TRAVAUX ET RECHERCHES

8

ASYMPTOTIC THEORY FOR NON I.I.D. PROCESSES

Proceedings of the 5th Franco-Belgian Meeting of Statisticians, november 1984

EDITED BY
J.-P. FLORENS
ASSISTED BY
M. MOUCHART, J.-P. RAOULT,
J.-M. ROLIN and L. SIMAR

Bruxelles
Publications des Facultés universitaires Saint-Louis
Boulevard du Jardin botanique 43
1986

TESTING FOR A CHANGE-POINT IN LINEAR SYSTEMS WITH INCOMPLETE OBSERVATION

Fabien CAMPILLO

INRIA, Valbonne, France

Abstract

The aim of this work is to develop a test procedure for a change-point detection problem in a partially observed linear Gaussian continuous time model. We study the asymptotic behaviour of the test we propose. We also investigate the properties of the maximum likelihood estimators of the associated parameters.

1. INTRODUCTION

The detection of changes in statistical models has inspired many works for years, as proved by the numerous publications on the theme (cf. the survey papers [11 , 17] and the survey included in [1]). The special importance of these methods lies in their ability to comply with sudden parametric model variations. Classical methods were unable of that.

One of the most often used techniques is based on the maximum likelihood statistic. This approach has been mainly developed by Willsky and Jones [18], in the failure detection problem in linear systems with incomplete observation. Their works lead to a practicable algorithm for the implementation of the likelihood ratio test procedure. However, the level of this test is not well defined. What is more, to this day no result has been presented concerning the optimality of this method and the convergence of the associated estimators.

An interesting approach to this problem is given by the asymptotical statistic theory (cf. Lecam [13], and also Ibragimov-Has'Minskii [10]), already applied to the failure detection by Deshayes and Picard [5 ,9]. For statistical model with change-point, they study the asymptotic behaviour of the likelihood method. With an invariance theorem on the likelihood process, their works emphasize "edges-problems", then they overcome these by using weighted likelihood ratio statistics.

Most works already published in this field concern the failure detection in discrete time dynamical systems. Let us consider the continuous time linear system :

$$\left.\begin{aligned} dX_t &= (AX_t + a)dt + B\,dW_t , \quad X_o = x_o \\ dY_t &= (CX_t + c)dt + dV_t \quad , \quad Y_o = 0 \end{aligned}\right\} \tag{1.1}$$

in which X represents the state of a non-observable physical system and Y represents its observation. W and V are two independent standard Wiener processes, with values in $\mathbb{R}^n$ and $\mathbb{R}^d$ respectively; X and Y take values in $\mathbb{R}^n$ and $\mathbb{R}^d$ respectively.

This paper deals with the test for sudden changes in parameter a . The problem of detection of abrupt changes in parameter A is more difficult and in Campillo [3] we investigate this situation.

Let us consider the state equation :

$$dX_t = \{AX_t + a + \Gamma(\tau,t)\gamma\}dt + B\,dW_t , \tag{1.2}$$

where $\Gamma(\tau, t) = 0$ if $\tau < t$, 1 otherwise; τ denotes the unknown time of failure and γ ($\neq 0$) the amplitude of the jump. We suppose that the coefficients A, a, B, C, c and the initial condition x_o are known. $\theta = \begin{pmatrix} \tau \\ \gamma \end{pmatrix}$ denotes the unknown parameter.

Given the observation $\{Y_t \ ; \ t \leq 1\}$ (to simplify the notations we will take 1 as final instant), we now have to decide whether there has been failure (i.e. $\tau \geq 1$), or not (i.e. $\tau < 1$). In the first case, we also want to estimate the failure parameter θ.

In Section 2, we shall present the likelihood ratio test associated to this situation; because of the edges-problems, we shall have to reformulate the test problem. Then we shall present an asymptotical study (Section 3) : We shall elaborate a new test procedure and prove its asymptotical optimality in the local sense (Section 3.2). We shall also investigate the asymptotic properties of the maximum likelihood estimator of the failure parameter θ (Section 3.3).

2. THE LIKELIHOOD RATIO TEST (LRT)

In this section, we present the LRT of the "no failure" hypothesis against the "failure" hypothesis. The derivation of the likelihood ratio that we propose is closely related to that of Willsky and Jones in the discrete time case. At the end of this section difficulties will come up, that are related with the implementation of the likelihood ratio test (LRT) procedure.

Let us consider the system :

$$\left.\begin{aligned} dX_t &= \{AX_t + a + \Gamma(\tau,t)\gamma\}dt + B\, dW_t \ , \quad X_o = x_o \\ dY_t &= (CX_t + c)dt + dV_t \qquad\qquad\quad , \quad Y_o = 0 \end{aligned}\right\} . \tag{2.1}$$

Let H_o and $H_1(\theta)$ be respectively the hypotheses of "no failure" and "failure of parameter θ". We define :

$$\hat{X}_t = E_o[X_t \mid Y_s \ ; \ s \leq t] \quad .$$

$$R(t) = E_o[(X_t - \hat{X}_t)\ (X_t - \hat{X}_t)^*] \ ,$$

(E_o denotes the expectation with respect to H_o and M^* the transpose of M). $\hat{X}_t$ and $R(t)$ are given by the following Kalman "H_o-filter" (cf. [15]) :

$$d\hat{X}_t = (A\hat{X}_t + a)dt + R(t)C^* \{dY_t - (C\hat{X}_t + c)dt\}$$
$$\dot{R}(t) = AR(t) + R(t)A^* + BB^* - R(t)C^*CR(t) \tag{2.2}$$

with $\hat{X}_o = x_o$ and $R(0) = 0$. The innovation ν of this filter is defined by :

$$\nu_t = Y_t - \int_0^t (C\hat{X}_s + c)ds \ . \tag{2.3}$$

It is well known that (cf. [15]) :

$$\text{Under } H_o : \nu \text{ is a standard Wiener process.} \tag{2.4}$$

We now want to compute the law of the process ν under $H_1(\theta)$. To this end, we introduce $S(\theta,\cdot)$ - the θ-failure signature - :

$$S(\theta,t) := S_t^\tau \gamma := C \int_0^t \Psi(t,s)\ \Gamma(\tau,s)ds\ \gamma \tag{2.5}$$

where $\Psi(t,s)$ is the fundamental matrix :

$$\frac{\partial\Psi(t,s)}{\partial t} = \{A - R(t)C^*C\}\ \Psi(t,s) \ , \quad \Psi(s,s) = I_{n\times n} \ . \tag{2.6}$$

$S(\theta,\cdot)$ provides us with an explicit description of how various failures propagate through the system and the filter, in the sens that :

$$\text{Under } H_1(\theta) : \left\{\nu_t - \int_0^t S(\theta,u)du\right\} \text{ is a standard Wiener process ,} \tag{2.7}$$

which follows from the fact

$$\nu_t^{(\theta)} := \nu_t - \int_0^t S(\theta,u)du \tag{2.8}$$

is the innovation of the Kalman filter based on the $H_1(\theta)$ hypothesis.

Let $\Omega = C([0,1]\,;\mathbb{R}^d)$ and $\underline{\underline{F}}$ = Borel field over Ω . We define P_o (resp. P_θ) the measure induced by ν on $(\Omega,\underline{\underline{F}})$ under the hypothesis H_o (resp. $H_1(\theta)$). In fact P_o is the Wiener measure on $(\Omega,\underline{\underline{F}})$. Using Girsanov's theorem, it follows from (2.4) and (2.7) that $P_o \sim P_\theta$ and :

$$dP_\theta = L(\theta)\ dP_o$$

where $$L(\theta) = L(\tau,\gamma) = \exp \int_0^1 \left\{S(\theta,t)^*\ d\nu_t - \frac{1}{2}\,|S(\theta,t)|^2\right\} .$$

The likelihood ratio for "no failure" against "failure" is :

$$L := \max \left\{ L(\tau,\gamma) \ ; \ 0 \leqslant \tau < 1 \ , \ \ \gamma \neq 0 \right\} .$$

Let :

$$M^{(\tau)} := \int_0^1 (S_t^\tau)^* \, d\nu_t \ , \quad \langle M^{(\tau)} \rangle := \int_0^1 (S_t^\tau)^* \, S_t^\tau \, dt \ .$$

We can easily check that the maximum likelihood estimator $\hat{\theta} = \begin{pmatrix} \hat{\tau} \\ \hat{\gamma} \end{pmatrix}$ of the parameter θ :

$$L(\hat{\theta}) = \max \left\{ L(\tau,\gamma) \ ; \ 0 \leqslant \tau < 1 \ , \ \ \gamma \neq 0 \right\}$$

is given by :

$$\hat{\tau} \in \operatorname{Arg} \max_{0 \le \tau < 1} \left\{ M^{(\tau)*} \langle M^{(\tau)} \rangle^{-1} M^{(\tau)} \right\} \ , \quad \hat{\gamma} = \langle M^{(\hat{\tau})} \rangle^{-1} \ M^{(\hat{\tau})} \ .$$

The LRT for failure detection is :

$$\left. \begin{array}{lll} L \leqslant \sigma & : & \text{"no failure"} \\ & & \\ L > \sigma & : & \text{"failure"} \end{array} \right\} , \tag{2.9}$$

for some decision threshold σ to be chosen.

We now come up against two difficulties :

- Though the distribution of the statistic $L(\tau,\gamma)$ under H_0 is known, the one of $\max \{L(\tau,\gamma) \ ; \ 0 \leqslant \tau < 1 \ , \ \gamma \neq 0\}$ is not, because the dependence of $L(\tau,\gamma)$ with respect to (τ,γ) is too complicated. So the level of the test (2.9) is not defined.
- The LRT consists in testing all the failure hypotheses from 0 to 1. In order to test the failure hypothesis at time τ, the only useful information is given by $\{Y_t \ ; \ \tau \leqslant t < 1\}$. When τ is too close to 1, there is a lack of information and the test procedure might fail. These "edges"problems" have been pointed out from both a numerical (cf. [2 ,18]) and theoretical (cf. [5 ,9]) viewpoints.

In order to overcome these difficulties we shall reformulate the initial test problem : Instead of "$\tau = 1$" vs "$\tau < 1$", we consider the new test problem :

$$\left. \begin{array}{l} H_0 : \text{"no failure occurs before } \tau_0 \ (\tau \geqslant \tau_0)\text{"} \\ \text{against} \\ H_1 : \text{"a failure occurs before } \tau_0 \ (\tau < \tau_0)\text{"} \end{array} \right\} \tag{2.10}$$

where $\tau_0 < 1$.

This method is related to the weighted likelihood ratio statistics introduced by Deshayes and Picard. We shall derive a test procedure for H_0 vs H_1 and prove its optimality in a local sense for a well suited mode of convergence.

3. ASYMPTOTICAL STUDY

Several asymptotics can be considered for this problem, e.g. "amplitude of the jump $\to \infty$", "noises intensity $\to 0$", "period of observation $\to \infty$". In this last case we have not yet obtained significant results. On the other hand, one can prove that the first two asymptotics are equivalent (cf. Campillo [3]).

We are now dealing with the first asymptotic. Let us consider the following system :

$$\left.\begin{array}{ll} dX_t = \left\{AX_t + a + \frac{1}{\varepsilon}\,\Gamma(\tau,t)\gamma\right\}dt + B\,dW_t\ , & X_0 = x_0 \\ dY_t = (CX_t + c)dt + dV_t & ,\quad Y_0 = 0 \end{array}\right\}, \tag{3.1}$$

$\varepsilon \in\,]0,1]$. We suppose that the hypotheses of the previous sections are fulfilled.

Consider ν the innovation process (2.2) - (2.3). As in the preceding section :

$$\text{Under } H_0 \quad : \text{"}\{\nu_t\} \text{ is a standard Wiener process"} \tag{3.2}$$

$$\text{Under } H_1(\theta) : \text{"}\left\{\nu_t - \frac{1}{\varepsilon}\int_0^1 S(\theta,u)du\right\} \text{ is a standard Wiener process"}, \tag{3.3}$$

where $S(\theta,\cdot)$ is defined by (2.5) - (2.6).

Let P_0 (resp. $P_\theta^{(\varepsilon)}$) be the law of the process ν on $(\Omega,\underline{\underline{F}})$ when H_0 (resp. $H_1(\theta)$) holds. (3.2) - (3.3) imply that P_0 is the Wiener measure on $(\Omega,\underline{\underline{F}})$ and :

$$\frac{dP_\theta^{(\varepsilon)}}{dP_0} = \exp \int_0^1 \left\{\frac{1}{\varepsilon}\, S(\theta,t)^*\, d\nu_t - \frac{1}{2\varepsilon^2}\, |S(\theta,t)|^2\, dt\right\} . \tag{3.4}$$

From now on, ν will denote the canonical process on $(\Omega,\underline{\underline{F}}) = (C([0,1]\,,\mathbb{R}^d)$, Borel field) (i.e. $\nu_t(\omega) = \omega(t)$) . From Girsanov's theorem it follows that the process :

$$\nu_t^{(\varepsilon,\theta)} := \nu_t - \frac{1}{\varepsilon}\int_0^1 S(\theta,u)\,du \tag{3.5}$$

is a standard Wiener process over $(\Omega,\underline{\underline{F}},P_\theta^{(\varepsilon)})$. E_0 and $E_\theta^{(\varepsilon)}$ will denote the expectation with respect to P_0 and $P_\theta^{(\varepsilon)}$ respectively.

In Section 3.1, we establish that the family $\{P_\theta^{(\varepsilon)}\}$ satisfies a condition of local asymptotic normality (LAN) as $\varepsilon \to 0$. In Section 3.2, this property allows us to derive an asymptotically optimal test procedure for H_0 vs H_1. The kind of test we propose is quite new in the failure detection problem, and is related to Neyman's $C(\alpha)$-test (cf. [4 , 16]).

In Section 3.3, we use Ibragimov-Has'Minskii's work [10] so as to investigate the asymptotical properties of the maximum likelihood estimator of the failure parameter.

3.1. Local Asymptotic Normality (LAN) of $\{P_\theta^{(\varepsilon)}\}$

Let us consider the following notations :

$\nabla(\theta,t)$ is the mean square derivative of $S(\theta,t)$ with respect to θ,

$$\text{i.e.}\quad \lim_{\delta\to 0} |\delta|^{-2}\int_0^1 |S(\theta+\delta,t) - S(\theta,t) - \nabla(\theta,t)\delta|^2\,dt = 0\ ; \tag{3.6)'$$

$F(\theta)$ is the Fisher information matrix at point θ :

$$F(\theta) = \int_0^1 \nabla(\theta,t)^*\,\nabla(\theta,t)\,dt\ ; \tag{3.6)"$$

$$\Delta_{\varepsilon,\theta} := \int_0^1 \nabla(\theta,t)^*\,\left\{d\nu_t - \frac{1}{\varepsilon}S(\theta,t)\,dt\right\}. \tag{3.6)'''$$

It is not difficult to prove that $S(\theta,\cdot)$ is differentiable with respect to θ in $L^2(0,1)$. In fact, we have :

$$\nabla(\theta,t) = \left(\dot S_t^\tau\,\gamma \mid S_t^\tau\right) \quad \text{where}\quad \dot S_t^\tau := \partial S_t^\tau/\partial\tau\ .$$

<u>*THEOREM 3.1.*</u> *Let* $\theta \in \Theta = \left\{\begin{pmatrix}\tau\\ \gamma\end{pmatrix} ;\ 0 < \tau < 1\ ,\ \gamma \neq 0\right\}$. *The family* $\left\{P_\theta^{(\varepsilon)}\ ;\ \theta\in\Theta\right\}$ *satisfies a property of LAN at point* θ *as* $\varepsilon \to 0$. *Namely, for the following normalized likelihood ratio :*

$$Z_{\varepsilon,\theta}(U) = dP_{\theta+\varepsilon U}^{(\varepsilon)} / dP_\theta^{(\varepsilon)}\ ,\quad U \in D_{\varepsilon,\theta} := \left\{U = \begin{pmatrix}u\\ v\end{pmatrix} \in \mathbb{R}^{1+n}\ ;\ \theta+\varepsilon U \in \Theta\right\},$$

the representation :

$$Z_{\varepsilon,\theta}(U) = \exp\left\{U^* \Delta_{\varepsilon,\theta} - \frac{1}{2} U^* F(\theta) U + 0_{\varepsilon,\theta}(U)\right\}$$

is valid, where

(i) *Law* $(\Delta_{\varepsilon,\theta} \mid P_\theta^{(\varepsilon)}) = N(0,F(\theta))$;

(ii) *for any bounded sequence* $\{U_\varepsilon\}$ $(U_\varepsilon \in D_{\varepsilon,\theta})$:

$$P_\theta^{(\varepsilon)} - \lim_{\varepsilon\to 0} 0_{\varepsilon,\theta}(U_\varepsilon) = 0$$

(i.e. $P_\theta^{(\varepsilon)}\ (|0_{\varepsilon,\theta}(U_\varepsilon)| > \eta) \to 0$ *as* $\varepsilon \to 0,\ \forall\, \eta > 0$).

<u>*Proof*</u> *:* From (3.5) it follows that :

$$\Delta_{\varepsilon,\theta} := \int_0^1 \nabla(\theta,t)^*\, d\nu_t^{(\varepsilon,\theta)} .$$

Under $P_\theta^{(\varepsilon)}$, $\nu^{(\varepsilon,\theta)}$ is a standard Wiener process, thus $\Delta_{\varepsilon,\theta}$ is Gaussian with mean 0 and covariance matrix $F(\theta)$. So (i) is proved.

We now define :

$$0_{\varepsilon,\theta}(U) = \log\left(Z_{\varepsilon,\theta}(U)\right) - U^*\Delta_{\varepsilon,\theta} + \frac{1}{2} U^*F(\theta)U$$

$$= \log\left(dP_{\theta+\varepsilon U}^{(\varepsilon)} / dP_0\right) - \log\left(dP_\theta^{(\varepsilon)} / dP_0\right) - U^*\Delta_{\varepsilon,\theta} + \frac{1}{2} U^*F(\theta)U .$$

This, together with (3.4), yields

$$0_{\varepsilon,\theta}(U) = \frac{1}{\varepsilon}\int_0^1 \left(S(\theta+\varepsilon U,t) - S(\theta,t) - \nabla(\theta,t)\varepsilon U\right)^* d\nu_t^{(\varepsilon,\theta)}$$

$$- \frac{1}{2\varepsilon^2}\int_0^1 \left(|S(\theta+\varepsilon U,t) - S(\theta,t)|^2 - |\nabla(\theta,t)\varepsilon U|^2\right) dt .$$

Let $\{U_\varepsilon\}$ be a bounded sequence. From this last relation we deduce that $0_{\varepsilon,\theta}(U_\varepsilon)$ is a Gaussian variable $N(\mu_\varepsilon,\sigma_\varepsilon^2)$ under $P_\theta^{(\varepsilon)}$. The definition of $\nabla(\theta,t)$ implies that $\mu_\varepsilon \to 0$ and $\sigma_\varepsilon^2 \to 0$ as $\varepsilon \to 0$, so the assertion (ii) of the theorem is proved.

□

3.2. Testing for a Change-Point

Let $\theta_0 = \begin{pmatrix}\tau_0\\ \gamma_0\end{pmatrix} \in \Theta$ such that $F(\theta_0) > 0$, in [3] one can find an example of sufficient condition for the existence of such a point. We substitute variable $U = \begin{pmatrix}u\\ v\end{pmatrix} = (\theta - \theta_0)/\varepsilon$ and instead of "$\tau \geq \tau_0$" vs "$\tau < \tau_0$" we examine the test problem :

$$H_0 : \text{"}u \geq 0\text{"} \quad \text{vs} \quad H_1 : \text{"}u < 0\text{"} . \tag{3.7}$$

We remark that v is a nuisance parameter. Let Δ_ε, F, $0_\varepsilon(U)$ denote $\Delta_{\varepsilon,\theta_0}$, $F(\theta_0)$, $0_{\varepsilon,\theta_0}(U)$ respectively; $\Delta_\varepsilon = \begin{pmatrix}\Delta_\varepsilon^1\\ \Delta_\varepsilon^2\end{pmatrix}$ where Δ_ε^1 (resp. Δ_ε^2) takes values in $\mathbb{R}$ (resp. $\mathbb{R}^n$). Furthermore, we define :

$$P_U^{(\varepsilon)} = P_{\theta_0+\varepsilon U}^{(\varepsilon)} , \quad E_U^{(\varepsilon)} = E_{\theta_0+\varepsilon U}^{(\varepsilon)} ,$$

$$D_{\varepsilon,K}^0 = \left\{U = \begin{pmatrix}u\\ v\end{pmatrix} ; \theta_0 + \varepsilon U \in \Theta , |U| \leq K , u \geq 0\right\} ,$$

$$D_{\varepsilon,K}^1 = \left\{U = \begin{pmatrix}u\\ v\end{pmatrix} ; \theta_0 + \varepsilon U \in \Theta , |U| \leq K , u < 0\right\} ,$$

$$D_{\varepsilon,K} = D_{\varepsilon,K}^0 \,\dot{\cup}\, D_{\varepsilon,K}^1 .$$

Let $\Phi^* : \mathbb{R} \times \mathbb{R}^n \to \{0,1\}$ be the test function :

$$\Phi^*(\delta^1,\delta^2) := \begin{cases} 1 & \text{if } \delta^1 \leq \rho(\delta^2) \\ 0 & \text{if } \delta^1 > \rho(\delta^2) \end{cases} \tag{3.8}$$

where $\rho(\cdot)$ is chosen so that

$$E_0^{(\varepsilon)} [\Phi^*(\Delta_\varepsilon^1,\delta^2) \mid \Delta_\varepsilon^2 = \delta^2] = \alpha, \quad \delta^2\text{-a.e.} \tag{3.9}$$

(α is a given false alarm probability). It follows from Theorem 3.1 that law $(\Delta_\varepsilon \mid P_0^{(\varepsilon)}) = N(0,F)$, so we can easily prove that the relation (3.9) determines without ambiguity the function $\rho(\cdot)$, namely :

LEMMA 3.2.

$$\rho(\delta^2) = \frac{q(\alpha)}{G_1^{\frac{1}{2}}} - \frac{g^*\delta^2}{G_1} , \quad \delta^2 \in \mathbb{R}^n ,$$

where $q(\alpha)$ *denotes the* $(1-\alpha)$*-quantile of the Gauss law* $N(0,1)$, *and*

$G = \left(\begin{array}{c|c} G_1 & g^* \\ \hline g & G_2 \end{array}\right)$ $(G_1, g, G_2 \in \mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n\times n})$ *denotes the inverse matrix of* F .

□

This is the main result of the section :

THEOREM 3.3. The test-sequence $\{\Phi^*_\varepsilon\}$ *defined by* $\Phi^*_\varepsilon = \Phi^*(\Delta^1_\varepsilon,\Delta^2_\varepsilon)$, *is asymptotically locally most powerful unbiased of level* α .

In order to prove this result we need new notations and a lemma.

NOTATIONS 3.4.

Let $\Delta = \begin{pmatrix}\Delta^1 \\ \Delta^2\end{pmatrix}$ be a $N(0,F)$ variable defined on a probability space $(\Xi,\underline{\underline{A}},\tilde{P}_0)$. We define :

(i) $N(d\delta^1,d\delta^2)$ the distribution of Δ ;

(ii) $N^1(d\delta^1 \mid \delta^2)$ the conditional distribution of Δ^1 given $\Delta^2 = \delta^2$;

(iii) $N^2(d\delta^2)$ the marginal distribution of Δ^2 .

Let $\tilde{P}_U$ $(U = \begin{pmatrix}u \\ v\end{pmatrix})$ be the probability over $(\Xi,\underline{\underline{A}})$ such that :

$$d\tilde{P}_U = \exp(U^*\Delta)\ c(U)\ d\tilde{P}_0$$

($c(U)$ is a factor of normalization).

On the space $(\Xi,\underline{\underline{A}},\tilde{P}_U)$, we define :

(i) $N_U(d\delta^1,d\delta^2)$ the distribution of Δ ;

(ii) $N^1_U(d\delta^1 \mid \delta^2)$ the conditional distribution of Δ^1 given $\Delta^2 = \delta^2$;

(iii) $N^2_U(d\delta^2)$ the marginal distribution of Δ^2 .

We have the following classical results (cf. Lehmann [14]) :

(i) $N_U(d\delta^1,d\delta^2) = \exp(U^*\delta)\ c(U)\ N(d\delta^1,d\delta^2)$;

(ii) $N^1_U(d\delta^1 \mid \delta^2) = \exp(u\delta^1)\ c_{\delta^2}(u)\ N^1(d\delta^1 \mid \delta^2)$;

(iii) $N^2_U(d\delta^2) = \exp(v^*\delta^2)\ (c_{\delta^2}(u))^{-1}\ c(U)\ N^2(d\delta^2)$;

where $\delta = \begin{pmatrix} \delta^1 \\ \delta^2 \end{pmatrix}$, $U = \begin{pmatrix} u \\ v \end{pmatrix}$. $c(U)$ and $c_{\delta^2}(u)$ are normalizing coefficients. In particular, we note that $N_U^1(d\delta^1 \mid \delta^2) = N_u^1(d\delta^1 \mid \delta^2)$ is independent of v .

LEMMA 3.5. Let $\phi_\varepsilon \in L^\infty(\Omega, \underline{\underline{F}}, P_0^{(\varepsilon)})$ *be such that* $|\phi_\varepsilon| \leqslant L$ a.s. $(\forall\ \varepsilon)$. *Then for any given bounded sequence* $\{U_\varepsilon\}$ *in* $\mathbb{R} \times \mathbb{R}^n$:

$$E_{U_\varepsilon}^{(\varepsilon)}\ \phi_\varepsilon \underset{\varepsilon \to 0}{\sim} \tilde{E}_{U_\varepsilon}\ \psi_\varepsilon(\Delta^1, \Delta^2) ,$$

where $\psi_\varepsilon(\delta^1, \delta^2) := E_0^{(\varepsilon)}\ [\phi_\varepsilon \mid \Delta_\varepsilon^1 = \delta^1,\ \Delta_\varepsilon^2 = \delta^2]$, $\tilde{E}_U$ *denotes the expectation with respect to* $\tilde{P}_U$ $(A_\varepsilon \underset{\varepsilon \to 0}{\sim} B_\varepsilon$, *i.e.* $|A_\varepsilon - B_\varepsilon| \to 0$ *as* $\varepsilon \to 0)$.

Proof : First we prove that

$$E_0^{(\varepsilon)}\ \phi_\varepsilon \exp\left[U_\varepsilon^* \Delta_\varepsilon + 0_\varepsilon(U_\varepsilon)\right] c(U_\varepsilon) \underset{\varepsilon \to 0}{\sim} E_0^{(\varepsilon)}\ \phi_\varepsilon \exp(U_\varepsilon^* \Delta_\varepsilon)\ c(U_\varepsilon) . \tag{3.10}$$

The absolute value of the difference of these two terms is bounded by :

$$L\ E_0^{(\varepsilon)} \exp(U_\varepsilon^* \Delta_\varepsilon)\ c(U_\varepsilon)\ |1 - \exp(0_\varepsilon(U_\varepsilon))| . \tag{3.11}$$

Law $(\Delta_\varepsilon \mid P_0^{(\varepsilon)}) = N(0,F)$ so

$$E_0^{(\varepsilon)} \exp(U_\varepsilon^* \Delta_\varepsilon)\ c(U_\varepsilon) = E_0^{(\varepsilon)} \exp(U_\varepsilon^* \Delta_\varepsilon - \tfrac{1}{2}\ U^* F U) = 1 .$$

Hence, we can define a probability law $\hat{P}_0^{(\varepsilon)}$ by :

$$d\hat{P}_0^{(\varepsilon)} = \exp(U_\varepsilon^* \Delta_\varepsilon - \tfrac{1}{2}\ U_\varepsilon^* F U_\varepsilon)\ dP_0^{(\varepsilon)} .$$

Using $\hat{P}_0^{(\varepsilon)}$, (3.11) is equal to

$$L \cdot \hat{E}_0^{(\varepsilon)}\ |1 - \exp\left[0_\varepsilon(U_\varepsilon)\right]| . \tag{3.12}$$

$\exp\left[0_\varepsilon(U_\varepsilon)\right] \geqslant 0$, so a sufficient condition for (3.12) tending to 0, is :

$$\hat{E}_0^{(\varepsilon)} \exp\left[0_\varepsilon(U_\varepsilon)\right] = 1 \quad (\forall\ \varepsilon) , \tag{3.13}$$

$$\hat{P}_0^{(\varepsilon)} - \lim_{\varepsilon \to 0} \exp\left[0_\varepsilon(U_\varepsilon)\right] = 1 . \tag{3.14}$$

(3.13) follows from the definition of $\hat{P}_0^{(\varepsilon)}$ and $Z_{\varepsilon,\theta_0}(U_\varepsilon)$ (cf. Theorem 3.1). As $P_0^{(\varepsilon)} - \hat{P}_0^{(\varepsilon)}$, the convergence (3.14) is equivalent to :

$$P_0^{(\varepsilon)} - \lim_{\varepsilon \to 0} \exp\left(0_\varepsilon(U_\varepsilon)\right) = 1 ,$$

which follows from Theorem 3.1 : Thus (3.10) is proved. Finally, (3.10) implies that

$$E_{U_\varepsilon}^{(\varepsilon)} \phi_\varepsilon = E_0^{(\varepsilon)} \phi_\varepsilon \exp\left(U_\varepsilon^* \Delta_\varepsilon + 0_\varepsilon(U_\varepsilon)\right) c(U_\varepsilon)$$

$$\underset{\varepsilon \to 0}{\sim} E_0^{(\varepsilon)} \phi_\varepsilon \exp(U_\varepsilon^* \Delta_\varepsilon) c(U_\varepsilon)$$

$$= E_0^{(\varepsilon)} E_0^{(\varepsilon)} [\phi_\varepsilon \mid \Delta_\varepsilon^1, \Delta_\varepsilon^2] \exp(U_\varepsilon^* \Delta_\varepsilon) c(U_\varepsilon)$$

$$= E_0^{(\varepsilon)} \psi_\varepsilon(\Delta_\varepsilon^1, \Delta_\varepsilon^2) \exp(U_\varepsilon^* \Delta_\varepsilon) c(U_\varepsilon)$$

$$= \tilde{E}_0 \psi_\varepsilon(\Delta^1, \Delta^2) \exp(U_\varepsilon^* \Delta) c(U_\varepsilon) \qquad \text{(with notations 3.4)}$$

$$= \tilde{E}_{U_\varepsilon} \psi_\varepsilon(\Delta^1, \Delta^2) .$$

□

Proof of Theorem 3.3. : We want to prove that

1. $\{\phi_\varepsilon^*\}$ is asymptotically locally (a.l.) unbiased of level α, i.e.

$$\overline{\lim_{\varepsilon \to 0}} \sup_{U \in D_{\varepsilon,K}^0} E_U^{(\varepsilon)} \phi_\varepsilon^* \leq \alpha \leq \underline{\lim_{\varepsilon \to 0}} \inf_{U \in D_{\varepsilon,K}^1} E_U^{(\varepsilon)} \phi_\varepsilon^* \quad (\forall K > 0). \tag{3.15}$$

2. For any test-sequence $\{\Phi_\varepsilon\}$ a.l. unbiased of level α, we have :

$$\underline{\lim_{\varepsilon \to 0}} \inf_{U \in D_{\varepsilon,K}^1} E_U^{(\varepsilon)} (\Phi_\varepsilon^* - \Phi_\varepsilon) \geq 0 \qquad (\forall K > 0) . \tag{3.16}$$

Let us first prove the second statement. Let $\{\Phi_\varepsilon\}$ be an a.l. unbiased test-sequence of level α. From the continuity of the power function $U \to E_U^{(\varepsilon)} \Phi_\varepsilon$ and (3.14), it is not difficult to prove that for any bounded sequence $\{v_\varepsilon\}$ in $\mathbb{R}^n$,

$$\lim_{\varepsilon \to 0} E_{0,v_\varepsilon}^{(\varepsilon)} \Phi_\varepsilon = \alpha . .$$

($E_{u,v}^{(\varepsilon)}$ stands for $E_U^{(\varepsilon)}$ with $U = \begin{pmatrix} u \\ v \end{pmatrix}$).

Hence, if $\Psi_\varepsilon(\delta^1, \delta^2) := E_0^{(\varepsilon)} [\Phi_\varepsilon \mid \Delta_\varepsilon^1 = \delta^1, \Delta_\varepsilon^2 = \delta^2]$, it follows from Lemma 3.5 :

$$\lim_{\varepsilon\to 0} \tilde{E}_{0,v_\varepsilon} \Psi_\varepsilon(\Delta^1,\Delta^2) = \alpha ,$$

($\tilde{E}_{u,v}$ stands for $\tilde{E}_U$ with $U = \begin{pmatrix} u \\ v \end{pmatrix}$). Thus for any bounded sequence $\{v_\varepsilon\}$ we deduce that :

$$\lim_{\varepsilon\to 0} \int \left\{\tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2] - \alpha\right\} \exp(v_\varepsilon^*\delta^2)\, N^2(d\delta^2) = 0 . \qquad (3.17)$$

We now consider a sequence $\{U_\varepsilon\}$, $U_\varepsilon \in \overline{D^1_{\varepsilon,K}}$, such that :

$$E^{(\varepsilon)}_{U_\varepsilon} (\Phi^*_\varepsilon - \Phi_\varepsilon) = \inf \{E^{(\varepsilon)}_U (\Phi^*_\varepsilon - \Phi_\varepsilon) \ ; \ U \in D^1_{\varepsilon,K}\} .$$

From Lemma 3.5, we get :

$$E^{(\varepsilon)}_{U_\varepsilon} (\Phi^*_\varepsilon - \Phi_\varepsilon) \underset{\varepsilon\to 0}{\sim} \tilde{E}_{U_\varepsilon}(\Phi^* - \Psi_\varepsilon)(\Delta^1,\Delta^2)$$

$$\geqslant \inf_{U\in D^1_{\varepsilon,K}} \tilde{E}_U(\Phi^* - \Psi_\varepsilon)(\Delta^1,\Delta^2) .$$

So in order to establish (3.16), it is enough to prove :

$$\underline{\lim}_{\varepsilon\to 0} \inf_{U\in D^1_{\varepsilon,K}} \tilde{E}_U(\Phi^* - \Psi_\varepsilon)(\Delta^1,\Delta^2) \geqslant 0 \qquad (\forall\, K > 0) . \qquad (3.18)$$

Let $u < 0$,

$$\int \left(\Phi^*(\delta^1,\delta^2) - \Psi_\varepsilon(\delta^1,\delta^2)\right) \left(\exp(u\delta^1) - \exp(u\rho(\delta^2))\right) c_{\delta^2}(u)\, N^1(d\delta^1 \mid \delta^2)$$

$$= \int_{\delta^1\leq\rho(\delta^2)} \left(1 - \Psi_\varepsilon(\delta^1,\delta^2)\right) \left(\exp(u\delta^1) - \exp(u\rho(\delta^2))\right) c_{\delta^2}(u)\, N^1(d\delta^1 \mid \delta^2)$$

$$+ \int_{\delta^1>\rho(\delta^2)} \left(0 - \Psi_\varepsilon(\delta^1,\delta^2)\right) \left(\exp(u\delta^1) - \exp(u\rho(\delta^2))\right) c_{\delta^2}(u)\, N^1(d\delta^1 \mid \delta^2)$$

$$\geqslant 0 .$$

This relation together with (3.9) implies that

$$\tilde{E}_U [(\Phi^* - \Psi_\varepsilon)(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]$$

$$\geqslant \exp(u\rho(\delta^2))\, c_{\delta^2}(u)\, \tilde{E}_0[(\Phi^* - \Psi_\varepsilon)(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]$$

$$= \exp(u\rho(\delta^2))\, c_{\delta^2}(u) \left\{\alpha - \tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\right\} .$$

So, we have

$$\inf_{U \in D^1_{\varepsilon,K}} \tilde{E}_U(\Phi^* - \Psi_\varepsilon)(\Delta^1,\Delta^2)$$

$$= \inf_{U \in D^1_{\varepsilon,K}} \int \tilde{E}_U[(\Phi^* - \Psi_\varepsilon)(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\, N^2_U(d\delta^2)$$

$$\geq \inf_{U \in D^1_{\varepsilon,K}} \int \left\{\alpha - \tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\right\} \exp(u\rho(\delta^2))\; c_{\delta^2}(u)\; N^2_U(d\delta^2)$$

$$= \inf_{U \in D^1_{\varepsilon,K}} \int \left\{\alpha - \tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\right\} \exp(u\rho(\delta^2) + v^*\delta^2))\; c(U)\; N^2(d\delta^2)$$

$$= \int \left\{\alpha - \tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\right\} \exp(\tilde{u}_\varepsilon\rho(\delta^2) + \tilde{v}^*_\varepsilon\delta^2)\; c(\tilde{U}_\varepsilon)\; N^2(d\delta^2)$$

for some bounded sequence $\{\tilde{U}_\varepsilon\}$.

Lemma 3.2 implies that $\exp(\tilde{u}_\varepsilon\rho(\delta^2) + \tilde{v}^*_\varepsilon\delta^2)$ is of the form $\exp(a_\varepsilon + b^*_\varepsilon\delta^2)$, where both $\{a_\varepsilon\}$ and $\{b_\varepsilon\}$ are bounded sequences. So we have the following relations :

$$\varliminf_{\varepsilon \to 0} \inf_{U \in D^1_{\varepsilon,K}} \tilde{E}_U(\Phi^* - \Psi_\varepsilon)(\Delta^1,\Delta^2)$$

$$\geq \varliminf_{\varepsilon \to 0} \exp(a_\varepsilon) \int \left\{\alpha - \tilde{E}_0[\Psi_\varepsilon(\Delta^1,\delta^2) \mid \Delta^2 = \delta^2]\right\} \exp(b^*_\varepsilon\delta^2)\; N^2(d\delta^2)$$

$$= 0 ,$$

where the last equality follows from (3.17). Relation (3.18) (and (3.16)) is thus verified.

Using a similar argument, we can prove that $\{\Phi^*_\varepsilon\}$ is a.1. of level α, more precisely

$$\lim_{\varepsilon \to 0} \sup_{U \in D^0_{\varepsilon,K}} E^{(\varepsilon)}_U\, \Phi^*_\varepsilon = \alpha \qquad (\forall\, K > 0) .$$

Taking $\phi_\varepsilon = \alpha$ a.e. in (3.16), implies that $\{\phi^*_\varepsilon\}$ is a.1. unbiased. So the desired relations are proved.

□

3.3. Properties of the Maximum Likelihood Estimator (MLE)

Having detected a failure, we want to estimate the parameter θ. The problem is related to the estimation of diffusion process parameters (cf. [10, 12]). We consider the MLE (with the notations of the beginning of Section 3) :

$$\hat{\theta}_\varepsilon \in \operatorname{Arg} \max_{\theta \in \Theta^1} dP_\theta^{(\varepsilon)} / dP_0 ,$$

where $\Theta^1 = \left\{ \begin{pmatrix} \tau \\ \gamma \end{pmatrix} ; 0 \leqslant \tau < \tau_0 , \gamma \neq 0 \right\}$. We make the following assumptions :

(A1) $\quad \forall\, \theta, \theta' \in \Theta^1, \quad \theta \neq \theta', \quad \|S(\theta) - S(\theta')\| > 0 ;$

(A2) $\quad \forall\, K \subset \Theta^1$ compact, $\exists\, \alpha_K, \beta_K > 0$ s.t. $\alpha_K I \leqslant F(\theta) \leqslant \beta_K I, \forall\, \theta \in K ;$

where $\|S(\theta)\| := \left(\int_0^1 |S(\theta,t)|^2 \, dt \right)^{\frac{1}{2}}$ and I is the unit matrix. An example of sufficient condition for (A1), (A2) to hold can be found in Campillo [3].

We can easily check that $\hat{\theta}_\varepsilon = \begin{pmatrix} \hat{\tau}_\varepsilon \\ \hat{\gamma}_\varepsilon \end{pmatrix}$ is well defined $\left((\varepsilon,\theta) \to dP_\theta^{(\varepsilon)}/dP_0 \right.$ admits a continuous version$\left.\right)$ and :

$$\hat{\tau}_\varepsilon \in \operatorname{Arg} \max_{0 \le \tau < \tau_0} \left(\int_0^1 (S_t^\tau)^* d\nu_t \right)^* \left(\int_0^1 (S_t^\tau)^* S_t^\tau dt \right)^{-1} \int_0^1 (S_t^\tau)^* \, d\nu_t$$

$$\hat{\gamma}_\varepsilon = \left\{ \left(\int_0^1 (S_t^\tau)^* S_t^\tau \, dt \right)^{-1} \int_0^1 (S_t^\tau)^* \, d\nu_t \right\} \Big|_{\tau = \hat{\tau}_\varepsilon} .$$

The definition of $F(\theta)$ together with (A2) imply that $\int_0^1 (S_t^\tau)^* S_t^\tau \, dt$ is a regular matrix $(0 \leqslant \tau \leqslant \tau_0)$.

We consider the following conditions : for any compact set K included in Θ^1 .

(C1) $\quad \sup_{\theta^1, \theta^2 \in K} |F(\theta^1)^{-\frac{1}{2}} F(\theta^2) F(\theta^1)^{-\frac{1}{2}}| < \infty ;$

(C2) the function $S(\theta,t)$ is continuously differentiable with respect to θ in $L^2(0,1)$;

(C3) $\quad \eta_\varepsilon^1(K) = \sup_{\theta \in K} \sup_{\substack{|U| < \varepsilon^{-\frac{1}{2}} \\ U \in D_{\varepsilon,\theta}(K)}} |U|^{-2} \, \| \{ \nabla(\theta + \varepsilon F(\theta)^{-\frac{1}{2}} U) - \nabla(\theta) \} \, F(\theta)^{-\frac{1}{2}} U \|^2 \to 0$ as $\varepsilon \to 0$;

(C4) $$\eta_\varepsilon^2(K) = \inf_{\theta\in K} \inf_{\substack{|U|>\varepsilon^{-\frac12} \\ U\in D_{\varepsilon,\theta}(K)}} |F(\theta)^{\frac12}|^{-1} \|S(\theta+\varepsilon F(\theta)^{-\frac12}U) - S(\theta)\| > 0 \ ;$$

where $D_{\varepsilon,\theta}(K) = \left\{U = \begin{pmatrix} u \\ v \end{pmatrix} ; \theta + \varepsilon F(\theta)^{-\frac12} , u \in K\right\}$. Then we have the following result (cf. Ibragimov-Has'Minskii [10], th. 5.1, p. 203) :

<u>THEOREM 3.6.</u> *If the conditions* (C1) - (C4) *are fulfilled, then for any compact set* K *in* Θ^1, *the following assertions are valid :*

(i) The MLE $\hat\theta_\varepsilon$ *is consistent ;*

(ii) The MLE $\hat\theta_\varepsilon$ *is asymptotically normal :*

$$\text{law}\left(\frac{1}{\varepsilon}(\hat\theta_\varepsilon - \theta) \mid P_\theta^{(\varepsilon)}\right) \underset{\varepsilon\to 0}{\to} N\left(0, F(\theta)^{-\frac12}\right) ;$$

(iii) All the moments of the random variable $(\hat\theta_\varepsilon - \theta)/\varepsilon$ *converge as* $\varepsilon \to 0$ *to the corresponding moments of the normal distribution* $N\left(0, F(\theta)^{-\frac12}\right)$.

□

Under the assumptions (A1), (A2), we now prove the conditions (C1) - (C4). Let K be a compact set in Θ^1.

(<u>C1</u>) $$|F(\theta^1)^{-\frac12} F(\theta^2) F(\theta^1)^{-\frac12}| \leq |F(\theta^2)| \cdot |F(\theta^1)^{-\frac12}|^2 \leq \beta_K / \alpha_K .$$

(<u>C2</u>) This condition follows from the fact that there exists a constant C_k which depends only on K such that the inequality :

$$\|\nabla(\theta^1) - \nabla(\theta^2)\|^2 \leq C_k\left(|\theta^1 - \theta^2| + |\theta^1 - \theta^2|^2\right)$$

is valid for any $\theta^1, \theta^2 \in K$. This result can be proved without difficulty.

(<u>C3</u>) $$\eta_\varepsilon^1(K) \leq \sup_{\theta\in K} \sup_{\substack{|U|<\varepsilon^{-\frac12} \\ U\in D_{\varepsilon,\theta}(K)}} |U|^{-2} |F(\theta)^{-\frac12}|^2 |U|^2 \|\nabla\left(\theta + \varepsilon F(\theta)^{-\frac12}U\right) - \nabla(\theta)\|^2$$

$$\leq C_K^2 \alpha_K^{-1} \sup_{\theta\in K} \sup_{\substack{|U|<\varepsilon^{-\frac12} \\ U\in D_{\varepsilon,\theta}(K)}} |\varepsilon F(\theta)^{-\frac12}U|$$

$$\leq C_K^2\, \alpha_K^{-1} \sup_{\theta\in K} \sup_{\substack{|U|<\varepsilon^{-\frac{1}{2}} \\ U\in D_{\varepsilon,\theta}(K)}} (\varepsilon/\alpha_K)^{\frac{1}{2}} \underset{\varepsilon\to 0}{\to} 0 .$$

(C4) $$\eta_\varepsilon^2(K) \geq \beta_K^{-\frac{1}{2}} \inf_{\substack{\theta^1,\theta^2\in K \\ |\theta^1-\theta^2|\geq\varepsilon\beta_K}} \| S(\theta^1) - S(\theta^2)\| .$$

The function $(\theta^1,\theta^2) \to \|S(\theta^1) - S(\theta^2)\|$ is continuous on the compact set $\{(\theta^1,\theta^2) \in K \times K ; |\theta^1 - \theta^2| \geq \varepsilon\beta_K\}$, so we can find $\theta^1(\varepsilon,K) \neq \theta^2(\varepsilon,K)$ such that :

$$\inf_{\substack{\theta^1,\theta^2\in K \\ |\theta^1-\theta^2|\geq\varepsilon\beta_K}} \|S(\theta^1) - S(\theta^2)\| = \|S\left(\theta^1(\varepsilon,K)\right) - S\left(\theta^2(\varepsilon,K)\right)\| ,$$

and it follows from (A1) that this last term is strictly positive.

□

*

* *

REFERENCES

[1] Basseville, M. (1982), "Contributions à la détection séquentielle de ruptures de modèles statistiques". Thèse d'Etat, Université de Rennes.

[2] Basseville, M. and A. Benveniste (1981), "An example of failure detection : design and comparative study of some algorithms". Rapport INRIA n° 73.

[3] Campillo, F. (1984), "Filtrage et détection de ruptures de processus partiellement observés". Thèse de troisième cycle, Université de Provence.

[4] Chibisov, D.M. (1973), "Asymptotic expansions for Neyman's C(α)-test", in *Proceedings of the Second Japan-USSR Symposium of Probabilitu Theory,* G. Maryama, Yu. Prokhorov, eds., Lecture Notes in Mathematics, 330, 16-45.

[5] Deshayes, J. and D. Picard (1982), "Test de rupture de régression : comparaison asymptotique", *Teoryia Ver. Prim.,* 95-108.

[6] Deshayes, J. and D. Picard (1983), "Ruptures de modèles en statistique". Thèse d'Etat, Orsay.

[7] Deshayes, J. and D. Picard (1984), "Principe d'invariance sur les processus de vraisemblance", *Annales de l'Institut Henri Poincaré,* 20(1), 1-20.

[8] Deshayes, J. and D. Picard (1984), "Méthodes globales de test et d'estimation de ruptures : point de vue asymptotique". Séminaire "Détection de ruptures dans les modèles dynamiques de signaux et systèmes", CNRS (RCP567), Paris.

[9] Deshayes, J. and D. Picard (1984), "Comment utiliser les statistiques de vraisemblance dans les séries chronologiques". Séminaire "Détection de ruptures dans les modèles dynamiques de signaux et systèmes", CNRS (RCP567), Paris.

[10] Ibragimov, I.A. and R.Z. Has'Minskii (1981), *Statistical Estimation Asymptotic Theory,* Springer Verlag, New York.

[11] Kligene, K. and L. Tel'Ksnis (1984), "Methods of detecting instants of change of random process properties", *Automation and Remote Control,* traduit de *Automatika i Telemekhanika,* 10(1983), 5-56.

[12] Kutoyants, Yu.A. (1984), *Parameter Estimation for Stochastic Processes,* Heldermann Verlag, Berlin.

[13] LeCam, L. (1969), "Théorie asymptotique de la décision statistique", Université de Montréal.

[14] Lehman, E.L. (1959), *Testing Statistical Hypotheses,* Wiley, New York.

[15] Liptser, R.S. and A.N. Shiryayev (1977), *Statistics of Random Processes,* Springer Verlag, New York.

[16] Michel, R. (1979), "On the asymptotic efficiency of conditional tests for exponential families", *The Annals of Statistics,* 7(6), 1256-1263.

[17] Willsky, A.S. (1976), "A survey of design method for failure detection in dynamic systems", *Automatica,* 12, 601-611.

[18] Willsky, A.S. and H.L. Jones (1974), "A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes", *Proceedings of the IEEE Conference on Decision and Control.*