HHS Public Access

Author manuscript

*IEEE Trans Med Imaging*. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

*IEEE Trans Med Imaging*. 2016 September ; 35(9): 2164–2173. doi:10.1109/TMI.2016.2553001.

# Optimal joint detection and estimation that maximizes ROC-type curves

**Adam Wunderlich**,

Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD 20993 USA.

**Bart Goossens**, and

Dept. of Telecommunications and Information Processing, TELIN-IPI-iMinds, Ghent University, Ghent, Belgium.

**Craig K. Abbey**

Dept. of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106 USA.

## Abstract

Combined detection-estimation tasks are frequently encountered in medical imaging. Optimal methods for joint detection and estimation are of interest because they provide upper bounds on observer performance, and can potentially be utilized for imaging system optimization, evaluation of observer efficiency, and development of image formation algorithms. We present a unified Bayesian framework for decision rules that maximize receiver operating characteristic (ROC)-type summary curves, including ROC, localization ROC (LROC), estimation ROC (EROC), free-response ROC (FROC), alternative free-response ROC (AFROC), and exponentially-transformed FROC (EFROC) curves, succinctly summarizing previous results. The approach relies on an interpretation of ROC-type summary curves as plots of an expected utility versus an expected disutility (or penalty) for signal-present decisions. We propose a general utility structure that is flexible enough to encompass many ROC variants and yet sufficiently constrained to allow derivation of a linear expected utility equation that is similar to that for simple binary detection. We illustrate our theory with an example comparing decision strategies for joint detection-estimation of a known signal with unknown amplitude. In addition, building on insights from our utility framework, we propose new ROC-type summary curves and associated optimal decision rules for joint detection-estimation tasks with an unknown, potentially-multiple, number of signals in each observation.

### Index Terms

Receiver operating characteristic; signal detection theory; utility theory; ideal observer



Adam Wunderlich is now with the Communications Technology Laboratory, National Institute of Standards and Technology, Boulder, CO 80305 USA

## I. Introduction

Joint detection and estimation tasks arise often in medical imaging. For example, a common task is to detect the presence of focal lesions and estimate their respective locations [1], [2]. Other types of combined detection-estimation tasks are found in cancer imaging, where lesion detection may be followed by estimation of tumor size or functional biomarkers such as standardized uptake value [3], [4]. Also, measures of tumor heterogeneity and texture can be estimated [5], [6]. Further examples of joint detection and estimation tasks arise in cardiac imaging (stenosis detection and quantification) [7], [8], neuroimaging (detection and severity estimation of traumatic brain injury) [9], [10], and osteoporosis assessment (fracture detection and bone density estimation) [11], [12].

Observer performance on a binary detection task is commonly summarized with the receiver operating characteristic (ROC) curve [13]. Because ROC curves only summarize performance for detection and not estimation, several modifications of the ROC curve have been proposed to measure performance on joint detection and estimation tasks; most of this work has been motivated by lesion detection-localization tasks in medical imaging. Namely, for joint detection and localization of a single signal (or target), the localization ROC (LROC) curve [14] can be used. When there are an unknown number of signals, detection and localization performance can be summarized with the free-response ROC (FROC) curve [15], two versions of alternative free-response ROC (AFROC) curves [16], [17], or the exponentially-transformed FROC (EFROC) curve [18]. In addition, the estimation ROC (EROC) curve [19] has been proposed to summarize performance for general joint detection and estimation tasks.

Optimal methods for simultaneous detection and estimation are of theoretical interest because they give upper bounds on observer performance. Such optimal strategies, commonly called “ideal observers” in the medical imaging literature [20], can potentially be utilized for imaging system optimization [21], [22], evaluation of observer efficiency [23]–[25], and development of image formation algorithms [23], [24]. For the above ROC-type curves, a higher summary curve implies better performance. Hence, the optimal decision rule can be defined as the strategy that maximizes the height of the summary curve for each value of the abscissa (horizontal axis).[1] For a simple binary detection task, the optimal ROC decision rule is the likelihood ratio test, which consists of comparing the likelihood ratio statistic to a threshold [26], [27]. For joint detection and estimation tasks, optimal decision rules have been derived by Khurd et al. that maximize the LROC curve [28] and the FROC and AFROC curves [29], respectively. In addition, Clarkson [19] presented the decision rule that maximizes the EROC curve.

In this paper, we present a unified, decision-theoretic derivation of optimal decision rules associated with ROC-type curves for joint binary detection and estimation tasks. Namely, building on findings in [30], a common utility framework is presented that captures key properties associated with ROC-type summary curves. In particular, because it explicitly describes how decision outcomes are rewarded or penalized, this utility framework provides

[1]Throughout this work, for simplicity, we assume that each summary curve is defined for every value of its abscissa.

a means to adapt the performance criterion to the problem at hand [30]. The decision rule that maximizes expected utility is derived and shown to reduce to the optimal ROC, LROC, EROC, FROC, AFROC, and EFROC decision rules as special cases, putting the results of [19], [28], [29] in a common Bayesian framework. In addition, based on insights from this utility framework, we propose new performance summary curves and optimal decision rules for joint detection and estimation tasks with an unknown, potentially-multiple, number of signals.

In [19], [28], [29], optimal LROC, EROC, FROC, and AFROC decision rules are derived by generalizing a proof for the Neyman-Pearson lemma that is commonly found in the engineering literature, e.g., [26], [27], which directly solves a constrained maximization problem. In addition, [28], [29] give secondary decision-theoretic derivations of the optimal LROC, FROC, and AFROC decision rules, respectively. However, these secondary derivations are not sufficient to show that the summary curves are maximized. By contrast, our approach relies on the structure of the expected utility equation to show that maximizing expected utility maximizes the corresponding summary curve. Moreover, the utility structures used here for FROC and AFROC curves are more general than those in [29], and are seen to be consistent with the requirement that utilities lie on an interval scale [31], a property of classical theories of utility [32], [33].

A key concept underlying our approach is the observation that the ordinate (vertical axis) and abscissa (horizontal axis) of an ROC-type summary curve can be interpreted as an expected utility and an expected disutility (or penalty), respectively, for signal-present decisions. Namely, for the ROC curve, the true-positive rate can be seen as an expected utility and the false-positive rate can be seen as an expected disutility. This interpretation allows us to put other ROC-types curves, such as LROC, EROC, FROC, and AFROC into a common utility framework. Furthermore, it leads to a generalization of the classical likelihood ratio test for binary detection to joint detection-estimation problems.

## II. Preliminaries: the ROC curve and its progeny

This section provides necessary background and notation that will be used in later sections. We start with basic aspects of binary detection and the ROC curve. Subsequently, we review the definitions of LROC, EROC, FROC, AFROC, and EFROC curves. Throughout the paper, we denote vectors with boldface letters and scalars with non-boldface letters.

### A. Binary detection

**1) The ROC curve—**A binary signal detection task consists of deciding whether or not a signal is present. Let $\mathcal{H} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ be a binary random variable denoting signal presence, where $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond to signal-absent and signal-present cases, respectively. Given an observation, $\mathbf{g}$, a decision rule for this task produces an estimate $\hat{\mathcal{H}}(\mathbf{g})$ of $\mathcal{H}$. In this setting, there are four possible outcomes: true-positive (TP), false-negative (FN), true-negative (TN), and false-positive (FP), which lead to the definition of four rates (or conditional probabilities): $R_{\mathrm{TP}} = P(\hat{\mathcal{H}} = \mathcal{H}_1 | \mathcal{H} = \mathcal{H}_1)$, $R_{\mathrm{TN}} = P(\hat{\mathcal{H}} = \mathcal{H}_0 | \mathcal{H} = \mathcal{H}_0)$, $R_{\mathrm{FN}} = P(\hat{\mathcal{H}} = \mathcal{H}_0 | \mathcal{H} = \mathcal{H}_1)$, and $R_{\mathrm{FP}} = P(\hat{\mathcal{H}} = \mathcal{H}_1 | \mathcal{H} = \mathcal{H}_0)$, where $P(E)$ denotes the probability of event $E$. In the engineering literature, e.g., [26], [27], the true-positive rate, $R_{\mathrm{TP}}$, false-

negative rate, $R_{FN}$, and false-positive rate, $R_{FP}$, are commonly called the "detection" (or "hit"), "miss" and "false-alarm" probabilities, respectively. From the above definitions, it follows that $R_{TP} + R_{FN} = 1$ and $R_{TN} + R_{FP} = 1$. Consequently, for a given decision threshold, binary detection performance is fully described by the pair ($R_{TP}$, $R_{FP}$). The ROC curve is defined as the plot of $R_{TP}$ versus $R_{FP}$ over all decision thresholds [26]; see Fig. 1.

**2) Expected utility—**Throughout the paper, to be consistent with [19] and [30], we work with expected utility instead of Bayes risk. Nonetheless, all results can alternatively be derived with Bayes risk by defining costs as negative utilities [34]. Let $U_E$ denote the utility of event $E$. Also, to simplify our notation, we will express $P(\mathcal{H} = \mathcal{H}_i)$ as $P(\mathcal{H}_i)$. The total expected utility for a binary detection task is the sum of each event utility weighted by its corresponding probability [26],

$$\mathscr{U} = U_{TP} R_{TP} P(\mathscr{H}_1) + U_{FN} R_{FN} P(\mathscr{H}_1) + U_{TN} R_{TN} P(\mathscr{H}_0) + U_{FP} R_{FP} P(\mathscr{H}_0). \quad (1)$$

Above, note that a script letter, $\mathscr{U}$, denotes expected utility. Throughout the paper, expected utilities will be denoted with script letters to distinguish them from utilities of events. Using the identities $R_{TP} + R_{FN} = 1$ and $R_{TN} + R_{FP} = 1$, the expected utility can be rewritten in terms of $R_{TP}$ and $R_{FP}$ as

$$\mathscr{U} = (U_{FP} - U_{TN}) P(\mathscr{H}_0) R_{FP} + (U_{TP} - U_{FN}) P(\mathscr{H}_1) R_{TP} + U_{TN} P(\mathscr{H}_0) + U_{FN} P(\mathscr{H}_1). \quad (2)$$

Rearranging terms, we arive at the iso-utility equation [35],

$$R_{TP} = \frac{(U_{TN} - U_{FP}) P(\mathscr{H}_0)}{(U_{TP} - U_{FN}) P(\mathscr{H}_1)} R_{FP} + \frac{\mathscr{U} - U_{TN} P(\mathscr{H}_0) - U_{FN} P(\mathscr{H}_1)}{(U_{TP} - U_{FN}) P(\mathscr{H}_1)}, \quad (3)$$

which defines iso-utility lines in ROC coordinates.

We make two observations from the above equations that will be important for later developments. First, under the reasonable assumption that correct decisions have greater utility than incorrect decisions, i.e., $U_{TP} > U_{FN}$ and $U_{TN} > U_{FP}$, (2) shows that expected utility has the form $\mathscr{U} = AR_{FP} + BR_{TP} + C$ with $A < 0$ and $B > 0$. So a higher ROC curve implies higher expected utility, and vice-versa. Therefore, maximizing expected utility is equivalent to maximizing $R_{TP}$ for each $R_{FP}$. This observation clarifies why both the Bayes criterion and Neyman-Pearson criterion lead to the likelihood ratio test for simple binary detection [26].

Second, the iso-utility equation (3) is invariant under positive affine transformations of utility. This outcome is consistent with the fact that classical theories of utility [32], [33] have the property that utility lies on an interval scale [31]. Specifically, for any constants $\alpha > 0$ and $\beta$, if the total expected utility is transformed as $\mathscr{U}' = \alpha\mathscr{U} + \beta$, and the event utilities

are transformed as $U'_E = \alpha U_E + \beta$ for each event $E$, e.g., $U'_{\mathrm{TP}} = \alpha U_{\mathrm{TP}} + \beta$, etc., then (3) is unchanged.

### B. EROC and LROC curves

Our focus in this work is a joint detection-estimation task involving a binary determination of signal-presence and an estimation of a random signal-parameter vector, $\boldsymbol{\theta} \in \Theta$. Namely, given an observation, $\mathbf{g}$, the aim is to obtain an estimate $\hat{\mathscr{H}}(\mathbf{g})$ for $\mathscr{H}$, and if $\hat{\mathscr{H}} = \mathscr{H}_1$, to estimate $\boldsymbol{\theta}$ as $\hat{\boldsymbol{\theta}}(\mathbf{g})$. Note that the parameter vector $\boldsymbol{\theta}$ is assumed to be relevant only when the signal is present, i.e., we do not consider estimation of parameters associated with signal absence. For the LROC curve, $\boldsymbol{\theta}$, is the coordinate vector for the signal location, whereas for the EROC curve, $\boldsymbol{\theta}$ is a general parameter vector. Since the LROC curve is a special case of the EROC curve, we start with the EROC definition.

Let $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ be a utility function for the parameter estimate in the case of a TP decision. It is assumed that the utility function is higher for an accurate parameter estimate and lower for an inaccurate estimate. Also, let $E[\cdot]$ denote mathematical expectation, and let $\mathscr{U}(S)$ be the indicator function that equals one when the proposition $S$ is true and equals zero otherwise. The EROC curve [19] plots the expected utility of the parameter estimate for a TP decision (called the expected utility of a TP for short), defined as

$$\mathscr{U}_{\mathrm{TP}} = E\left[\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})\mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_1) \middle| \mathscr{H} = \mathscr{H}_1\right], \quad (4)$$

versus the false positive rate, $R_{\mathrm{FP}}$, over all decision thresholds. (Note that this expected utility is denoted with a script letter to distinguish it from the utility for a TP event.) Above, the expectation is over $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\mathscr{H}}$, conditional on the event $\mathscr{H} = \mathscr{H}_1$. To make the EROC curve easier to interpret, and without loss of generality, the utility function can be scaled to fall in the interval [0, 1] so that the EROC ordinate is between zero and one [36]. An example of an EROC curve is given in Fig. 1.

The ROC curve can be viewed as a special case of the EROC curve when the utility function is identically equal to one, i.e, $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \equiv 1$, since $E[\mathscr{U}(\hat{\mathscr{H}} = \mathscr{H}_1) | \mathscr{H} = \mathscr{H}_1] = P(\hat{\mathscr{H}} = \mathscr{H}_1 | \mathscr{H} = \mathscr{H}_1) = R_{\mathrm{TP}}$. Another special case of the EROC curve is the LROC curve [14], which applies when there is at most one signal per observation and $\boldsymbol{\theta}$ is the signal location. The LROC curve is obtained when the utility function is the indicator function for correct localization, i.e., $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \mathscr{U}(\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\| \leq R)$, where $R$ is an acceptable radius for correct localization. The LROC curve plots the correctly-localized true-positive rate,

$$R_{\mathrm{TPL}} = P\left(\left[\hat{\mathscr{H}} = \mathscr{H}_1\right] \;\&\; \left[\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\| \leq R\right] \middle| \mathscr{H} = \mathscr{H}_1\right), \quad (5)$$

versus the false-positive rate, $R_{\mathrm{FP}}$. In words, $R_{\mathrm{TPL}}$ is the probability of a TP decision for which the signal is correctly localized; see [37] for additional details on LROC curves.

### C. FROC, AFROC, and EFROC curves

The FROC curve, and its close relatives, AFROC and EFROC curves, are designed to assess performance for a detection-location task with an unknown, random number of signals. Namely, suppose that each observation contains $n \geq 0$ signals, where $n$ is random, and that the task is to decide if the observation contains signals ($\mathcal{H} = \mathcal{H}_1$), and if so, to detect and localize all signals. For this task, the signal-parameter vector is $\boldsymbol{\theta} = [n, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n]$, where $\mathbf{r}_i$ is the coordinate vector of the $i$th signal. The number of (correctly localized) TP detections for an observation is modeled as the sum of $n$ Bernoulli random variables, each with success probability, $R_{\mathrm{TP}}^{\mathrm{F}}$, which is called the signal detection fraction (the superscript "F" denotes FROC).[2] In general, the Bernoulli random variables can be statistically dependent, so that their sum is not necessarily Binomially distributed. The FROC curve plots the signal detection fraction versus the mean number of false-positive detections per observation; see Fig. 1.[3]

Before covering AFROC and EFROC curves, we establish some notation and take a closer look at the definition of the FROC curve. First, we introduce counting functions, $M_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ and $M_{\mathrm{FP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, for the number of TP and FP detections in a signal-present observation, respectively [29]. At this point, to keep our treatment general, we will not assume specific forms for the counting functions; one choice of definitions is presented in Section VI. In addition, let the estimated number of signals in an observation be $\hat{n}$, and denote the number of FP detections in an observation as $N_{\mathrm{FP}}$, which is equal to $\hat{n}$ for a signal-absent observation and $M_{\mathrm{FP}}$ for a signal-present observation.

The FROC curve coordinates can be expressed in terms of the TP and FP counting functions. Specifically, the abscissa, the mean number of FPs per observation, is $E[N_{\mathrm{FP}}] = E[\hat{n}|\mathcal{H} = \mathcal{H}_0]P(\mathcal{H}_0) + E[M_{\mathrm{FP}}|\mathcal{H} = \mathcal{H}_1]P(\mathcal{H}_1)$. In addition, the ordinate, the signal detection fraction, can be written as either $R_{\mathrm{TP}}^{\mathrm{F}} = E[M_{\mathrm{TP}}]/E[n]$ or $R_{\mathrm{TP}}^{\mathrm{F}} = E[M_{\mathrm{TP}}/n|\mathcal{H} = \mathcal{H}_1]$. These alternative expressions for the ordinate were first presented in [29]. They can be derived by using the law of iterated expectations and the fact that $M_{\mathrm{TP}}$ is the sum of $n$ (generally dependent) Bernoulli trials, each with probability of success $R_{\mathrm{TP}}^{\mathrm{F}}$; the details are omitted for brevity. In our previous work on utility theory [30], we used the first expression. However, in this paper, we will use a utility structure based on the second expression for consistency with [29].

Because the horizontal range of the FROC curve is potentially infinite, the area under the curve is generally undefined and not suitable as a summary figure of merit. For this reason, three variants of the FROC curve have been proposed that result in a finite area under the curve. Each of these variants uses the FROC ordinate (vertical axis) and a modification of the FROC abscissa (horizontal axis). The first variant is the AFROC curve [16], which

[2]Note that $R_{\mathrm{TP}}^{\mathrm{F}}$ is assumed to be the same for all signals. Although this is a fairly standard assumption, it may be violated in some circumstances due to, e.g., limited search area, signal cuing, or satisfaction of search [38].

[3]In the FROC literature, starting with [15], the FROC ordinate (vertical axis) is typically defined as an empirical quantity, the total number of TP detections divided by the total number of signals in the sample. A more appropriate definition is the population mean of this empirical quantity, which is easily seen to be $R_{\mathrm{TP}}^{\mathrm{F}}$ [30].

defines the abscissa to be the probability of at least one false-positive per observation, $P(N_{FP} > 0) = P(\hat{n} > 0 \mid \mathcal{H} = \mathcal{H}_0)P(\mathcal{H}_0) + P(M_{FP} > 0 \mid \mathcal{H} = \mathcal{H}_1)P(\mathcal{H}_1)$. The second variant [17], which following [29], we call the signal-absent-abscissa AFROC (SAA-AFROC) curve, defines the abscissa as the probability that there is at least one false-positive in a signal-absent obervation, $P(N_{FP} > 0 \mid \mathcal{H} = \mathcal{H}_0)$. The third variant, called the EFROC curve [18], defines the abscissa as $1-\exp(-E[N_{FP}])$. Since the EFROC abscissa is a strictly increasing transformation of the FROC abscissa, it follows that the decision rule that maximizes the FROC curve for each value of the abscissa also maximizes the EFROC curve. For this reason, our development below for FROC curves also applies to EFROC curves, and we will not discuss the EFROC case separately.

## III. The general Bayes decision rule

Before considering optimal strategies that maximize ROCtype summary curves, it is useful to start with the general Bayes decision rule [34] for a combined binary detection and estimation task, which maximizes expected utility (or equivalently, minimizes Bayes risk). We choose to work with expected utility instead of Bayes risk to be consistent with [19] and [30], which are formulated in terms of utility functions. Equivalent formulations based on Bayes risk can be found in [20, p. 907–908] and [29, p. 381–382].

Using the notation introduced in Section II, let $\mathcal{H} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ denote signal presence and let $\boldsymbol{\theta} \in \Theta$ be a parameter vector associated with the signal(s). Given an observation, $\mathbf{g}$, we seek estimates $\hat{\mathcal{H}}(\mathbf{g})$ and $\hat{\boldsymbol{\theta}}(\mathbf{g})$ for $\mathcal{H}$ and $\boldsymbol{\theta}$, respectively.

Below, we write the joint probability density function (pdf) for a vector of random variables $\mathbf{x}$ as $p(\mathbf{x})$. Generally, random variables are allowed to be either continuous or discrete. When they are discrete, probability density functions should be interpreted as probability mass functions and integrals over probability densities become summations; these modifications are clear from the context. To streamline our notation, we write $p(\mathbf{g} \mid \mathcal{H} = \mathcal{H}_i)$ as $p(\mathbf{g} \mid \mathcal{H}_i)$. In addition, since the signal-parameter vector is only relevant when the signal is present, $p(\boldsymbol{\theta} \mid \mathcal{H} = \mathcal{H}_1)$ will be written simply as $p(\boldsymbol{\theta})$.

Expressing the utility function for an estimate $(\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}})$ as $U(\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}}, \mathcal{H}, \boldsymbol{\theta})$, the posterior expected utility is

$$\mathcal{U}_p(\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}} | \mathbf{g}) = E_{(\mathcal{H}, \boldsymbol{\theta}) | \mathbf{g}} \left[ U(\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}}, \mathcal{H}, \boldsymbol{\theta}) | \mathbf{g} \right], \quad (6)$$

where the expectation is over $(\mathcal{H}, \boldsymbol{\theta})$ given $\mathbf{g}$. Above, we use $\mathcal{U}_p$ to denote posterior expected utility, in order to distinguish it from expected utility, $\mathcal{U}$. The Bayes decision rule, denoted $(\mathcal{H}^*, \boldsymbol{\theta}^*)$, maximizes $\mathcal{U}_p$ for each $\mathbf{g}$, i.e.,

$$(\hat{\mathcal{H}}^*, \boldsymbol{\theta}^*) = \arg \max_{\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}}} \mathcal{U}_p(\hat{\mathcal{H}}, \hat{\boldsymbol{\theta}} | \mathbf{g}). \quad (7)$$

Equivalently, the Bayes decision rule maximizes the total expected utility, $\mathscr{U}$, since by the law of iterated expectations, $\mathscr{U}$ is equal to the expectation of $\mathscr{U}_p$ with respect to **g**; this is similar to the equivalence of minimizing either posterior expected loss or Bayes risk [34, Sec. 4.4].

For a binary detection task, there are four possible outcomes: TP, FN, TN, and FP. Consequently, the utility function for a joint binary detection-estimation task can be decomposed in terms of four utilities: $V_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, $V_{\mathrm{FN}}(\boldsymbol{\theta})$, $V_{\mathrm{TN}}$, and $V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})$. The letter "V" is used to distinguish these utilities for the joint detection-estimation task from the binary detection utilities, $U_{\mathrm{TP}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TN}}$, and $U_{\mathrm{FP}}$, introduced in Section II-A2. More generally, the utilities can also depend on parameters that are not associated with signal presence [20, p. 908]. However, the level of generality used here is sufficient to derive the optimal decision rules that maximize ROC-type curves.

With the above utilities, it follows that

$$\begin{aligned}\mathscr{U}_p(\hat{\mathscr{H}}, \hat{\boldsymbol{\theta}}|\mathbf{g}) = \mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_0)\,[V_{\mathrm{TN}} P(\mathscr{H}_0|\mathbf{g}) \\ + \int_{\Theta} V_{\mathrm{FN}}(\boldsymbol{\theta}) p(\mathscr{H}_1, \boldsymbol{\theta}|\mathbf{g}) d\boldsymbol{\theta}\Big] + \mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_1) \\ \times \Big[V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) P(\mathscr{H}_0|\mathbf{g}) + \int_{\Theta} V_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) p(\mathscr{H}_1, \boldsymbol{\theta}|\mathbf{g}) d\boldsymbol{\theta}\Big]. \end{aligned} \tag{8}$$

Define the conditional likelihood ratio as

$$\Lambda(\mathbf{g}|\boldsymbol{\theta}) = \frac{p(\mathbf{g}|\mathscr{H}_1, \boldsymbol{\theta})}{p(\mathbf{g}|\mathscr{H}_0)}. \tag{9}$$

In the Appendix, we show that the Bayes decision rule is

$$\boldsymbol{\theta}^*(\mathbf{g}) = \arg \max_{\hat{\boldsymbol{\theta}}} \left\{ \frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)} V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) + \int_{\Theta} V_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \Lambda(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \right\}, \tag{10a}$$

$$T(\mathbf{g}) = \frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)} (V_{\mathrm{FP}}(\boldsymbol{\theta}^*) - V_{\mathrm{TN}}) + \int_{\Theta} (V_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*) - V_{\mathrm{FN}}(\boldsymbol{\theta})) \Lambda(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{10b}$$

$$\mathscr{H}^* = \begin{cases} \mathscr{H}_0 & \text{if } T(\mathbf{g}) \leq 0 \\ \mathscr{H}_1 & \text{if } T(\mathbf{g}) > 0. \end{cases} \tag{10c}$$

The optimal decision rule finds the estimate $\boldsymbol{\theta}^*$, calculates the decision statistic, $\mathcal{T}(\mathbf{g})$, and decides in favor of signal-presence if $\mathcal{T}(\mathbf{g})$ is positive. Note that (10) is equivalent to equation (23) in Khurd et al. [29].

Observe that the above formulation implicitly includes the possibility of nuisance parameters, which are not of interest in themselves, and therefore do not affect expected utility. Specifically, the signal-parameter vector can be partitioned as $\boldsymbol{\theta} = [\boldsymbol{\phi}, \boldsymbol{\nu}]$, where $\boldsymbol{\phi}$ is a parameter vector of interest, and $\boldsymbol{\nu}$ is a vector of nuisance parameters. In this case, the utilities $V_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, $V_{\mathrm{FN}}(\boldsymbol{\theta})$, $V_{\mathrm{TN}}$, and $V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})$ do not depend on $\boldsymbol{\nu}$, but the conditional likelihood ratio, $\Lambda(\mathbf{g} \mid \boldsymbol{\theta})$ and the prior density, $p(\boldsymbol{\theta})$, do depend on $\boldsymbol{\nu}$. Consequently, the decision rule (10) marginalizes over the nuisance parameters. In the extreme case that all of the parameters are nuisance parameters, i.e., $\boldsymbol{\theta} = \boldsymbol{\nu}$, the task is pure binary detection, and (10) reduces to the classical likelihood ratio test [26].

## IV. The optimal ROC-type decision rule

We will use the optimal Bayes rule given in the previous section to derive decision rules that maximize ROC-type curves. For this purpose, we introduce a special case of the general utility structure in the previous section.

### A. Utility framework

First, recall that $U_{\mathrm{TN}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TP}}$, and $U_{\mathrm{FP}}$ are constituent utilities for outcomes of the binary detection task. Second, let $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})$ and $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ be disutility functions corresponding to parameter estimates for FP and TP decisions, respectively. These functions will be used to penalize incorrect and inaccurate parameter estimates, where larger values indicate a greater penalty. Last, let $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ be a utility function corresponding to parameter estimates for TP decisions, where a larger value corresponds to a greater reward. For our purposes, the above disutility and utility functions for parameter estimates will always be taken to be positive, but this requirement is not necessary. Summarizing, $U_{\mathrm{TN}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TP}}$, and $U_{\mathrm{FP}}$ are utilities for the binary detection task, and $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})$, $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, and $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ are disutility and utility functions, respectively, for the estimation task.

Now, define the following utility structure for the joint detection-estimation task:

$$V_{\mathrm{TN}} = U_{\mathrm{TN}}, \quad V_{\mathrm{FN}}(\boldsymbol{\theta}) = U_{\mathrm{FN}} \tag{11a}$$

$$V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) = \Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) U_{\mathrm{FP}} + (1 - \Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})) U_{\mathrm{TN}} \tag{11b}$$

$$V_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) U_{\mathrm{TP}} + (1 - \Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})) U_{\mathrm{FN}} + \Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})(U_{\mathrm{FP}} - U_{\mathrm{TN}}). \tag{11c}$$

Since this structure is a special case of the utility structure of Section III, some generality is lost, e.g, $V_{\mathrm{FN}}$ does not depend on $\boldsymbol{\theta}$. However, we will see that (11) is flexible enough to encompass many variants of ROC-type curves and yet sufficiently constrained so as to allow a derivation of a linear expected utility equation similar to that for binary detection.

The structure in (11) has two properties worth noting. First, observe that when $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) \equiv 1$, $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \equiv 0$, and $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \equiv 1$, (11) reduces to the utility structure for simple binary detection in Section II-A2. Second, (11) decomposes $V_{\mathrm{TN}}$, $V_{\mathrm{FN}}$, $V_{\mathrm{FP}}$, and $V_{\mathrm{TP}}$ as affine combinations of $U_{\mathrm{TN}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TP}}$, and $U_{\mathrm{FP}}$, i.e., the coefficients sum to one. Because affine combinations commute with affine transformations, this second property implies that any positive affine transformation of the constituent utilities $U_{\mathrm{TN}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TP}}$, and $U_{\mathrm{FP}}$ will translate into the same positive affine transformation of $V_{\mathrm{TN}}$, $V_{\mathrm{FN}}$, $V_{\mathrm{FP}}$, and $V_{\mathrm{TP}}$, and vice-versa. Therefore, (11) is consistent with interval-scale utility [32], [33].

We define the expected FP and TP disutilities as

$$\mathscr{D}_{\mathrm{FP}} = E[\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})\mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_1) | \mathscr{H} = \mathscr{H}_0] \quad (12)$$

and

$$\mathscr{D}_{\mathrm{TP}} = E[\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})\mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_1) | \mathscr{H} = \mathscr{H}_1], \quad (13)$$

and the expected disutility for a positive decision as

$$\mathscr{D}_{\mathrm{P}} = P(\mathscr{H}_0)\mathscr{D}_{\mathrm{FP}} + P(\mathscr{H}_1)\mathscr{D}_{\mathrm{TP}}. \quad (14)$$

Similarly, we define the expected utility of a TP as

$$\mathscr{U}_{\mathrm{TP}} = E[\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})\mathscr{I}(\hat{\mathscr{H}} = \mathscr{H}_1) | \mathscr{H} = \mathscr{H}_1]. \quad (15)$$

Under the utility structure in (11), the total expected utility can be expressed as

$$\mathscr{U} = (U_{\mathrm{FP}} - U_{\mathrm{TN}})\mathscr{D}_{\mathrm{P}} + (U_{\mathrm{TP}} - U_{\mathrm{FN}})P(\mathscr{H}_1)\mathscr{U}_{\mathrm{TP}} + U_{\mathrm{TN}}P(\mathscr{H}_0) + U_{\mathrm{FN}}P(\mathscr{H}_1). \quad (16)$$

Next, rearranging terms leads to the iso-utility equation,

$$\mathscr{U}_{\mathrm{TP}} = \frac{(U_{\mathrm{FP}} - U_{\mathrm{TN}})}{(U_{\mathrm{TP}} - U_{\mathrm{FN}})P(\mathscr{H}_1)}\mathscr{D}_{\mathrm{P}} + \frac{\mathscr{U} - U_{\mathrm{TN}}P(\mathscr{H}_0) - U_{\mathrm{FN}}P(\mathscr{H}_1)}{(U_{\mathrm{TP}} - U_{\mathrm{FN}})P(\mathscr{H}_1)}. \quad (17)$$

In the special case that $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \equiv 0$, the expected disutility of a TP, $\mathscr{D}_{\mathrm{TP}}$, is zero, and (17) becomes

$$\mathscr{U}_{\mathrm{TP}}=\frac{(U_{\mathrm{FP}}-U_{\mathrm{TN}})P(\mathscr{H}_0)}{(U_{\mathrm{TP}}-U_{\mathrm{FN}})P(\mathscr{H}_1)}\mathscr{D}_{\mathrm{FP}}+\frac{\mathscr{U}-U_{\mathrm{TN}}P(\mathscr{H}_0)-U_{\mathrm{FN}}P(\mathscr{H}_1)}{(U_{\mathrm{TP}}-U_{\mathrm{FN}})P(\mathscr{H}_1)}. \quad (18)$$

At this stage, it is instructive to pause and compare the above results with the iso-utility equation for pure binary detection, (3). In particular, we see that (18), has the same form as (3), with $\mathscr{U}_{\mathrm{TP}}$ and $\mathscr{D}_{\mathrm{FP}}$ corresponding to $R_{\mathrm{TP}}$ and $R_{\mathrm{FP}}$, respectively. Likewise, (17) is also very similar to (3), with $\mathscr{D}_p$ corresponding to $R_{\mathrm{FP}}$, the only difference being the absence of $P(\mathscr{H}_0)$ in the term with $\mathscr{D}_{\mathrm{P}}$. As we observed for (3), (17) and (18) are invariant with respect to positive affine transformations of $\mathscr{U}$, $U_{\mathrm{TP}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TN}}$, and $U_{\mathrm{FP}}$, which provides additional evidence that utility structure in (11) is consistent with interval-scale utility.

Table I lists specical cases for which $\mathscr{U}_{\mathrm{TP}}$ corresponds to the ordinate of an ROC-type summary curve and either $\mathscr{D}_{\mathrm{P}}$ or $\mathscr{D}_{\mathrm{P}}/P(\mathscr{H}_0)$ corresponds to the abscissa. For example, in the fourth row, the expressions for the FROC axes in Section II-C imply that $\mathscr{D}_{\mathrm{P}}$ is equal to the mean number of FPs per observation, $E[N_{\mathrm{FP}}]$, and $\mathscr{U}_{\mathrm{TP}}$ is equal to the signal detection fraction, $R_{\mathrm{TP}}^{\mathrm{F}}$. Each of the summary curves discussed in Section II is included in Table I, with the exception of EFROC. (Recall from Section II that since maximizing the FROC curve is equivalent to maximizing the EFROC curve, it is not necessary to treat the EFROC curve separately.)

Given the above observations, we can make the notion of an "ROC-type" summary curve more precise. Namely, any curve that plots $\mathscr{U}_{\mathrm{TP}}$ versus a strictly increasing function of $\mathscr{D}_{\mathrm{P}}$ can be said to be a member of the "ROC-type" family of performance summary curves. Later, we examine possibilities for new types of summary curves from this family. It is useful to give the plot of $\mathscr{U}_{\mathrm{TP}}$ versus $\mathscr{D}_{\mathrm{P}}$ a name, so we call it the utility ROC (UROC) curve.

## B. Optimal decision rule

Under the reasonable assumption that $U_{\mathrm{TP}} > U_{\mathrm{FN}}$, the expected utility equation (16) implies that increasing $\mathscr{U}_{\mathrm{TP}}$ for any fixed value of $\mathscr{D}_{\mathrm{P}}$ leads to higher expected utility. (This is simply a mathematical statement that a higher curve is better.) Therefore, maximizing expected utility, $\mathscr{U}$, is equivalent to maximizing $\mathscr{U}_{\mathrm{TP}}$ for each $\mathscr{D}_{\mathrm{P}}$. Hence, we can obtain the decision rule that maximizes any ROC-type summary curve by inserting the utility structure (11) into the general Bayes rule (10). Doing this, we obtain the following decision rule:

$$\Lambda(\mathbf{g}|\boldsymbol{\theta})=\frac{p(\mathbf{g}|\mathscr{H}_1,\boldsymbol{\theta})}{p(\mathbf{g}|\mathscr{H}_0)} \quad (19a)$$

$$T_0=\frac{P(\mathscr{H}_0)}{P(H_1)}\frac{(U_{\mathrm{TN}}-U_{\mathrm{FP}})}{(U_{\mathrm{TP}}-U_{\mathrm{FN}})} \quad (19b)$$

$$\boldsymbol{\theta}^*(\mathbf{g})=\arg\max_{\hat{\boldsymbol{\theta}}}\left\{-T_0\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}})+\int_\Theta\left[\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})\ -\Delta_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})T_0\frac{P(\mathscr{H}_1)}{P(\mathscr{H}_0)}\right]\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\right\}$$

(19c)

$$T(\mathbf{g})=-T_0\Delta_{\mathrm{FP}}(\boldsymbol{\theta}^*)+\int_\Theta\left[\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta},\boldsymbol{\theta}^*)\ -\Delta_{\mathrm{TP}}(\boldsymbol{\theta},\boldsymbol{\theta}^*)T_0\frac{P(\mathscr{H}_1)}{P(\mathscr{H}_0)}\right]\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} \quad (19d)$$

$$\mathscr{H}^*=\begin{cases}\mathscr{H}_0 & \text{if } T(\mathbf{g})\le 0\\ \mathscr{H}_1 & \text{if } T(\mathbf{g})>0.\end{cases} \quad (19e)$$

Above, $T_0$ can be interpreted as a decision threshold. For example, in the EROC case of Table I, since $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) = 1$ and $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = 0$, signal-presence is determined by comparing the integral in (19d) to $T_0$. More generally, we see that the decision threshold is included within the integrand in (19d).

A notable aspect of the above decision rule is that the constituent utilities, $U_{\mathrm{TP}}$, $U_{\mathrm{FN}}$, $U_{\mathrm{TN}}$, and $U_{\mathrm{FP}}$ are absorbed by the decision threshold, $T_0$. Therefore, knowledge of these utilities is not required if the threshold can be determined by other means. Namely, $T_0$ can be adjusted to achieve a desired value of the summary curve abscissa, and the resulting decision rule will maximize the ordinate.

For the special cases in Table I, the decision rule in (19) yields the optimal decision strategies [19], [26], [28], [29] that maximize ROC, LROC, EROC, FROC, AFROC, and SAA-AFROC curves, respectively. Thus, the above theory unifies previous results within a single decision-theoretic framework.

### C. Some Special Cases

To better appreciate the generality of the optimal decision rule (19), it is informative to consider the EROC-optimal decision rule, and special cases that correspond to the commonly-used generalized likelihood ratio test (GLRT) and maximum *a posteriori* (MAP) decision rules [26], [27]. From Table I, the EROC curve corresponds to the choices $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) = 1$ and $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = 0$, so the EROC-optimal decision rule is

$$\boldsymbol{\theta}^*(\mathbf{g})=\arg\max_{\hat{\boldsymbol{\theta}}}\left\{\int_\Theta\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\right\} \quad (20a)$$

$$T(\mathbf{g}) = \int_{\Theta} \Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \Lambda(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (20b)$$

$$\mathscr{H}^* = \begin{cases} \mathscr{H}_0 & \text{if } T(\mathbf{g}) \leq T_0 \\ \mathscr{H}_1 & \text{if } T(\mathbf{g}) > T_0. \end{cases} \quad (20c)$$

Note that the parameter estimate, $\boldsymbol{\theta}^*(\mathbf{g})$, in (20a) is simply the Bayes estimator for $\boldsymbol{\theta}$ with utility function $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*)$. In addition, the decision statistic is essentially a weighted average of the likelihood ratio, where the weighting function is $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*)$. Since the decision statistic is smaller for inaccurate parameter estimates, the decision rule is less likely to decide in favor of signal-presence when the parameter estimate is poor. By contrast, the likelihood ratio test, which corresponds to $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*) \equiv 1$, simply marginalizes the likelihood ratio over the parameter vector.

If the utility function for the parameter estimate is a Dirac delta function, i.e., $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \delta(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$, and the prior on the parameter vector is uniform over a set $\Theta$ with finite Lebesgue measure $\mu(\Theta)$, so that $p(\theta) = 1/\mu(\Theta)$, then (20) is equivalent to the decision rule

$$\boldsymbol{\theta}^*(\mathbf{g}) = \arg \max_{\boldsymbol{\theta}} \{p(\mathbf{g}|\mathscr{H}_1, \boldsymbol{\theta})\} \quad (21a)$$

$$\mathscr{H}^* = \begin{cases} \mathscr{H}_0 & \text{if } \Lambda(\mathbf{g}|\boldsymbol{\theta}^*) \leq T_0 \mu(\Theta) \\ \mathscr{H}_1 & \text{if } \Lambda(\mathbf{g}|\boldsymbol{\theta}^*) > T_0 \mu(\Theta) \end{cases} \quad (21b)$$

In this case, (21a) is the maximum likelihood estimator (MLE) for $\boldsymbol{\theta}$ and (21b) is the GLRT [26], [27].

Similarly, taking $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \delta(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$, assuming that $U_{\mathrm{TN}} - U_{\mathrm{FP}} = U_{\mathrm{TP}} - U_{\mathrm{FN}}$, and applying Bayes' rule, it is straightforward show that (20) becomes

$$\boldsymbol{\theta}^*(\mathbf{g}) = \arg \max_{\boldsymbol{\theta}} \{p(\mathscr{H}_1, \boldsymbol{\theta}|\mathbf{g})\} \quad (22a)$$

$$\mathscr{H}^* = \begin{cases} \mathscr{H}_0 & \text{if } p(\mathscr{H}_1, \boldsymbol{\theta}^*|\mathbf{g}) \leq P(\mathscr{H}_0|\mathbf{g}) \\ \mathscr{H}_1 & \text{if } p(\mathscr{H}_1, \boldsymbol{\theta}^*|\mathbf{g}) > P(\mathscr{H}_0|\mathbf{g}), \end{cases} \quad (22b)$$

which is the MAP decision rule for joint detection and estimation [39].

## V. Example: Known Signal with Unknown Amplitude

We illustrate the theory in the previous section with an example involving detection of a known signal with unknown amplitude in white Gaussian noise (WGN). This task can be viewed as a surrogate for tumor detection in positron emission tomography (PET), where the task is to detect a lesion and estimate its maximum standardized uptake value (SUVmax) [4]. We compare the classical GLRT-MLE detection-estimation strategy to two decision rules that maximize ROC-type curves.

Below, 2-D images are writen as column vectors. A $p \times 1$ random vector, **x**, following a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix, Σ, will be denoted as $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \Sigma)$. Similarly, $x \sim \mathcal{N}(\mu, \sigma^2)$ will denote a scalar random variable with mean $\mu$ and variance $\sigma^2$.

Let **g** be an $m \times 1$ observation vector, which under the signal-absent and signal-present hypotheses, takes the form

$$\mathscr{H}_0: \mathbf{g} = \mathbf{w} \quad (23a)$$

$$\mathscr{H}_1: \mathbf{g} = A\mathbf{s} + \mathbf{w}. \quad (23b)$$

Above, **w** is an $m \times 1$ noise vector, **s** is a known $m \times 1$ signal vector, and $A$ is an unknown scalar amplitude. Given a measurement, **g**, the task is to decide if the signal, **s**, is present and if so, to estimate the amplitude, $A$. For PET tumor detection, **s** is the lesion and $A$ is SUVmax. The noise vector, **w**, is assumed to be zero-mean, WGN with variance $\sigma_w^2$, i.e., $\mathbf{w} \sim \mathcal{N}_m(\mathbf{0}, \sigma_w^2 I)$, where $I$ is the $m \times m$ identity matrix. Also, we assume that the amplitude, $A$, is random, statistically independent of **w**, and follows a normal distribution with mean $\mu_A$ and variance $\sigma_A^2$, i.e., $A \sim \mathcal{N}(\mu_A, \sigma_A^2)$.

A convenient aspect of the above detection-estimation problem is that closed-form results are known for some cases. Namely, the MLE for $A$ and the GLRT are [27, p. 254–255]

$$\hat{A}_{\mathrm{MLE}} = \frac{\mathbf{s}^T \mathbf{g}}{\mathbf{s}^T \mathbf{s}} \quad (24a)$$

$$T_{\mathrm{G}}(\mathbf{g}) = (\hat{A}_{\mathrm{MLE}})^2 \quad (24b)$$

$$\hat{\mathscr{H}}_{\mathrm{G}}=\begin{cases}\mathscr{H}_0 & \text{if } T_{\mathrm{G}}(\mathbf{g})\leq\tau\\ \mathscr{H}_1 & \text{if } T_{\mathrm{G}}(\mathbf{g})>\tau,\end{cases} \quad (24c)$$

where $\tau$ is the decision threshold. In addition, the optimal ROC decision strategy, the likelihood ratio test, marginalizes over $A$, and the decision statistic is [27, p. 257–258]

$$T_L(\mathbf{g})=\mu_A(\mathbf{s}^T\mathbf{g})+\frac{\sigma_A^2}{2\sigma_w^2}(\mathbf{s}^T\mathbf{g})^2, \quad (25)$$

which is compared to a threshold to make the decision.

To define an EROC curve, we assume a Gaussian utility function, $\Upsilon_{\mathrm{TP}}(A,\hat{A})=\exp[-(A-\hat{A})^2/(2\sigma_u^2)]$, where $\sigma_u$ is a parameter to be specified. In the context of PET tumor detection, this utility function rewards estimates of SUVmax that are closer to the true value. In a supplementary document available under the multimedia tab for this paper on the journal website, it is shown that the optimal EROC decision rule is

$$\hat{A}_{\mathrm{E}}(\mathbf{g})=\frac{\sigma_A^2\mathbf{s}^T\mathbf{g}+\sigma_w^2\mu_A}{\sigma_w^2+\sigma_A^2\mathbf{s}^T\mathbf{s}} \quad (26a)$$

$$T_{\mathrm{E}}(\mathbf{g})=T_L(\mathbf{g}). \quad (26b)$$

$$\hat{\mathscr{H}}_{\mathrm{E}}=\begin{cases}\mathscr{H}_0 & \text{if } T_{\mathrm{E}}(\mathbf{g})\leq\tau\\ \mathscr{H}_1 & \text{if } T_{\mathrm{E}}(\mathbf{g})>\tau.\end{cases} \quad (26c)$$

The above decision rule possesses three interesting properties. First, it can be shown that the amplitude estimate (26a) is the mean of the posterior distribution, $p(A|\,\mathbf{g},\mathscr{H}_1)$. Second, (26) is independent of the utility function parameter, $\sigma_u$. Lastly, the decision statistic (26b) is identical to that for the likelihood ratio test, so it is also the optimal ROC decision rule.

Also, we define a UROC curve by taking the utility function $\Upsilon_{\mathrm{TP}}(A\,\hat{A})$ to be a Gaussian as above, and by defining the disutility functions as $\Delta_{\mathrm{TP}}(A,\hat{A})\equiv 0$ and $\Delta_{\mathrm{FP}}(\hat{A})=10^{|\hat{A}|}$. In contrast to the EROC curve, this UROC curve penalizes amplitude estimates for FP decisions nonuniformly. Namely, amplitude estimates that are larger in magnitude result in a larger penalty for FP decisions. This penalty is desireable in the PET tumor detection setting, since FP decisions with larger estimates of SUVmax may result in more aggressive treatment.

In the aforementioned supplementary document accompanying this paper on the journal website, we show that the EROC-optimal objective function is

$$F(\mathbf{g},\hat{A})=\int_{-\infty}^{\infty}\Upsilon_{\mathrm{TP}}(A,\hat{A})\Lambda(\mathbf{g}|A)p(A)\mathrm{d}A \quad (27)$$

$$=\frac{\tilde{\sigma}_w\sigma_u}{\sqrt{(\tilde{\sigma}_w^2+\sigma_A^2)(\sigma_u^2+\sigma_p^2)}}\exp\left[\frac{(\hat{A}_{\mathrm{MLE}})^2}{2\tilde{\sigma}_w^2}-\frac{(\hat{A}_{\mathrm{MLE}}-\mu_A)^2}{2(\tilde{\sigma}_w^2+\sigma_A^2)}-\frac{(\hat{A}-\mu_p)^2}{2(\sigma_u^2+\sigma_p^2)}\right], \quad (28)$$

where $\tilde{\sigma}_w = \sigma_w/\|s\|$. From (19), it follows that the optimal UROC decision rule is

$$\hat{A}_{\mathrm{U}}=\arg\max_{\hat{A}}\left\{-T_0 10^{|\hat{A}|}+F(\mathbf{g},\hat{A})\right\} \quad (29a)$$

$$T_{\mathrm{U}}(\mathbf{g})=-T_0 10^{|\hat{A}|}+F(\mathbf{g},\hat{A}_{\mathrm{U}}) \quad (29b)$$

$$\hat{\mathscr{H}}_{\mathrm{U}}=\begin{cases}\mathscr{H}_0 & \text{if } T_{\mathrm{U}}(\mathbf{g})\le 0\\ \mathscr{H}_1 & \text{if } T_{\mathrm{U}}(\mathbf{g})>0,\end{cases} \quad (29c)$$

where $T_0$ is the decision threshold.

We performed a Monte Carlo evaluation to compare the GLRT-MLE, EROC-optimal, and UROC-optimal decision rules. A 32 × 32 image region of interest (ROI) was used, so that $m$ = 1024, and the signal, **s**, was taken to be a Gaussian, with $k$th entry $s_k=\exp[(-x_k^2-y_k^2)/(2\sigma_s^2)]$, where $\sigma_s$ = 3 and $x_k = y_k = -16 + (k-1)$ for $k$ = 1, 2, …, 32. The remaining parameters were $\sigma_w$ = 2, $\mu_A$ = 1, $\sigma_A$ = 0.5, $\sigma_u$ = 0.2. Fig. 2 shows empirical ROC, EROC, and UROC curves estimated from 10, 000 Monte Carlo trials, in which half of the images contained a signal and half did not. The EROC-optimal and UROC-optimal decision rules both outperformed the GLRT-MLE strategy for each summary curve, which was expected since the GLRT-MLE does not utilize the prior distribution for $A$. In adddition, the ROC curves were observed to be nearly the same for the EROC-optimal and UROC-optimal strategies, which indicates that the UROC-optimal decision rule did not sacrifice accuracy for the detection component of the task. As expected, the EROC-optimal and UROC-optimal decision rules resulted in higher EROC and UROC curves, respectively.

## VI. New performance summary curves for detection tasks with an unknown number of signals

The theory in Section IV is suggestive of new classes of summary curves and associated optimal decision rules for joint detection-estimation with an unknown number of signals. Below, we present examples of new types of summary curves when signal-localization is an important factor.

New summary curves can be defined by combining the EROC paradigm with aspects of either the FROC or AFROC paradigms. Specifically, suppose that the task consists of detecting and correctly localizing $n \geq 0$ signals, where $n$ is random and unknown, together with estimation of additional parameters associated with each signal. Let the parameter vector for the $i$th signal be of the form $\boldsymbol{\theta}_i = [\mathbf{r}_i, \boldsymbol{\eta}_i]$, where $\mathbf{r}_i$ is the signal location (coordinate vector) and $\boldsymbol{\eta}_i$ is a vector of additional parameters of interest, such as amplitude, width, orientation, etc. Similarly, let the $j$th signal-parameter estimate be $\hat{\boldsymbol{\theta}}_j = [\hat{\mathbf{r}}_j, \hat{\boldsymbol{\eta}}_j]$. We write the overall parameter vector for a signal-present observation as $\boldsymbol{\theta} = [n, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_n]$, and denote its estimate as $\hat{\boldsymbol{\theta}} = [\hat{n}, \hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \ldots, \hat{\boldsymbol{\theta}}_{\hat{n}}]$. Above, note that the indices of $\hat{\boldsymbol{\theta}}_j$ do not necessarily correspond to the indices of $\boldsymbol{\theta}_i$.

To assess signal-localization, we first define

$$f(i) = \arg\min_j \|\mathbf{r}_i - \hat{\mathbf{r}}_j\|, \tag{30}$$

which is the index of the signal-location estimate closest to signal $i$, and

$$g(j) = \arg\min_i \|\mathbf{r}_i - \hat{\mathbf{r}}_j\|, \tag{31}$$

which is the index of the signal closest to signal-location estimate $j$. With these index functions, the TP and FP counting functions for a signal-present observation can be defined as

$$M_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} \mathscr{I}(\|\hat{\mathbf{r}}_{f(i)} - \mathbf{r}_i\| \leq R) \tag{32}$$

and

$$M_{\mathrm{FP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \sum_{j=1}^{\hat{n}} \mathscr{I}(\|\hat{\mathbf{r}}_j - \mathbf{r}_{g(j)}\| > R), \tag{33}$$

respectively, where $R$ is an acceptable radius for correct localization. Next, motivated by the definition of the EROC ordinate, we let $u(\hat{\boldsymbol{\eta}}_{f(i)}, \boldsymbol{\eta}_i)$ be a utility function for the estimate of $\boldsymbol{\eta}_i$, and define the utility function

$$\gamma(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \frac{1}{n}\sum_{i=1}^{n} u(\hat{\boldsymbol{\eta}}_{f(i)}, \boldsymbol{\eta}_i)\mathscr{I}(\|\hat{\mathbf{r}}_{f(i)} - \mathbf{r}_i\| \leq R). \tag{34}$$

Finally, setting $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \gamma(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, we see that the resulting $\mathscr{U}_{\mathrm{TP}}$ combines aspects of the EROC and FROC ordinates into a single expression. We call the plot of $\mathscr{U}_{\mathrm{TP}}$ versus the FROC absicssa, $E[N_{\mathrm{FP}}]$, the multiple estimate FROC curve (MEFROC). Similarly, we call the plot of $\mathscr{U}_{\mathrm{TP}}$ versus the AFROC abscissa, $P(N_{\mathrm{FP}} > 0)$ the multiple estimate AFROC (MEAFROC) curve. Note that when $u(\hat{\boldsymbol{\eta}}_{f(i)}, \boldsymbol{\eta}_i) \equiv 1$, MEFROC and MEAFROC curves reduce to FROC and AFROC curves respectively, since in this case, $\gamma(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = M_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})/n$.

The MEFROC and MEAFROC curves defined above only penalize poor estimates of the parameters $\boldsymbol{\eta}_i$ for TP detections. More generally, we can define a summary curve that penalizes parameter estimates $\hat{\boldsymbol{\eta}}_j$ for FP detections. Specifically, let $w(\hat{\boldsymbol{\eta}}_j)$ be a disutility function for a parameter estimate arising from a FP detection, let

$$\alpha(\hat{\boldsymbol{\theta}}) = \sum_{j=1}^{\hat{n}} w(\hat{\boldsymbol{\eta}}_j), \tag{35}$$

be a disutility function for FP detections in signal-absent observations, and let

$$\beta(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \sum_{j=1}^{\hat{n}} w(\hat{\boldsymbol{\eta}}_j)\mathscr{I}(\|\hat{\mathbf{r}}_j - \mathbf{r}_{g(j)}\| > R) \tag{36}$$

be a disutility function for FP detections in signal-present observations. Defining $\Delta_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) = \alpha(\hat{\boldsymbol{\theta}})$, $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \beta(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ and $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \gamma(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, we call the resulting summary curve plotting $\mathscr{U}_{\mathrm{TP}}$ versus $\mathscr{D}_{\mathrm{P}}$ the generalized FROC (GFROC) curve. Note that the GFROC curve reduces to the MEFROC curve in the special case that $w(\hat{\boldsymbol{\eta}}_j) \equiv 1$.

The disutility and utility functions for MEFROC, MEAFROC, and GFROC curves are summarized in Table II. The decision rule (19) maximizes the height of each curve.

## VII. Discussion and Conclusions

Theoretically-optimal approaches for combined detection and estimation tasks are of interest in medical imaging because they give upper bounds on observer performance and can potentially be utilized for imaging system optimization, evaluation of observer efficiency, and development of image formation algorithms. This paper presents a unified, decision-theoretic derivation of optimal decision rules for joint detection and estimation that maximize ROC-type performance summary curves. A key component of our approach is

that ROC-type summary curves can be interpreted as plots of an expected utility versus an expected disutility for signal-present decisions. Our derivation relies on a general utility structure that concisely unifies previous work on utility theory for ROC-type curves [30]. Using this structure, decision rules that maximize ROC, LROC, EROC, FROC, AFROC, SAA-AFROC, and EFROC curves are placed in a common framework, succinctly summarizing previous results in [19], [28], [29]. In addition, our general utility framework is suggestive of new performance summary curves and associated optimal decision rules. For example, Section V utilizes a UROC curve that nonuniformly penalizes parameter estimates for FP decisions. In Section VI, aspects of EROC and FROC curves are combined to introduce MEFROC, MEAFROC, and GFROC curves for joint detection-estimation problems involving an unknown number of signals in each observation.

An advantage of our utility-based formulation of performance summary curves and decision rules is that rewards and penalties for correct and incorrect decisions are made explicit. This characteristic of our theory facilitates adaptation of the performance criterion to a given problem. For example, consider the task from Section VI that requires detecting an unknown number of signals and estimation of both location and additional supplementary parameters. For this task, because the FROC utility structure does not reward or penalize estimation of supplementary parameters, it is preferable to use a performance criterion that does, such as the GFROC curve. Hence, a careful consideration of utility forms the rationale for the choice of performance criterion [30].

A more focused application of the theory presented in this paper to practical problems is an important topic for future work. One difficulty is that computational evaluation of the optimal decision rules given here may be highly complex, particularly for problems involving an unknown number of signals, because it generally requires evaluation of high-dimensional data likelihoods, multi-dimensional integration and nonconvex optimization. Nonetheless, computationally tractable methods might be possible in some scenarios, and techniques utilized in [21], [40]–[42] may be a good starting point for such research.

## Acknowledgments

B.G. acknowledges support from a postdoctoral fellowship of the Research Foundation - Flanders (FWO, Belgium). C.K.A. was supported by the National Institutes of Health under grant number R21-EB018939.

## Appendix

Here we derive the general Bayes decision rule in (10). Starting with (8), we apply Bayes rule, recall (9), and use the fact that $\mathcal{Q}(\hat{\mathcal{H}} = \mathcal{H}_0) = 1 - \mathcal{Q}(\hat{\mathcal{H}} = \mathcal{H}_1)$ to rewrite the posterier expected utility as

$$\begin{aligned}\mathscr{U}_p(\hat{\mathscr{H}},\hat{\boldsymbol{\theta}}|\mathbf{g}) = \frac{p(\mathbf{g}|\mathscr{H}_0)P(\mathscr{H}_1)}{p(\mathbf{g})}\Bigg\{ & V_{\mathrm{TN}}\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)} \\ & + \int_{\Theta} V_{\mathrm{FN}}(\boldsymbol{\theta})\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} \\ & + \mathscr{I}(\hat{\mathscr{H}}=\mathscr{H}_1)\left[\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)}\left(V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) - V_{\mathrm{TN}}\right)\right. \\ & \left. + \int_{\Theta}\left(V_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) - V_{\mathrm{FN}}(\boldsymbol{\theta})\right)\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\right]\Bigg\}, \end{aligned} \tag{37}$$

where $p(\boldsymbol{\theta}|\mathscr{H}_1)$ is written as $p(\boldsymbol{\theta})$ for notational simplicity. Now, define the decision statistic as

$$T(\mathbf{g}) = \max_{\hat{\boldsymbol{\theta}}}\left\{\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)}\left(V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) - V_{\mathrm{TN}}\right) + \int_{\Theta}\left(V_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) - V_{\mathrm{FN}}(\boldsymbol{\theta})\right)\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}.\right. \tag{38}$$

To maximize the posterior expected utility in (37), the optimal strategy is as follows. First, if $T(\mathbf{g}) \leq 0$, decide $\mathscr{H}_0$ (Since $V_{\mathrm{TN}}$ and $V_{\mathrm{FN}}$ do not depend on $\hat{\boldsymbol{\theta}}$, it is inconsequential.) Otherwise, if $T(\mathbf{g}) > 0$, decide $\mathscr{H}_1$ and estimate $\boldsymbol{\theta}$ as

$$\begin{aligned}\boldsymbol{\theta}^*(\mathbf{g}) = \arg\max_{\hat{\boldsymbol{\theta}}} \Bigg\{ & V_{\mathrm{TN}}\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)} + \int_{\Theta} V_{\mathrm{FN}}(\boldsymbol{\theta})\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} \\ & + \frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)}\left(V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) - V_{\mathrm{TN}}\right) \\ & + \int_{\Theta}\left(V_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) - V_{\mathrm{FN}}(\boldsymbol{\theta})\right)\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\Bigg) \end{aligned} \tag{39}$$

$$= \arg\max_{\hat{\boldsymbol{\theta}}}\left\{\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)}V_{\mathrm{FP}}(\hat{\boldsymbol{\theta}}) + \int_{\Theta} V_{\mathrm{TP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})\Lambda(\mathbf{g}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\right\}. \tag{40}$$

Finally, note that the decision statistic (38) can be rewritten in terms of $\boldsymbol{\theta}^*$ to obtain (10b).

## References


1. Seltzer SE, Swensson RG, Nawfel RD, Lentin JF, Kazda I, Judy PF. Visualization and detection-localization on computed tomographic images. Invest. Radiol. 1991; 26(4):285–294. [PubMed: 2032815]
2. Lladó X, Ganiler O, Oliver A, Martí R, Freixenet J, Valls L, Vilanova JC, Rami#x000F3;-Torrentà L, Rovira À. Automated detection of multiple sclerosis lesions in serial brain MRI. Neuroradiology. 2012; 54(8):787–807. [PubMed: 22179659]
3. Zhao B, Schwartz LH, Larson SM. Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers. J. Nucl. Med. 2009; 50(2):239–249. [PubMed: 19164218]

4. Rezai P, Pisaneschi MJ, Feng C, Yaghmai V. A radiologist's guide to treatment response criteria in oncologic imaging: Functional, molecular, and disease-specific imaging biomarkers. Am. J. Roentgenol. 2013; 201(2):246–256. [PubMed: 23883206]
5. Hatt M, Cheze-le Rest C, Van Baardwijk A, Lambin P, Pradier O, Visvikis D. Impact of tumor size and tracer uptake heterogeneity in 18F-FDG PET and CT non–small cell lung cancer tumor delineation. J. Nucl. Med. 2011; 52(11):1690–1697. [PubMed: 21990577]
6. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. Radiology. 2013; 269(1):8–14. [PubMed: 24062559]
7. Scoblionko DP, Brown BG, Mitten S, Caldwell JH, Kennedy JW, Bolson EL, Dodge HT. A new digital electronic caliper for measurement of coronary arterial stenosis: comparison with visual estimates and computer-assisted measurements. Am. J. Cardiol. 1984; 53(6):689–693. [PubMed: 6702614]
8. Kirişli H, Schaap M, Metz C, Dharampal A, Meijboom WB, Papadopoulou S, Dedic A, Nieman K, De Graaf M, Meijs M, et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. Med. Image Anal. 2013; 17(8):859–876. [PubMed: 23837963]
9. Hofman PA, Stapert SZ, van Kroonenburgh MJ, Jolles J, de Kruijk J, Wilmink JT. MR imaging, single-photon emission CT, and neurocognitive performance after mild traumatic brain injury. American Journal of Neuroradiology. 2001; 22(3):441–449. [PubMed: 11237964]
10. Shenton M, Hamoda H, Schneiderman J, Bouix S, Pasternak O, Rathi Y, Vu M-A, Purohit M, Helmer K, Koerte I, et al. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. Brain Imaging Behav. 2012; 6(2):137–192. [PubMed: 22438191]
11. Guglielmi G, Muscarella S, Bazzocchi A. Integrated imaging approach to osteoporosis: state-of-the-art review and update. Radiographics. 2011; 31(5):1343–1364. [PubMed: 21918048]
12. Link TM. Osteoporosis imaging: state of the art and advanced imaging. Radiology. 2012; 263(1): 3–17. [PubMed: 22438439]
13. Metz CE. Basic principles of ROC analysis. Semin. Nucl. Med. 1978; 8:283–298. [PubMed: 112681]
14. Starr SJ, Metz CE, Lusted LB, Goodenough DJ. Visual detection and localization of radiographic images. Radiology. Sep; 1975 116(3):533–538. [PubMed: 1153755]
15. Bunch P, Hamilton J, Sanderson G, Simmons A. Free-response approach to the measurement and characterization of radiographic observer performance. J. Appl. Photogr. Eng. 1978; 4(4):166–171.
16. Chakraborty DP, Winter LHL. Free-response methodology: Alternate analysis and a new observer-performance experiment. Radiology. Mar; 1990 174(3):873–881. [PubMed: 2305073]
17. Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: Modeling, analysis, and validation. Med. Phys. Aug; 2004 31(8):2313–2330. [PubMed: 15377098]
18. Popescu LM. Nonparametric signal detectability evaluation using an exponential transformation of the FROC curve. Med. Phys. Oct; 2011 38(10):5690–5702. [PubMed: 21992384]
19. Clarkson E. Estimation receiver operating characteristic curve and ideal observers for combined detection/estimation tasks. J. Opt. Soc. Am. A. Dec.2007 24(12):B91–B98.
20. Barrett, HH., Myers, KJ. Foundations of Image Science. Hoboken, NJ: John Wiley & Sons; 2004.
21. Zhou L, Khurd P, Kulkarni S, Rangarajan A, Gindi G. Aperture optimization in emission imaging using ideal observers for joint detection and localization. Phys. Med. Biol. 2008; 53(8):2019–2034. [PubMed: 18364551]
22. Zhou L, Gindi G. Collimator optimization in SPECT based on a joint detection and localization task. Phys. Med. Biol. 2009; 54(14):4423. [PubMed: 19556684]
23. Abbey CK, Zemp RJ, Liu J, Lindfors KK, Insana MF. Observer efficiency in discrimination tasks simulating malignant and benign breast lesions imaged with ultrasound. IEEE Trans. Med. Imag. 2006; 25(2):198–209.
24. Nguyen NQ, Abbey CK, Insana MF. An adaptive filter to approximate the Bayesian strategy for sonographic beamforming. IEEE Trans. Med. Imag. 2011; 30(1):28–37.
25. Abbey CK, Eckstein MP. Observer efficiency in free-localization tasks with correlated noise. Front. Psychol. 2014; 5

26. Van Trees, HL. Detection, Estimation, and Modulation Theory, Part I. New York: John Wiley & Sons; 1968.
27. Kay, SM. Fundamentals of Statistical Signal Processing, Volume II: Detection Theory. Upper Saddle River, NJ: Pentice Hall; 1998.
28. Khurd P, Gindi G. Decision strategies that maximize the area under the LROC curve. IEEE Trans. Med. Imag. Dec.2005 24(12):1626–1636.
29. Khurd P, Liu B, Gindi G. Ideal AFROC and FROC observers. IEEE Trans. Med Imag. Feb.2010 29(2):375–386.
30. Wunderlich A, Abbey CK. Utility as a rationale for choosing observer performance assessment paradigms for detection tasks in medical imaging. Med. Phys. Nov.2013 40(11):111903. [PubMed: 24320436]
31. Stevens SS. On the theory of scales of measurement. Science. 1946; 103(2684):677–680.
32. von Neumann, J., Morgenstern, O. Theory of Games and Economic Behavior. 2. Princeton University Press; 1947.
33. Savage, LJ. The Foundations of Statistics. New York: John Wiley & Sons; 1954.
34. Berger, JO. Statistical Decision Theory and Bayesian Analysis. 2. New York: Springer; 1993.
35. Abbey CK, Eckstein MP, Boone JM. An equivalent relative utility metric for evaluating screening mammography. Med. Decis. Making. Jan-Feb;2010 30:113–122. [PubMed: 19706880]
36. Wunderlich A, Goossens B. Nonparametric estimation receiver operating characteristic analysis for performance evaluation on combined detection and estimation tasks. J. Med. Imaging. Oct.2014 1(3):031002.
37. Wunderlich A, Noo F. A nonparametric procedure for comparing the areas under correlated LROC curves. IEEE Trans. Med. Imag. Nov; 2012 31(11):2050–2061.
38. Berbaum KS, Franken EA Jr, Dorfman DD, Rooholamini SA, Kathol MH, Barloon TJ, Behlke FM, Sato Y, Lu CH, El-Khoury GY, Flickinger FW, Montgomery WJ. Satisfaction of search in diagnostic radiology. Invest Radiol. 1990; 25(2):133–140. [PubMed: 2312249]
39. Olmo G, Magli E, Presti LL. Joint statistical signal detection and estimation. Part I: Theoretical aspects of the problem. Signal Processing. 2000; 80(1):57–73.
40. Kupinski MA, Hoppin JW, Clarkson E, Barrett HH. Ideal-observer computation in medical imaging with use of Markov-chain Monte Carlo techniques. J. Opt. Soc. Am. A. 2003; 20(3):430–438.
41. Zhang L, Cavaro-Ménard C, Le Callet P, Tanguy J-Y. A perceptually relevant channelized joint observer (PCJO) for the detection-localization of parametric signals. IEEE Trans. Med. Imag. 2012; 31(10):1875–1888.
42. Zhang L, Goossens B, Cavaro-Ménard C, Callet PL, Ge D. Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size. J. Opt. Soc. Am. A. 2013; 30(11):2422–2432.

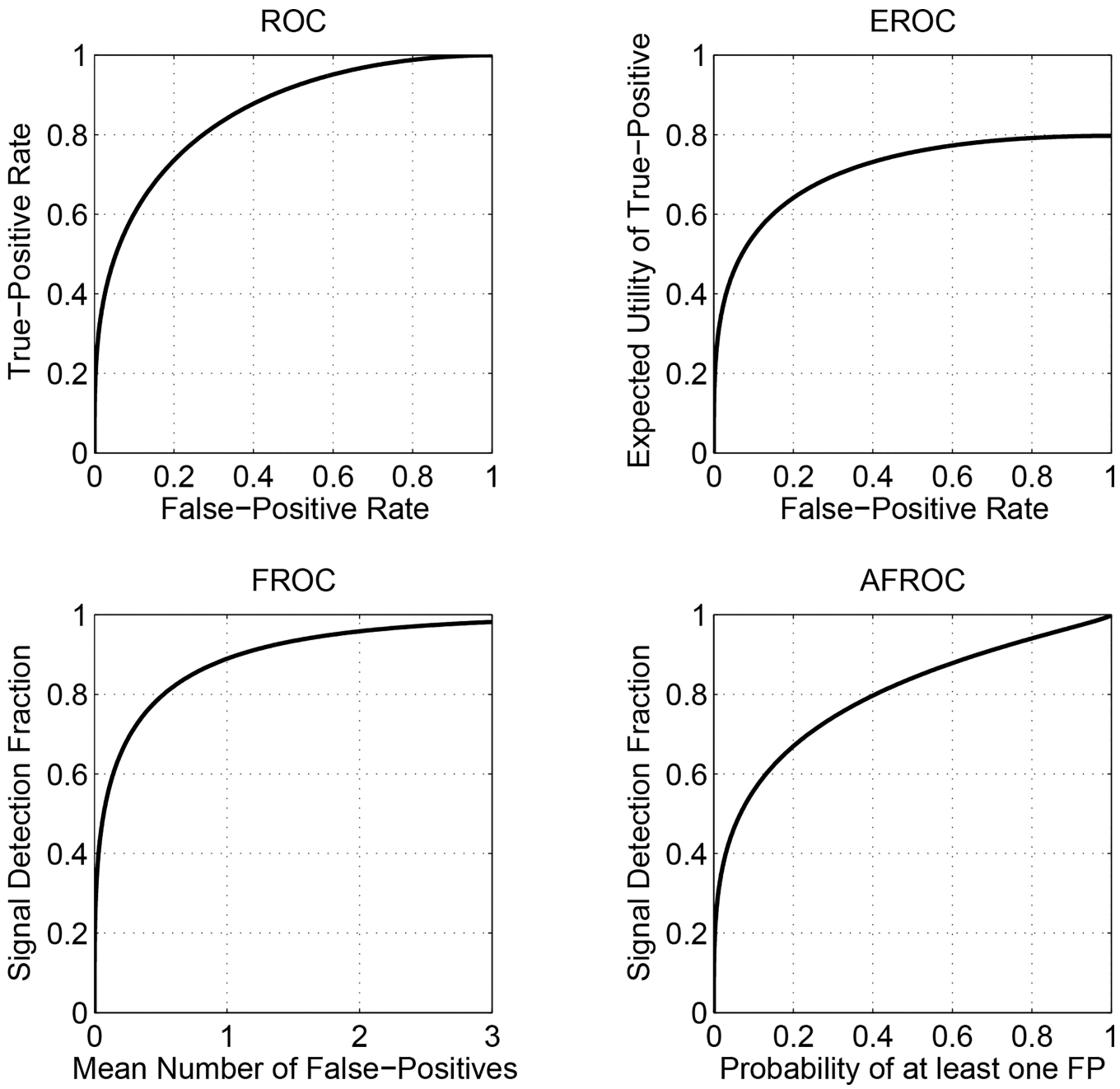


**Fig. 1.**

Examples of ROC-type summary curves. Note that the EROC curve has a utility function limited to the interval [0, 1], and that the FROC abscissa has a potentially infinite extent.

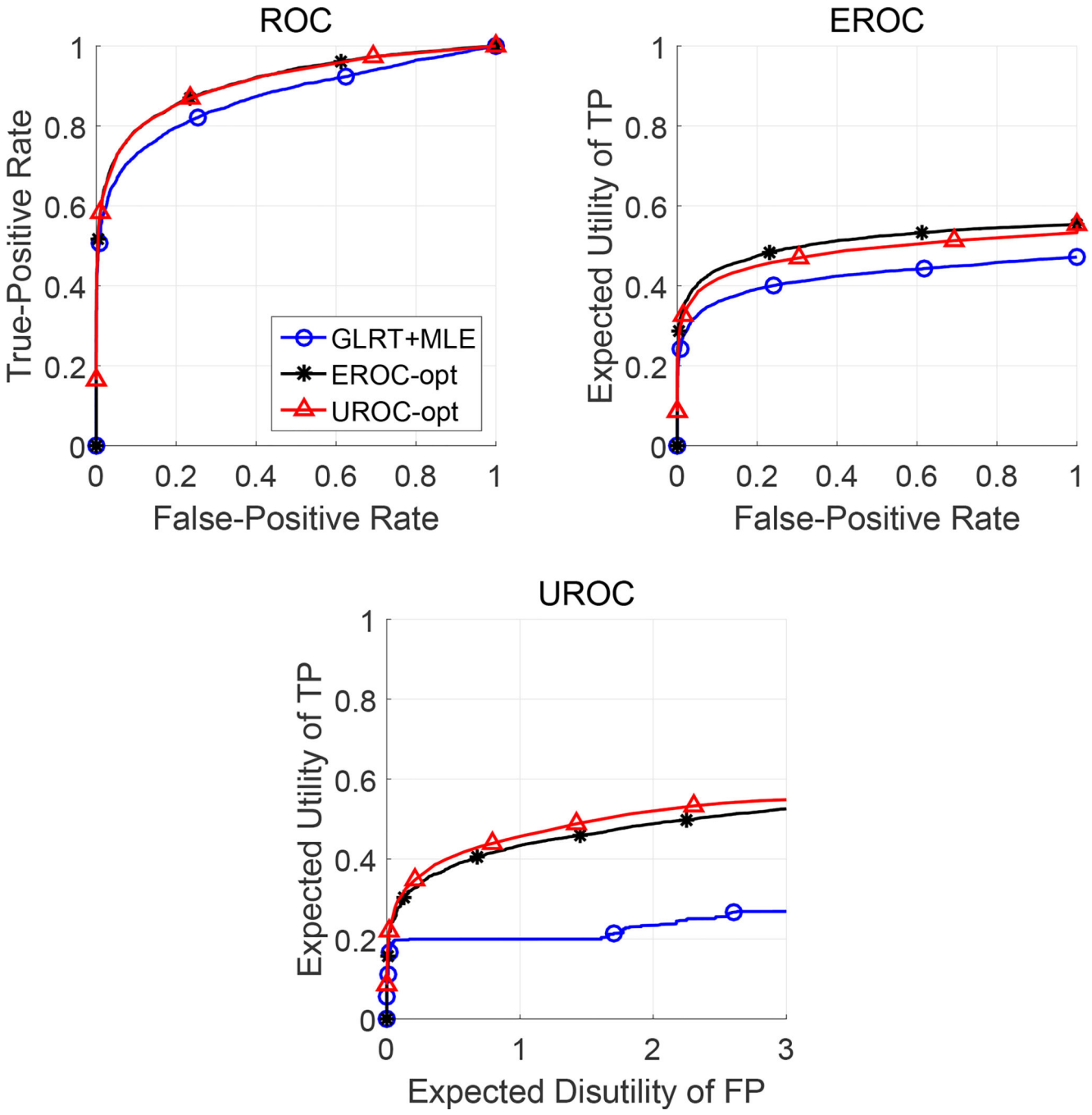


**Fig. 2.**

ROC, EROC, and UROC curves for the example, which compares GLRT-MLE, EROC-optimal, and UROC-optimal decision rules. Note that the UROC abscissa has an infinite extent, and has been truncated to [0 3] for display purposes.

**TABLE I**

Special cases of the disutility and utility functions that correspond to ROC-type summary curves; see Section II for definitions and notation. Note that $\mathscr{D}_{\mathrm{P}}$ and $\mathscr{U}_{\mathrm{TP}}$ are derived quantities defined by (14) and (15), respectively.

| curve | $\Delta_{\mathrm{FP}}(\boldsymbol{\theta}\,\hat{})$ | $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta},\boldsymbol{\theta}\,\hat{})$ | $\mathscr{D}_{\mathrm{P}}$ | $\mathscr{U}_{\mathrm{TP}}$ |
|---|---|---|---|---|---|
| ROC | 1 | 0 | 1 | $R_{\mathrm{FP}}P(\mathcal{H}_0)$ | $R_{\mathrm{TP}}$ |
| LROC | 1 | 0 | $\mathbb{I}(\|\boldsymbol{\theta}\,\hat{} - \boldsymbol{\theta}\| \le R)$ | $R_{\mathrm{FP}}P(\mathcal{H}_0)$ | $R_{\mathrm{TPL}}$ |
| EROC | 1 | 0 | $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $R_{\mathrm{FP}}P(\mathcal{H}_0)$ | $\mathscr{U}_{\mathrm{TP}}$ |
| FROC | $n\,\hat{}$ | $M_{\mathrm{FP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $M_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})/n$ | $E[N_{\mathrm{FP}}]$ | $R_{\mathrm{TP}}^{\mathrm{F}}$ |
| AFROC | 1 | $\mathbb{I}(M_{\mathrm{FP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{}) > 0)$ | $M_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})/n$ | $P(N_{\mathrm{FP}} > 0)$ | $R_{\mathrm{TP}}^{\mathrm{F}}$ |
| SAA-AFROC | 1 | 0 | $M_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})/n$ | $P(N_{\mathrm{FP}} > 0 \mid \mathcal{H} = \mathcal{H}_0)P(\mathcal{H}_0)$ | $R_{\mathrm{TP}}^{\mathrm{F}}$ |

*IEEE Trans Med Imaging*. Author manuscript; available in PMC 2017 September 01.

**TABLE II**

Disutility and utility functions that define new types of summary curves. Each curve plots $\mathscr{U}_{\mathrm{TP}}$ versus $\mathscr{D}_{\mathrm{P}}$.

| curve | $\Delta_{\mathrm{FP}}(\boldsymbol{\theta}\,\hat{})$ | $\Delta_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $\Upsilon_{\mathrm{TP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ |
|---|---|---|---|
| MEFROC | $n\,\hat{}$ | $M_{\mathrm{FP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $\gamma(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ |
| MEAFROC | 1 | $\mathscr{U}(M_{\mathrm{FP}}(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{}) > 0)$ | $\gamma(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ |
| GFROC | $\alpha(\boldsymbol{\theta}\,\hat{})$ | $\beta(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ | $\gamma(\boldsymbol{\theta}, \boldsymbol{\theta}\,\hat{})$ |