



**HAL**  
open science

## Time varying extremes for monitoring aquatic biosensors

Gilles Durrieu, Ion Grama, Quang-Khoai Pham, Jean-Marie Tricot

► **To cite this version:**

Gilles Durrieu, Ion Grama, Quang-Khoai Pham, Jean-Marie Tricot. Time varying extremes for monitoring aquatic biosensors. 2013. hal-00918833

**HAL Id: hal-00918833**

**<https://hal.science/hal-00918833>**

Preprint submitted on 16 Dec 2013

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

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

# TIME VARYING EXTREMES FOR MONITORING AQUATIC BIOSENSORS

DURRIEU, G., GRAMA, I., PHAM, Q. K., AND TRICOT, J. M.

Université de Bretagne Sud, LMBA, UMR CNRS 6205, F-56000 Vannes, France.

ABSTRACT. Measurement of mollusks bivalves activity is a way to record the animal behavior and so to evaluate possible changes in the water quality. In the framework of ecological time series data at times  $0 < t_1 < \dots < t_n \leq T$ , we observe independent observations  $X_{t_1}, \dots, X_{t_n}$  where each  $X_{t_i}$  is distributed according to the distribution function  $F_{t_i}$ . For each  $t \in [0, T]$ , we propose a non parametric adaptive estimator for tail probabilities and extreme quantiles of  $F_t$ . The idea of our approach is to adjust the tail of the distribution function  $F_t$  with a Pareto distribution with parameter  $\theta_{t,\tau}$  starting from a threshold  $\tau$ . The parameter  $\theta_{t,\tau}$  is estimated using a non parametric kernel estimator of bandwidth  $h$  based on the observations larger than  $\tau$ . Under some regularity assumptions, we prove that the proposed adaptive estimator of  $\theta_{t,\tau}$  is consistent and we determine its rate of convergence. We also propose a sequential testing based procedure for the automatic choice of the threshold  $\tau$  when the bandwidth  $h$  is fixed. Finally, we study the properties of this procedure by simulation and on real data set to estimate global changes (pollution, temperature change) and so to help in the survey of aquatic systems.

## 1. INTRODUCTION

Protection of the aquatic environment is a top priority for marine managers, policy makers, and the general public. Human activities are responsible for significant discharges of pollutants into environment. These pollutants lead to the degradation of many habitats disturbing ecosystems and also causing problems in terms of public health. Surveillance and protection of aquatic systems is thus fundamental and it is of great interest to be able to inform in real time people of water conditions. Due to an increasing interest in the health of aquatic systems, there is a compelling need for the use of remote online sensors to instantly and widely distribute information on a daily basis. Regulations and controls

on water quality have already been established. Among these controls, bioindicators are increasingly used because they can be effective in their ability to reveal the presence of traces (very low concentrations) of contaminants through accumulation in tissues of aquatic animals (see [Tran et al., 2003, 2004, 2007](#)). In this paper, we focus on the activity of oysters. The ability of oyster to permanently “taste” their environment is one of the possible ways to monitor the quality of our coastal waters and read throughout the year the health of both the oysters and their environment.

The interest in investigating the bivalve’s activities by recording their valve movements (valvometry) has been explored in ecotoxicology. The basic idea of valvometry is to use the bivalve’s ability to close its shells when exposed to a contaminant as an alarm signal ([Doherty et al. \(1987\)](#), [Sow et al. \(2011\)](#), [Nagai et al. \(2006\)](#) among others). Nowadays, valvometers are available on the market and use the principle of electromagnetic induction [Sloff et al. \(1983\)](#); [Jenner et al. \(1989\)](#) such as the Mossel Monitor in [Kramer et al. \(1989\)](#) or the Dreissena Monitor in [Borcherding and Volpers \(1994\)](#). There has been a clear research interest in the recent years to measuring the bivalve’s behaviors directly in real conditions ([Robson et al. \(2007\)](#); [Tran et al. \(2003\)](#); [Sow et al. \(2011\)](#)).

These noninvasive valvometric techniques provide high-frequency data and different statistical models were proposed to study their behavior in their natural habitat and constantly monitor water quality when faced with stress such as a pollutant: valves can suddenly close or express abnormal movements indicating a change in water quality ([Sow et al. \(2011\)](#); [Schmitt et al. \(2011\)](#); [Jou and Liao \(2006\)](#); [Coudret et al. \(2013\)](#); [Azais et al.](#)). When faced with pollution or poor quality water, the oyster closes its shells, or in extreme situations the activity of animals can change dramatically after exposure to even very low levels of pollution.

Ecological and genetical studies often focus on average and median effects of environmental factors such as temperature, precipitation, salinity etc., but ecological dynamic is strongly affected by environmental extremes events [Denny et al. \(2009\)](#); [Durrieu and](#)

Briollais (2009). As long as extreme thresholds are not exceeded, the performance of individuals, the behavior of populations, and composition of communities are often well described by mean conditions (Brown et al. (2004)). However, biological thresholds are often exceeded in an environmental context (Gaines and Denny (1993)) such as an excursion of a climate variable like temperature outside of some extreme thresholds (Katz et al. (2005)). Extreme levels of characteristics variables can lead to impairment of function or outright mortality of individuals, with important implications for populations, communities, and ecosystems. For instance, extreme events play an important role in ecology (Gutschick and BassiriRad (2003)) influencing community dynamics and biodiversity (Altman and Whitlatch (2007); Dayton (1971); Connell (1978); Gross et al. (2005); Gutschick and BassiriRad (2003); Sousa (1979)). Extreme events can cause dramatic ecological change that recovery is not possible. Such effects arise when populations are pushed below some minimum density threshold (e.g., the Allee effect) or when a community or ecosystem enters an alternate stable state (Allee et al. (1949); Folke et al. (2004)). Mortalities of Pacific oysters during the summer months have been documented throughout the world and can affect between 10 and 50% of the juveniles (Samain and McCombie (2008)), with extreme cases involving  $> 90\%$  mortality (Burge et al. (2007)).

In this article, we propose a new statistical method for the estimation of extreme conditional probabilities and extreme conditional quantiles in the framework of time series data collected in an ecological study. The paper is organized as follows. Section 2 describes the model and estimators. Our approach is based on adjusting a Pareto correction to the weighted distribution function beyond a given threshold  $\tau$ . The asymptotic results of the estimators are stated in Section 3, with proofs given in the supplementary materials. We determine rates of convergence of the corresponding estimators of the parameters in the adjusted model when the threshold and the bandwidth are deterministic. As the threshold is unknown in practice, we propose in Section 4 a selection method based on the maximal propagation of the Pareto fit. The convergence results are then extended to the resulting

adaptive estimator. Section 5 illustrates the performance of the proposed procedure via simulation studies and an application to ecological data sets is provided in Section 6. The aim of these experiments is to propose a water quality monitoring system through the observation of extreme oyster's behavior. Finally, Section 7 concludes with a general discussion of our approach and its feasibility and applicability in practice.

## 2. MODEL AND ESTIMATOR

Let  $0 < T < \infty$  and  $(F_t)_{t \in [0, T]}$  be a family of distribution functions indexed in  $[0, T]$ . We observe independent random variables  $X_{t_1}, \dots, X_{t_n}$  associated to a sequence of times  $0 < t_1 < \dots < t_n \leq T$  where for each  $t_i$  the random variable  $X_{t_i}$  has a distribution function  $F_{t_i}$  supported on the interval  $[x_0, \infty)$ ,  $x_0 \geq 0$  with a strictly positive density  $f_{t_i}$ . Given  $x > x_0$  and  $p \in (0, 1)$ , the main aim is to provide a pointwise estimate of the tail probability  $S_t(x) = 1 - F_t(x)$  and the extreme  $p$ -quantile  $F_t^{-1}(p)$  processes on  $[0, T]$ .

The empirical survival function is routinely used to estimate  $S_t(x)$ , but this estimator does not provide a reliable estimation for large values of  $x$ , due to the lack of observations in this range. Otherwise, a parametric model is fitted to data and the values  $S_t(x)$  and  $F_t^{-1}(p)$  are inferred from the corresponding local fit. However this may cause a severe bias if the fitted model is misspecified. We combine the flexible empirical distribution function and the parametric fit in one model. The idea is to adjust, for some  $\tau \geq x_0$ , the excess distribution function

$$F_{t, \tau}(x) = 1 - \frac{1 - F_t(x)}{1 - F_t(\tau)}, \quad x \in [\tau, \infty)$$

by a parametric model which has good prediction properties. Here, we choose a Pareto distribution defined by

$$(2.1) \quad G_{\tau, \theta}(x) = 1 - \left(\frac{x}{\tau}\right)^{-\frac{1}{\theta}}, \quad x \in [\tau, \infty),$$

where  $\theta > 0$  is a parameter and  $\tau \geq x_0$  is a threshold value. Before the threshold  $\tau$  we estimate  $F_t$  by the empirical distribution function, while beyond  $\tau$  we use the adjusted probability  $G_{\tau,\theta}$  where  $\theta$  has to be estimated. This choice is justified by the extreme value theory: in the case when  $F_t$  belongs to the domain of attraction of the Fréchet distribution it is known that  $F_{t,\tau}(x)$  is approximated by the Pareto distribution  $G_{\tau,\theta}$  as  $\tau \rightarrow \infty$ .

We study the asymptotic properties of the weighted maximum quasi-likelihood estimator of  $\theta$  and propose a selection rule to determine the threshold  $\tau$  for a given value of  $t$ . Consider the semi-parametric model defined by

$$(2.2) \quad F_{t,\tau,\theta}(x) = \begin{cases} F_t(x) & \text{if } x \in [x_0, \tau], \\ 1 - (1 - F_t(\tau))(1 - G_{\tau,\theta}(x)) & \text{if } x > \tau. \end{cases}$$

Let  $K(\cdot)$  be a kernel function assumed to be continuous, non-negative, symmetric with support on the real line such that  $K(x) \leq 1$  and define the weights  $W_{t,h}(t_i) = K\left(\frac{t_i - t}{h}\right)$ , where  $h > 0$  is a bandwidth parameter. For any  $t \in [0, T]$ , the weighted quasi-log-likelihood function is

$$\begin{aligned} \mathcal{L}_{t,h}(\tau, \theta) &= \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t,\tau,\theta}}{dx}(X_{t_i}) \\ &= \sum_{i=1}^n 1_{\{X_{t_i} \leq \tau\}} W_{t,h}(t_i) \log f_t(X_{t_i}) \\ &\quad + \sum_{i=1}^n 1_{\{X_{t_i} > \tau\}} W_{t,h}(t_i) \log \left( \frac{1}{\tau\theta} \left( \frac{X_{t_i}}{\tau} \right)^{-\frac{1}{\theta}-1} \right), \end{aligned}$$

where  $1_A$  takes the values 1 when the condition  $A$  is verified and 0 otherwise. Maximizing  $\mathcal{L}_{t,h}(\tau, \theta)$  with respect to  $\theta$ , we obtain the estimator

$$(2.3) \quad \hat{\theta}_{t,h,\tau} = \frac{1}{\hat{n}_{t,h,\tau}} \sum_{i=1}^n W_{t,h}(t_i) 1_{\{X_{t_i} > \tau\}} \log \left( \frac{X_{t_i}}{\tau} \right)$$

where

$$(2.4) \quad \hat{n}_{t,h,\tau} = \sum_{i=1}^n W_{t,h}(t_i) 1_{\{X_{t_i} > \tau\}}$$

is the weighted number of the observations beyond the threshold  $\tau$ . The distribution function  $F_t(x)$  at time  $t$  is then estimated by

$$(2.5) \quad \widehat{F}_{t,h,\tau}(x) = \begin{cases} \widehat{F}_{t,h}(x) & \text{if } x \in [x_0, \tau], \\ 1 - \left(1 - \widehat{F}_{t,h}(\tau)\right) \left(\frac{x}{\tau}\right)^{-\frac{1}{\widehat{\theta}_{t,h,\tau}}} & \text{if } x > \tau, \end{cases}$$

which combines the weighted empirical distribution function

$$\widehat{F}_{t,h}(x) = \frac{1}{\sum_{j=1}^n W_{t,h}(t_j)} \sum_{i=1}^n W_{t,h}(t_i) 1_{\{X_{t_i} \leq x\}}$$

and the fitted Pareto law. For any  $p \in (0, 1)$ , the estimator of the  $p$ -quantile of  $X_t$  is defined by

$$(2.6) \quad \widehat{q}_p(t) = \begin{cases} \widehat{F}_{t,h}^{-1}(p) & \text{if } p < \widehat{p}_\tau, \\ \tau \left(\frac{1-\widehat{p}_\tau}{1-p}\right)^{\widehat{\theta}_{t,h,\tau}} & \text{otherwise,} \end{cases}$$

where  $\widehat{p}_\tau = \widehat{F}_{t,h}(\tau)$ .

### 3. ASYMPTOTIC PROPERTIES

**3.1. Main results.** Let  $\mathcal{K}(P, Q) = \int \log \frac{dP}{dQ} dP$  be the Kullback-Leibler entropy between two equivalent measures  $P$  and  $Q$ . The  $\chi^2$  entropy between  $P$  and  $Q$  is defined by  $\chi^2(P, Q) = \int \frac{dP}{dQ} dP - 1$ . By Jensen's inequality we have  $\chi^2(P, Q) \geq 0$ . For any non-negative random variables  $A_n$  and  $B_n$ , the notation  $A_n = O_{\mathbb{P}}(B_n)$  as  $n \rightarrow \infty$  means that there exists a constant  $c > 0$  such that  $\mathbb{P}(A_n \leq cB_n) \rightarrow 1$  as  $n \rightarrow \infty$ .

For any  $t \in [0, T]$  denote

$$\bar{n}_{t,h,\tau} = \sum_{i=1}^n W_{t,h}(t_i) (1 - F_{t_i}(\tau)).$$

The number  $\bar{n}_{t,h,\tau}$  can be interpreted as the mean number of weighted observations exceeding the threshold  $\tau$  associated to  $t$ . If the kernel  $K$  has a finite support  $\text{supp}K$  then  $\bar{n}_{t,h,\tau}$  is the mean number of weighted observations  $X_{t_i}$  exceeding the threshold  $\tau$  with  $t_i \in \text{supp}K$ .

The main result of the paper is the following theorem which provides an oracle inequality for the estimator  $\widehat{\theta}_{t,h,\tau}$ .

**Theorem 3.1.** *Assume that  $\{\tau_n\}$  and  $\{h_n\}$  are two sequences such that  $\tau_n \geq x_0$  and*

$$(3.1) \quad \bar{n}_{t,h_n,\tau_n} \rightarrow \infty \text{ as } n \rightarrow \infty.$$

*Then, for any sequence of positive numbers  $\{\theta_n\}$ , we have as  $n \rightarrow \infty$ ,*

$$(3.2) \quad \mathcal{K}(\widehat{\theta}_{t,h_n,\tau_n}, \theta_n) = O_{\mathbb{P}} \left( \frac{\log n}{\bar{n}_{t,h_n,\tau_n}} + \frac{1}{\bar{n}_{t,h_n,\tau_n}} \sum_{i=1}^n W_{t,h_n}(t_i) \chi^2(F_{t_i}, F_{t_i,\tau_n,\theta_n}) \right).$$

*Proof.* See Section 8.2 in the supplementary materials. □

The first term in the bound (3.2) is referred as the stochastic error while the second one is the weighted square modelling bias induced by the use of the local parametric tail instead of the true one in the neighborhood of the estimation point  $t$ . In the case when  $K$  has the compact support  $[-1, 1]$ , the bias term can be bounded as follows:

$$\frac{1}{\bar{n}_{t,h_n,\tau_n}} \sum_{i=1}^n W_{t,h}(t_i) \chi^2(F_{t_i}, F_{t_i,\tau_n,\theta_n}) \leq \sup_{s \in [t-h, t+h]} \chi^2(F_{s,\tau_n}, G_{\tau_n,\theta_n}).$$

To ensure that the second term in the right hand side of (3.2) is at least of the same order as the first one, we shall assume that the family  $(F_t)_{t \in [0, T]}$  satisfies the following small modeling bias condition (cf. Spokoiny (2009)):

**C1.** For any  $t \in [0, T]$  there exist sequences  $\{\theta_n\}$ ,  $\{\tau_n\}$  and  $\{h_n\}$  (generally depending on  $t$ ) such that

$$(3.3) \quad \sum_{i=1}^n W_{t,h_n}(t_i) \chi^2(F_{t_i}, F_{t_i,\tau_n,\theta_n}) = O(\log n) \text{ as } n \rightarrow \infty.$$

The best approximation order in (3.2) is attained when the sequences  $\{\theta_n\}$ ,  $\{\tau_n\}$  and  $\{h_n\}$  are chosen such that

$$(3.4) \quad \sum_{i=1}^n W_{t,h_n}(t_i) \chi^2(F_{t_i}, F_{t_i,\tau_n,\theta_n}) \asymp \log n,$$



where  $a_n \asymp b_n$  means  $0 < c_1 \leq \frac{a_n}{b_n} \leq c_2 < \infty$ , for any  $n$  and some constants  $c_1$  and  $c_2$ . If (3.4) is satisfied, we say that  $\{\theta_n\}$  (respectively  $\{\tau_n\}$  and  $\{h_n\}$ ) is the oracle parameter (respectively oracle threshold and oracle bandwidth). We shall see in the next section that under some regularity assumptions, (3.4) is true for  $\theta_n = \theta_{t,\tau_n}$ , where

$$(3.5) \quad \theta_{t,\tau} = \arg \min_{\theta > 0} \mathcal{K}(F_{t,\tau}, G_{\tau,\theta})$$

is the best fitted Pareto parameter given by

$$(3.6) \quad \theta_{t,\tau} = \int_{\tau}^{\infty} \log \frac{x}{\tau} \frac{f_t(x) dx}{1 - F_t(\tau)}, \tau \geq x_0.$$

From Theorem 3.1, we have:

**Theorem 3.2.** *Assume that the family  $(F_t)_{t \in [0, T]}$  satisfies condition C1 and*

$$\bar{n}_{t, h_n, \tau_n} \rightarrow \infty \text{ as } n \rightarrow \infty.$$

*Then, we have,*

$$\mathcal{K}(\hat{\theta}_{t, h_n, \tau_n}, \theta_n) = O_{\mathbb{P}}\left(\frac{\log n}{\bar{n}_{t, h_n, \tau_n}}\right) \text{ as } n \rightarrow \infty.$$

*Proof.* This is a consequence of Theorem 3.1 and condition C1. □

The class of distributions  $(F_t)_{t \in [0, T]}$  satisfying condition C1 is very large. For instance, this is the case when  $(F_t)_{t \in [0, T]}$  is a time varying mixture of Pareto models or the time varying Hall model which generalizes the model studied in Hall (1982), Hall and Welsh (1984). Note that the Hall model includes the class of stable non-normal distributions. In the next section, we determine the explicit rate of convergence for these two models.

**3.2. Time varying Hall model.** The family  $(F_t)_{t \in (0, T]}$  is a time varying Hall model if there exists positive finite constants  $c_{\min}, c_{\max}, \gamma_{\min}, \gamma_{\max}, A_{\max} > 0$  and  $\rho > 0$  such that, for each  $t \in [0, T]$ , the distribution function  $F_t$  satisfies  $F_t(x_0) = 0$  and

$$(3.7) \quad f_t(x) = \frac{c_t}{\gamma_t} x^{-\frac{1}{\gamma_t} - 1} (1 + r_t(x)), \quad |r_t(x)| \leq A_t x^{-\frac{\rho}{\gamma_t}} \text{ as } x \rightarrow \infty,$$

where  $\gamma_t$ ,  $c_t$  and  $A_t$  are some time depending functions satisfying  $c_{\min} \leq c_t \leq c_{\max}$ ,  $\gamma_{\min} \leq \gamma_t \leq \gamma_{\max}$ ,  $A_t \leq A_{\max}$ .

For simplicity, we shall assume in this section that the kernel function  $K$  has the compact support  $[-1, 1]$ .

**Proposition 3.3.** *Assume the time varying Hall model given by (3.7). Let  $\theta_{t,\tau}$  be the best fitted Pareto parameter defined by (3.5). Suppose that there exist constants  $0 < \beta \leq 1$  and  $L > 0$ , such that for any  $0 \leq t, s \leq T$ ,*

$$(3.8) \quad |\gamma_t - \gamma_s| \leq L|t - s|^\beta.$$

Then the family  $(F_t)_{t \in [0, T]}$  verifies condition C1 with

$$h_n \asymp \left( \frac{\log n}{n} \right)^{\frac{1}{1+\beta(2+1/\rho)}}, \quad \tau_n \asymp \left( \frac{n}{\log n} \right)^{\frac{\gamma_t \beta / \rho}{1+\beta(2+1/\rho)}} \quad \text{and} \quad \theta_n = \theta_{t, \tau_n}.$$

*Proof.* See Section 8.3 in the supplementary materials. □

The following theorem gives an explicit rate of convergence of the estimator  $\widehat{\theta}_{t, h_n, \tau_n}$ .

**Theorem 3.4.** *Under the assumptions of Proposition 3.3, we have*

$$\sqrt{\mathcal{K} \left( \widehat{\theta}_{t, h_n, \tau_n}, \theta_{t, \tau_n} \right)} = O_{\mathbb{P}} \left( \left( \frac{\log n}{n} \right)^{\frac{\beta}{1+\beta(2+1/\rho)}} \right) \quad \text{as } n \rightarrow \infty.$$

*Proof.* This Theorem is a consequence of Proposition 3.3 and Theorem 3.2. □

**3.3. Mixture of two Pareto distributions.** We consider that  $F_t$  is a mixture of two Pareto distributions defined by

$$(3.9) \quad F_t(x) = C(1 - x^{-1/\gamma_t}) + (1 - C)(1 - x^{-1/\delta_t}), \quad x \geq 1,$$

where  $\delta_{\min} \leq \delta_t < \gamma_t \leq \gamma_{\max}$  and  $C \in (0, 1)$ .

As in the previous section, we shall assume that the kernel function  $K$  has the compact support  $[-1, 1]$ .

**Proposition 3.5.** *Assume the time varying mixture of two Pareto distributions given by (3.9). Suppose that there exist constants  $\beta \in (0, 1]$  and  $L > 0$ , such that for any*

$0 \leq t, s \leq T$ ,

$$|\gamma_t - \gamma_s| \leq L|t - s|^\beta,$$

and

$$|\delta_t - \delta_s| \leq L|t - s|^\beta.$$

Then the family  $(F_t)_{t \in [0, T]}$  verifies condition C1 with

$$(3.10) \quad h_n \asymp \left( \frac{\log n}{n} \right)^{\frac{1}{1+\beta(2+1/\rho_t)}}, \quad \tau_n \asymp \left( \frac{n}{\log n} \right)^{\frac{\gamma_t \beta / \rho_t}{1+\beta(2+1/\rho_t)}}$$

and

$$\theta_n = \theta_{t, \tau_n} = \frac{\gamma_t C \tau_n^{-1/\gamma_t} + \delta_t (1 - C) \tau_n^{-1/\delta_t}}{C \tau_n^{-1/\gamma_t} + (1 - C) \tau_n^{-1/\delta_t}},$$

where  $\rho_t = \frac{\gamma_t}{\delta_t} - 1 > 0$ .

*Proof.* See Section 8.4 in the supplementary materials. □

The next theorem gives the explicit rate of convergence of the estimator  $\widehat{\theta}_{t, h_n, \tau_n}$ .

**Theorem 3.6.** *Under the assumptions of Proposition 3.5, we have*

$$\sqrt{\mathcal{K} \left( \widehat{\theta}_{t, h_n, \tau_n}, \theta_{t, \tau_n} \right)} = O_{\mathbb{P}} \left( \left( \frac{\log n}{n} \right)^{\frac{\beta}{1+\beta(2+1/\rho_t)}} \right) \text{ as } n \rightarrow \infty.$$

*Proof.* This Theorem is a consequence of Proposition 3.5 and Theorem 3.2. □

If  $\rho_t \rightarrow \infty$  we obtain the rate  $\left( \frac{\log n}{n} \right)^{\frac{\beta}{2\beta+1}}$  which is the usual rate of convergence in the non-parametric estimation under the Lipschitz condition (see Stone (1982), Gardes and Girard (2008)). This shows that the rate of convergence provided by Theorems 3.1, 3.4 and 3.6 are exact up to a  $\log n$  multiple.

#### 4. AUTOMATIC SELECTION OF THE THRESHOLD $\tau$

**4.1. Maximum Likelihood Maximal Propagation (MLMP) procedure.** An important problem concerns the choice of the threshold  $\tau$ . Here, we propose a selection procedure to determine  $\tau$  assuming that the bandwidth  $h$  is fixed. The idea of the procedure has similarities with the propagation approach proposed by Spokoiny (2009) where a

sequence of likelihood ratio test is used to detect the maximal length local parametric fit for the tail. As soon as it is detected the next step of our procedure consists in maximizing the penalized likelihood. This second part is different from the approach in [Spokoiny \(2009\)](#) and is inspired by [Grama and Spokoiny \(2008\)](#), [Grama et al. \(2011\)](#) and [Grama et al. \(2013a,b\)](#).

Let  $Y_1 \geq \dots \geq Y_{M_{t,h}}$  be the order statistics pertaining to the observations  $\mathcal{Y}_{t,h} = \{X_{t_i} : \frac{t_i-t}{h} \in \text{supp}K\}$ , where  $\text{supp}K$  is the support of the kernel  $K$  and  $M_{t,h} = \text{card}(\mathcal{Y}_{t,h})$ . We choose the threshold  $\tau$  in the set  $\mathcal{Y}_{t,h}$  by maximizing the quasi-log-likelihood function

$$\max_{\theta} \mathcal{L}_{t,h}(\tau, \theta) - \text{Pen}_{t,h}(\tau, \hat{\theta}_{t,h,\hat{s}}) = \mathcal{L}_{t,h}(\tau, \hat{\theta}_{t,h,\tau}) - \text{Pen}_{t,h}(\tau, \hat{\theta}_{t,h,\hat{s}}),$$

where the penalty function is defined by

$$(4.1) \quad \text{Pen}_{t,h}(\tau, \theta) = \mathcal{L}_{t,h}(\tau, \theta)$$

and  $\hat{s}$  is a break time to be determined from a multiple goodness-of-fit testing procedure below.

Consider the null hypothesis  $\mathcal{H}_0(\tau) : F_t = F_{t,\tau,\theta}$  where the distribution function  $F_{t,\tau,\theta}$  is defined by [\(2.2\)](#) and the alternative hypothesis  $\mathcal{H}_1(s, \tau) : F_t = F_{t,\mu,s,\theta,\tau}$ , where  $F_{t,\mu,s,\theta,\tau}$  is the Pareto change point distribution

$$(4.2) \quad F_{t,\mu,s,\theta,\tau}(x) = \begin{cases} F_t(x) & \text{if } x \in [x_0, s], \\ 1 - (1 - F_t(s))(1 - G_{s,\mu}(x)) & \text{if } x \in (s, \tau), \\ 1 - (1 - F_t(s))(1 - G_{s,\mu}(\tau))(1 - G_{\tau,\theta}(x)) & \text{if } x \in (\tau, \infty), \end{cases}$$

where  $\mu, \theta > 0$  and  $x_0 \leq s \leq \tau$ . We proceed by consecutive testing for the null hypothesis  $\mathcal{H}_0(Y_k)$  against the alternatives  $\mathcal{H}_1(Y_k, Y_l)$ , for all  $k \in [k_0, M_{t,h}]$  and  $l$  such that  $\delta'(k-1) \leq l-1 \leq (1-\delta'')(k-1)$ , where  $k_0 = \delta_0 M_{t,h} \geq 3$  is a constant interpreted as the initial value of  $k$ , and  $\delta_0, \delta', \delta''$  are constants satisfying  $0 < \delta_0, \delta', \delta'' < \frac{1}{2}$ . The break time  $\hat{s}$  is the first time  $Y_k$  for which  $\mathcal{H}_0(Y_k)$  is rejected.

Recall that  $\widehat{\theta}_{t,h,\tau}$  is the maximum likelihood estimator of  $\theta$  given by (2.3). In the same way we obtain the maximum likelihood estimator of  $\mu$  :

$$\widehat{\mu}_{t,h,s,t} = \frac{\widehat{n}_{t,h,s}}{\widehat{n}_{t,h,s,\tau}} \widehat{\theta}_{t,h,s} - \frac{\widehat{n}_{t,h,\tau}}{\widehat{n}_{t,h,s,\tau}} \widehat{\theta}_{t,h,\tau},$$

where  $\widehat{n}_{t,h,s,\tau} = \sum_{i=1}^n W_{t,h}(t_i) 1_{\{s < X_{t_i} \leq \tau\}}$ . The log-likelihood ratio test statistic for testing  $\mathcal{H}_0(s)$  against  $\mathcal{H}_1(s, \tau)$  is given by

$$(4.3) \quad LR_{t,h}(s, \tau) = \widehat{n}_{t,h,s,\tau} \mathcal{K} \left( \widehat{\mu}_{t,h,s,\tau}, \widehat{\theta}_{t,h,s} \right) + \widehat{n}_{t,h,\tau} \mathcal{K} \left( \widehat{\theta}_{t,h,\tau}, \widehat{\theta}_{t,h,s} \right).$$

Taking into account (4.1), we have

$$\mathcal{L}_{t,h} \left( \tau, \widehat{\theta}_{t,h,\tau} \right) - \text{Pen}_{t,h}(\tau, \theta) = \widehat{n}_{t,h,\tau} \mathcal{K} \left( \widehat{\theta}_{t,h,\tau}, \theta \right),$$

which implies that the second term in (4.3) can be viewed as the penalized quasi-log-likelihood

$$\mathcal{L}_{t,h}^{\text{Pen}}(s, \tau) = \mathcal{L}_{t,h} \left( \tau, \widehat{\theta}_{t,h,\tau} \right) - \text{Pen}_{t,h} \left( \tau, \widehat{\theta}_{t,h,s} \right).$$

We denote by  $D > 0$  the critical value in the testing procedure below. To speed up the calculations, we take  $k = k_0 + i k_{step}$ ,  $i = 0, \dots, M_{grid}$ , where  $k_{step} = \lceil M_{t,h}/M_{grid} \rceil$  is an increment for  $k$  and  $M_{grid}$  is fixed. The values  $\delta_0, k_{step}, \delta', \delta''$  and  $D$  are the parameters of the procedure to be determined empirically.

The procedure of the adaptive choice of  $\tau$  is as follows:

**Step 1.** Set  $k = k_0$ .

**Step 2.** Compute the test statistic

$$\mathbb{Z}(Y_k) = \max_{\delta'(k-1) \leq l-1 \leq (1-\delta'')(k-1)} LR_{t,h}(Y_k, Y_l).$$

**Step 3.** If  $k \leq n - k_{step}$  and  $\mathbb{Z}(Y_k) \leq D$ , increase  $k$  by  $k_{step}$  and go to Step 2. If  $k > n - k_{step}$  or  $\mathbb{Z}(Y_k) > D$ , let  $\widehat{k} = k$ ,

$$(4.4) \quad \widehat{l} - 1 = \arg \max_{\delta'(k-1) \leq l-1 \leq (1-\delta'')(k-1)} \mathcal{L}_{t,h}^{\text{Pen}}(Y_{\widehat{k}}, Y_l),$$

take the adaptive threshold as  $\widehat{\tau}_n = Y_{\widehat{l}}$  and exit.

By definition, the adaptive estimator is set to  $\widehat{\theta}_{t,h,\widehat{\tau}_n}$ .

**4.2. Propagation property of the test statistic.** We shall prove that if  $Y_k \geq \tau_n$  then the test statistic

$$\mathbb{Z}(Y_k) = \sup_{\delta'(k-1) \leq l-1 \leq (1-\delta'')(k-1)} LR_{t,h_n}(Y_k, Y_l)$$

does not exceed

$$(4.5) \quad D = D(n) = c^* \log n$$

with high probability for some constant  $c^* > 0$ .

**Theorem 4.1.** *Assume that the family  $(F_t)_{t \in [0,T]}$  satisfies condition C1. Then, there exists a constant  $c^* > 0$  in (4.5), such that*

$$\mathbb{P} \left( \sup_{Y_k \geq \tau_n} \mathbb{Z}(Y_k) > D(n) \right) \leq \frac{4}{n} \text{ as } n \rightarrow \infty.$$

*Proof.* See Section 8.5 of the supplementary materials. □

Since  $P(\sup_{Y_k \geq \tau_n} \mathbb{Z}(Y_k) \leq D(n)) \leq P(\widehat{\tau}_n < \tau_n)$ , from Theorem 4.1 it follows that under condition C1,

$$\mathbb{P}(\widehat{\tau}_n < \tau_n) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

The meaning of this assertion is that the oracle threshold  $\tau_n$  is detected by our selection procedure with high probability as  $n \rightarrow \infty$ .

Since condition C1 means that  $F_t$  has a Pareto like tail on  $[\tau_n, \infty)$ , our selection procedure is equivalent to performing a goodness-of-fit test for testing the null hypothesis  $\mathcal{H}_0(\widehat{\tau}_n) : F_t = F_{t,\widehat{\tau}_n,\theta}$ .

**4.3. Rates of convergence of the adaptive estimator.** We first compare the performance of the adaptive estimator  $\widehat{\theta}_{t,h,\widehat{\tau}_n}$  with that of the non adaptive estimator  $\widehat{\theta}_{t,h_n,\tau_n}$ .

**Theorem 4.2.** *Assume that the family  $(F_t)_{t \in [0,T]}$  satisfies condition C1. Then, there exists a constant  $c^* > 0$  in 4.5, such that as  $n \rightarrow \infty$*

$$\mathcal{K}(\widehat{\theta}_{t,h_n,\widehat{\tau}_n}, \widehat{\theta}_{t,h_n,\tau_n}) = O_{\mathbb{P}}\left(\frac{\log n}{\bar{n}_{t,h_n,\tau_n}}\right).$$

*Proof.* See Section 8.6 of the supplementary materials. □

The previous theorem allows to extend the results of the non adaptive setting to the adaptive one. The following theorem gives the rate of convergence of the adaptive estimator  $\widehat{\theta}_{t,h_n,\widehat{\tau}_n}$ .

**Theorem 4.3.** *Assume that the family  $(F_t)_{t \in [0,T]}$  satisfies condition C1. Then, there exists a constant  $c^* > 0$  in 4.5, such that as  $n \rightarrow \infty$*

$$\mathcal{K}(\widehat{\theta}_{t,h_n,\widehat{\tau}_n}, \theta_n) = O_{\mathbb{P}}\left(\frac{\log n}{\bar{n}_{t,h_n,\tau_n}}\right).$$

*Proof.* Combining Theorem 4.2 and Theorem 3.1 we obtain Theorem 4.3. □

Recall that the adaptive estimator of the excess distribution function  $F_{t,\tau_n}$  is given by  $G_{\tau_n,\widehat{\theta}_{t,h_n,\widehat{\tau}_n}}$  (see 2.1). We give now the rate of convergence of the adaptive estimator  $G_{\tau_n,\widehat{\theta}_{t,h_n,\widehat{\tau}_n}}$  to  $F_{t,\tau_n}$  in terms of the Kullback-Leibler divergence.

**Theorem 4.4.** *Assume that the family  $(F_t)_{t \in [0,T]}$  satisfies condition C1 with  $\theta_n = \theta_{t,\tau_n}$ . Moreover, assume that as  $n \rightarrow \infty$*

$$(4.6) \quad \chi^2(F_{t,\tau_n}, G_{\tau_n,\theta_n}) = O\left(\frac{\log n}{\bar{n}_{t,h_n,\tau_n}}\right).$$

*Then, there exists a constant  $c^* > 0$  in 4.5, such that as  $n \rightarrow \infty$*

$$\mathcal{K}(F_{t,\tau_n}, G_{\tau_n,\widehat{\theta}_{t,h_n,\widehat{\tau}_n}}) = O_{\mathbb{P}}\left(\frac{\log n}{\bar{n}_{t,h_n,\tau_n}}\right).$$

*Proof.* See Section 8.7 of the supplementary materials. □

In the particular case of the Hall model we obtain an explicit rate of convergence of the adaptive estimator  $\widehat{\theta}_{t,h_n,\widehat{\tau}_n}$  (cf. Theorem 3.4).

**Theorem 4.5.** *Under the assumptions of Proposition 3.3, there exists a constant  $c^* > 0$  in 4.5, such that*

$$\sqrt{\mathcal{K}(\widehat{\theta}_{t,h_n,\widehat{\tau}_n}, \theta_{t,\tau_n})} = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{\frac{\beta}{1+\beta(2+1/\rho)}}\right) \text{ as } n \rightarrow \infty,$$

where  $h_n \asymp \left(\frac{\log n}{n}\right)^{\frac{1}{1+\beta(2+1/\rho)}}$ .

*Proof.* This Theorem is a consequence of Theorem 4.2 and Theorem 3.4.  $\square$

In the case of mixture of two Pareto distributions, we obtain similar rate of convergence for the adaptive estimator.

## 5. SIMULATIONS

We first give arguments on the choice of the parameters of the selection procedure given in Section 4. The proposed procedure depends on the initial proportion  $\delta_0$ , the parameters  $\delta'$ ,  $\delta''$ , the grid length  $M_{grid}$  and the critical value  $D$ . The parameters  $\delta_0$ ,  $\delta'$  and  $\delta''$  should be large enough to prevent from large variability in the first several iterations of the algorithm. We fix  $\delta_0 = \frac{1}{20}$ ,  $\delta' = \frac{1}{4}$  and  $\delta'' = \frac{1}{20}$ . We observe in the simulation that the procedure is not very sensitive to the choice of the parameter  $M_{grid}$ . We choose  $M_{grid} = 200$  to reduce the computation time.

According to the Wilks phenomenon under the null hypothesis which specifies that the distribution functions  $F_{t_i}$ ,  $i = 1, \dots, n$ , are i.i.d. Pareto  $G_{1,\theta}$ , the test statistic  $2LR_{t,h}(s, \tau)$  is asymptotically  $\chi^2$  with 1 degree of freedom. Our Theorem 4.1 can be regarded as a non asymptotic version of the Wilks phenomenon. Since our setting is non asymptotic, to determine practically the critical value  $D$ , we simulate the values of the test statistic  $T_n = \sup_{k_0 \leq k = k_0 + ik_{step} \leq n} \mathbb{Z}(X_{(k)})$  under the null hypothesis stated above. The value  $D$  is chosen as the 0.99-empirical quantile to ensure a 0.01 type I error. This is motivated



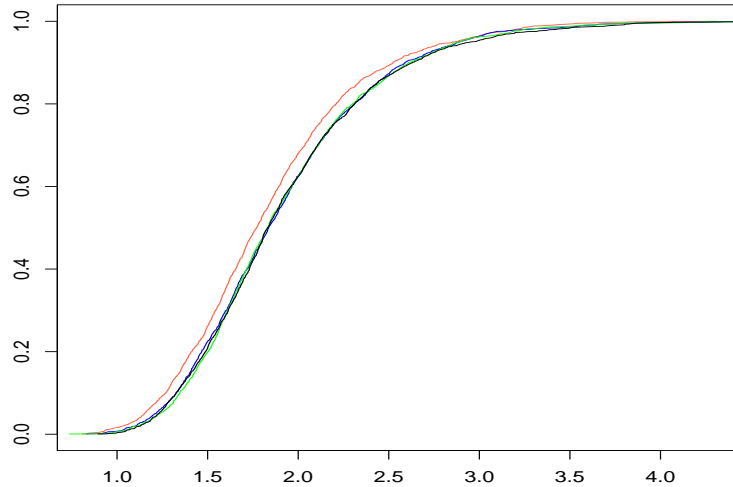


FIGURE 1. Empirical distribution function of the statistic  $T_n$  (red line:  $2nh = 200$ ; blue line:  $2nh = 500$ ; green line:  $2nh = 1000$ ; black line:  $2nh = 2000$ ).

by the propagation property of the test statistic under the null hypothesis: The selection procedure should choose the smallest possible threshold  $\tau = X_{(k)}$  (largest  $k$ ) i.e. the test statistic  $T_n$  should not exceed  $D$  with high probability. One can verify that the test statistic  $T_n$  does not depend on  $\theta$  and therefore we can fix  $\theta$  equal to 1.

The empirical distribution functions of  $T_n$  for the Gaussian kernel and various sample sizes in the window  $[t - h, t + h]$  is given in Figure 1. The value  $D = 3.6382$  corresponds to the empirical quantile of order  $p = 0.99$  from the sampled values.

We now analyze the behavior of the adaptive estimator under the alternative hypothesis. The performance of the proposed procedure is tested using the mixture model (see Section 3.3)

$$(5.1) \quad F_t(x) = C(1 - x^{-1/\gamma_t}) + (1 - C)(1 - x^{-1/\gamma_t - 5}), \quad x \geq 1, 0 \leq t \leq 1$$

where  $C = 0.75$  and

$$\gamma_t = 0.5 + 0.25 \sin(2\pi t).$$

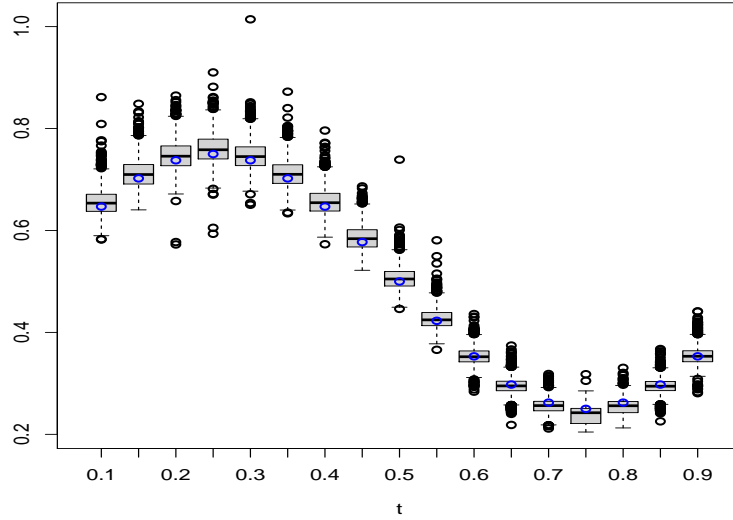


FIGURE 2. Boxplots of adaptive estimators  $\hat{\theta}_{t_i, h_n, \hat{\tau}_n}$  at  $t_i = 0.1, 0.15, \dots, 0.9$  for the Hall model from 2000 realizations with  $n = 50000$ .

We suppose that  $n = 50000$  and the value of  $h_n$  is fixed to 0.034. We generate  $B = 2000$  replicates of  $X_{t_1}, \dots, X_{t_n}$  from the mixture model (5.1). We focus on the estimation of  $\theta(t_i)$  for  $t_i = 0.1, 0.15, \dots, 0.9$ . The estimators  $\hat{\theta}_{t_i, h_n, \hat{\tau}_n}$  are computed using the adaptive procedure with the parameters  $\delta_0, \delta', \delta'', M_{grid}$  and  $D$  as fixed above. In Figure 2, we display the boxplots of 2000 realizations of  $\hat{\theta}_{t_i, h_n, \hat{\tau}_n}$ . Figures 3 and 4 show respectively the empirical mean square errors of  $\hat{\theta}_{t_i, h_n, \hat{\tau}_n}$  and  $\hat{q}_{0.999}(t_i)$ .

Our simulations also show that the threshold  $\tau_n = X_{(\hat{l})}$  in (4.4) is not very sensitive to the choice of the critical value  $D$  in the sense that the adaptive choice  $\hat{l}$  remains constant with respect to relatively large variations of  $D$ .

## 6. APPLICATION TO THE REAL DATA

We first describe the experimental site and the animal species. Then, we give some details on evaluation of valve activity. Afterwards, we provide information on data collection and transmission. Finally, exploring typical features of the valvometric environmental data samples (i.e. measurements of distances between the two parts of the shell of bivalves)

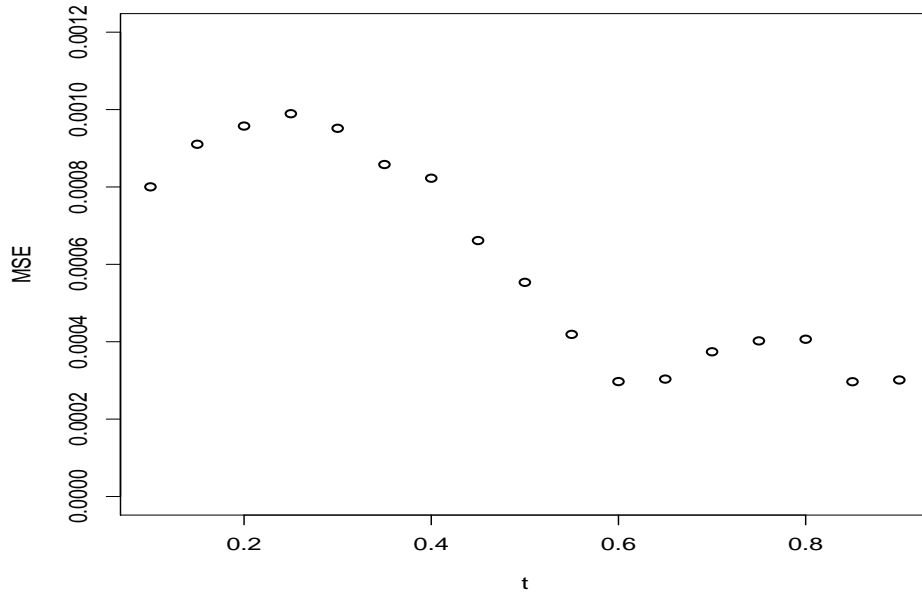


FIGURE 3. MSE of adaptive estimators  $\hat{\theta}_{t_i, h_n, \hat{\tau}_n}$  at  $t_i = 0.1, 0.15, \dots, 0.9$  for the Hall model from 2000 realizations with  $n = 50000$ .

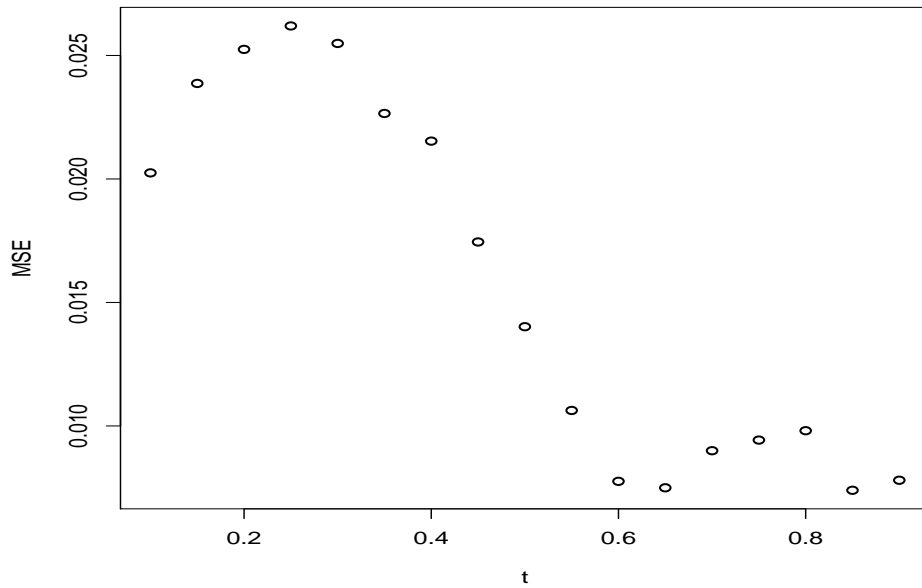


FIGURE 4. MSE of adaptive estimators  $\hat{q}_{0.999}(t_i)$  at  $t_i = 0.1, 0.15, \dots, 0.9$  for the Hall model from 2000 realizations with  $n = 50000$ .

collected by a laboratory called Environnements et Paléoenvironnements Océaniques et Continentaux (EPOC, <http://www.epoc.u-bordeaux.fr/>), we explain which inferences are valuable from the biological point of view.

**6.1. Data acquisition.** The monitoring site we considered is located in France at Locmariaquer (Latitude: 47°57 N, Longitude: 2°94 W). A group of sixteen Pacific oysters, *Crassostrea gigas*, measuring from 8 to 10 cm length, are installed on each site. Every oyster has almost the same age (1.5 years old). They were placed in a traditional oyster farmer bag.

The electronic principle of monitoring was described by [Tran et al. \(2003\)](#) and further modified by [Chambon et al. \(2007\)](#). Some information about these specific aspects can be found on <http://molluscan-eye.epoc.u-bordeaux1.fr>. The main challenge was to ensure the complete autonomy of the system without in-situ human intervention for at least one full year. In brief, each animal is equipped with two light coils (sensors), of approximately 53 mg each (unembedded), fixed on the edge of each valve. These coils measure 2.5×2.5×2 mm and were coated with a resin sealing before fixation on the valves. One of the coils emits a high-frequency, sinusoidal signal which is received by the other coil. For each sixteen animals, one measurement is received every 0.1s (10 Hz). This means that each animal's behavior is measured every 1.6s. Every day, a data set with 864,000 pairs of values (1 distance value, 1 stamped time value) is generated. A first electronic card manages the electrodes and is in a waterproof case next to the animals. A second electronic card handling the data acquisition and the programmed emission is also in the field but outside the water on a pier. This unit is equipped with a GSM/GPRS modem and uses Linux operating system for driving the first control unit, managing the data storage, accessing the Internet, and transferring the data.

A self-developed software module runs on mobile phone technology. After each 24h period (or any other programmed period of time), the data collected are transmitted to a

remote central workstation server where they are stored in text files. Every day, files from each site are inserted in a SQL database. This database is accessible with the software R (R Development Core Team, 2012) or a text terminal, via Internet or directly from the storage server.

**6.2. The biological issue.** These measurements produce some characteristic features that can be examined in Figure 5. As argued in Tran et al. (2003, 2010); Sow et al. (2011); Coudret et al. (2013), pollution can affect the activity of oysters and in particular the shells opening and closing speed. For instance in an inhospitable extreme environment, oysters will close more rapidly its shells. Thus, detecting extreme changes of the closing speed can provide insights about the health of oysters and so give an insight about the water quality.

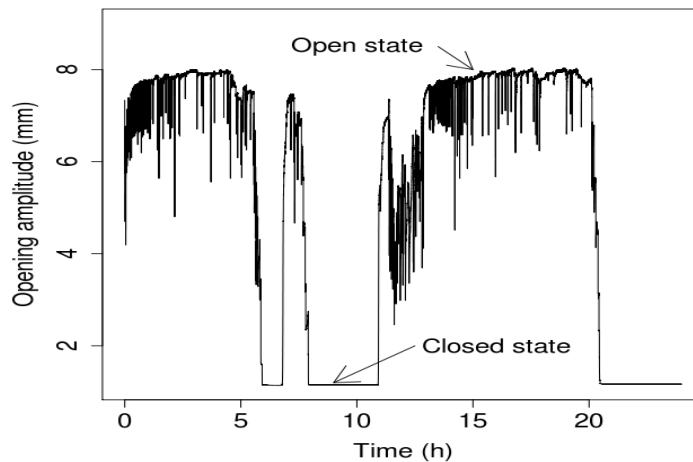


FIGURE 5. An example of valvometric data for one oyster

**6.3. Results.** We consider dataset associated to movement speeds which are considered as an indicator of the animal stress activity since its movements are associated to aquatic system perturbations. For instance, a stressed oyster in the presence of pollution or environmental perturbations exhibits irregular and numerous micro-closings and opening periods with high speed.

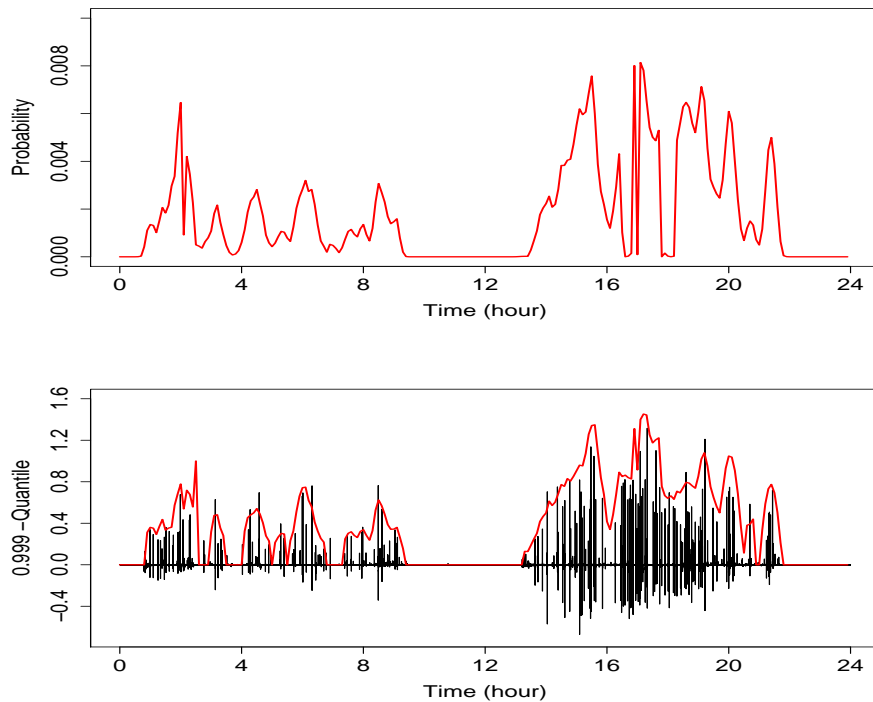


FIGURE 6. The red line on the top figure displays the estimated tail probabilities  $P(X_t > 0.3)$  on April 18, 2011. The red line on the bottom figure displays the 0.999-quantile process the same day. The black lines represent the speeds of valve closings.

Figure 6 shows for the 18th of April 2011 the plot of probability  $P(X_t > 0.3)$  and 0.999-quantile estimators of the valve closing speed for one oyster in the Locmariaquer site. For an easier data visualization of the extreme quantiles of the closing velocities of the 16 oysters through the period starting from 4th of March to 21th of August 2011, we use in Figure 7 a customized color table (gray color associated to the smallest quantiles class, yellow to the intermediate quantiles class and red to the largest quantiles class) to match computed extreme quantile values.

The 4th of March and the 21th of August 2011 correspond respectively to the 63th days and to the 233th days of the 2011 year. For each day, there are 16 lines of colored points representing the extreme 0.999-quantiles values at each  $t \in [0, 24]$  hour of each oyster's velocity. The advantage of this representation is to give the extreme quantiles values for

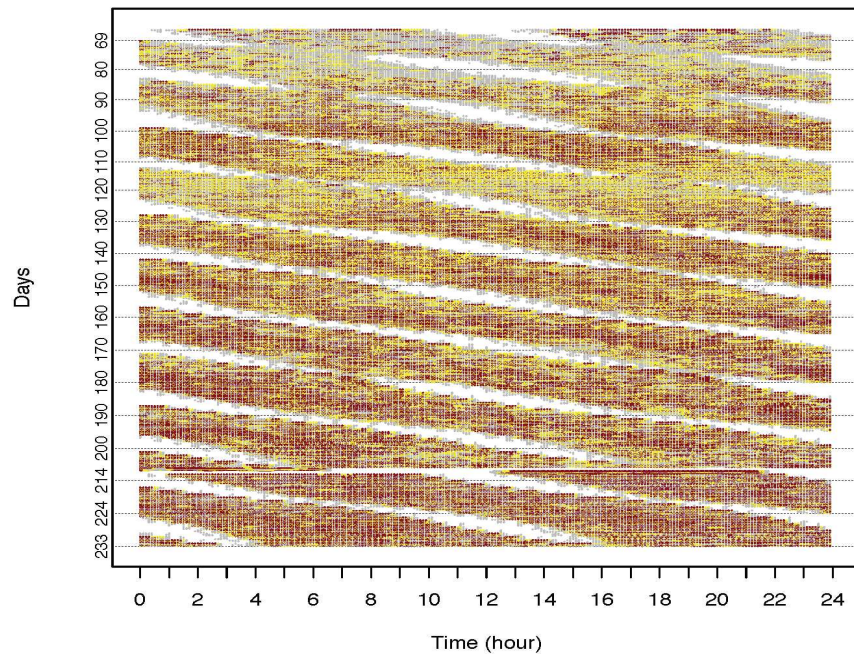


FIGURE 7. Representation of the extreme 0.999-quantile estimator of the closing velocity between the 4th of March and the 21th of August 2011 considering the 16 oysters in Locmariaquer. The x-axis represents the time in a 24 hour time period and the y-axis represents the number of days since the 1st of January 2011

each of the 16 oysters for a given time period in one single graph. Figure 7 shows that the closing activity is highly correlated with the tidal amplitude and that the closing state is synchronized with the low tide period. This is confirmed by [Sow et al. \(2011\)](#); [Coudret et al. \(2013\)](#) using non parametric methods.

We notice particularly a yellow zone (between the 110th and 125th days) explained by a sudden change in temperature collected by a temperature sensor installed near the oysters (data not shown) and a red colored area showing a more intense activity including spawning activity (days  $\geq 210$ ). These results thus contribute to the development of a tool for monitoring water quality based on the analysis of continuous behavior of bivalves (bio-indicator of pollution in the field). This velocity information provides an important

indication of the change in behavior of oysters such as a spawn or a period of abnormal stress characterized by rapid partial closures, openings).

## 7. CONCLUDING REMARKS

*The model:* this article deals with estimation of the tail probabilities and extreme quantiles in the framework of time series data  $X_{t_1}, \dots, X_{t_n}$ . Our approach is based on adjusting a Pareto correction to the weighted empirical distribution function beyond a given threshold  $\tau$  for observations in the neighborhood of each time  $t$  with given bandwidth  $h$ . The choice of the adjusted Pareto model is justified by the extreme value theory.

*Theoretical results:* we determine rates of convergence of the corresponding estimators of the parameters in the adjusted model when the threshold and the kernel bandwidth are deterministic. These results are then extended to the estimator with adaptive threshold.

*Adaptive estimator:* in applications the threshold usually is not known. We propose a Maximum Likelihood Maximal Propagation (MLMP) selection procedure based on the maximal propagation of a parametric adjustment and a subsequent choice of the threshold using penalized maximum likelihood.

*Model validity:* the construction of the adaptive estimator is based on a testing procedure which can be viewed as a goodness-of-fit test for the parametric-based part of the model. So the question of the model validation for the adjusted tail is answered by the MLMP procedure: at each step of the procedure the adjusted tail is tested and if it is not rejected the sample is enlarged and tested again until the parametric model is rejected. The choice of the "optimal" threshold is made among the already tested models and therefore the adaptive adjusted tail is validated as well. If the test rejects the parametric tail fit from the very beginning, the Pareto tail adjustment is not significant. On the opposite, if all the tests accept the parametric Pareto fit then the underlying distribution is significantly Pareto.



*Simulations:* we study the behavior of the estimators under the null and the alternative hypotheses. Under the null hypothesis, assuming i.i.d. standard Pareto observations, we compute the critical value in the MLMP procedure and we show that it remains stable with respect to the number of observations in the window. This is done for the Gaussian kernel, but the critical values for other kernels can be determined in the same way. We perform numerical simulations with the adaptive estimators of the Pareto tail parameter in order to show that under the alternative hypothesis the mean squared error is small. We analyze also the relative mean squared error of the extreme quantile estimator.

*Application:* we apply the developed procedure in the context of an ecological study. The objective is to determine extreme environmental disturbances through high frequency measurements of oysters activity considered as a bioindicators of pollution.

#### REFERENCES

- W. C. Allee, A. E. Emerson, O. Park, T. Park, and K. P. Schmidt. *Principles of animal ecology*. Springer-Verlag, Philadelphia, 1949. ISBN 978-0721611204.
- S. Altman and R. B. Whitlatch. Effects of small-scale disturbance on invasion success in marine communities. *Journal of Experimental Marine Biology and Ecology*, 342:15–29, 2007.
- R. Azaïs, G. Coudret, and G. Durrieu. A hidden renewal model for monitoring aquatic systems biosensors. submitted.
- J. Borchering and M. Volpers. The dreissena monitor - first results on the application of this biological early warning system in the continuous monitoring of water quality. *Water Science Technology*, 29:199–201, 1994.
- J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. Toward a metabolic theory of ecology. *Ecology*, 85:1171–1789, 2004.
- C.A. Burge, L.R. Judah, L.L. Conquest, F.J. Griffin, et al. Summer seed mortality of the pacific oyster, *crassostrea gigas thunberg* grown in tomales bay, california, usa:

- the influence of oyster stock, planting time, pathogens, and environmental stressors. *Journal of Shellfish Research*, 26(1):163–172, 2007.
- C. Chambon, A. Legeay, G. Durrieu, P. Gonzalez, P. Ciret, and J.-C. Massabuau. Influence of the parasite worm *Polydora* sp. on the behaviour of the oyster *Crassostrea gigas*: a study of the respiratory impact and associated oxidative stress. *Marine Biology*, 152(2):329–338, 2007.
- J. H. Connell. Diversity in tropical rain forests and coral reefs. *Science*, 199:1302–1310, 1978.
- R. Coudret, G. Durrieu, and J. Saracco. Comparison of kernel density estimators with assumption on number of modes. *Communication in Statistics - Simulation and Computation*, In press, 2013.
- P. K. Dayton. Competition, disturbance, and community organization: the provision and subsequent utilization of space in a rocky intertidal community. *Ecological Monographs*, 41:351–389, 1971.
- M.W. Denny, L.J.H. Hunt, L.P. Miller, and C.D.G. Harley. On the prediction of extreme ecological events. *Ecological Monographs*, 79(3):397–421, 2009.
- F. G. Doherty, D. S. Cherry, and J. Cairns. Valve closure responses of the asiatic clam *C. fluminea* exposed to cadmium and zinc. *Hydrobiologia*, 153:159–167, 1987.
- G. Durrieu and L. Briollais. Sequential determination of sample size for robust linear regression: application to microarray experimental designs. *Journal of the American Statistical Association*, 104(486):650–660, 2009.
- C. Folke, S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling. Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology and Systematics*, 35:557–581, 2004.
- S. D. Gaines and M.W. Denny. The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. *Ecology*, 74:1677–1692, 1993.
- L. Gardes and S. Girard. A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99:2368–2388, 2008.

- I. Grama and V. Spokoiny. Statistics of extremes by oracle estimation. *Ann. Statist.*, 36(4):1619–1648, 2008.
- I. Grama, J.M. Tricot, and J.F. Petiot. Estimation of the survival probabilities by adjusting a cox model to the tail. *C.R. Acad. Sci. Paris, Ser. I.*, 349(13-14):807–811, 2011.
- I. Grama, J.M. Tricot, and J.F. Petiot. Estimation of the extreme survival probabilities from censored data. *Buletinul Academiei De Stiinte a Republicii Moldova. Matematica*, 2013a.
- I. Grama, J.M. Tricot, and J.F. Petiot. Long term survival probabilities and Kaplan-Meier estimator. *Proceedings of the Joint Statistical Meeting, Montreal, Canada, 2013*, 2013b.
- K. L. Gross, G. G. Mittelbach, and H. L. Reynolds. Grassland invasibility and diversity: responses to nutrients, seed input. *Ecology*, 86:476–486, 2005.
- V. P. Gutschick and H. BassiriRad. Extreme events as shaping physiology, ecology, and evolution of plants: toward a unified definition and evaluation of their consequences. *New Phytologist*, 160(1):21–42, 2003.
- P. Hall. On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(1):37–42, 1982.
- P. Hall and A. H. Welsh. Best attainable rates of convergence for estimates of parameters of regular variation. *Annals of Statistics*, 12(3):1079–1084, 1984.
- H.A. Jenner, F. Noppert, and T. Sikking. A new system for the detection of valve movement rene of bivalves. *Kema Scientific and Technical Reports*, 7(2):91–98, 1989.
- L. J. Jou and C. M. Liao. A dynamic artificial clam (*corbicula fluminea*) allows parcimony on-line measurement of waterborne metals. *Environmental Pollution*, 144:172–183, 2006.
- R.W. Katz, G.S. Brush, and M.B. Parlange. Statistics of extremes: modeling ecological disturbances. *Ecology*, 86(5):1124–1134, 2005.

- K.J.M. Kramer, H.A. Jenner, and D. De Zwart. The valve movement response of mussels: a tool in biological monitoring. *Hydrobiologia*, 188/189:433–443, 1989.
- K. Nagai, T. Honjo, J. Go, H. Yamashita, and S.J. Oh. Detecting the shellfish killer *heterocapsa circularisquama* (dinophyceae) by measuring the bivalve valve activity with hall element sensor. *Aquaculture*, 255:395–401, 2006.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- A. Robson, R. Wilson, and C. Garcia de Leaniz. Mussels flexing their muscles: a new method for quantifying bivalve behavior. *Marine Biology*, 151:1195–1204, 2007.
- J.F. Samain and H. McCombie. *Summer mortality of the Pacific oyster *Crassostrea gigas*: the Morest Project*. Quae, Versailles, 2008. ISBN 978-2-7592-0054-2.
- F. G. Schmitt, M. De Rosa, G. Durrieu, M. Sow, P. Ciret, D. Tran, and J. C. Massabuau. Statistical analysis of bivalve high frequency microclosing behavior: scaling properties and shot noise modeling. *International Journal of Bifurcation and Chaos*, 21(12):3565–3576, 2011.
- W. Sloff, D. De Zwart, and J.M. Marquenie. Detection limits of a biological monitoring system for chemical water pollution based on mussel activity. *Environmental Contamination and Toxicology*, 30(4):400–405, 1983.
- W. P. Sousa. Disturbance in marine intertidal boulder fields: the non-equilibrium maintenance of species diversity. *Ecology*, 60:1225–1239, 1979.
- M. Sow, G. Durrieu, and L. Briollais. Water quality assessment by means of HFNI valvometry and high-frequency data modeling. *Environmental Monitoring and Assessment*, 182(1-4):155–170, 2011.
- V. Spokoiny. Multiscale local change point detection with applications to value-at-risk. *Ann. Statist.*, 37(3):1405–1436, 2009.
- C.J. Stone. Optimal global rates of convergence for nonparametric regression. *Annals of Statistics*, 10(4):1040–1053, 1982.

- D. Tran, P. Ciret, A. Ciutat, G Durrieu, and J.-C. Massabuau. Estimation of potential and limits of bivalve closure response to detect contaminants: application to cadmium. *Environmental Toxicology and Chemistry*, 22(4):914–920, 2003.
- D. Tran, E. Fournier, G. Durrieu, and J.-C. Massabuau. Copper detection in the Asiatic clam *Corbicula fluminea*: Optimum valve closure response. *Aquatic Toxicology*, 66: 333–343, 2004.
- D. Tran, E. Fournier, G. Durrieu, and J.-C. Massabuau. Inorganic mercury detection by valve closure response in the freshwater clam *corbicula fluminea*: integration of time and water metal concentration changes. *Environmental Toxicology and Chemistry*, 26: 1545–1551, 2007.
- D. Tran, H. Haberkorn, P. Soudant, P. Ciret, and J.-C. Massabuau. Behavioral responses of *Crassostrea gigas* exposed to the harmful algae *Alexandrium minutum*. *Aquaculture*, 298:338–345, 2010.
- D. Tran, A. Nadau, G. Durrieu, P. Ciret, J.-P. Parisot, and J.-C. Massabuau. Field chronobiology of a molluscan bivalve: How the moon and the sun cycles interact to drive oyster activity rhythms. *Chronobiology International*, 28(4):307–317, 2011.

## 8. SUPPLEMENTARY MATERIALS

8.1. **Auxiliary results.** We consider a semiparametric Pareto model with two change points  $v$  and  $\tau$  defined by

$$F_{t,\theta,s,\mu,v,\theta',\tau}(x) = \begin{cases} F_t(x) & \text{if } x \in [x_0, s], \\ 1 - (1 - F_t(s))(1 - G_{s,\theta}(x)) & \text{if } x \in (s, v], \\ 1 - (1 - F_t(s))(1 - G_{s,\theta}(v))(1 - G_{v,\mu}(x)) & \text{if } x \in (v, \tau], \\ 1 - (1 - F_t(s))(1 - G_{s,\theta}(v))(1 - G_{v,\mu}(\tau))(1 - G_{\tau,\theta'}(x)) & \text{if } x > \tau, \end{cases}$$

where  $\theta', \theta, \mu > 0$ ,  $\tau \geq v \geq s \geq x_0$ . Let

$$\mathcal{L}_{t,h}(F_{t,v,\theta}) = \mathcal{L}_{t,h}(v, \theta) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t,v,\theta}}{dx}(X_{t_i}),$$

$$\mathcal{L}_{t,h}(F_{t,\mu,v,\theta,\tau}) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t,\mu,v,\theta,\tau}}{dx}(X_{t_i}),$$

$$Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) = \mathcal{L}_{t,h}(F_{t,v,\theta'}) - \mathcal{L}_{t,h}(F_{t,v,\theta}) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t,v,\theta'}}{dF_{t,v,\theta}}(X_{t_i}),$$

and

$$Z_{t,h}(F_{t,\mu,v,\theta,\tau}, F_{t,v,\theta}) = \mathcal{L}_{t,h}(F_{t,\mu,v,\theta,\tau}) - \mathcal{L}_{t,h}(F_{t,v,\theta}) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t,\mu,v,\theta,\tau}}{dF_{t,v,\theta}}(X_{t_i}).$$

The following proposition gives the exponential bounds for the quasi-log-likelihood ratios  $Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta})$  and  $Z_{t,h}(F_{t,\mu,v,\theta,\tau}, F_{t,v,\theta})$ . We introduce a measure of discrepancy between the family of distributions  $(F_u)_{u \in [0,T]}$  and the adjusted models  $(F_{u,s,\theta})_{u \in [0,T]}$  at time  $t$  by

$$d_{t,h,s,\theta} = \sum_{i=1}^n W_{t,h}(t_i) \chi^2(F_{t_i}, F_{t_i,s,\theta}).$$

**Proposition 8.1.** *For any  $y > 0$ ,  $\tau \geq v \geq s \geq x_0$ , and any  $\mu, \theta, \theta' > 0$ , we have*

$$(8.1) \quad \mathbb{P}(Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) > y) \leq \exp\left(-\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2}\right),$$

$$(8.2) \quad \mathbb{P}(Z_{t,h}(F_{t,\mu,v,\theta,\tau}, F_{t,v,\theta}) > y) \leq \exp\left(-\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2}\right).$$

*Proof.* We first prove (8.1). Let  $v \geq s \geq x_0$ . Since  $\frac{dF_{t,v,\theta'}}{dF_{t,v,\theta}} = \frac{dF_{t_i,\theta,s,\theta',v}}{dF_{t_i,s,\theta}}$  for  $i = 1, \dots, n$ , by the definition of  $Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta})$ , we have

$$Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t_i,\theta,s,\theta',v}}{dF_{t_i,s,\theta}}(X_{t_i}).$$

Denote for brevity  $H_{t_i} = F_{t_i,\theta,s,\theta',v}$ . Applying exponential Chebyshev's inequality, we have

$$(8.3) \quad \mathbb{P}(Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) > y) \leq \exp\left(-\frac{y}{2}\right) \mathbb{E}\left(\exp\left(\frac{1}{2}Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta})\right)\right).$$

Since  $0 \leq W_{t,h}(t_i) \leq 1$  for all  $i = 1, \dots, n$ , we deduce by Hölder's inequality

$$\begin{aligned} & \log \mathbb{E}\left(\exp\left(\frac{1}{2}Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta})\right)\right) \\ &= \sum_{i=1}^n \log \mathbb{E}\left(\exp\left(\frac{1}{2}1_{\{X_{t_i}>s\}} \log \frac{dH_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right)^{W_{t,h}(t_i)} \\ &\leq \sum_{i=1}^n W_{t,h}(t_i) \log \mathbb{E}\left(\exp\left(\frac{1}{2}1_{\{X_{t_i}>\tau\}} \log \frac{dH_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right) \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E}\left(\exp\left(\frac{1}{2}1_{\{X_{t_i}>s\}} \log \frac{dH_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right) \\ &= \mathbb{E}\left(\exp\left(\frac{1}{2}1_{\{X_{t_i}>s\}} \log \frac{dH_{t_i}}{dF_{t_i}}(X_{t_i})\right) \exp\left(\frac{1}{2}1_{\{X_{t_i}>s\}} \log \frac{dF_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right) \\ &\leq \sqrt{\mathbb{E}\left(\exp\left(1_{\{X_{t_i}>s\}} \log \frac{dH_{t_i}}{dF_{t_i}}(X_{t_i})\right)\right)} \sqrt{\mathbb{E}\left(\exp\left(1_{\{X_{t_i}>s\}} \log \frac{dF_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right)}. \end{aligned}$$

Using the fact that, for  $i = 1, \dots, n$

$$\mathbb{E}\left(\exp\left(1_{\{X_{t_i}>s\}} \log \frac{dH_{t_i}}{dF_{t_i}}(X_{t_i})\right)\right) = 1$$

and

$$\mathbb{E}\left(\exp\left(1_{\{X_{t_i}>s\}} \log \frac{dF_{t_i}}{dF_{t_i,s,\theta}}(X_{t_i})\right)\right) = 1 + \chi^2(F_{t_i}, F_{t_i,s,\theta}),$$

we obtain

$$\begin{aligned}
\log \left( \mathbb{E} \left( \exp \left( \frac{1}{2} Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) \right) \right) \right) &\leq \frac{1}{2} \sum_{i=1}^n W_{t,h}(t_i) \log (1 + \chi^2(F_{t_i}, F_{t_i,s,\theta})) \\
(8.4) \qquad \qquad \qquad &\leq \frac{1}{2} \sum_{i=1}^n W_{t,h}(t_i) \chi^2(F_{t_i}, F_{t_i,s,\theta}) = \frac{d_{t,h,s,\theta}}{2}.
\end{aligned}$$

Combining (8.3) and (8.4), we deduce that

$$(8.5) \qquad \mathbb{P}(Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) > y) \leq \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right),$$

which prove that (8.1) is satisfied.

Since, for  $i = 1, \dots, n$

$$\frac{dF_{t,\mu,v,\theta,\tau}}{dF_{t,v,\theta}}(X_{t_i}) = \frac{dF_{t_i,\theta,s,\mu,v,\theta,\tau}}{dF_{t_i,s,\theta}}(X_{t_i}),$$

we have

$$Z_{t,h}(F_{t,\mu,v,\theta,\tau}, F_{t,v,\theta}) = \sum_{i=1}^n W_{t,h}(t_i) \log \frac{dF_{t_i,\theta,s,\mu,v,\theta,\tau}}{dF_{t_i,s,\theta}}(X_{t_i}).$$

Now (8.2) is proved in the same way as (8.1).  $\square$

Next, we give an exponential bound for the maximum quasi-log-likelihood ratio which permits to obtain a rate of convergence of nonadaptive estimator  $\hat{\theta}_{t,h,\tau_n}$ .

**Proposition 8.2.** *For any  $y > 0$ ,  $\tau \geq v \geq s \geq x_0$ , and  $\theta > 0$ , we have*

$$\begin{aligned}
\mathbb{P} \left( [\hat{n}_{t,h,v}] \mathcal{K}(\hat{\theta}_{t,h,v}, \theta) > y \right) &\leq 2n \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right), \\
\mathbb{P} \left( [\hat{n}_{t,h,v,\tau}] \mathcal{K}(\hat{\mu}_{t,h,v,\tau}, \theta) > y \right) &\leq 2n \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right),
\end{aligned}$$

where  $[u]$  is the integer part of  $u$ .

*Proof.* We shall prove only the first inequality, the second one being proved in the same way. We start with the obvious relation  $Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) = \hat{n}_{t,h,v} \Lambda(\theta')$ , where  $\Lambda(u) = \log \frac{\theta}{u} - (\frac{1}{u} - \frac{1}{\theta}) \hat{\theta}_{t,h,v}$ . Denote for brevity  $g(u, k) = (\log \frac{\theta}{u} - \frac{y}{k}) / (\frac{1}{u} - \frac{1}{\theta})$ . Note that for  $k > 0$  and  $0 < u < \theta$ , the inequality  $k\Lambda(u) > y$  is equivalent to  $g(u, k) > \hat{\theta}_{t,h,v}$  for  $k > 0$ . Similarly, for  $u > \theta$  the inequality  $k\Lambda(u) > y$  is equivalent to  $g(u, k) < \hat{\theta}_{t,h,v}$ . Moreover with  $k > 0$  fixed, the function  $g(u, k)$  has a maximum for  $0 < u < \theta$  and a minimum for



$u > \theta$ . Let  $\theta^+(k) = \arg \max_{0 < u < \theta} g(u, k)$  and  $\theta^-(k) = \arg \min_{u > \theta} g(u, k)$ . One can see that

$$\begin{aligned} \left\{ [\widehat{n}_{t,h,v}] \Lambda(\widehat{\theta}_{t,h,v}) > y, \widehat{\theta}_{t,h,v} < \theta \right\} &= \left\{ g(\widehat{\theta}_{t,h,v}, [\widehat{n}_{t,h,v}]) > \widehat{\theta}_{t,h,v}, \widehat{\theta}_{t,h,v} < \theta \right\} \\ &\subseteq \left\{ g(\theta^+([\widehat{n}_{t,h,v}]), [\widehat{n}_{t,h,v}]) > \widehat{\theta}_{t,h,v}, \widehat{\theta}_{t,h,v} < \theta \right\} \\ &= \left\{ [\widehat{n}_{t,h,v}] \Lambda(\theta^+([\widehat{n}_{t,h,v}])) > y, \widehat{\theta}_{t,h,v} < \theta \right\} \\ &\subseteq \left\{ [\widehat{n}_{t,h,v}] \Lambda(\theta^+([\widehat{n}_{t,h,v}])) > y \right\}. \end{aligned}$$

In the same way, we have

$$\left\{ [\widehat{n}_{t,h,v}] \Lambda(\widehat{\theta}_{t,h,v}) > y, \widehat{\theta}_{t,h,v} > \theta \right\} \subseteq \left\{ [\widehat{n}_{t,h,v}] \Lambda(\theta^-([\widehat{n}_{t,h,v}])) > y \right\}.$$

Since  $\Lambda(\widehat{\theta}_{t,h,v}) = \mathcal{K}(\widehat{\theta}_{t,h,v}, \theta)$  for any  $\theta > 0$  and  $\mathcal{K}(\widehat{\theta}_{t,h,v}, \theta) = 0$  for  $\theta = \widehat{\theta}_{t,h,v}$ , these inclusions imply

$$\left\{ [\widehat{n}_{t,h,v}] \mathcal{K}(\widehat{\theta}_{t,h,v}, \theta) > y \right\} \subseteq \left\{ [\widehat{n}_{t,h,v}] \Lambda(\theta^+([\widehat{n}_{t,h,v}])) > y \right\} \cup \left\{ [\widehat{n}_{t,h,v}] \Lambda(\theta^-([\widehat{n}_{t,h,v}])) > y \right\}.$$

Hence,

$$\begin{aligned} &\mathbb{P} \left( [\widehat{n}_{t,h,v}] \mathcal{K}(\widehat{\theta}_{t,h,v}, \theta) > y \right) \\ &\leq \mathbb{P} \left( [\widehat{n}_{t,h,v}] \Lambda(\theta^+([\widehat{n}_{t,h,v}])) > y \right) + \mathbb{P} \left( [\widehat{n}_{t,h,v}] \Lambda(\theta^-([\widehat{n}_{t,h,v}])) > y \right) \\ &\leq \sum_{k=1}^{[n_{t,h}]} \mathbb{P} \left( [\widehat{n}_{t,h,v}] \Lambda(\theta^+(k)) > y \right) + \sum_{k=1}^{[n_{t,h}]} \mathbb{P} \left( [\widehat{n}_{t,h,v}] \Lambda(\theta^-(k)) > y \right), \end{aligned}$$

where  $n_{t,h} = \sum_{i=1}^n W_{t,h}(t_i)$ . From Proposition 8.1, and the inclusion  $\{[\widehat{n}_{t,h,v}] \Lambda(\theta') > y\} \subseteq \{\widehat{n}_{t,h,v} \Lambda(\theta') > y\}$ , we have, for any  $\theta' > 0$

$$\begin{aligned} \mathbb{P}([\widehat{n}_{t,h,v}] \Lambda(\theta') > y) &\leq \mathbb{P}(\widehat{n}_{t,h,v} \Lambda(\theta') > y) \\ &= \mathbb{P}(Z_{t,h}(F_{t,v,\theta'}, F_{t,v,\theta}) > y), \\ &\leq \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right). \end{aligned}$$

We deduce

$$\mathbb{P} \left( [\widehat{n}_{t,h,v}] \mathcal{K}(\widehat{\theta}_{t,h,v}, \theta) > y \right) \leq 2[n_{t,h}] \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right).$$

Since  $[n_{t,h}] \leq n$ , we have

$$\mathbb{P} \left( [\widehat{n}_{t,h,v}] \mathcal{K}(\widehat{\theta}_{t,h,v}, \theta) > y \right) \leq 2n \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right).$$

□

**Proposition 8.3.** *For any  $y > 0$ ,  $s \geq x_0$  and  $\theta > 0$ , we have*

$$\mathbb{P} \left( \sup_{s \leq v} [\widehat{n}_{t,h,v}] \mathcal{K}(\widehat{\mu}_{t,h,v}, \theta) > y \right) \leq 2n^4 \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right) + \frac{1}{n}$$

and

$$\mathbb{P} \left( \sup_{s \leq v \leq \tau} [\widehat{n}_{t,h,v,\tau}] \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) > y \right) \leq n^7 \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right) + \frac{1}{n}.$$

*Proof.* The proof of the proposition is similar that of Proposition 7.4 in [Grama and Spokoiny \(2008\)](#). □

The following Proposition gives an exponential bound for the statistic  $LR_{t,h}(v, \tau)$  (see [\(4.3\)](#) ).

**Proposition 8.4.** *For any  $y > 0$ ,  $s \geq x_0$  and  $\theta > 0$ , we have*

$$\mathbb{P} \left( \sup_{s \leq v \leq \tau} LR_{t,h}(v, \tau) > 4y \right) \leq 2n^7 \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right) + \frac{2}{n}.$$

*Proof.* From the fact that

$$LR_{t,h}(v, \tau) = \max_{\mu, \theta' > 0} \mathcal{L}_{t,h}(F_{t,\mu,v,\theta',\tau}) - \max_{\theta > 0} \mathcal{L}_{t,h}(F_{t,v,\theta})$$

and

$$\max_{\theta > 0} \mathcal{L}_{t,h}(F_{t,v,\theta}) \geq \mathcal{L}_{t,h}(F_{t,v,\theta}),$$

it follows that

$$LR_{t,h}(v, \tau) \leq \max_{\mu, \theta' > 0} \mathcal{L}_{t,h}(F_{t,\mu,v,\theta',\tau}) - \mathcal{L}_{t,h}(F_{t,v,\theta}).$$

Proceeding as in the proof of [\(4.3\)](#), we see that

$$\begin{aligned} \max_{\mu, \theta' > 0} \mathcal{L}_{t,h}(F_{t,\mu,v,\theta',\tau}) - \mathcal{L}_{t,h}(F_{t,v,\theta}) &= \widehat{n}_{t,h,v,\tau} \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) + \widehat{n}_{t,h,\tau} \mathcal{K}(\widehat{\theta}_{t,h,\tau}, \theta) \\ &\leq 2[\widehat{n}_{t,h,v,\tau}] \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) + 2[\widehat{n}_{t,h,\tau}] \mathcal{K}(\widehat{\theta}_{t,h,\tau}, \theta). \end{aligned}$$

We deduce

$$LR_{t,h}(v, \tau) \leq 2[\widehat{n}_{t,h,v,\tau}] \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) + 2[\widehat{n}_{t,h,\tau}] \mathcal{K}(\widehat{\theta}_{t,h,\tau}, \theta)$$

and

$$\begin{aligned} \left\{ \sup_{s \leq v \leq \tau} LR_{t,h}(v, \tau) > 4y \right\} &\subseteq \left\{ \sup_{s \leq v \leq \tau} [\widehat{n}_{t,h,v,\tau}] \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) > y \right\} \\ &\cup \left\{ \sup_{s \leq \tau} [\widehat{n}_{t,h,\tau}] \mathcal{K}(\widehat{\theta}_{t,h,\tau}, \theta) > y \right\}. \end{aligned}$$

From Proposition 8.3 and the previous inclusion, we obtain

$$\begin{aligned} &\mathbb{P} \left( \sup_{s \leq v \leq \tau} LR_{t,h}(v, \tau) > 4y \right) \\ &\leq \mathbb{P} \left( \sup_{s \leq v \leq \tau} [\widehat{n}_{t,h,v,\tau}] \mathcal{K}(\widehat{\mu}_{t,h,v,\tau}, \theta) > y \right) + \mathbb{P} \left( \sup_{s \leq \tau} [\widehat{n}_{t,h,\tau}] \mathcal{K}(\widehat{\theta}_{t,h,\tau}, \theta) > y \right) \\ &\leq 2n^7 \exp \left( -\frac{y}{2} + \frac{d_{t,h,s,\theta}}{2} \right) + \frac{2}{n}. \end{aligned}$$

□

Denote

$$\bar{n}_{t,h_n,\tau_n} = \sum_{i=1}^n W_{t,h_n}(t_i)(1 - F_{t_i}(\tau_n)).$$

**Lemma 8.5.** *If the sequence  $(\tau_n)$  is such that  $\bar{n}_{t,h_n,\tau_n} \rightarrow \infty$  as  $n \rightarrow \infty$ , then  $\widehat{n}_{t,h_n,\tau_n} \stackrel{\mathbb{P}}{\asymp} \bar{n}_{t,h_n,\tau_n}$  as  $n \rightarrow \infty$ . Moreover,  $[\widehat{n}_{t,h_n,\tau_n}] \stackrel{\mathbb{P}}{\asymp} \bar{n}_{t,h_n,\tau_n}$  as  $n \rightarrow \infty$ .*

*Proof.* By Chebyshev's exponential inequality, for any  $u > 0$  and  $\epsilon \in (0, 1)$ ,

$$(8.6) \quad \mathbb{P} \left( \frac{\widehat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} < 1 - \epsilon \right) \leq \exp \left( u(1 - \epsilon) \bar{n}_{t,h_n,\tau_n} + \log \mathbb{E} \left( e^{-u \widehat{n}_{t,h_n,\tau_n}} \right) \right).$$

Applying Hölder's inequality, we have

$$\begin{aligned} \log \mathbb{E} e^{-u \widehat{n}_{t,h_n,\tau_n}} &= \sum_{i=1}^n \log \mathbb{E} \left( \left( e^{-u 1_{\{X_{t_i} > \tau_n\}}} \right)^{W_{t,h_n}(t_i)} \right) \\ &\leq \sum_{i=1}^n W_{t,h_n}(t_i) \log \mathbb{E} \left( e^{-u 1_{\{X_{t_i} > \tau_n\}}} \right). \end{aligned}$$

Using the fact that

$$\log \mathbb{E} \left( e^{-u 1_{\{X_{t_i} > \tau_n\}}} \right) = \log \left( 1 - (1 - F_{X_{t_i}}(\tau_n))(1 - e^{-u}) \right) \leq - (1 - F_{X_{t_i}}(\tau_n)) (1 - e^{-u}),$$

we have

$$(8.7) \quad u\bar{n}_{t,h_n,\tau_n} + \log \mathbb{E} \left( e^{-u\hat{n}_{t,h,\tau_n}} \right) \leq \bar{n}_{t,h_n,\tau_n} (e^{-u} + u - 1) \leq n_{t,h_n,\tau_n} \frac{u^2}{2}.$$

From (8.6) and (8.7) it follows that

$$(8.8) \quad \mathbb{P} \left( \frac{\hat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} < 1 - \epsilon \right) \leq \exp \left( \bar{n}_{t,h_n,\tau_n} \left( -u\epsilon + \frac{u^2}{2} \right) \right).$$

In the same way, we have

$$(8.9) \quad \mathbb{P} \left( \frac{\hat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} > 1 + \epsilon \right) \leq \exp \left( \bar{n}_{t,h_n,\tau_n} (e^u - (1 + \epsilon)u - 1) \right).$$

Note that, there exist  $\delta > 0$  such that, for all  $0 < u < \delta$  it holds  $\frac{e^u - u - 1}{u} < \epsilon$ . Taking  $u = \min \left\{ \epsilon, \frac{\delta}{2} \right\}$ , we obtain

$$(8.10) \quad (e^u - (1 + \epsilon)u - 1) < 0, \text{ and } -u\epsilon + \frac{u^2}{2} < 0.$$

Then from (8.8), (8.9) and (8.10), we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{\hat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} < 1 - \epsilon \right) = \lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{\hat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} > 1 + \epsilon \right) = 0,$$

which is equivalent to

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \left| \frac{\hat{n}_{t,h_n,\tau_n}}{\bar{n}_{t,h_n,\tau_n}} - 1 \right| > \epsilon \right) = 0,$$

or  $\hat{n}_{t,h_n,\tau_n} \stackrel{\mathbb{P}}{\asymp} \bar{n}_{t,h_n,\tau_n}$  as  $n \rightarrow \infty$ . On the other hand, since  $\bar{n}_{t,h_n,\tau_n} \rightarrow \infty$  as  $n \rightarrow \infty$ , we have

$$[\hat{n}_{t,h_n,\tau_n}] \leq \hat{n}_{t,h,\tau_n} \leq 2[\hat{n}_{t,h_n,\tau_n}],$$

which implies that  $[\hat{n}_{t,h_n,\tau_n}] \stackrel{\mathbb{P}}{\asymp} \bar{n}_{t,h_n,\tau_n}$ , as  $n \rightarrow \infty$ . This completes the proof.  $\square$

**Lemma 8.6.** *For any given integer positive  $k_0$  and any sequence  $\tau_n, n = 1, 2, \dots$ , satisfying  $\bar{n}_{t,h_n,\tau_n} \rightarrow \infty$  as  $n \rightarrow \infty$ , it holds  $\lim_{n \rightarrow \infty} \mathbb{P}(Y_{k_0} > \tau_n) = 1$ .*

*Proof.* Since  $\bar{n}_{t,h_n,\tau_n} \rightarrow \infty$  as  $n \rightarrow \infty$ , by Lemma 8.5, there exist constants  $C_1, C_2 > 0$  such that, as  $n \rightarrow \infty$ ,

$$\mathbb{P}(C_1 \bar{n}_{t,h_n,\tau_n} \leq \hat{n}_{t,h_n,\tau_n} \leq C_2 \bar{n}_{t,h_n,\tau_n}) \rightarrow 1,$$

and, for  $M = k_0/C_1$ , there exist  $n_0 > 0$  such that, for any  $n > n_0$ ,

$$\bar{n}_{t,h_n,\tau_n} > M.$$

So that, for any  $n > n_0$ ,

$$\{C_1 \bar{n}_{t, h_n, \tau_n} \leq \hat{n}_{t, h_n, \tau_n}\} \subseteq \{\hat{n}_{t, h_n, \tau_n} > k_0\}.$$

Since  $0 \leq W_{t, h_n}(t_i) \leq 1, i = 1, \dots, n$ , we have

$$\{\hat{n}_{t, h_n, \tau_n} > k_0\} \subseteq \{Y_{k_0} > \tau_n\}.$$

It follows that

$$\mathbb{P}(Y_{k_0} > \tau_n) \rightarrow 1,$$

which ends the proof.  $\square$

**8.2. Proof of Theorem 3.1.** Letting  $s = v = \tau_n, h = h_n, \theta = \theta_n$  and  $y = 4 \log n + d_{\tau_n}$ .

Using the first inequality of Proposition 8.2, we have, as  $n \rightarrow \infty$ ,

$$\mathcal{K}(\hat{\theta}_{t, h_n, \tau_n}, \theta_n) = O_{\mathbb{P}}\left(\frac{1}{[\hat{n}_{t, h_n, \tau_n}]}(4 \log n + d_{\tau_n})\right).$$

By lemma 8.5,  $[\hat{n}_{t, h_n, \tau_n}] \stackrel{\mathbb{P}}{\asymp} \bar{n}_{t, h_n, \tau_n}$ . Therefore,

$$\mathcal{K}(\hat{\theta}_{t, h_n, \tau_n}, \theta_n) = O_{\mathbb{P}}\left(\frac{1}{\bar{n}_{t, h_n, \tau_n}}(4 \log n + d_{\tau_n})\right),$$

as  $n \rightarrow \infty$ .

**8.3. Proof of Proposition 3.3.** We fix some notations. For any distribution function  $F$  supported on the interval  $[x_0, \infty)$ ,  $x_0 \geq 0$  and having a strictly positive density  $f_F$  define

$$\alpha_F(x) = \frac{1}{x \lambda_F(x)}, \quad x \geq x_0,$$

where  $\lambda_F(x) = \frac{f_F(x)}{1-F(x)}$  is the hazard rate function corresponding to  $F$ . Consider the distance

$$\rho_*(x, y) = \max\left\{\left|\log \frac{x}{y}\right|, \left|\frac{1}{x} - \frac{1}{y}\right|\right\}, \quad x, y > 0.$$

Denote  $U_{t, h} = [t - h, t + h]$ .

Without loss of generality, we can assume that  $A_{max} > 1$ . For any  $t \in [0, T]$ , we shall determine two sequences  $\{\tau_n\}$  and  $\{h_n\}$  such that

$$(8.11) \quad \sup_{s \in U_{t, h_n}} \sup_{x \geq \tau_n} \rho_*(\alpha_{F_s, \tau_n}, \theta_{t, \tau_n})^2 = O\left(\frac{n_{t, h_n}}{n_{t, h_n}(1 - F_t(\tau_n))}\right) \rightarrow 0 \text{ as } n \rightarrow \infty$$

and

$$(8.12) \quad \sup_{s \in U_{t, h_n}} \int_{\tau_n}^{\infty} \left(1 + \log \frac{x}{\tau_n}\right)^2 \left(\frac{x}{\tau_n}\right)^{\epsilon_0} \frac{f_s(x) dx}{1 - F_s(\tau_n)} \leq \epsilon_1, \text{ for any } n \geq N,$$

where  $\epsilon_0, \epsilon_1$  and  $N$  are some constants. To do this, let  $0 < \epsilon_0 < \frac{1}{\gamma_{max}}$  and  $t \in [0, T]$ .

Denote

$$I_t(\tau) = \int_{\tau}^{\infty} \left(1 + \log \frac{x}{\tau}\right)^2 \left(\frac{x}{\tau}\right)^{\epsilon_0} \frac{f_t(x) dx}{1 - F_t(\tau)}, \tau \geq x_0.$$

We first prove that

$$(8.13) \quad I_t(\tau) \leq \epsilon_1,$$

for all  $t \in [0, T]$  and  $\tau \geq N$  with some constants  $\epsilon_1 > 0$  and  $N > 0$ . Since  $\rho > 0$ ,  $0 < \gamma_{min} \leq \gamma_t \leq \gamma_{max} < \infty$ ,  $|r_t(x)| \leq A_t x^{-\frac{\rho}{\gamma_t}}$  and  $|R_t(x)| \leq A_t x^{-\frac{\rho}{\gamma_t}}$ ,  $x \geq x_0$ , we have

$$(8.14) \quad A_t x^{-\frac{\rho}{\gamma_t}} \leq A_{max} x^{-\frac{\rho}{\gamma_{max}}} \leq \frac{1}{2}$$

for any  $x \geq \max\left(x_0, (2A_{max})^{\frac{\gamma_{max}}{\rho}}\right) = N$ . Hence,

$$(8.15) \quad \frac{1 + r_t(x)}{1 + R_t(\tau)} \leq 1 + \left| \frac{1 + r_t(x)}{1 + R_t(\tau)} - 1 \right| \leq 1 + 4A\tau^{-\frac{\rho}{\gamma_{max}}} \leq 3,$$

for any  $x \geq \tau \geq N$ . Since  $0 < \epsilon_0 < \frac{1}{\gamma_{max}}$ , there exists an integer  $k > 0$  such that  $\epsilon_0 + \frac{1}{k} < \frac{1}{\gamma_{max}}$ . Using (8.15) and  $\log \frac{x}{\tau} \leq k \left(\frac{x}{\tau}\right)^{1/2k}$ , for all  $x \geq \tau$ , we deduce, for any  $\tau \geq N$ ,

$$\begin{aligned} I_t(\tau) &= \int_{\tau}^{\infty} \left(1 + \log \frac{x}{\tau}\right)^2 \left(\frac{x}{\tau}\right)^{\epsilon_0 - \frac{1}{\gamma_t} - 1} \frac{(1 + r_t(x)) dx}{1 + R_t(\tau)} \frac{1}{\tau} \\ &\leq 3 \int_{\tau}^{\infty} \left(1 + \log \frac{x}{\tau}\right)^2 \left(\frac{x}{\tau}\right)^{\epsilon_0 - \frac{1}{\gamma_{max}} - 1} \frac{dx}{\tau} \\ &\leq 3 \int_{\tau}^{\infty} 4k^2 \left(\frac{x}{\tau}\right)^{\epsilon_0 + \frac{1}{k} - \frac{1}{\gamma_{max}} - 1} \frac{dx}{\tau}. \end{aligned}$$

By taking  $\epsilon_1 = \frac{12k^2}{\frac{1}{\gamma_{max}} - \epsilon_0 - \frac{1}{k}}$ , the inequality (8.13) is satisfied.

Next, we prove that

$$(8.16) \quad \sup_{x \geq \tau} \rho_*(\alpha_{F_{s,\tau}(x)}, \theta_{t,\tau}) \leq C_\alpha \left( 12A_t \gamma_{max} \tau^{-\frac{\rho}{\gamma_t}} + 4A_t \gamma_{max} \tau^{-\frac{\rho}{\gamma_s}} + L_\gamma h^\beta \right).$$

for any  $s \in U_{t,h}$  and  $\tau \geq N$  with some constant  $C_\alpha > 0$ . Indeed, we have

$$\alpha_{F_t}(x) = \left( \frac{1 + R_t(x)}{1 + r_t(x)} \right) \gamma_t.$$

From (8.14), for all  $x \geq N$

$$\left| \frac{1 + R_t(x)}{1 + r_t(x)} - 1 \right| = \frac{|R_t(x) - r_t(x)|}{1 + r_t(x)} \leq \frac{|R_t(x)| + |r_t(x)|}{1 + r_t(x)} \leq 4A_t x^{-\frac{\rho}{\gamma_t}} \leq 2.$$

This implies for any  $x, \tau \geq N$

$$(8.17) \quad \sup_{x \geq \tau} |\alpha_{F_t}(x) - \gamma_t| \leq 4A_t \gamma_t \tau^{-\frac{\rho}{\gamma_t}},$$

and

$$0 < \alpha_{min} \leq \alpha_{F_t}(x) \leq \alpha_{max} < \infty,$$

where  $\alpha_{min} = \gamma_{min}$  and  $\alpha_{max} = 3\gamma_{max}$ . Integrating by parts in (3.6), we obtain

$$(8.18) \quad \begin{aligned} \theta_{t,\tau} &= \log \frac{x}{\tau} \frac{1 - F_t(x)}{1 - F_t(\tau)} \Big|_\tau^\infty + \int_\tau^\infty \frac{1 - F_t(x)}{1 - F_t(\tau)} \frac{dx}{x} \\ &= \int_\tau^\infty \left( \frac{x}{\tau} \right)^{-\frac{1}{\gamma_t} - 1} \frac{1 + R_t(x)}{1 + R_t(\tau)} \frac{dx}{\tau}. \end{aligned}$$

From (8.14), for any  $x \geq N$

$$(8.19) \quad \left| \frac{1 + R_t(x)}{1 + R_t(\tau)} - 1 \right| \leq 4A_t \tau^{-\frac{\rho}{\gamma_t}} \leq 2.$$

Combining (8.18) and (8.19) gives

$$(8.20) \quad |\theta_{t,\tau} - \gamma_t| \leq \int_\tau^\infty \left( \frac{x}{\tau} \right)^{-\frac{1}{\gamma_t} - 1} \left| \frac{1 + R_t(x)}{1 + R_t(\tau)} - 1 \right| \frac{dx}{\tau} \leq 4A_t \gamma_t \tau^{-\frac{\rho}{\gamma_t}} \leq 2\gamma_t.$$

Therefore  $\gamma_{min} \leq \theta_{t,\tau} \leq 3\gamma_{max}$ . Using  $\alpha_{F_t}(x) = \alpha_{F_{t,\tau}}(x)$  for any  $x \geq \tau \geq x_0$ , from (8.17) and (8.20) we have, for any  $x \geq \tau \geq N$ ,

$$\begin{aligned} \sup_{x \geq \tau} \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}) &\leq C_\alpha \sup_{x \geq \tau} |\alpha_{F_{t,\tau}}(x) - \theta_{t,\tau}| \\ &\leq C_\alpha \left( \sup_{x \geq \tau} |\alpha_{F_t}(x) - \gamma_t| + |\theta_{t,\tau} - \gamma_t| \right) \\ &\leq 8C_\alpha A_t \gamma_t \tau^{-\frac{\rho}{\gamma_t}}, \end{aligned}$$

where  $C_\alpha = \max(\alpha_{min}^{-1}, \alpha_{min}^{-2})$ . Taking into account that  $\gamma_t \leq \gamma_{max}$ , we conclude

$$(8.21) \quad \sup_{x \geq \tau} \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}) \leq 8C_\alpha A_t \gamma_{max} \tau^{-\frac{\rho}{\gamma_t}}.$$

Since for any  $t, s \in [0, T]$ ,  $x \geq N$ ,

$$\begin{aligned} |\alpha_{F_t}(x) - \alpha_{F_s}(x)| &= \left| \left( \frac{1 + R_t(x)}{1 + r_t(x)} - 1 \right) \gamma_t + (\gamma_t - \gamma_s) + \left( 1 - \frac{1 + R_s(x)}{1 + r_s(x)} \right) \gamma_s \right| \\ &\leq \gamma_{max} \left| \frac{1 + R_t(x)}{1 + r_t(x)} - 1 \right| + |\gamma_t - \gamma_s| + \gamma_{max} \left| \frac{1 + R_s(x)}{1 + r_s(x)} - 1 \right|, \\ &\leq 4A_t \gamma_{max} \left( x^{-\frac{\rho}{\gamma_t}} + x^{-\frac{\rho}{\gamma_s}} \right) + L_\gamma |t - s|^\beta \end{aligned}$$

and

$$\rho_*(\alpha_{F_t}(x), \alpha_{F_s}(x)) \leq C_\alpha |\alpha_{F_t}(x) - \alpha_{F_s}(x)|,$$

we obtain

$$(8.22) \quad \rho_*(\alpha_{F_t}(x), \alpha_{F_s}(x)) \leq C_\alpha \left( 4A_t \gamma_{max} (x^{-\frac{\rho}{\gamma_t}} + x^{-\frac{\rho}{\gamma_s}}) + L_\gamma |t - s|^\beta \right).$$

From (8.21), (8.22), we deduce, for any  $x \geq \tau \geq N$ ,  $s \in U_{t,h}$ ,

$$\begin{aligned} \rho_*(\alpha_{F_{s,\tau}}(x), \alpha_{F_{t,\tau}}(x)) &= \rho_*(\alpha_{F_t}(x), \alpha_{F_s}(x)) \\ &\leq C_\alpha \left( 4A_t \gamma_{max} (\tau^{-\frac{\rho}{\gamma_t}} + \tau^{-\frac{\rho}{\gamma_s}}) + L_\gamma h^\beta \right). \end{aligned}$$



Therefore,

$$\begin{aligned} \sup_{x \geq \tau} \rho_*(\alpha_{F_{s,\tau}(x)}, \theta_{t,\tau}) &\leq \sup_{x \geq \tau} \rho_*(\alpha_{F_{s,\tau}(x)}, \alpha_{F_{t,\tau}(x)}) + \sup_{x \geq \tau} \rho_*(\alpha_{F_{t,\tau}(x)}, \theta_{t,\tau}) \\ &\leq C_\alpha \left( 12A_t \gamma_{max} \tau^{-\frac{\rho}{\gamma_t}} + 4A_t \gamma_{max} \tau^{-\frac{\rho}{\gamma_s}} + L_\gamma h^\beta \right), \end{aligned}$$

which proves that (8.16) is satisfied.

From inequality (8.16), by taking the sequences  $\{\tau_n\}$ ,  $\{h_n\}$  such that  $\tau_n \rightarrow \infty$  and  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ , we have

$$\lim_{n \rightarrow \infty} \sup_{s \in U_{t,h_n}} \sup_{x \geq \tau_n} \rho_*(\alpha_{F_{s,\tau_n}(x)}, \theta_{t,\tau_n}) = 0.$$

Hence, for any  $0 < \epsilon_0 < \frac{1}{\gamma_{max}}$ , there exists  $n_0 > 0$  such that

$$(8.23) \quad \sup_{s \in U_{t,h_n}} \sup_{x \geq \tau_n} \rho_*(\alpha_{F_{s,\tau_n}(x)}, \theta_{t,\tau_n}) \leq \epsilon_0$$

for any  $n \geq n_0$ . Moreover, from (8.13), the inequality (8.12) is satisfied with  $\tau_n$  and  $h_n$  defined above. Therefore, by Proposition 8.6 in Grama and Spokoiny (2008), it follows that, for any  $n \geq n_0$ ,  $s \in U_{t,h_n}$ ,

$$\chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) \leq C(\epsilon_0, \epsilon_1) \sup_{x \geq \tau_n} \rho_*^2(\alpha_{F_{s,\tau_n}(x)}, \theta_{t,\tau_n}),$$

where  $C(\epsilon_0, \epsilon_1) = \epsilon_1 * \exp(\epsilon_0)$ . This implies

$$\begin{aligned} &\sup_{s \in U_{t,h_n}} \chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) \\ &\leq C(\epsilon_0, \epsilon_1) \max(\alpha_{min}^{-2}, \alpha_{min}^{-4}) \left( 12A_t \gamma_{max} \tau_n^{-\frac{\rho}{\gamma_t}} \left( 1 + \frac{1}{3} \tau_n^{\frac{\rho L_\alpha}{\gamma_{min}^2} h_n^\beta} \right) + L_\gamma h_n^\beta \right)^2. \end{aligned}$$

We now determine the location  $\tau_n$  and bandwidth  $h_n$  satisfying (8.11). The balance conditions are

$$\begin{aligned} h_n^{2\beta} &\asymp \tau_n^{-\frac{2\rho}{\gamma_t}}, \\ \tau_n^{-\frac{2\rho}{\gamma_t}} &\asymp \frac{\log n_{t,h_n}}{n_{t,h_n} \tau_n^{-\frac{1}{\gamma_t}} \left( 1 + A \tau_n^{-\frac{\rho}{\gamma_t}} \right)}. \end{aligned}$$

The optimal choice is given by

$$\tau_n \asymp \left( \frac{\log n_{t,h_n}}{n_{t,h_n}} \right)^{\frac{-\gamma_t}{1+2\rho}}.$$

Taking into account that  $n_{t,h_n} \asymp 2nh_n$ , we obtain as  $n \rightarrow \infty$

$$\begin{aligned} h_n &\asymp \left( \frac{\log n}{n} \right)^{\frac{1}{1+\beta(2+\rho^{-1})}}, \\ \tau_n &\asymp \left( \frac{\log n}{n} \right)^{\frac{-\gamma_t \beta \rho^{-1}}{1+\beta(2+\rho^{-1})}}, \\ \frac{\log n_{t,h_n}}{n_{t,h_n}(1 - F_t(\tau_n))} &\asymp \left( \frac{\log n}{n} \right)^{\frac{2\beta}{1+\beta(2+\rho^{-1})}}, \end{aligned}$$

and

$$\chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) = O\left( \frac{\log n_{t,h_n}}{n_{t,h_n}(1 - F_t(\tau_n))} \right)$$

uniformly in  $s \in U_{t,h_n}$ . This implies, as  $n \rightarrow \infty$ ,

$$\sup_{s \in U_{t,h_n}} \chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) = O\left( \frac{\log n_{t,h_n}}{n_{t,h_n}(1 - F_t(\tau_n))} \right).$$

It is easy to verify that, as  $n \rightarrow \infty$ ,

$$\bar{n}_{t,h_n,\tau_n} = \sum_{t_i \in U_{t,h_n}} W_{t,h_n}(t_i)(1 - F_{t_i}(\tau_n)) \asymp n_{t,h_n}(1 - F_t(\tau_n)),$$

which completes the proof.

**8.4. Proof of Proposition 3.5.** We shall determine sequences  $(\tau_n)$  and  $(h_n)$ , satisfying

$\tau_n \geq x_0$ ,  $\tau_n \rightarrow \infty$ ,  $h_n \rightarrow 0$  and

$$(8.24) \quad \sup_{x \geq \tau_n} \rho_*(\alpha_{F_s, \tau_n}(x), \theta_{t,\tau_n}) = o(1),$$

$$(8.25) \quad \int_{\tau_n}^{\infty} \left( 1 + \log \frac{x}{\tau_n} \right)^2 \left( \frac{x}{\tau_n} \right)^{\epsilon_0(n)} F_{s,\tau_n}(dx) = O(1)$$

as  $n \rightarrow \infty$ , uniformly in  $s \in U_{t,h_n}$ .

We prove first (8.24). Indeed, by straightforward calculations, we have

$$\begin{aligned}\alpha_{F_t}(x) &= \alpha(x, \gamma_t, \delta_t) = \frac{Cx^{-1/\gamma_t} + (1-C)x^{-1/\delta_t}}{\gamma_t^{-1}Cx^{-1/\gamma_t} + \delta_t^{-1}(1-C)x^{-1/\delta_t}} \\ \theta_{t,\tau} &= \theta(\tau, \gamma_t, \delta_t) = \frac{\gamma_t C \tau^{-1/\gamma_t} + \delta_t (1-C) \tau^{-1/\delta_t}}{C \tau^{-1/\gamma_t} + (1-C) \tau^{-1/\delta_t}}.\end{aligned}$$

It is easy to see that there exist a constant  $N > 0$  such that, for any  $x \geq N$ ,

$$\left| \frac{\partial \alpha}{\partial \gamma_t}(x, \gamma_t, \delta_t) \right| \leq C_1(t) \quad \text{and} \quad \left| \frac{\partial \alpha}{\partial \delta_t}(x, \gamma_t, \delta_t) \right| \leq C_1(t),$$

uniformly in  $\gamma_t$  and  $\delta_t$  with some constant  $C_1(t)$  depending on  $t$ . Therefore, by Taylor's expansion,

$$\begin{aligned}\alpha_{F_t}(x) - \alpha_{F_s}(x) &= \frac{\partial \alpha}{\partial \gamma_t}(x, \gamma_s + \varkappa(\gamma_t - \gamma_s), \delta_s + \varkappa(\delta_t - \delta_s))(\gamma_t - \gamma_s) + \\ &+ \frac{\partial \alpha}{\partial \delta_t}(x, \gamma_s + \varkappa(\gamma_t - \gamma_s), \delta_s + \varkappa(\delta_t - \delta_s))(\delta_t - \delta_s),\end{aligned}$$

where  $\varkappa \in (0, 1)$ . So that, for any  $x \geq N$ ,

$$|\alpha_{F_t}(x) - \alpha_{F_s}(x)| \leq L_\alpha |t - s|^\beta,$$

where  $L_\alpha = C_1(t)(L_\gamma + L_\delta)$ . Let  $h > 0$ . Since  $\rho_*(\alpha_{F_t}(x), \alpha_{F_s}(x)) \leq \max(\delta_0^{-1}, \delta_0^{-2})|\alpha_{F_t}(x) - \alpha_{F_s}(x)|$ , we have, for any  $s \in U_{t,h}$ ,  $x \geq N$ ,

$$\rho_*(\alpha_{F_t}(x), \alpha_{F_s}(x)) \leq \max(\delta_0^{-1}, \delta_0^{-2})L_\alpha h^\beta,$$

and, for any  $\tau \geq N$ ,

$$\begin{aligned}\rho_*(\alpha_{F_{s,\tau}}(x), \theta_{t,\tau}) &\leq \rho_*(\alpha_{F_{s,\tau}}(x), \alpha_{F_{t,\tau}}(x)) + \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}) \\ &= \rho_*(\alpha_{F_s}(x), \alpha_{F_t}(x)) + \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}) \\ &\leq \max(\delta_0^{-1}, \delta_0^{-2})L_\alpha h^\beta + \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}).\end{aligned}$$

It follows that, for any  $\tau \geq N$ ,

$$\sup_{x \geq \tau} \rho_*(\alpha_{F_{s,\tau}}(x), \theta_{t,\tau}) \leq \max(\delta_0^{-1}, \delta_0^{-2})L_\alpha h^\beta + \sup_{x \geq \tau} \rho_*(\alpha_{F_{t,\tau}}(x), \theta_{t,\tau}),$$

uniformly in  $s \in U_{t,h}$ . On the other hand, in the same way as in Section 8.3 (cf. bound (8.14)), there exist a constant  $C_2(t) > 0$  such that for any sequence  $(\tau_n)$  satisfying  $\tau_n \geq x_0$ , it holds

$$\sup_{x \geq \tau_n} \rho_*(\alpha_{F_{t,\tau_n}}(x), \theta_{t,\tau_n}) \leq C_2(t) \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}}.$$

By choosing  $\tau_n \rightarrow \infty$  and  $h = h_n$ , where  $(h_n)$  is a sequence satisfying  $h_n \rightarrow 0$ , we obtain

$$(8.26) \quad \sup_{x \geq \tau_n} \rho_*(\alpha_{F_{s,\tau_n}}(x), \theta_{t,\tau_n}) \leq \max(\delta_0^{-1}, \delta_0^{-2}) L_\alpha h_n^\beta + C_2(t) \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}} = \epsilon_0(n).$$

From this, we obtain (8.24).

We now prove (8.25). Denote  $I_s(n) = \int_{\tau_n}^\infty \left(1 + \log \frac{x}{\tau_n}\right)^2 \left(\frac{x}{\tau_n}\right)^{\epsilon_0(n)} F_{s,\tau_n}(dx)$ . As  $\epsilon_0(n) \rightarrow 0$  as  $n \rightarrow \infty$ , there exist  $n_0 > 0$  such that, for any  $n > n_0$ ,  $\epsilon_0(n) < \gamma_{\max}^{-1}$ . By straightforward calculations

$$I_s(n) = \frac{C \gamma_s^{-1}}{C + (1-C) \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}}} g(\gamma_s^{-1} - \epsilon_0(n)) + \frac{(1-C) \delta_s^{-1} \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}}}{C + (1-C) \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}}} g(\delta_s^{-1} - \epsilon_0(n)),$$

where  $g(x) = \frac{1}{x} + \frac{2}{x^2} + \frac{2}{x^3}$ ,  $x \neq 0$  and  $n \geq n_0$ . It is easy to see that, for any  $n \geq n_0$  and  $s \in U_{t,h_n}$ ,

$$I_s(n) \leq 2 \delta_{\min}^{-1} g(\gamma_{\max}^{-1} - \epsilon_0(n)) = \epsilon_1(n).$$

This implies (8.25).

Combining (8.24) and (8.25), from Proposition 8.6 in Grama and Spokoiny (2008), we have

$$(8.27) \quad \chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) \leq \epsilon_1(n) \exp(\epsilon_0(n)) \sup_{x \geq \tau_n} \rho_*(\alpha_{F_{s,\tau_n}}(x), \theta_{t,\tau_n}),$$

uniformly in  $s \in U_{t,h_n}$ , for all  $n \geq n_0$ . Note that  $\epsilon_1(n) \exp(\epsilon_0(n)) \leq C_3$  for large  $n$ , say  $n \geq n_1$ , where  $n_1$  is a constant. From (8.26) and (8.27), we obtain

$$\sup_{s \in U_{t,h_n}} \chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) \leq C_3 \left( \max(\delta_0^{-1}, \delta_0^{-2}) L_\alpha h_n^\beta + C_2(t) \tau_n^{\frac{1}{\gamma_t} - \frac{1}{\delta_t}} \right)^2$$

for any  $n \geq \max\{n_0, n_1\}$ . From this, we have the balance conditions for determining the oracle location  $\tau_n$ , and oracle bandwidth  $h_n$  :

$$(8.28) \quad \begin{aligned} \tau_n^{\frac{2}{\gamma_t} - \frac{2}{\delta_t}} &\asymp \frac{\log n_{t,h_n}}{n_{t,h_n}(C\tau_n^{-1/\gamma_t} + (1-C)\tau_n^{-1/\delta_t})}, \text{ and} \\ h_n^{2\beta} &\asymp \tau_n^{\frac{2}{\gamma_t} - \frac{2}{\delta_t}}. \end{aligned}$$

Optimizing in  $\tau_n$  gives  $\tau_n \asymp \left(\frac{n_{t,h_n}}{\log n_{t,h_n}}\right)^{\gamma_t \delta_t / (2\gamma_t - \delta_t)}$ , where  $\delta_t = \frac{\gamma_t}{1+\nu_t}$ . Taking into account (8.28) with  $n_{t,h_n} \asymp 2nh_n$ , we obtain, as  $n \rightarrow \infty$ ,

$$\begin{aligned} h_n &\asymp \left(\frac{\log n}{n}\right)^{\frac{1}{1+\beta(2+\nu_t^{-1})}}, \tau_n \asymp \left(\frac{\log n}{n}\right)^{\frac{-\beta\gamma_t\nu_t^{-1}}{1+\beta(2+\nu_t^{-1})}}, \\ \frac{\log n_{t,h}}{n_{t,h}(1-F_t(\tau_n))} &= O\left(\left(\frac{\log n}{n}\right)^{\frac{2\beta}{1+\beta(2+\nu_t^{-1})}}\right), \end{aligned}$$

and

$$\chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) = O\left(\frac{\log n_{t,h}}{n_{t,h}(1-F_t(\tau_n))}\right), \text{ uniformly in } s \in U_{t,h_n},$$

This implies, as  $n \rightarrow \infty$ ,

$$\sup_{s \in U_{t,h_n}} \chi^2(F_{s,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) = O\left(\frac{\log n_{t,h_n}}{n_{t,h_n}(1-F_t(\tau_n))}\right).$$

Finally, it is easy to verify that as  $n \rightarrow \infty$ ,

$$\bar{n}_{t,h_n,\tau_n} = \sum_{t_i \in U_{t,h_n}} W_{t,h_n}(t_i)(1-F_{t_i}(\tau_n)) \asymp n_{t,h_n}(1-F_t(\tau_n)),$$

which completes the proof.

**8.5. Proof of Theorem 4.1.** First we prove that there exists  $c^*$  such that, for  $n$  sufficiently large,

$$(8.29) \quad \mathbb{P}\left(\sup_{\tau_n \leq v \leq \tau} \text{LR}_{t,h_n}(v, \tau) > c^* \log n\right) \leq \frac{4}{n}.$$

By Proposition 8.4, we have, for any  $y > 0$ ,

$$\mathbb{P}\left(\sup_{\tau_n \leq v \leq \tau} \text{LR}_{t,h_n}(v, \tau) > 4y\right) \leq 2n^7 \exp\left(-\frac{y}{2} + \frac{d_{\tau_n}}{2}\right) + \frac{2}{n},$$

where  $d_{\tau_n} = \sum_{i=1}^n W_{t,h_n}(t_i) \chi^2(F_{t_i}, F_{t_i, \tau_n}, \theta_n)$ . Letting  $y = 16 \log n + \frac{d_{\tau_n}}{2}$ , we obtain

$$\mathbb{P} \left( \sup_{\tau_n \leq v \leq \tau} LR_{t,h_n}(v, \tau) > 4y \right) \leq \frac{4}{n}.$$

Moreover, by condition C1, we can choose  $4y < D = c^* \log n$ , for some constant  $c^*$  and  $n$  sufficiently large, which implies (8.29). Now the assertion of the theorem comes from (8.29) and the inclusion

$$\left\{ \sup_{X_{(k)} \geq \tau_n} \mathbb{Z}(X_{(k)}) > D \right\} \subseteq \left\{ \sup_{\tau_n \leq v \leq \tau} LR_{t,h_n}(v, \tau) > D \right\}.$$

**8.6. Proof of Theorem 4.2.** We use the notations  $\hat{n}_r = \hat{n}_{t,h_n,Y_r} = r - 1$ ,  $\hat{\theta}_r = \hat{\theta}_{t,h_n,Y_r}$  and

$$\mathbb{Z}(\tau_n) = \max_{\delta' \hat{n}_{t,h_n,\tau_n} \leq \hat{n}_l \leq (1-\delta'') \hat{n}_{t,h_n,\tau_n}} LR_{t,h}(\tau_n, Y_l),$$

for  $r = 2, \dots, n$ . Let  $\Omega_{t,h_n,\tau_n} = \bigcap_{Y_r \geq \tau_n} \{\mathbb{Z}(Y_r) \leq D\} \cap \{\mathbb{Z}(\tau_n) \leq D\}$  and  $\Omega_{t,h_n,\tau_n}^* = \Omega_{t,h_n,\tau_n} \cap \{\mathbb{Z}(Y_{k_0}) \leq D\}$ . Obviously, we have

$$\Omega_{t,h,\tau_n} \cap \{Y_{k_0} \geq \tau_n\} \subseteq \Omega_{t,h_n,\tau_n}^*.$$

Since  $\Omega_{t,h_n,\tau_n}^c \subseteq \{\sup_{\tau_n \leq v \leq \tau} LR_{t,h_n}(v, \tau) > D\}$ , by Theorem 4.1, it follows that

$$\lim_{n \rightarrow \infty} (\Omega_{t,h_n,\tau_n}^c) = 0.$$

From this and Lemma 8.6, we obtain,

$$(8.30) \quad \lim_{n \rightarrow \infty} (\Omega_{t,h_n,\tau_n}^*) = 1$$

By the definition of  $\hat{k}$ , on the set  $\Omega_{t,h,\tau_n}^*$ , it holds  $\hat{n}_{\hat{k}-1} \geq \hat{n}_{t,h_n,\tau_n}$ .

First we compare  $\hat{\theta}_{\hat{k}-1}$  and  $\hat{\theta}_{t,h_n,\tau_n}$ . To this end define the sequence of natural numbers  $m_i, i = 0, \dots, i^*$ , such that  $m_0 = \hat{k} - 1$  and  $\delta' \hat{n}_{m_i-1} \leq \hat{n}_{m_i} \leq \frac{1}{2} \hat{n}_{m_i-1} \leq (1 - \delta'') \hat{n}_{m_i-1}$ , for  $i = 1, \dots, i^*$ , where  $i^*$  such that  $\delta' \hat{n}_{m_{i^*}} \leq \hat{n}_{t,h_n,\tau_n} \leq (1 - \delta'') \hat{n}_{m_{i^*}}$ . Denote  $\hat{n}_{m_{i^*}+1} = \hat{n}_{t,h_n,\tau_n}$  and  $\hat{\theta}_{m_{i^*}+1} = \hat{\theta}_{t,h_n,\tau_n}$ . Since, on the set  $\Omega_{t,h,\tau_n}^*$ ,

$$\mathbb{Z}(\tau_n) \leq D = c^* \log n$$

and

$$\mathbb{Z}(X_{(r)}) \leq D = c^* \log n, \quad \text{for } k_0 \leq r \leq \widehat{k} - 1,$$

by (4.3) with  $s = Y_{m_{i-1}} \leq \tau = Y_{m_i}$ , we have

$$\widehat{n}_{m_i} \mathcal{K}(\widehat{\theta}_{m_i}, \widehat{\theta}_{m_{i-1}}) \leq LR_{t, h_n}(s, \tau) \leq D, \quad i = 1, \dots, i^*.$$

In the same way, with  $s = Y_{m_{i^*}} \leq \tau = \tau_n$ , we have

$$\widehat{n}_{t, h_n, \tau_n} \mathcal{K}(\widehat{\theta}_{t, h_n, \tau_n}, \widehat{\theta}_{m_{i^*}}) \leq LR_{t, h_n}(s, \tau) \leq D.$$

This, in turn, implies

$$\sum_{i=1}^{i^*+1} \sqrt{\mathcal{K}(\widehat{\theta}_{m_i}, \widehat{\theta}_{m_{i-1}})} \leq D^{1/2} \sum_{i=1}^{i^*+1} \widehat{n}_{m_i}^{-1/2}.$$

Taking into account that  $\widehat{n}_{m_i} \leq \widehat{n}_{m_{i-1}}$ , for  $i = 1, \dots, i^* + 1$ , we obtain

$$\sum_{i=1}^{i^*+1} \widehat{n}_{m_i}^{-1/2} \leq \widehat{n}_{m_{i^*+1}}^{-1/2} \sum_{i=1}^{i^*+1} 2^{-(i^*-i+1)/2} \leq (2 + \sqrt{2}) \widehat{n}_{m_{i^*+1}}^{-1/2}.$$

According to Lemma 8.1 and 8.2 in Gama and Spokoiny (2008), for  $n$  sufficiently large, it holds

$$\sqrt{\mathcal{K}(\widehat{\theta}_{m_{i^*+1}}, \widehat{\theta}_{m_0})} \leq \frac{3}{2} \sum_{i=1}^{i^*+1} \sqrt{\mathcal{K}(\widehat{\theta}_{m_i}, \widehat{\theta}_{m_{i-1}})} \leq \frac{3}{2} (2 + \sqrt{2}) D^{1/2} \widehat{n}_{m_{i^*+1}}^{-1/2},$$

and

$$\begin{aligned} \sqrt{\mathcal{K}(\widehat{\theta}_{\widehat{k}-1}, \widehat{\theta}_{t, h_n, \tau_n})} &= \sqrt{\mathcal{K}(\widehat{\theta}_{m_0}, \widehat{\theta}_{m_{i^*+1}})} \\ &\leq \frac{9}{4} (2 + \sqrt{2}) D^{1/2} \widehat{n}_{m_{i^*+1}}^{-1/2} \\ (8.31) \qquad \qquad \qquad &= \frac{9}{4} (2 + \sqrt{2}) \sqrt{c^* \frac{\log n}{\widehat{n}_{t, h_n, \tau_n}}}. \end{aligned}$$

Now we shall compare  $\widehat{\theta}_{\widehat{k}-1} = \widehat{\theta}_{m_0}$  and  $\widehat{\theta}_{\widehat{l}}$ . Recall that  $\widehat{l}$  satisfies

$$\delta' \widehat{n}_{t, h_n, \tau_n} \leq \delta' \widehat{n}_{\widehat{k}-1} < \delta' \widehat{n}_{\widehat{k}} \leq \widehat{n}_{\widehat{l}} \leq (1 - \delta'') \widehat{n}_{\widehat{k}}.$$

Since, on the set  $\Omega_{t,h_n,\tau_n}^*$ , we have  $LR_{t,h}(Y_{\widehat{k}_{-1}}, Y_l) \leq D = c^* \log n$ , it follows that

$$(8.32) \quad \sqrt{\mathcal{K}(\widehat{\theta}_{\widehat{l}}, \widehat{\theta}_{\widehat{k}_{-1}})} \leq \sqrt{c^* \frac{\log n}{\widehat{n}_{\widehat{l}}}} \leq \sqrt{\frac{c^* \log n}{\delta' \widehat{n}_{t,h_n,\tau_n}}}$$

Combining (8.31) and (8.32), by Lemma 8.2 in Gama and Spokoiny (2008), it follows that, on the set  $\Omega_{t,h_n,\tau_n}^*$ ,

$$\sqrt{\mathcal{K}(\widehat{\theta}_{\widehat{l}}, \widehat{\theta}_{t,h_n,\tau_n})} \leq \sqrt{cc^* \frac{\log n}{\widehat{n}_{t,h_n,\tau_n}}},$$

where  $c$  is a positive constant. Taking into account (8.30), we have

$$\mathbb{P} \left( \mathcal{K}(\widehat{\theta}_{\widehat{l}}, \widehat{\theta}_{t,h_n,\tau_n}) \leq cc^* \frac{\log n}{\widehat{n}_{t,h_n,\tau_n}} \right) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

Hence, by Lemma 8.5, we obtain

$$\mathbb{P} \left( \mathcal{K}(\widehat{\theta}_{\widehat{l}}, \widehat{\theta}_{t,h_n,\tau_n}) \leq cc^* \frac{\log n}{\widehat{n}_{t,h_n,\tau_n}} \right) \rightarrow 1 \text{ as } n \rightarrow \infty,$$

the result follows.

**8.7. Proof of Theorem 4.4.** From the decomposition

$$\mathcal{K}(F_{t,\tau}, G_{\tau,\theta}) = \mathcal{K}(F_{t,\tau}, G_{\tau,\theta_{t,\tau}}) + \int_{\tau}^{\infty} \log \frac{dG_{\tau,\theta_{t,\tau}}}{dG_{\tau,\theta}} dF_{t,\tau}$$

and the identity  $\int_{\tau}^{\infty} \log \frac{dG_{\tau,\theta_{t,\tau}}}{dG_{\tau,\theta}} dF_{t,\tau} = \mathcal{K}(\theta_{t,\tau}, \theta)$ , we have, for any  $\theta > 0$  and any  $\tau \geq x_0$ ,

$$(8.33) \quad \mathcal{K}(F_{t,\tau}, G_{\tau,\theta}) = \mathcal{K}(F_{t,\tau}, G_{\tau,\theta_{t,\tau}}) + \mathcal{K}(\theta_{t,\tau}, \theta).$$

Using (8.33) with  $\tau = \tau_n$ ,  $h = h_n$  and  $\theta = \widehat{\theta}_{t,h_n,\widehat{\tau}_n}$ , we have

$$(8.34) \quad \mathcal{K}(F_{t,\tau_n}, G_{\tau_n,\widehat{\theta}_{t,h_n,\widehat{\tau}_n}}) = \mathcal{K}(F_{t,\tau_n}, G_{\tau_n,\theta_{t,\tau_n}}) + \mathcal{K}(\theta_{t,\tau_n}, \widehat{\theta}_{t,h_n,\widehat{\tau}_n}).$$

From Theorem 4.3 and Lemma 8.1 in Grama and Spokoiny (2008), we have, for  $n$  sufficiently large,

$$(8.35) \quad \mathcal{K}(\theta_{t,\tau_n}, \widehat{\theta}_{t,h_n,\widehat{\tau}_n}) \leq \frac{9}{4} \mathcal{K}(\widehat{\theta}_{t,h_n,\widehat{\tau}_n}, \theta_{t,\tau_n}) = O_{\mathbb{P}} \left( \frac{\log n}{\widehat{n}_{t,h_n,\tau_n}} \right)$$



as  $n \rightarrow \infty$ . Using condition (4.6) and the bounds

$$\mathcal{K}(F_{t,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) \leq \log(1 + \chi^2(F_{t,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}})) \leq \chi^2(F_{t,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}),$$

we obtain

$$(8.36) \quad \mathcal{K}(F_{t,\tau_n}, G_{\tau_n, \theta_{t,\tau_n}}) = O\left(\frac{\log n_{t,h_n}}{\bar{n}_{t,h_n,\tau_n}}\right).$$

Combining (8.34), (8.35) and (8.36), it follows that as  $n \rightarrow \infty$

$$\mathcal{K}\left(F_{t,\tau_n}, G_{\tau_n, \hat{\theta}_{t,h_n, \hat{\tau}_n}}\right) = O_{\mathbb{P}}\left(\frac{\log n_{t,h_n}}{\bar{n}_{t,h_n,\tau_n}}\right).$$