Signal Processing 82 (2002) 649–661
Keywords: Abrupt change detection; Photometric signal; Extra solar planet detection

Parametric modeling of photometric signals

A. Ferrari[a,*], J.Y. Tourneret[b], G. Alengrin[a]

[a] *UMR 6525 Astrophysique, Université de Nice Sophia-Antipolis, Parc Valrose, 06108 Nice cedex 2, France*
[b] *ENSEEIHT/TÉSA, 2 rue Camichel, BP 7122, 31071 Toulouse cedex 7, France*

Abstract

This paper studies a new model for photometric signals under high flux assumption. Photometric signals are modeled by Gaussian autoregressive processes having the same mean and variance denoted Constraint Gaussian Autoregressive Processes (CGARP's). The estimation of the CGARP parameters is discussed. The Cramér Rao lower bounds for these parameters are studied and compared to the estimator mean square errors. The CGARP is intended to model the signal received by a satellite designed for extrasolar planets detection. A transit of a planet in front of a star results in an abrupt change in the mean and variance of the CGARP. The Neyman–Pearson detector for this changepoint detection problem is derived when the abrupt change parameters are known. Closed form expressions for the Receiver Operating Characteristics (ROC) are provided. The Neyman–Pearson detector combined with the maximum likelihood estimator for CGARP parameters allows to study the generalized likelihood ratio detector. ROC curves are then determined using computer simulations.

1. Introduction

The detection of extrasolar planets is a challenging problem in astronomy (see the extrasolar planets encyclopedia [18]). Among the methods currently pursued to detect extrasolar planets, the transit method (also referred as photometric method or occultation method) may be the only one to find earth class planets in the near future. The transit method is based on the detection of photometric flux variations which results from the transit of a planet in front of a star. Several space-based projects have been proposed to achieve this goal including the US Kepler project, the Eddington european project and the COROT (COnvection and ROTation) french project. For a star like the Sun, the typical relative flux variation for a Jupiter-like giant planet is 10^{-2} (during 25 h) whereas for a Earth-like planet it is 10^{-4} (during 13 h). Consequently, the photometric flux variations caused by terrestrial planets are difficult to detect using the conventional transit method and new detectors have to be investigated.

The first problem addressed in this paper is the modeling of the photometric signals received by the satellite. Based on two realistic assumptions, the photometric signals are modeled by Gaussian autoregressive processes having the same mean and variance, denoted Constraint Gaussian Autoregressive Processes (CGARP's). The second problem addressed in this paper is the detection of extrasolar planets by using standard estimation and detection tools. The theoretical distribution of photometric signals being

* Corresponding author. Tel.: +33-9207-6349; fax: +33-9207-6321.
E-mail address: andrea.ferrari@unice.fr (A. Ferrari).

generally intractable, most extrasolar planet detectors reduce to measure the brightness drop of a star which results from the transit of one of its planets across its disk [6]. This paper proposes to detect a variation in the received photometric flux from the CGARP modeling and conventional likelihood ratio based detectors.

The paper is organized as follows:

- Section 2 studies the theoretical model for photometric signals. The concept of CGARP is introduced from the asymptotic distribution, of the data.
- Section 3 studies the maximum likelihood estimation (MLE) of the CGARP parameters and the corresponding Cramér Rao lower bounds (CRLB's) for large sample size. These bounds are compared to the mean square errors (MSE's) of the estimates obtained through Monte Carlo simulations.
- Finally, the Neyman–Pearson detector (NPD) for the abrupt change (AC) detection is developed in Section 4. The exact distribution of the test statistic is obtained, allowing computation of Receiver Operating Characteristics (ROC). The practical application where the AC parameters are unknown is then investigated. The generalized likelihood ratio detector (GLRD) is derived and its performance is studied from Monte Carlo simulations.

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. The model derivation relies on the semiclassical theory of photodetection: the light propagation to the detector is described by diffraction theory and the photocount only occurs during the sensor photodetection. When light with a fixed intensity over time is incident on a photodetector, the joint probability of registering the successive photocounts $x_n,\ n=1,\ldots,N$ on a single pixel is distributed according to an i.i.d. Poisson distribution:

$$P_{\text{ideal}}(\boldsymbol{X}) = \prod_{n=1}^{N} \frac{\mathrm{e}^{-\lambda}\lambda^{x_n}}{x_n!}, \tag{1}$$

where $\boldsymbol{X} = (x_1,\ldots,x_N)^{\mathrm{t}}$ and λ is the light intensity (square of the wave amplitude) integrated between two successive measurements [8, p. 466].

The problem is more complicated when the light wave incident on the photosurface has stochastic attributes. In this case, the distribution (1) is regarded as a conditional probability distribution and the deterministic parameter λ is replaced by a random variable λ_n with mean $E[\lambda_n]=\lambda$ [8, p. 467], [16, p. 419]. The distribution of $\boldsymbol{X} = (x_1,\ldots,x_N)^{\mathrm{t}}$ conditioned on $\boldsymbol{\Lambda}=(\lambda_1,\ldots,\lambda_N)^{\mathrm{t}}$ is then an independent sequence with Poisson distribution of parameter $\boldsymbol{\Lambda}$. Consequently, the unconditioned distribution of $\boldsymbol{X}$ is

$$P(\boldsymbol{X}) = \int_0^{\infty} \prod_{n=1}^{N} \frac{\mathrm{e}^{-\lambda_n}\lambda_n^{x_n}}{x_n!}\, p(\boldsymbol{\Lambda})\,\mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_N, \tag{2}$$

where $p(\boldsymbol{\Lambda})$ is the joint distribution of $(\lambda_1,\ldots,\lambda_N)$. This transform relating the photocount probability to the integrated intensity probability is often referred to as the Poisson Mandel transform of $p(\boldsymbol{\Lambda})$ [20]. This transform was first derived in 1958 using classical arguments and rederived in 1964 using a semiclassical method [16]. It is worthy to note that Eq. (2) can also handle the case where the detector is not perfect. If the quantum efficiency of the detector varies with time [16,8], Eq. (1) can be regarded as the ideal sensor signal. As previously, the statistical properties of the observed signal in the real case are described by Eq. (2). In the monovariate case, x_1 is known in the actuarial literature as a mixed Poisson process [9].

The ideal case defined in Eq. (1) can be obtained from (2) by choosing for $p(\boldsymbol{\Lambda})$ a product of Dirac delta functions. This case corresponds to a well-stabilized single-mode laser radiation. Theoretical expressions of $P(\boldsymbol{X})$ can also be derived in the case of polarized thermal light [8]. Unfortunately, an a priori distribution for $\boldsymbol{\Lambda}$ is generally very difficult to choose. Moreover, an analytic expression of (2) cannot be computed for most probability density functions (pdf's) $p(\boldsymbol{\Lambda})$. For these reasons, few studies use the stochastic model described by Eq. (2) for the detection of photometric flux variation.

This paper proposes a simple model for $\boldsymbol{X}$ using two realistic assumptions:

$$\mathsf{A1}: E[\lambda_n] = \lambda \gg 1, \quad \mathsf{A2}:\ \mathrm{var}[\lambda_n] \ll \lambda.$$

A1 expresses a high flux assumption and A2 conveys the fact that the variations of the integrated light intensity and the variations of the sensor are small. Based on these assumptions, this paper proposes to model x_n

as a Gaussian AR process subjected to the constraint $E[x_n] = \mathrm{var}[x_n] = \lambda$, denoted *Constraint Gaussian AR Process* (*CGARP*). The Gaussian distribution for $\boldsymbol{X}$ is justified by the following proposition.

Proposition 1. *If the probability distribution of* $\boldsymbol{X}$ *verifies* (2):

(1) *the mean and covariance matrix of* $\boldsymbol{X}$ *are*:

$$E[\boldsymbol{X}] = \lambda \boldsymbol{u}, \quad C_{XX} = C_{\Lambda\Lambda} + \lambda I_N, \tag{3}$$

where $\boldsymbol{u}$ *is a* $N \times 1$ *vector of ones*, I_N *is the* $N \times N$ *identity matrix*, λ *the mean of* λ_n *and* $C_{\Lambda\Lambda}$ *the covariance matrix of* $\boldsymbol{\Lambda}$.

(2) *Define* $C_{\Lambda\Lambda} = VD^2V^{\mathrm{t}}$ *the eigen decomposition of* $C_{\Lambda\Lambda}$, $\gamma_k(\boldsymbol{\tau})$ *the* kth *order cumulant of* $(\lambda_{\tau_1}, \ldots, \lambda_{\tau_k})$ *where* $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_k)$ *and the standardized vector*:

$$\boldsymbol{Y}_\lambda = (D + \sqrt{\lambda} I_N)^{-1} V^{\mathrm{t}} (\boldsymbol{X} - \lambda \boldsymbol{u}), \tag{4}$$

$(E[\boldsymbol{Y}_\lambda] = 0,\ C_{Y_\lambda Y_\lambda} = I_N)$. *If* $\gamma_k(\tau) = o(\lambda^{k/2})$ *for* $k \geqslant 2$, $\boldsymbol{Y}_\lambda$ *converges in distribution to the multivariate Gaussian distribution* $\mathcal{N}(0, I_N)$ *when* $\lambda \to +\infty$.

Proof. See Appendix A.

The vector $\boldsymbol{Y}_\lambda$ converges in distribution (when $\lambda \to +\infty$) to the multivariate Gaussian distribution $\mathcal{N}(0, I_N)$. Consequently, the distribution of $\boldsymbol{Y}_\lambda$ can be well approximated for large λ by its asymptotic distribution (see for instance [4, p. 204]). Hence, Eq. (4) and assumption A1, A2 imply that the distribution of $\boldsymbol{X}$ can be approximated by the multivariate Gaussian distribution $\mathcal{N}(\lambda \boldsymbol{u}, C_{\Lambda\Lambda} + \lambda I_N)$. Eq. (3) shows that second order stationarity for $\boldsymbol{\Lambda}$ implies second order stationarity for $\boldsymbol{X}$. Moreover, assumption A2 implies that the variance of x_n, which equals $\lambda + \mathrm{var}[\lambda_n]$, can be approximated by λ.

Finally, *parametric AR modeling* for $\boldsymbol{X}$ is motivated in the stationary case by the fact that for any continuous spectral density $S(f)$, an AR process can be found with a spectral density arbitrary close to $S(f)$ [4, p. 132]. Classical justifications for this model in the stationary Gaussian context such as the Wold decomposition can be also found for example in [17].

Based on the previous comments, $\boldsymbol{X}$ is modeled by a pth order CGARP defined by

$$x_n = -\sum_{k=1}^{p} a_k x_{n-k} + \lambda \sum_{k=0}^{p} a_k + e_n, \tag{5}$$

where $a_0 = 1$ and e_n is an i.i.d. zero mean Gaussian sequence. The variance of e_n in model (5) is such that $\mathrm{var}[x_n] = \lambda$. It is denoted $\sigma_e^2(\boldsymbol{a}, \lambda)$ in order to take into account its dependence toward the a_k and λ. An analytic expression of $\sigma_e^2(\boldsymbol{a}, \lambda)$ is difficult to obtain. However, a formal expression can be obtained by rewriting the Yule Walker equations [19] as a linear system where the unknowns are the signal covariances, $\boldsymbol{c} = (\lambda, c(1), \ldots, c(p))^{\mathrm{t}}$. This leads to

$$(A_1 + A_2)\boldsymbol{c} = (\sigma_e^2(\boldsymbol{a}, \lambda), 0, \ldots, 0)^{\mathrm{t}}, \tag{6}$$

where

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_1 & 1 & 0 & \cdots & 0 \\ a_2 & a_1 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ a_p & a_{p-1} & \cdots & a_1 & 1 \end{pmatrix},$$
$$A_2 = \begin{pmatrix} 0 & a_1 & \cdots & a_{p-1} & a_p \\ 0 & a_2 & \cdots & a_p & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & a_p & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{7}$$

Defining the vector $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\mathrm{t}}$, we then obtain:

$$\sigma_e^2(\boldsymbol{a}, \lambda) = \frac{\lambda}{\boldsymbol{e}_1^{\mathrm{t}}(A_1 + A_2)^{-1}\boldsymbol{e}_1} = \lambda \sigma_e^2(\boldsymbol{a}, 1). \tag{8}$$

Note that $\sigma_e^2(\boldsymbol{a}, \lambda)$ can be computed using the recursive algorithm studied in [1, p. 117] for the computation of the power of a linear process.

It is important to note that in the ideal case, the high flux assumption implies that the Poisson distribution can be approximated by a Gaussian distribution with same mean and variance [5]. This particular case corresponds to $a_k = 0,\ \forall k > 0$ in (5). Consequently, a major effect of the detector imperfections is to correlate the signal measurements.

3. CGARP parameter estimation

3.1. Maximum likelihood estimation

This section is devoted to the maximum likelihood estimation of $\boldsymbol{a}$ and λ. Denote $\mathscr{L}(X|\boldsymbol{a},\lambda)$ the log-likelihood function of $\{x_{p+1},\ldots,x_N\}$ conditioned on $\{x_1,\ldots,x_p\}$:

$$\begin{aligned}\mathscr{L}(X|\boldsymbol{a},\lambda) &= -\frac{N-p}{2}\log(2\pi\lambda\sigma_e^2(\boldsymbol{a},1)) \\ &\quad -\frac{1}{2\sigma_e^2(\boldsymbol{a},1)\lambda}\sum_{n=p+1}^{N}\left(\sum_{k=0}^{p}a_k x_{n-k}-\lambda\sum_{k=0}^{p}a_k\right)^2.\end{aligned} \tag{9}$$

After dropping some constants, the maximization of (9) with respect to λ for a known AR parameter vector $\boldsymbol{a}$ leads to

$$\begin{aligned}&\left(\sum_{k=0}^{p}a_k\right)^2\lambda^2+\sigma_e^2(\boldsymbol{a},1)\lambda \\ &\quad -\frac{1}{N-r}\sum_{n=p+1}^{N}\left(\sum_{k=0}^{p}a_k x_{n-k}\right)^2=0.\end{aligned} \tag{10}$$

The two roots of this second order polynomial being obviously of opposite sign, an analytic expression of $\hat{\lambda}(\boldsymbol{a})$ is given by the positive root of (10).

In order to obtain the MLE's of $(\lambda,\boldsymbol{a})$, $\hat{\lambda}(\boldsymbol{a})$ is replaced in (9) and the resulting criterion is maximized over $\boldsymbol{a}$ using a classical optimization algorithm. A critical point of this step is the optimization initialization. We propose to take as initial condition the estimates obtained with classical AR identification algorithms (e.g. [19]) after removing the estimated mean.

These results suggest the following remarks:

- In the Poisson i.i.d. case, the MLE of λ equals the sample mean. Here, λ could obviously also be estimated by the sample mean (method of moments) or the sample variance. However, the MLE of λ has to be preferred because of its asymptotical good properties [11].
- In order to take into account the constraint $E[x_n]=\mathrm{var}[x_n]=\lambda$, an estimator of λ could also be constructed by estimating the second order moment of x_n i.e. $E[x_n^2]=\lambda+\lambda^2$. This estimator is the positive root of the following second order equation:

$$\lambda^2+\lambda-\frac{1}{N-r}\sum_{n=r+1}^{N}x_n^2=0. \tag{11}$$

It can be easily checked that in the uncorrelated case ($a_k=0,\ \forall k\geqslant 1$), Eq. (10) reduces to Eq. (11). Consequently, the second order moment based estimator does not take into account the correlations between the observed samples x_n.

Next section derives the CRLB's for λ and $\boldsymbol{a}$ which are compared to the MLE mean square errors.

3.2. Cramér Rao lower bounds

CRLB's are convenient tools for determining the achievable accuracy of estimators. Unfortunately, the CRLB's cannot always be obtained in a simple closed form expression. In these cases, asymptotic expressions can be used to approximate the CRLB's for large values of the number of samples. This section derives the asymptotic CRLB's for the parameters $\boldsymbol{a}$ and λ of the model (5).

Proposition 2. *The asymptotic Cramér Rao lower bounds for the parameters of a CGARP defined in Eq.* (5) *verify*:

$$\begin{aligned}&\lim_{N\to+\infty}N.\mathrm{CRLB}(\lambda) \\ &\quad=\left[\frac{2\lambda^2}{\sigma_e^2(\boldsymbol{a},\lambda)}+\nabla_a(\lambda)^{\mathrm{t}}C_p^{-1}\nabla_a(\lambda)\right. \\ &\qquad\left.-\mu\left(\frac{2\lambda^2}{\sigma_e^2(\boldsymbol{a},\lambda)}+\nabla_a(\lambda)^{\mathrm{t}}C_p^{-1}\nabla_a(\lambda)\right)^2\right]\sigma_e^2(\boldsymbol{a},\lambda),\end{aligned} \tag{12}$$

$$\begin{aligned}&\lim_{N\to+\infty}N.\mathrm{CRLB}(\boldsymbol{a}) \\ &\quad=[C_p^{-1}-\mu C_p^{-1}\nabla_a(\lambda)\nabla_a(\lambda)^{\mathrm{t}}C_p^{-1}]\sigma_e^2(\boldsymbol{a},\lambda),\end{aligned} \tag{13}$$

where μ *is given by* (50), *the components of* $\nabla_a(\lambda)$ *by* (52) *and* C_p^{-1} *is computed using the Gohberg–Semencul formula* [19].

Proof. See Appendix B.

The theoretical expressions of the CRLB's derived in Eqs. (12) and (13) have been compared to the

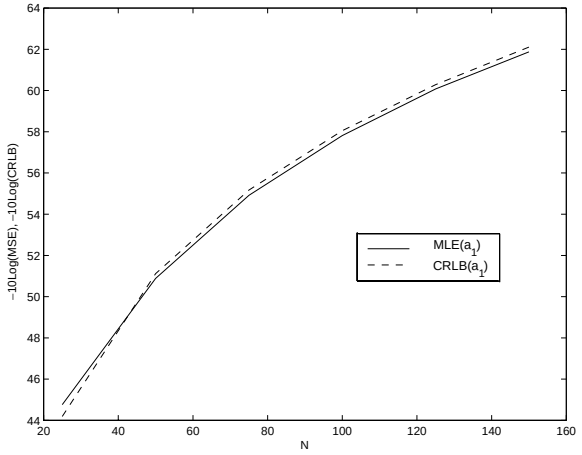

Fig. 1. CRLB of a_1 and estimated MSE of the MLE of $a_1 = 0.8$ for $p = 1$ and $\lambda = 1000$.

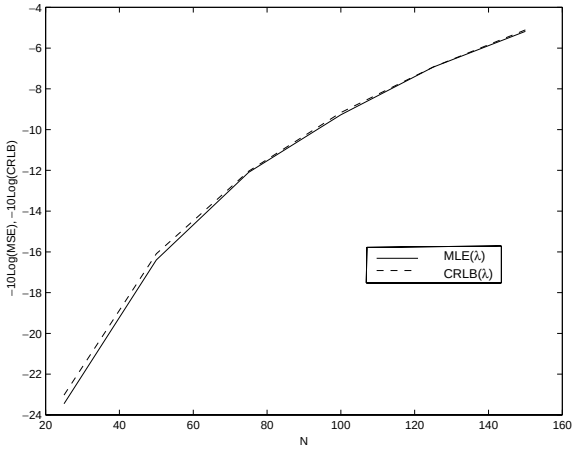

Fig. 2. CRLB of λ and estimated MSE of the MLE of $\lambda = 1000$ for $p = 1$ and $a_1 = 0.8$.

MSE's of the parameter MLE's. For this purpose, 1000 independent realizations of the signal (5) with $p = 1$, $a_1 = 0.8$, $\lambda = 1000$ have been generated for different values of N. The parameter MLE's have been determined for each realization and the corresponding MSE's have been computed. A comparison between the estimated MSE's and the CRLB's is depicted in Figs. 1 and 2. These figures suggest the following comments:

- for large values of N, a perfect adequacy between (12), (13) and the MSE's is observed: expressions (12), (13) are a good approximation for both the parameter CRLB's and the estimate MSE's,
- for small values of N, the "loss of efficiency" of the MLE and the unverified asymptotic assumption for CRLB's result in a increasing difference between the two curves,
- the validity domain for expressions (12), (13) is approximately $N \geqslant 40$, which is in agreement with the results obtained in [7].

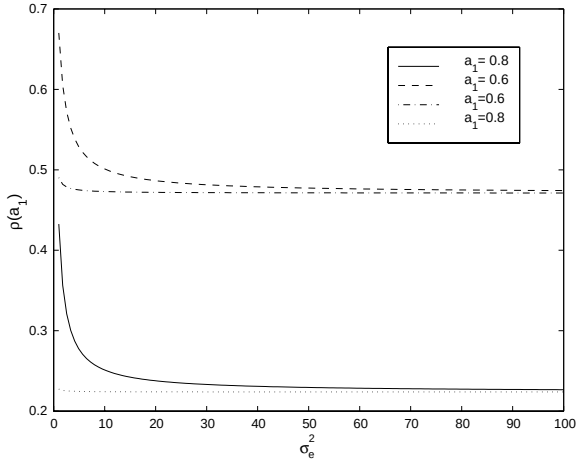

Fig. 3. Comparison between the CRLB of a_1 for the CGARP and the GARP.

Next simulations compare the CRLB's for the CGARP parameters corresponding to model (5) with the CRLB's for a standard unconstraint Gaussian AR Process (GARP) with mean λ, AR parameter vector $\boldsymbol{a}$ and driving noise variance σ_e^2. The later are denoted $\mathrm{CRLB}_u(\boldsymbol{\theta})$, where $_u$ stands for unconstraint and $\boldsymbol{\theta} = (\lambda, \sigma_e^2, \boldsymbol{a})$. The unconstraint Fisher Information Matrix for $\boldsymbol{\theta}$ is block diagonal and does not depend of λ. Consequently, the asymptotic CRLB's of $(\sigma_e^2, \boldsymbol{a})$ are given by the first terms of (45) and the asymptotic CRLB of λ is

$$\mathrm{CRLB}_u(\lambda) = \frac{\sigma_e^2}{\mathrm{N}(\sum_{k=0}^{p} a_k)^2}. \tag{14}$$

The CRLB's are compared for a first order model by means of the following ratios:

$$\rho(a_1) = \frac{\mathrm{CRLB}(a_1)}{\mathrm{CRLB}_u(a_1)}, \quad \rho(\lambda) = \frac{\mathrm{CRLB}(\lambda)}{\mathrm{CRLB}_u(\lambda)}. \tag{15}$$

Figs. 3 and 4 represent $\rho(a_1)$ and $\rho(\lambda)$ for different values of a_1. The bounds corresponding to the

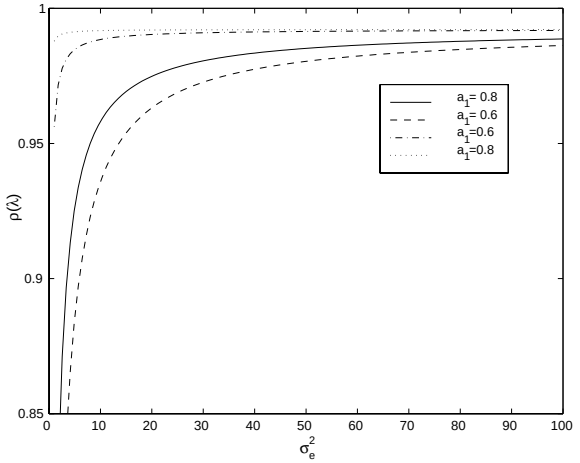

Fig. 4. Comparison between the CRLB of λ for the CGARP and the GARP.

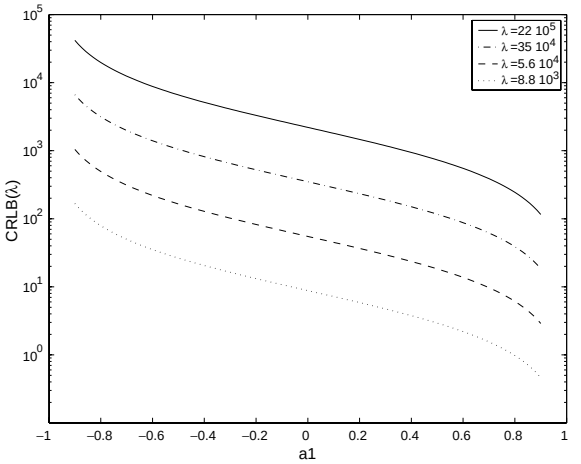

Fig. 5. CRLB of λ for the first order model ($N = 1000$).

CGARP parameters are functions of λ, contrarily to GARP's parameters. For this reason, the results are plotted as functions of σ_e^2. This choice guarantees that $\rho(a_1)$ and $\rho(\lambda)$ are computed for two processes having the same variance. The results suggest the following remarks:

- The asymptotic CRLB's are lower for the CGARP's than for the GARP's, since $\rho(a_1)$ and $\rho(\lambda)$ are < 1. This result can be easily generalized for the regression coefficients noticing that the first term of (13) is the asymptotic CRLB of the GARP coefficients and the second a positive definite matrix.
- When a_1^2 tends to 1, the variances of the CGARP's and GARP's tend to $+\infty$. Consequently, CRLB(a_1) and CRLB$_u(a_1)$ decrease to 0. Fig. 3 reveals that in this case CRLB(a_1) decreases faster than CRLB$_u(a_1)$.
- Fig. 4 shows that the limit of $\rho(\lambda)$ when σ_e^2 tends to $+\infty$ is not a function of a_1, contrarily to $\rho(a_1)$.

These results prove the better identifiability of CGARP's with respect to unconstraint GARP's (specially for the regression coefficients). Indeed, it is well known that the MLE is asymptotically efficient. In other words, the variances of ML estimates are close to the corresponding CRLB's for large number of samples. Consequently, the qualitative behavior of CRLB's for CGARP's and GARP's is similar to the qualitative behavior of CGARP and GARP parameter estimates. Regarding the detection algorithm, the performance of the GLRD is closely related to the variance of the estimated value of λ_1, which is smaller for the CGARP model than for the GARP model. Consequently, the GLRD based on CGARP modeling has to be preferred to the GLRD based on GARP modeling.

Finally, Fig. 5 depicts the behavior of CRLB(λ) as a function of a_1 for different values of λ. The objective is to evaluate the achievable precision on the star flux estimation. The parameters of the simulation have been computed from Eddington specifications [6]: a sampling time of 30 s, a collecting area of 55 cm^2, a bandwidth of 3250 Å, an optical transmission of 0.92^3 and a detector quantum efficiency of 0.84. The different star magnitudes (degrees of brightness) used in the simulation are $M_v = 8, 10, 12$ and 14 and the observation time has been chosen equal to 8.3 hours ($N = 1000$). These values correspond to a flux of $\lambda = 22 \times 10^5$, 35×10^4, 5.6×10^4 and 8.8×10^3. Fig. 5 shows that CRLB(λ) increases when λ increases and when a_1 decreases. Eq. (14) proves that this last property is also verified by the GARP's.

4. Abrupt change detection

This section is devoted to the major application of the model under scope: the detection of a decrease in the photometric flux. In the search for terrestrial planets by occultation this photometric flux variation

results from the decrease of the star diffracted intensity during the transit of a planet. This can be modeled using the notations of Section 2 as AC (at instant r) in the parameter λ:

$$\begin{aligned} &\forall n \in S_0 = \{1,\dots,r\} && \lambda = \lambda_0, \\ &\forall n \in S_1 = \{r+1,\dots,N\} && \lambda = \lambda_1 \ (< \lambda_0). \end{aligned} \tag{16}$$

The planet detection problem can then be formulated as the following AC detection problem:

$$\begin{aligned} &H_0 : S_1 = \emptyset \text{ (no jump)}, \\ &H_1 : S_1 \neq \emptyset \text{ (jump)}. \end{aligned} \tag{17}$$

In the case of i.i.d. Poisson distributed data x_n, the NPD for problem (17) can be easily derived and yields:

$$H_0 \text{ rejected if } \frac{1}{N-r}\sum_{n=r+1}^{N} x_n < \zeta. \tag{18}$$

The problem is obviously more complicated when the observations x_n are correlated. The next section derives the NPD for problem (17), when x_n is the CGARP defined in (5).

AC detection and estimation for AR processes have been studied for long time (see [3,14] 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 and variance of a CGARP. The study is restricted to off-line change point detection algorithm [3]. A similar problem was studied in [21], for nonzero mean AR processes 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 checked 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.

4.1. The Neyman–Pearson detector

After dropping the constant terms, the log-likelihood function of $\{x_{p+1},\dots,x_N\}$ conditioned on $\{x_1,\dots,x_p\}$ under hypothesis H_1 can be written:

$$\begin{aligned} \mathscr{L}(X|H_1) &= -(N-p)\log\sigma_e^2(\boldsymbol{a},1) \\ &\quad -(r-p)\log\lambda_0 - (N-r)\log\lambda_1 \\ &\quad -\frac{1}{\sigma_e^2(\boldsymbol{a},1)}\left(\frac{1}{\lambda_0}\sum_{n=p+1}^{r}\mathrm{e}_{n,0}^2 + \frac{1}{\lambda_1}\sum_{n=r+1}^{N}\mathrm{e}_{n,1}^2\right), \end{aligned} \tag{19}$$

where

$$e_{n,i} = \sum_{k=0}^{p} a_k x_{n-k} - \lambda_i\left(\sum_{k=0}^{p} a_k\right). \tag{20}$$

The log likelihood function under hypothesis H_0 is readily obtained from (19). After dropping the constant terms, we obtain:

$$\begin{aligned} \mathscr{L}(X|H_0) &= -(N-p)\log\sigma_e^2(\boldsymbol{a},1) - (N-p)\log\lambda_0 \\ &\quad -\frac{1}{\sigma_e^2(\boldsymbol{a},1)\lambda_0}\sum_{n=p+1}^{N}\mathrm{e}_{n,0}^2. \end{aligned} \tag{21}$$

Using the hypothesis $\lambda_1 < \lambda_0$, the NPD reduces to:

$$T = \frac{1}{N-r}\sum_{n=r+1}^{N}\left(\sum_{k=0}^{p} a_k x_{n-k}\right)^2 \underset{H_1}{\overset{H_0}{\gtrless}} \zeta \tag{22}$$

which suggests the following comments:

- the Neyman Pearson test is a uniformly most powerful test, since the test statistic T does not depend on λ_0 and λ_1,
- in the constraint i.i.d. case ($a_k = 0$, $\forall k > 0$), T reduces to the estimated power of x_n, whereas in the unconstraint i.i.d. Poisson case, T is the estimated mean (18).

When a jump occurs independently on the mean and the variance of the process as in [21], the test statistic T is the difference between two positive definite quadratic forms which is generally indefinite. The exact distribution of T is then very difficult to study and has been approximated by a Gaussian distribution in [21]. In this paper, because of the constraint (8), T reduces to a single positive definite quadratic form whose distribution can be determined under hypothesis H_0 and H_1: under hypothesis H_i, T is the sum

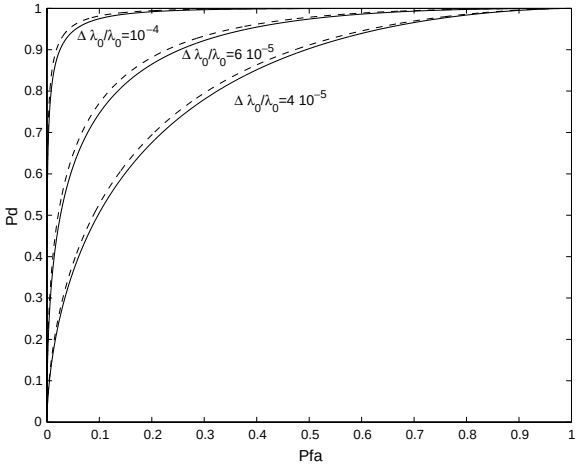

Fig. 6. ROC curves for the NPD as a function of $\Delta\lambda_0/\lambda_0$ for a fixed value of λ_0 ($\lambda_0 = 35 \times 10^4$, $N - r = 1000$). Continuous line: $p = 1$, $a_1 = 0.2$. Dashed line: $p = 2$, $a_1 = 0.2$, $a_2 = 0.05$.

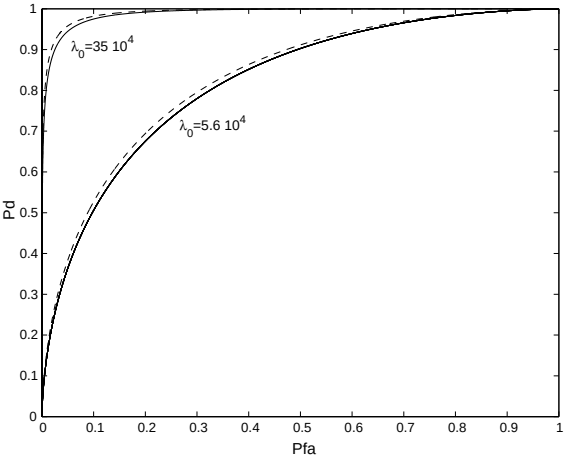

Fig. 7. ROC curves for the NPD as a function of λ_0 for a fixed value of $\Delta\lambda_0/\lambda_0$ ($\Delta\lambda_0/\lambda_0 = 10^{-4}$, $N - r = 1000$). Continuous line: $p = 1$, $a_1 = 0.2$. Dashed line: $p = 2$, $a_1 = 0.2$, $a_2 = 0.05$.

of the square of $N - r$ independent Gaussian random variables with:

- mean $\lambda_i(\sum_{k=0}^{p} a_k)/\sqrt{N-r}$,
- variance $\sigma_e^2(\boldsymbol{a}, \lambda_i)/(N - r)$.

Consequently, $(N - r)T/\sigma_e^2(\boldsymbol{a}, \lambda_i)$ is distributed as a noncentral χ^2 distribution with $N - r$ degrees of freedom and noncentrality parameter:

$$\mu_i = (N-r)\,\frac{\lambda_i^2(\sum_{k=0}^{p} a_k)^2}{\sigma_e^2(a,\lambda_i)} = (N-r)\,\frac{\lambda_i(\sum_{k=0}^{p} a_k)^2}{\sigma_e^2(a,1)}. \tag{23}$$

The pdf of T can be expressed as

$$p_T(t|H_i) = \frac{N-r}{\sigma_e^2(\boldsymbol{a},\lambda_i)}\, f_i\left(\frac{N-r}{\sigma_e^2(\boldsymbol{a},\lambda_i)} - t\right), \tag{24}$$

where $f_i(t)$ is a mixture of central χ^2 pdf's:

$$f_i(t) = \frac{1}{2}\,(t/\mu_i)^{(N-r-2)/4} I_{(N-r-2)/2}(\sqrt{\mu_i t})\mathrm{e}^{-(\mu_i+t)/2}, \quad t \geqslant 0, \tag{25}$$

and $I_v(x)$ is the modified Bessel function of the first kind of order v [10].

Eq. (24) allows us to plot ROC curves, in order to evaluate the NPD performance and the influence of the various parameters. Figs. 6 and 7 show the ROC curves computed for a transit of an earth-like planet in front of a magnitude 10 star, as a function of the ratio $\Delta\lambda_0/\lambda_0$ where $\Delta\lambda_0 = \lambda_0 - \lambda_1$ and λ_0. As can be expected, the NPD performance increases when $\Delta\lambda_0/\lambda_0$ and λ_0 increase. These figures also show that the qualitative behavior of the NPD is very similar for a first order CGARP ($p = 1$, continuous line) and an higher-order CGARP ($p = 2$, dashed line). Based on these comments, next simulations have been carried out for first order CGARP's for simplicity. Fig. 8

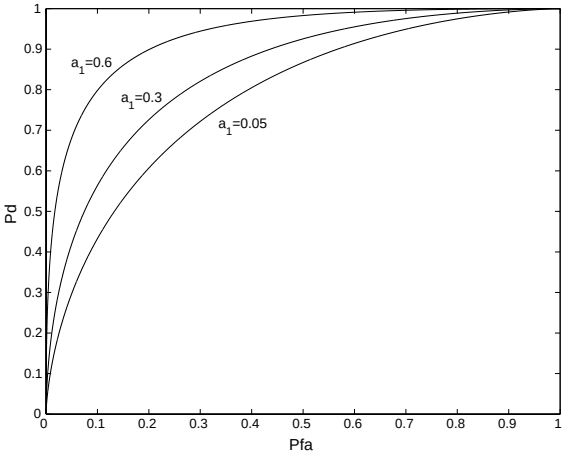

Fig. 8. ROC curves for the NPD as a function of a_1 ($p = 1$, $a_1 = 0.2$, $N - r = 1000$, $\lambda_0 = 5.6 \times 10^4$, $\Delta\lambda_0/\lambda_0 = 10^{-4}$).

studies the effect of the signal correlation on the detection performance. The detector performance increases with the signal correlation, as it could be predicted.

4.2. Generalized likelihood ratio detector

The optimal NPD provides 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, $k = 1,\ldots,p$ and λ_0 are assumed to be known). The GLRD has received much attention in practical applications, where the abrupt change parameters are unknown [2,12]. The GLRD is the ratio of the supremum of the likelihood function with respect to the unknown parameters under both hypothesis. The resulting likelihood ratio is compared to a suitable threshold which depends on the Probability of False Alarm (PFA).

The GLRD requires the computation of the MLE's of λ_1 and r. When r is known, the MLE of λ_1, denoted $\hat{\lambda}_1(r)$, is defined as in Section 3.1 by the positive root of

$$\left(\sum_{k=0}^{p} a_k\right)^2 \lambda_1^2 + \sigma_e^2(\boldsymbol{a},1)\lambda_1 - \frac{1}{N-r}\sum_{n=r+1}^{N}\left(\sum_{k=0}^{p} a_k x_{n-k}\right)^2 = 0. \tag{26}$$

In order to obtain the MLE of r, the expression of $\hat{\lambda}_1(r)$ obtained from (26) is replaced in (19) and the resulting criterion is evaluated for $r \in \{p+1,\ldots,N-1\}$. The global maximizer of this criterion defines the MLE of r denoted $\hat{r}$ and the MLE of λ_1 is $\hat{\lambda}_1(\hat{r})$.

The GLRD statistics is then obtained by replacing r and λ_1 in $\mathcal{L}(X|H_1) - \mathcal{L}(X|H_0)$ by their MLE's. A closed form expression of the GLRD statistic distribution is clearly difficult to derive. Consequently, ROC curves have been computed using Monte Carlo simulations.

Fig. 9 shows the ROC curve of the GLRD for $p=1$, $a_1 = 0.2$, $N = 6000$, $r = 3000$, $\lambda_0 = 14 \times 10^6$ ($M_v = 6$) and $\Delta\lambda_0/\lambda_0 = 1.5 \times 10^{-2}$. Note that this figure has been obtained from 100 independent signal realizations. A comparison between Figs. 6 and 9 shows the well-known loss of performance of the GLRD compared to the NPD. However, the GLRD shows satisfactory performance depending on the signal correlation which is measured here by the value of a_1.

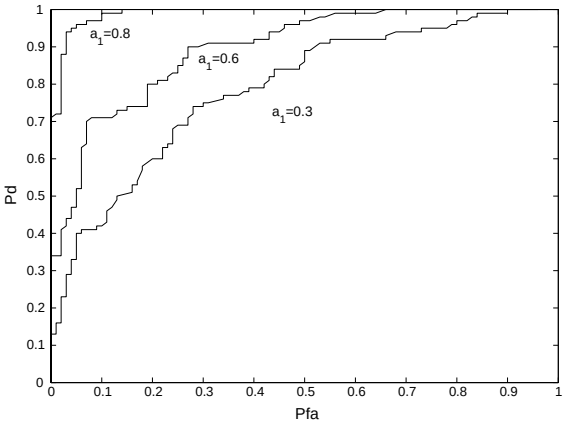

Fig. 9. ROC curves for the GLRD. $p=1$, $N-r=6000$, $r=3000$, $\lambda_0 = 14 \times 10^6$ and $\Delta\lambda_0/\lambda_0 = 1.5 \times 10^{-2}$.

It is important to note that the parameters λ_0 and a_k, $k = 1,\ldots,p$, have been assumed to be known in this section. These parameters are obviously unknown in practical applications and have to be estimated. Their estimation relies on a calibration procedure that is performed independently of the detection using the method described in Section 3.1. However, it is interesting to note that the MLE's of parameters λ_0 and a_k, $k = 1,\ldots,p$ could also be included in the likelihood ratio. This might significantly modify the GLRD performance.

5. Conclusions

This paper studied a new model denoted CGARP (for Constraint Gaussian Autoregressive Process) model for the analysis of high flux photometric signal. The estimation of CGARP parameters using the maximum likelihood method was addressed. A comparison between the parameter estimates and the corresponding Cramér Rao lower bounds was provided. CGARP modeling was shown to be a useful tool for the detection of extrasolar planets by occultation. Indeed, the extrasolar planet detection problem was formulated as the detection of abrupt changes in the CGARP parameters. The Neyman–Pearson detector for this detection problem with known

parameters was studied. Closed form expressions of the Neyman–Pearson detector ROC's were derived. The generalized likelihood ratio detector was then studied for practical applications in which the parameters are unknown. ROC's for the generalized likelihood ratio detector were computed by using Monte Carlo simulations and provided the detection performance.

The abrupt change detection problem addressed in this paper assumed that the flux variations due to the presence of a planet were observed on the last data samples. Within the framework of extrasolar planet detection, this problem occurs when photometric signals are recorded during a short time interval. When signals are recorded during a longer time interval, the photometrix flux decreases when the planet is in front of the star and takes its original value when the planet has moved. These flux variations can even be observed during successive occultations. It is interesting to note that the proposed detection algorithm can be extended to this situation, by defining an appropriate set of indices S_1 in the AC detection problem (16). In this case, the NPD test statistics T can be obtained from Eq. (22), where the summation over n is replaced by $n \in S_1$ and $N - r$ is replaced by the cardinality of S_1. Consequently, the pdf of T and the ROC curves can be computed very similarly to those obtained in this paper. Of course, the GLRD performance increases when the number of samples under hypothesis H_1 increases. Consequently, the detection performance should improves, when successive planet occultations can be observed. This generalization is currently under investigation.

Finally, the AC model used in this paper does not take into account the ingress/egress of the planet on the star and the limb-darkening. These effects can be easily incorporated in the formalism proposed in the article at the price of an increased number of unknown parameters. However, it is worthy to note that they will only have a little effect on the detection performance [22].

Appendix Asymptotic normality of $\boldsymbol{Y_\lambda}$

1. Mean and covariance of $\boldsymbol{X}$

The first and second-order statistics of $\boldsymbol{X}$ can be easily computed using conditional expectations:

$$E[x_n] = E[E[x_n/\lambda_n]] = E[\lambda_n] = \lambda, \tag{A.1}$$

$$\operatorname{cov}[x_i, x_j] = E[E[x_i x_j|\lambda_i, \lambda_j]] - \lambda^2 \tag{A.2}$$

$$= \begin{cases} \operatorname{cov}[\lambda_i, \lambda_j] & \text{if } i \neq j, \\ \lambda + \operatorname{var}[\lambda_i] & \text{if } i = j. \end{cases} \tag{A.3}$$

2. Asymptotic normality of $\boldsymbol{Y_\lambda}$

In order to prove the asymptotic normality of $\boldsymbol{Y}_\lambda$, we study its second characteristic function denoted $\Psi_{Y_\lambda}(\boldsymbol{\Omega}) = \log \Phi_{Y_\lambda}(\boldsymbol{\Omega})$, where $\Phi_{Y_\lambda}(\boldsymbol{\Omega}) = E[\mathrm{e}^{j\boldsymbol{\Omega}^{\mathrm{t}}\boldsymbol{Y}_\lambda}]$ and $\boldsymbol{\Omega} = (\omega_1, \ldots, \omega_N)^{\mathrm{t}}$. Denote $\boldsymbol{\varphi}_\lambda = V(D + \sqrt{\lambda} I_N)^{-1}\boldsymbol{\Omega}$ where I_N is the $N \times N$ identity matrix. The components of the vector $\boldsymbol{\varphi}_\lambda$ are

$$\varphi_k(\lambda) = \sum_{q=1}^{N} \frac{v_{kq}\omega_q}{d_q + \sqrt{\lambda}} \tag{A.4}$$

using obvious notations. Note that assumption A2 and the orthogonality of V implies that

$$\varphi_k(\lambda) = O\left(\frac{1}{\sqrt{\lambda}}\right). \tag{A.5}$$

A straightforward computation yields

$$\Psi_{Y_\lambda}(\boldsymbol{\Omega}) = -j\lambda \sum_{k=1}^{N} \varphi_k + \log \Phi_X(\boldsymbol{\varphi}_\lambda). \tag{A.6}$$

Using conditional expectations, the first characteristic function of $\boldsymbol{X}$ can be written as follows:

$$\Phi_X(\boldsymbol{\Omega}) = E[E[\mathrm{e}^{j\boldsymbol{\Omega}^{\mathrm{t}}X}|\boldsymbol{\Lambda}]] = E\left[\prod_{k=1}^{N} \mathrm{e}^{(\mathrm{e}^{j\omega_k}-1)\lambda_k}\right] \tag{A.7}$$

$$= E[\mathrm{e}^{\sum_{k=1}^{N} \lambda_k(\mathrm{e}^{j\omega_k}-1)}]. \tag{A.8}$$

Consequently, by denoting $\Upsilon_{\boldsymbol{\Lambda}}(\boldsymbol{S}) = E[\mathrm{e}^{\boldsymbol{\Lambda}^{\mathrm{t}}\boldsymbol{S}}]$ the moment generating function of $\boldsymbol{\Lambda}$, Eq. (A.8) reads:

$$\Phi_X(\boldsymbol{\varphi}_\lambda) = \Upsilon_{\boldsymbol{\Lambda}}(\boldsymbol{S}), \tag{A.9}$$

with $\boldsymbol{S} = (s_1, \ldots, s_N)^{\mathrm{t}}$ and $s_k = \mathrm{e}^{j\varphi_k(\lambda)} - 1$.

Assuming that Λ satisfies some regularity conditions (see [15, p. 198]) and denoting $\gamma_k(\boldsymbol{\tau})$ the kth order cumulant of $(\lambda_{\tau_1},\ldots,\lambda_{\tau_k})$ where $\boldsymbol{\tau}=(\tau_1,\ldots,\tau_k)$, the second-moment generating function of Λ can be expanded as

$$\log \Upsilon_{\Lambda}(\boldsymbol{S}) = \lambda \sum_{\tau_1=1}^{N} s_{\tau_1} + \frac{1}{2} \sum_{\tau_1,\tau_2=1}^{N} \gamma_2(\tau_1,\tau_2) s_{\tau_1} s_{\tau_2} + \frac{1}{3!} \sum_{\tau_1,\tau_2,\tau_3=1}^{N} \gamma_3(\tau_1,\tau_2,\tau_3) s_{\tau_1} s_{\tau_2} s_{\tau_3} + \cdots \tag{A.10}$$

for $\lambda > \rho$ ($\rho > 0$ is defined such that the series in Eq. (A.10) represents a function which is regular for $\lambda > \rho$ [15, p. 198]). The second characteristic function of Y_λ can then be computed by replacing (A.10) in (A.6):

$$\Psi_{Y_\lambda}(\boldsymbol{\Omega}) = -j\lambda \sum_{\tau_1=1}^{N} \varphi_{\tau_1}(\lambda) + \lambda \sum_{\tau_1=1}^{N} (\mathrm{e}^{j\varphi_{\tau_1}(\lambda)} - 1) + \frac{1}{2} \sum_{\tau_1,\tau_2=1}^{N} \gamma_2(\tau_1,\tau_2)(\mathrm{e}^{j\varphi_{\tau_1}(\lambda)} - 1) \times (\mathrm{e}^{j\varphi_{\tau_2}(\lambda)} - 1) + \cdots . \tag{A.11}$$

Consequently, by substituting $\mathrm{e}^{j\varphi_k(\lambda)}$ by its power series expansion and using Eq. (A.5) and the hypothesis $\gamma_k(\boldsymbol{\tau}) = o(\lambda^{k/2})$, $k \geqslant 2$, we obtain:

$$\Psi_{Y_\lambda}(\boldsymbol{\Omega}) = -\frac{\lambda}{2} \sum_{\tau_1=1}^{N} \varphi_{\tau_1}(\lambda)^2 + o(1). \tag{A.12}$$

Finally, the limit off $\Psi_{Y_\lambda}(\boldsymbol{\Omega})$ can be computed noticing that the sum in the first term is the norm of $\boldsymbol{\varphi}_\lambda$:

$$\lim_{\lambda\to+\infty} \Psi_{Y_\lambda}(\boldsymbol{\Omega}) = \lim_{\lambda\to+\infty} -\frac{\lambda}{2} \sum_{q=1}^{N} \frac{\omega_q^2}{(d_q + \sqrt{\lambda})^2} = -\frac{1}{2} \sum_{q=1}^{N} \omega_q^2. \tag{A.13}$$

Eq. (A.13) shows that the second characteristic function of Y_λ converges to the second characteristic function of the multivariate Gaussian distribution $\mathcal{N}(0, I_N)$ when $\lambda \to \infty$. Consequently, Y_λ converges in distribution to the multivariate Gaussian distribution $\mathcal{N}(0, I_N)$.

Appendix B. Asymptotic CRLB's for the CGARP parameters

The unknown parameter vector for the CGARP defined by (5) is $\boldsymbol{\alpha} = (\lambda, \boldsymbol{a})$. Since there is a one-to-one transformation between $\boldsymbol{\alpha} = (\lambda, \boldsymbol{a})$ and $\boldsymbol{\theta} = (\sigma_e^2, \boldsymbol{a})$, the CRLB's for $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ are linked by the following relation [13]:

$$\mathrm{CRLB}(\boldsymbol{\alpha}) = \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathrm{CRLB}(\boldsymbol{\theta}) \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^{\mathrm{t}}, \tag{B.1}$$

where $\partial g(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ is the Hessian of the transformation:

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \lambda/\sigma_e^2 & \nabla_a(\lambda)^{\mathrm{t}} \\ 0 & I_p \end{pmatrix}. \tag{B.2}$$

In this expression $\lambda/\sigma_e^2 = 1/\sigma_e^2(\boldsymbol{a}, 1)$ is the partial derivation of λ with respect to σ_e^2 (according to (8)) and $\nabla_a(\lambda)$ is the gradient of λ with respect to $\boldsymbol{a}$.

B.1. Asymptotic CRLB for $\boldsymbol{\theta} = (\sigma_e^2, \boldsymbol{a})$

The coefficients of the Fisher Information Matrix (FIM) for a Gaussian process are the sum of a term that only depends on the covariances of the process and a term that takes into account its mean [19]. The first term of the asymptotic FIM for $\boldsymbol{\theta}$ is the well-known asymptotic form given for example in [7]. The mean of the process being $\lambda \boldsymbol{u}$, $\boldsymbol{u} = (1,\ldots,1)^{\mathrm{t}}$, the coefficients of the second term are

$$\left(\frac{\partial \lambda}{\partial \theta_k} \boldsymbol{u}^{\mathrm{t}}\right) C_N^{-1} \left(\frac{\partial \lambda}{\partial \theta_l} \boldsymbol{u}\right) = \boldsymbol{u}^{\mathrm{t}} C_N^{-1} \boldsymbol{u} \frac{\partial \lambda}{\partial \theta_k} \frac{\partial \lambda}{\partial \theta_l}, \tag{B.3}$$

where C_N is the order N signal covariance matrix. Consequently:

$$\mathrm{FIM}(\boldsymbol{\theta}) = N \begin{pmatrix} (2\sigma_e^4)^{-1} & 0 \\ 0 & (\sigma_e^2)^{-1} C_p \end{pmatrix} + \boldsymbol{u}^{\mathrm{t}} C_N^{-1} \boldsymbol{u} \nabla_{\boldsymbol{\theta}}(\lambda) \nabla_{\boldsymbol{\theta}}(\lambda)^{\mathrm{t}}, \tag{B.4}$$

where

$$\nabla_{\boldsymbol{\theta}}(\lambda)^{\mathrm{t}} = \left(\frac{\lambda}{\sigma_e^2}, \nabla_a(\lambda)^{\mathrm{t}}\right). \tag{B.5}$$

The inversion of (B.4) using the inversion lemma and products of bloc matrices yields:

$$\mathrm{CRLB}(\boldsymbol{\theta}) = \frac{\sigma_e^2}{N}\begin{pmatrix} 2\sigma_e^2 & 0 \\ 0 & C_p^{-1} \end{pmatrix} - \frac{\sigma_e^2}{N}\mu_N \begin{pmatrix} 2\lambda \\ C_p^{-1}\nabla_a(\lambda) \end{pmatrix} (2\lambda \ \nabla_a(\lambda)^{\mathrm{t}} C_p^{-1}), \tag{B.6}$$

with

$$\mu_N = \frac{\sigma_e^2 \boldsymbol{u}^{\mathrm{t}} C_N^{-1} \boldsymbol{u}}{N}\left(1 + \frac{\sigma_e^2}{N}\boldsymbol{u}^{\mathrm{t}} C_N^{-1} \times \boldsymbol{u}\left(\frac{2\lambda^2}{\sigma_e^2} + \nabla_a(\lambda)^{\mathrm{t}} C_p^{-1}\nabla_a(\lambda)\right)\right)^{-1}. \tag{B.7}$$

The inverse of C_N is computed using the Gohberg–Semencul formula [19] which yields:

$$\boldsymbol{u}^{\mathrm{t}} C_N^{-1}\boldsymbol{u} = \frac{1}{\sigma_e^2}\left((N-p)\left(\sum_{k=0}^{p} a_k\right)^2 + \sum_{q=1}^{p}\left(\sum_{k=0}^{p-q} a_k\right)^2 - \sum_{q=1}^{p}\left(\sum_{k=q}^{p} a_k\right)^2\right). \tag{B.8}$$

Hence

$$\lim_{N\to+\infty} \frac{\boldsymbol{u}^{\mathrm{t}} C_N^{-1}\boldsymbol{u}}{N} = \frac{1}{\sigma_e^2}\left(\sum_{k=0}^{p} a_k\right)^2. \tag{B.9}$$

Replacing this expression in (B.7) yields:

$$\mu = \lim_{N\to+\infty} \mu_N \tag{B.10}$$

$$= \left(\sum_{k=0}^{p} a_k\right)^2 \left(1 + \left(\sum_{k=0}^{p} a_k\right)^2 \times \left(\frac{2\lambda^2}{\sigma_e^2} + \nabla_a(\lambda)^{\mathrm{t}} C_p^{-1}\nabla_a(\lambda)\right)\right)^{-1}. \tag{B.11}$$

B.2. Asymptotic CRLB for $\boldsymbol{\alpha} = (\lambda, \boldsymbol{a})$

The substitution of Eqs. (B.2), (B.6) in (B.1) leads to Eqs. (13), (12). Note that $\nabla_a(\lambda)$ can be computed from (8) as follows:

$$\frac{\partial\lambda}{\partial a_k} = \sigma_e^2 \frac{\partial(\sigma_e^2(\boldsymbol{a}, 1))^{-1}}{\partial a_k} = \sigma_e^2 \frac{\partial \boldsymbol{e}_1^{\mathrm{t}}(A_1 + A_2)^{-1}\boldsymbol{e}_1}{\partial a_k} \tag{B.12}$$

$$= -\sigma_e^2 \boldsymbol{e}_1^{\mathrm{t}}(A_1 + A_2)^{-1} \frac{\partial(A_1 + A_2)}{\partial a_k} \times (A_1 + A_2)^{-1}\boldsymbol{e}_1. \tag{B.13}$$

References

[1] K. Åström, Introduction to Stochastic Control Theory, Academic Press, New York, 1970.

[2] M. Basseville, A. Benveniste, Detection of Abrupt Changes in Signals and Dynamical Systems, Lecture Notes in Control and Information, Springer, Berlin, 1986.

[3] M. Basseville, I. Nikiforov, Detection of Abrupt Changes, Theory and Application, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[4] P.J. Brockwell, R. Davis, Time Series: Theory and Methods, Springer, Berlin, 2nd Edition, 1990.

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

[6] Eddington, Assessment Study Report, ESA-SCI (2000)8, July 2000.

[7] B. Friedlander, B. Porat, The Exact Cramer–Rao Bound for Gaussian Autoregressive Processes, IEEE Trans. Aerospace Elect. Systems, 25 (4) 3–8.

[8] J. Goodman, Statistical Optics, Wiley Interscience, New York, 1985.

[9] J. Grandell, Mixed Poisson Processes, Chapman & Hall, London, 1997.

[10] N.L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, Vol. 2, Wiley, New York, 1994.

[11] S.M. Kay, Modern Spectral Estimation, Signal Processing Series, Prentice-Hall, Englewood Cliffs, NJ, 1987.

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

[13] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Signal Processing Series, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[14] M. Lavielle, Optimal segmentation of random processes, IEEE Trans. Signal Process. 46 (5) (1998) 1365–1373.

[15] E. Lukacs, Characteristic Functions, Griffin, London, 2nd Edition, 1970.

[16] C. Mehta, Progress in Optics: Theory of photoelectron counting, Vol. 8, North-Holland, Amsterdam, 1970.

[17] M.B. Priestley, Spectral Analysis and Time Series, Monographs and Textbooks, 8th Edition, Academic Press, New York, 1994.

[18] J. Schneider, The extrasolar planets encyclopaedia, http://www.obspm.fr/planets.
[19] P. Stoica, R. Moses, Introduction to Spectral Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1997.
[20] F. Sultani, Méthodes d'Inversion de la Transformation de Poisson, Ph.D. Thesis, Université de Nice-Sophia Antipolis, France, 1995.
[21] J.Y. Tourneret, Detection and estimation of abrupt changes contaminated by multiplicative Gaussian noise, EURASIP Signal Process. 68 (3) (1998) 259–270.
[22] J.Y. Tourneret, A. Ferrari, A. Swami, Cramer–Rao lower bounds for abrupt changes in additive and multiplicative noise, IEEE Trans. Signal Process., to be submitted.