lambda_{\text{s}}s^2(\boldsymbol{\phi})}{\left(\alpha\tilde{s}(\boldsymbol{\phi}) + \lambda_b\right)^2}. \tag{24}$$

In the absence of model mismatch, $\tilde{\mathsf{F}}(\boldsymbol{\theta})$ would simplify to the time-normalized Fisher information matrix $\mathsf{F}(\boldsymbol{\theta})$ for the model (2). Let

$$\mathsf{B}(\boldsymbol{\theta}) \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \alpha w(\boldsymbol{\theta}) \end{bmatrix} \tag{25}$$

where

$$w(\boldsymbol{\theta}) \triangleq \nabla_{\boldsymbol{\phi}}^2 \tilde{s}(\boldsymbol{\phi})\left(\frac{\lambda_{\text{s}}}{\tilde{\lambda}(\boldsymbol{\theta})} - 1\right).$$

The matrix $\mathsf{B}(\boldsymbol{\theta})$ depends on the degree of model mismatch and is zero when $\tilde{\lambda}(\boldsymbol{\theta}) = \lambda_{\text{s}}$. By algebra, one can show that for gamma-ray imaging detectors,

$$\tilde{\mathsf{G}}(\boldsymbol{\theta}) = \tilde{\mathsf{F}}(\boldsymbol{\theta}), \tag{26}$$

and

$$\tilde{\mathsf{H}}(\boldsymbol{\theta}) = \tilde{\mathsf{F}}(\boldsymbol{\theta}) - \mathsf{B}(\boldsymbol{\theta}). \tag{27}$$

The expression in (27) shows that model mismatch can decrease the Fisher information. This decrease results in increased variance of the parameter estimates $\tilde{\boldsymbol{\theta}}_\tau$.

In the next section, we evaluate these expressions at $\boldsymbol{\theta} = \tilde{\boldsymbol{\mu}}$ to explore the detection performance of various detectors.

V. NUMERICAL RESULTS

We applied the asymptotic distributions of test statistics derived in the previous section to detection performance prediction for gamma-ray measurement systems. We considered a position-sensitive Compton detector as an example application of the theory and to show that the theory accurately predicts empirical performance. Throughout, we refer to the performance predicted using the asymptotic theory as "predicted performance." We used the SIT (15) in our experiments because of its simplicity and superior performance in the absence of model mismatch compared to the GLRT as shown experimentally in [5].

We used the asymptotic distributions to create ROC curves that one can use to set the test threshold and to estimate the probability of detection for a given false alarm rate. The ROC curve is a plot of the probability of detection as a function of the probability of false alarm. Each point on the ROC is a pair of probability of false alarm and probability of detection values resulting from a particular choice of the threshold γ in (15) for the SIT or (17) for the GLRT. The ROC curve allows one to choose a threshold value γ that results in the desired false alarm rate.

We also state our results in terms of area under the ROC curve (AUC) [18], which is a measure of detector performance. The AUC ranges from 0.5 to 1, where 0.5 is the poorest performance and 1 is the best performance. The AUC is a measure of overall detectability that does not depend on the threshold γ in (15) or (17).

Throughout this section, we compare predicted performance accounting for and ignoring model mismatch. The purpose of this comparison is to illustrate the pitfalls of applying the asymptotic formulas that do not account for model mismatch in [5] to problems where model mismatch exists.

In the absence of model mismatch, there exists a parameter vector $\boldsymbol{\theta}_{\mathrm{true}}$ such that $\lambda_{\mathsf{s}} = \tilde{\lambda}(\boldsymbol{\theta}_{\mathrm{true}})$ and $\mathsf{p}(x;\boldsymbol{\theta}_{\mathrm{true}}) = \mathsf{p}(x)$ for all $x \in \mathcal{R}$. In this case, the theory from [5] states that the maximum likelihood estimator is a consistent estimator of $\boldsymbol{\theta}_{\mathrm{true}}$. When applying the theory in [5] in this section, we used the incorrect assumption that the estimates are normally distributed with mean $\boldsymbol{\theta}_{\mathrm{true}}$ rather than $\tilde{\boldsymbol{\mu}}$. This assumption is incorrect in the presence of model mismatch, and is one factor that contributes to the failure of the theory from [5] in the results that follow. In the absence of model mismatch, $\Sigma(\boldsymbol{\theta}) = \tilde{\mathsf{G}}^{-1}(\boldsymbol{\theta}) = \tilde{\mathsf{H}}^{-1}(\boldsymbol{\theta})$, so we make the substitution $\Sigma(\boldsymbol{\theta}) = \tilde{\mathsf{G}}^{-1}(\boldsymbol{\theta})$ to capture the naïve use of the methods in [5].

To compute the predicted performance, we sampled event attributes from the true distribution. We generated one list of events with a source present and another without a source present. We used these lists to compute the asymptotic means $\tilde{\boldsymbol{\mu}}^{(1)}$ and $\tilde{\boldsymbol{\mu}}^{(0)}$ by numerically solving (3). This is justified by the convergence result of Theorem 1, provided that the number of events in the list is sufficiently large. Although the list must be large enough to obtain an accurate estimate of the asymptotic mean, the asymptotic method often requires fewer simulated events to compute the ROC than empirical methods. We also used the lists of recorded events to evaluate (9) and (10) by Monte Carlo integration to obtain $\Sigma(\tilde{\boldsymbol{\mu}}^{(1)})$ and $\Sigma(\tilde{\boldsymbol{\mu}}^{(0)})$. In the experiments that follow, we used an equal number of background events in the two lists. Let N_{back} be the number of sampled background events in each scan. The list of recorded events without a source present consists of N_{back} events and the list of recorded events with a source present consists of N_{back} background events and $N_{\mathrm{back}}\alpha^{\mathsf{t}}s^{\mathsf{t}}/\lambda_b^{\mathsf{t}}$ source events. We use N_{pred} to report the total number of recorded events used to predict the detection performance using approximations based on asymptotics, where

$$ N_{\mathrm{pred}} = N_{\mathrm{back}}\left(2 + \alpha^{\mathsf{t}}s^{\mathsf{t}}/\lambda_b^{\mathsf{t}}\right), \tag{28} $$

and the $\alpha^{\mathsf{t}}s^{\mathsf{t}}/\lambda_b^{\mathsf{t}}$ term is due to the source events in the list that contains both source and background events. We use the relation in (28) because N_{pred} is the expected number of total measurements used to compute the ROC using the proposed method.

We compared the performance predicted by asymptotics to the empirical performance to show the accuracy of the proposed method. To calculate the empirical performance, we simulated a large number of scans with and without the source present. To simulate a scan, we first drew the number of emitted source counts J_{src} from a Poisson distribution with mean $\tau\alpha^{\mathsf{t}}$. We then simulated J_{src} emitted source photons, recording the attributes
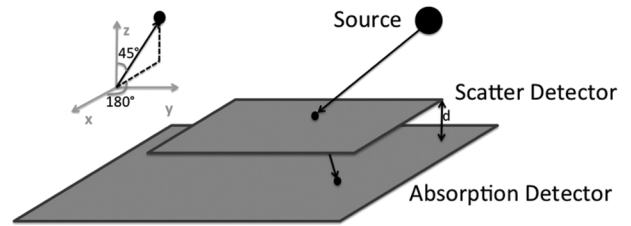

Fig. 1.　Illustration of parallel-plate Compton detector used for simulation.

of each photon that interacts with the detector. To simulate the background events for the scan, we drew the number of background counts J_{back} from a Poisson distribution with mean $\tau\lambda_b^{\mathsf{t}}$ and simulated background photons until we record the attributes for J_{back} interactions. For each scan, we computed the QMLE $\tilde{\alpha}$ by solving (3).

We computed the empirical ROC and AUC using the source intensity estimates for all simulated scans with and without a source. For each value of the threshold in (15), the fraction of source intensity estimates above that threshold when a source is present is the probability of detection and the fraction of source intensity estimates above that threshold when a source is absent is the probability of false alarm. We computed the empirical AUC using the Wilcoxon-Mann-Whitney statistic [19]. Let N_{scans} be the number of scans used to compute the ROC. We simulated N_{scans} scans with a source present and N_{scans} scans without a source present. Let N_{emp} be the mean number of recorded sample events needed to complete all $2N_{\mathrm{scans}}$ scans. We report the mean number of events used to compute an empirical ROC or AUC as N_{emp}, where

$$ N_{\mathrm{emp}} = N_{\mathrm{scans}}\tau\left(2\lambda_b^{\mathsf{t}} + \alpha^{\mathsf{t}}s^{\mathsf{t}}\right). \tag{29} $$

A. Source Intensity Test for Compton Detector—Monoenergetic Source and Background

1) Setup: We used the expressions derived in Section IV-E to evaluate the AUC for a Compton imaging detector when the source position $\boldsymbol{\phi}$ and background spatial distribution are assumed known, but the background intensity λ_b is unknown. The detector is an idealized parallel-plate Compton camera [20] illustrated in Fig. 1. We assumed that the position uncertainty and attenuation due to the finite thickness of the detector plates are negligible. The detector records positions and energies of interacting photons that undergo Compton scatter in the first, or scattering detector, and either Compton scatter or photoelectric absorption in the second, or absorption detector. The photoelectric and Compton cross sections for this simulation are equal to that of CdZnTe. The absorption detector is infinitely large, the scatter detector is 1 cm by 1 cm, and the plates are separated by a distance $d = 0.1\ \mathrm{cm}$. The interaction positions in both detectors are quantized into 0.1 cm $\times$ 0.1 cm pixels. We also added Gaussian noise with zero mean and standard deviation of 1 keV to the recorded energies to simulate the effect of electronic noise on the recorded energy.

The source is a 662 keV point-source located at an azimuth angle of 180 degrees and an elevation angle of 45 degrees in the far-field. For this experiment, let $\boldsymbol{\phi}$ be the azimuth and elevation

angles of the source. We estimated the true sensitivity s^{t} for the chosen source position by simulating 10^4 emitted source photons and divided the number of recorded photons by the number of simulated photons. This is justified because one can show that in the absence of background, $\bar{J}/\tau$ converges in probability to $\lambda_{\mathrm{s}} = \alpha^{\mathrm{t}} s^{\mathrm{t}}$ as $\tau \to \infty$. In this experiment, we assumed that every photon strikes the scatter detector, but only some interacting photons result in two-interaction events. We found the true sensitivity under this assumption to be $s^{\mathrm{t}} = 0.44$. We model the sensitivity by its true value.

The true background is monoenergetic with the same energy as the source. Although a monoenergetic background is not realistic, it is useful to study this case where energy alone does not differentiate between source and background photons. Let $\boldsymbol{\phi} = (\theta, \phi)$ be the angular position about the detector in spherical coordinates with inclination angle θ and azimuth angle ϕ. When the background is spatially uniform in the hemisphere above the detector, the origin direction of a photon emitted from the background obeys the density

$$f(\theta, \phi) = \frac{1}{2\pi} \sin \theta \quad \theta \in [0, \pi/2] \quad \phi \in [0, 2\pi].$$

We model the background as spatially uniform, but simulate a nonuniform background. We draw the origin direction of each background photon from the density

$$f(\theta, \phi) = \frac{\sin(2\phi + \pi/2) + 1.5}{3\pi} \sin \theta$$
$$\theta \in [0, \pi/2] \quad \phi \in [0, 2\pi].$$

The difference between the modeled and true background is a source of model mismatch.

In this example, the attribute vector x for an interacting photon is a 6×1 vector containing the x-coordinate, y-coordinate, and recorded energy. When the source is in the far field, the x and y-coordinates of the interaction in the scatter (top) detector are uniformly distributed, and the measured values are quantized to 0.1 cm because the detector is pixelated. The energy deposited in each detector follows the Klein-Nishina formula [21, p. 51]. The coordinates of the interaction in the second detector are a function of the interaction location in the first detector and the Compton scattering angle of the first interaction. The second interaction is also quantized to 0.1 cm. This true distribution [4] is computationally difficult to compute, so we use the model proposed in [4], which provides a Gaussian approximation to the uncertainties induced by detector pixelization.

2) Results: We first examined the agreement between the ROC predicted with asymptotics and the empirical ROC with the parallel-plate Compton detector. Fig. 2 shows the ROC for the true mean number of emitted source counts $\alpha^{\mathrm{t}}\tau = 60$ and known true mean number of recorded background counts $\lambda_b^{\mathrm{t}}\tau = 300$. We used $N_{\mathrm{pred}} = 208{,}782$ events to compute the predicted ROC and $N_{\mathrm{emp}} = 313{,}174$ events to compute the empirical ROC. The predicted ROC accounting for model mismatch is much closer to the empirical ROC than the predicted ROC ignoring model mismatch. Even with this small number of source
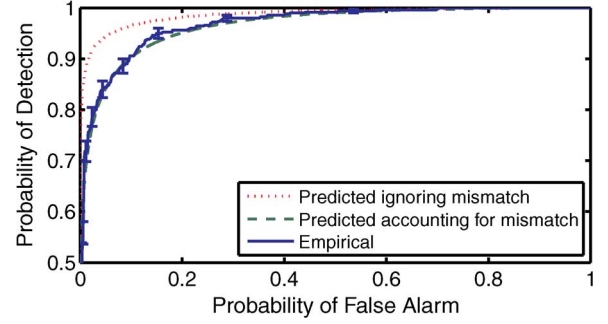

Fig. 2. Empirical and predicted ROC curves of the SIT with $\alpha^{\mathrm{t}}\tau = 60 \ \mathrm{counts}$ and $\lambda_b^{\mathrm{t}}\tau = 300 \ \mathrm{counts}$. Error bars denote standard error.
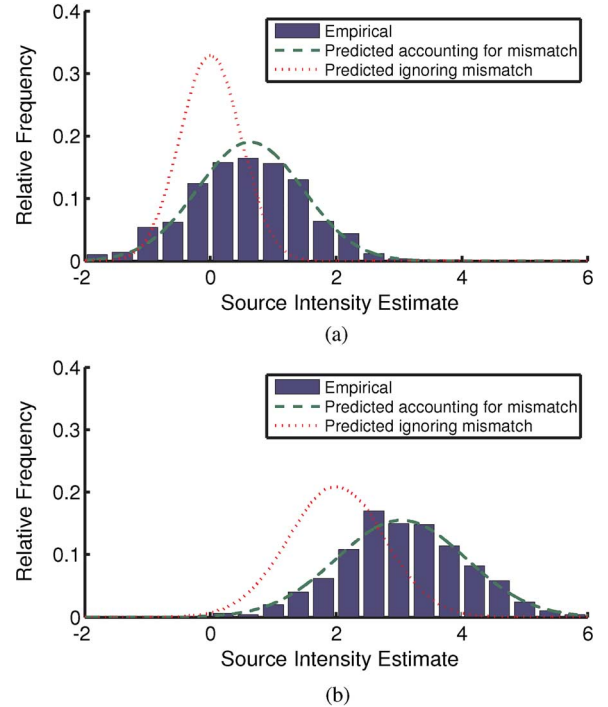

Fig. 3. Histograms and scaled asymptotic probability density functions of source intensity estimates where $\alpha^{\mathrm{t}}\tau = 60 \ \mathrm{counts}$ and $\lambda_b^{\mathrm{t}}\tau = 300 \ \mathrm{counts}$. (a) Source absent ($H_0$). (b) Source present ($H_1$).

counts per scan, the asymptotic approximation appears to be reasonable. This result shows that the proposed asymptotic analysis gives a reasonable approximation of the true ROC in this case.

The asymptotic distributions of the source intensity estimates used to calculate the ROC in Fig. 2 are also reasonably accurate approximations of the empirical distributions. Fig. 3 shows histograms of the empirical source intensity estimates and the asymptotic probability density functions (we applied scaling to the density functions to make the units consistent with the histogram). We allowed the source intensity estimates to be negative because the modeled likelihood in (2) is well-defined if α and λ_b are chosen such that $\tilde{\mathrm{p}}(x; \boldsymbol{\theta})$ is nonnegative for any observed event x, even it α is negative. The asymptotic distributions that account for mismatch fit the empirical distributions reasonably well.

We also examined the agreement between the predicted AUC and the empirical AUC as a function of scan time. We
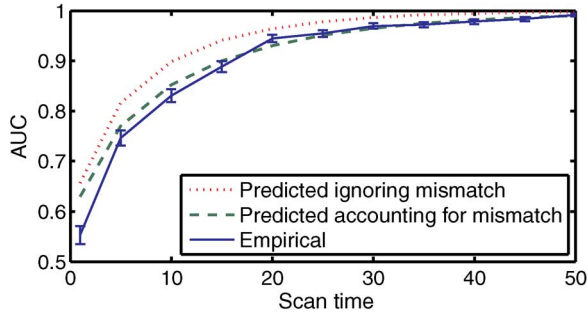
Fig. 4. Empirical and asymptotic AUC of SIT vs. τ for $\alpha^{\mathrm{t}} = 2\ \mathrm{counts/sec}$ and $\lambda_b^{\mathrm{t}} = 10\ \mathrm{counts/sec}$. Error bars denote standard error.
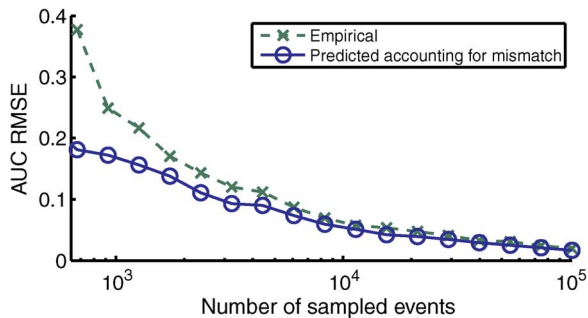

Fig. 5. RMSE of the empirical and predicted AUC versus the number of samples from the true distribution where $\alpha\tau = 20\ \mathrm{counts}$ and $\lambda_b^{\mathrm{t}}\tau = 200\ \mathrm{counts}$. The true AUC was computed empirically using 1,021,956 sampled events.

expect the agreement between the empirical and predicted AUC to improve as scan time increases because Theorems 1 and 2 give convergence results as scan time goes to infinity. Fig. 4 shows the empirical and asymptotic AUC versus scan time for source intensity $\alpha^{\mathrm{t}} = 2\ \mathrm{counts/sec}$ and background intensity $\lambda_b^{\mathrm{t}} = 10\ \mathrm{counts/sec}$. The numbers of events used to compute the asymptotic and empirical AUC are given by (28) and (29), respectively, where $N_{\mathrm{back}} = 10{,}000\ \mathrm{counts}$ and $N_{\mathrm{scans}} = 500\ \mathrm{counts}$. In this example, the predicted AUC approximates the empirical AUC well even for small scan times. In both Figs. 2 and 4, the predicted performance using the proposed expressions that account for model mismatch are much closer to the empirical performance than the predictions that ignore model mismatch.

To support the claim that the proposed asymptotic method for predicting the ROC and AUC requires less data than empirical methods, we computed the root mean square error (RMSE) [22, p. 261] of the predicted and empirical AUC with various numbers of events sampled from the true distribution. Recall that the number of events used to compute the ROC asymptotically N_{pred}, given by (28), is the number of events used to simulate a list of events from the background and another list of events containing source and background events. The number of events N_{emp} used to compute the ROC empirically is given by (29).

Fig. 5 shows the RMSE of the empirical and predicted AUC as a function of the number of photon interaction events $N_{\mathrm{pred}} = N_{\mathrm{emp}}$ sampled from the true distribution. We estimated the RMSE at each point with 1000 trials and drew N_{pred} sampled events without replacement from a pool of $10 N_{\mathrm{pred}}$ for each trial. This resampling scheme is similar to the jackknife

[23]. Fig. 5 shows that the predicted AUC using the proposed asymptotic method is more reliable in terms of RMSE than the empirical AUC, especially for a modest number of samples.

VI. CONCLUSION AND FUTURE WORK

We extended the asymptotic theory of maximum likelihood estimators under model mismatch to the case where the number of measurements is Poisson. We used this theory to develop specialized expressions for the asymptotic distributions of QML estimates and test statistics for gamma-ray imaging systems, namely (16) and (19)–(27). One can evaluate the expressions using Monte Carlo methods where the events are drawn from the true distribution of recorded photon interaction events. In practice, one can evaluate these expressions for any system model using a modest amount of data generated from the true distribution.

We provided an application of the proposed theory to detection of gamma-ray sources. The results showed that the proposed method of computing the ROC using asymptotic approximations under model mismatch agrees with empirical ROC methods in the scenarios considered. When the asymptotic mean $\tilde{\mu}_{[1]}^{(0)}$ of the source intensity estimates in the absence of a source is less than zero, the approximation may not be as accurate for small scan times. If the asymptotic approximation is poor, one can improve the approximation by considering a larger scan time. We provide examples illustrating the quality of the approximation in different conditions in [24].

Future work will apply this theory to other types of detectors and detection problems. The proposed theory is general enough to incorporate source energy in the parameter $\boldsymbol{\phi}$. We plan to use this theory to study detection performance in polyenergetic backgrounds. We also plan to study the performance of physical detectors using real data.

APPENDIX

A. Proof of Theorem 1

Let the probability measure $v(H) \triangleq \Pr(x \in H)$ be the true distribution of a single recorded attribute. Since v is typically unknown, we model it with the probability measure $\tilde{v}_{\boldsymbol{\theta}}$. We now make the following assumptions:

Assumption 1: The attributes of distinct photon interaction events are i.i.d.

Assumption 2: Assume that $\tilde{v}_{\boldsymbol{\theta}}$ has a Radon-Nikodym density with respect to a base measure μ.

By Assumption 2, there exists a function $\tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ such that

$$\tilde{v}(H) = \int_H \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) d\mu(x).$$

The true mean number of recorded photons is typically unknown, so we model it as

$$J_\tau \sim \mathrm{Poisson}\left(\tilde{\lambda}(\boldsymbol{\theta})\tau\right).$$

One can show that Assumptions 1 and 2 imply that the Radon-Nikodym density of the modeled distribution of J_τ observed attributes and J_τ is given by (2). The Radon-Nikodym density in

(2) is the likelihood of $\boldsymbol{\theta}$ given the number of recorded attributes and their values. The same formula is given in [13].

We have now shown that the modeled likelihood function exists, but we have yet to show the existence of a maximizer over the parameter space $\boldsymbol{\theta}$. To show existence, we introduce three more assumptions:

Assumption 3: $\tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$.

Assumption 4: $\tilde{\lambda}(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$.

Assumption 5: Θ is a compact subset of $\mathbb{R}^d$.

Assumptions 3, 4, and 5 and the existence of the Radon-Nikodym density of the modeled distribution implies that

$$\boldsymbol{\theta}_\tau \triangleq \arg \max_{\boldsymbol{\theta} \in \Theta} \tilde{\mathsf{p}}(x_1, x_2, \ldots, x_J, J; \boldsymbol{\theta})$$

exists.

Additional suitable regularity conditions for Theorem 1 are:

Assumption 6: $|\tilde{\mathsf{p}}(x; \boldsymbol{\theta})| < m(x)$ for all $\boldsymbol{\theta} \in \Theta$ for some $m(x)$ that is integrable with respect to the true attribute distribution, i.e., $\int_{\mathcal{R}} m(x)\mathsf{p}(x)dx$ is finite.

1) Assumption 7: The solution to (6) is unique.

In this proof, Assumptions 1–7 are assumed satisfied.

Proof: We use the result for strong consistency under model mismatch with a nonrandom number of measurements in [6] to prove strong consistency when the number of measurements in Poisson.

We divide the scan of length τ into n intervals of unit duration. We can assume without loss of generality that τ is an integer because our choice of scan interval duration is arbitrary. Treating the scan of duration $n = \tau$ as n independent scans of unit duration leads to the following restatement of the log-likelihood that differs only in terms constant with respect to $\boldsymbol{\theta}$:

$$\log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}) \equiv \left(J_k \log \left(\tilde{\lambda}(\boldsymbol{\theta}) \right) - \tilde{\lambda}(\boldsymbol{\theta}) - \log(J_k!) \right.$$
$$\left. + \sum_{j=1}^{J_k} \log \tilde{\mathsf{p}}(x_{kj}; \boldsymbol{\theta}) \right) \quad (30)$$

where $J_k \sim \text{Poisson}(\lambda_s)$, $J_\tau = \sum_{k=1}^n J_k$ by the summation property of independent Poisson random variables [22, p. 196], x_{kj} is the attribute vector of the jth event to occur during the kth scan. Since the modeled mean number of total received counts is $\tilde{J}(\boldsymbol{\theta})$, the modeled mean number of counts per scan interval is $\frac{\tilde{J}(\boldsymbol{\theta})}{\tau} = \tilde{\lambda}(\boldsymbol{\theta})$ by the summation property of independent Poisson random variables.

Assumption 6 implies that

$$\log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}) = J_\tau \log \left(\tilde{\lambda}(\boldsymbol{\theta}) \right) - \tau \tilde{\lambda}(\boldsymbol{\theta}) - \log(J_\tau!) + \sum_{j=1}^{J_\tau} \log \tilde{\mathsf{p}}(x_j; \boldsymbol{\theta})$$
$$\leq J_\tau \log \left(\tilde{\lambda}(\boldsymbol{\theta}) \right) - \tau \tilde{\lambda}(\boldsymbol{\theta}) - \log(J_\tau!) + \sum_{j=1}^{J_k} m(x_j)$$
$$\triangleq f(\boldsymbol{x}),$$

so $\log \tilde{\mathsf{p}}(\boldsymbol{x}; \boldsymbol{\theta}) < f(\boldsymbol{x})$ for all $\boldsymbol{\theta} \in \Theta$ and $f(\boldsymbol{x})$ is integrable with respect to the true distribution because it is the sum of integrable functions.

Strong convergence of $\tilde{\boldsymbol{\theta}}_\tau$ to $\tilde{\boldsymbol{\mu}}$ follows by [25, Theorem 2.1].

To show that $\tilde{\boldsymbol{\mu}}$ is the member of the parameter space that minimizes the Kullback-Leibler divergence between the model and the true distributions, we make the following additional assumptions:

Assumption 8: The probability measure v characterizing the true distribution is absolutely continuous with respect to the base measure μ, with Radon-Nikodym density $\mathsf{p}(x)$ for $x \in \mathcal{R}$.

Assumption 9: $\mathsf{E}[\log \mathsf{p}(x)]$, the logarithm of the true density of a single attribute, exists and is finite.

Assumption 8 guarantees that the true distribution of the list of recorded attributes is a Radon-Nikodym density. Assumption 9 implies that $\mathsf{E}[\log \mathsf{p}(\boldsymbol{x})]$, the logarithm the density of a list of attributes, exists by the linearity of expectation. Thus, the Kullback-Leibler divergence $\mathsf{E}\left[\log\left(\frac{\mathsf{p}(\boldsymbol{x})}{\tilde{\mathsf{p}}(\boldsymbol{x};\boldsymbol{\theta})}\right)\right]$ exists.

The existence of the QMLE and Assumptions 6–9 satisfy the conditions of Theorem 2.2 of [6], which gives strong convergence of the estimates to the element $\tilde{\boldsymbol{\mu}} \in \Theta$ that minimizes $\mathsf{E}\left[\log\left(\frac{\mathsf{p}(\boldsymbol{x})}{\tilde{\mathsf{p}}(\boldsymbol{x};\boldsymbol{\theta})}\right)\right]$. ∎

B. Proof of Theorem 2

We state the following theorem, which is used in the proof of Theorem 2:

Theorem 3: Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \in \mathbb{R}^d$ be an i.i.d. sequence of random vectors with $\mathsf{E}[\boldsymbol{x}_k] = \boldsymbol{m}$ and $\mathsf{E}[(\boldsymbol{x}_k - \boldsymbol{m})(\boldsymbol{x}_k - \boldsymbol{m})^T] = \Sigma$ for $k = 1, 2, \ldots$, and let $J_\tau \sim \text{Poisson}(\lambda\tau)$. Then as $\tau \to \infty$

$$S_\tau \triangleq \frac{1}{\sqrt{\tau}} \left(\sum_{i=1}^{J_\tau} \boldsymbol{x}_i - \lambda\tau\boldsymbol{m} \right) \xrightarrow{d} N\left(0, \lambda(\Sigma + \boldsymbol{m}\boldsymbol{m}^T)\right).$$

We omit the proof of Theorem 3 for brevity. The proof shows that the characteristic function of the random sum converges pointwise to the characteristic function of a normal random vector with mean zero and covariance $\lambda(\Sigma + \boldsymbol{m}\boldsymbol{m}^T)$.

We make the following assumptions to guarantee the asymptotic normality of the QMLE.

Assumption 10: $\log \tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ is continuously differentiable in $\boldsymbol{\theta}$ for each $x \in \mathcal{R}$.

Assumption 11: $\log \tilde{\lambda}(\boldsymbol{\theta})$ and $\tilde{\lambda}(\boldsymbol{\theta})$ are continuously differentiable in $\boldsymbol{\theta}$.

Assumption 12: $\nabla_{\boldsymbol{\theta}} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ and $\nabla_{\boldsymbol{\theta}}^2 \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ are measurable in x for each $\boldsymbol{\theta} \in \Theta$.

Assumption 13: The elements of AA^T, where $A \triangleq \nabla_{\boldsymbol{\theta}} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) + \nabla_{\boldsymbol{\theta}} \log \tilde{\lambda}(\boldsymbol{\theta})$, are dominated by functions integrable with respect to the true distribution.

Assumption 14: The elements of $\nabla_{\boldsymbol{\theta}}^2 \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ are dominated by functions integrable with respect to the true distribution.

Assumption 15: $\tilde{\boldsymbol{\mu}}$ is an interior point of Θ.

Assumption 16: $\tilde{\mathsf{G}}(\tilde{\boldsymbol{\mu}})$ is nonsingular.

Assumption 17: $\tilde{\mathsf{H}}(\boldsymbol{\theta})$ is invertible in an open neighborhood around $\tilde{\boldsymbol{\mu}}$.

Assumption 18: $\tilde{\boldsymbol{\theta}}_\tau \xrightarrow{p} \tilde{\boldsymbol{\mu}}$.

Proof: Let

$$\tilde{g}_\tau(\boldsymbol{\theta}) \triangleq \nabla_{\boldsymbol{\theta}} \log \mathsf{p}(\boldsymbol{x}; \boldsymbol{\theta})$$
$$= J_\tau \nabla_{\boldsymbol{\theta}} \log \left(\tilde{\lambda}(\boldsymbol{\theta}) \right) - \tau \nabla_{\boldsymbol{\theta}} \tilde{\lambda}(\boldsymbol{\theta}) + \sum_{j=1}^{J_\tau} \nabla_{\boldsymbol{\theta}} \log \mathsf{p}(x_j; \boldsymbol{\theta}),$$

and

$$\tilde{\mathsf{H}}_\tau(\boldsymbol{\theta}) \triangleq \nabla_{\boldsymbol{\theta}}^2 \log \mathsf{p}(\boldsymbol{x}; \boldsymbol{\theta})$$

$$= J_\tau \nabla_{\boldsymbol{\theta}}^2 \log \tilde{\lambda}(\boldsymbol{\theta}) - \tau \nabla_{\boldsymbol{\theta}}^2 \tilde{\lambda}(\boldsymbol{\theta}) + \sum_{j=1}^{J_\tau} \nabla_{\boldsymbol{\theta}}^2 \log \mathsf{p}(x_j; \boldsymbol{\theta}),$$

which both exist by Assumptions 10, 11, and 12. By the mean value theorem,

$$\tilde{g}_\tau(\check{\boldsymbol{\theta}}_\tau) - \tilde{g}_\tau(\tilde{\boldsymbol{\mu}}) = \tilde{\mathsf{H}}_\tau(\check{\boldsymbol{\theta}}_\tau) \left(\check{\boldsymbol{\theta}}_\tau - \tilde{\boldsymbol{\mu}} \right),$$

for some $\check{\boldsymbol{\theta}}_\tau$ such that $\check{\boldsymbol{\theta}}_\tau$ lies on the line segment between $\tilde{\boldsymbol{\mu}}$ and $\check{\boldsymbol{\theta}}_\tau$. Since $\tilde{g}_\tau(\check{\boldsymbol{\theta}}_\tau) = 0$ by the definition of the QMLE,

$$\sqrt{\tau} \left(\check{\boldsymbol{\theta}}_\tau - \tilde{\boldsymbol{\mu}} \right) = \left(-\frac{1}{\tau} \tilde{\mathsf{H}}_\tau(\check{\boldsymbol{\theta}}_\tau) \right)^{-1} \left(\frac{1}{\sqrt{\tau}} \tilde{g}_\tau(\tilde{\boldsymbol{\mu}}) \right). \quad (31)$$

From (6), we have

$$\nabla_{\boldsymbol{\theta}} \tilde{\lambda}(\tilde{\boldsymbol{\mu}}) / \lambda_{\mathsf{s}} = \mathsf{E} \left[\nabla_{\boldsymbol{\theta}} \log \mathsf{p}(x_j; \tilde{\boldsymbol{\mu}}) \right] + \nabla_{\boldsymbol{\theta}} \log \left(\tilde{\lambda}(\tilde{\boldsymbol{\mu}}) \right). \quad (32)$$

By (32) and Theorem 3

$$\frac{1}{\sqrt{\tau}} \tilde{g}_\tau(\tilde{\boldsymbol{\mu}}) \xrightarrow{d} \mathcal{N} \left(0, \tilde{\mathsf{G}}(\tilde{\boldsymbol{\mu}}) \right) \text{ as } \tau \to \infty$$

where $\tilde{\mathsf{G}}(\boldsymbol{\theta})$, which exists by Assumption 13, is defined in (9). By an extension of the weak law of large numbers to Poisson random sums and by Assumptions 15 and 18, we have that

$$-\frac{1}{\tau} \tilde{\mathsf{H}}_\tau(\check{\boldsymbol{\theta}}_\tau) \xrightarrow{p} \tilde{\mathsf{H}}(\tilde{\boldsymbol{\mu}}) \text{ as } \tau \to \infty$$

where $\tilde{\mathsf{H}}(\boldsymbol{\theta})$, which exists by Assumption 14, is defined in (10). Combining (9), (10), (31), and using Slutsky's theorem [26, p. 39],

$$\sqrt{n} \left(\check{\boldsymbol{\theta}}_\tau - \tilde{\boldsymbol{\mu}} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma(\tilde{\boldsymbol{\mu}}) \right)$$

where $\Sigma(\tilde{\boldsymbol{\mu}})$, which exists and is nonsingular by Assumptions 16 and 17, is defined in (8). ∎

C. Verification of Regularity Conditions

This Appendix describes the methods by which one could verify Assumptions 1–7 needed for Theorem 1 and Assumptions 10–18 required for Theorem 2 using the example system in Section V.

The recorded attributes are i.i.d. because of the chosen system setup, and one can verify that Assumptions 2–4, 6, and 10–12 are satisfied. Assumption 5 is satisfied if one places an upper bound on the set of possible source and background intensity estimates. Assumption 7 is satisfied in all but degenerate cases where there are multiple sources of exactly the same intensity, which are of little practical interest. Assumptions 13 and 14 are difficult to verify directly without an expression for the true distribution of recorded attributes. However, we expect that the true distribution is well behaved so that the continuous gradient and Hessian of the log-likelihood are integrable with respect to the

true distribution. One can verify Assumptions 15–17 for a particular case after computing $\tilde{\boldsymbol{\mu}}$, $\tilde{\mathsf{G}}(\boldsymbol{\theta})$, and $\tilde{\mathsf{H}}(\boldsymbol{\theta})$. Assumption 18 follows from Theorem 1.

D. Sample Calculations of $\mathsf{G}(\tilde{\boldsymbol{\mu}})$ and $\tilde{\mathsf{H}}(\tilde{\boldsymbol{\mu}})$

We provide calculations for the [1, 1] elements of $\tilde{\mathsf{G}}(\tilde{\boldsymbol{\mu}})$ and $\tilde{\mathsf{H}}(\tilde{\boldsymbol{\mu}})$ in (26) and (27). Calculations of the other elements of these matrices follows from similar calculation.

1) Calculation of $\tilde{\mathsf{G}}(\boldsymbol{\theta}_{[1,1]})$: From the definition of $\tilde{\lambda}(\boldsymbol{\theta})$ in (13)

$$\frac{\partial}{\partial \alpha} \log \tilde{\lambda}(\boldsymbol{\theta}) = \frac{\tilde{s}(\boldsymbol{\phi})}{\alpha \tilde{s}(\boldsymbol{\phi}) + \lambda_b}. \quad (33)$$

From the mixture-model definition of $\tilde{\mathsf{p}}(x; \boldsymbol{\theta})$ in (14)

$$\frac{\partial}{\partial \alpha} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) = \frac{\tilde{s}(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi})}{\alpha \tilde{s}(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi}) + \lambda_b \tilde{\mathsf{p}}_{\mathsf{B}}(x)} - \frac{\tilde{s}(\boldsymbol{\phi})}{\alpha \tilde{s}(\boldsymbol{\phi}) + \lambda_b}. \quad (34)$$

Substituting (33) and (34) into the definition of $\tilde{\mathsf{G}}(\tilde{\boldsymbol{\mu}})$ in (9)

$$\tilde{\mathsf{G}}(\boldsymbol{\theta}) = \lambda_{\mathsf{s}} \mathsf{E} \left[\frac{\tilde{s}^2(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}^2(x; \boldsymbol{\phi})}{(\alpha \tilde{s}(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi}) + \lambda_b \tilde{\mathsf{p}}_{\mathsf{B}}(x))^2} \right]$$

$$= \frac{\lambda_{\mathsf{s}} \tilde{s}^2(\boldsymbol{\phi})}{\tilde{\lambda}^2(\boldsymbol{\theta})} \mathsf{E} \left[\frac{\tilde{\mathsf{p}}_{\mathsf{S}}^2(x; \boldsymbol{\phi})}{\tilde{\mathsf{p}}^2(x; \boldsymbol{\theta})} \right]. \quad (35)$$

2) Calculation of $\tilde{\mathsf{H}}(\boldsymbol{\theta}_{[1,1]})$: Taking derivatives with respect to α

$$\frac{\partial^2}{\partial \alpha^2} \log \tilde{\lambda}(\boldsymbol{\theta}) = -\frac{\tilde{s}(\boldsymbol{\phi})}{(\alpha \tilde{s}^2(\boldsymbol{\phi}) + \lambda_b)} \quad (36)$$

$$\frac{\partial^2}{\partial \alpha^2} \tilde{\lambda}(\boldsymbol{\theta}) = 0, \quad (37)$$

and

$$\frac{\partial^2}{\partial \alpha^2} \log \tilde{\mathsf{p}}(x; \boldsymbol{\theta}) = \frac{-\tilde{s}^2(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}^2(x; \boldsymbol{\phi})}{(\alpha \tilde{s}(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi}) + \lambda_b \tilde{\mathsf{p}}_{\mathsf{B}}(x))^2}$$
$$+ \frac{\tilde{s}^2(\boldsymbol{\phi})}{(\alpha \tilde{s}^2(\boldsymbol{\phi}) + \lambda_b)^2}. \quad (38)$$

Substituting (36), (37), and (38) into (10), we obtain

$$\tilde{\mathsf{H}}(\boldsymbol{\theta}) = \lambda_{\mathsf{s}} \mathsf{E} \left[\frac{\tilde{s}^2(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}^2(x; \boldsymbol{\phi})}{(\alpha \tilde{s}(\boldsymbol{\phi}) \tilde{\mathsf{p}}_{\mathsf{S}}(x; \boldsymbol{\phi}) + \lambda_b \tilde{\mathsf{p}}_{\mathsf{B}}(x))^2} \right]$$

$$= \frac{\lambda_{\mathsf{s}} \tilde{s}^2(\boldsymbol{\phi})}{\tilde{\lambda}^2(\boldsymbol{\theta})} \mathsf{E} \left[\frac{\tilde{\mathsf{p}}_{\mathsf{S}}^2(x; \boldsymbol{\phi})}{\tilde{\mathsf{p}}^2(x; \boldsymbol{\theta})} \right].$$

REFERENCES

[1] S. J. Wilderman, N. H. Clinthorne, J. A. Fessler, C.-H. Hua, and W. L. Rogers, "List mode EM reconstruction of Compton scatter camera images in 3-D," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2000, vol. 2, pp. 15 292–15 295.

[2] L. Mihailescu, K. Vetter, M. Burks, E. Hull, and W. Craig, "SPEIR: A Ge Compton camera," *Nucl. Instrum. Methods in Phys. Res. Sect. A*, vol. 570, no. 1, pp. 89–100, Jan. 2007.

[3] Z. He, W. Li, G. F. Knoll, D. K. Wehe, J. Berry, and C. M. Stahle, "3-D position sensitive CdZnTe gamma-ray spectrometers," *Nucl. Instrum. Methods in Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Assoc. Equip.*, pp. 173–178, Feb. 1999.

[4] D. Xu and Z. He, "Gamma-ray energy-imaging integrated spectral deconvolution," *Nucl. Instr. Meth. Phys. Res. A.*, vol. 574, no. 1, pp. 98–109, April 2007.

[5] D. J. Lingenfelter, J. A. Fessler, C. D. Scott, and Z. He, "Benefits of position-sensitive detectors for radioactive source detection," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4473–4483, Sep. 2010.

[6] H. White, "Maximum likelihood estimation of mispecified models," *Econometrica*, vol. 50, no. 1, pp. 1–25, 1982.

[7] R. D. Evans, *The Atomic Nucleus*. New York: McGraw-Hill, 1955.

[8] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.

[9] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[10] *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later*, T. B. Fomby and R. C. Hill, Eds. Kidlington, U.K.: Elsevier, 2003, vol. 17 (Adv. Econometr.).

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. New York: Prentice-Hall, 1998.

[12] L. C. Parra and H. H. Barrett, "List-mode likelihood-EM algorithm and noise estimation demonstrated on 2D-PET," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 228–235, Apr. 1998.

[13] H. H. Barrett, T. White, and L. C. Parra, "List-mode likelihood," *J. Opt. Soc. Amer. A*, vol. 14, no. 1, pp. 2914–2923, Nov. 1997.

[14] R. Billingsley, *Probability and Measure*, 2nd ed. New York: Wiley, 1986.

[15] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. New York: Prentice-Hall, 1993.

[17] D. J. C. MacKay, "Introduction to Monte Carlo Methods," in *Learning in Graphical Models*, M. I. Jordan, Ed. Dordrencht: Kluwer, 1998, pp. 175–204.

[18] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, April 1982.

[19] L. Yan, R. Dodier, M. C. Mozer, and R. Wolniewicz, "Optimizing classifier performance via an approximation to the Wilcoxon-Mann-Whitney statistic," in *Proc. 20th Int. Conf. Mach. Learn.*, Aug. 2003.

[20] V. Maxim, M. Frandeș, and R. Prost, "Analytical inversion of the Compton transform using the full set of available projections," *Inverse Probl.*, vol. 25, no. 9, p. 095001+, Sep. 2009.

[21] G. F. Knoll, *Radiation Detection and Measurement*, 3rd ed. New York: Wiley, 2000.

[22] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 2002.

[23] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: SIAM, 1982.

[24] D. J. Lingenfelter, "Source detection and image reconstruction with position-sensitive gamma-ray detectors," Ph.D. dissertation, Univ. Michigan, Ann Arbor, to be published.

[25] H. White, "Consequences and detection of mispecified nonlinear regression models," *J. Amer. Statist. Assoc.*, vol. 76, no. 374, pp. 419–433, 1981.

[26] T. S. Ferguson, *A Course in Large Sample Theory*. London, U.K.: Chapman and Hall, 1996.

Daniel J. Lingenfelter (S'09) received the B.S.E. degree in electrical engineering in 2006 and the M.S.E. degree in electrical engineering systems in 2009, both from the University of Michigan, Ann Arbor.

He is currently a graduate student with the Department of Electrical Engineering and Computer Science, University of Michigan. His research interests include statistical signal processing, particularly imaging and detection theory.

Jeffrey A. Fessler (S'83–M'90–SM'00–F'06) received the B.S.E.E. degree from Purdue University, West Lafayette, IN, in 1985, the M.S.E.E. and M.S. degrees from Stanford University, Stanford, CA, in 1986 and 1989, respectively. From 1985 to 1988, he was a National Science Foundation Graduate Fellow at Stanford, where he received the Ph.D. degree in electrical engineering in 1990.

He has been with the University of Michigan since 1990. From 1991 to 1992, he was a Department of Energy Alexander Hollaender Postdoctoral Fellow in the Division of Nuclear Medicine. From 1993 to 1995, he was an Assistant Professor in the Nuclear Medicine and the Bioengineering Program. He is now a Professor in the Departments of Electrical Engineering and Computer Science, Radiology, and Biomedical Engineering. His research interests are in statistical aspects of imaging problems, and he has supervised doctoral research in PET, SPECT, X-ray CT, MRI, and optical imaging problems.

Dr. Fessler received the Francois Erbsmann award for his IPMI93 presentation. He became a Fellow of the IEEE in 2006 for contributions to the theory and practice of image reconstruction. He has served as an Associate Editor for the IEEE TRANSACTIONS ON MEDICAL IMAGING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and the IEEE SIGNAL PROCESSING LETTERS. He was Co-Chair of the 1997 SPIE Conference on Image Reconstruction and Restoration, Technical Program Co-Chair of the 2002 IEEE International Symposium on Biomedical Imaging (ISBI), and General Chair of ISBI 2007.

Clayton D. Scott (S'99–M'04) received the A.B. degree in mathematics from Harvard University, Cambridge, MA, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 2000 and 2004, respectively.

He was a Postdoctoral Fellow with the Department of Statistics at Rice University, and is currently an Assistant Professor in the Departments of Electrical Engineering and Computer Science and of Statistics, University of Michigan, Ann Arbor. His research interests include pattern recognition, machine learning, statistical signal processing, and applications.

Zhong He (M'98–SM'00) received the B.S. degree in physics from Tsinghua University, and the M.S. degree at High Energy Physics Institute, Beijing, China. He received the Ph.D. degree from Southampton University, U.K., in 1993.

Since 1994, he has been with the University of Michigan, initially as a Research Scientist and joined the teaching faculty in 1998. His research during 1986 to 1995 focused on scintillation detectors, such as NaI, CsI, BGO, LSO, coupled with position-sensitive PM tubes, or silicon photodoides. Since 1995, his research has focused mostly on the development of coplanar-grid and three-dimensional position-sensitive wide band-gap semiconductor gamma-ray imaging spectrometers. He pioneered the three-dimensional position-sensitive single polarity charge sensing technology which has received increasing interests in homeland security, nuclear nonproliferation, astrophysics, and planetary sciences. He has also worked on high-pressure Xe gamma-ray detectors. His research has covered all three major types of radiation sensors, semiconductor, scintillation and gas detectors. Since 1999, his group has developed very low noise, room-temperature operation charge sensing application specific integrated circuitries (ASICs) in collaboration with Gamma-Medica-Ideas AS in Norway, and Brookhaven National Laboratory, to measure signal amplitude and charge drift-time simultaneously in semiconductors. He has graduated 11 Ph.D. students.