BAYESIAN DETECTION WITH THE POSTERIOR DISTRIBUTION OF THE LIKELIHOOD RATIO

I. Smith and A. Ferrari

Laboratoire H. Fizeau,
UNS, CNRS, OCA
Campus Valrose, F-06108 Nice cedex, France
isabelle.smith@unice.fr, ferrari@unice.fr

ABSTRACT

This paper focuses on simple versus composite hypothesis testing in Bayesian settings. The Posterior distribution of the Likelihood Ratio (PLR) provides an interesting alternative to the classical Bayes Factor (BF). First its general properties are studied and reveal its relationships with the BF, the Fractional BF and the Generalized Likelihood Ratio. Then the PLR is proved to be equal to a frequentist p-value for an invariant model and the corresponding invariant prior.

A practical implementation of the test can be performed using a simple Monte Carlo Markov Chain. Performances of the PLR used as a test are illustrated on extra-solar planet detection using direct imaging. Finally, the possibility to use different parametrizations of the test is illustrated by the study of their operating characteristics.

1. INTRODUCTION

Simple versus composite hypothesis testing is a general statistical issue in parametric modeling. It consists for a given dataset x in choosing among the hypotheses $\mathrm{H}_0 : \theta = \theta_0$ and $\mathrm{H}_1 : \theta \neq \theta_0$ with θ unknown. Under the Bayesian approach adopted here, the simple versus composite hypothesis test may be expressed in terms of the underlying priors: $\theta = \theta_0$ or $\theta \sim \pi(\theta)$ where $\pi(\theta) \neq \delta(\theta - \theta_0)$ is a given multivariate prior describing the uncertainty and constraints on the parameter of interest θ. We assume that the data model $p(x|\theta)$ has the same expression under H_0 and H_1 and does not depend on other parameters than θ.

This paper tackles this decision problem using the Posterior distribution of the Likelihood Ratio $p(x|\theta_0)p(x|\theta)^{-1}$ This approach has been initially proposed in [7] and studied, to our knowledge, only in [1, 2]. It is organized as follows:

- Section 2 introduces the frequentist and Bayesian hypothesis testing tools involved in the rest of the paper.
- Section 3 develops new theoretical properties of the Posterior distribution of the Likelihood Ratio. First its definition is given and a practical implementation procedure is proposed. Then, its general properties are derived, including its relationships with other classical tests. Finally, it is proved to be equal to a frequentist p-value for an invariant model and the corresponding invariant prior.
- Section 4 presents numerical simulations. First the PLR is applied to a realistic test case, then a Monte Carlo simulation using a simpler data model is used to characterize the frequentist properties of the test.

2. CLASSICAL HYPOTHESIS TESTING PROCEDURES

We first recall some classical notations used in frequentist decision theory that will be usefull in the sequel. Of particular interest ([16]) is the Likelihood Ratio (LR) defined by

$$\mathrm{LR}(x,\theta) = \frac{p(x|\theta_0)}{p(x|\theta)} \tag{1}$$

$\mathrm{LR}(x,\theta)$ evaluated at $\theta = \hat{\theta}_{\mathrm{ML}}(x)$ is the usual frequentist Generalized Likelihood Ratio (GLR):

$$\mathrm{GLR}(x) = \mathrm{LR}(x, \hat{\theta}_{\mathrm{ML}}(x)) = \frac{p(x|\theta_0)}{\max_\theta p(x|\theta)} \tag{2}$$

The GLR test consists in thresholding $\mathrm{GLR}(x)$.

We now recall basic notions of Bayesian testing under the hypothesis test

$$\mathrm{H}_0 : \theta = \theta_0 \qquad \mathrm{H}_1 : \theta \sim \pi(\theta) \tag{3}$$

Using the additional piece of prior information $\Pr(\mathrm{H}_0)$, the Posterior Odds Ratio

$$\mathrm{POR}(x) = \frac{\Pr(\mathrm{H}_0|x)}{\Pr(\mathrm{H}_1|x)} \tag{4}$$

minimizes the Bayesian risk under the 0-1 loss function associated to the estimation of the indicator function $I_{\theta_0}(\theta)$ [11]. Then, practical Bayesian hypothesis testing often consists in giving the POR, and thresholding it if a 0-1 decision is required: Reject H_0 if $\mathrm{POR}(x) \leq \zeta$. The POR equals the classical Bayes Factor (BF) [12]

$$\mathrm{BF}(x) = \frac{p(x|\mathrm{H}_0)}{p(x|\mathrm{H}_1)} = \frac{p(x|\theta_0)}{\int p(x|\theta)\pi(\theta)d\theta} \tag{5}$$

up to the multiplicative prior odds ratio $\Pr(\mathrm{H}_0)\Pr(\mathrm{H}_1)^{-1}$ which does not depend on x. The BF is also classicaly used for 0-1 decision and its threshold can be interpreted on its own grounds [12].

An important issue of the POR and the BF is that they are not uniquely defined if the prior $\pi(\theta)$ is improper, ie if $\int \pi(\theta)d\theta = \infty$. However, even though the prior is improper the posterior distribution $\pi^*(\theta|x)$ is in general proper. Taking advantage of this fact, Partial Bayes Factors have been proposed as alternatives to the Bayes Factor. Schematically, they assume that there exists a "minimal training sample" y chosen from the whole sample x such that $\pi^*(\theta|y)$ is proper. Then, a ("Partial") Bayes Factor can be uniquely defined from the rest of the data, using $\pi^*(\theta|y)$ as the proper prior. In particular, a Fractional Bayes Factor $\mathrm{FBF}(x,b)$, $b \in (0,1)$ has been proposed in [14]:

$$\mathrm{FBF}(x,b) = \frac{p(x|\theta=\theta_0)}{\int p(x|\theta)\pi(\theta)d\theta}\left(\frac{p(x|\theta=\theta_0)^b}{\int p(x|\theta)^b\pi(\theta)d\theta}\right)^{-1} \tag{6}$$

Note that $\mathrm{FBF}(x,0) = \mathrm{BF}(x)$.

 ISSN 2076-1465

Hypothesis test practitioners in general expect some "predata" information about the operating characteristics of the test. In frequentist procedures, the Probability of False Alarm of the test (PFA) and the Probability of good Detection (PD) may be used for calibration and/or performance assessment. Calibration under the PFA is in general attributed to the Neyman Pearson paradigm and opposed to the Bayesian one, but quoting [17] for example, "A Bayesian is calibrated if his probability statements have their asserted coverage in repeated experience". In particular, these quantities are also of interest in the (Bayesian) exoplanet detection frame presented in section 4.

Some "postdata" information is in general also expected. The most classical frequentist post-data measure about the significance of a decision of the form "Reject H_0 if $T(x) \leq \zeta$" is the p-value [13] defined by

$$p_{\mathrm{val}}\{T(x)\} = \Pr\{T(y) \leq T(x)|\mathrm{H}_0, x\} \quad (7)$$

It can be easily verified that the distribution of the p-value under H_0 is uniform in $(0,1)$ if $T(x)$ is a continuous random variable. Although p-values are still widely studied and generalized even in the Bayesian frame, the POR (and relatives) remain the standard Bayesian post-data evidence.

3. TESTING WITH THE POSTERIOR DISTRIBUTION OF THE LIKELIHOOD RATIO

3.1 Definition and motivation

A. Dempster proposed in 1974 [7] the first Bayesian test for (3) that relies on the Likelihood Ratio $\mathrm{LR}(x,\theta)$ defined in Eq. (1). For a given dataset x, the test consists in rejecting H_0 if the probability that the data are "much more" likely under $\theta \neq \theta_0$ than θ_0 is "high enough":

$$\text{Reject } \mathrm{H}_0 \text{ if } \quad \mathrm{PLR}(x,\zeta) > p \quad (8)$$

$$\text{with} \quad \mathrm{PLR}(x,\zeta) = \Pr\{\mathrm{LR}(x,\theta) \leq \zeta|x\} \quad (9)$$

The PLR (Posterior of the Likelihood Ratio) is simply the posterior cumulative distribution of the Likelihood Ratio. Fig. 4 illustrates a case where H_0 is rejected (left plot) and a case where H_0 is accepted (right plot).

Dempster's motivation in [7] is based on the role of the Likelihood Ratio in statistical inference, also emphasized by [16]. The PLR has only been studied as such by M. Aitkin in 1997 [1, 8] and in 2005 [2], on case studies mostly.

3.2 Implementation and optimization of the PLR

Note that implementation of this test may appear complicated. This complexity can be highly reduced using a Monte Carlo Markov Chain (MCMC) algorithm:

1. generate $\{\theta^{[j]} \sim \pi^*(\theta|x)\}_j$ using a MCMC algorithm
2. compute the chain $\{\mathrm{LR}(x,\theta^{[j]})\}$
3. compute the PLR (9) as the empirical cumulative distribution of the LR chain
4. if H_0 is rejected, use the chain $\{\theta^{[j]}\}$ for estimation

Interestingly enough, the PLR is a family of tests parametrized by two parameters: (ζ, p). Unlike detectors defined from a single threshold, it is possible to optimize the test inside this family. We propose to do it using the Receiver Operating Characteristics (ROC) curve, which displays $\mathrm{PD}(\zeta,p)$ as a function of $\mathrm{PFA}(\zeta,p)$. The principle is

1. compute $\mathrm{PFA}(\zeta,p)$ and $\mathrm{PD}(\zeta,p)$ $\forall(\zeta,p)$
2. fix a PFA_0 and obtain $\{(\zeta,p) : \mathrm{PFA}(\zeta,p) = \mathrm{PFA}_0\}$
   - choose from this set $(\zeta^*(\mathrm{PFA}_0), p^*(\mathrm{PFA}_0))$ that maximizes $\mathrm{PD}(\zeta,p)$.

This procedure is performed numerically using the proposed MCMC algorithm. For each hypothesis, a matrix of size $N_I \times N_J$ containing on line i the Markov chain obtained from dataset i is built. Each matrix is reordered by sorting in increasing order each line (cumulative posterior distribution for a given dataset) and then each column (cumulative frequentist distribution for a given ζ). For a number N_I of datasets and a chain length N_J both sufficiently large, approximate $\mathrm{PFA}(\zeta,p)$ and $\mathrm{PD}(\zeta,p)$ can be read from each matrix. For example, for $\zeta(i,j) = \mathrm{LR}(i,j)$ (i.e. the $(i,j)^{\mathrm{th}}$ component of the matrix) and $p(i,j) = j/N_J$, the PFA is approximately given by $\mathrm{PFA}(\zeta(i,j), p(i,j)) = i/N_I$. Therefore, the approximate optimal parameters can be easily obtained from the two matrices.

Finally, note that an analog procedure can be used for the FBF: the FBF is computed from the Markov chain $\{\mathrm{LR}(x,\theta^{[j]})\}$ using an importance sampling procedure and the test is optimize with respect to the threshold and b.

3.3 General properties of the PLR

The first result illustrates the deep connections existing between the FBF and the posterior distribution of the LR. The proof is straightforward but no reference has been found.

Proposition 1 *If the prior π is proper, the FBF for the simple versus composite hypothesis test equals the fractional posterior moment of the LR:*

$$\forall b,\ FBF(x,b) = \mathsf{E}[LR(x,\theta)^{1-b}|x] \quad (10)$$

This result shows that the BF ($b=0$) is the posterior mean of the LR. The BF can then be interpreted as the mean square error estimator of $\mathrm{LR}(x,\theta)$. A standard uncertainty on this inference would be given by the the posterior standard deviation of the LR:

$$\begin{aligned} \widehat{\mathrm{LR}}(x,\theta) &= \mathrm{BF}(x) \pm \mathrm{std}(\mathrm{LR}(x,\theta)) \quad (11) \\ &= \mathrm{BF}(x) \pm \sqrt{\mathrm{FBF}(x,-1) - \mathrm{BF}(x)^2} \end{aligned}$$

However, although this uncertainty is natural for an estimation of LR, it is not relevant for its thresholding. It will be illustrated in the simulations.

The next result gives two general properties about the posterior density $p_{\mathrm{LR}|x}$ of the LR. These results assume that the hypotheses are nested: $\theta_0 \in \mathrm{Sup}(\pi)$. This case is of particular interest in many practical applications, such as the problem addressed in the last section of the paper.

Proposition 2 *The posterior density of the LR of a nested simple versus composite hypothesis verifies:*

- *The minimum of its support is $GLR(x)$:*

$$\min_{\zeta}\{\zeta : p_{LR|x}(\zeta|x) > 0\} = GLR(x) \quad (12)$$

- *Under regularity assumptions that get stronger as L (the length of θ) increases, the function $\zeta \to p_{LR|x}(\zeta|x)$ diverges for $\zeta \to GLR(x)^+$.*

Illustrations of these results are shown in figure 3.

Whereas the first property is a direct consequence of (2), the proof of the second is much more delicate. A usual transform to infer the distribution of $\mathrm{LR}(x,\theta)$ is

$$\phi : \theta \to (\mathrm{LR}(x,\theta), \check{\theta}_1) \text{ with } \check{\theta}_1 = (\theta_2, .., \theta_L)$$

The usual variables transformation gives:

$$p_{\mathrm{LR}(x,\theta),\check{\theta}_1}(\zeta, \check{u}_1) = \sum_k \frac{\pi^*(v^k|x)}{|J(v^k)|},\ J(\theta) = \frac{\partial \mathrm{LR}(x,\theta)}{\partial \theta_1}$$

where the v^k are all the solutions of $\phi(v^k) = (\zeta, \check{u}_1)$. For $L = 1$ (θ is scalar), the result is straightforward if $\theta \to \mathrm{LR}(x, \theta)$ is continuously differentiable. For $L > 1$, $L-1$ integrations are required to marginalize out $\check{\theta}_1$. We show [19] that if locally there exists $\alpha > L$ and $(\alpha_1, .., \alpha_L) \in \mathbb{R}^L_{+*}$ such that for all θ close enough to $\hat{\theta}_{\mathrm{ML}}(x)$

$$\mathrm{GLR}(x) < \mathrm{LR}(x, \theta) \leq \mathrm{GLR}(x) + \sum_{\ell=1}^{L} \alpha_\ell (\theta - \hat{\theta}_{\mathrm{ML}}(x))_\ell^\alpha$$

then $\epsilon^{-1} \Pr(\mathrm{GLR}(x) < \mathrm{LR}(x, \theta) \leq \mathrm{GLR}(x) + \epsilon | x) \to \infty$ when $\epsilon \to 0$.

3.4 PLR under a general invariant case

Dempster noticed that *when testing the mean of a normal distribution with a constant prior density*, the PLR is equal to 1 minus the p-value introduced in Eq. (7). Aitkin generalized this result with an additional nuisance parameter in the mean.

The next theorem states that the result of Dempster is true in some general invariant cases. Invariance is a central framework to unify the frequentist, Fisherian pivotal and Bayesian paradigms [9, 10]. It relies on two assumptions on the model: the likelihood belongs to an invariant distribution family and the prior distribution is invariant under the transformation group defining the likelihood family.

The next definition specifies the invariance of a family of densities under a group of transformations.

Definition 1 *A family $\mathcal{F}_\Theta = \{f(.|\theta), \theta \in \Theta\}$ of densities wrt a measure μ on $\mathcal{X}$ is said to be* invariant under the transformation group $\mathcal{G}$ *if, for every $g \in \mathcal{G}$ there exists a unique $\theta^* \in \Theta$ such that if the random variable X has density $f(.|\theta)$, $Y = g(X)$ has density $f(.|\theta^*) \in \mathcal{F}_\Theta$. We define $\bar{\mathcal{G}}$ as the set of all functions $\bar{g}$ induced by the group $\mathcal{G}$ and defined as $\theta^* = \bar{g}(\theta)$. $\bar{\mathcal{G}}$ is a group.*

In Bayesian models, it is often required to define a non-informative prior related to some specific property [4]. In particular, model invariance can be accounted for using the right Haar prior [4, 9]. It is in general improper.

Definition 2 (Haar measure) *A right invariant Haar measure, to be denoted H^r, under a group of transformations $\mathcal{G}$ is a measure which, for all measurable functions κ on $\mathcal{G}$ and for all $g_0 \in \mathcal{G}$ satisfies*

$$\int_{\mathcal{G}} \kappa(g) H^r(dg) = \int_{\mathcal{G}} \kappa(g g_0) H^r(dg)$$

A left invariant Haar measure H^l is defined replacing gg_0 by $g_0 g$.

The group $\mathcal{G}$ is related to the parameters space Θ assuming that the function $\phi_\theta : \mathcal{G} \to \Theta$, $\phi_\theta(g) = g(\theta)$ is bijective for all $g \in \mathcal{G}$. The measure induced by H^r is then defined for all $A \subset \Theta$ by $\Pr(\theta \in A) = H^r(\phi_a^{-1} A)$. The measure induced by H^r can be defined in the same way on $\mathcal{X}$ if the function ϕ_x: $\mathcal{G} \to \mathcal{X}$, $\phi_x(g) = g(x)$ is bijective.

The main contribution of this paper is an expression of the PLR, under invariance assumptions, as a frequentist integral involving the modulus of the group of transformations.

Definition 3 *The modulus of $\mathcal{G}$ is the function Δ defined on $\mathcal{G}$ to $(0, \infty)$ which, for all measurable functions κ on $\mathcal{G}$ satisfies*

$$\int_{\mathcal{G}} \kappa(g g_0^{-1}) H^l(dg) = \Delta(g_0) \int_{\mathcal{G}} \kappa(g) H^l(dg)$$

Theorem 1 *Let $\mathcal{F}_\Theta = \{f(.|\theta), \theta \in \Theta\}$ be a family of probability densities wrt a measure μ on $\mathcal{X}$. Assume that:*

1. *$\mathcal{F}_\Theta$ is invariant under the group $\mathcal{G}$.*
2. *ϕ_θ and ϕ_x are bijective. $\mathcal{X}$ and Θ are isomorphic.*
3. *The prior measure on Θ is the measure induced by H^r from ϕ_θ.*
4. *The measure μ on $\mathcal{X}$ is the measure induced by H^r from ϕ_x.*

Then the PLR defined in Eq. (9) can be reexpressed as the frequentist integral:

$$PLR(\zeta, x) = \Pr\left\{ \frac{f(x|\theta_0)}{\Delta(\phi_c^{-1}(x))} \leq \zeta \frac{f(y|\theta_0)}{\Delta(\phi_c^{-1}(y))} | H_0, x \right\} \quad (13)$$

for all $c \in \mathcal{X}$.

The proof of this theorem, which is well beyond the scope of this article, is available in [19].

Setting $\zeta = 1$ in (13) leads directly to the following corollary where the second equality follows from the uniform distribution of the p-value.

Corollary 1 *Under the same hypothesis as theorem 1*

$$PLR(1, x) = 1 - p_{val} \left\{ \frac{f(x|\theta_0)}{\Delta(\phi_c^{-1}(x))} \right\} \quad (14)$$

Consequently, the PFA of the test (8) for $\zeta = 1$ equals $1 - p$.

The assumption of an isomorphism between $\mathcal{X}$ and Θ is rather restrictive. This constraint is relaxed considering a family of probability densities where the invariance is defined on $\tau(X) \in \mathcal{T}$, a sufficient statistics for $\theta \in \Theta$, such that $\mathcal{T}$, $\mathcal{G}$, Θ and $\bar{\mathcal{G}}$ are all isomorphic. In this case the proof of theorem 1 can be extended. It leads to (13,14) where the statistics used in the p-value is now $f_s(\tau|\theta_0)/\Delta(\phi_c^{-1}(\tau))$ where $f_s(\tau|\theta_0)$ is the marginal distribution of $\tau(X)$ under H_0.

Many studies try to conciliate or compare frequentist and Bayesian hypotheses tests procedures. For the simple vs composite hypothesis test see for example [3, 11]. This result contributes to this issue, but in the frame of measure of evidence using the LR [16, 7] and not measure of accuracy of a set estimation I_{θ_0} [11].

It also has some practical interest. For example, for $\zeta = 1$ the threshold $1 - p$ equals the significance level of the test.

4. APPLICATION TO EXOPLANET DETECTION WITH DIRECT IMAGING

The detection procedure presented in section 3 is realistically applied to the detection of exoplanets from direct imaging using the future VLT instrument SPHERE [5].

4.1 Statistical model for exoplanet detection in direct imaging

A hierarchical Bayesian model precisely related to our context has been developed in [18] and is summed up here. The θ vector of the hypothesis test (3) refers to the exoplanet intensity in the different channels.

- The dataset is made of K successive sets of L images associated to different spectral bands, where each image is a $M \times 1$ vector $i_\ell(k)$. The $x_k^t = (i_1(k)^t, .., i_L(k)^t)$ are assumed to be conditionally independent and described by:

$$x_k | \mu, \Sigma, \theta \sim \mathcal{N}_{LM}(A_k \theta + \mu, \Sigma) \quad (15)$$

Matrix A_k contains source profiles $p_\ell(k)$ assumed known.

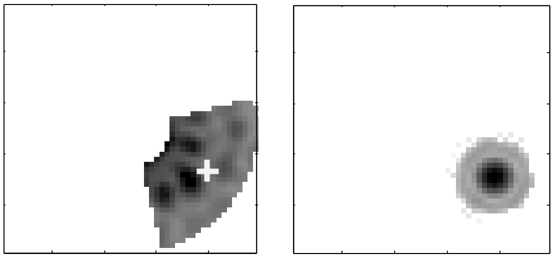

Figure 1: Simulated data from CAOS-SPHERE with a contrast of 10^6 between the star and the planet. Left: data $x_2(20)^{0.2}$. Right: source response $p_2(20)^{0.2}$.

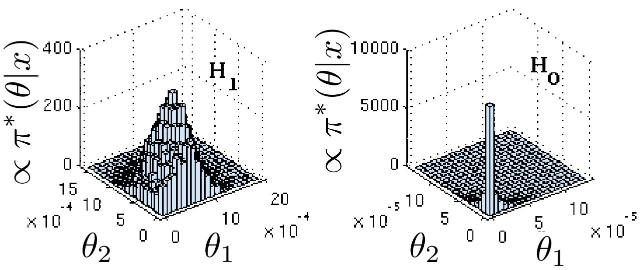

Figure 2: Histograms of two Markov chains $\{\theta^{[j]} \sim \pi^*(\theta|x)\}$ resulting from the data with and without a planet.

- Due to the high dynamics involved in possible low signal to noise ratio cases, and due to the inherent wide range of sources intensities, the distribution of θ has to spread out on several order of magnitudes. It is therefore natural to assume that $(\ln\theta_1, \dots, \ln\theta_L)^t$ is jointly gaussian. This so-called multivariate log-normal distribution describes high dynamics signals, has a positive suppor and is proper:

$$\theta|m, B \sim \log\mathcal{N}(m, B) \tag{16}$$

Conjugate priors (Normal - inverse Wishart in both cases) are assumed for the unknown parameters. This first level likelihood is marginalized and leads to an explicit form of $\pi^*(\theta|x)$ where $x = \{x_k\}_{k=1,..,K}$. The Markov chain $\{\theta^{[j]} \sim \pi^*(\theta|x)\}_j$ necessary to compute the test statistics is obtained from a slice sampling method [15].

4.2 Application of the detection procedure on a realistic dataset

The simulation of realistic astrophysical datasets is performed by the dedicated physical step-by-step Software Package SPHERE [5] developed and used within the CAOS environment [6]. A dataset x is simulated under H_1 with a luminosity contrast of 10^6 between the star and the exoplanet (corresponding to an intensity θ_{H1}), and another under H_0, obtained from an area adjacent to the one under H_1. The data under H_1, of size $(K, L, M) = (20, 2, 425)$ are illustrated on figure 1. Note that it is impossible to simulate many datasets.

The detection procedure described in section 3 and used with the realistic statistical model summarized in section 4.1 is finally applied to these two datasets. The hyperparameters are chosen simply ($\nu = 2M$, $\Sigma_0 = \widehat{\sigma^2} I_{2M}$) or unfavourable ($m_0 = \ln(1000\theta_{H1})$). $N_I = 10^5$ samples are computed for each chain $\{\theta^{[j]}\}$.

Figure 2 shows the histograms of the Markov chains resulting from these two cases. Under the H_1 case, the Bayes Factor (5) seems to indicate with no ambiguity a

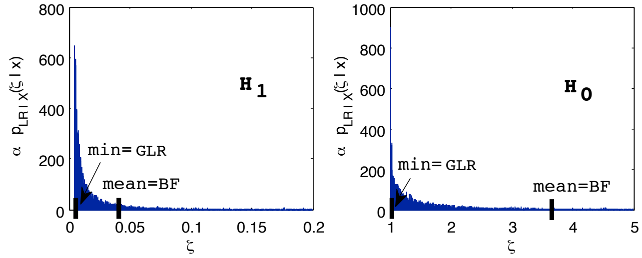

Figure 3: Histograms of the $\{\mathrm{LR}(x, \theta^{[j]})\}$ chains, computed from the chains $\{\theta^{[j]}\}$ shown in figure 2. The GLR and the BF are indicated.

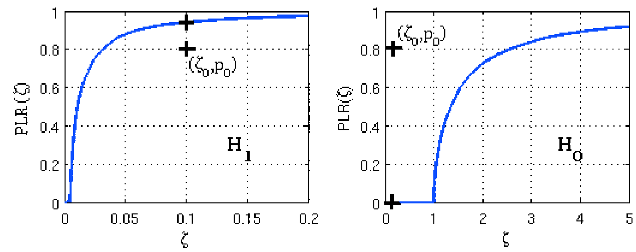

Figure 4: A posteriori empirical cumulative distributions of LR, displayed from the chains $\{\mathrm{LR}(x, \theta^{[j]})\}$ shown in Fig. 3.

detection: $\mathrm{BF} = 0.04 < \zeta_0$ for $\zeta_0 = 0.1$[1]. The measure $\mathrm{PLR}(x, \zeta_0) = 0.94 > 0.8$ confirms the absence of ambiguity of the BF result. Similarly, in the H_0 case, BF = 3.7 indicates again with no ambiguity that there is no exoplanet. This is confirmed by the quantile $\mathrm{PLR}(x, \zeta_0) = 0$. For a more complete information, the empirical posterior distributions of $\mathrm{LR}(x, \theta^{[n]})$ are presented on figure 3 and figure 4. They also illustrate the properties given in section 3.

Finally, estimation can be performed for the data where a signal has been detected (ie data simulated under H_1). The posterior distribution is shown on figure 2 (left). The signal is estimated by the posterior mean and its uncertainty by the posterior standard deviation: $\theta = (6.2 \pm 2.8\ ; 4.6 \pm 2.6).10^{-4}$ for a true $\theta_{H1} = (8\ ; 0.5).10^{-5}$.

4.3 Comparison with a frequentist GLR test

The proposed procedure is compared to a Generalized Likelihood Ratio Test. The GLR as defined in (2) on the marginalized density $p(x|\theta)$ being analytically intractable, the GLR has been derived directly on the first level likelihood (15) assuming that the covariance matrix is proportional to identity: $\Sigma = \sigma^2 I_{LM}$. This last assumption is required to avoid a complex constrained optimization and to obtain a closed form expression of the test. Then,

$$\mathrm{GLR}_2(x) = \frac{\max_{\mu,\sigma}\{\prod_k p(x_k|\mu, \sigma^2 I_{LM}, \theta = 0)\}}{\max_{\mu,\sigma,\theta}\{\prod_k p(x_k|\mu, \sigma^2 I_{LM}, \theta)\}} \tag{17}$$

The analytical maximization of the likelihood under H_1 for $L > 1$ generalizes a computation in [20] where $L = 1$, and leads to:

$$\mathrm{GLR}_2(x) = \left(\frac{\widehat{\sigma}^2_{H1}}{\widehat{\sigma}^2_{H0}}\right)^{\frac{KLM}{2}} \tag{18}$$

$$\text{where} \quad \widehat{\sigma}^2_{Hi} = \frac{\sum_k \|x_k - A_k\widehat{\theta}_{Hi} - \widehat{\mu}_{Hi}\|^2}{KLM}$$

[1] The uncertainty defined in equation (11) gives "LR = 0.04 ± 0.34". As mentioned, this measure of uncertainty is not relevant for detection.

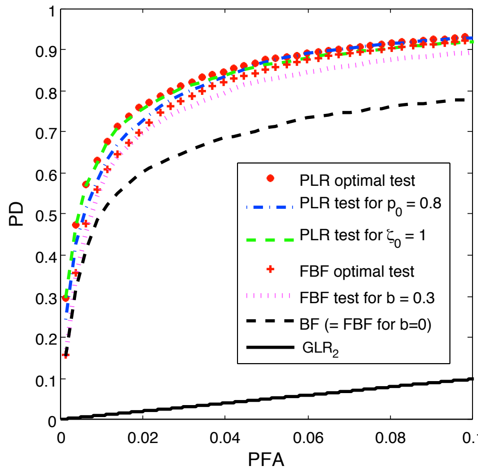

Figure 5: ROC curves of the PLR, the FBF and the GLR_2.

where $\widehat{\theta}_{H0} = 0$ and $(\widehat{\mu}^t_{H1}, \widehat{\theta}^t_{H1})$ and $\widehat{\mu}_{H0}$ minimize least square criteria obtained from the model (15).

Note that since the hypotheses are nested, contrary to $\mathrm{LR}(x, \theta^{[n]})$ $\mathrm{GLR}_2(x)$ has always a value inferior or equal to 1. Here, $\ln(\mathrm{GLR}_2(x)) = -4350$ for the data simulated under H_1 and $\ln(\mathrm{GLR}_2(x)) = -1300$ under H_0. Since it is not numerically possible to realistically simulate a large number of datasets, it is impossible to relate numerically the threshold of the GLR test to its PFA. The model (15) is not identically distributed, so the classical results on the asymptotic distribution of the GLR neither apply. It is consequently difficult to choose the threshold ζ.

In any case, the values of $\mathrm{GLR}_2(x)$ applied to areas closed but distinct from the precedent cases indicate that the $\mathrm{GLR}_2(x)$ discriminates with difficulty H_0 and H_1.

4.4 Illustration of the PLR optimization procedure

Other interesting properties of the PLR are now illustrated on an astrophysical context totally similar to the previous one, but the data are now simulated from the statistical model and not the physical one, so that a long run performance analysis can be performed. The data are simulated from the marginalized likelihood presented in [18] for $KLM = 80$. The data under H_1 are characterized by a fixed $\theta = \theta_{H1}$.

Fig. 5 illustrates the ROC curves obtained for some intuitive parametrizations ($\zeta = 1$ etc) and the optimal ones. The optimal ROC curves are computed using the procedure discribed in section 3.2. We note that:

- The classical Bayes Factor is uniformly less performant than the other FBF and the PLR. For PFA = 0.1, the performances of the PLR overpass the ones of the Bayes Factor by 15%.
- The tests with fixed parametrization have performances very close to the optimal ones. It strenghtens their use.
- The bad performances of the GLR_2 test (18) where Σ was wrongly assumed to be proportionnal to identity are confirmed here: it is equivalent to a heads or tails test.

REFERENCES

[1] M. Aitkin. The calibration of p-values, posterior Bayes factors and the AIC from the posterior distribution of the likelihood. *Statistics and Computing*, 1997.

[2] M. Aitkin, R. J. Boys, and T. Chadwick. Bayesian point null hypothesis testing via the posterior likelihood ratio. *Statistics and Computing*, 2005.

[3] J. Berger and T. Selke. Testing a point null hypothesis: the irreconcilability of P values and evidence (with discussion). *Journal of the American Statistical Association*, 1987.

[4] J. O. Berger. *Statistical decision theory and Bayesian analysis.* Springer-Verlag, 2nd edition, 1985.

[5] M. Carbillet, A. Boccaletti, C. Thalmann, et al. The software package SPHERE: a numerical tool for end-to-end simulations of the VLT instrument SPHERE. In N. Hubin, C. E. Max, and P. L. Wizinowich, editors, *Adaptive Optics Systems.* SPIE, 2008.

[6] M. Carbillet et al. CAOS - a numerical simulation tool for astronomical adaptive optics (and beyond). In *Advancements in Adaptive Optics*, SPIE Proceedings, 2004.

[7] A. P. Dempster. The direct use of likelihood for significance testing. In *Proceedings of Conference on Foundational Questions in Statistical Inference*, 1974.

[8] A. P. Dempster. Commentary on the paper by Murray Aitkin, and on discussion by Mervyn Stone. *Statistics and Computing*, 1997.

[9] M. Eaton. *Group invariance applications in Statistics.* Regional Conference Series in Probability and Statistics, 1989.

[10] M. Eaton and W. Sudderth. Consistency and strong inconsistency of group-invariant predictive inferences. *Bernoulli*, 1999.

[11] J. Hwang, G. Casella, C. Robert, M. Wells, and H. Farrell. Estimation of accuracy in testing. *Annals of Statistics*, 1992.

[12] R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 1995.

[13] E. L. Lehmann and J. P. Romano. *Testing statistical hypotheses.* Springer, 3rd edition, 2005.

[14] A. O'Hagan. Fractional Bayes factors for model comparison. *Journal of the Royal Statistical Society*, 1995.

[15] C. Robert and G. Casella. *Monte Carlo Statistical Methods.* Springer-Verlag, 1999.

[16] R. Royall *Statistical evidence: a likelihood paradigm.* Chapman and Hall / CRC Press, 1997.

[17] D. Rubin. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics*, 1984.

[18] I. Smith and A. Ferrari. Detection from a multi-channel sensor using a hierarchical Bayesian model. *ICASSP*, 2009.

[19] I. Smith and A. Ferrari. Posterior distribution of the likelihood ratio for hypothesis testing: new theoretical results. Technical report, Nice University, 2010.

[20] I. Smith, A. Ferrari, and M. Carbillet. Detection of a moving source in speckle noise. Application to exoplanet detection. *IEEE Transactions on Signal Processing*, 57(3), 2009.