On the High-Snr Receiver Operating Characteristic of Glrt for The Conditional Signal Model

Simone Urbano, Eric Chaumette, Goupil Philippe, Jean-Yves Tourneret

▶ **To cite this version:**

Simone Urbano, Eric Chaumette, Goupil Philippe, Jean-Yves Tourneret. On the High-Snr Receiver Operating Characteristic of Glrt for The Conditional Signal Model. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Apr 2018, Calgary, Canada. 10.1109/ICASSP.2018.8462406 . hal-03023437

HAL Id: hal-03023437
https://hal.science/hal-03023437

Submitted on 26 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON THE HIGH-SNR RECEIVER OPERATING CHARACTERISTIC OF GLRT FOR THE CONDITIONAL SIGNAL MODEL

Simone Urbano[(1)(4)]*, Eric Chaumette*[(2)(4)]*, Philippe Goupil*[(1)] *and Jean-Yves Tourneret*[(3)(4)]

[(1)] Airbus Flight Control System Department, Airbus, Toulouse, France ([simone.urbano,philippe.goupil]@airbus.com)

[(2)] University of Toulouse/Isae-Supaero, 10 av. Edouard Belin, Toulouse, France (eric.chaumette@isae.fr)

[(3)] University of Toulouse/INP-ENSEEIHT/IRIT, 2 Rue Charles Camichel, Toulouse, France (Jean-Yves.Tourneret@enseeiht.fr)

[(4)] Cooperative research laboratory TéSA, 7 Boulevard de la Gare, Toulouse, France

ABSTRACT

This paper studies the performance of the generalized likelihood ratio test (GLRT) for the conditional signal model. By conditional signal model, we mean that under both hypotheses, the observations are a linear superposition of unknown deterministic signals corrupted by additive noise, with a mixing matrix depending on an unknown deterministic parameter vector. The contribution of this work is the derivation of closed form expressions for the probabilities of false alarm and of detection of the GLRT at high signal-to-noise ratio, allowing the receiver operating characteristic of the GLRT to be computed analytically. The most general case is tackled, i.e., when the number of unknown signals and the number of unknown deterministic parameters of the mixing matrix are allowed to be different under the two hypotheses.

Index Terms— Generalized likelihood ratio test, receiver operating characteristic, maximum likelihood estimation, conditional signal model.

1. INTRODUCTION

In a two hypothesis testing problem, if no a priori information about the probability of each hypothesis is available, the likelihood ratio test (LRT) in the Neyman-Pearson sense maximizes the probability of detection for a given probability of false alarm [1]. Unfortunately, optimal statistical tests such as the LRT cannot always be implemented since there are often some unknown parameters in the observation model, leading to the so-called composite hypothesis testing problem (CHTP) [1] (also referred to as joint detection estimation problem [2]). A very common approach in this situation is to replace the unknown parameters in the LRT by their maximum likelihood estimators (MLEs), following the ideas of the generalized likelihood ratio test (GLRT) [1]. It is known that the GLRT does not generally preserve the optimal properties of the LRT [1]. However, the GLRT approach has several advantages. Firstly, the test is often easy to derive and sometimes its expression and distribution can be determined analytically [3][4][5][6]. Secondly, the GLRT is known to be the uniformly most powerful (UMP) test for some restricted classes of problems [9][10]. Last but not least, in the limit of large sample support[1], the distribution of the GLRT statistics can be determined analytically for a class of linear observation models with a known mixing matrix and it is known to perform as well as the LRT in this particular case [3][4][5][6]. In this paper we are interested in the performance of GLRT for linear observation models where the mixing matrix is no longer known, but has a known parametric form. More specifically, the observations $\mathbf{y}$ are formed from a linear superposition of M unknown deterministic signals $\boldsymbol{\beta}$ corrupted by an additive noise $\mathbf{w}$ [4][7][8][11], i.e., $\mathbf{y} = \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta} + \mathbf{w}$, where the mixing matrix depends on an unknown deterministic parameter vector $\boldsymbol{\alpha}$.

Regarding the estimation of unknown quantities $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, this problem has received considerable attention during the last fifty years, both for time series analysis [4] and array processing [7][8][11]. These two problems have been merged into the framework of modern array processing [8] where mostly two different Gaussian signal models are considered: the conditional signal model (CSM) which considers the M individual signals $\boldsymbol{\beta}$ as unknown deterministic variables, and the unconditional signal model (USM) [12][13] which assumes these M individual signals to be jointly Gaussian random variables. Therefore the two different signal models have led to two kinds of maximum likelihood estimators (MLEs), namely the conditional MLE (CMLE) and the unconditional MLE (UMLE). In both cases, numerous works [8][14][15][16] have shown that in non-linear estimation problems three distinct regions of operation of the MLE can be observed. In the asymptotic region, the MSE of the MLE is small and, in many cases, close to the Cramér-Rao bound (CRB). In the *a priori* performance region where the number of independent samples and/or the signal-to-noise ratio (SNR) are very low, the observations provide little information and the MSE of the MLE is close to the one obtained from the prior knowledge about the problem. Between these two extremes, there is the transition region where the MSE of the MLE usually deteriorates rapidly with respect to the CRB, and exhibits a threshold behaviour corresponding to a "performance breakdown".

In this paper, we consider the less constrained CSM framework. Note that assuming that the M individual signals are Gaussian is a strong hypothesis that could fail in many real-life applications. Regarding the detection theory, it appears that the most studied CHTP related to CSM is the detection of signals with unknown parameters in additive Gaussian noise [8][5, §9.5][6, §28] (the restricted case where one hypothesis consists of Gaussian noise only), or the detection of two different communication signals [5, §9.3][6, §26]. As a consequence, the study of the general CHTP where under both hypotheses the observations are signals with unknown parameters in additive Gaussian noise has been somewhat overlooked. Therefore, to the best of our knowledge, the derivation of the receiver operating characteristic (ROC) of the associated GLRT, where under both hypotheses the CMLEs operate in the asymptotic region in terms of SNR, is new. It is the main contribution of this paper. We tackle the most general case where the number of signal sources and the number of unknown deterministic parameters are allowed to be different

[1] When the distribution of the MLE can be approximated by its asymptotic probability density function [3][4][5][6][7]

under the two hypotheses.

2. GENERALIZED LIKELIHOOD RATIO TEST

A wide class of detection problems can be addressed using the following two hypothesis test based on the CSM

$$\begin{cases} \mathcal{H}_0 : \mathbf{y} = \mathbf{H}_0(\boldsymbol{\alpha}_0)\boldsymbol{\beta}_0 + \mathbf{w} \\ \mathcal{H}_1 : \mathbf{y} = \mathbf{H}_1(\boldsymbol{\alpha}_1)\boldsymbol{\beta}_1 + \mathbf{w} \end{cases} \quad (1)$$

where $\mathbf{y} \in \mathbb{R}^N$ is the vector of observations, $\mathbf{w} \in \mathbb{R}^N \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ is a Gaussian noise vector with known variance σ^2, and $\mathbf{H}_0(\boldsymbol{\alpha}_0) \in \mathbb{R}^{N\times M_0}$, $\mathbf{H}_1(\boldsymbol{\alpha}_1) \in \mathbb{R}^{N\times M_1}$ are two different observation matrices that depend on the unknown parameter vectors $\boldsymbol{\alpha}_0 \in \mathbb{R}^{P_0}$, $\boldsymbol{\alpha}_1 \in \mathbb{R}^{P_1}$. By definition of the CSM [13], all the unknown parameter vectors $\boldsymbol{\beta}_0 \in \mathbb{R}^{M_0}$, $\boldsymbol{\alpha}_0 \in \mathbb{R}^{P_0}$, $\boldsymbol{\beta}_1 \in \mathbb{R}^{M_1}$ and $\boldsymbol{\alpha}_1 \in \mathbb{R}^{P_1}$ are deterministic. Note that the detection problem under consideration in (1) is more general than the classical linear model studied in [3], where the same known observation matrix is considered for the two hypotheses and the unknown vector reduces to ${\boldsymbol{\beta}_1}^2$, leading to a different detector.

Let us recall that the GLRT requires to estimate the unknown parameter vectors $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ under both hypotheses (i.e., for $i = 0, 1$) using the ML principle. Considering the hypothesis of additive white Gaussian noise of known variance σ^2 ($w(n) \sim \mathcal{N}(0, \sigma^2)$), we have

$$p(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2}\left\|\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}\right\|^2\right).$$

As a consequence, the GLRT for the detection problem (1) can be written

$$\frac{p(\mathbf{y}; \widehat{\boldsymbol{\alpha}}_1, \widehat{\boldsymbol{\beta}}_1|\mathcal{H}_1)}{p(\mathbf{y}; \widehat{\boldsymbol{\alpha}}_0, \widehat{\boldsymbol{\beta}}_0|\mathcal{H}_0)} = \frac{\exp\left(-\frac{1}{2\sigma^2}\left\|\mathbf{y} - \mathbf{H}_1(\widehat{\boldsymbol{\alpha}}_1)\widehat{\boldsymbol{\beta}}_1\right\|^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\left\|\mathbf{y} - \mathbf{H}_0(\widehat{\boldsymbol{\alpha}}_0)\widehat{\boldsymbol{\beta}}_0\right\|^2\right)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma'.$$

The estimation problem, considering a general notation that is true for both the hypotheses $\mathcal{H}_0, \mathcal{H}_1$, reduces to the minimization of the following least squares (LS) criterion [7]

$$J(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}\|^2 = (\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}). \quad (2)$$

Introducing the orthogonal projection matrix $\Pi_{\mathbf{H}} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$, that projects a vector onto the columns of $\mathbf{H}$, and $\Pi^{\perp}_{\mathbf{H}} = \mathbf{I} - \Pi_{\mathbf{H}}$, that projects a vector onto the space orthogonal to the columns of $\mathbf{H}$, the LS criterion can be rewritten as

$$\begin{aligned} J(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \|\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}\|^2 \\ &= \left\|\boldsymbol{\Pi}_{\mathbf{H}(\boldsymbol{\alpha})}(\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta})\right\|^2 + \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}(\mathbf{y} - \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta})\right\|^2, \end{aligned}$$

leading to

$$J(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \left\|\mathbf{H}(\boldsymbol{\alpha})\left(\left(\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{H}(\boldsymbol{\alpha})\right)^{-1}\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{y} - \boldsymbol{\beta}\right)\right\|^2 + \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2. \quad (3)$$

[2] In the classical linear model, there is a known relationship between $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$.

Thus, the value of $\boldsymbol{\beta}$ that minimizes $J(\boldsymbol{\alpha}, \boldsymbol{\beta})$ for a given value of $\boldsymbol{\alpha}$ is classically obtained as

$$\boldsymbol{\beta} = (\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{H}(\boldsymbol{\alpha}))^{-1}\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{y}$$

which corresponds to the classical unconstrained LS estimator of $\boldsymbol{\beta}$ for a known vector $\boldsymbol{\alpha}$. The MLE of $\boldsymbol{\alpha}$ is then obtained as the solution of the following optimization problem

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2 = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\|\boldsymbol{\Pi}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2 \\ &= \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} (\mathbf{y}^T\mathbf{H}(\boldsymbol{\alpha})(\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{H}(\boldsymbol{\alpha}))^{-1}\mathbf{H}^T(\boldsymbol{\alpha})\mathbf{y}) \quad (4) \end{aligned}$$

where we have used the properties of projection matrices[3]. Of course, there is no analytical solution for the MLE of $\boldsymbol{\alpha}$ in the general case. However, this MLE can be obtained numerically using for example a grid search approach (in general via a multivariate optimization problem). Once the MLE of $\boldsymbol{\alpha}$ has been determined, $\boldsymbol{\beta}$ is estimated as follows

$$\widehat{\boldsymbol{\beta}} = (\mathbf{H}^T(\widehat{\boldsymbol{\alpha}})\mathbf{H}(\widehat{\boldsymbol{\alpha}}))^{-1}\mathbf{H}^T(\widehat{\boldsymbol{\alpha}})\mathbf{y}.$$

Therefore, the GLRT associated with the CHTP (1) is

$$T = \frac{\left\|\mathbf{y} - \mathbf{H}_0(\widehat{\boldsymbol{\alpha}}_0)\widehat{\boldsymbol{\beta}}_0\right\|^2 - \left\|\mathbf{y} - \mathbf{H}_1(\widehat{\boldsymbol{\alpha}}_1)\widehat{\boldsymbol{\beta}}_1\right\|^2}{\sigma^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \lambda \quad (5a)$$

$$= \frac{\left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}_0(\widehat{\boldsymbol{\alpha}}_0)}\mathbf{y}\right\|^2 - \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}_1(\widehat{\boldsymbol{\alpha}}_1)}\mathbf{y}\right\|^2}{\sigma^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \lambda \quad (5b)$$

It is unlikely to obtain an analytical tractable expression of the ROC of the GLRT (5a-5b) in all regions of operation of the MLE. However, in the asymptotic region, the ROC can be derived by invoking the high SNR consistency and efficiency of the CMLE [17]. It is the objective of the next section.

3. HIGH SNR CMLE APPROXIMATION

Using a general notation for the two hypotheses, one can use a first order Taylor series expansion of $\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}$ where $\boldsymbol{\alpha}^T = (\alpha_1, \ldots, \alpha_P)$ and $\mathbf{e}_p$ denotes the pth vector of the natural basis of $\mathbb{R}^P$

$$\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\alpha_p\mathbf{e}_p)} \approx \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} + d\alpha_p \frac{\partial \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p},$$

leading to

$$\begin{aligned} \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\alpha_p\mathbf{e}_p)}\mathbf{y}\right\|^2 &\approx \left\|\left(\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} + d\alpha_p \frac{\partial \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p}\right)\mathbf{y}\right\|^2 \\ \approx \left\|\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} + d\alpha_p \frac{\partial \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p}\mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta} + d\alpha_p \frac{\partial \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p}\mathbf{w}\right\|^2. \quad (6) \end{aligned}$$

However, since $\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{H}(\boldsymbol{\alpha}) = \mathbf{0}$, we obtain

$$\frac{\partial \boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p}\mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta} = -\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\frac{\partial \mathbf{H}(\boldsymbol{\alpha})}{\partial \alpha_p}\boldsymbol{\beta} = -\boldsymbol{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\frac{\partial \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}}{\partial \alpha_p}$$

[3] $\boldsymbol{\Pi}$ is symmetric ($\boldsymbol{\Pi}^T = \boldsymbol{\Pi}$) and idempotent ($\boldsymbol{\Pi}^2 = \boldsymbol{\Pi}$)

which yields the following equivalent form of (6)

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\alpha_p\mathbf{e}_p)}\mathbf{y}\right\|^2 \approx \left\| \begin{array}{l} \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} - d\alpha_p \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} \frac{\partial \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}}{\partial \alpha_p} \\ +d\alpha_p \frac{\partial \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}}{\partial \alpha_p}\mathbf{w} \end{array} \right\|^2. \tag{7}$$

Moreover, since $\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\alpha_p\mathbf{e}_p)}\mathbf{w} \to \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{w}$ when $d\alpha_p \to 0$, it is sensible to assume that $d\alpha_p \partial \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}/\partial\alpha_p \mathbf{w}$ is asymptotically negligible in (7), leading to

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\alpha_p\mathbf{e}_p)}\mathbf{y}\right\|^2 \underset{d\alpha_p\to 0}{\to} \left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} - d\alpha_p \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} \frac{\partial \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}}{\partial \alpha_p}\right\|^2$$

i.e., in a more compact form

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\boldsymbol{\alpha})}\mathbf{y}\right\|^2 \underset{\|d\boldsymbol{\alpha}\|\to 0}{\to} \left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} - \mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) d\boldsymbol{\alpha}\right\|^2,$$
$$\mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) = \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} \frac{\partial \mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\beta}}{\partial \boldsymbol{\alpha}^T} \tag{8a}$$

or equivalently

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\boldsymbol{\alpha})}\mathbf{y}\right\|^2 \underset{\|d\boldsymbol{\alpha}\|\to 0}{\to} \left\|\mathbf{\Pi}^{\perp}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2 + \left\|\mathbf{\Pi}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} - \mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) d\boldsymbol{\alpha}\right\|^2. \tag{8b}$$

Due to the high SNR consistency of the CMLE of $\boldsymbol{\alpha}$ (denoted as $\widehat{\boldsymbol{\alpha}}$), in the asymptotic region, $\widehat{\boldsymbol{\alpha}} \simeq \boldsymbol{\alpha} + d\widehat{\boldsymbol{\alpha}}$, where $\|d\widehat{\boldsymbol{\alpha}}\| \ll 1$ [17], and (8a-8b) holds, leading to the LS minimization

$$\widehat{d\boldsymbol{\alpha}} = \arg\min_{d\boldsymbol{\alpha}} \left\{ \left\|\mathbf{\Pi}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y} - \mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) d\boldsymbol{\alpha}\right\|^2 \right\} \tag{9a}$$

which solution is:

$$\widehat{d\boldsymbol{\alpha}} = \left(\mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)^T \mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)\right)^{-1} \mathbf{D}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)^T \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}. \tag{9b}$$

Moreover, it can be easily shown that $E\left[\widehat{d\boldsymbol{\alpha}}\right] = \mathbf{0}$ and that

$$E\left[\widehat{d\boldsymbol{\alpha}}\widehat{d\boldsymbol{\alpha}}^T\right] = \mathbf{\Gamma}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)^{-1},$$
$$\left(\mathbf{\Gamma}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)\right)_{p,p'} = \frac{1}{\sigma^2}\boldsymbol{\beta}^T \frac{\partial \mathbf{H}\left(\boldsymbol{\alpha}\right)}{\partial \alpha_p}^T \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} \frac{\partial \mathbf{H}\left(\boldsymbol{\alpha}\right)}{\partial \alpha_{p'}}\boldsymbol{\beta} \tag{10}$$

where $\mathbf{\Gamma}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)^{-1}$ is actually the CRB associated with the unknown parameter vector $\boldsymbol{\alpha}$ [7], which proves that (8a-8b) provides a relevant high SNR approximation of $\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\widehat{\boldsymbol{\alpha}})}\mathbf{y}\right\|^2$. Therefore, the following high SNR approximation

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\widehat{\boldsymbol{\alpha}})}\mathbf{y}\right\|^2 = \left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha}+d\widehat{\boldsymbol{\alpha}})}\mathbf{y}\right\|^2 \approx \left\|\mathbf{\Pi}^{\perp}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2, \tag{11}$$

is relevant as well. Last but not least, it is worth noting that

$$\left\|\mathbf{\Pi}^{\perp}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{y}\right\|^2 = \mathbf{y}^T\mathbf{L}^{\perp}\mathbf{y},\ \mathbf{L}^{\perp} = \mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{\Pi}^{\perp}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})},$$

where $\mathbf{L}^{\perp}$ is an orthogonal projection matrix since $\mathbf{L}^{\perp}$ is symmetric and idempotent (one can easily verify this statement by observing that $\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{\Pi}^{\perp}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}\mathbf{\Pi}^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} = \Pi^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})} - \mathbf{\Pi}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}$ and $\Pi^{\perp}_{\mathbf{H}(\boldsymbol{\alpha})}\mathbf{\Pi}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})} = \mathbf{\Pi}_{\mathbf{D}(\boldsymbol{\alpha},\boldsymbol{\beta})}$). Surprisingly, to the best of our knowledge, high SNR approximation of the CMLE $\widehat{\boldsymbol{\alpha}}$ (9a) and of $\left\|\mathbf{\Pi}^{\perp}_{\mathbf{H}(\widehat{\boldsymbol{\alpha}})}\mathbf{y}\right\|^2$ (11) are new, which may explain why the ROC of the GLRT (5a) in the high SNR region of CSMs are not available in the open literature [3][4][5][6][7][8][11].

4. HIGH SNR ROC OF THE GLRT

From (11), the statistic T (5a-5b) can be reformulated at high SNR as

$$T = \frac{\mathbf{y}^T}{\sigma}\left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\frac{\mathbf{y}}{\sigma} \tag{12}$$

where $\mathbf{L}_0^{\perp}$ and $\mathbf{L}_1^{\perp}$ are orthogonal projection matrices with $\mathbf{L}_i^{\perp} = \mathbf{\Pi}^{\perp}_{\mathbf{H}_{\hat{i}}(\boldsymbol{\alpha}_i)}\mathbf{\Pi}^{\perp}_{\mathbf{D}_{\hat{i}}(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i)}\mathbf{\Pi}^{\perp}_{\mathbf{H}_{\hat{i}}(\boldsymbol{\alpha}_i)}$, for $i = 0, 1$. Let $\mathcal{S} = span\left\{\mathbf{A}\right\}$ denotes the linear span of the set of the column vectors of $\mathbf{A}$, $\mathbf{\Pi}_{\mathcal{S}}$ denotes the orthogonal projection matrix onto $\mathcal{S}$, and $\mathcal{S}^{\perp}$ denotes the orthogonal complement of the subspace $\mathcal{S}$ for the canonical inner product. Let $\mathcal{S}_0 = span\{\mathbf{L}_0^{\perp}\}$ and $\mathcal{S}_1 = span\{\mathbf{L}_1^{\perp}\}$, then the following results can be obtained

$$\begin{array}{l} \forall \mathbf{y} \in (\mathcal{S}_0 \cap \mathcal{S}_1) \ : \left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\mathbf{y} = \mathbf{0} \\ \forall \mathbf{y} \in (\mathcal{S}_0^{\perp} \cap \mathcal{S}_1) \ : \left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\mathbf{y} = -\mathbf{y} \\ \forall \mathbf{y} \in (\mathcal{S}_0 \cap \mathcal{S}_1^{\perp}) \ : \left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\mathbf{y} = \mathbf{y} \\ \forall \mathbf{y} \in (\mathcal{S}_0^{\perp} \cap \mathcal{S}_1^{\perp}) : \left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\mathbf{y} = \mathbf{0} \end{array}$$

where

$$(\boldsymbol{S}_0 \cap \boldsymbol{S}_1) \cup (\boldsymbol{S}_0^{\perp} \cap \boldsymbol{S}_1) \cup (\boldsymbol{S}_0 \cap \boldsymbol{S}_1^{\perp}) \cup (\boldsymbol{S}_0^{\perp} \cap \boldsymbol{S}_1^{\perp}) = \mathbb{R}^N.$$

Moreover $\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}$ is symmetric. Thus $\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}$ is diagonalizable and admits three potential eigenvalues $\{-1, 0, 1\}$ with multiplicities denoted as

$$\begin{array}{rcll} m\{-1\} & = & \dim\{\mathcal{S}_0^{\perp} \cap \mathcal{S}_1\} = \nu & (13a) \\ m\{1\} & = & \dim\{\mathcal{S}_0 \cap \mathcal{S}_1^{\perp}\} = r & (13b) \\ m\{0\} & = & N - m\{-1\} - m\{1\} = N - \nu - r & \end{array}$$

leading to

$$T = \frac{\mathbf{y}^T}{\sigma}\left(\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}\right)\frac{\mathbf{y}}{\sigma} = \left\|\mathbf{U}_1^T\mathbf{y}/\sigma\right\|^2 - \left\|\mathbf{U}_{-1}^T\mathbf{y}/\sigma\right\|^2$$

where $\mathbf{U}_1$ is a matrix whose r column vectors form an orthonormal basis of $\mathcal{S}_0 \cap \mathcal{S}_1^{\perp}$, and $\mathbf{U}_{-1}$ is a matrix whose ν column vectors form an orthonormal basis of $\mathcal{S}_0^{\perp} \cap \mathcal{S}_1$. Therefore, at high SNR, the probability of false alarm (PFA) and the probability of detection (PD) of the GLRT (5a-5b) can be expressed as:

$$\begin{array}{l} P_{FA} = P\left(\left\|\mathbf{U}_1^T\mathbf{y}/\sigma\right\|^2 - \left\|\mathbf{U}_{-1}^T\mathbf{y}/\sigma\right\|^2 > \lambda | \mathcal{H}_0\right) \\ P_D = P\left(\left\|\mathbf{U}_1^T\mathbf{y}/\sigma\right\|^2 - \left\|\mathbf{U}_{-1}^T\mathbf{y}/\sigma\right\|^2 > \lambda | \mathcal{H}_1\right). \end{array} \tag{14a}$$

It is then worth noticing that: a) $\mathcal{S}_0 \cap \mathcal{S}_1^{\perp} \subset span\left\{\mathbf{H}_0\left(\boldsymbol{\alpha}\right)\right\}^{\perp}$, and b) $\mathcal{S}_1 \cap \mathcal{S}_0^{\perp} \subset span\left\{\mathbf{H}_1\left(\boldsymbol{\alpha}_1\right)\right\}^{\perp}$. As a consequence: a) under $\mathcal{H}_0$, $\mathbf{U}_1^T\mathbf{y} = \mathbf{U}_1^T\mathbf{w}$, and b) under $\mathcal{H}_1 : \mathbf{U}_{-1}^T\mathbf{y} = \mathbf{U}_{-1}^T\mathbf{w}$. Finally, (14a) can be computed as

$$\begin{array}{l} P_{FA} = P\left(\chi_r^2\left(0\right) - \chi_{\nu}^2\left(\left\|\mathbf{U}_{-1}^T\mathbf{H}_0\left(\boldsymbol{\alpha}_0\right)\boldsymbol{\beta}_0/\sigma\right\|^2\right) > \lambda\right) \\ P_D = P\left(\chi_r^2\left(\left\|\mathbf{U}_1^T\mathbf{H}_1\left(\boldsymbol{\alpha}_1\right)\boldsymbol{\beta}_1/\sigma\right\|^2\right) - \chi_{\nu}^2\left(0\right) > \lambda\right) \end{array} \tag{14b}$$

where $\chi_l^2(\zeta)$ denotes a chi-square random variable with l degrees of freedom and non-centrality parameter ζ, whose probability density function (p.d.f.) is given by [3]

$$p\left(x; l, \zeta\right) = \frac{1}{2}e^{-(x+\zeta)/2}\left(\frac{x}{\zeta}\right)^{l/4-1/2} I_{l/2-1}\left(\sqrt{\zeta x}\right) \tag{14c}$$

where $I_{\upsilon}\left(z\right)$ is the modified Bessel function of the first kind. Thus, the ROC of the GLRT at high SNR depends simply on the distribution of the difference between two independent chi-square random

variables whose p.d.f.s are known (14c). As a result, the computation of (P_{FA}, P_D) (14b) can be obtained easily as follows. Firstly, one computes numerically the eigenvalues and the eigenvectors of $\mathbf{L}_0^{\perp} - \mathbf{L}_1^{\perp}$ which yields r, ν (13a-13b) and $\mathbf{U}_1$, $\mathbf{U}_{-1}$. Secondly, since the two chi-square random variables are independent, the p.d.f. of their difference is the correlation between the individual p.d.f.s, which can be computed numerically as well.
Let us notice that above we have implicitly considered the most general case where $\nu \geq 1$ and $r \geq 1$. However if $\mathcal{S}_1 \subset \mathcal{S}_0$, then $\mathcal{S}_0^{\perp} \subset \mathcal{S}_1^{\perp}$ and $\mathcal{S}_0^{\perp} \cap \mathcal{S}_1 = \varnothing$. In that particular case $\nu = 0$, leading to a simplified form of (P_{FA}, P_D)

$$\begin{aligned} P_{FA} &= P\left(\chi_r^2(0) > \lambda\right) \\ P_D &= P\left(\chi_r^2\left(\left\|\mathbf{U}_1^T \mathbf{H}_1(\boldsymbol{\alpha}_1)\boldsymbol{\beta}_1/\sigma\right\|^2\right) > \lambda\right). \end{aligned} \tag{15}$$

Likewise, if $\mathcal{S}_0 \subset \mathcal{S}_1$ then $r = 0$, which leads to another simplified form of (P_{FA}, P_D).

5. SIMULATION RESULTS

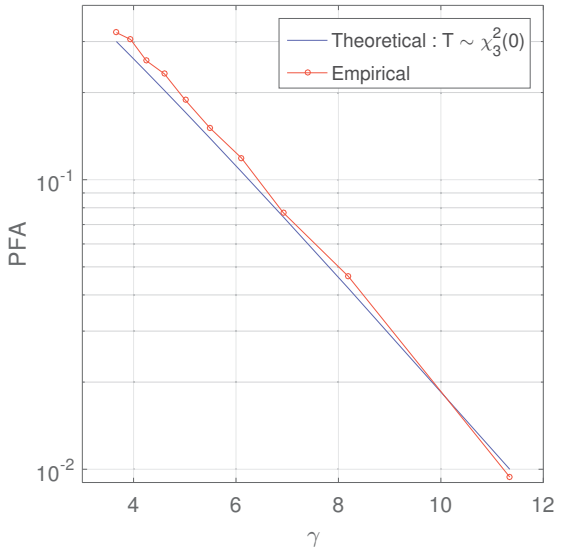

Fig. 1. Empirical and theoretical P_{FA} versus threshold γ.

Consider the problem of detecting a spurious sinusoidal signal at the output of a linear process[4]:

$$\begin{cases} \mathcal{H}_0 : y(n) = \lambda x(n-p) + w(n) \\ \mathcal{H}_1 : y(n) = \lambda x(n-p) + l_1 \cos(2\pi f n) + l_2 \sin(2\pi f n) + w(n) \end{cases}$$

where $n = 0, \ldots, N-1$. This problem is a particular case of (1) with the following unknown parameter vectors $\alpha_0 = p$, $\beta_0 = \lambda$, $\boldsymbol{\alpha}_1^T = (p, f)$, $\boldsymbol{\beta}_1^T = (\lambda, l_1, l_2)$ and mixing matrices $\mathbf{H}_0(\boldsymbol{\alpha}_0) \triangleq \mathbf{x}_p$, $\mathbf{H}_1(\boldsymbol{\alpha}_1) = [\mathbf{x}_p\ \mathbf{c}_f\ \mathbf{s}_f]$ where $\mathbf{x}_p^T = (\cdots, x(n-p), \cdots)$, $\mathbf{c}_f^T = (\cdots, \cos(2\pi f n), \cdots)$ and $\mathbf{s}_f^T = (\cdots, \sin(2\pi f n), \cdots)$. The

[4]The process is approximated by a delay p and an attenuation/amplification effect λ on a known signal x.

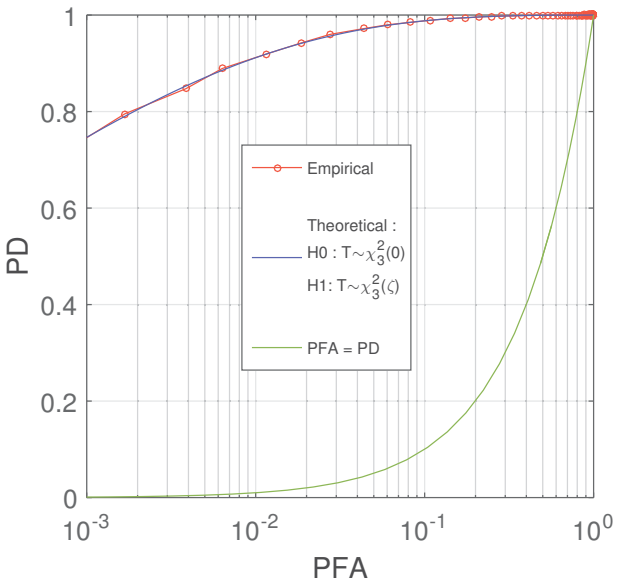

Fig. 2. Empirical and theoretical ROC curves.

CMLEs and the test statistic T for this particular case are straightforward to derive

$$\widehat{p}_0 = \underset{p}{\operatorname{argmax}} \left\{ (\mathbf{x}_p^T \mathbf{y})^2 / \|\mathbf{x}_p\|^2 \right\}$$

$$\begin{pmatrix} \widehat{p}_1 \\ \widehat{f} \end{pmatrix} = \underset{p,f}{\operatorname{argmax}} \left\{ \frac{(\mathbf{x}_p^T \mathbf{y})^2 - 4\mathbf{x}_p^T \mathbf{y}\, \mathrm{Re}\{I_{xy}(f)\} + 2\|\mathbf{x}_p\|^2 I_y(f) - 4\,\mathrm{Im}\{I_{xy}(f)\}^2}{\|\mathbf{x}_p\|^2 - 2I_x(f)} \right\}$$

$$T = \frac{(\mathbf{x}_{\widehat{p}_1}^T \mathbf{y})^2 - 4\mathbf{x}_{\widehat{p}_1}^T \mathbf{y}\, \mathrm{Re}\left\{I_{xy}(\widehat{f})\right\} + 2\|\mathbf{x}_{\widehat{p}_1}\|^2 I_y(\widehat{f}) - 4\,\mathrm{Im}\left\{I_{xy}(\widehat{f})\right\}^2}{\sigma^2(\|\mathbf{x}_{\widehat{p}_1}\|^2 - 2I_{\mathbf{x}}(\widehat{f}))} - \frac{(\mathbf{x}_{\widehat{p}_0}^T \mathbf{y})^2}{\sigma^2 \|\mathbf{x}_{\widehat{p}_0}\|^2} \tag{16}$$

where the term $I_{\mathbf{z}}(f) = \left((\mathbf{c}_f^T \mathbf{z})^2 + (\mathbf{s}_f^T \mathbf{z})^2\right)/N$ is the periodogram of $\mathbf{z}$, $\mathbf{z}^T = (z(0), \cdots, z(N-1))$, and $I_{xy}(f) = \left((\mathbf{c}_f^T \mathbf{x} - i\mathbf{s}_f^T \mathbf{x})(\mathbf{c}_f^T \mathbf{y} + i\mathbf{s}_f^T \mathbf{y})\right)/N$ is the cross-periodogram of $\mathbf{x}$ and $\mathbf{y}$. Note also that a large sample approximation $(N \to \infty)$ has been considered for the derivation of $\widehat{\alpha}_1^T = \left(\widehat{p}_1, \widehat{f}\right)$. Last but not least, since $span\{\mathbf{H}_0(\boldsymbol{\alpha}_0)\} \subset span\{\mathbf{H}_1(\boldsymbol{\alpha}_1)\}$, then $\mathcal{S}_1 \subset \mathcal{S}_0$ and we can resort to (15) to compute (P_{FA}, P_D). We consider the case where $x(n) \sim \mathcal{N}(0, \lambda^2/\sigma^2)$ is a Gaussian process, $N = 200$, $f = 4/100$, $p = 6$, leading to $r = 3$. Figures 1 and 2 compare the theoretical and empirical P_{FA} and ROC curves of the GLRT defined as $T \lessgtr \lambda$, where the estimation of α_0 and $\boldsymbol{\alpha}_1$ is carried out considering a grid search optimization where $\Delta p = 1e^{-5}s$ and $\Delta f = 1e^{-4}Hz$. The number of Monte-Carlo trials used for this experiment was 10^4. The SNRs under both hypotheses, i.e., $N\lambda^2/\sigma^2 = l_1^2 N/\sigma^2 = l_2^2 N/\sigma^2 = 13dB$, are high enough to assume that all CMLEs operate in their asymptotic region, which has been checked by comparing their MSEs with the associated CRBs. The almost perfect match between the empirical and theoretical results confirms the accuracy of analytical expression (14b) of the ROC of the High-SNR GLRT (12) derived in this paper.

6. REFERENCES

[1] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*, New York, Wiley, 1968

[2] J. Galy, E. Chaumette, P. Larzabal, "Joint Detection Estimation Problem of Monopulse Angle Measurement", IEEE Trans. on AES, 46(1): 397-413, 2010

[3] S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998

[4] L. L. Scharf, *Statistical signal processing : detection, estimation, and time series analysis*, Addison-Wesley, 2002

[5] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*, Springer Science+Business Media, LLC, 2008

[6] A. Lapidoth, *A Foundation in Digital Communication,* Cambridge University Press, 2009

[7] S.M. Kay, *Fundamentals of Statistical Signal Processing: estimation theory*, Prentice-Hall, 1993

[8] H.L. Van Trees, *Optimum Array Processing*, New-York, Wiley-Interscience, 2002

[9] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" IEEE Trans. on IT, 38(5): 1597-1602, 1992

[10] L. L. Scharf and B. Friedlander, "Matched subspace detectors," IEEE Trans. on SP, 42(8): 2146-2157, 1994

[11] P.J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press 2010

[12] F.C.Schweppe, "Sensor array data processing for multiple signal sources", IEEE Trans. on IT, 14: 294-305, 1968

[13] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," IEEE Trans. on SP, 38(10): 1783-1795, 1990.

[14] D.C. Rife, R.R. Boorstyn, "Single tone parameter estimation from discrete-time observations", IEEE Trans. on IT, 20(5): 591-598, 1974

[15] F. Athley, "Threshold region performance of MaximumLikelihood direction of arrival estimators", IEEE Trans. on SP, 53(4): 1359-1373, 2005

[16] C. D. Richmond, "Capon Algorithm Mean Squared Error Threshold SNR Prediction and Probability of Resolution", IEEE Trans. on SP, 53(8): 2748-2764, 2005

[17] A. Renaux, P. Forster, E. Chaumette, P. Larzabal, "On the High SNR CML Estimator Full Statistical Characterization", IEEE Trans. on SP, 54(12): 4840-4843, 2006