Systems & Control Letters 75 (2015) 69–76

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Performance bounds for mismatched decision schemes with Poisson process observations

Chetan D. Pahlajani [a,*], Jianxin Sun [b], Ioannis Poulakakis [b], Herbert G. Tanner [b]

[a] Mathematics, Indian Institute of Technology Gandhinagar, Chandkheda, Ahmedabad 382424, India

[b] Department of Mechanical Engineering, University of Delaware, Newark, DE 19716, USA

ARTICLE INFO

Article history:
Received 16 February 2014
Received in revised form
10 November 2014
Accepted 13 November 2014
Available online 4 December 2014

Keywords:
Decision making
Poisson process
Model mismatch
Nuclear detection

ABSTRACT

This paper develops a framework for analyzing the performance loss in fixed time interval decision algorithms that are based on observations of time-inhomogeneous Poisson processes, when some parameters characterizing the observation process are not known exactly. Key to the development is the formulation of an analytically computable performance metric which can be used in lieu of the true, but intractable, error probabilities. The proposed metric is obtained by identifying analytical upper bounds on the error probabilities in terms of the uncertain parameters. Using these tools, it is shown that performance degrades gracefully as long as the true values of the parameters remain within a neighborhood of the nominal values used in decision making. The results find direct application to problems of detecting illicit nuclear materials in transit.

1. Introduction

Many physical processes of interest are characterized by sequences of discrete events occurring randomly in time, modeled mathematically as *point processes* [1,2]; an important class of the latter is the collection of Poisson processes, used to capture the underlying physics in queueing theory [1], optical communications [3], neuroscience [4], and nuclear detection [5]. Problems of decision making between two hypotheses on the basis of Poisson (and more general point) process observations have long been studied [1,2,6–9]. For the Poisson case, the optimal Neyman–Pearson rule is known to be given by a Likelihood Ratio Test (LRT), where the decision is based on comparing a likelihood ratio formed by the observations against a suitable threshold. The functional form of the likelihood ratio is determined by the intensities of the Poisson process under the two hypotheses.

In many situations, however, these intensities are subject to uncertainty due to incomplete knowledge of the model. For instance, Poisson process intensities under the two hypotheses may be specified in terms of a parameter vector whose exact value – assumed to be the same under both hypotheses – may not be accurately known. One approach to ensuring acceptable performance of decision algorithms over a *range* of parameter values is to apply robust techniques [7,10,11]. Then, to identify the parameters most crucial for robustness, one needs to understand the relative impact of parameter uncertainty on decision-making performance. The challenge now is that performance is measured by error probabilities, the analytical computation of which is extremely difficult, if not impossible. This observation sets the stage for the present research, which aims at establishing an alternate analytically tractable performance metric which can shed light on the above problem. We note that although the setting described is, at first glance, reminiscent of composite hypothesis testing [11], there are some subtle differences. Indeed, the parameter vector in the framework above takes the *same*, albeit imperfectly known, value under both hypotheses; Remark 1 describes how this is a natural assumption in some instances. In contrast, composite hypothesis testing typically assumes that the parameter vector takes *different* values under the two hypotheses in disjoint subsets of parameter space.

The mathematical models and techniques described above find natural application in the field of *nuclear detection*. A particularly challenging instance of the problem of nuclear detection is that of detecting illicit Special Nuclear Material (SNM) in transit [5,12–14]. Assuming that the moving target is identified, one is asked to decide whether that target is a carrier of an SNM radiation source, using radiation count data from a spatially dispersed network of inexpensive Geiger counters or scintillators. The critical question is whether the photons recorded by the counters are solely due

* Corresponding author.
E-mail addresses: cdpahlajani@iitgn.ac.in (C.D. Pahlajani), jxsun@udel.edu (J. Sun), poulakas@udel.edu (I. Poulakakis), btanner@udel.edu (H.G. Tanner).

http://dx.doi.org/10.1016/j.sysconle.2014.11.006
0167-6911/

to ubiquitous background radiation or whether they also contain emissions from a moving source. Since both background and source photon arrivals at a sensor can be modeled by Poisson processes, one is faced with a problem of detecting a Poisson signal buried inside another signal of similar nature and magnitude, within a small time interval. Furthermore, one of these processes is actually time-inhomogeneous, since the perceived source intensity incident at a sensor varies with the inverse square of the distance between source and sensor [5].

This decision problem has been studied in a fixed interval framework [15,16], i.e., when data is collected by sensors over a fixed time interval, at the end of which a decision is made. The likelihood ratio has been identified in terms of the problem parameters, including the motion of the source [15]. Chernoff upper bounds [17–19] on the error probabilities for the corresponding LRT have been computed [16], identifying the analytical dependence of the bounds on the problem parameters. To fully exploit these insights in a field-deployable nuclear detector network system, however, one needs to recognize and account for the presence of model uncertainty, a dominant source of which is *radiation clutter* [13]: the myriad "nuisance sources" and spatiotemporal environmental variations whose cumulative effect is to create a dynamic and imperfectly modeled background.

In this paper, we study the effect of imperfectly known intensities on a class of decision problems for Poisson processes which include, as special cases, several scenarios encountered in nuclear detection. Working with a parametrized family of models, where each value of the parameter vector corresponds to a specific choice of intensities, we obtain Chernoff upper bounds on the error probabilities for decision schemes with *mismatch* [20,21]. By the latter, we mean that the decision rule is an LRT based on some nominal model which may be different from the true model governing the stochastic processes of interest. The Chernoff bounds, or equivalently, the exponents in the bounds, now furnish a performance measure which can be analytically characterized in terms of the problem parameters under the true and nominal models. Further, the exponents are seen to vary smoothly when the true model is a sufficiently small perturbation about the nominal one, implying that at least locally, performance degrades gracefully as parameters deviate from their nominal (known) values.

2. Background

We start with a binary hypothesis testing problem. The probabilistic setup consists of a measurable space $(\Omega, \mathscr{F})$ supporting a k-dimensional counting process $\boldsymbol{N}_t \triangleq (N_t(1), \ldots, N_t(k))$, $t \in [0, T]$, together with probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$, with $\mathbb{P}_1 \ll \mathbb{P}_0$, i.e., $\mathbb{P}_1$ is absolutely continuous with respect to $\mathbb{P}_0$. Here, $\mathbb{P}_j$ denotes the probability measure under hypothesis H_j, $j \in \{0, 1\}$. We assume that the components $N_t(i)$, $i \in \{1, \ldots, k\}$, of $\boldsymbol{N}_t$ are independent Poisson processes under each $\mathbb{P}_j$, $j \in \{0, 1\}$, having intensity $\beta_i(t)$ with respect to $\mathbb{P}_0$, and intensity $\beta_i(t) + \nu_i(t)$ with respect to $\mathbb{P}_1$. The functions $\beta_i(\cdot)$ and $\nu_i(\cdot)$ are assumed to be positive, continuous, and bounded away from zero. The problem is to decide, based on the observed sample path of $\boldsymbol{N}_t$ over $t \in [0, T]$, between hypotheses H_0 and H_1.

Let $\mu_i(t)$ be the ratio of intensities for $N_t(i)$ under hypothesis H_1 versus H_0, i.e., $\mu_i(t) \triangleq 1 + \nu_i(t)/\beta_i(t)$. With $(\tau_n(i) : n \geq 1)$ denoting the jump times of $N_t(i)$, and after defining

$$L_t(i) \triangleq \exp\left(-\int_0^t \nu_i(s)ds\right) \prod_{n=1}^{N_t(i)} \mu_i(\tau_n(i)), \tag{1}$$

let $\{L_t : t \in [0, T]\}$ be the stochastic process

$$L_t \triangleq \prod_{i=1}^{k} L_t(i). \tag{2}$$

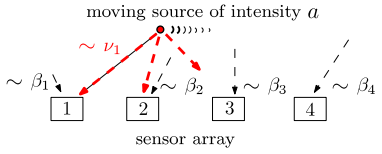


Fig. 1. Setup for a basic networked fixed-interval moving source detection scenario. Sensors are indexed by $\{1, 2, \ldots\}$ and receive photons that can be attributed either to background (thin dashed arrows) or to source radiation (thick red dashed arrows). Background intensity at sensor i location is characterized by β_i, and the intensity of the source is determined by the parameter a. The intensity of this source ν_i, as perceived at a sensor i, depends on the distance between sensor and source, r_i. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

By convention, $\prod_{n=1}^{0}(\cdot) = 1$. The optimal Neyman–Pearson test for deciding between H_0 and H_1 is an LRT given by comparing L_T to a suitably chosen threshold $\gamma > 0$, deciding H_1 if $L_T \geq \gamma$, and H_0 if $L_T < \gamma$ [15]. The performance of the LRT can be measured in terms of the corresponding error probabilities; that is, the probability of false alarm $P_F \triangleq \mathbb{P}_0(L_T \geq \gamma)$ and the probability of miss $P_M \triangleq \mathbb{P}_1(L_T < \gamma)$. More often than not, computing P_F and P_M is analytically intractable, thereby motivating the need for easily computable upper bounds that can be used as proxies for the corresponding probabilities at the expense of some sharpness. It can be shown [16] that if one explicitly computes

$$\begin{aligned}\Lambda(p) &\triangleq \log \mathbb{E}_0[L_T^p] \\ &= \sum_{i=1}^{k} \int_0^T \left[\mu_i(s)^p - p\mu_i(s) + p - 1\right] \beta_i(s)ds,\end{aligned}$$

for $p \in \mathbb{R}$, then P_F and P_M admit the Chernoff bounds

$$\begin{aligned} P_F &\leq \exp\left(\inf_{p>0} \left[\Lambda(p) - p \log \gamma\right]\right), \\ P_M &\leq \exp\left(\inf_{p<1} \left[\Lambda(p) + (1-p) \log \gamma\right]\right). \end{aligned} \tag{3}$$

The availability of the bounds (3) in analytical form greatly facilitates the implementation of the test in many practical situations. For example, these bounds can be used [16] to devise a procedure for selecting the threshold γ so that the LRT $\{L_T \geq \gamma\}$ conforms with desired performance requirements, typically characterized by the probability of false alarm P_F being less than or equal to a desired level α.

To motivate the general treatment which follows, we begin with a concrete example of using the framework described above to detect a moving nuclear source (see Fig. 1). At the initial time $t = 0$, a moving vehicle (target) which may be a source of SNM with minimum activity $a > 0$, is identified. The target's trajectory over a fixed time interval $[0, T]$ is assumed to be known. This target is within sensing range of a spatially dispersed network of k radiation sensors, some of which may be mobile. For $i \in \{1, \ldots, k\}$, $N_t(i)$ represents the number of counts registered at sensor i up to and including time $t \in [0, T]$, while $\beta_i(t)$ represents the intensity at time t due to background at the spatial location of sensor i. In keeping with the inverse square fall-off with distance for source intensity – as is common in the relevant literature [5] – we take $\chi > 0$ as a sensor-specific cross section coefficient and $r_i(t)$ to be the distance at time t between the target and sensor i, and define the *perceived* source intensity at sensor i at time t as

$$\nu_i(t) = \frac{\chi a}{2\chi + r_i(t)^2}. \tag{4}$$

The goal is to decide, at the fixed time T, whether the counts recorded by the collection of sensors correspond solely to background radiation (hypothesis H_0), or whether they also contain

emissions from SNM carried by the target (hypothesis H_1). To achieve this goal, each sensor locally processes its observations to form $L_t(i)$ via (1), which is transmitted once, at $t = T$, to a fusion center. The latter combines the transmitted information by computing the product (2) to form L_T, which is then used to optimally decide which of the two hypotheses H_0 or H_1 is correct based on the LRT $\{L_T \geq \gamma\}$. The details of the test can be found elsewhere [15]; here, we emphasize that the performance of the test depends on the degree to which the physical parameters that participate in the computations are accurately known. These parameters may include the background intensities β_i, the intensity of the source a, the cross-section coefficient χ, and the parameters that affect the distance r_i between the target and the sensor i – e.g. the velocity of the target – that participate in ν_i computed by (4).

3. Problem formulation

A key consideration in the design and analysis of decision making systems is their performance in the presence of modeling uncertainties. In such instances, one may have a decision rule based on some nominal model of the system which is different from the true system model. For situations where the model uncertainty is caused by imperfect knowledge of problem parameters, it is of interest to assess the effect on performance of deviations of each parameter from its best known value. In the present paper, we study this question for a class of decision problems involving Poisson processes which include, as special cases, some scenarios encountered in nuclear detection.

The setting for our problem is as described in Section 2, except that we now consider a *family* of models parametrized by $\boldsymbol{\theta} \in \Theta$, where Θ is an open subset of $\mathbb{R}^d$, $d \geq 1$. Thus, we now have two families of probability measures $\{\mathbb{P}_0^{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ and $\{\mathbb{P}_1^{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$, where $\mathbb{P}_1^{\boldsymbol{\theta}} \ll \mathbb{P}_0^{\boldsymbol{\theta}}$, such that the components $N_t(i)$ of $\boldsymbol{N}_t$ now have intensities $\beta_i(t, \boldsymbol{\theta})$ and $\beta_i(t, \boldsymbol{\theta}) + \nu_i(t, \boldsymbol{\theta})$ under $\mathbb{P}_0^{\boldsymbol{\theta}}$ and $\mathbb{P}_1^{\boldsymbol{\theta}}$ respectively. The interpretation is that knowing the true value of $\boldsymbol{\theta}$ – which is the *same* under both hypotheses – is tantamount to complete knowledge of the model parameters, viz., the intensities and the underlying probability measures.

In addition to the assumptions in Section 2, we require below that the functions β_i and ν_i are smooth in $\boldsymbol{\theta}$ with uniformly bounded partial derivatives. These help control the change in model parameters in terms of changes in $\boldsymbol{\theta}$ and play a role in the proof of Proposition 3.

Assumption 1. There exist positive numbers $0 < \beta_{\min} < \beta_{\max} < \infty$ such that $\beta_i(t, \boldsymbol{\theta}) \in [\beta_{\min}, \beta_{\max}]$ for all $t \in [0, T]$, $1 \leq i \leq k$, $\boldsymbol{\theta} \in \Theta$. Further, for each $\boldsymbol{\theta} \in \Theta$, $1 \leq i \leq k$, the map $t \mapsto \beta_i(t, \boldsymbol{\theta})$ is continuous on $[0, T]$. Finally, for $t \in [0, T]$, $1 \leq i \leq k$, the map $\boldsymbol{\theta} \mapsto \beta_i(t, \boldsymbol{\theta})$ is C^1 with $\sup_{1\leq i\leq k} \sup_{t\in[0,T]} \sup_{\boldsymbol{\theta}\in\Theta} \|\nabla_{\boldsymbol{\theta}} \beta_i(t, \boldsymbol{\theta})\| < \infty$, where $\nabla_{\boldsymbol{\theta}}$ denotes the gradient with respect to $\boldsymbol{\theta}$.

Assumption 2. There exist positive numbers $0 < \nu_{\min} < \nu_{\max} < \infty$ such that $\nu_i(t, \boldsymbol{\theta}) \in [\nu_{\min}, \nu_{\max}]$ for all $t \in [0, T]$, $1 \leq i \leq k$, $\boldsymbol{\theta} \in \Theta$. Further, for each $\boldsymbol{\theta} \in \Theta$, $1 \leq i \leq k$, the map $t \mapsto \nu_i(t, \boldsymbol{\theta})$ is continuous on $[0, T]$. Finally, for $t \in [0, T]$, $1 \leq i \leq k$, the map $\boldsymbol{\theta} \mapsto \nu_i(t, \boldsymbol{\theta})$ is C^1 with $\sup_{1\leq i\leq k} \sup_{t\in[0,T]} \sup_{\boldsymbol{\theta}\in\Theta} \|\nabla_{\boldsymbol{\theta}} \nu_i(t, \boldsymbol{\theta})\| < \infty$, where $\nabla_{\boldsymbol{\theta}}$ denotes the gradient with respect to $\boldsymbol{\theta}$.

We assume that our decision scheme is an LRT (with threshold $\gamma > 0$) based on some nominal value $\hat{\boldsymbol{\theta}} \in \Theta$, while the true intensities and probability measures governing the statistics of $\boldsymbol{N}_t$ correspond to the true parameter value $\boldsymbol{\theta} \in \Theta$, which is in general different from $\hat{\boldsymbol{\theta}}$. We would like to understand how this mismatch propagates through the detection process, and specifically how it impacts the performance of the decision making algorithm.

Remark 1. We note that since our parameter $\boldsymbol{\theta}$ takes the same value under both hypotheses, our setting is different from composite hypothesis testing. The framework here has been tailored to the problem of nuclear detection in the presence of radiation clutter [13] where the dominant source of model uncertainty is the background intensity, which is the same under both the null and alternative hypotheses.

For example, in the particular case illustrated in the nuclear detection application of Section 2, we would include the parameters that are relevant to our detection problem in the array $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{k+1})$. The first k elements of this array capture the background intensities at sensors $i \in \{1, \ldots, k\}$, which, without significant loss of generality, are assumed to be constant; i.e., $\beta_i(t, \boldsymbol{\theta}) \triangleq \theta_i$. The component θ_{k+1} corresponds to the value of the intensity of the source a; namely, $a = \theta_{k+1}$.[1] We assume that the vector $\boldsymbol{\theta}$ of *true* values of the aforementioned quantities is not exactly known; instead, an estimate $\hat{\boldsymbol{\theta}}$ of these parameters is available, which may differ from $\boldsymbol{\theta}$. After stating our theoretical results in Section 4, we demonstrate them numerically in Section 5.

4. Results

The contributions are as follows. First, for any $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, Theorem 1 gives Chernoff upper bounds on the error probabilities. These exponential bounds provide a performance measure for detection with the exponents expressible in terms of the problem parameters under the true ($\boldsymbol{\theta}$) and nominal ($\hat{\boldsymbol{\theta}}$) models. Proposition 2 shows where the tightest bounds are attained and Proposition 3 establishes that for $\boldsymbol{\theta}$ near $\hat{\boldsymbol{\theta}}$, the bounds are C^1 in $\boldsymbol{\theta}$. The latter implies a form of robustness for decision making: it assures us that conservative approximations of decision performance given by the Chernoff bounds vary smoothly with respect to small perturbations in the underlying model. It is important to note that Theorem 1 and Proposition 2 are *global* in that they hold for any $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, while Proposition 3 is *local* and holds only for $\boldsymbol{\theta}$ in the vicinity of $\hat{\boldsymbol{\theta}}$. The proofs of these results will be given in the Appendix.

For $1 \leq i \leq k$, $t \in [0, T]$, $\boldsymbol{\theta} \in \Theta$, let

$$\mu_i(t, \boldsymbol{\theta}) \triangleq 1 + \frac{\nu_i(t, \boldsymbol{\theta})}{\beta_i(t, \boldsymbol{\theta})}$$

be the ratio of intensities under hypothesis H_1 versus H_0. Also, for a stochastic process C_t, and with $\tau_n(i)$ for $n \geq 1$ and $1_{(\tau_n(i)\leq t)}$ denoting the jump times of $N_t(i)$ and the indicator function on interval $(\tau_n(i) \leq t)$, respectively, let $\int_0^t C_s dN_s(i) \triangleq \sum_{n\geq 1} C_{\tau_n(i)} 1_{(\tau_n(i)\leq t)}$ for $t \in [0, T]$. For $\boldsymbol{\theta} \in \Theta$, let $\{L_t^{\boldsymbol{\theta}} : t \in [0, T]\}$ be the stochastic process

$$L_t^{\boldsymbol{\theta}} \triangleq \exp\left\{-\sum_{i=1}^k \int_0^t \nu_i(s, \boldsymbol{\theta}) ds + \sum_{i=1}^k \int_0^t \log \mu_i(s, \boldsymbol{\theta}) dN_s(i)\right\}. \quad (5)$$

Note that $L_t^{\boldsymbol{\theta}} = \prod_{i=1}^k \left\{\exp\left(-\int_0^t \nu_i(s, \boldsymbol{\theta}) ds\right) \prod_{n=1}^{N_t(i)} \mu_i(\tau_n(i), \boldsymbol{\theta})\right\}$, where $\prod_{n=1}^0(\cdot) = 1$ by convention. Eq. (5) is thus in accordance with (2) and (1). As indicated earlier, we will decide between H_0 and H_1 using the LRT $\{L_T^{\hat{\boldsymbol{\theta}}} \geq \gamma\}$ which compares the likelihood ratio based on the nominal value $\hat{\boldsymbol{\theta}}$ against a threshold $\gamma > 0$, deciding H_1 if $L_T^{\hat{\boldsymbol{\theta}}} \geq \gamma$, and H_0 if $L_T^{\hat{\boldsymbol{\theta}}} < \gamma$. Since the true probability measures correspond to the (true) parameter value $\boldsymbol{\theta} \neq \hat{\boldsymbol{\theta}}$, the probabilities of false alarm and miss are now given by

$$P_F^{(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})}(\gamma) \triangleq \mathbb{P}_0^{\boldsymbol{\theta}}\left(L_T^{\hat{\boldsymbol{\theta}}} \geq \gamma\right), \qquad P_M^{(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})}(\gamma) \triangleq \mathbb{P}_1^{\boldsymbol{\theta}}\left(L_T^{\hat{\boldsymbol{\theta}}} < \gamma\right), \quad (6)$$

[1] We note that uncertainty in position measurements can also be incorporated using more parameters. For instance, if the distance $r_i(t)$ between the target and sensor i depends on a parameter θ_{k+2} (related to the accuracy of the range measurement) and can be expressed as $f_i(t, \theta_{k+2})$, then the intensity (4) due to the source can be expressed as $\nu_i(t, \boldsymbol{\theta}) \triangleq \frac{\chi\theta_{k+1}}{2\chi + f_i(t,\theta_{k+2})^2}$.

respectively. Finally, for $p, q \in \mathbb{R}$ and $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, define the quantities $\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p)$ and $\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q)$ by

$$\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) \triangleq \log \mathbb{E}_0^{\boldsymbol{\theta}}\left[(L_T^{\hat{\boldsymbol{\theta}}})^p\right], \qquad \Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q) \triangleq \log \mathbb{E}_1^{\boldsymbol{\theta}}\left[(L_T^{\hat{\boldsymbol{\theta}}})^q\right].$$

Theorem 1. *For $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$ and $\gamma > 0$, the Chernoff bounds on $P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and $P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ are given by*

$$\begin{aligned} P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &\le \exp\left[\inf_{p>0}\left(\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) - p\log\gamma\right)\right], \\ P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &\le \exp\left[\inf_{q<0}\left(\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q) - q\log\gamma\right)\right], \end{aligned} \tag{7}$$

where $\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p)$ and $\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q)$ are explicitly computable via

$$\begin{aligned} \Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) &= \sum_{i=1}^{k}\int_0^T \Big\{\Big[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\Big]\beta_i(s,\boldsymbol{\theta}) \\ &\quad + p\Big[1-\mu_i(s,\hat{\boldsymbol{\theta}})\Big]\beta_i(s,\hat{\boldsymbol{\theta}})\Big\}\,ds, \\ \Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q) &= \sum_{i=1}^{k}\int_0^T \Big\{\Big[\mu_i(s,\hat{\boldsymbol{\theta}})^q - 1\Big]\cdot[\beta_i(s,\boldsymbol{\theta}) \\ &\quad + \nu_i(s,\boldsymbol{\theta})] + q\Big[1-\mu_i(s,\hat{\boldsymbol{\theta}})\Big]\beta_i(s,\hat{\boldsymbol{\theta}})\Big\}\,ds. \end{aligned} \tag{8}$$

The proof of the theorem is found in the Appendix.

Remark 2. For the case $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, we have $d\mathbb{P}_1^{\hat{\boldsymbol{\theta}}}/d\mathbb{P}_0^{\hat{\boldsymbol{\theta}}} = L_T^{\hat{\boldsymbol{\theta}}}$ on the σ-algebra $\mathscr{F}_T^{\boldsymbol{N}} \triangleq \sigma(\boldsymbol{N}_s : 0 \le s \le T)$, implying that the LRT $\{L_T^{\hat{\boldsymbol{\theta}}} \ge \gamma\}$ is optimal (in the Neyman–Pearson sense) for deciding between H_0 and H_1 [15]. Further, in this case, we have $\mathbb{E}_1^{\hat{\boldsymbol{\theta}}}\big[(L_T^{\hat{\boldsymbol{\theta}}})^q\big] = \mathbb{E}_0^{\hat{\boldsymbol{\theta}}}\big[(L_T^{\hat{\boldsymbol{\theta}}})^{q+1}\big]$ for all $q \in \mathbb{R}$. Hence, $\Lambda_1^{(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}})}(q) = \Lambda_0^{(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}})}(q+1)$, and both bounds in (7) can be expressed in terms of $\Lambda_0^{(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}})}$. Taking $p = q+1$, the bound on $P_M^{(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}})}(\gamma)$ can be expressed as an infimum over $p < 1$, as in (3). The resulting bounds are seen to match those in the literature [16].

Since $\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(0) = \Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(0) = 0$, we have $\inf_{p>0}(\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) - p\log\gamma) \le 0$ and $\inf_{q<0}(\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q) - q\log\gamma) \le 0$ for any choice of $\gamma > 0$. Thus, in order for the bounds in Theorem 1 to be non-trivial, we need these infima to be strictly negative. Proposition 2 describes how γ should be chosen to ensure non-triviality of the bounds, and also identifies where the infima are attained. To this end, let $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, pick $\gamma > 0$, and let

$$\begin{aligned} R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &\triangleq \inf_{p>0}\left(\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) - p\log\gamma\right), \\ R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &\triangleq \inf_{q<0}\left(\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q) - q\log\gamma\right) \end{aligned} \tag{9}$$

denote the exponents in the Chernoff bounds. We now have

Proposition 2. *For $\log\gamma \in \big(\frac{\partial\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial p}(0), \frac{\partial\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial q}(0)\big)$, there exist unique $p^* = p^*(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) > 0$ and $q^* = q^*(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) < 0$ such that*

$$\frac{\partial\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial p}(p^*) = \log\gamma = \frac{\partial\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial q}(q^*). \tag{10}$$

Moreover,

$$\begin{aligned} R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &= \Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p^*) - p^*\frac{\partial\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial p}(p^*) < 0, \\ R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &= \Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(q^*) - q^*\frac{\partial\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial q}(q^*) < 0. \end{aligned} \tag{11}$$

Hence, the tightest bounds on $P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and $P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ are given by

$$P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) \le \exp\big(R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big), \qquad P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) \le \exp\big(R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big), \tag{12}$$

and these bounds are non-trivial.

The proof is given in the Appendix.

To summarize, for any $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$ and $\gamma > 0$, Theorem 1 provides us with the performance metrics $\exp\big(R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big)$ and $\exp\big(R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big)$, while Proposition 2 describes how the exponents $R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and $R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ can be evaluated for γ properly chosen to tighten those bounds. The following result in Proposition 3 establishes that for $\boldsymbol{\theta}$ near $\hat{\boldsymbol{\theta}}$, the exponents $R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and $R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ vary smoothly in $\boldsymbol{\theta}$, thereby ensuring the smoothness in $\boldsymbol{\theta}$ of their exponentials $\exp\big(R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big)$ and $\exp\big(R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)\big)$. In order to avail of Proposition 2, we will restrict $\boldsymbol{\theta}$ to a small enough ball $B(\hat{\boldsymbol{\theta}}, \delta)$ of radius $\delta > 0$ centered at $\hat{\boldsymbol{\theta}}$, and require that $\log\gamma$ be chosen from an interval (l, r) small enough that $(l, r) \subset \big(\frac{\partial\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial p}(0), \frac{\partial\Lambda_1^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}}{\partial q}(0)\big)$ for all $\boldsymbol{\theta} \in B(\hat{\boldsymbol{\theta}}, \delta)$.

Proposition 3. *Fix $\hat{\boldsymbol{\theta}} \in \Theta$. There exist $\delta > 0$ and an interval $(l, r) \subset \mathbb{R}$ such that for all $\log\gamma \in (l, r)$, the maps $\boldsymbol{\theta} \mapsto R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and $\boldsymbol{\theta} \mapsto R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ are continuously differentiable (i.e., C^1) on the open ball $B(\hat{\boldsymbol{\theta}}, \delta)$.*

The proof is found in the Appendix.

5. Numerical visualization

To help understand how Proposition 3 applies in practice, we would like to visualize $R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma)$ and other relevant quantities in the following scenario. Our workspace is the (x, y)-plane. We have four static Geiger counters located on the x-axis at $(0, 0)$, $(2, 0)$, $(4, 0)$, $(6, 0)$ (all lengths are in meters), while a mobile target moves parallel to the x-axis, starting from $(0, 5)$ with velocity equal to 3 m/min (see Fig. 1). The total detection period is set to 2 min.

The parameter vector is taken to be $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_5)$, where θ_i, $1 \le i \le 4$, represents background radiation level at sensor i, and θ_5 is the source activity. We fix the estimated model parameters $\hat{\boldsymbol{\theta}}$ as follows: $\hat{\theta}_i = 10$, $1 \le i \le 4$, $\hat{\theta}_5 = 120$. All the position readings are assumed to be accurate. The true model parameter vector $\boldsymbol{\theta}$ is now varied around $\hat{\boldsymbol{\theta}}$. We let the true source activity θ_5 change from 90 to 150 and let the true background radiation level θ_1 at sensor 1 change from 7 to 13. All other true model parameters are the same as their estimated counterparts. With this set of $(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ pairs, Fig. 2 shows $\exp\big(R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)\big)$ and $\exp\big(R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)\big)$; as predicted by Proposition 3, both are differentiable. Fig. 3 shows a top-down view.

To probe the tightness of the proposed Chernoff bounds, we ran Monte Carlo simulations on the same scenario to compute an estimate of $P_M^{(\theta,\hat{\theta})}(1)$ and compared it with the Chernoff bounds. We collected 18 444 i.i.d. sample sets for each mesh point (θ_1, θ_5) and ran an LRT over these samples to get an estimate of $P_M^{(\theta,\hat{\theta})}(1)$ with 1% precision and 95% confidence. The results are shown in Fig. 4.

6. Conclusions

This paper brings into focus the impact of parameter mismatch on the performance of a fixed interval binary detection scheme based on time-inhomogeneous Poisson process observations. A framework is proposed, within which deciding between the two hypotheses is achieved through a likelihood ratio test (LRT). However, the test is based on a nominal model regarding the statistics

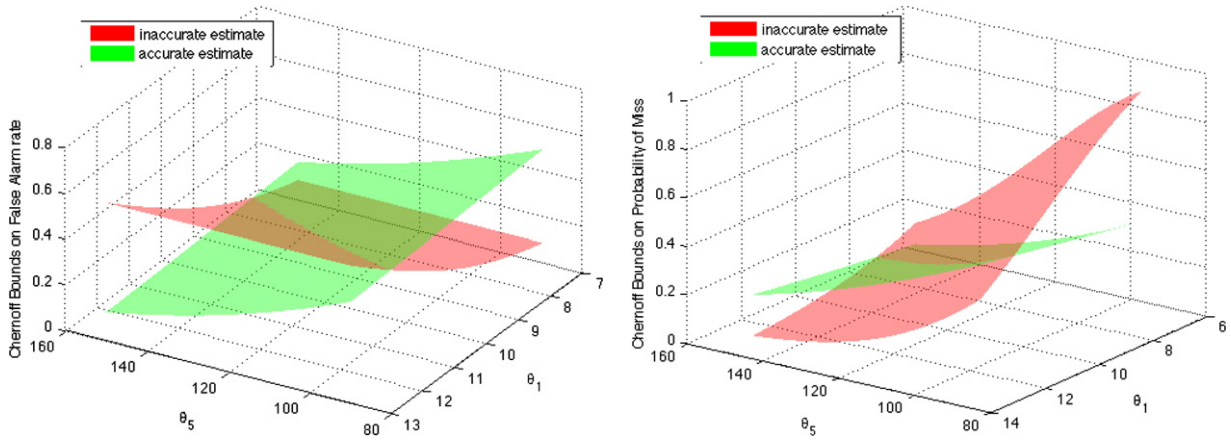
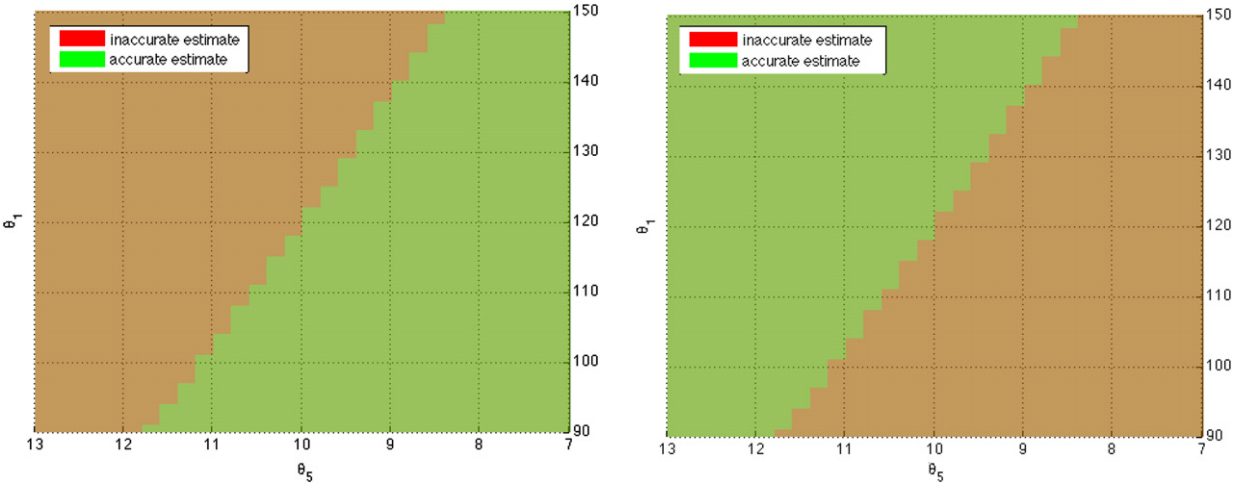


Fig. 2. Surface of Chernoff bounds on $P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)$ and $P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)$ when estimate model is inaccurate ($\hat{\theta}_1 = 10$, $\hat{\theta}_5 = 120$, in red) and accurate ($\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$, in green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Top-down view of the Chernoff bounds on probability of false alarm.

(b) Top-down view of the Chernoff bounds on probability of miss.

Fig. 3. Although for some $\boldsymbol{\theta}$ values, the inaccurate model yields lower values for the bound on probability of miss $\exp\left(R_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)\right)$ compared to the accurate model, the former also has higher values for the bound on probability of false alarm $\exp\left(R_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)\right)$ in these regions. The flipped color region in the above top-down view figures indicate exactly that. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

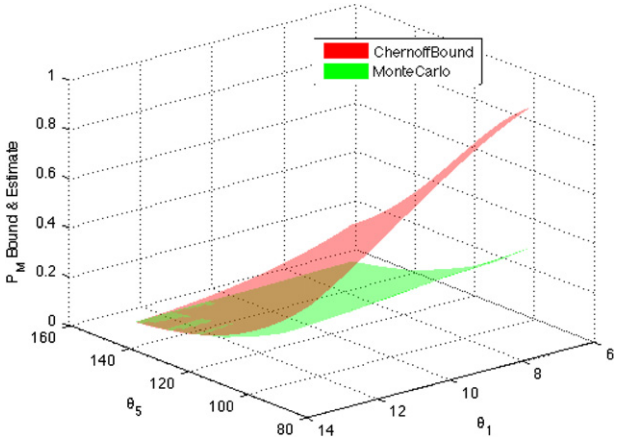


Fig. 4. Comparison of the Chernoff bound and Monte Carlo estimate of $P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1)$. In some regions, the Monte Carlo estimate is slightly larger than the Chernoff bound, because the precision of Monte Carlo in these regions where $P_M^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(1) \ll 1\%$ is not significantly high.

of the underlying Poisson processes, which may differ from the true one due to imperfectly known parameters. At the core of our approach is the derivation of analytically tractable upper bounds on the error probabilities associated with the performance of the mismatched LRT. Under the assumption that the true parameter values are in a neighborhood of the (nominal) values used in computing the corresponding likelihood ratio, it is shown that the bounds that capture the performance of the test vary smoothly, implying a degree of robustness to parameter variations. These results are directly applicable to problems associated with the detection of radioactive material in transit, and are relevant to a number of other applications that involve distinguishing between two (possibly time-inhomogeneous) Poisson processes with parameters that are not accurately known. The framework proposed in this paper provides analytically tractable performance metrics that can inform about the effect of parameter uncertainty on decision-making performance.

Acknowledgments

This work was initiated while the first author was a Visiting Assistant Professor in the Department of Mathematical Sciences at the University of Delaware. The first author would like to thank Petr Plechac for insightful discussions that stimulated the current work. The second and fourth authors were partially supported by ARL MAST CTA # W911NF-08-2-0004.

Appendix

Proof of Theorem 1

The proof of Theorem 1 proceeds through the following steps. We start with Lemma 4 which collects some useful facts from the martingale theory of point processes [1]. Next, Lemma 5 establishes the integral equation (13) which plays a pivotal role in the proofs of the ensuing Lemmas 6 and 7. Finally, we use the last two lemmas to prove Theorem 1.

To exploit various martingales associated with $\boldsymbol{N}_t$, we let $\{\mathscr{F}_t^{\boldsymbol{N}} : t \in [0,T]\}$ be the filtration generated by the process $\boldsymbol{N}_t$. Thus, for $t \in [0,T]$, $\mathscr{F}_t^{\boldsymbol{N}} \triangleq \sigma(\boldsymbol{N}_s : 0 \le s \le t)$ is the smallest σ-algebra on $(\Omega, \mathscr{F})$ with respect to which all the k-dimensional random variables $\boldsymbol{N}_s$, $0 \le s \le t$, are measurable.

Lemma 4. *Let $\boldsymbol{\theta} \in \Theta$, $i \in \{1,\dots,k\}$. Then,*

1. *$M_t^{\boldsymbol{\theta}}(i) \triangleq N_t(i) - \int_0^t \beta_i(s,\boldsymbol{\theta})ds$ is a $(\mathbb{P}_0^{\boldsymbol{\theta}}, \mathscr{F}_t^{\boldsymbol{N}})$-martingale for $t \in [0,T]$. Further, for any $\mathscr{F}_t^{\boldsymbol{N}}$-predictable[2] process X_t satisfying $\mathbb{E}_0^{\boldsymbol{\theta}}\left[\int_0^T |X_s|\beta_i(s,\boldsymbol{\theta})ds\right] < \infty$, the process $\int_0^t X_s dM_s^{\boldsymbol{\theta}}(i)$ is a zero-mean $(\mathbb{P}_0^{\boldsymbol{\theta}}, \mathscr{F}_t^{\boldsymbol{N}})$-martingale for $t \in [0,T]$.*
2. *$\tilde{M}_t^{\boldsymbol{\theta}}(i) \triangleq N_t(i) - \int_0^t [\beta_i(s,\boldsymbol{\theta}) + \nu_i(s,\boldsymbol{\theta})]\,ds$ is a $(\mathbb{P}_1^{\boldsymbol{\theta}}, \mathscr{F}_t^{\boldsymbol{N}})$-martingale for $t \in [0,T]$. Further, for any $\mathscr{F}_t^{\boldsymbol{N}}$-predictable process X_t satisfying $\mathbb{E}_1^{\boldsymbol{\theta}}\left[\int_0^T |X_s|[\beta_i(s,\boldsymbol{\theta}) + \nu_i(s,\boldsymbol{\theta})]ds\right] < \infty$, the process $\int_0^t X_s d\tilde{M}_s^{\boldsymbol{\theta}}(i)$ is a zero-mean $(\mathbb{P}_1^{\boldsymbol{\theta}}, \mathscr{F}_t^{\boldsymbol{N}})$-martingale for $t \in [0,T]$.*

Proof. Direct application of [1, Theorem II.3.T8].

Lemma 5. *For any $p \in \mathbb{R}$, $t \in [0,T]$, $\boldsymbol{\theta} \in \Theta$,*

$$(L_t^{\boldsymbol{\theta}})^p = 1 + \sum_{i=1}^k \int_0^t (L_{s-}^{\boldsymbol{\theta}})^p \left[\mu_i(s,\boldsymbol{\theta})^p - 1\right] dN_s(i) + p\sum_{i=1}^k \int_0^t (L_{s-}^{\boldsymbol{\theta}})^p \left[1 - \mu_i(s,\boldsymbol{\theta})\right]\beta_i(s,\boldsymbol{\theta})ds. \tag{13}$$

Proof. Since the calculations are similar to those of [16, Lemma 11], we simply provide a brief sketch. Fix $p \in \mathbb{R}$, $\boldsymbol{\theta} \in \Theta$. For $t \in [0,T]$, we write $(L_t^{\boldsymbol{\theta}})^p = x(t)y(t)$, where $x(t) \triangleq \exp\left(p\sum_{i=1}^k \int_0^t \log\mu_i(s,\boldsymbol{\theta})dN_s(i)\right)$ and $y(t) \triangleq \exp\left(-p\sum_{i=1}^k \int_0^t \nu_i(s,\boldsymbol{\theta})ds\right)$. Using the Product Formula [1, Theorem A4.T2], we write $x(t)y(t) = x(0)y(0) + \int_0^t x(s-)dy(s) + \int_0^t y(s)dx(s)$, and we now reason as in the proof of [16, Lemma 11] to get (13).

Lemmas 6 and 7, stated next, will be used in the proof of Theorem 1 to establish (8). In the proofs of both lemmas, we start with (13) for $(L_t^{\hat{\boldsymbol{\theta}}})^p$, use Lemma 4 to express the right hand side in terms of martingales with respect to the appropriate probability measure, take expectations, and solve the resulting deterministic integral equation.

Lemma 6. *For $p \in \mathbb{R}$, $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, $t \in [0,T]$, we have*

$$\mathbb{E}_0^{\boldsymbol{\theta}}\left[(L_t^{\hat{\boldsymbol{\theta}}})^p\right] = \exp\left(\sum_{i=1}^k \int_0^t \left\{\left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right]\beta_i(s,\boldsymbol{\theta}) + p\left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})\right\} ds\right).$$

Proof. Fix $p \in \mathbb{R}$, $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$. By (13), we have

$$\begin{aligned}(L_t^{\hat{\boldsymbol{\theta}}})^p &= 1 + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right] dN_s(i) + p\sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})ds \\ &= 1 + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right] dM_s^{\boldsymbol{\theta}}(i) + p\sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})ds \\ &\quad + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right]\beta_i(s,\boldsymbol{\theta})ds.\end{aligned}$$

Hence,

$$\begin{aligned}(L_t^{\hat{\boldsymbol{\theta}}})^p &= 1 + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right] dM_s^{\boldsymbol{\theta}}(i) \\ &\quad + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^p \left\{\left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right]\beta_i(s,\boldsymbol{\theta}) + p\left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})\right\} ds.\end{aligned}$$

Note that $(L_{t-}^{\hat{\boldsymbol{\theta}}})^p$ is $\mathscr{F}_t^{\boldsymbol{N}}$-predictable, being as it is left-continuous and $\mathscr{F}_t^{\boldsymbol{N}}$-adapted. Taking expectations with respect to $\mathbb{P}_0^{\boldsymbol{\theta}}$, and using Lemma 4, we get

$$\begin{aligned}\mathbb{E}_0^{\boldsymbol{\theta}}\left[(L_t^{\hat{\boldsymbol{\theta}}})^p\right] &= 1 + \sum_{i=1}^k \int_0^t \mathbb{E}_0^{\boldsymbol{\theta}}\left[(L_{s-}^{\hat{\boldsymbol{\theta}}})^p\right] \left\{\left[\mu_i(s,\hat{\boldsymbol{\theta}})^p - 1\right] \right. \\ &\quad \left. \times\ \beta_i(s,\boldsymbol{\theta}) + p\left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})\right\} ds.\end{aligned}$$

The equation above can be solved as in [1, Theorem A4.T4] to get the stated claim.

Lemma 7. *For $q \in \mathbb{R}$, $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$, $t \in [0,T]$, we have*

$$\mathbb{E}_1^{\boldsymbol{\theta}}\left[(L_t^{\hat{\boldsymbol{\theta}}})^q\right] = \exp\left(\sum_{i=1}^k \int_0^t \left\{\left[\mu_i(s,\hat{\boldsymbol{\theta}})^q - 1\right]\cdot[\beta_i(s,\boldsymbol{\theta}) + \nu_i(s,\boldsymbol{\theta})] + q\left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})\right\} ds\right).$$

Proof. The proof is very similar to that of Lemma 6. We use (13) to get

$$\begin{aligned}(L_t^{\hat{\boldsymbol{\theta}}})^q &= 1 + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^q \left[\mu_i(s,\hat{\boldsymbol{\theta}})^q - 1\right] d\tilde{M}_s^{\boldsymbol{\theta}}(i) \\ &\quad + \sum_{i=1}^k \int_0^t (L_{s-}^{\hat{\boldsymbol{\theta}}})^q \left\{\left[\mu_i(s,\hat{\boldsymbol{\theta}})^q - 1\right]\right. \\ &\quad \left. \cdot\ [\beta_i(s,\boldsymbol{\theta}) + \nu_i(s,\boldsymbol{\theta})] + q\left[1 - \mu_i(s,\hat{\boldsymbol{\theta}})\right]\beta_i(s,\hat{\boldsymbol{\theta}})\right\} ds\end{aligned}$$

for $q \in \mathbb{R}$, $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in \Theta$. Now, using Lemma 4, we take expectations with respect to $\mathbb{P}_1^{\boldsymbol{\theta}}$, and solve the resulting deterministic integral equation to get the stated claim.

Proof of Theorem 1. Write $\eta = \log\gamma$. By the Markov inequality, we have for $p > 0$, $q < 0$,

$$\begin{aligned}P_F^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(\gamma) &= \mathbb{P}_0^{\boldsymbol{\theta}}\left((L_T^{\hat{\boldsymbol{\theta}}})^p \ge \exp(p\eta)\right) \le \exp(-p\eta)\mathbb{E}_0^{\boldsymbol{\theta}}\left[(L_T^{\hat{\boldsymbol{\theta}}})^p\right] \\ &= \exp\left(\Lambda_0^{(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}(p) - p\eta\right),\end{aligned}$$

[2] See [1, Section I.3]. For our purposes, it will be enough to note that if a process X_t is $\mathscr{F}_t^{\boldsymbol{N}}$-adapted and left-continuous, then X_t is $\mathscr{F}_t^{\boldsymbol{N}}$-predictable.

$$P_M^{(\theta,\hat\theta)}(\gamma) = \mathbb{P}_1^{\theta}\left((L_T^{\hat\theta})^q > \exp(q\eta)\right) \le \exp(-q\eta)\mathbb{E}_1^{\theta}\left[(L_T^{\hat\theta})^q\right] = \exp\left(\Lambda_1^{(\theta,\hat\theta)}(q) - q\eta\right).$$

Taking infima over $p > 0$, $q < 0$, and noting that $x \mapsto \exp(x)$ is strictly increasing, we get (7). An application of Lemmas 6 and 7 at $t = T$, followed by taking logarithms, yields (8).

Proof of Propositions 2 and 3

The next two lemmas record properties of $\Lambda_0^{(\theta,\hat\theta)}$ and $\Lambda_1^{(\theta,\hat\theta)}$ which are used in the proof of Proposition 2.

Lemma 8. *For $\theta, \hat\theta \in \Theta$, we have that $p \mapsto \Lambda_0^{(\theta,\hat\theta)}(p)$ is C^2 with $\Lambda_0^{(\theta,\hat\theta)}(0) = 0$ and*

$$\frac{\partial \Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p) = \sum_{i=1}^{k}\int_0^T \left\{\mu_i(s,\hat\theta)^p\left(\log\mu_i(s,\hat\theta)\right)\beta_i(s,\theta) + \left[1-\mu_i(s,\hat\theta)\right]\beta_i(s,\hat\theta)\right\} ds, \tag{14}$$

$$\frac{\partial^2 \Lambda_0^{(\theta,\hat\theta)}}{\partial p^2}(p) = \sum_{i=1}^{k}\int_0^T \mu_i(s,\hat\theta)^p\left(\log\mu_i(s,\hat\theta)\right)^2\beta_i(s,\theta)ds.$$

Further, $\partial^2\Lambda_0^{(\theta,\hat\theta)}/\partial p^2 > 0$, implying that $p \mapsto \Lambda_0^{(\theta,\hat\theta)}(p)$ is strictly convex. Finally, for any $p \in \mathbb{R}$, $p \neq 0$, we have

$$\Lambda_0^{(\theta,\hat\theta)}(p) - p\frac{\partial \Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p) < 0. \tag{15}$$

Proof. Since the integrand in the expression for $\Lambda_0^{(\theta,\hat\theta)}(p)$ given by (8) is smooth in p for each fixed $s \in [0,T]$, it is easily shown using Assumptions 1 and 2 that one can take arbitrarily many derivatives of $\Lambda_0^{(\theta,\hat\theta)}(p)$ with respect to p by simply differentiating under the integral sign; this yields (14). It is also easily seen that $\partial^2\Lambda_0^{(\theta,\hat\theta)}/\partial p^2 > 0$. By the ensuing strict convexity of $p \mapsto \Lambda_0^{(\theta,\hat\theta)}(p)$, we have that for any $p, \tilde p$ in $\mathbb{R}$ with $p \neq \tilde p$,

$$\Lambda_0^{(\theta,\hat\theta)}(\tilde p) > \Lambda_0^{(\theta,\hat\theta)}(p) + (\tilde p - p)\frac{\partial \Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p).$$

Setting $\tilde p = 0$, we get (15).

Lemma 9. *For $\theta, \hat\theta \in \Theta$, we have that $\Lambda_1^{(\theta,\hat\theta)}(q)$ is C^2 with $\Lambda_1^{(\theta,\hat\theta)}(0) = 0$ and*

$$\frac{\partial \Lambda_1^{(\theta,\hat\theta)}}{\partial q}(q) = \sum_{i=1}^{k}\int_0^T \left\{\mu_i(s,\hat\theta)^q\left(\log\mu_i(s,\hat\theta)\right)\cdot\left[\beta_i(s,\theta)+\nu_i(s,\theta)\right] + \left[1-\mu_i(s,\hat\theta)\right]\beta_i(s,\hat\theta)\right\} ds, \tag{16}$$

$$\frac{\partial^2 \Lambda_1^{(\theta,\hat\theta)}}{\partial q^2}(q) = \sum_{i=1}^{k}\int_0^T \mu_i(s,\hat\theta)^q\left(\log\mu_i(s,\hat\theta)\right)^2\cdot\left[\beta_i(s,\theta)+\nu_i(s,\theta)\right]ds.$$

Further, $\partial^2\Lambda_1^{(\theta,\hat\theta)}/\partial q^2 > 0$, implying that $q \mapsto \Lambda_1^{(\theta,\hat\theta)}(q)$ is strictly convex. Finally, for any $q \in \mathbb{R}$, $q \neq 0$, we have

$$\Lambda_1^{(\theta,\hat\theta)}(q) - q\frac{\partial \Lambda_1^{(\theta,\hat\theta)}}{\partial q}(q) < 0. \tag{17}$$

Proof. The proof is very similar to that of Lemma 8.

Proof of Proposition 2. We make use here of Lemmas 8 and 9. Note that by (14), (16), we have $\partial\Lambda_0^{(\theta,\hat\theta)}/\partial p(0) < \partial\Lambda_1^{(\theta,\hat\theta)}/\partial q(0)$. Fix $\log\gamma \in \left(\frac{\partial\Lambda_0^{(\theta,\hat\theta)}}{\partial p}(0), \frac{\partial\Lambda_1^{(\theta,\hat\theta)}}{\partial q}(0)\right)$. Since $p \mapsto \partial\Lambda_0^{(\theta,\hat\theta)}/\partial p$ and $q \mapsto \partial\Lambda_1^{(\theta,\hat\theta)}/\partial q$ are strictly increasing and continuous, there exist unique $p^* = p^*(\theta,\hat\theta) > 0$ and $q^* = q^*(\theta,\hat\theta) < 0$ such that (10) holds. By (strict) convexity of $p \mapsto \Lambda_0^{(\theta,\hat\theta)}(p) - p\log\gamma$, and $q \mapsto \Lambda_1^{(\theta,\hat\theta)}(q) - q\log\gamma$, it follows that the infima in (9) are in fact attained at p^* and q^*. Now using (15), (17), we get (11). Then, application of Theorem 1 yields (12).

Proof of Proposition 3. We will make repeated use of the fact that, on account of Assumptions 1 and 2, and Eqs. (8), (14), (16), the functions $\Lambda_0^{(\theta,\hat\theta)}(p)$, $\Lambda_1^{(\theta,\hat\theta)}(q)\frac{\partial\Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p)$, $\frac{\partial\Lambda_1^{(\theta,\hat\theta)}}{\partial q}(q)$ are C^1 in θ. Since Θ is an open subset of $\mathbb{R}^d$, there now exists $\delta_0 > 0$ small enough that $B(\hat\theta,\delta_0) \subset \Theta$ and

$$l \triangleq \sup_{\theta\in B(\hat\theta,\delta_0)}\frac{\partial\Lambda_0^{(\theta,\hat\theta)}}{\partial p}(0) < \inf_{\theta\in B(\hat\theta,\delta_0)}\frac{\partial\Lambda_1^{(\theta,\hat\theta)}}{\partial q}(0) \triangleq r.$$

Let $\log\gamma \in (l,r)$. Then, for $\theta \in B(\hat\theta,\delta_0)$, Proposition 2 applies with $R_F^{(\theta,\hat\theta)}(\gamma)$ and $R_M^{(\theta,\hat\theta)}(\gamma)$ given by (11), where the corresponding $p^* = p^*(\theta,\hat\theta) > 0$ and $q^* = q^*(\theta,\hat\theta) < 0$ satisfy (10).

We next show that $\theta \mapsto p^*(\theta,\hat\theta)$ and $\theta \mapsto q^*(\theta,\hat\theta)$ are continuously differentiable in a neighborhood of $\hat\theta$. Let $F(\theta,p) \triangleq \frac{\partial\Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p) - \log\gamma$. Note that by (10), $p^*(\theta,\hat\theta)$ is defined implicitly through $F(\theta,p^*(\theta,\hat\theta)) = 0$. Clearly, $F(\hat\theta,p(\hat\theta,\hat\theta)) = 0$. It now follows from the implicit function theorem that there exist $\delta_1 > 0$, $\varepsilon_1 > 0$ and a unique continuously differentiable function $f : B(\hat\theta,\delta_1) \to (p^*(\hat\theta,\hat\theta)-\varepsilon_1, p^*(\hat\theta,\hat\theta)+\varepsilon_1)$ such that $F(\theta,f(\theta)) = 0$. Of course, $p^*(\theta,\hat\theta) = f(\theta)$. Similarly, letting $G(\theta,q) \triangleq \frac{\partial\Lambda_1^{(\theta,\hat\theta)}}{\partial q}(q) - \log\gamma$, one can show that there exist $\delta_2 > 0$, $\varepsilon_2 > 0$ and a unique continuously differentiable function $g : B(\hat\theta,\delta_2) \to (q^*(\hat\theta,\hat\theta)-\varepsilon_2, q^*(\hat\theta,\hat\theta)+\varepsilon_2)$ such that $G(\theta,g(\theta)) = 0$. Letting $\delta \triangleq \min\{\delta_0,\delta_1,\delta_2\}$, we see that $\theta \mapsto p^*(\theta,\hat\theta)$ and $\theta \mapsto q^*(\theta,\hat\theta)$ are continuously differentiable on $B(\hat\theta,\delta)$.

Let us now show that the maps $\theta \mapsto R_F^{(\theta,\hat\theta)}(\gamma)$ and $\theta \mapsto R_M^{(\theta,\hat\theta)}(\gamma)$ are continuously differentiable on the open ball $B(\hat\theta,\delta)$. Let $\mathscr{E}_F(\theta,p) \triangleq \Lambda_0^{(\theta,\hat\theta)}(p) - p\frac{\partial\Lambda_0^{(\theta,\hat\theta)}}{\partial p}(p)$ and let $\mathscr{E}_M(\theta,q) \triangleq \Lambda_1^{(\theta,\hat\theta)}(q) - q\frac{\partial\Lambda_1^{(\theta,\hat\theta)}}{\partial q}(q)$. Note that $\mathscr{E}_F$ and $\mathscr{E}_M$ are continuously differentiable. If we let $\varphi(\theta) \triangleq (\theta, p^*(\theta,\hat\theta))$ and $\psi(\theta) \triangleq (\theta, q^*(\theta,\hat\theta))$, then $R_F^{(\theta,\hat\theta)}(\gamma) = \mathscr{E}_F(\varphi(\theta))$ and $R_M^{(\theta,\hat\theta)}(\gamma) = \mathscr{E}_M(\psi(\theta))$. The result now follows by noting that the composition of continuously differentiable functions is also continuously differentiable.

References

[1] P. Brémaud, Point Processes and Queues. Martingale Dynamics, Springer-Verlag, 1981.
[2] D.L. Snyder, Random Point Processes, Wiley-Interscience, 1975.
[3] R.M. Gagliardi, S. Karp, Optical Communications, second ed., Wiley-Interscience, 1995.
[4] H.C. Tuckwell, Stochastic Processes in the Neurosciences, SIAM, 1989.
[5] R.J. Nemzek, J.S. Dreicer, D.C. Torney, T.T. Warnock, Distributed sensor networks for detection of mobile radioactive sources, IEEE Trans. Nucl. Sci. 51 (4) (2004) 1693–1700.
[6] R. Boel, P. Varaiya, E. Wong, Martingales on jump processes II: Applications, SIAM J. Control 13 (5) (1975) 1022–1061.
[7] E.A. Geraniotis, H.V. Poor, Minimax discrimination for observed Poisson processes with uncertain rate functions, IEEE Trans. Inform. Theory 31 (5) (1985) 660–669.

[8] S. Verdú, Asymptotic error probability of binary hypothesis testing for Poisson point-process observations, IEEE Trans. Inform. Theory 32 (1) (1986) 113–115.
[9] S. Verdú, Multiple-access channels with point-process observations: Optimum demodulation, IEEE Trans. Inform. Theory 32 (5) (1986) 642–651.
[10] S.A. Kassam, H.V. Poor, Robust techniques for signal processing: a survey, Proc. IEEE 73 (1985) 433–481.
[11] H.V. Poor, An Introduction to Signal Detection and Estimation, second ed., Springer-Verlag, 1994.
[12] R. Byrd, J. Moss, W. Priedhorsky, C. Pura, G.W. Richter, K. Saeger, W. Scarlett, S. Scott, R. Wagner Jr., Nuclear detection to prevent or defeat clandestine nuclear attack, IEEE Sens. J. 5 (4) (2005) 593–609.
[13] R.C. Runkle, L.E. Smith, A.J. Peurrung, The photon haystack and emerging radiation detection technology, J. Appl. Phys. 106 (041101) (2009) 1–21.
[14] L. Wein, M. Atkinson, The last line of defense: designing radiation detection-interdiction systems to protect cities form a nuclear terrorist attack, IEEE Trans. Nucl. Sci. 54 (3) (2007) 654–669.
[15] C.D. Pahlajani, I. Poulakakis, H.G. Tanner, Networked decision making for Poisson processes with applications to nuclear detection, IEEE Trans. Automat. Control 59 (1) (2014) 193–198.
[16] C.D. Pahlajani, J. Sun, I. Poulakakis, H.G. Tanner, Error probability bounds for nuclear detection: Improving accuracy through controlled mobility, Automatica 50 (10) (2014) 2470–2481.
[17] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Statist. 23 (1952) 493–507.
[18] J.L. Hibey, D.L. Snyder, J.H. van Schuppen, Error-probability bounds for continuous-time decision problems, IEEE Trans. Inform. Theory 24 (5) (1978) 608–622.
[19] C.M. Newman, B.W. Stuck, Chernoff bounds for discriminating between two Markov processes, Stochastics 2 (1979) 139–153.
[20] D. Kazakos, Signal detection under mismatch, IEEE Trans. Inform. Theory 28 (4) (1982).
[21] Y. Lee, Y. Sung, Generalized Chernoff information for mismatched Bayesian detection and its application to energy detection, IEEE Signal Process. Lett. 19 (11) (2012) 753–756.