DETECTION OF EXTRA SOLAR PLANETS USING PARAMETRIC MODELING

A. Ferrari[†,⋆], *J.Y. Tourneret*[‡] *and F.-X. Schmider*[⋆]

[⋆] UMR 6525 Astrophysique, Université de Nice-Sophia-Antipolis, F-06108 Nice Cedex 2
[†] I3S UPRES-A 6070, Université de Nice-Sophia-Antipolis, F-06560 Valbonne
[‡]ENSEEIHT/GAPSE, 2 Rue Camichel, 31071 Toulouse, France
ferrari@unice.fr, tournere@len7.enseeiht.fr, fxs@irisalfa.unice.fr

ABSTRACT

We present an algorithm for the detection of extra-solar planets by occultation on the satellite COROT. Under high flux assumption, the signal is modeled as an autoregressive process having equal mean and variance. A transit of a planet in front of a star will produce an abrupt jump in the mean/variance of the process. The Neyman-Pearson detector is derived when the abrupt change parameters are known. The theoretical distribution of the test statistic is obtained allowing the computation of the ROC curves. The generalized likelihood ratio detector is then studied for the practical case were the change parameters are unknown. This detector requires the maximum likelihood estimates of the parameters. ROC curves are then determined using computer simulations.

1. INTRODUCTION

The small French satellite COROT (CNES) is dedicated to stellar asteroseismology and extra-solar planetary detection. The further is based on the possible transit of a planet in front of a star, which will produce a decrease of the photometric signal from the star, proportional to the ratio of the planet to the star surface, during the time of the transit. The photometric signal is recorded on a CCD camera. This communication addresses the planet detection problem.

Section 2 presents the signal model. A simplified model is then developed under the high flux assumption. For this model, the Neyman-Pearson detector for the abrupt change (AC) detection is studied in section 3. The exact distribution of the test statistic is obtained, allowing computation of ROC curves. Section 4 deals with the practical application where the AC parameters are unknown. The generalized likelihood ratio detector is derived. Consequently maximum likelihood estimators of the AC parameters are determined.

2. SIGNAL MODEL DERIVATION

We assume that the signal is dominated by the photon noise i.e. the read-out noise and the thermal noise for the electronic are negligible. In the ideal sensor case, the problem consists of detecting a jump in the mean λ of an iid signal x_n having a Poisson distribution.

Denote as r the step instant:

$$\begin{cases} \forall n \in S_0 = [1, r] & : \quad \lambda = \lambda_0, \\ \forall n \in S_1 = [r+1, N] & : \quad \lambda = \lambda_1 \ (< \lambda_0). \end{cases} \tag{1}$$

The AC detection problem is:

$$\begin{cases} H_0 & : \quad S_1 = \emptyset \ \text{(no jump)}, \\ H_1 & : \quad S_1 \neq \emptyset \ \text{(jump)}. \end{cases} \tag{2}$$

The Neyman-Pearson detector (NPD) can be easily computed for this problem and yields:

$$H_0 \text{ rejected if } \frac{1}{N-r} \sum_{n=r+1}^{N} x_n < \zeta. \tag{3}$$

The problem is more complicated when the detector is not perfect, and in particular if its sensitivity changes with time. In this case, the parameter λ is replaced by a random variable λ_n having a mean λ, subjected to a jump. The distribution of $X = \{x_1, \ldots, x_N\}$ conditioned on $\Lambda = \{\lambda_1, \ldots, \lambda_N\}$ is an iid sequence with Poisson distribution of parameter Λ. The unconditioned distribution of X is:

$$p(X) = \int_\Lambda p(X/\Lambda) p(\Lambda) d\Lambda \tag{4}$$

$$= \int_\Lambda \prod_{n=1}^{N} \frac{e^{-\lambda_n} \lambda_n^{x_n}}{x_n!} p(\Lambda) d\Lambda, \tag{5}$$

referred as the Poisson Mandel transform of $p(\Lambda)$, [7].

An a priori distribution for Λ is very difficult to choose. Moreover, the derivation of the test for (2) requires the computation of the resulting distribution of X using (5) which is generally a very complicated task. For these reasons, a simpler model for X is proposed using two realistic assumptions:

$$var[\lambda_n] \ll E[\lambda_n], \ E[\lambda_n] = \lambda \gg 1. \tag{6}$$

The first assumption conveys the fact that the variations of the sensor are small. The second traduces a high flux assumption. A fundamental consequence is that under mild assumptions on the distribution of Λ, X is approximately Gaussian distributed. This property can be proved by relating the characteristic function of X to the characteristic function of Λ using (5) (see [7]). For example when Λ is Gaussian distributed with mean m, the characteristic function of X tends to the Gaussian characteristic function when m tends to infinity, [3].

Using conditional expectation, the mean and variance of x_n can be determined:

$$E[x_n] = E[E[x_n/\lambda_n]] = E[\lambda_n] = \lambda. \tag{7}$$

$$\begin{aligned} var[x_n] &= E[x_n^2] - \lambda^2 = E[E[x_n^2/\lambda_n]] - \lambda^2 \\ &= E[\lambda_n + \lambda_n^2] - \lambda^2 = \lambda + var[\lambda_n] \approx \lambda, \end{aligned} \tag{8}$$

the last approximation coming from the above assumption on the sensor.

Consequently, the distribution of X can be approximated by a corelated Gaussian distribution having same variance and mean λ. This paper proposes to model x_n as an autoregressive process. Classical justifications for this model in the stationary Gaussian context can be found for example in [5]. Assume that x_n is a p^{th} order autoregressive process with equal mean and variance λ:

$$x_n = -\sum_{k=1}^{p} a_k x_{n-k} + \lambda(1 + \sum_{k=1}^{p} a_k) + e_n, \qquad (9)$$

where e_n is an iid zero mean Gaussian sequence.

It is important to note that in the perfect sensor case, the high flux assumption implies that the Poisson distribution tends to a Normal distribution with same mean and variance, [2]. This particular case corresponds to $\forall k$, $a_k = 0$ in (9). A major effect of the detector imperfections will be to correlate the signal measurements.

The variance of e_n in model (9) is such that $var[x_n] = \lambda$. In order to take into account its dependence toward the a_k and λ, it will be denoted $\sigma_e^2(a, \lambda)$ in the sequel. An analytic expression of $\sigma_e^2(a, \lambda)$ is very difficult to obtain. A formal expression can be obtained by rewriting the Yule Walker equations, [6], as a linear system where the unknowns are the signal covariances, $\mathbf{c} = (\lambda, c(1), \ldots, c(p))^t$. This leads to:

$$(A_1 + A_2)\mathbf{c} = (\sigma_e^2(a, \lambda), 0, \ldots, 0)^t, \qquad (10)$$

with,

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_1 & 1 & 0 & \cdots & 0 \\ a_2 & a_1 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ a_p & a_{p-1} & \cdots & a_1 & 1 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & a_1 & \cdots & a_{p-1} & a_p \\ 0 & a_2 & \cdots & a_p & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & a_p & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

We then obtain:

$$\sigma_e^2(a, \lambda) = \lambda \sigma_e^2(a, 1), \qquad (11)$$

where $\sigma_e^2(a, 1)$ is the inverse of the upper left element of the matrix $(A_1 + A_2)^{-1}$.

3. THE NEYMAN PEARSON DETECTOR

This section derives the Neyman Pearson Detector (NPD) for problem (2), when x_n is an AR process defined in (9). AC detection and estimation for linear models has been studied for long time (see [1, 4] and references therein for an overview). The new contribution here is the development of a detection scheme in the particular case where the jumps occurs on the mean/variance of the signal. The study is restricted to off-line change detection algorithms, [1].

In [8], the author studies the case of a non zero mean autoregressive signal multiplied by a sigmoidal function modeling a jump. However, even if this model represents a jump in the mean and the power of an AR process, it can be easily check that it cannot handle the case where mean and variance are equal. This constraint, as it will be shown below simplifies substantially the test statistic.

Denote $\mathcal{L}(X/H_1)$ the log-likelihood function of $\{x_{p+1}, \ldots, x_N\}$ conditioned on $\{x_1, \ldots, x_p\}$ under hypothesis H_1. After dropping the constant terms, we have:

$$\begin{aligned}\mathcal{L}(X/H_1) = &-(N-p)\log \sigma_e^2(a,1) \\ &- (r-p)\log \lambda_0 - (N-r)\log \lambda_1 \\ &- \frac{1}{\sigma_e^2(a,1)}\left(\frac{1}{\lambda_0}\sum_{n=p+1}^{r} e_{n,0}^2 + \frac{1}{\lambda_1}\sum_{n=r+1}^{N} e_{n,1}^2\right), \end{aligned} \qquad (12)$$

where:

$$e_{n,i} = x_n + \sum_{k=1}^{p} a_k x_{n-k} - \lambda_i(1 + \sum_{k=1}^{p} a_k). \qquad (13)$$

The log likelihood function under hypothesis H_0 is readily obtained from (12):

$$\begin{aligned}\mathcal{L}(X/H_0) = &-(N-p)\log \sigma_e^2(a,1) - (N-p)\log \lambda_0 \\ &- \frac{1}{\sigma_e^2(a,1)\lambda_0}\sum_{n=p+1}^{N} e_{n,0}^2. \end{aligned} \qquad (14)$$

After simplification and using the hypothesis $\lambda_1 < \lambda_0$, the Neyman-Pearson test reduces to:

$$T = \frac{1}{N-r}\sum_{n=r+1}^{N}(x_n + \sum_{k=1}^{p} a_k x_{n-k})^2 \underset{H_1}{\overset{H_0}{\gtrless}} \zeta. \qquad (15)$$

It is important to note that when a jump occurs independently on the mean and the variance of the process, [8], the test statistic T is the difference of two positive definite quadratic forms that is generally indefinite. The exact distribution of T is very difficult to study in this case and has been approximated by a Gaussian distribution in [8]. In this paper, because of the constraint (11), T reduces to a single positive definite quadratic form.

Another noteworthy point is that T is not function of λ_0 and λ_1. Moreover, in the iid case, T reduces to the estimated power of x_n whereas in the Poisson case T is the estimated mean (3).

Under hypothesis H_i, T is the sum of the square of $N - r$ independent Gaussian random variables with:

- mean $\lambda_i(1 + \sum_{k=1}^{p} a_k)/\sqrt{N-r}$,
- variance $\sigma_e^2(a, \lambda_i)/(N-r)$.

Consequently $(N-r)T/\sigma_e^2(a, \lambda_i)$ is distributed as a non central χ^2 distribution with $N - r$ degrees of freedom and non-centrality parameter:

$$\mu_i = \frac{(N-r)(1 + \sum_{k=1}^{p} a_k)}{\sigma_e(a,1)}. \qquad (16)$$

The distribution of T is:

$$p_T(t/H_i) = \frac{N-r}{\sigma_e^2(a, \lambda_i)} f(\frac{N-r}{\sigma_e^2(a, \lambda_i)} t), \qquad (17)$$

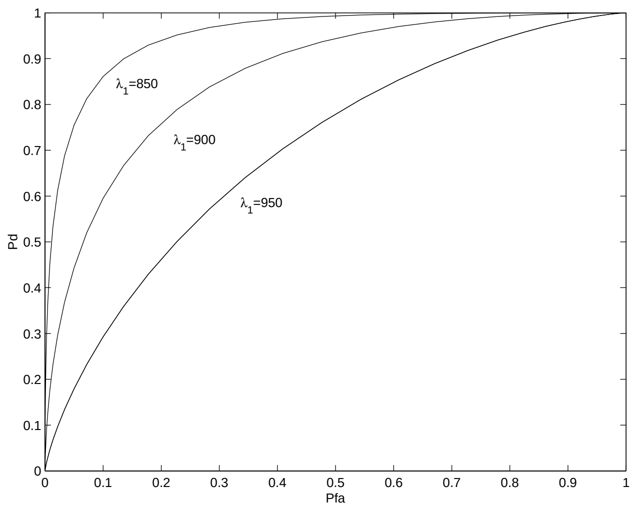

Figure 1: ROC curves for the NPD. $\lambda_0 = 1000$, $p = 1$, $a_1 = 0.2$, $N - r = 10$ and $\lambda_1 = 850$, 900, 950.

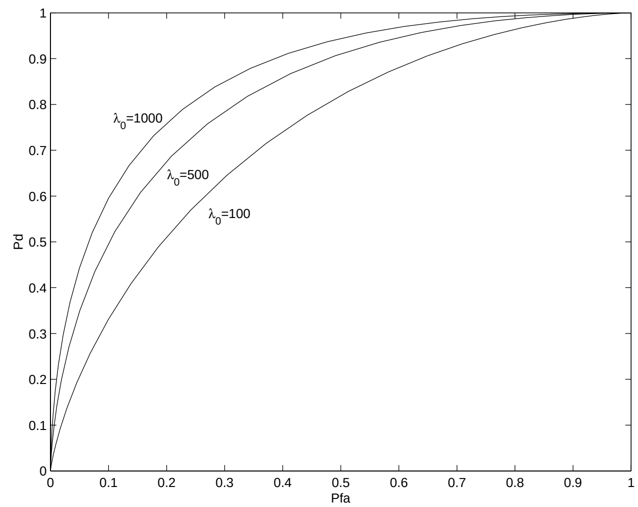

Figure 2: ROC curves for the NPD. $p = 1$, $a_1 = 0.2$, $N - r = 10$ and $(\lambda_0 = 1000, \lambda_1 = 900)$, $(\lambda_0 = 500, \lambda_1 = 450)$, $(\lambda_0 = 100, \lambda_1 = 90)$.

where:

$$f(t) = \frac{1}{2}(t/\mu_i)^{\frac{N-r-2}{4}} I_{\frac{N-r-2}{2}}(\sqrt{\mu_i t})e^{-\frac{\mu_i + t}{2}} \qquad (18)$$

and $I_\nu(x)$ is the modified Bessel function. It is worthy to note that μ_i and consequently $f(t)$ do not depend of λ_i. The dependence versus λ_i occurs only through $\sigma_e^2(a, \lambda_i)$ in (17).

In order to evaluate the performances of the NPD and the influence of the various parameters, ROC curves have been plotted from (17). Figure (1) shows the ROC curves for $\lambda_0 = 1000$, $p = 1$, $a_1 = 0.2$, $N - r = 10$ and different values of λ_1. Figure (2) shows the ROC curves when λ_0 varies and the amplitude of the jump equals 10% of λ_0. The curves show that the performances increase with λ_0. To explain this result define the "snr" as the ratio between the mean squared and the variance of the signal. The signal model implies that this snr equals λ_i under each hypothesis. Consequently, in this simulation, as λ_0 increases the overall "snr" and the width of the jump when both increase. Finally, figure (3) studies the effect of the signal correlation on the detection performances. The performances increase with the signal correlation, as it could be predicted.

4. GENERALIZED LIKELIHOOD RATIO DETECTOR

The optimal NPD gives a reference to which suboptimal detectors can be compared. However, it requires a priori knowledge of the abrupt change parameters λ_1 and r (the parameters a_k, $k = 1, \dots, p$ and λ_0 are assumed to be known). Otherwise, it can be replaced by the Generalized Likelihood detector (GLRD) obtained by the ratio of the supremum of the likelihood function with respect to the unknown parameters under both hypothesis.

This solution requires the computation of the Maximum Likelihood Estimatior (MLE) of λ_1 and r. When r is known the MLE of λ_1, denoted $\hat{\lambda}_1(r)$, is obtained by setting the partial derivative of $\mathcal{L}(X/H_1)$ with respect to λ_1 to zero. Straighforward calculations yield:

$$(1 + \sum_{k=1}^{p} a_k)^2 \lambda_1^2 + \sigma_e^2(a, 1)\lambda_1 - \frac{1}{N-r}\sum_{n=r+1}^{N}(x_n + \sum_{k=1}^{p} a_k x_{n-k})^2 = 0. \qquad (19)$$

The two roots of this second order polynomial being obviously of opposite sign, an analytic expression of $\hat{\lambda}_1(r)$ is given by the positive root of (19).

This result suggests the following remarks:

- In the Poisson iid case, the MLE of λ_1 equals the sample mean. In our case, λ_1 could obviously also be estimated by the sample mean or the sample variance of the process with a variance of the estimate in both case that is $O(1/N)$, [5]. For example, for the process (9) the corresponding mean MLE estimator is:

$$\hat{\lambda}_1 = \frac{1}{N-r}\sum_{n=r+1}^{N}\frac{x_n + \sum_{k=1}^{p} a_k x_{n-k}}{1 + \sum_{k=1}^{p} a_k} \qquad (20)$$

  However, these solutions do not take into account the particular constraint $E[x_n] = var[x_n]$, see (7,8).

- An estimator of the variance λ_1 is the difference between estimated signal second order moment and the mean squared: λ^2, in this case. Reordering the terms, we obtain that an estimate of λ_1 is given by the positive solution of

$$\lambda_1^2 + \lambda_1 - \frac{1}{N-r}\sum_{n=r+1}^{N} x_n^2 = 0. \qquad (21)$$

  It can be easily checked that in the uncorrelated case, $\forall k$, $a_k = 0$, equation (19) reduces to (21).

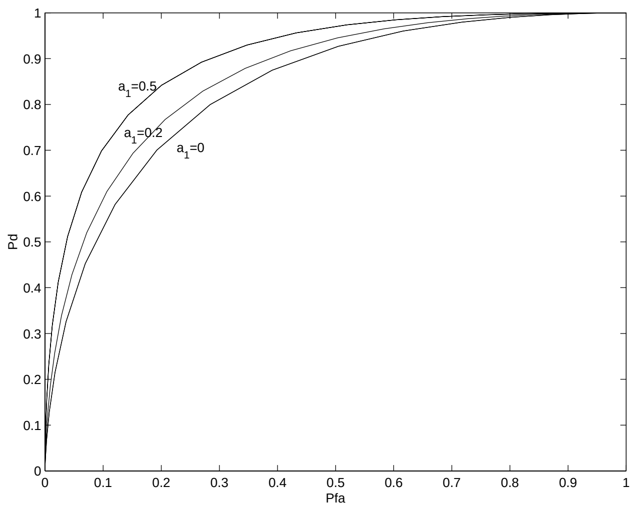

Figure 3: ROC curves for the NPD. $p = 1$, $N - r = 10$, $\lambda_0 = 1000$, $\lambda_1 = 900$, and $a_1 = 0$, 0.2, 0.5.

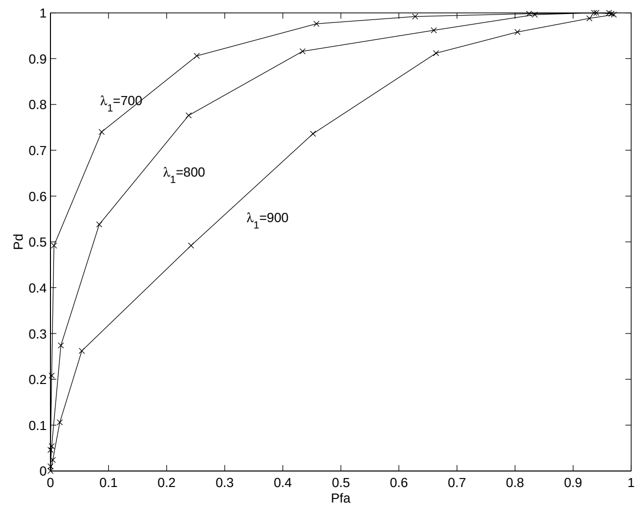

Figure 4: ROC curves for the GLRD. $p = 1$, $a_1 = 0.2$, $N = 512$, $r = 412$, $\lambda_0 = 1000$ and $\lambda_1 = 700$, 800, 900.

Finally, in order to obtain $\hat{r}$, the expression of $\hat{\lambda}_1(r)$ obtained from (19) is replaced in (12). The resulting criterion is evaluated for r in $[p+1, N-1]$. The global maximizer is retained as $\hat{r}$ and $\hat{\lambda}_1$ equals $\hat{\lambda}_1(\hat{r})$. This two quantities are then replaced in $\mathcal{L}(X/H_1) - \mathcal{L}(X/H_0)$ and the resulting value is compared to a threshold.

The GLRD ROC curves are depicted in figure 4 for $p = 1$, $a_1 = 0.2$, $N = 512$, $r = 412$, $\lambda_0 = 1000$ and for different values of λ_1 (700, 800 and 900). This figure is obtained using for each value of the threshold, corresponding to a cross, 500 independent signal realizations. Comparison between figure 1 and 4 shows that the well known loss of performances of the GLRD compared to the NPD, can be compensated by an increase of the samples under hypothesis H_1.

We have assumed through all this section that the parameters λ_0 and a_k, $k = 1, \dots, p$, are known. They could be included in the GLR test. However, their estimation relies on a calibration procedure that is performed independently of the detection. For this scope, the solution that as been retained is the MLE. In the same manner as above, the log likelihood function under H_0 is first maximized with respect to λ_0 and the resulting criterion is maximized with respect to parameters a_k with a numerical method. Simulations have shown that good performances are obtained when the initial conditions for the a_k are the unconstraint Yule Walker solution.

5. SUMMARY AND CONCLUSIONS

This communication proposed an algorithm for detection of extrasolar planets by occultation under high flux assumption. We demonstrated that this problem can be modeled as the detection of abrupt change in the variance/mean of a Gaussian autoregressive process having equal mean and variance. For this model, contrary to the general unconstraint case, the theoretical distribution of the test statistic for the Neyman-Pearson detector can be obtained. Maximum likelihood estimators for the abrupt changes parameters were derived in order to implement a generalized likelihood detector. Performances of this detector were obtained by computer simulations.

6. REFERENCES

[1] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes, Theory and Application*. Prentice Hall, 1993.

[2] H. Cramèr. *Mathematical Methods of Statistics*. Princeton, Princeton University Press, 1971.

[3] A. Ferrari. Limiting properties of the Poisson Mandel transform. Technical note, July 1998.

[4] M. Lavielle. Optimal Segmentation of Random Processes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 46(5):1365–1373, 1998.

[5] M. B. Priestley. *Spectral Analysis And Time Series*. Monographs and Textbooks. Academic Press, 8th edition, 1994.

[6] Petre Stoica and Randolph Moses. *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[7] F. Sultani. *Méthodes d'Inversion de la Transformation de Poisson*. PhD thesis, Université de Nice-Sophia Antipolis - France, 1995.

[8] J.Y. Tourneret. Detection and Estimation of Abrupt Changes Contaminated by Multiplicative Gaussian Noise. *EURASIP Signal Processing*, 68(3):259–270, 1998.