Peaks-Over-Threshold modeling under random censoring

Jan Beirlant, *University of Leuven*

Armelle Guillou, *Université de Strasbourg*

Gwladys Toulemonde, *Université Montpellier 2*

Abstract. Recently, the topic of extreme value under random censoring has been considered. Different estimators for the index have been proposed (see Beirlant *et al.*, 2007). All of them are constructed as the classical estimators (without censoring) divided by the proportion of non-censored observations above a certain threshold. Their asymptotic normality have been established by Einmahl *et al.* (2008). An alternative approach consists of using the Peaks-Over-Threshold method (Balkema and de Haan, 1974; Smith, 1987) and to adapt the likelihood to the context of censoring. This leads to $ML-$estimators whose asymptotic properties are still unknown. The aim of this paper is to propose one-step approximations, based on the Newton-Raphson algorithm. Based on a small simulation study, the one-step estimators are shown to be close approximations to the $ML-$estimators. Also the asymptotic normality of the one-step estimators has been established, whereas in case of the $ML-$estimators it is still an open problem. The proof of our result, whose approach is new in the Peaks-Over-Threshold context, is in the spirit of Lehmann's theory (1991).

AMS 2000 subject classifications. 62G05, 62G20, 62G32, 62N02.

Keywords. Asymptotic normality, extreme value index, random censoring, Newton-Rapshon algorithm.

Short title. Statistics of censored extremes

Corresponding author. Armelle Guillou, Université de Strasbourg, IRMA, 7 rue René Descartes, 67084 Strasbourg cedex, France. E-mail: guillou@math.u-strasbg.fr

1 Introduction

Let X_i, $i = 1, ..., n$, be independent and identically distributed (i.i.d.) random variables with common continuous distribution function F, and let Y_i, $i = 1, ..., n$, be a second i.i.d. sequence with continuous distribution function G independent of the $X-$sequence. One now observes $Z_i = X_i \wedge Y_i$ and $\delta_i = 1\!\!1_{X_i \leq Y_i}$, $i = 1, ..., n$. Let H be the distribution function of Z_1 with $\tau_H = \sup\{x : H(x) < 1\}$ the supremum of the support of H. Furthermore, let $\overline{H}^1(x) = \mathbb{P}(Z > z, \delta = 1) = \mathbb{P}(z < X \leq Y)$.

Aside from the classical random right censoring model, we will assume that both F and G are in the domain of attraction of an extreme value distribution, with extreme value indices γ_1 and γ_2 respectively. Remark that this semi-parametric assumption is very general and hardly puts any restriction on the applicability of the results. Moreover this implies that the extreme value index of H exists and it will be denoted in the sequel by γ.

The aim in this paper is to estimate the extreme value index γ_1, although no observations from F are available. This topic has already been studied in Beirlant *et al.* (2007), where different estimators for the extreme value index have been introduced. For instance generalizations of the well-known Hill estimator (Hill, 1975), moment estimator (Dekkers *et al.*, 1989) and $UH-$estimator (Beirlant *et al.*, 1996) have been proposed in the context of censored observations. All these estimators are constructed in the same way: the classical non-censored estimator based on the $Z-$observations is divided by the proportion of non-censored observations above the threshold t, that is

$$\widehat{\gamma}_{Z,t}^{(c,.)} := \frac{\widehat{\gamma}_{Z,t}^{(.)}}{\frac{1}{N_t}\sum_{j=1}^{n} \delta_j 1\!\!1_{\{Z_j > t\}}} \tag{1}$$

with N_t the number of absolute excesses over t. $\widehat{\gamma}_{Z,t}^{(.)}$ could be any estimator not adapted to censoring, in particular those previously mentioned: $\widehat{\gamma}_{Z,t}^{(H)}, \widehat{\gamma}_{Z,t}^{(M)}$, or $\widehat{\gamma}_{Z,t}^{(UH)}$. Recently, Einmahl *et al.* (2008) have established, in a unified way, the asymptotic normality of any extreme value index estimator of the form (1) in case where $t = Z_{n-k,n}$, the $(k+1)$-largest observation among $Z_1, ..., Z_n$.

However, another approach to obtain estimators in the case of censoring is to adapt the likelihood to this context. Indeed, recall that the maximum likelihood method relies on the results given by Balkema and de Haan (1974) and Pickands (1975), stating that the limit distribution of the absolute exceedances $E_j = X_j - t$, given $X_j > t$, over a threshold t when $t \to \tau_F$, is given by the generalized Pareto distribution. In the case of censoring, setting now $E_j = Z_j - t$, given $Z_j > t$, we can easily adapt the likelihood (see e.g. Andersen *et al.*, 1993, p. 411) to

$$L(\gamma_1, \sigma_{1,t}) = \prod_{j=1}^{N_t} \left[g_{\gamma_1,\sigma_{1,t}}(E_j)\right]^{\delta_j} \left[1 - G_{\gamma_1,\sigma_{1,t}}(E_j)\right]^{1-\delta_j} \tag{2}$$

where $1 - G_{\gamma_1,\sigma_{1,t}}(x) = \left(1 + \frac{\gamma_1}{\sigma_{1,t}} x\right)^{-\frac{1}{\gamma_1}}$ and $g_{\gamma_1,\sigma_{1,t}}$ is the associated density.

As usual in this context, we cannot explicitly deduce the expression of the two estimators of γ_1 and $\sigma_{1,t}$, but we can obtain them numerically. However, it is very hard to obtain the asymptotic normality of these estimators. Recall that in the case of no-censoring, this topic has been studied by several authors, even recently (see Smith, 1987; Drees, 1998; Drees *et al.*, 2004). In the case of censoring, this normality is of course more difficult to derive and at this moment the asymptotic normality is an open problem. In this paper, we propose an alternative approach which consists of solving the $ML-$equations based on a one-step Newton-Raphson approximation (as discussed for instance in Lehmann, 1991, Sections 6.3 and 6.4). This leads to new estimators adapted to censoring for $(\gamma_1, \sigma_{1,t})$, denoted in the sequel by $(\widehat{\gamma}_{Z,t}^{(c,OS)}, \widehat{\sigma}_{Z,t}^{(c,OS)})$. More precisely, in the spirit of Theorem 4.2, Chapter 6 in Lehmann (1991), the proposed one-step estimators are defined by the following way

$$\begin{pmatrix} \widehat{\gamma}_{Z,t}^{(c,OS)} \\ \\ \frac{\widehat{\sigma}_{Z,t}^{(c,OS)}}{\sigma_{1,t}} \end{pmatrix} = \begin{pmatrix} \widehat{\gamma}_{Z,t}^{(c,.)} \\ \\ \frac{\widehat{\sigma}_{Z,t}^{(c,.)}}{\sigma_{1,t}} \end{pmatrix} - \begin{pmatrix} L''_{11} & \sigma_{1,t}L''_{12} \\ \\ \sigma_{1,t}L''_{12} & \sigma_{1,t}^2 L''_{22} \end{pmatrix}^{-1} \begin{pmatrix} L'_1 \\ \\ \sigma_{1,t}\, L'_2 \end{pmatrix} \tag{3}$$

where L'_i and L''_{ij}, $i = 1,2, j = 1,2$ denote the first, resp. second, derivatives evaluated at $\left(\widehat{\gamma}_{Z,t}^{(c,.)}, \widehat{\sigma}_{Z,t}^{(c,.)}\right)$ of the log-likelihood $\log L(\gamma_1, \sigma_{1,t})$. The estimators $\widehat{\gamma}_{Z,t}^{(c,.)}$ and $\widehat{\sigma}_{Z,t}^{(c,.)}$ are any initial estimators for γ_1 and $\sigma_{1,t}$ such that $\sqrt{N_t}\Big(\widehat{\gamma}_{Z,t}^{(c,.)} - \gamma_1 \quad \frac{\widehat{\sigma}_{Z,t}^{(c,.)}}{\sigma_{1,t}} - 1\Big)'$ is asymptotically normal. For instance, it could be any of the estimators mentioned in Section 1.

Note that, in the same way, it is also possible to construct two-steps estimators.

In the sequel, we assume that $\text{sgn}(\gamma_1) = \text{sgn}(\gamma_2)$ and $\tau_F = \tau_G$ which corresponds to one of the following cases:

$$\begin{cases} \text{case 1:} \quad \gamma_1 > 0, \gamma_2 > 0, \\ \text{case 2:} \quad \gamma_1 < 0, \gamma_2 < 0,\ \tau_F = \tau_G, \\ \text{case 3:} \quad \gamma_1 = \gamma_2 = 0,\ \tau_F = \tau_G. \end{cases}$$

The other possibilities are not very interesting. Typically they are very close to the « completely censored situation » where estimation is impossible (this holds in particular when $\gamma_1 > 0$ and $\gamma_2 < 0$) or the « uncensored case » which has been studied in detail in the literature (this holds in particular when $\gamma_1 < 0$ and $\gamma_2 > 0$). Note also that in case 1, F and G are of Pareto-type which means that $1-F(x) = x^{-1/\gamma_1}\ell_F(x)$ and $1-G(x) = x^{-1/\gamma_2}\ell_G(x)$ with ℓ_F and ℓ_G two slowly varying functions at infinity. Similarly, we have in case 2, $1-F(x) = (\tau_F - x)^{-1/\gamma_1}\ell_F((\tau_F - x)^{-1})$ and $1 - G(x) = (\tau_G - x)^{-1/\gamma_2}\ell_G((\tau_G - x)^{-1})$. As a consequence, in these two first cases, it is easy to check that $\gamma = \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}$ since $1 - H(x) = (1 - F(x))(1 - G(x))$. In the last one, $\gamma = 0$. For convenience we always use the notation $\gamma =: \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}$.

The aim of this paper is to generalize the Peaks-Over-Threshold method to the case of randomly right censored data, and more specifically to establish the asymptotic normality of the pair $(\widehat{\gamma}_{Z,t}^{(c,OS)}\ \widehat{\sigma}_{Z,t}^{(c,OS)})'$ correctly normalized. This is the goal of our Section 2. As such we also provide a new method for deriving this result, offering an alternative to the existing approaches for proving such results in the uncensored i.i.d. case. Then various examples are detailed in Section 3 with a small simulation study in order to illustrate the similar behaviour between the one-step estimators and the $ML-$estimators, whose asymptotic behaviour is unknown. The proof of our theorem is postponed to Section 4 and the technical details are given in the appendix.

2 Main Result

We denote by U_F (resp. by U_G) the tail quantile function of F (resp. of G), that is $U_F(x) := F^{\leftarrow}\left(1-\frac{1}{x}\right) = \inf\{y : F(y) \geq 1-\frac{1}{x}\}$ (resp. $U_G(x) := G^{\leftarrow}\left(1-\frac{1}{x}\right)$). We assume that there exists two positive auxiliary functions a_F and a_G such that

$$\lim_{x\to\infty} \frac{U_F(xu)-U_F(x)}{a_F(x)} = \int_1^u v^{\gamma_1-1}dv =: h_{\gamma_1}(u) \quad \text{for } u>0 \tag{4}$$

$$\lim_{x\to\infty} \frac{U_G(xu)-U_G(x)}{a_G(x)} = \int_1^u v^{\gamma_2-1}dv =: h_{\gamma_2}(u) \quad \text{for } u>0. \tag{5}$$

We will say that the tail quantile function U_F (resp. U_G) satisfies the extreme value condition $C_{\gamma_1}(a_F)$ (resp. $C_{\gamma_2}(a_G)$) if (4) (resp. (5)) holds with the auxiliary function a_F (resp. a_G). In order to establish the asymptotic normality of our estimators, we need a second order condition, which can be expressed as follows. Using the theory of generalized regular variation of second order outlined in de Haan and Stadtmüller (1996), one assumes that for all $u>0$:

$$\frac{U_F(xu)-U_F(x)}{a_F(x)} - h_{\gamma_1}(u) \sim a_{2,F}(x)k_F(u), \quad x\uparrow\infty \tag{6}$$

$$\frac{U_G(xu)-U_G(x)}{a_G(x)} - h_{\gamma_2}(u) \sim a_{2,G}(x)k_G(u), \quad x\uparrow\infty \tag{7}$$

where $a_{2,F}, a_{2,G} \to 0$ are regularly varying $\mathcal{R}_{\rho_1}, \mathcal{R}_{\rho_2}$ functions respectively, with $\rho_1 \le 0$ and $\rho_2 \le 0$ and

$$\begin{aligned} k_F(u) &= A_F\, h_{\gamma_1+\rho_1}(u) + c_F \int_1^u t^{\gamma_1-1} h_{\rho_1}(t)\, dt \\ k_G(u) &= A_G\, h_{\gamma_2+\rho_2}(u) + c_G \int_1^u t^{\gamma_2-1} h_{\rho_2}(t)\, dt, \end{aligned}$$

for suitable constants A_F, A_G, c_F and c_G.

Note that if $\rho_1 < 0$ and $\rho_2 < 0$, then an appropriate choice of the auxiliary functions a_F and a_G results in a simplification of the limit functions, namely $k_F(u) = (\rho_1^{-1} c_F + A_F) h_{\gamma_1+\rho_1}(u)$ and $k_G(u) = (\rho_2^{-1} c_G + A_G) h_{\gamma_2+\rho_2}(u)$ (see Proposition 3.2 in Beirlant *et al.*, 2004).

We transform the remainder condition that has been stated above in terms of U_F (resp. U_G) towards a statement based on the distribution function F (resp. G). Sometimes this can be more easy to handle. Assuming that U_F and U_G are continuous and satisfy (4)-(7), then we have that (see for instance Theorem 3.3 in Beirlant *et al.*, 2004)

$$\frac{1}{\chi_1(t)} \left\{ \frac{1 - F(t + v\,\sigma_{1,t})}{1 - F(t)} - \eta_{\gamma_1}(v) \right\} \longrightarrow \quad \psi_1(v), \quad t \to \tau_F \tag{8}$$

$$\frac{1}{\chi_2(t)} \left\{ \frac{1 - G(t + v\,\sigma_{2,t})}{1 - G(t)} - \eta_{\gamma_2}(v) \right\} \longrightarrow \quad \psi_2(v), \quad t \to \tau_G \tag{9}$$

where $\chi_1(t) = a_{2,F}\left(U_F^{\leftarrow}(t)\right), \chi_2(t) = a_{2,G}\left(U_G^{\leftarrow}(t)\right), \eta_\gamma(v) = (1 + \gamma v)^{-1/\gamma}$, $\psi_1(v) = \eta_{\gamma_1}^{1+\gamma_1}(v) k_F\left(\eta_{\gamma_1}^{-1}(v)\right)$ and $\psi_2(v) = \eta_{\gamma_2}^{1+\gamma_2}(v) k_G\left(\eta_{\gamma_2}^{-1}(v)\right)$.

Before stating our main result, we need to introduce some notations. Let

$$B_{1,t}(\gamma_1, \sigma_{1,t}) := \begin{cases} \dfrac{1}{\gamma_1 \sigma_{1,t}} \dfrac{1}{\overline{H}(t)} \displaystyle\int_t^{\tau_H} \dfrac{\overline{H}(x)}{1 + \frac{\gamma_1}{\sigma_{1,t}}(x-t)} dx - \dfrac{1}{\gamma_1 \sigma_{1,t}} \dfrac{1}{\overline{H}(t)} \displaystyle\int_t^{\tau_H} \dfrac{\overline{H}(x)}{(1 + \frac{\gamma_1}{\sigma_{1,t}}(x-t))^2} dx \\ \qquad\qquad - \dfrac{1}{\sigma_{1,t}} \dfrac{1}{\overline{H}(t)} \displaystyle\int_t^{\tau_H} \dfrac{\overline{H}^1(x)}{(1 + \frac{\gamma_1}{\sigma_{1,t}}(x-t))^2} dx & \text{if } \gamma_1 \neq 0 \\ \\ \dfrac{1}{\sigma_{1,t}^2} \dfrac{1}{\overline{H}(t)} \displaystyle\int_t^{\tau_H} (x-t)\, \overline{H}(x) dx - \dfrac{1}{\sigma_{1,t}} \dfrac{1}{\overline{H}(t)} \displaystyle\int_t^{\tau_H} \overline{H}^1(x) dx & \text{if } \gamma_1 = 0 \end{cases}$$

and

$$B_{2,t}(\gamma_1,\sigma_{1,t}) := \begin{cases} -\dfrac{\overline{H}^1(t)}{\overline{H}(t)} + \dfrac{1}{\sigma_{1,t}}\dfrac{1}{\overline{H}(t)}\displaystyle\int_t^{\tau_H} \dfrac{\overline{H}(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx \\ \qquad\qquad + \dfrac{\gamma_1}{\sigma_{1,t}}\dfrac{1}{\overline{H}(t)}\displaystyle\int_t^{\tau_H} \dfrac{\overline{H}^1(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx & \text{if } \gamma_1 \neq 0 \\ -\dfrac{\overline{H}^1(t)}{\overline{H}(t)} + \dfrac{1}{\sigma_{1,t}}\dfrac{1}{\overline{H}(t)}\displaystyle\int_t^{\tau_H} \overline{H}(x)dx & \text{if } \gamma_1 = 0. \end{cases}$$

The functions $B_{1,t}$ and $B_{2,t}$ are computed from expected values of L_1' and L_2'. In a classical parametric setting, $B_{1,t}$ and $B_{2,t}$ are equal to zero, while here in a semiparametric extreme value setting they tend to zero which leads to some bias in the estimates. We have now the following result.

Theorem. *Suppose that U_F and U_G are continuous and satisfy (4)-(7). Moreover, if $\gamma < 0$, we assume that the slowly varying function ℓ_F is normalised. Assuming when $t \to \tau_H$,*

$$\sqrt{N_t}\, B_{1,t}(\gamma_1,\sigma_{1,t}) \xrightarrow{\mathbb{P}} \alpha_1 \in \mathbb{R}, \tag{10}$$

$$\sqrt{N_t}\, B_{2,t}(\gamma_1,\sigma_{1,t}) \xrightarrow{\mathbb{P}} \alpha_2 \in \mathbb{R}, \tag{11}$$

$$\textit{and} \quad \frac{\sigma_{2,t}}{\sigma_{1,t}+\sigma_{2,t}} \longrightarrow C \in (0,1) \quad \textit{if} \quad \gamma_1 = \gamma_2 = 0 \quad \textit{and} \quad \tau_F = \tau_G, \tag{12}$$

then we have, in case $\gamma > -\frac{1}{2}$,

$$\sqrt{N_t}\begin{pmatrix} \widehat{\gamma}_{Z,t}^{(c,OS)} - \gamma_1 \\ \\ \dfrac{\widehat{\sigma}_{Z,t}^{(c,OS)}}{\sigma_{1,t}} - 1 \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mu,\Sigma) \tag{13}$$

where

$$\mu := \begin{cases} \begin{pmatrix} \left(\dfrac{\gamma_1}{\gamma}\right)^2 (1+\gamma)\left[\dfrac{\gamma_1}{\gamma}(1+\gamma)\alpha_1 - \alpha_2\right] \\ \dfrac{\gamma_1}{\gamma}(1+\gamma)\left[-\dfrac{\gamma_1}{\gamma}\alpha_1 + 2\alpha_2\right] \end{pmatrix} & \textit{if} \quad \gamma_1\gamma_2 > 0 \quad \textit{and} \quad \tau_F = \tau_G \\ \\ \begin{pmatrix} C^{-3}\left[\alpha_1 - \alpha_2\, C\right] \\ C^{-2}\left[-\alpha_1 + 2\alpha_2\, C\right] \end{pmatrix} & \textit{if} \quad \gamma_1 = \gamma_2 = 0 \quad \textit{and} \quad \tau_F = \tau_G \end{cases}$$

and

$$\Sigma := \begin{cases} \begin{pmatrix} \left(\dfrac{\gamma_1}{\gamma}\right)^3 (1+\gamma)^2 & -\left(\dfrac{\gamma_1}{\gamma}\right)^2 (1+\gamma) \\ & \\ -\left(\dfrac{\gamma_1}{\gamma}\right)^2 (1+\gamma) & 2\,\dfrac{\gamma_1}{\gamma}(1+\gamma) \end{pmatrix} & \textit{if} \quad \gamma_1\gamma_2 > 0 \quad \textit{and} \quad \tau_F = \tau_G \\ & \\ \begin{pmatrix} C^{-3} & -C^{-2} \\ -C^{-2} & 2\,C^{-1} \end{pmatrix} & \textit{if} \quad \gamma_1 = \gamma_2 = 0 \quad \textit{and} \quad \tau_F = \tau_G. \end{cases}$$

3 Examples and a small simulation study

In this section, we consider three examples:

- a Burr distribution censored by another Burr distribution;
- a ReverseBurr distribution censored by another ReverseBurr distribution;
- a Logistic distribution censored by a Logistic distribution.

Remark that these three examples cover the three cases considered in this paper. For each example, we evaluate explicitly the main bias terms $B_{1,t}(\gamma_1, \sigma_{1,t})$ and $B_{2,t}(\gamma_1, \sigma_{1,t})$. This leads to a reformulation of conditions (10)-(11) of our theorem and also we prove that (12) is fulfilled in case 3. Then, in a second part of this section, we provide a small simulation study to illustrate the correspondance, for these three distributions, between the one-step estimators and the $ML-$estimators obtained by direct optimization of the likelihood (2).

3.1 Various examples

- **Example 1:** $X \sim \text{Burr}(\beta_1, \tau_1, \lambda_1)$ and $Y \sim \text{Burr}(\beta_2, \tau_2, \lambda_2)$, $\beta_i, \tau_i, \lambda_i > 0$, $i = 1, 2$, and $\tau_1 \neq 1$.

In that case

$$\begin{aligned}
1-F(x) &= \left(\frac{\beta_1}{\beta_1+x^{\tau_1}}\right)^{\lambda_1} = x^{-\tau_1\lambda_1}\beta_1^{\lambda_1}\left(1+\beta_1 x^{-\tau_1}\right)^{-\lambda_1}, x>0;\\
1-G(x) &= \left(\frac{\beta_2}{\beta_2+x^{\tau_2}}\right)^{\lambda_2} = x^{-\tau_2\lambda_2}\beta_2^{\lambda_2}\left(1+\beta_2 x^{-\tau_2}\right)^{-\lambda_2}, x>0.
\end{aligned}$$

We can infer that $\gamma_1 = \dfrac{1}{\lambda_1\tau_1}, \gamma = \dfrac{1}{\lambda_1\tau_1+\lambda_2\tau_2}$ and $\sigma_{1,t} = \gamma_1\, t\left(1+\dfrac{\beta_1}{\tau_1}t^{-\tau_1}(1+o(1))\right)$.

Direct, but tedious computations, lead to

$$B_{1,t}(\gamma_1,\sigma_{1,t}) = \begin{cases} -\dfrac{\gamma^2}{\gamma_1^2}\dfrac{\beta_1(1-\tau_1)(1+\gamma-\gamma^2\tau_1)}{\tau_1(1+\gamma\tau_1)(1+\gamma\tau_1+\gamma)(1+\gamma)(1+2\gamma)}\,t^{-\tau_1} & \text{if } \tau_1\le\tau_2\\ \\ o\left(t^{-\tau_2}\right) & \text{if } \tau_1>\tau_2 \end{cases}$$

$$B_{2,t}(\gamma_1,\sigma_{1,t}) = \begin{cases} -\dfrac{\gamma}{\gamma_1}\dfrac{\beta_1(1-\tau_1)(1+\gamma)}{\tau_1(1+\gamma\tau_1+\gamma)(1+2\gamma)}\,t^{-\tau_1} & \text{if } \tau_1\le\tau_2\\ \\ o\left(t^{-\tau_2}\right) & \text{if } \tau_1>\tau_2. \end{cases}$$

Now, if we recall that N_t is by definition the number of absolute excesses over t, it is distributed according to a Binomial $(n,\overline{H}(t))$ distribution. Then $\frac{N_t}{n\overline{H}(t)}\to 1$ in probability as $t\to\infty$ and therefore, a sufficient condition for (10) and (11) is

$$\sqrt{n\overline{H}(t)}\,t^{-\tau}\longrightarrow\alpha \quad\text{or equivalently}\quad \sqrt{n}\,t^{-\tau-\frac{1}{2\gamma}}\longrightarrow\alpha,\quad \text{as } t\to\infty,$$

where $\alpha\in\mathbb{R}$ and $\tau=\min(\tau_1,\tau_2)$.

- **Example 2:** $X\sim$ ReverseBurr $(\beta_1,\tau_1,\lambda_1,\tau_H)$ and $Y\sim$ ReverseBurr $(\beta_2,\tau_2,\lambda_2,\tau_H)$, $\beta_i,\tau_i,\lambda_i,\tau_H>0,\ i=1,2$.

In that case

$$\begin{aligned}
1-F(x) &= \left(\frac{\beta_1}{\beta_1+(\tau_H-x)^{-\tau_1}}\right)^{\lambda_1} = (\tau_H-x)^{\tau_1\lambda_1}\beta_1^{\lambda_1}\left(1+\beta_1(\tau_H-x)^{\tau_1}\right)^{-\lambda_1}, x<\tau_H;\\
1-G(x) &= \left(\frac{\beta_2}{\beta_2+(\tau_H-x)^{-\tau_2}}\right)^{\lambda_2} = (\tau_H-x)^{\tau_2\lambda_2}\beta_2^{\lambda_2}\left(1+\beta_2(\tau_H-x)^{\tau_2}\right)^{-\lambda_2}, x<\tau_H
\end{aligned}$$

leading to $\gamma_1 = -\dfrac{1}{\tau_1\lambda_1}, \gamma = -\dfrac{1}{\tau_1\lambda_1 + \tau_2\lambda_2}$ and $\sigma_{1,t} = -\gamma_1(\tau_H - t)\left(1 - \dfrac{\beta_1}{\tau_1}(\tau_H - t)^{\tau_1}(1 + o(1))\right)$.

Direct computations yield

$$B_{1,t}(\gamma_1, \sigma_{1,t}) = \begin{cases} \dfrac{\gamma^2}{\gamma_1^2} \dfrac{\beta_1(1+\tau_1)(1+\gamma+\gamma^2\tau_1)}{\tau_1(1-\gamma\tau_1)(1-\gamma\tau_1+\gamma)(1+\gamma)(1+2\gamma)} (\tau_H - t)^{\tau_1} & \text{if } \tau_1 \leq \tau_2 \\ \\ o\left((\tau_H - t)^{\tau_2}\right) & \text{if } \tau_1 > \tau_2 \end{cases}$$

$$B_{2,t}(\gamma_1, \sigma_{1,t}) = \begin{cases} \dfrac{\gamma}{\gamma_1} \dfrac{\beta_1(1+\tau_1)(1+\gamma)}{\tau_1(1-\gamma\tau_1+\gamma)(1+2\gamma)} (\tau_H - t)^{\tau_1} & \text{if } \tau_1 \leq \tau_2 \\ \\ o\left((\tau_H - t)^{\tau_2}\right) & \text{if } \tau_1 > \tau_2. \end{cases}$$

Consequently, a sufficient condition for (10) and (11) is

$$\sqrt{n\overline{H}(t)}\,(\tau_H - t)^{\tau} \longrightarrow \alpha \quad \text{or equivalently} \quad \sqrt{n}\,(\tau_H - t)^{\tau - \frac{1}{2\gamma}} \longrightarrow \alpha, \quad \text{as } t \to \tau_H,$$

where $\alpha \in \mathbb{R}$ and $\tau = \min(\tau_1, \tau_2)$.

• **Example 3:** $X, Y \sim$ Logistic.

In that case

$$1 - F(x) = 1 - G(x) = \frac{2}{1 + e^x},\ x > 0.$$

Hence $\gamma_1 = \gamma = 0$ and $\sigma_{1,t} = 1$. This implies that

$$B_{1,t}(\gamma_1, \sigma_{1,t}) = \frac{1}{9}\, e^{-t}\,(1 + o(1))$$

and

$$B_{2,t}(\gamma_1, \sigma_{1,t}) = \frac{1}{3}\, e^{-t}\,(1 + o(1)).$$

Therefore a necessary and sufficient condition for (10) and (11) is

$$\sqrt{n\overline{H}(t)}\, e^{-t} \longrightarrow \alpha \quad \text{or equivalently} \quad \sqrt{n}\, e^{-2t} \longrightarrow \alpha, \quad \text{as } t \to \infty,$$

where $\alpha \in \mathbb{R}$. Also, since $\sigma_{1,t} = \sigma_{2,t} = 1$, (12) is clearly fullfilled.

3.2 Simulations

In order to illustrate these three examples, we simulate 300 samples of size 500 from the following distributions:

- a Burr $(10, 4, 1)$ censored by a Burr $(10, 1, 0.5)$: $\gamma_1 = 0.25, \gamma_2 = 2$;
- a ReverseBurr $(1, 8, 0.5, 10)$ censored by a ReverseBurr $(10, 1, 2.5, 10)$: $\gamma_1 = -0.25, \gamma_2 = -0.4$;
- a Logistic censored by another Logistic: $\gamma_1 = \gamma_2 = 0$.

These three specific choices of parameters lead to three different expected percentages of censoring in the right-tail: 11% for the Burr example, 38% for the ReverseBurr example and 50% for the last one.

The moment estimators adapted to censoring have been used as the initial estimators. They are defined as follows:

$$\begin{cases} \widehat{\gamma}_{Z,t}^{(c,M)} := \frac{\overline{H}_n(t)}{\overline{H}_n^1(t)} \widehat{\gamma}_{Z,t}^{(M)} := \frac{\overline{H}_n(t)}{\overline{H}_n^1(t)} \left(M_{Z,t}^{(1)} + 1 - \frac{1}{2} \left(1 - \frac{(M_{Z,t}^{(1)})^2}{M_{Z,t}^{(2)}} \right)^{-1} \right) \\ \\ \widehat{\sigma}_{Z,t}^{(c,M)} := \frac{\overline{H}_n(t)}{\overline{H}_n^1(t)} \frac{t\sqrt{3(M_{Z,t}^{(1)})^2 - M_{Z,t}^{(2)}}}{\sqrt{3[\rho_1(\widehat{\gamma}_{Z,t}^{(M)})]^2 - \rho_2(\widehat{\gamma}_{Z,t}^{(M)})}} \end{cases}$$

where

$$\rho_1(\gamma) := \begin{cases} 1 & \text{if } \gamma \geq 0 \\ \frac{1}{1-\gamma} & \text{if } \gamma < 0 \end{cases}$$

$$\rho_2(\gamma) := \begin{cases} 2 & \text{if } \gamma \geq 0 \\ \frac{2}{(1-\gamma)(1-2\gamma)} & \text{if } \gamma < 0 \end{cases}$$

$$M_{Z,t}^{(1)} := \frac{\sum_{i=1}^n \log \frac{Z_i}{t} 1\!\!1_{\{Z_i > t\}}}{n\overline{H}_n(t)}$$

$$M_{Z,t}^{(2)} := \frac{\sum_{i=1}^n (\log \frac{Z_i}{t})^2 1\!\!1_{\{Z_i > t\}}}{n\overline{H}_n(t)}$$

with $\overline{H}_n(t)$ and $\overline{H}_n^1(t)$ the empirical versions of $\overline{H}(t)$ and $\overline{H}^1(t)$ respectively.

In all the figures, the panel (a) represents the median whereas panel (b) shows the empirical mean squared errors (MSE) based on these 300 samples. Figure 1 displays the result for the estimators of γ_1 whereas Figure 2 shows the results for $\frac{\widehat{\sigma}_{Z,t}^{(c,.)}}{\sigma_{1,t}}$. The full line corresponds to the maximum likelihood estimator (obtained by direct optimization of the likelihood (2)) and the dashed line corresponds to the one-step estimator. The horizontal line is always the true value of the estimated parameter. All the plotted estimators are adapted to censoring. These graphs correspond successively to the three distributions: Burr (top), ReverseBurr (middle), Logistic (bottom).

We observe that the one-step estimator has a similar behaviour to the $ML-$estimator. In particular, it is very hard to differentiate the two graphs in the case of the Burr $(10, 4, 1)$ censored by a Burr $(10, 1, 0.5)$. In terms of the MSE, we can say that the minima of the two curves are very close, except in the case of the Logistic distribution for the scale parameter where the difference between the two graphs is significant (also for the median), the one-step estimator being considerably better. In this case the percentage of censoring is quite large (50%). Remark that in the case of no-censoring the two graphs are much closer as illustrated on Figure 3.

4 Proof of our Theorem

The proof of our theorem is in the spirit of Lehmann's theory (1991). This proof is very technical. Therefore, to facilitate its understanding, we first sketch the proof, the details being postponed to the appendix.

Step 1: The system of equations (3) can be rewritten as

$$\sqrt{N_t}\begin{pmatrix} \widehat{\gamma}_{Z,t}^{(c,OS)} - \widehat{\gamma}_{Z,t}^{(c,.)} \\ \\ \dfrac{\widehat{\sigma}_{Z,t}^{(c,OS)}}{\sigma_{1,t}} - \dfrac{\widehat{\sigma}_{Z,t}^{(c,.)}}{\sigma_{1,t}} \end{pmatrix}$$

$$= \frac{1}{\left(\frac{1}{N_t}L''_{11}\right)\left(\frac{\sigma_{1,t}^2}{N_t}L''_{22}\right) - \left(\frac{\sigma_{1,t}}{N_t}L''_{12}\right)^2} \begin{pmatrix} \dfrac{\sigma_{1,t}^2}{N_t}L''_{22} & -\dfrac{\sigma_{1,t}}{N_t}L''_{12} \\ \\ -\dfrac{\sigma_{1,t}}{N_t}L''_{12} & \dfrac{1}{N_t}L''_{11} \end{pmatrix} \begin{pmatrix} -\dfrac{1}{\sqrt{N_t}}L'_1 \\ \\ -\dfrac{\sigma_{1,t}}{\sqrt{N_t}}L'_2 \end{pmatrix} \tag{14}$$

where to alleviate the notations, we have not specified that all the derivatives are evaluated at $\left(\widehat{\gamma}_{Z,t}^{(c,.)}, \widehat{\sigma}_{Z,t}^{(c,.)}\right)$.

Step 2: According to (14) we have to compute the limit in probability of the second derivatives correctly normalized. Direct computations lead to

$$-\frac{1}{N_t}L''_{11}(\gamma_1,\sigma_{1,t}) \quad \xrightarrow{\mathbb{P}} \quad \begin{cases} 2\left(\dfrac{\gamma}{\gamma_1}\right)^3 \dfrac{1}{(1+\gamma)(1+2\gamma)} & \text{if } \gamma_1 \neq 0 \\ \\ 2\,C^3 & \text{if } \gamma_1 = 0 \end{cases}$$

$$-\frac{\sigma_{1,t}^2}{N_t}L''_{22}(\gamma_1,\sigma_{1,t}) \quad \xrightarrow{\mathbb{P}} \quad \begin{cases} \dfrac{\gamma}{\gamma_1}\dfrac{1}{1+2\gamma} & \text{if } \gamma_1 \neq 0 \\ \\ C & \text{if } \gamma_1 = 0 \end{cases}$$

$$-\frac{\sigma_{1,t}}{N_t}L''_{12}(\gamma_1,\sigma_{1,t}) \quad \xrightarrow{\mathbb{P}} \quad \begin{cases} \left(\dfrac{\gamma}{\gamma_1}\right)^2 \dfrac{1}{(1+\gamma)(1+2\gamma)} & \text{if } \gamma_1 \neq 0 \\ \\ C^2 & \text{if } \gamma_1 = 0. \end{cases}$$

However, the preceding convergences hold in case where the derivatives are evaluated at $(\gamma_1, \sigma_{1,t})$. In our case (see (14)), we need a similar result but when the derivatives are evaluated at

$(\widehat{\gamma}_{Z,t}^{(c,.)}, \widehat{\sigma}_{Z,t}^{(c,.)})$. In fact, it is fairly easy to deduce these results taken the fact that $\dfrac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} \leq 1$ into account. For instance, we have

$$
\begin{aligned}
&\frac{1}{N_t}\frac{1}{\widehat{\sigma}_{Z,t}^{(c,.)}}\sum_{j=1}^{N_t}\frac{E_j}{1+\frac{\widehat{\gamma}_{Z,t}^{(c,.)}}{\widehat{\sigma}_{Z,t}^{(c,.)}}E_j} - \frac{1}{N_t}\frac{1}{\sigma_{1,t}}\sum_{j=1}^{N_t}\frac{E_j}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} \\
&= \quad \frac{1}{N_t}\frac{1}{\widehat{\sigma}_{Z,t}^{(c,.)}}\sum_{j=1}^{N_t}\left(\frac{E_j}{1+\frac{\widehat{\gamma}_{Z,t}^{(c,.)}}{\widehat{\sigma}_{Z,t}^{(c,.)}}E_j} - \frac{E_j}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j}\right) + \frac{1}{N_t}\left(\frac{1}{\widehat{\sigma}_{Z,t}^{(c,.)}} - \frac{1}{\sigma_{1,t}}\right)\sum_{j=1}^{N_t}\frac{E_j}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} \\
&= \quad -\frac{1}{\widehat{\gamma}_{Z,t}^{(c,.)}}\left(\frac{\widehat{\tau}_{1,t}}{\tau_{1,t}} - 1\right)\frac{\tau_{1,t}}{\widehat{\widehat{\tau}}_{1,t}}\frac{\widehat{\tau}_{1,t}}{\widehat{\widehat{\tau}}_{1,t}}\frac{1}{N_t}\sum_{j=1}^{N_t}\left(\frac{\widehat{\widehat{\tau}}_{1,t}E_j}{1+\widehat{\widehat{\tau}}_{1,t}E_j}\right)^2 + \frac{1}{\gamma_1}\left(\frac{\sigma_{1,t}}{\widehat{\sigma}_{Z,t}^{(c,.)}} - 1\right)\frac{1}{N_t}\sum_{j=1}^{N_t}\frac{\tau_{1,t}E_j}{1+\tau_{1,t}E_j} \\
&\xrightarrow{\mathbb{P}} 0,
\end{aligned}
$$

where $\tau_{1,t} = \dfrac{\gamma_1}{\sigma_{1,t}}, \widehat{\tau}_{1,t} = \dfrac{\widehat{\gamma}_{Z,t}^{(c,.)}}{\widehat{\sigma}_{Z,t}^{(c,.)}}$ and $\widehat{\widehat{\tau}}_{1,t}$ is between $\tau_{1,t}$ and $\widehat{\tau}_{1,t}$.

Step 3: Now expanding the first derivatives of the log-likelihood appearing in the right-side of (3) about $(\gamma_1, \sigma_{1,t})$ (as in the proof of Theorem 4.2, Chapter 6 in Lehmann, 1991) and taking the conditions (10) and (11) into account, we have to establish the asymptotic normality of

$$
V := \begin{pmatrix} -\sqrt{N_t}\left[\dfrac{1}{N_t}L_1'(\gamma_1,\sigma_{1,t}) - B_{1,t}(\gamma_1,\sigma_{1,t})\right] \\ \\ -\sqrt{N_t}\left[\dfrac{\sigma_{1,t}}{N_t}L_2'(\gamma_1,\sigma_{1,t}) - B_{2,t}(\gamma_1,\sigma_{1,t})\right] \end{pmatrix}.
$$

According to Appendix 2, this follows from the asymptotic normality of

$$
W := \sqrt{\frac{n}{\overline{H}(t)}}\begin{pmatrix} \overline{H}_n(t) - \overline{H}(t) \\ \overline{H}_n^1(t) - \overline{H}^1(t) \\ \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\log\left(1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j - t)\right)\mathbb{1}_{\{Z_j>t\}} - \frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}(x)}{1+\frac{\gamma_1}{\sigma_{1,t}}(x-t)}dx \\ \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)\mathbb{1}_{\{Z_j>t\}} - \frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx \\ \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\delta_j\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)\mathbb{1}_{\{Z_j>t\}} - \frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}^1(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx \end{pmatrix}
$$

in the case $\gamma_1 \neq 0$ and

$$W := \sqrt{\frac{n}{\overline{H}(t)}} \begin{pmatrix} \overline{H}_n(t) - \overline{H}(t) \\ \overline{H}_n^1(t) - \overline{H}^1(t) \\ \frac{1}{\sigma_{1,t}} \left[\frac{1}{n} \sum_{j=1}^{n} (Z_j - t) \mathbb{1}_{\{Z_j > t\}} - \int_t^{\tau_H} \overline{H}(x) dx \right] \\ \frac{1}{\sigma_{1,t}^2} \left[\frac{1}{2n} \sum_{j=1}^{n} (Z_j - t)^2 \mathbb{1}_{\{Z_j > t\}} - \int_t^{\tau_H} (x - t) \overline{H}(x) dx \right] \\ \frac{1}{\sigma_{1,t}} \left[\frac{1}{n} \sum_{j=1}^{n} \delta_j (Z_j - t) \mathbb{1}_{\{Z_j > t\}} - \int_t^{\tau_H} \overline{H}^1(x) dx \right] \end{pmatrix}$$

otherwise.

Step 4: Using the multivariate central limit theorem, the pairs $(Z_i, \delta_i)_{i=1,\dots,n}$ being independent, the following asymptotic variance-covariance matrix for W follows in case $\gamma_1 \neq 0$:

$$\begin{pmatrix} 1 & \frac{\gamma}{\gamma_1} & \gamma & \frac{\gamma_1}{\gamma} a & a \\ \frac{\gamma}{\gamma_1} & \frac{\gamma}{\gamma_1} & \frac{\gamma^2}{\gamma_1} & a & a \\ \gamma & \frac{\gamma^2}{\gamma_1} & 2\gamma^2 & b & \frac{\gamma}{\gamma_1} b \\ \frac{\gamma_1}{\gamma} a & a & b & c & \frac{\gamma}{\gamma_1} c \\ a & a & \frac{\gamma}{\gamma_1} b & \frac{\gamma}{\gamma_1} c & \frac{\gamma}{\gamma_1} c \end{pmatrix},$$

where $a = \gamma \gamma_1^{-1} \left(1 - (1+\gamma)^{-1}\right)$, $b = \gamma \left(1 - (1+\gamma)^{-2}\right)$, and $c = 1 - 2(1+\gamma)^{-1} + (1+2\gamma)^{-1}$.

Similarly, in case $\gamma_1 = 0$, using the assumption (12), we obtain the following variance-covariance matrix for W:

$$\begin{pmatrix} 1 & C & C & C^2 & C^2 \\ C & C & C^2 & C^3 & C^2 \\ C & C^2 & 2\,C^2 & 3\,C^3 & 2\,C^3 \\ C^2 & C^3 & 3\,C^3 & 6\,C^4 & 3\,C^4 \\ C^2 & C^2 & 2\,C^3 & 3\,C^4 & 2\,C^3 \end{pmatrix}.$$

It is important to note that the computation of these matrices is quite complex. It requires to split each integral into several parts and to use the convergence results (8) and (9) combined with the dominated convergence theorem in order to conclude. To illustrate these techniques, we give two examples of such integrals in the Appendix 3, one in the case $\gamma = 0$, and another one in the case $\gamma < 0$, the case of a positive one being more easy to handle.

Step 5: The preceding step combined with Appendix 2 leads to the following variance-covariance matrix for V in case $\gamma_1 \neq 0$:

$$\begin{pmatrix} 2\left(\dfrac{\gamma}{\gamma_1}\right)^3 \dfrac{1}{(1+\gamma)(1+2\gamma)} & \left(\dfrac{\gamma}{\gamma_1}\right)^2 \dfrac{1}{(1+\gamma)(1+2\gamma)} \\ & \\ \left(\dfrac{\gamma}{\gamma_1}\right)^2 \dfrac{1}{(1+\gamma)(1+2\gamma)} & \dfrac{\gamma}{\gamma_1} \dfrac{1}{1+2\gamma} \end{pmatrix}$$

whereas, in the case $\gamma_1 = 0$, we obtain the following matrix:

$$\begin{pmatrix} 2\,C^3 & C^2 \\ & \\ C^2 & C \end{pmatrix}.$$

Concerning now the bias term, we have to use the conditions (10) and (11) combining with Step 2, to conclude that it is equal to

$$\begin{pmatrix} \left(\dfrac{\gamma_1}{\gamma}\right)^2 (1+\gamma) \left[\dfrac{\gamma_1}{\gamma}(1+\gamma)\alpha_1 - \alpha_2\right] \\ \dfrac{\gamma_1}{\gamma}(1+\gamma) \left[-\dfrac{\gamma_1}{\gamma}\alpha_1 + 2\alpha_2\right] \end{pmatrix}$$

in the case $\gamma_1 \neq 0$ and

$$\begin{pmatrix} C^{-3}\left[\alpha_1 - \alpha_2\, C\right] \\ \\ C^{-2}\left[-\alpha_1 + 2\alpha_2\, C\right] \end{pmatrix}$$

in the case $\gamma_1 = 0$.

Combining all these steps, our theorem follows.

5 Appendix

5.1 Appendix 1: The first and second derivatives

Direct computations lead to the following first derivatives of the log-likelihood:

$$\begin{cases} L_1'(\gamma_1,\sigma_{1,t}) := \dfrac{\partial \log L(\gamma_1,\sigma_{1,t})}{\partial \gamma_1} & = \dfrac{1}{\gamma_1^2}\displaystyle\sum_{j=1}^{N_t}\log\Big(1+\frac{\gamma_1}{\sigma_{1,t}}E_j\Big) - \frac{1}{\gamma_1}\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} \\ \\ L_2'(\gamma_1,\sigma_{1,t}) := \dfrac{\partial \log L(\gamma_1,\sigma_{1,t})}{\partial \sigma_{1,t}} & = -\dfrac{1}{\sigma_{1,t}}\displaystyle\sum_{j=1}^{N_t}\delta_j + \frac{1}{\sigma_{1,t}}\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} \end{cases}$$

and to the following second derivatives:

$$\begin{cases} L_{11}''(\gamma_1,\sigma_{1,t}) & := \dfrac{\partial^2 \log L(\gamma_1,\sigma_{1,t})}{\partial^2 \gamma_1} \\ & = -\dfrac{2}{\gamma_1^3}\displaystyle\sum_{j=1}^{N_t}\log\Big(1+\frac{\gamma_1}{\sigma_{1,t}}E_j\Big) + \frac{2}{\gamma_1^3}\sum_{j=1}^{N_t}\frac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} + \frac{1}{\gamma_1^2}\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{(\gamma_1 E_j/\sigma_{1,t})^2}{(1+\frac{\gamma_1}{\sigma_{1,t}}E_j)^2} \\ \\ L_{22}''(\gamma_1,\sigma_{1,t}) & := \dfrac{\partial^2 \log L(\gamma_1,\sigma_{1,t})}{\partial^2 \sigma_{1,t}} \\ & = \dfrac{1}{\sigma_{1,t}^2}\displaystyle\sum_{j=1}^{N_t}\delta_j - \frac{2}{\sigma_{1,t}^2}\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j} + \frac{1}{\sigma_{1,t}^2}\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{(\gamma_1 E_j/\sigma_{1,t})^2}{(1+\frac{\gamma_1}{\sigma_{1,t}}E_j)^2} \\ \\ L_{12}''(\gamma_1,\sigma_{1,t}) & := \dfrac{\partial^2 \log L(\gamma_1,\sigma_{1,t})}{\partial \sigma_{1,t}\partial \gamma_1} \\ & = -\dfrac{1}{\gamma_1}\dfrac{1}{\sigma_{1,t}}\displaystyle\sum_{j=1}^{N_t}\Big(\frac{1}{\gamma_1}+\delta_j\Big)\frac{(\gamma_1 E_j/\sigma_{1,t})^2}{(1+\frac{\gamma_1}{\sigma_{1,t}}E_j)^2} + \frac{1}{\gamma_1}\frac{1}{\sigma_{1,t}}\sum_{j=1}^{N_t}\delta_j\,\frac{\gamma_1 E_j/\sigma_{1,t}}{1+\frac{\gamma_1}{\sigma_{1,t}}E_j}. \end{cases}$$

In case where $\gamma_1 = 0$, the first derivatives can be read as follows

$$\begin{cases} L_1'(0,\sigma_{1,t}) & := \dfrac{1}{2}\displaystyle\sum_{j=1}^{N_t}\frac{E_j^2}{\sigma_{1,t}^2} - \sum_{j=1}^{N_t}\delta_j\frac{E_j}{\sigma_{1,t}} \\ \\ L_2'(0,\sigma_{1,t}) & := -\dfrac{1}{\sigma_{1,t}}\displaystyle\sum_{j=1}^{N_t}\delta_j + \frac{1}{\sigma_{1,t}^2}\sum_{j=1}^{N_t}E_j \end{cases}$$

whereas the second are given by:

$$\begin{cases} L''_{11}(0,\sigma_{1,t}) & := -\dfrac{2}{3}\displaystyle\sum_{j=1}^{N_t}\left(\dfrac{E_j}{\sigma_{1,t}}\right)^3+\sum_{j=1}^{N_t}\delta_j\left(\dfrac{E_j}{\sigma_{1,t}}\right)^2 \\ \\ L''_{22}(0,\sigma_{1,t}) & := \dfrac{1}{\sigma_{1,t}^2}\displaystyle\sum_{j=1}^{N_t}\delta_j-\dfrac{2}{\sigma_{1,t}^3}\sum_{j=1}^{N_t}E_j \\ \\ L''_{12}(0,\sigma_{1,t}) & := -\dfrac{1}{\sigma_{1,t}}\displaystyle\sum_{j=1}^{N_t}\left(\dfrac{E_j}{\sigma_{1,t}}\right)^2+\dfrac{1}{\sigma_{1,t}}\sum_{j=1}^{N_t}\delta_j\dfrac{E_j}{\sigma_{1,t}}. \end{cases}$$

5.2 Appendix 2: Justification of the matrix W

Using the preceding appendix, it is clear that in case $\gamma_1 \neq 0$, we have

$$\begin{aligned}
&-\sqrt{N_t}\left[\frac{1}{N_t}L'_1(\gamma_1,\sigma_{1,t})-B_{1,t}(\gamma_1,\sigma_{1,t})\right] \\
&= \quad -\sqrt{\frac{\overline{H}(t)}{\overline{H}_n(t)}}\sqrt{\frac{n}{\overline{H}(t)}}\Bigg\{\frac{1}{\gamma_1^2}\left(\frac{1}{n}\sum_{j=1}^{n}\log\left(1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}(z)}{1+\frac{\gamma_1}{\sigma_{1,t}}(z-t)}dz\right) \\
&\qquad -\frac{1}{\gamma_1^2}\left(\frac{1}{n}\sum_{j=1}^{n}\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\right) \\
&\qquad -\frac{1}{\gamma_1}\left(\frac{1}{n}\sum_{j=1}^{n}\delta_j\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}^1(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\right) \\
&\qquad +[\overline{H}_n(t)-\overline{H}(t)]\Bigg[-\frac{1}{\gamma_1\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\frac{\overline{H}(z)}{1+\frac{\gamma_1}{\sigma_{1,t}}(z-t)}dz+\frac{1}{\gamma_1\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\frac{\overline{H}(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz \\
&\qquad\qquad +\frac{1}{\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\frac{\overline{H}^1(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\Bigg]\Bigg\}
\end{aligned}$$

and

$$
\begin{aligned}
&-\sqrt{N_t}\left[\frac{\sigma_{1,t}}{N_t}L_2'(\gamma_1,\sigma_{1,t})-B_{2,t}(\gamma_1,\sigma_{1,t})\right]\\
&\quad= \quad -\sqrt{\frac{\overline{H}(t)}{\overline{H}_n(t)}}\sqrt{\frac{n}{\overline{H}(t)}}\Bigg\{-\left(\overline{H}_n^1(t)-\overline{H}^1(t)\right)\\
&\qquad+\frac{1}{\gamma_1}\left(\frac{1}{n}\sum_{j=1}^n\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\right)\\
&\qquad+\left(\frac{1}{n}\sum_{j=1}^n\delta_j\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}^1(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\right)\\
&\qquad+[\overline{H}_n(t)-\overline{H}(t)]\Bigg[\frac{\overline{H}^1(t)}{\overline{H}(t)}-\frac{1}{\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\frac{\overline{H}(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\\
&\qquad\qquad-\frac{\gamma_1}{\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\frac{\overline{H}^1(z)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(z-t))^2}dz\Bigg]\Bigg\}
\end{aligned}
$$

whereas, in case $\gamma_1=0$, we have

$$
\begin{aligned}
&-\sqrt{N_t}\left[\frac{1}{N_t}L_1'(\gamma_1,\sigma_{1,t})-B_{1,t}(\gamma_1,\sigma_{1,t})\right]\\
&\quad= \quad -\sqrt{\frac{\overline{H}(t)}{\overline{H}_n(t)}}\sqrt{\frac{n}{\overline{H}(t)}}\Bigg\{\frac{1}{\sigma_{1,t}^2}\left(\frac{1}{2n}\sum_{j=1}^n(Z_j-t)^2 1\!\!1_{\{Z_j>t\}}-\int_t^{\tau_H}(x-t)\overline{H}(x)dx\right)\\
&\qquad-\frac{1}{\sigma_{1,t}}\left(\frac{1}{n}\sum_{j=1}^n\delta_j(Z_j-t)1\!\!1_{\{Z_j>t\}}-\int_t^{\tau_H}\overline{H}^1(x)dx\right)\\
&\qquad-\left(\overline{H}_n(t)-\overline{H}(t)\right)\left(\frac{1}{\sigma_{1,t}^2}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}(x-t)\overline{H}(x)dx-\frac{1}{\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\overline{H}^1(x)dx\right)\Bigg\}
\end{aligned}
$$

and

$$
\begin{aligned}
&-\sqrt{N_t}\left[\frac{\sigma_{1,t}}{N_t}L_2'(\gamma_1,\sigma_{1,t})-B_{2,t}(\gamma_1,\sigma_{1,t})\right]\\
&\quad= \quad -\sqrt{\frac{\overline{H}(t)}{\overline{H}_n(t)}}\sqrt{\frac{n}{\overline{H}(t)}}\Bigg\{-\left(\overline{H}_n^1(t)-\overline{H}^1(t)\right)+\frac{1}{\sigma_{1,t}}\left(\frac{1}{n}\sum_{j=1}^n(Z_j-t)1\!\!1_{\{Z_j>t\}}-\int_t^{\tau_H}\overline{H}(x)dx\right)\\
&\qquad\qquad+\left(\overline{H}_n(t)-\overline{H}(t)\right)\left(\frac{\overline{H}^1(t)}{\overline{H}(t)}-\frac{1}{\sigma_{1,t}}\frac{1}{\overline{H}(t)}\int_t^{\tau_H}\overline{H}(x)dx\right)\Bigg\}.
\end{aligned}
$$

5.3 Appendix 3: Two examples of integral to be computed

We have to compute several integrals. To this aim, we have to divide them into several parts.

Below, we illustrate two such examples.

5.3.1. An example in case where γ_1 and γ_2 are negative and $\tau_F = \tau_G$

We illustrate our technique with the following integral

$$\mathrm{Var}\left(\sqrt{\frac{n}{\overline{H}(t)}}\left[\frac{1}{n}\sum_{j=1}^{n}\delta_j\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}^1(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx\right]\right)$$

on the first-hand because it is the most difficult one to handle and on the other-hand because the required assumption $\gamma > -1/2$ appears clearly. According to de Haan and Ferreira (2006, Theorem 1.2.5, p. 21), if the convergence

$$\lim_{t\uparrow\tau_H}\frac{\overline{F}(t+\sigma_{1,t}x)}{\overline{F}(t)} = (1+\gamma_1 x)^{-1/\gamma_1}$$

holds for some $\sigma_{1,t} > 0$, then it also holds for $\sigma_{1,t} = -\gamma_1(\tau_H - t)$. Therefore, in all the sequel we assume that $\sigma_{1,t}$ is such defined.

The quantity of interest can then be rewritten as:

$$\begin{aligned}
&\mathrm{Var}\left(\sqrt{\frac{n}{\overline{H}(t)}}\left[\frac{1}{n}\sum_{j=1}^{n}\delta_j\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(Z_j-t)}\right)1\!\!1_{\{Z_j>t\}}-\frac{\gamma_1}{\sigma_{1,t}}\int_t^{\tau_H}\frac{\overline{H}^1(x)}{(1+\frac{\gamma_1}{\sigma_{1,t}}(x-t))^2}dx\right]\right)\\
&= -\int_t^{\tau_H}\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(x-t)}\right)^2\frac{d\overline{H}^1(x)}{\overline{H}(t)} - \overline{H}(t)\left(\int_t^{\tau_H}\left(1-\frac{1}{1+\frac{\gamma_1}{\sigma_{1,t}}(x-t)}\right)\frac{d\overline{H}^1(x)}{\overline{H}(t)}\right)^2\\
&= -\int_t^{\tau_H}\left(1-\frac{\tau_H-t}{\tau_H-x}\right)^2\frac{d\overline{H}^1(x)}{\overline{H}(t)}+o(1)\\
&= \int_t^{\tau_H}\frac{\overline{H}^1(x)}{\overline{H}(t)}d\left(\left(1-\frac{\tau_H-t}{\tau_H-x}\right)^2\right)+o(1)\\
&= -\int_t^{\tau_H}\int_x^{\tau_H}\frac{\overline{G}(u)}{\overline{H}(t)}d\overline{F}(u)d\left(\left(1-\frac{\tau_H-t}{\tau_H-x}\right)^2\right)+o(1)\\
&= -\int_t^{\tau_H}\frac{\overline{G}(u)}{\overline{H}(t)}\int_t^{u}d\left(\left(1-\frac{\tau_H-t}{\tau_H-x}\right)^2\right)d\overline{F}(u)+o(1)\\
&= -\int_t^{\tau_H}\frac{\overline{G}(u)}{\overline{H}(t)}\left(1-\frac{\tau_H-t}{\tau_H-u}\right)^2 d\overline{F}(u)+o(1)\\
&= -\int_t^{\tau_H}\frac{\overline{G}(u)}{\overline{H}(t)}\left(1-\frac{\tau_H-t}{\tau_H-u}\right)^2\left[\frac{1}{\gamma_1}(\tau_H-u)^{-\frac{1}{\gamma_1}-1}\ell_F\left(\frac{1}{\tau_H-u}\right)\right.\\
&\qquad\qquad \left.+(\tau_H-u)^{-\frac{1}{\gamma_1}}\ell_F\left(\frac{1}{\tau_H-u}\right)\frac{\ell_F'(\frac{1}{\tau_H-u})}{\ell_F(\frac{1}{\tau_H-u})}\frac{1}{(\tau_H-u)^2}\right]du+o(1)
\end{aligned}$$

$$
\begin{aligned}
&= -\int_t^{\tau_H} \frac{\overline{H}(u)}{\overline{H}(t)} \left(1 - \frac{\tau_H - t}{\tau_H - u}\right)^2 \frac{1}{\gamma_1} (\tau_H - u)^{-1} du \\
&\quad - \int_t^{\tau_H} \frac{\overline{H}(u)}{\overline{H}(t)} \left(1 - \frac{\tau_H - t}{\tau_H - u}\right)^2 (\tau_H - u)^{-1} \varepsilon\left(\frac{1}{\tau_H - u}\right) du + o(1)
\end{aligned}
$$

with $\varepsilon\left(\frac{1}{\tau_H-u}\right) := \frac{\ell_F'(\frac{1}{\tau_H-u})}{\ell_F(\frac{1}{\tau_H-u})} \frac{1}{\tau_H-u}$ using the fact that ℓ_F is normalized. Recall that, when $u \to \tau_H$, $\varepsilon\left(\frac{1}{\tau_H-u}\right)$ converges to 0, (see Bingham *et al.*, 1987, p. 15).

By changing variables, we obtain that

$$
\begin{aligned}
&\operatorname{Var}\left(\sqrt{\frac{n}{\overline{H}(t)}} \left[\frac{1}{n}\sum_{j=1}^n \delta_j \left(1 - \frac{1}{1 + \frac{\gamma_1}{\sigma_{1,t}}(Z_j - t)}\right) 1\!\!1_{\{Z_j > t\}} - \frac{\gamma_1}{\sigma_{1,t}} \int_t^{\tau_H} \frac{\overline{H}^1(x)}{(1 + \frac{\gamma_1}{\sigma_{1,t}}(x-t))^2} dx\right]\right) \\
&= -\frac{1}{\gamma_1} \int_1^\infty \frac{\overline{H}(\tau_H - \frac{\tau_H - t}{x})}{\overline{H}(t)} (1-x)^2 x^{-1} dx \\
&\quad - \int_1^\infty \frac{\overline{H}(\tau_H - \frac{\tau_H - t}{x})}{\overline{H}(t)} (1-x)^2 x^{-1} \varepsilon\left(\frac{x}{\tau_H - t}\right) dx + o(1) \\
&=: T_{1,t} + T_{2,t}.
\end{aligned}
$$

Recall now that $\frac{\overline{H}(\tau_H - \frac{\tau_H - t}{x})}{\overline{H}(t)} \to x^{\frac{1}{\gamma}}$ as t tends to τ_H. Moreover, using Potter's bounds (see e.g. Bingham *et al.*, 1987, p. 25), for any chosen constants $A > 1$ and $\eta > 0$ we have

$$
\frac{\overline{H}(\tau_H - \frac{\tau_H - t}{x})}{\overline{H}(t)} = x^{\frac{1}{\gamma}} \frac{\ell_H(\frac{x}{\tau_H - t})}{\ell_H(\frac{1}{\tau_H - t})} \le A x^{\frac{1}{\gamma} + \eta}, \quad x > 1
$$

for t sufficiently large.

Also if $\gamma > -1/2$, choosing η sufficiently small, the integral

$$
-\frac{1}{\gamma_1} \int_1^\infty x^{\frac{1}{\gamma} + \eta} (1-x)^2 x^{-1} dx
$$

is clearly finite. This allows us to use the dominated convergence theorem and to deduce that

$$
\lim_{t \to \tau_H} T_{1,t} = -\frac{1}{\gamma_1} \int_1^\infty x^{\frac{1}{\gamma}} (1-x)^2 x^{-1} dx = \frac{\gamma}{\gamma_1} \left(1 - \frac{2}{1+\gamma} + \frac{1}{1+2\gamma}\right).
$$

Concerning now $T_{2,t}$, using the fact that $\varepsilon\left(\frac{x}{\tau_H - t}\right)$ tends to 0 as $t \to \tau_H$ and the preceding argument, we can also use the dominated convergence theorem in order to deduce that $T_{2,t}$ tends to 0.

5.3.2 An example in case where $\gamma_1 = \gamma_2 = 0$ and $\tau_F = \tau_G$

Similarly as in the preceding case, we select the « most difficult » integral that have to be computed. It is the following:

$$\begin{aligned}
&\mathrm{Var}\left(\sqrt{\frac{n}{\overline{H}(t)}}\frac{1}{\sigma_{1,t}}\left[\frac{1}{n}\sum_{j=1}^{n}\delta_j(Z_j-t)\mathbb{1}_{Z_j>t}-\int_t^{\tau_H}\overline{H}^1(x)dx\right]\right)\\
&\qquad = \frac{1}{\sigma_{1,t}^2}\frac{-1}{\overline{H}(t)}\int_t^{\tau_H}(x-t)^2d\overline{H}^1(x)-\overline{H}(t)\left(\frac{1}{\sigma_{1,t}}\frac{-1}{\overline{H}(t)}\int_t^{\tau_H}(x-t)d\overline{H}^1(x)\right)^2\\
&\qquad = \frac{2\sigma_{H,t}^2}{\sigma_{1,t}^2}\int_0^{\frac{\tau_H-t}{\sigma_{H,t}}} z\,\frac{\overline{H}^1(t+\sigma_{H,t}z)}{\overline{H}(t)}dz+o(1).
\end{aligned}$$

This integral can be rewritten as:

$$\begin{aligned}
\int_0^{\frac{\tau_H-t}{\sigma_H}} z\,\frac{\overline{H}^1(t+\sigma_{H,t}z)}{\overline{H}(t)}dz &= -\int_0^{\frac{\tau_H-t}{\sigma_{H,t}}}\frac{z}{\overline{H}(t)}\int_{t+z\sigma_{H,t}}^{\tau_H}\overline{G}(u)d\overline{F}(u)dz\\
&= -\int_t^{\tau_H}\frac{\overline{G}(u)}{\overline{H}(t)}\int_0^{\frac{u-t}{\sigma_{H,t}}} z\,dz\,d\overline{F}(u)\\
&= -\frac{1}{2}\left(\frac{\sigma_{2,t}}{\sigma_{H,t}}\right)^2\int_0^{\frac{\tau_H-t}{\sigma_{2,t}}}\frac{\overline{G}(t+\sigma_{2,t}z)}{\overline{G}(t)}z^2 d\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}\\
&= -\frac{1}{2}\left(\frac{\sigma_{2,t}}{\sigma_{H,t}}\right)^2\left[\int_0^{\frac{\tau_H-t}{\sigma_{2,t}}}\left\{\frac{\overline{G}(t+\sigma_{2,t}z)}{\overline{G}(t)}-e^{-z}\right\}z^2 d\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}\right.\\
&\qquad -\int_0^{\frac{\tau_H-t}{\sigma_{2,t}}}\left\{\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}-e^{-\frac{\sigma_{2,t}}{\sigma_{1,t}}z}\right\}d\left(z^2\,e^{-z}\right)\\
&\qquad \left.-\int_0^{\frac{\tau_H-t}{\sigma_{2,t}}} e^{-\frac{\sigma_{2,t}}{\sigma_{1,t}}z}d\left(z^2\,e^{-z}\right)\right]\\
&=:\; T_{3,t}+T_{4,t}+T_{5,t}.
\end{aligned}$$

We treat the three terms separately. Direct computations lead to

$$\frac{2\sigma_{H,t}^2}{\sigma_{1,t}^2}T_{5,t}=\frac{2\sigma_{H,t}^3}{\sigma_{1,t}^3}+o(1),$$

which is the expected result for the variance that we would like to compute. Therefore, we have now to prove the negligibility of the two other terms.

First concerning $T_{4,t}$, we have

$$T_{4,t} \;=\; \frac{1}{2}\left(\frac{\sigma_{2,t}}{\sigma_{H,t}}\right)^2\int_0^{\frac{\tau_H-t}{\sigma_{2,t}}}\left\{\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}-e^{-\frac{\sigma_{2,t}}{\sigma_{1,t}}z}\right\}\left(2ze^{-z}-z^2e^{-z}\right)dz$$

$$= \frac{1}{2}\left(\frac{\sigma_{2,t}}{\sigma_{H,t}}\right)^2 \frac{\sigma_{1,t}}{\sigma_{2,t}} \int_0^{\frac{\tau_H - t}{\sigma_{1,t}}} \left\{ \frac{\overline{F}(t+\sigma_{1,t}z)}{\overline{F}(t)} - e^{-z} \right\} \left(2\frac{\sigma_{1,t}}{\sigma_{2,t}}\, z e^{-\frac{\sigma_{1,t}}{\sigma_{2,t}}z} - \left(\frac{\sigma_{1,t}}{\sigma_{2,t}}\right)^2 z^2 e^{-\frac{\sigma_{1,t}}{\sigma_{2,t}}z} \right) dz.$$

Consequently

$$|T_{4,t}| \leq \frac{1}{2}\frac{\sigma_{2,t}\sigma_{1,t}}{\sigma_{H,t}^2} \sup_{z\in\left(0,\frac{\tau_H-t}{\sigma_{1,t}}\right)} \left| \frac{\overline{F}(t+\sigma_{1,t}z)}{\overline{F}(t)} - e^{-z} \right| \times 4\,\frac{\sigma_{2,t}}{\sigma_{1,t}}\,(1+o(1))$$

which leads to

$$\frac{2\sigma_{H,t}^2}{\sigma_{1,t}^2}|T_{4,t}| \leq 4\left(\frac{\sigma_{2,t}}{\sigma_{1,t}}\right)^2 (1+o(1)) \sup_{z\in\left(0,\frac{\tau_H-t}{\sigma_{1,t}}\right)} \left| \frac{\overline{F}(t+\sigma_{1,t}z)}{\overline{F}(t)} - e^{-z} \right| \longrightarrow 0,$$

by condition (12).

Second, concerning $T_{3,t}$, we have

$$\begin{aligned}
\left| 2\,\frac{\sigma_{H,t}^2}{\sigma_{1,t}^2} T_{3,t} \right| &= \left(\frac{\sigma_{2,t}}{\sigma_{1,t}}\right)^2 \left| \int_0^{\frac{\tau_H-t}{\sigma_{2,t}}} \left\{ \frac{\overline{G}(t+\sigma_{2,t}z)}{\overline{G}(t)} - e^{-z} \right\} z^2\, d\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)} \right| \\
&\leq -\left(\frac{\sigma_{2,t}}{\sigma_{1,t}}\right)^2 \sup_{z\in\left(0,\frac{\tau_H-t}{\sigma_{2,t}}\right)} \left| \frac{\overline{G}(t+\sigma_{2,t}z)}{\overline{G}(t)} - e^{-z} \right| \int_0^{\frac{\tau_H-t}{\sigma_{2,t}}} z^2\, d\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}.
\end{aligned}$$

Since the supremum tends to 0, we only have to prove that

$$Q_t := -\left(\frac{\sigma_{2,t}}{\sigma_{1,t}}\right)^2 \int_0^{\frac{\tau_H-t}{\sigma_{2,t}}} z^2\, d\frac{\overline{F}(t+\sigma_{2,t}z)}{\overline{F}(t)}$$

is bounded. To this aim, we are going to use (8). We first remark that

$$\begin{aligned}
Q_t &= 2\int_0^{\frac{\tau_H-t}{\sigma_{1,t}}} \frac{\overline{F}(t+\sigma_{1,t}z)}{\overline{F}(t)} z dz \\
&= 2\,\chi_1(t) \int_0^{\frac{\tau_H-t}{\sigma_{1,t}}} z\psi_1(z)\frac{1}{\psi_1(z)\chi_1(t)} \left[\frac{\overline{F}(t+\sigma_{1,t}z)}{\overline{F}(t)} - e^{-z} \right] dz + 2\int_0^{\frac{\tau_H-t}{\sigma_{1,t}}} z\, e^{-z}\, dz \\
&= 2\,\chi_1(t) \int_0^{\infty} f_t(z) dz + 2 + o(1).
\end{aligned}$$

Since $f_t(z) \to z\psi_1(z)$ and $|f_t(z)| \leq 2\,z|\psi_1(z)|$ which is integrable, we deduce by the dominated convergence theorem that $Q_t \to 2$.

6 References

Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. (1993). *Statistical models based on counting processes.* New York: Springer.

Balkema, A., de Haan, L. (1974). Residual life at great age. Ann. Probab. 2:792-804.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). *Statistics of extremes, theory and applications*. New York: Wiley.

Beirlant, J., Guillou, A., Dierckx, G., Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring. Extremes 10:151-174.

Beirlant, J., Vynckier, P., Teugels, J.L. (1996). Tail index estimation, Pareto quantile plots, and regression diagnostics. J. Amer. Statist. Assoc. 91:1659-1667.

Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular Variation*. Cambridge, U.K.: Cambridge University Press.

Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. Ann. Statist. 17:1833-1855.

Drees, H. (1998). On smooth statistical tail functionals. Scand. J. Statist. 25:187-210.

Drees, H., Ferreira, A., de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. Ann. Applied Probab. 14:1179-1201.

Einmahl, J.H.J., Fils-Villetard, A., Guillou, A. (2008). Statistics of extremes under random censoring. Bernoulli 14(1):207-227.

de Haan, L., Ferreira, A. (2006). *Extreme value theory: An introduction*. New York: Springer.

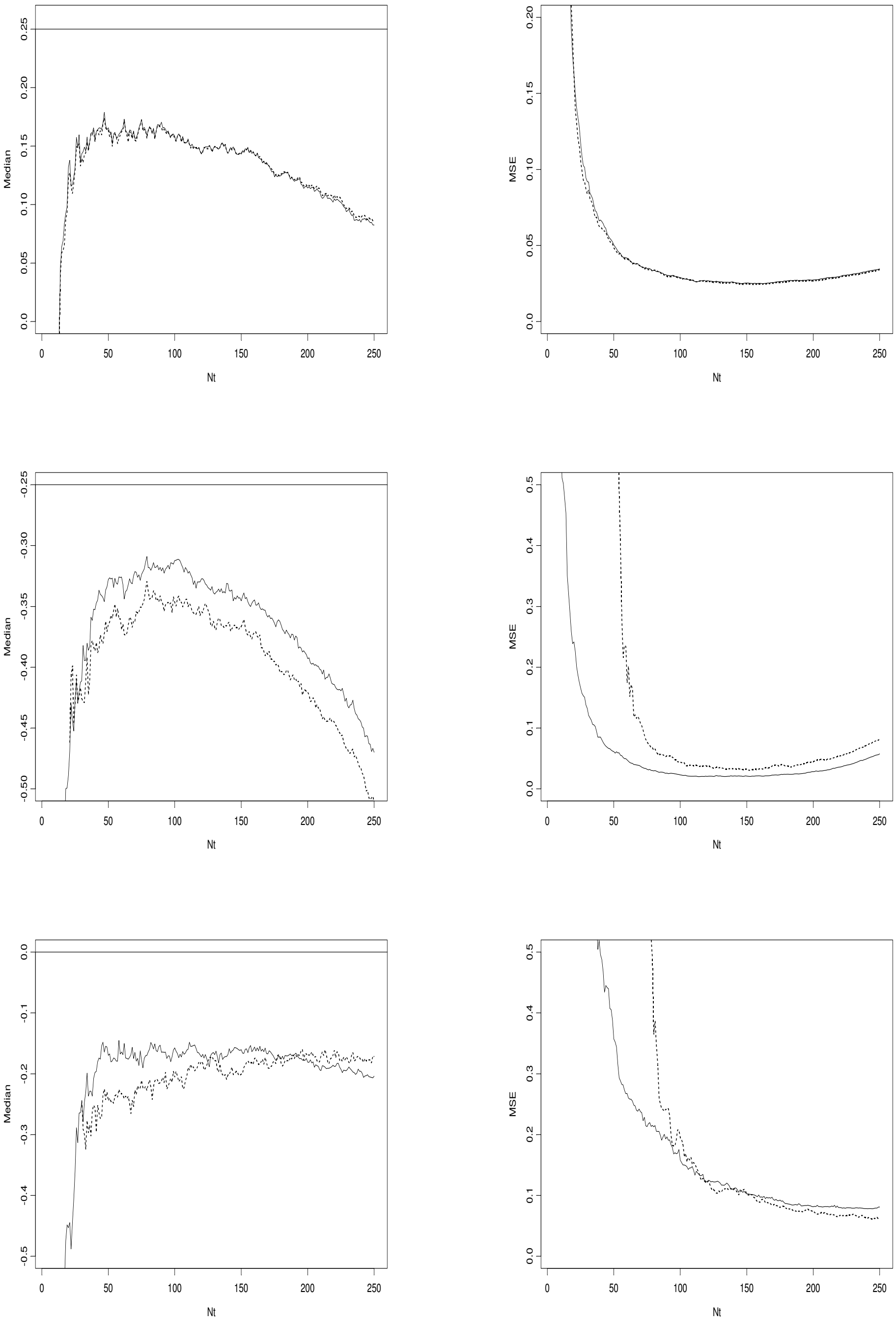
de Haan, L., Stadtmüller, U. (1996). Generalized regular variation of second order. J. Austral. Math. Soc. Ser. A 61:381-395.

Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. Ann. Statist. 3:1163-1174.

Lehmann, E.L. (1991). *Theory of point estimation*. Brooks-Cole publishing company.

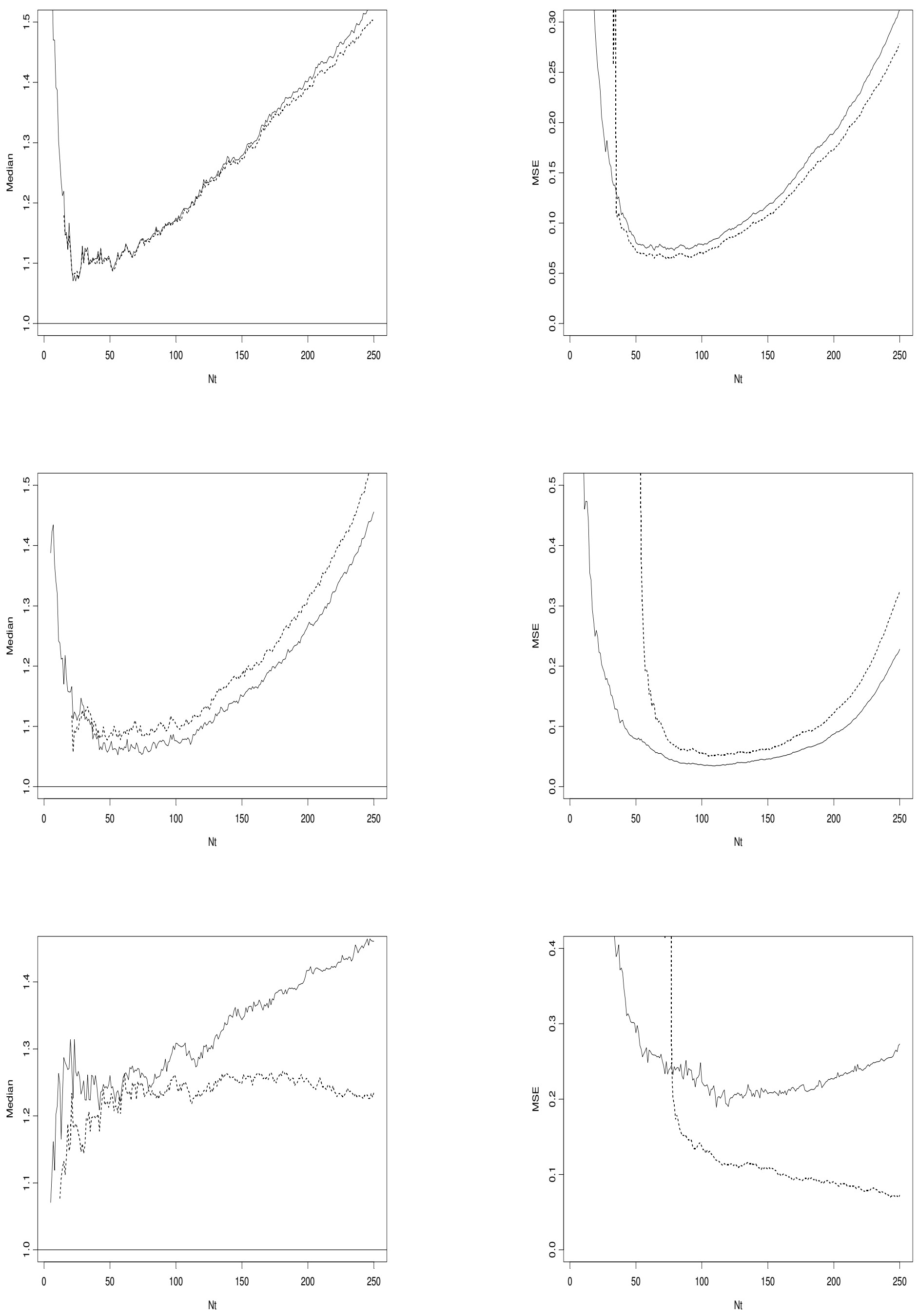
Pickands III, J. (1975). Statistical inference using extreme order statistics. Ann. Statist. 3:119-131.

Smith, R.L. (1987). Estimating tails of probability distributions. Ann. Statist. 15:1174-1207.



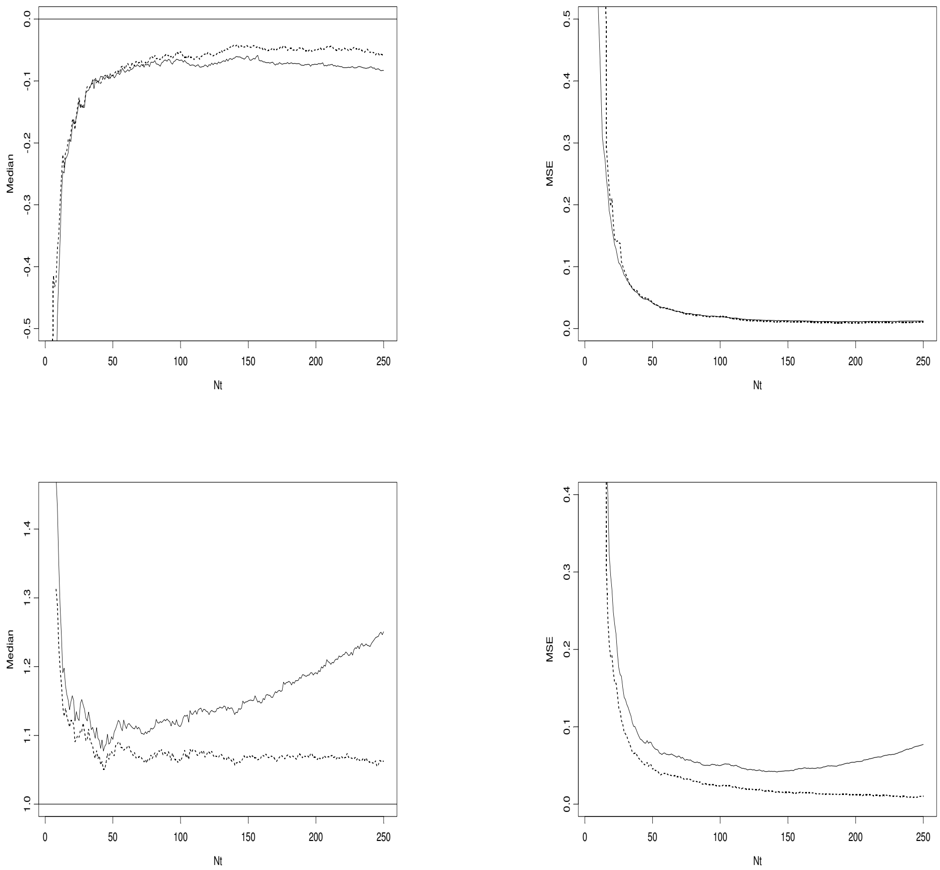
(a) (b)

Figure 1: (a) Median for the extreme value index estimators and (b) the associated empirical mean squared errors for a Burr $(10, 4, 1)$ censored by a Burr $(10, 1, 0.5)$ (on the top), for a ReverseBurr $(1, 8, 0.5, 10)$ censored by a ReverseBurr $(10, 1, 2.5, 10)$ (in the middle) and for a Logistic censored by another Logistic (on the bottom). The $ML-$estimator is in full line whereas the one-step estimator is in dashed line. The horizontal line corresponds to γ_1.



(a) (b)

Figure 2: Median for $\frac{\widehat{\sigma}_{Z,t}^{(c,.)}}{\sigma_{1,t}}$ and (b) the associated empirical mean squared errors for a Burr $(10,4,1)$ censored by a Burr $(10,1,0.5)$ (on the top), for a ReverseBurr $(1,8,0.5,10)$ censored by a ReverseBurr $(10,1,2.5,10)$ (in the middle) and for a Logistic censored by another Logistic (on the bottom). The $ML-$estimator is in full line whereas the one-step estimator is in dashed line.



(a) (b)

Figure 3: The case of no-censoring: at the top: (a) median for the estimators of γ_1 and (b) the associated empirical mean squared errors and at the bottom: (a) median for the estimators of the scale parameter correctly normalized and (b) the associated empirical mean squared errors for a Logistic distribution. The $ML-$estimator is in full line whereas the one-step estimator is in dashed line. The horizontal line is the true value of the estimated parameter.