

Supplementary Material for “Tail risk inference via expectiles in heavy-tailed time series”

A. C. Davison

Institute of Mathematics, Ecole Polytechnique Fédérale de Lausanne

S. A. Padoan

Department of Decision Sciences, Bocconi University of Milan

and

G. Stupfler

Université de Rennes, Ensai, CNRS, CREST

May 6, 2022

Abstract

This document of supplementary material contains a detailed discussion of the technical conditions, further numerical details and results, and all proofs. References below of the form (\cdot^*) refer to an element (condition, theorem, equation, etc.) in the main paper.

A Discussion of technical conditions

On Condition B* Condition B(ii)* provides a way to quantify the extremal dependence between the observations at different time points, through the so-called tail copula function (Schmidt & Stadtmüller 2006). Condition B(iii)* is a slightly more precise version of condition (C3) in Drees (2003), who used the latter to obtain bounds on cluster sizes of exceedances. In particular it holds if the ρ -mixing coefficients of (Y_t) (in the sense of Bradley 2005) are summable, and then the coefficients $\rho(t)$ can indeed be taken as the ρ -mixing coefficients of the time series (Y_t) , because

$$\mathbb{P}(A \cap B) \leq \text{corr}(\mathbb{1}\{A\}, \mathbb{1}\{B\})\sqrt{\mathbb{P}(A)\mathbb{P}(B)} + \mathbb{P}(A)\mathbb{P}(B)$$

for any pair of events (A, B) . Although at first sight this assumption seems to restrict our β -mixing setup, Drees (2003) shows that it holds in general ARMA, ARCH and GARCH models that are of interest in financial applications.

On the assumptions of Theorem 3.1*. The condition $\sum_{l \geq 1} [\beta(l)]^{\delta/(2+\delta)} < \infty$ implies, and can be replaced by, $\sum_{l \geq 1} [\alpha(l)]^{\delta/(2+\delta)} < \infty$, where the α -mixing coefficients of the sequence (Y_t) are defined by

$$\forall l \geq 1, \alpha(l) := \sup_{m \geq 1} \sup_{A \in \mathcal{F}_{1,m}} \sup_{B \in \mathcal{F}_{m+l,\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

The time series (Y_t) is said to be α -mixing (or strongly mixing) if $\alpha(l) \rightarrow 0$ as $l \rightarrow \infty$. As Section 1.1 in Doukhan (1994) points out, α -mixing is weaker than β -mixing because $\alpha(l) \leq \beta(l)$ for any l . In fact, the condition $\sum_{l \geq 1} [\alpha(l)]^{\delta/(2+\delta)} < \infty$ is a standard assumption for central limit theorems in strong mixing frameworks; see Ibragimov (1962) and Rio (2017). It is also the right framework for the general theory of tail array sums provided by Rootzén et al. (1998), which is key to our analysis of the LAWS estimator in Theorem 3.1*. [For clarity, throughout the paper we adopted a formulation in terms of β -mixing coefficients for consistency with our β -mixing framework.]

The condition $r_n(1 - \tau_n) \rightarrow 0$ as $n \rightarrow \infty$ plays a key role in the calculation of the asymptotic variance of the estimator $\tilde{\xi}_{\tau_n}$, and also appears in Drees (2002, 2003), Rootzén (2009) and Drees & Rootzén (2010). Finally, condition $r_n(r_n/\sqrt{n(1 - \tau_n)})^\delta \rightarrow 0$ as $n \rightarrow \infty$ is used to check the Lindeberg condition of Theorem 4.1 in Rootzén et al. (1998), which is a central limit theorem tailored to our expectile estimation problem under the strong mixing setting. This last assumption on r_n could seem strong, due to its impact on the assumed mixing rate via condition $n\beta(l_n)/r_n \rightarrow 0$ as $n \rightarrow \infty$ in Condition B(i)*. For δ small it is indeed stronger than analogous conditions used in extreme quantile estimation: in that context,

Rootzén (2009) uses the assumption $r_n/[n(1 - \tau_n)]^{1/2-\delta} \rightarrow 0$, while Drees (2003) assumes the even weaker condition $r_n \log^2(n(1 - \tau_n))/\sqrt{n(1 - \tau_n)} \rightarrow 0$ as $n \rightarrow \infty$. As we point out in the main article, the assumption $r_n(r_n/\sqrt{n(1 - \tau_n)})^\delta \rightarrow 0$ as $n \rightarrow \infty$ and all the other assumptions of Theorem 3.1* are very mild when $\beta(l)$ converges to 0 geometrically fast as $l \rightarrow \infty$. In that case, our conditions are satisfied with, for instance, $l_n = \lfloor C \log n \rfloor$, $r_n = \lfloor \log^2(n) \rfloor$ and $\tau_n = 1 - n^{-\tau}$, for any $\tau \in (0, 1)$ and sufficiently large C .

B Additional finite-sample results

B.1 Confidence interval for intermediate expectiles

Here we use the asymptotic distribution of the direct LAWS estimator given in Theorem 3.1* to derive a two-sided confidence interval for the expectile at the intermediate level, with $(1 - \alpha)100\%$ nominal coverage probability. We first propose an estimator for the asymptotic variance

$$V(\gamma, R) = \frac{2\gamma^3}{1 - 2\gamma}(1 + \sigma^2(\gamma, R)) \quad (\text{B.1})$$

of Theorem 3.1*, where the expression of $\sigma^2(\gamma, R)$ is given in Theorem 3.1* and $R = (R_t, t = 1, 2, \dots)$.

Lemma C.2 implies that

$$\lim_{n \rightarrow \infty} \text{Var} \left(\frac{\xi_{\tau_n}^{-1}}{\sqrt{r_n(1 - \tau_n)}} \sum_{t=1}^{r_n} [(Y_t - \xi_{\tau_n}) \mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})] \right) = \frac{V(\gamma, R)}{\gamma^2}. \quad (\text{B.2})$$

We then estimate the limiting quantity in (B.2) by computing the empirical counterpart of the left-hand side in (B.2). Specifically, we split the data into big blocks of size r_n , separated by small blocks of size l_n , leading us to define the random variables

$$\widehat{S}_j = \frac{\widetilde{\xi}_{\tau_n}^{-1}}{\sqrt{r_n(1 - \tau_n)}} \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \left((Y_t - \widetilde{\xi}_{\tau_n}) \mathbb{1}\{Y_t > \widetilde{\xi}_{\tau_n}\} - \frac{1}{n} \sum_{i=1}^n [(Y_i - \widetilde{\xi}_{\tau_n}) \mathbb{1}\{Y_i > \widetilde{\xi}_{\tau_n}\}] \right)$$

for $j = 0, 1, \dots, m_n - 1$, where $m_n = \lfloor n/\ell_n \rfloor$ and $\ell_n = r_n + l_n$. We then approximate the variance in (B.2) by computing the sample variance Ω_n of $(\widehat{S}_0, \dots, \widehat{S}_{m_n-1})$. Then an estimator of the asymptotic variance $V(\gamma, R)$ is given by $\widehat{\gamma}_n^2 \Omega_n$, where $\widehat{\gamma}_n$ is an estimator of the tail index γ .

Numerical experiments suggest that the normal distribution in Theorem 3.1* may not be a very good approximation to the sample distribution of the rescaled expectile estimator, especially when the marginal distribution tail is quite heavy and there is strong serial dependence, so we consider two improvements.

First, to compensate for the poor precision offered by the Gaussian approximation in Theorem 3.1*, the estimator of the asymptotic variance is inflated according to the serial dependence and the heaviness of the marginal distribution tail. We then propose an alternative, possibly inflated, estimator of the asymptotic variance,

$$\widehat{V}_n(\gamma, R) = \widehat{\gamma}_n^\mu \Omega_n, \quad 1/2 \leq \mu \leq 2. \quad (\text{B.3})$$

When $\mu = 2$ the original estimator $\widehat{\gamma}_n^2 \Omega_n$ is recovered. This should be the natural option when the serial dependence is weak/moderate and the marginal distribution tail is not too heavy. Otherwise, with strong serial dependence and a heavier marginal distribution tail, a value $\mu < 2$ can be selected and an amplified estimate of the asymptotic variance obtained. Choosing μ is delicate: in Section B.2 below we discuss a sensible criterion for selecting μ . Secondly, we note that the asymptotic result in Theorem 3.1* can be equivalently rephrased as

$$\sqrt{n(1 - \tau_n)} \log \frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} \xrightarrow{d} \mathcal{N}(0, V(\gamma, R)), \quad n \rightarrow \infty.$$

We construct our confidence intervals based on this result, as simulation results suggest that the Gaussian approximation is more accurate on the log scale. We therefore propose the confidence interval

$$\left[\widetilde{\xi}_{\tau_n} \exp \left(z_{\alpha/2} \sqrt{\frac{\widehat{V}_n(\gamma, R)}{n(1 - \tau_n)}} \right), \widetilde{\xi}_{\tau_n} \exp \left(z_{1-\alpha/2} \sqrt{\frac{\widehat{V}_n(\gamma, R)}{n(1 - \tau_n)}} \right) \right], \quad (\text{B.4})$$

where $\widetilde{\xi}_{\tau_n}$ is the direct LAWS estimator, $\widehat{V}_n(\gamma, R)$ is the sample estimator in (B.3) and $z_{\alpha/2}$ and $z_{1-\alpha/2}$ are the $(\alpha/2)$ and $(1 - \alpha/2)$ standard normal quantiles, with $\alpha \in (0, 1)$. We call the estimator in (B.4) the LAWS-D-H estimator, where D stands for dependent, when $\widehat{\gamma}_n = \widehat{\gamma}_n^H$, and the LAWS-D-E estimator when $\widehat{\gamma}_n = \widehat{\gamma}_n^E$ in Equation (B.3).

Finally, we recall that for i.i.d. data (Daouia et al. 2018) the asymptotic variance of the direct LAWS estimator is $V(\gamma, R) = 2\gamma^3/(1 - 2\gamma)$. In this case, an asymptotic $(1 - \alpha)$ -confidence interval is

$$\left[\widetilde{\xi}_{\tau_n} \exp \left(z_{\alpha/2} \sqrt{\frac{2\widehat{\gamma}_n^3/(1 - 2\widehat{\gamma}_n)}{n(1 - \tau_n)}} \right), \widetilde{\xi}_{\tau_n} \exp \left(z_{1-\alpha/2} \sqrt{\frac{2\widehat{\gamma}_n^3/(1 - 2\widehat{\gamma}_n)}{n(1 - \tau_n)}} \right) \right].$$

We call this the LAWS-IID-H(E) confidence interval, depending on whether $\widehat{\gamma}_n^H$ or $\widehat{\gamma}_n^E$ is used.

B.2 Finite-sample performance of point and interval expectile estimation at intermediate levels

Here we analyse the finite-sample performance of the direct LAWS estimator proposed in Section 3.1* and the confidence interval discussed in Section B.1 by considering a selection of models that include those of Section 7.1*. We consider AR, ARMA, ARCH and GARCH models. Specifically, we first consider the AR(1) family $Y_{t+1} = \phi Y_t + \varepsilon_{t+1}$, where the innovations ε_t are i.i.d. and have a Student- t distribution with $\nu > 0$ degrees of freedom. We choose

$$(i) \quad \phi = 0.8, \nu = 3;$$

$$(ii) \quad \phi = 0.8, \nu = 4.$$

These models exhibit fairly strong linear dependence. We then consider the ARMA(1,1) family $Y_{t+1} = \phi Y_t + \varepsilon_{t+1} + \theta \varepsilon_t$, where the innovations ε_t are i.i.d. and have a symmetric Pareto distribution with shape parameter $\zeta > 0$. We consider

$$(iii) \quad \phi = 0.95, \theta = 0.9, \zeta = 3;$$

$$(iv) \quad \phi = 0.95, \theta = -0.6, \zeta = 3;$$

$$(v) \quad \phi = 0.95, \theta = -0.9, \zeta = 3;$$

$$(vi) \quad \phi = 0.3, \theta = 0.9, \zeta = 3.$$

Strong linear dependence is present in the first two models and weaker linear dependence appears in the second two models. We recall that for standard linear time series with a heavy-tailed innovation satisfying the so-called tail balance condition, which is the case for models (i)-(vi), the tail index of the time series is equal to the tail index of its innovations. Thus the tail index in models (i)-(vi) is always 1/3 except in model (ii), where it is 1/4. Finally, we consider the nonlinear GARCH(1,1) family $Y_{t+1} = \sigma_{t+1} \varepsilon_{t+1}$, where $\sigma_{t+1}^2 = \alpha_0 + \alpha_1 Y_t^2 + \beta \sigma_t^2$, and (ε_t) is a sequence of i.i.d. Gaussian innovations. We work on two ARCH models and two proper GARCH models:

$$(vii) \quad \alpha_0 = 0.0001, \alpha_1 = 0.9, \beta = 0;$$

$$(viii) \quad \alpha_0 = 0.4, \alpha_1 = 0.6, \beta = 0;$$

$$(ix) \quad \alpha_0 = 0.0001, \alpha_1 = 0.4, \beta = 0.5;$$

(x) $\alpha_0 = 0.1, \alpha_1 = 0.4, \beta = 0.4$.

The marginal distribution F_Y is heavy-tailed in such ARCH/GARCH models, under suitable conditions, see e.g., Chapter 8 of Embrechts et al. (1997). The actual value of the tail index, however, can only be calculated numerically, for instance using Theorem 2.1 in Mikosch & Stărică (2000). The tail index in models (vii)-(x) is respectively approximately 0.434, 0.262, 0.302 and 0.239. In addition, these four models feature quadratic serial dependence. The time series models (iii)-(vii) and (ix) have also been investigated by Drees (2003).

Our goal is to investigate the behaviour of our estimators with respect to changes in the sample size n and the intermediate level τ_n . Specifically, for each of the models (i)-(x) we simulate 10^4 samples of size $n = 250 \times m$, where 250 represents the number of trading days in a financial year, and m is the number of years. To obtain realistic sample sizes, we take $m = 10, 20, 30, 40$. Then, we consider the intermediate levels $\tau_n = 1 - 1/\sqrt{n} \approx 0.980, 0.986, 0.988, 0.990$. For each simulated sample we compute an estimate of ξ_{τ_n} and a confidence interval for it using the direct LAWS estimator $\tilde{\xi}_{\tau_n}$ and direct LAWS-D-E confidence interval in (B.4). Then, we compute a Monte Carlo approximation of the (actual) coverage probability and of the bias and relative variance of $\tilde{\xi}_{\tau_n}/\xi_{\tau_n} - 1$. The true value of ξ_{τ_n} is not explicit, but can be estimated to a high degree of accuracy by direct Monte-Carlo averaging, via the use of Equation (1*), using 1000 samples of size 5×10^7 from each time series model.

Details of the construction of our confidence intervals are as follows. The asymptotic variance $V(\gamma, R)$ in (B.1) is estimated with the estimator $\hat{V}_n(\gamma, R)$ in (B.3) obtained using big and small blocks of sizes $r_n = \lfloor \log^2(n) \rfloor$ and $l_n = \lfloor C \log n \rfloor$, respectively, for suitable values of the constant $C > 0$ (these are inspired by the conditions of Theorem 3.1*). The constant C is chosen such that l_n is greater than or equal to a lag after which the value of the sample autocorrelation is small, e.g., smaller than 0.1. To select μ , we recall that from simulations it appears that with small and moderate sample sizes the tail index can be severely underestimated when the data show strong serial dependence. We therefore compute the sample autocorrelation, and if it is smaller than 0.1 before the 25th lag we decide that the data show mild or weak serial dependence and set $\mu = 2$ if $\hat{\gamma}_n \leq 1/5$ (light heavy tail), $\mu = 1$ if $1/5 < \hat{\gamma}_n \leq 1/3$ (moderate heavy tail) and finally $\mu = 1/2$ if $\hat{\gamma}_n > 1/3$ (very heavy tail). Otherwise, if the lag 25 sample autocorrelation exceeds 0.1 we decide that the data show strong serial dependence and set $\mu = 2$ if $\hat{\gamma}_n \leq 1/8$, $\mu = 1$ if $1/8 < \hat{\gamma}_n \leq 1/5$ and $\mu = 1/2$ if $\hat{\gamma}_n > 1/5$. This may yield an inconsistent estimator of the asymptotic variance, but for our sample sizes this procedure seems to work reasonably well.

Tables B.2 (for $\hat{\gamma}_n = \hat{\gamma}_n^E$) and B.3 (for $\hat{\gamma}_n = \hat{\gamma}_n^H$) present the simulation results. According to

Model	$\tau_n = 0.980$	$\tau_n = 0.986$	$\tau_n = 0.988$	$\tau_n = 0.990$
(i)	2.666 [2.661, 2.671]	2.974 [2.968, 2.980]	3.165 [3.157, 3.172]	3.305 [3.297, 3.312]
(ii)	2.160 [2.158, 2.163]	2.373 [2.370, 2.376]	2.500 [2.497, 2.503]	2.591 [2.588, 2.594]
(iii)	4.821 [4.811, 4.832]	5.247 [5.235, 5.259]	5.496 [5.483, 5.509]	5.675 [5.660, 5.688]
(iv)	1.190 [1.187, 1.192]	1.297 [1.294, 1.299]	1.360 [1.358, 1.363]	1.406 [1.403, 1.409]
(v)	0.726 [0.725, 0.727]	0.816 [0.815, 0.817]	0.873 [0.872, 0.874]	0.916 [0.914, 0.917]
(vi)	3.737 [3.734, 3.741]	4.108 [4.103, 4.112]	4.342 [4.336, 4.347]	4.516 [4.510, 4.523]
(vii)	0.043 [0.043, 0.044]	0.051 [0.050, 0.051]	0.056 [0.055, 0.056]	0.060 [0.059, 0.060]
(viii)	1.566 [1.565, 1.568]	1.752 [1.750, 1.754]	1.867 [1.864, 1.869]	1.951 [1.949, 1.954]
(ix)	0.050 [0.050, 0.050]	0.057 [0.057, 0.057]	0.061 [0.061, 0.061]	0.064 [0.064, 0.064]
(x)	1.103 [1.102, 1.105]	1.227 [1.226, 1.229]	1.303 [1.301, 1.305]	1.359 [1.357, 1.361]

Table B.1: Approximate values for the expectile ξ_{τ_n} obtained with the models (i)-(x), with 95% confidence intervals in brackets.

the asymptotic theory of Theorem 3.1*, the variance and bias terms globally decrease as the sample size increases, with variations due to the dependence structure and the heavy-tail framework. When estimating γ using $\hat{\gamma}_n = \hat{\gamma}_n^E$, the coverage probabilities of our confidence intervals are satisfactory, with coverage around 85% in the worst case. The LAWS-D-E confidence interval also tends to perform better than its LAWS-D-H counterpart, as a comparison of Tables B.2 and B.3 shows. However, the results obtained with the AR(1) and ARMA(1,1) models (i)-(iv) indicate that the confidence intervals have a coverage probability that is too low in the presence of strong linear dependence. Similarly, the results for the ARCH(1) model (vii) and GARCH(1,1) model (ix) suggest that the intervals can also have a low coverage probability when the tail of the marginal distribution is quite heavy; recall that $\gamma \approx 0.432$ and 0.302 with these models. Although not perfect, our intervals take into account the serial dependence in the data and improve considerably on those rooted in i.i.d. theory, whose coverage probabilities are reported between brackets in Tables B.2 and B.3.

Finally, we repeated the first simulation experiment of this section but kept the sample size fixed in order to verify the robustness of our inferential procedure with respect to the violation of the condition $n(1 - \tau_n) \rightarrow \infty$. For each of the time series models (i)-(x) we simulated 10^4 samples of size $n = 2500$ and considered the intermediate levels $\tau = 0.980, 0.986, 0.988, 0.990$. The rest of the simulation settings

m	Model	Bias	Std dev	Cov prob	Model	Bias	Std dev	Cov prob
10	(i)	0.986	1.487	86.7 (64.1)	(vi)	1.282	0.810	95.8 (77.5)
20		1.966	1.322	90.1 (62.7)		0.384	0.653	96.6 (78.3)
30		0.393	1.157	91.0 (63.8)		1.373	0.652	96.5 (77.6)
40		4.172	1.556	92.4 (63.0)		0.569	0.576	97.5 (78.3)
10	(ii)	0.237	0.844	88.8 (66.4)	(vii)	61.458	56.798	87.6 (65.9)
20		0.180	0.629	92.7 (67.2)		18.135	2.911	89.9 (68.0)
30		0.041	0.539	93.6 (67.4)		9.866	2.399	91.3 (69.0)
40		0.001	0.454	94.9 (66.6)		16.809	3.234	91.5 (69.7)
10	(iii)	9.235	1.766	86.2 (33.6)	(viii)	0.611	0.974	95.0 (87.3)
20		1.611	1.553	87.2 (35.2)		0.110	0.640	95.3 (87.1)
30		0.003	1.229	88.8 (34.5)		0.120	0.603	95.9 (86.6)
40		0.178	1.176	88.4 (35.6)		0.021	0.502	96.4 (85.7)
10	(iv)	2.172	1.440	87.2 (43.7)	(ix)	4.932	2.243	84.9 (70.0)
20		0.111	1.395	90.5 (44.2)		1.314	1.334	87.0 (69.9)
30		0.441	0.930	90.6 (44.7)		0.841	1.309	88.7 (68.7)
40		0.113	0.845	91.2 (45.2)		1.372	1.374	89.6 (68.0)
10	(v)	0.633	0.878	95.0 (88.5)	(x)	0.494	1.074	92.6 (81.1)
20		0.106	0.624	95.4 (91.4)		0.086	0.683	93.8 (79.2)
30		0.016	0.590	96.9 (92.7)		0.001	0.576	94.9 (77.8)
40		0.076	0.546	97.0 (91.8)		0.046	0.514	95.4 (78.2)

Table B.2: Monte-Carlo simulations: bias $\times 10^5$ and standard deviation $\times 10$ of the relative estimator $(\tilde{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$ and coverage probability (in %) for the estimator in (B.4), with 95% nominal level, obtained with sample size $n = 250 \times m$, and estimating the tail index with $\hat{\gamma}_n^E$. The brackets contain the empirical coverage of the corresponding LAWS-IID-E interval.

remained unchanged.

Table B.4 collects the results, which remain satisfactory overall, leading to conclusions similar to those of the first experiment: the results deteriorate when the tail of the marginal distribution is heavier or

m	Model	Coverage	Model	Coverage
10	(i)	83.0 (58.9)	(vi)	91.9 (79.5)
20		87.3 (61.2)		95.1 (80.8)
30		89.9 (61.6)		96.1 (80.3)
40		90.9 (61.9)		96.9 (80.6)

10	(ii)	75.2 (52.9)	(vii)	86.4 (57.7)
20		77.0 (55.0)		89.4 (61.8)
30		78.7 (56.7)		90.4 (63.8)
40		79.3 (57.8)		91.9 (65.4)

10	(iii)	62.7 (20.9)	(viii)	93.8 (82.8)
20		69.6 (22.4)		94.6 (82.8)
30		73.3 (24.5)		95.8 (83.1)
40		75.4 (25.3)		95.7 (82.4)

10	(iv)	74.5 (32.0)	(ix)	83.6 (65.6)
20		81.0 (36.6)		86.7 (66.9)
30		85.0 (39.7)		87.7 (67.7)
40		86.6 (40.7)		88.6 (67.2)

10	(v)	94.7 (88.9)	(x)	89.7 (73.6)
20		96.4 (90.7)		91.9 (73.6)
30		97.1 (91.9)		92.8 (73.6)
40		97.4 (92.5)		94.2 (72.6)

Table B.3: Empirical coverages (%) of the interval (B.4) with $\alpha = 0.05$ when $n = 250 \times m$ and the tail index is estimated by $\hat{\gamma}_n^H$. The brackets contain the empirical coverages of the corresponding LAWS-IID-H interval.

the serial dependence is stronger. In contrast to the previous simulation results, here they also slightly deteriorate with increasing intermediate levels because expectile estimation is increasingly challenging as the intermediate level grows. Similar conclusions hold for the coverage probabilities.

τ	Model	Bias	Std dev	Cov Prob	Model	Bias	Std dev	Cov Prob
0.980	(i)	2.992	1.965	87.1 (64.0)	(vi)	0.886	0.759	95.4 (77.5)
0.986		2.296	1.933	85.6 (65.3)		3.580	1.067	93.4 (78.3)
0.988		4.017	1.927	83.6 (65.6)		4.465	1.145	92.3 (78.5)
0.990		11.161	2.623	82.3 (64.9)		3.568	1.214	90.5 (78.8)

0.980	(ii)	0.753	0.840	89.1 (66.1)	(vii)	0.036	3.483	87.9 (65.9)
0.986		0.304	0.973	88.9 (67.4)		0.091	4.247	84.0 (64.4)
0.988		0.007	0.998	88.1 (67.4)		0.013	4.871	82.1 (63.8)
0.990		0.007	1.008	84.2 (67.5)		0.146	7.865	79.4 (62.0)

0.980	(iii)	6.096	1.908	85.3 (34.4)	(viii)	0.221	0.838	94.5 (87.6)
0.986		3.230	2.073	80.4 (35.3)		0.222	1.143	93.8 (86.7)
0.988		3.786	2.107	77.4 (36.1)		0.137	1.000	93.6 (86.6)
0.990		13.853	1.978	74.5 (35.2)		0.259	1.148	92.7 (85.7)

0.980	(iv)	1.942	1.448	86.5 (42.9)	(ix)	0.010	1.694	84.6 (69.5)
0.986		1.353	1.728	83.3 (44.2)		0.010	2.075	82.2 (69.0)
0.988		2.166	1.597	81.6 (46.5)		0.010	2.933	80.4 (67.8)
0.990		0.727	1.685	79.9 (49.4)		0.020	2.362	79.7 (67.8)

0.980	(v)	0.071	0.768	94.7 (88.9)	(x)	0.002	0.852	92.8 (80.3)
0.986		0.338	0.931	94.8 (89.7)		0.726	1.035	91.0 (79.8)
0.988		0.299	0.967	94.7 (90.1)		2.351	1.051	90.6 (79.9)
0.990		2.672	1.188	94.3 (89.6)		1.044	1.103	90.4 (80.0)

Table B.4: Monte-Carlo simulations: bias $\times 10^5$ and standard deviation $\times 10$ of the relative estimator $(\tilde{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$ and coverage probability (in %) for the estimator in (B.4), with 95% nominal level, obtained with sample size $n = 2500$, and estimating the tail index with $\hat{\gamma}_n^E$. Next to the coverage probability, we report (between brackets) the coverage probability obtained using the LAWS-IID-E interval.

B.3 Further results

We provide approximate values of the true expectiles at level τ'_n and QMES_{X,α_n} (the latter used in Section 7.2*) in Table B.5 and graphs of non-coverage probabilities of our confidence intervals in models (i)-(x) at the extreme levels 0.9995 and 0.9999 in Figures B.1, B.2 and B.3. The models (a), (b), (c) and (d) considered in the main paper respectively correspond to models (i), (iii), (viii) and (x) of this document. We also report graphs of tail index estimates for our real data illustrations in Figures B.4 and B.5.

Model	$\tau'_n = 0.9995$	$\tau'_n = 0.9999$
(i)	8.120 [8.032, 8.219]	13.612 [13.253, 14.017]
(ii)	5.002 [4.978, 5.027]	7.152 [7.068, 7.238]
(iii)	10.290 [10.130, 10.461]	15.836 [15.140, 16.778]
(iv)	2.882 [2.845, 2.927]	4.722 [4.588, 4.911]
(v)	2.482 [2.468, 2.496]	4.240 [4.186, 4.303]
(vi)	10.955 [10.866, 11.044]	18.388 [18.057, 18.767]
(vii)	0.222 [0.217, 0.229]	0.447 [0.428, 0.473]
(viii)	4.546 [4.521, 4.571]	6.984 [6.903, 7.064]
(ix)	0.164 [0.163, 0.166]	0.269 [0.262, 0.276]
(x)	2.978 [2.963, 2.994]	4.419 [4.372, 4.469]

Model	$\tau'_n(\alpha_n) = 0.9995$	$\tau'_n(\alpha_n) = 0.9999$
(e)	9.674 [9.553, 9.823]	16.479 [15.988, 17.078]
(f)	13.137 [12.967, 13.346]	21.188 [20.262, 22.418]
(g)	6.101 [5.991, 6.226]	8.543 [8.166, 8.925]
(h)	6.887 [6.772, 7.022]	10.628 [10.217, 11.153]

Table B.5: Top: Approximate values for the expectile $\xi_{\tau'_n}$ obtained with the time series models (i)-(x), with 95% confidence intervals between brackets. Results obtained from Equation (1*) by Monte-Carlo averaging, using 1000 samples of size 5×10^7 from each model. Models (a), (b), (c) and (d) considered in Section 7.1* of the main paper respectively correspond to models (i), (iii), (viii) and (x) here. Bottom: Approximate values for the marginal expected shortfall $\text{QMES}_{X, \alpha_n}$ for the models (e), (f), (g) and (h) considered in Section 7.2*. Results obtained from Equation (4*) by Monte-Carlo averaging, using 1000 samples of size 5×10^7 from each model.

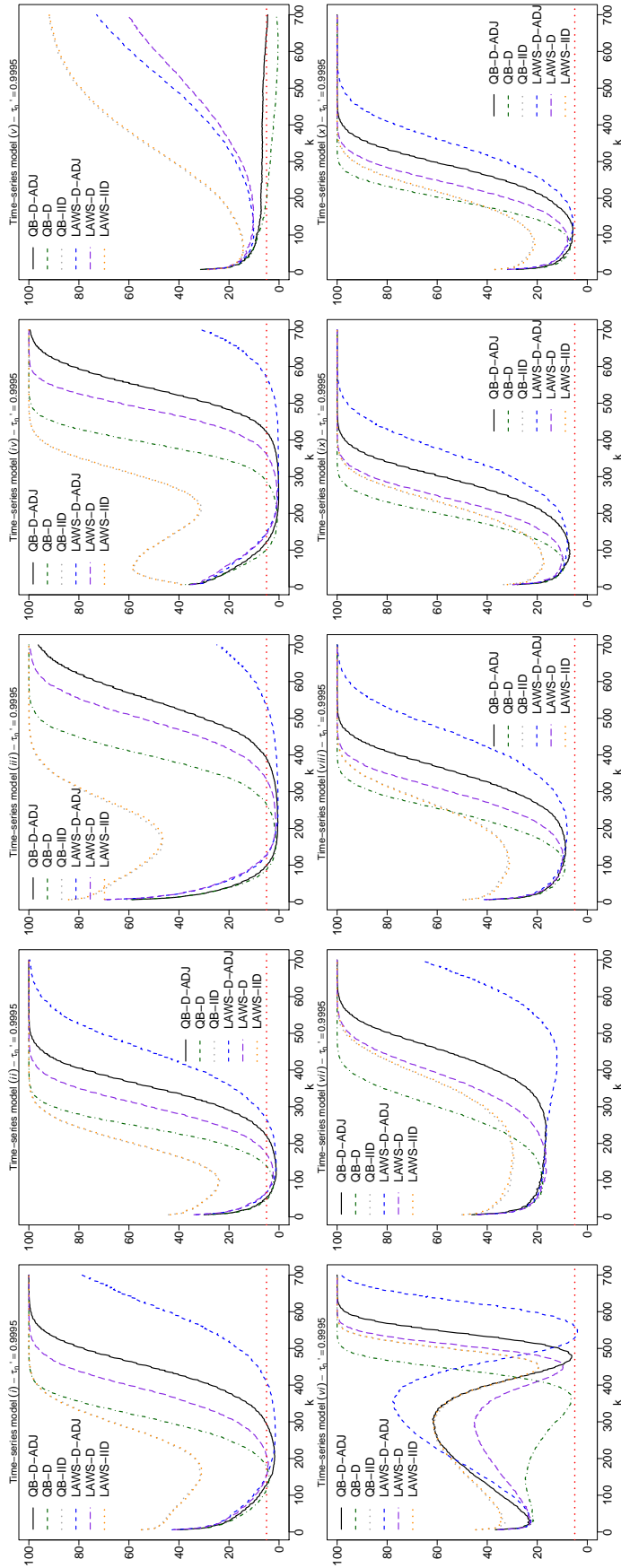


Figure B.1: Empirical error rates (%) for nominal 95% confidence intervals for the expectile $\xi_{\tau'_n}$ with $\tau'_n = 0.9995$. From left to right, top row: models (i)-(v); bottom row: models (vi)-(x). In each panel, the horizontal dotted red line represents the nominal 5% error rate.

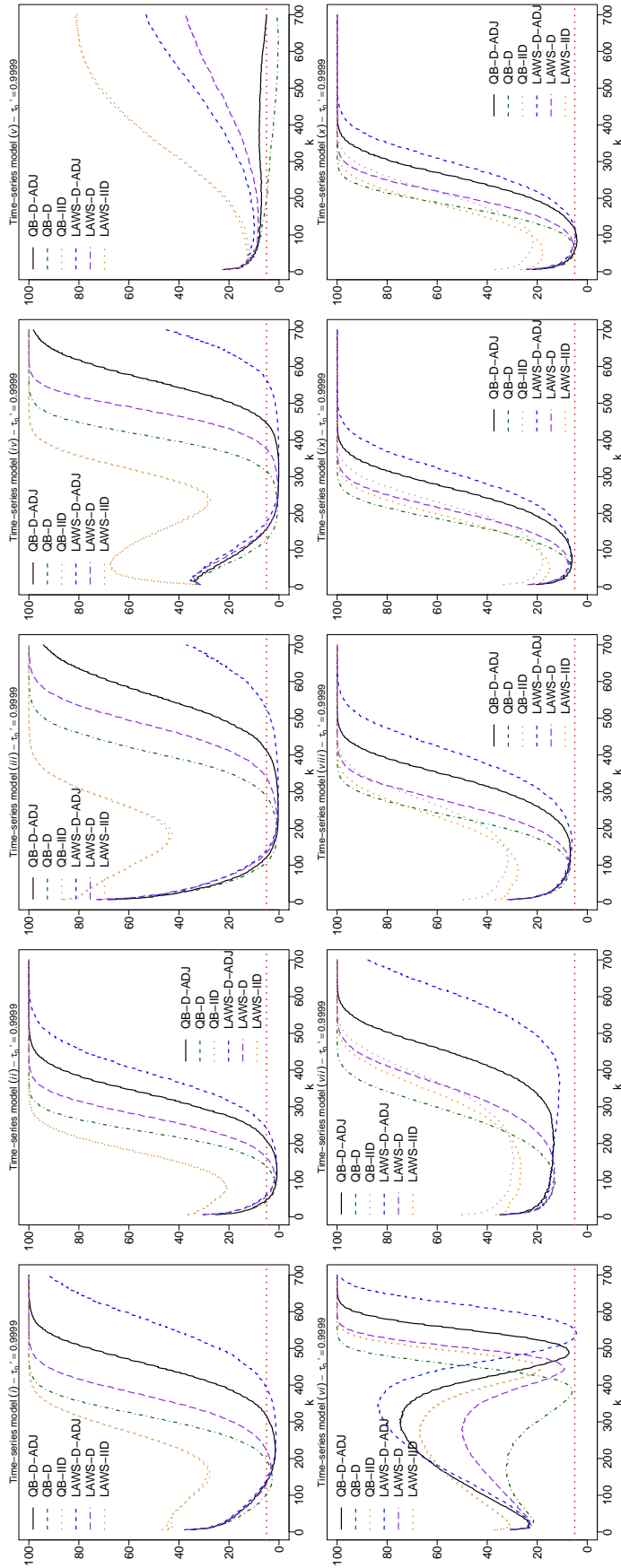


Figure B.2: Empirical error rates (%) for nominal 95% confidence intervals for the expectile $\xi_{\tau'_n}$ with $\tau'_n = 0.9999$. From left to right, top row: models (i)-(v); bottom row: models (vi)-(x). In each panel, the horizontal dotted red line represents the nominal 5% error rate.

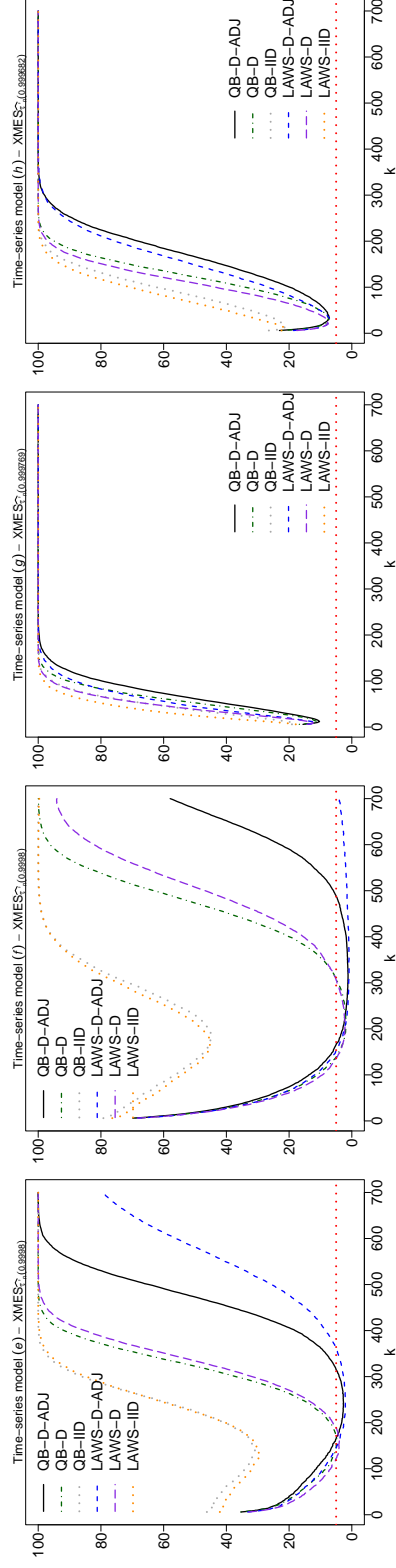


Figure B.3: Empirical error rates (%) for nominal 95% confidence intervals for the marginal expected shortfall $\text{QMES}_{X, \alpha_n}$ such that $\tau'_n(\alpha_n) = 0.9999$. From left to right: models (e)–(h).

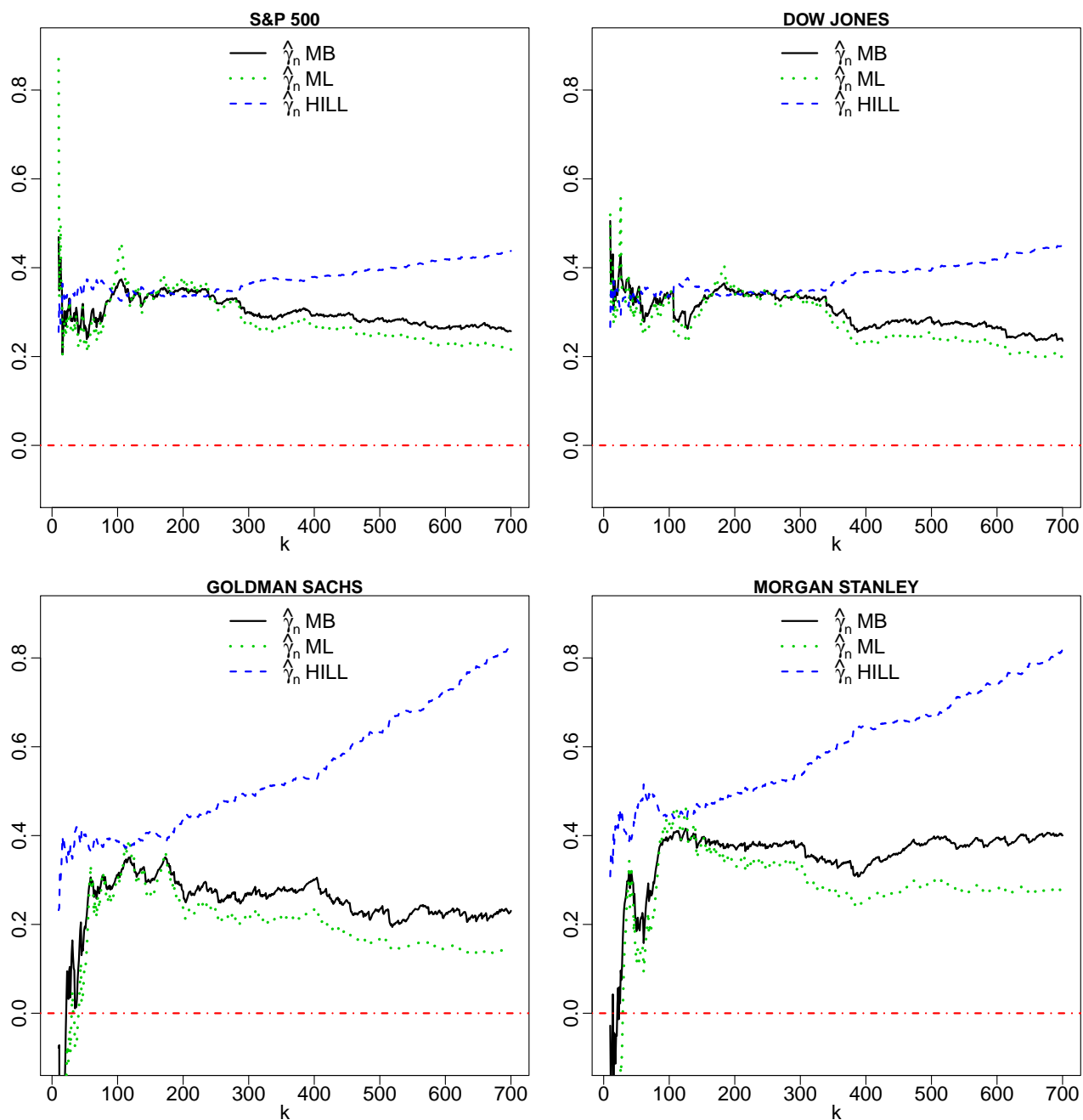


Figure B.4: Tail index estimation for the financial data in Section 8.1* and the two banks in Section 8.2*, showing point estimates of γ obtained with the Hill (blue dashed line), ML (green dotted line) and MB (black solid line) estimators. Top row: S&P 500 and Dow Jones Industrial Average, bottom row: Goldman Sachs and Morgan Stanley.

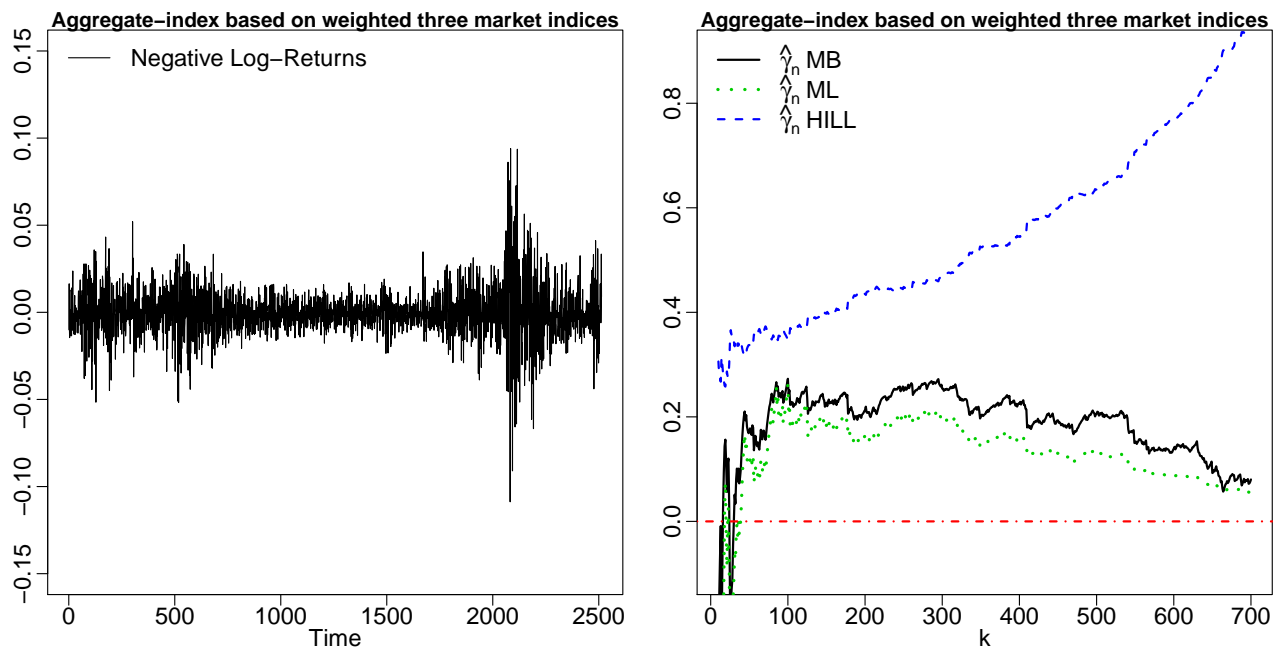


Figure B.5: Complement to the financial data analysis in Section 8.2*. Left panel: Daily log-returns of the aggregated market index. Right panel: Point estimates of the tail index γ_Y of these daily log-returns obtained with the Hill (blue dashed line), ML (green dotted line) and MB (black solid line) estimators.

C Proofs

We first state a couple of facts that will be used multiple times in our proofs. If Z , a random variable having survival function \bar{F} and tail quantile function U , is heavy-tailed with tail index γ , and if $\tau_n \uparrow 1$ and $z_n \rightarrow \infty$ as $n \rightarrow \infty$, then the convergences

$$\frac{\bar{F}(z_n)}{1 - \tau_n} \rightarrow z \in (0, \infty) \quad \text{and} \quad \frac{z_n}{U((1 - \tau_n)^{-1})} \rightarrow z^{-\gamma}, \quad \text{as } n \rightarrow \infty,$$

are equivalent. This local inversion property is a consequence of Definition B.1.8 and Proposition B.1.9.10 (pp. 366–367) in [de Haan & Ferreira \(2006\)](#). In particular, if $\mathbb{E}|Z_-| < \infty$ and $\gamma < 1$, then high expectiles ξ_{τ_n} of Z are well-defined and satisfy

$$\frac{\bar{F}(\xi_{\tau_n})}{1 - \tau_n} \rightarrow \gamma^{-1} - 1 \quad \text{and} \quad \frac{\xi_{\tau_n}}{U((1 - \tau_n)^{-1})} \rightarrow (\gamma^{-1} - 1)^{-\gamma} \quad \text{as } n \rightarrow \infty.$$

See for instance Proposition 1 in [Daouia et al. \(2019\)](#) in the case $p = 2$ (with the notation therein). In what follows, this will be referred to as the asymptotic proportionality of expectiles and quantiles.

Lemma C.1 establishes the convergence of a rescaled multivariate exceedance probability (above high thresholds) to the tail copula. It is a stronger version of Lemma 5 in [Stupfler \(2019\)](#).

Lemma C.1. *Assume that X_1, \dots, X_d are heavy-tailed random variables with continuous distribution functions F_1, \dots, F_d , tail quantile functions U_1, \dots, U_d and tail indices $\gamma_1, \dots, \gamma_d$. Suppose that there is a function R on $[0, \infty]^d \setminus \{(\infty, \dots, \infty)\}$ such that*

$$\forall (x_1, \dots, x_d) \in [0, \infty]^d \setminus \{(\infty, \dots, \infty)\}, \quad \lim_{s \rightarrow \infty} s \mathbb{P} \left(\bar{F}_j(X_j) \leq \frac{x_j}{s}, 1 \leq j \leq d \right) = R(x_1, \dots, x_d).$$

Let $\tau_n \uparrow 1$ and $a_{j,n}$, $1 \leq j \leq d$ be sequences such that $a_{j,n}/U_j((1 - \tau_n)^{-1}) \rightarrow a_j \in (0, \infty)$ as $n \rightarrow \infty$. Then for any $x_1, \dots, x_d > 0$,

$$\frac{1}{1 - \tau_n} \mathbb{P}(X_j > a_{j,n} x_j, 1 \leq j \leq d) \rightarrow R \left(a_1^{-1/\gamma_1} x_1^{-1/\gamma_1}, \dots, a_d^{-1/\gamma_d} x_d^{-1/\gamma_d} \right) \quad \text{as } n \rightarrow \infty.$$

Proof of Lemma C.1. Pick $x_1, \dots, x_d > 0$. By continuity of the F_j ,

$$\lim_{s \rightarrow \infty} s \mathbb{P}(X_j > U_j(s/x_j), 1 \leq j \leq d) = R(x_1, \dots, x_d).$$

Now pick an arbitrary $\varepsilon \in (0, 1)$. The regular variation of U_j yields

$$\frac{a_{j,n} x_j}{U_j((1 - \varepsilon)(a_j x_j)^{1/\gamma_j} (1 - \tau_n)^{-1})} \rightarrow (1 - \varepsilon)^{-\gamma_j} > 1.$$

We therefore find, for n large enough, that

$$\mathbb{P}(X_j > a_{j,n}x_j, 1 \leq j \leq d) \leq \mathbb{P}(X_j > U_j((1 - \varepsilon)(a_jx_j)^{1/\gamma_j}(1 - \tau_n)^{-1}), 1 \leq j \leq d).$$

By 1-homogeneity of R (a direct consequence of its definition), this yields

$$\limsup_{n \rightarrow \infty} \frac{1}{1 - \tau_n} \mathbb{P}(X_j > a_{j,n}x_j, 1 \leq j \leq d) \leq (1 - \varepsilon)^{-1} R\left(a_1^{-1/\gamma_1}x_1^{-1/\gamma_1}, \dots, a_d^{-1/\gamma_d}x_d^{-1/\gamma_d}\right)$$

with an analogous lower bound for the limit inferior with ε replaced by $-\varepsilon$. Letting $\varepsilon \downarrow 0$ completes the proof. \square

Lemma C.2 is important for the calculation of the asymptotic variance in Theorem 3.1*.

Lemma C.2. *Assume that Conditions A*, B(ii)* and B(iii)* are satisfied. Let $r_n \rightarrow \infty$ and $\tau_n \uparrow 1$ be such that $r_n(1 - \tau_n) \rightarrow 0$ and $n(1 - \tau_n) \rightarrow \infty$ as $n \rightarrow \infty$. Then, if $s_n \rightarrow \infty$ is a sequence of integers such that $s_n = O(r_n)$, we have, as $n \rightarrow \infty$,*

$$\text{Var}\left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} [\mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{P}(Y_1 > \xi_{\tau_n})]\right) \rightarrow (\gamma^{-1} - 1) \left(1 + 2 \sum_{t=1}^{\infty} R_t(1, 1)\right) < \infty.$$

If moreover $0 < \gamma < 1/2$, then we also have, as $n \rightarrow \infty$,

$$\begin{aligned} & \text{Var}\left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} \frac{1}{\xi_{\tau_n}} [(Y_t - \xi_{\tau_n}) \mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})]\right) \\ & \rightarrow \frac{2\gamma}{1 - 2\gamma} + 2(\gamma^{-1} - 1) \iint_{[1, \infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy < \infty. \end{aligned}$$

Proof of Lemma C.2. We only prove the second convergence result; the proof of the first uses the same ideas but is much simpler. By stationarity of (Y_t) ,

$$\begin{aligned} & \text{Var}\left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} \frac{1}{\xi_{\tau_n}} [(Y_t - \xi_{\tau_n}) \mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})]\right) \\ & = \frac{1}{(1 - \tau_n)\xi_{\tau_n}^2} \text{Var}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\}) - \frac{s_n - 1}{(1 - \tau_n)\xi_{\tau_n}^2} [\mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})]^2 \\ & + \frac{2}{s_n(1 - \tau_n)\xi_{\tau_n}^2} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E}((Y_1 - \xi_{\tau_n})(Y_{t+1} - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}, Y_{t+1} > \xi_{\tau_n}\}). \end{aligned} \quad (\text{C.5})$$

Equation (B.11) in the proof of Lemma 4 in Daouia et al. (2018), applied for both $a = 1$ and 2, yields

$$\begin{aligned} & \frac{1}{(1 - \tau_n)\xi_{\tau_n}^2} \text{Var}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\}) - \frac{s_n - 1}{(1 - \tau_n)\xi_{\tau_n}^2} [\mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})]^2 \\ & = \frac{2\gamma}{1 - 2\gamma} + o(1). \end{aligned} \quad (\text{C.6})$$

We now concentrate on the sum in (C.5). An integration by parts and a change of variables provide

$$\begin{aligned} & \frac{2}{s_n(1-\tau_n)\xi_{\tau_n}^2} \sum_{t=1}^{s_n-1} (s_n-t) \mathbb{E}((Y_1 - \xi_{\tau_n})(Y_{t+1} - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}, Y_{t+1} > \xi_{\tau_n}\}) \\ &= 2 \sum_{t=1}^{\infty} \frac{s_n-t}{s_n} \iint_{[1,\infty)^2} \frac{1}{1-\tau_n} \mathbb{P}(Y_1 > \xi_{\tau_n}x, Y_{t+1} > \xi_{\tau_n}y) dx dy \mathbb{1}\{t < s_n\}. \end{aligned} \quad (\text{C.7})$$

From the asymptotic proportionality of expectiles and quantiles, Condition B(ii)* and Lemma C.1,

$$\frac{s_n-t}{s_n} \frac{1}{1-\tau_n} \mathbb{P}(Y_1 > \xi_{\tau_n}x, Y_{t+1} > \xi_{\tau_n}y) \mathbb{1}\{t < s_n\} \rightarrow (\gamma^{-1} - 1)R_t(x^{-1/\gamma}, y^{-1/\gamma}) \quad (\text{C.8})$$

as $n \rightarrow \infty$, pointwise in $x, y \geq 1$ for any positive integer t . We then swap limit and sum of integrals in (C.7), noting that, by continuity of F , Potter bounds (see Proposition B.1.9.5 in de Haan & Ferreira 2006), Condition B(iii)* and the asymptotic proportionality of expectiles and quantiles, we have, for any $\varepsilon > 0$,

$$\begin{aligned} & \frac{s_n-t}{s_n} \frac{1}{1-\tau_n} \mathbb{P}(Y_1 > \xi_{\tau_n}x, Y_{t+1} > \xi_{\tau_n}y) \mathbb{1}\{t < s_n\} \\ & \leq \frac{1}{1-\tau_n} \mathbb{P}(\bar{F}(Y_1) \leq Cx^{-1/\gamma+2\varepsilon}\bar{F}(\xi_{\tau_n}), \bar{F}(Y_{t+1}) \leq Cy^{-1/\gamma+2\varepsilon}\bar{F}(\xi_{\tau_n})) \mathbb{1}\{t < s_n\} \\ & \leq C \frac{\bar{F}(\xi_{\tau_n})}{1-\tau_n} (\rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon} + \bar{F}(\xi_{\tau_n})x^{-1/\gamma+2\varepsilon}y^{-1/\gamma+2\varepsilon}) \mathbb{1}\{t < s_n\} \\ & \leq C (\rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon} + (1-\tau_n)x^{-1/\gamma+2\varepsilon}y^{-1/\gamma+2\varepsilon}) \mathbb{1}\{t < s_n\} \end{aligned} \quad (\text{C.9})$$

for n large enough, irrespective of the values of the positive integer t and $x, y \geq 1$; here C is an unimportant positive constant whose value may change from one line to the next. Now choose ε such that $(1+\varepsilon)2\gamma < 1$ (which is possible since $\gamma < 1/2$). The upper bound satisfies

$$(\rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon} + (1-\tau_n)x^{-1/\gamma+2\varepsilon}y^{-1/\gamma+2\varepsilon}) \mathbb{1}\{t < s_n\} \rightarrow \rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon}, \quad (\text{C.10})$$

as $n \rightarrow \infty$, pointwise in t, x, y . The functions $u \mapsto u^{-1/2\gamma+\varepsilon}$ and $u \mapsto u^{-1/\gamma+2\varepsilon}$ are both integrable on $[1, \infty)^2$, so

$$\begin{aligned} & \sum_{t=1}^{\infty} \iint_{[1,\infty)^2} (\rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon} + (1-\tau_n)x^{-1/\gamma+2\varepsilon}y^{-1/\gamma+2\varepsilon}) dx dy \mathbb{1}\{t < s_n\} \\ & \rightarrow \sum_{t=1}^{\infty} \iint_{[1,\infty)^2} \rho(t)x^{-1/2\gamma+\varepsilon}y^{-1/2\gamma+\varepsilon} dx dy < \infty, \end{aligned} \quad (\text{C.11})$$

as $n \rightarrow \infty$, by splitting the sum and using the convergence $s_n(1 - \tau_n) \rightarrow 0$. A combination of Theorem 1 in Pratt (1960) with (C.7), (C.8), (C.9), (C.10), (C.11) and the Tonelli theorem yields, as $n \rightarrow \infty$,

$$\begin{aligned} & \frac{2}{s_n(1 - \tau_n)\xi_{\tau_n}^2} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E}((Y_1 - \xi_{\tau_n})(Y_{t+1} - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}, Y_{t+1} > \xi_{\tau_n}\}) \\ & \rightarrow 2(\gamma^{-1} - 1) \sum_{t=1}^{\infty} \iint_{[1, \infty)^2} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy. \end{aligned} \quad (\text{C.12})$$

Combining (C.5), (C.6) and (C.12) completes the proof. \square

*Proof of Theorem 3.1**. Define a convex function ψ_n by

$$\psi_n(u) := \frac{1}{2\xi_{\tau_n}^2} \sum_{t=1}^n \left[\eta_{\tau_n} \left(Y_t - \xi_{\tau_n} - \frac{u \xi_{\tau_n}}{\sqrt{n(1 - \tau_n)}} \right) - \eta_{\tau_n}(Y_t - \xi_{\tau_n}) \right]$$

and define $\varphi_{\tau}(y) = |\tau - \mathbb{1}\{y \leq 0\}|y$. Use (B.12) in the proof of Theorem 2 in Daouia et al. (2018) (a purely analytical step that is not impacted by the time series context) to find

$$\sqrt{n(1 - \tau_n)} \left(\frac{\tilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) = \arg \min_{u \in \mathbb{R}} \psi_n(u) = \arg \min_{u \in \mathbb{R}} \{-uT_{1,n} + T_{2,n}(u)\} \quad (\text{C.13})$$

$$\text{with } T_{1,n} := \frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^n \frac{1}{\xi_{\tau_n}} \varphi_{\tau_n}(Y_t - \xi_{\tau_n})$$

$$\text{and } T_{2,n}(u) := -\frac{1}{\xi_{\tau_n}^2} \sum_{t=1}^n \int_0^{u\xi_{\tau_n}/\sqrt{n(1-\tau_n)}} (\varphi_{\tau_n}(Y_t - \xi_{\tau_n} - z) - \varphi_{\tau_n}(Y_t - \xi_{\tau_n})) dz.$$

We concentrate on $T_{1,n}$ first. Use the identities $\mathbb{E}(\varphi_{\tau_n}(Y_1 - \xi_{\tau_n})) = 0$ and $\varphi_{\tau_n}(x) = (1 - \tau_n)x + (2\tau_n - 1)x\mathbb{1}\{x > 0\}$ to write

$$\begin{aligned} T_{1,n} &= \frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^n \frac{1 - \tau_n}{\xi_{\tau_n}} [Y_t - \mathbb{E}(Y_1)] \\ &+ \frac{2\tau_n - 1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^n \frac{1}{\xi_{\tau_n}} [(Y_t - \xi_{\tau_n}) \mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})] \\ &=: T_{1,n}^{(1)} + (2\tau_n - 1)T_{1,n}^{(2)}. \end{aligned} \quad (\text{C.14})$$

We analyse the convergences of $T_{1,n}^{(1)}$ and $T_{1,n}^{(2)}$ separately. Note that $T_{1,n}^{(1)}$ is centred and satisfies, by stationarity of (Y_t) ,

$$\text{Var}(T_{1,n}^{(1)}) = \frac{(1 - \tau_n)}{\xi_{\tau_n}^2} \left(\text{Var}(Y_1) + 2 \sum_{l=1}^{\infty} \frac{n-l}{n} \text{Cov}(Y_1, Y_{l+1}) \mathbb{1}\{l < n\} \right).$$

Our assumptions ensure that $\mathbb{E}|Y_1|^{2+\delta} < \infty$. Using Lemma 1.3 in [Ibragimov \(1962\)](#), we can find a finite constant C such that $|\text{Cov}(Y_1, Y_{l+1})| \leq C[\alpha(l)]^{\delta/(2+\delta)}$. Therefore, by the dominated convergence theorem, one obtains, as $n \rightarrow \infty$,

$$\sum_{l=1}^{\infty} \frac{n-l}{n} \text{Cov}(Y_1, Y_{l+1}) \mathbb{1}\{l < n\} \rightarrow \sum_{l=1}^{\infty} \text{Cov}(Y_1, Y_{l+1}) < \infty$$

and in particular $\text{Var}(T_{1,n}^{(1)}) = O((1 - \tau_n)/\xi_{\tau_n}^2) \rightarrow 0$ as $n \rightarrow \infty$. From this we conclude that

$$T_{1,n}^{(1)} \xrightarrow{\mathbb{P}} 0. \quad (\text{C.15})$$

We now focus on the term $T_{1,n}^{(2)}$. To analyse its convergence, we use Theorem 4.1 in [Rootzén et al. \(1998\)](#), which is a central limit theorem for tail array sums. We first check the assumptions of this theorem. Recall, for $i < j$, the notation $\mathcal{F}_{i,j} = \sigma(Y_i, Y_{i+1}, \dots, Y_j)$ and note that for any $n \geq 2$ and $l < n$,

$$\alpha_n(l) := \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, A \in \mathcal{F}_{1,k}, B \in \mathcal{F}_{k+l,n}, 1 \leq k \leq n-l\} \leq \alpha(l).$$

Consequently, in the terminology of [Rootzén et al. \(1998\)](#), the time series $(Y_t, t \geq 1)$ is strongly mixing $(\alpha_n(l), l_n)$. Let further, for any positive sequence of integers (s_n) converging to infinity and such that $s_n \leq r_n$,

$$\sigma_{s_n}^2 := \text{Var} \left(\frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^{s_n} \frac{1}{\xi_{\tau_n}} [(Y_t - \xi_{\tau_n}) \mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})] \right)$$

and set

$$\Psi_n(x) := \sqrt{\frac{r_n}{n\sigma_{r_n}^2}} \frac{1}{\sqrt{n(1-\tau_n)}} \frac{1}{\xi_{\tau_n}} [x \mathbb{1}\{x > 0\} - \mathbb{E}((Y_1 - \xi_{\tau_n}) \mathbb{1}\{Y_1 > \xi_{\tau_n}\})].$$

With this notation,

$$T_{1,n}^{(2)} = \sqrt{\frac{n\sigma_{r_n}^2}{r_n}} \sum_{t=1}^n \Psi_n(Y_t - \xi_{\tau_n}) =: \sqrt{\frac{n\sigma_{r_n}^2}{r_n}} T_{1,n}^{(3)},$$

where $T_{1,n}^{(3)}$ is, in the terminology of [Rootzén et al. \(1998\)](#), a tail array sum satisfying the so-called basic assumptions of Section 2 therein. Furthermore

$$\mathbb{E}(\Psi_n(Y_1 - \xi_{\tau_n})) = 0 \quad \text{and} \quad \text{Var} \left(\sum_{t=1}^{r_n} \Psi_n(Y_t - \xi_{\tau_n}) \right) = \frac{r_n}{n}$$

by definition of Ψ_n and $\sigma_{r_n}^2$. Besides, by Lemma [C.2](#) and for any sequence (s_n) converging to infinity such that $s_n \leq r_n$,

$$\frac{n\sigma_{s_n}^2}{s_n} \rightarrow \frac{2\gamma}{1-2\gamma} + 2(\gamma^{-1} - 1) \iint_{[1,\infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy < \infty \quad \text{as } n \rightarrow \infty. \quad (\text{C.16})$$

Consequently, as $n \rightarrow \infty$,

$$\frac{n}{r_n} \text{Var} \left(\sum_{t=1}^{s_n} \Psi_n(Y_t - \xi_{\tau_n}) \right) = \frac{s_n}{r_n} \times \frac{n \sigma_{s_n}^2 / s_n}{n \sigma_{r_n}^2 / r_n} = O \left(\frac{s_n}{r_n} \right).$$

Recalling that $l_n/r_n \rightarrow 0$ and $r_n/n \rightarrow 0$ as $n \rightarrow \infty$, we find that assumption (2.3) of [Rootzén et al. \(1998\)](#) is satisfied. Note finally that, letting $Z_{n,1} := \sum_{t=1}^{r_n} \Psi_n(Y_t - \xi_{\tau_n})$, we have, for any $\varepsilon > 0$,

$$\frac{n}{r_n} \mathbb{E}(Z_{n,1}^2 \mathbb{1}\{|Z_{n,1}| > \varepsilon\}) \leq \varepsilon^{-\delta} \times \frac{n}{r_n} \mathbb{E}|Z_{n,1}|^{2+\delta} = O(n r_n^{1+\delta} \mathbb{E}|\Psi_n(Y_1 - \xi_{\tau_n})|^{2+\delta})$$

by Hölder's inequality. Combining Equation (B.11) in the proof of Lemma 4 in [Daouia et al. \(2018\)](#) with our Equation (C.16) and Hölder's inequality again, we find

$$\frac{n}{r_n} \mathbb{E}(Z_{n,1}^2 \mathbb{1}\{|Z_{n,1}| > \varepsilon\}) = O \left(r_n \left[\frac{r_n}{\sqrt{n(1-\tau_n)}} \right]^\delta \right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Theorem 4.1 in [Rootzén et al. \(1998\)](#) then entails $T_{1,n}^{(3)} \xrightarrow{d} N(0,1)$, or equivalently, by (C.16) and Slutsky's lemma,

$$(2\tau_n - 1)T_{1,n}^{(2)} \xrightarrow{d} \sqrt{\frac{2\gamma}{1-2\gamma} + 2(\gamma^{-1} - 1) \iint_{[1,\infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy} \times Z, \quad (\text{C.17})$$

where Z has a standard Gaussian distribution. We now turn to the control of $T_{2,n}(u)$, which we rewrite as

$$T_{2,n}(u) = T_{3,n}(u) - \frac{n}{\xi_{\tau_n}^2} \int_0^{u\xi_{\tau_n}/\sqrt{n(1-\tau_n)}} [\mathbb{E}(\varphi_{\tau_n}(Y_1 - \xi_{\tau_n} - z)) - \mathbb{E}(\varphi_{\tau_n}(Y_1 - \xi_{\tau_n}))] dz.$$

The second term on the right-hand side is purely deterministic, so it can be controlled just as in the i.i.d. case. By Equation (B.19) in the proof of Theorem 2 in [Daouia et al. \(2018\)](#),

$$T_{2,n}(u) = \frac{u^2}{2\gamma}(1 + o(1)) + T_{3,n}(u), \quad (\text{C.18})$$

with $T_{3,n}(u) := -\frac{1}{\xi_{\tau_n}^2} \sum_{t=1}^n \int_0^{u\xi_{\tau_n}/\sqrt{n(1-\tau_n)}} [\mathcal{S}_{n,t}(\xi_{\tau_n} + z) - \mathcal{S}_{n,t}(\xi_{\tau_n})] dz,$

where $\mathcal{S}_{n,t}(v) := \varphi_{\tau_n}(Y_t - v) - \mathbb{E}(\varphi_{\tau_n}(Y_1 - v))$. We finally control $T_{3,n}(u)$. By Lemma 3 in [Daouia et al. \(2018\)](#), one has

$$\begin{aligned} -[\mathcal{S}_{n,t}(\xi_{\tau_n} + z) - \mathcal{S}_{n,t}(\xi_{\tau_n})] &= [|\tau_n - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\}| - \mathbb{E}|\tau_n - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\}|] z \\ &\quad + (2\tau_n - 1)(Y_t - \xi_{\tau_n} - z)(\mathbb{1}\{Y_t \leq \xi_{\tau_n} + z\} - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\}) \\ &\quad - (2\tau_n - 1)\mathbb{E}[(Y_t - \xi_{\tau_n} - z)(\mathbb{1}\{Y_t \leq \xi_{\tau_n} + z\} - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\})]. \end{aligned}$$

Noting that $\mathbb{1}\{Y_t \leq \xi_{\tau_n} + z\} - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\} = \mathbb{1}\{z \geq Y_t - \xi_{\tau_n}, Y_t > \xi_{\tau_n}\}$ if $z > 0$ and $-\mathbb{1}\{z < Y_t - \xi_{\tau_n}, Y_t \leq \xi_{\tau_n}\}$ otherwise, straightforward calculations entail

$$\begin{aligned} T_{3,n}(u) &= (2\tau_n - 1) \frac{u^2}{2} \times \frac{1}{n(1 - \tau_n)} \sum_{t=1}^n [\mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{P}(Y_1 > \xi_{\tau_n})] \\ &\quad - \frac{2\tau_n - 1}{2\xi_{\tau_n}^2} \left\{ \sum_{t=1}^n \left(Y_t - \xi_{\tau_n} - \frac{u\xi_{\tau_n}}{\sqrt{n(1 - \tau_n)}} \right)^2 [\mathbb{1}\{Y_t \leq \xi_{\tau_n} + u\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}\} - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\}] \right. \\ &\quad \left. - n \mathbb{E} \left[\left(Y_1 - \xi_{\tau_n} - \frac{u\xi_{\tau_n}}{\sqrt{n(1 - \tau_n)}} \right)^2 [\mathbb{1}\{Y_1 \leq \xi_{\tau_n} + u\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}\} - \mathbb{1}\{Y_1 \leq \xi_{\tau_n}\}] \right] \right\}. \quad (\text{C.19}) \end{aligned}$$

The sum

$$\frac{1}{n(1 - \tau_n)} \sum_{t=1}^n [\mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{P}(Y_1 > \xi_{\tau_n})]$$

is again a centred tail array sum in the terminology of [Rootzén et al. \(1998\)](#) and satisfies the basic assumptions of Section 2 therein. Lemma [C.2](#) entails that, for any sequence (s_n) of integers tending to infinity with $s_n = O(r_n)$, we have

$$\frac{n}{s_n} \text{Var} \left(\frac{1}{n(1 - \tau_n)} \sum_{t=1}^{s_n} [\mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{P}(Y_1 > \xi_{\tau_n})] \right) = O \left(\frac{1}{n(1 - \tau_n)} \right) \rightarrow 0$$

as $n \rightarrow \infty$. This tail array sum thus satisfies the assumptions of Corollary 2.2 of [Rootzén et al. \(1998\)](#), and by choosing $s_n = r_n$ we find that this result implies

$$\frac{1}{n(1 - \tau_n)} \sum_{t=1}^n [\mathbb{1}\{Y_t > \xi_{\tau_n}\} - \mathbb{P}(Y_1 > \xi_{\tau_n})] \xrightarrow{\mathbb{P}} 0. \quad (\text{C.20})$$

Finally

$$\begin{aligned} &\frac{n}{\xi_{\tau_n}^2} \mathbb{E} \left| \left(Y_1 - \xi_{\tau_n} - \frac{u\xi_{\tau_n}}{\sqrt{n(1 - \tau_n)}} \right)^2 [\mathbb{1}\{Y_1 \leq \xi_{\tau_n} + u\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}\} - \mathbb{1}\{Y_1 \leq \xi_{\tau_n}\}] \right| \\ &\leq \frac{u^2}{1 - \tau_n} \left[\mathbb{P}(Y_1 > \xi_{\tau_n} - |u|\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}) - \mathbb{P}(Y_1 > \xi_{\tau_n} + |u|\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}) \right] \\ &= O \left(\frac{\mathbb{P}(Y_1 > \xi_{\tau_n}[1 - |u|/\sqrt{n(1 - \tau_n)})]}{\mathbb{P}(Y_1 > \xi_{\tau_n})} - \frac{\mathbb{P}(Y_1 > \xi_{\tau_n}[1 + |u|/\sqrt{n(1 - \tau_n)})]}{\mathbb{P}(Y_1 > \xi_{\tau_n})} \right) \end{aligned}$$

by asymptotic proportionality of expectiles and quantiles. Proposition B.1.10 in [de Haan & Ferreira \(2006\)](#) implies that for all $u_n \rightarrow \infty$ and $\varepsilon_n \rightarrow 0$, $\mathbb{P}(Y_1 > u_n[1 + \varepsilon_n])/\mathbb{P}(Y_1 > u_n) \rightarrow 1$ as $n \rightarrow \infty$, giving

$$\frac{n}{\xi_{\tau_n}^2} \mathbb{E} \left| \left(Y_1 - \xi_{\tau_n} - \frac{u\xi_{\tau_n}}{\sqrt{n(1 - \tau_n)}} \right)^2 [\mathbb{1}\{Y_1 \leq \xi_{\tau_n} + u\xi_{\tau_n}/\sqrt{n(1 - \tau_n)}\} - \mathbb{1}\{Y_1 \leq \xi_{\tau_n}\}] \right| = o(1). \quad (\text{C.21})$$

Consequently

$$\frac{1}{\xi_{\tau_n}^2} \mathbb{E} \left| \sum_{t=1}^n \left(Y_t - \xi_{\tau_n} - \frac{u \xi_{\tau_n}}{\sqrt{n(1-\tau_n)}} \right)^2 \left[\mathbb{1}\{Y_t \leq \xi_{\tau_n} + u \xi_{\tau_n} / \sqrt{n(1-\tau_n)}\} - \mathbb{1}\{Y_t \leq \xi_{\tau_n}\} \right] \right| = o(1). \quad (\text{C.22})$$

Combining (C.19), (C.20), (C.21) and (C.22) entails

$$T_{3,n}(u) \xrightarrow{\mathbb{P}} 0. \quad (\text{C.23})$$

We finally combine (C.13), (C.14), (C.15), (C.17), (C.18) and (C.23) with Theorem 5 in Knight (1999) to get

$$\begin{aligned} & \sqrt{n(1-\tau_n)} \left(\frac{\tilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \\ & \xrightarrow{d} \arg \min_{u \in \mathbb{R}} \left\{ -u \sqrt{\frac{2\gamma}{1-2\gamma} + 2(\gamma^{-1}-1) \iint_{[1,\infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy} \times Z + \frac{u^2}{2\gamma} \right\} \\ & = \sqrt{\frac{2\gamma^3}{1-2\gamma} + 2\gamma(1-\gamma) \iint_{[1,\infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy} \times Z. \end{aligned}$$

The proof is complete. \square

Under Condition C* we now provide a high-level result that makes it possible to quantify the bias incurred by using convergence (2*). This and Theorem 3.3*, are the key ingredients in the proof of Corollary 3.4*.

Proposition C.3. *Assume that $\mathbb{E}|Y_-| < \infty$ and Condition C* holds with $0 < \gamma < 1$. Assume further that $\tau_n \uparrow 1$ and $n(1-\tau_n) \rightarrow \infty$ as $n \rightarrow \infty$. Suppose also that there is a nondegenerate limiting random pair (Γ, Θ) such that*

$$\sqrt{n(1-\tau_n)} \left(\hat{\gamma}_n - \gamma, \frac{\hat{q}_{\tau_n}}{q_{\tau_n}} - 1 \right) \xrightarrow{d} (\Gamma, \Theta). \quad (\text{C.24})$$

If $\sqrt{n(1-\tau_n)} A((1-\tau_n)^{-1}) \rightarrow \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)} q_{\tau_n}^{-1} \rightarrow \lambda_2 \in \mathbb{R}$ as $n \rightarrow \infty$, then

$$\sqrt{n(1-\tau_n)} \left(\frac{\hat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \xrightarrow{d} m(\gamma)\Gamma + \Theta - \lambda$$

with $m(\gamma) := (1-\gamma)^{-1} - \log(\gamma^{-1}-1)$ and

$$\lambda := \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho} \right) \lambda_1 + \gamma(\gamma^{-1}-1)^\gamma \mathbb{E}(Y) \lambda_2.$$

This proposition makes no assumptions about the dependence within the time series (Y_t) and, unlike the related Theorem 1 in Daouia et al. (2018), does not require the distribution function F to be increasing.

Proof of Proposition C.3. An inspection of the proof of Theorem 1 in [Daouia et al. \(2018\)](#) reveals that (C.24) ensures the desired convergence (in particular, the independence assumption is unnecessary). To see why the assumption that F be increasing can be dropped, apply Proposition 1(i) in [Daouia et al. \(2020\)](#) rather than Corollary 1 in [Daouia et al. \(2018\)](#) to control the asymptotic bias of the estimator. \square

Proof of Theorem 3.3.* We prove this result in two steps: we first consider the case $s_0 = 1$ and we then show the result for an arbitrary $s_0 > 0$.

Step 1: the case $s_0 = 1$. For any $t \geq 1$, let $Z_t = 1/\bar{F}(Y_t)$. Since F is continuous, (Z_t) is clearly a strictly stationary time series whose marginal distribution is standard Pareto, *i.e.* with distribution function $G : z \mapsto 1 - z^{-1}$ on $[1, \infty)$. One readily checks that $(Y_t) \stackrel{d}{=} (U(Z_t))$. Besides, $(\bar{F}(Y_t)) = (1/Z_t) = (\bar{G}(Z_t))$, and in particular the sequence (Z_t) satisfies Conditions A* and B*.

Define $k = k_n = n(1 - \tau_n)$ and apply Theorem 2.1 of [Drees \(2003\)](#) to the sequence (Z_t) : there are versions of the process $(Z_{n-[ks],n})_{s \in (0,1]}$ such that

$$\frac{k}{n} Z_{n-[ks],n} = s^{-1} \left(1 + \frac{1}{\sqrt{k}} s^{-1} W(s) + o_{\mathbb{P}} \left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}} \right) \right) \quad (\text{C.25})$$

uniformly in $s \in (0, 1]$, with W a centred Gaussian process having covariance function r as in the statement of Theorem 3.3*, and an arbitrarily small $\varepsilon > 0$. In particular, the increments $W(t) - W(s)$ are Gaussian centred with variance $\mathbb{E}[(W(t) - W(s))^2] = r(t, t) - 2r(s, t) + r(s, s)$. Therefore, for any integer $p \geq 2$, there is a constant $C_p > 0$ with

$$\forall s, t \in [0, 1], \quad \mathbb{E}[(W(t) - W(s))^{2p}] = C_p (r(t, t) - 2r(s, t) + r(s, s))^p.$$

It is then a straightforward consequence of Condition B(ii)* and B(iii)* that there is another constant $C'_p > 0$ such that

$$\forall s, t \in [0, 1], \quad \mathbb{E}[(W(t) - W(s))^{2p}] \leq C'_p |t - s|^p.$$

Deduce from the Kolmogorov continuity criterion (see *e.g.* Theorem 2.1 p.26 of [Revuz & Yor 2005](#)) that W has a (Hölder) continuous modification \widetilde{W} that satisfies, for any $\varepsilon > 0$, $s^{-1/2+\varepsilon} \widetilde{W}(s) \rightarrow 0$ with probability 1 as $s \rightarrow 0$. This modification \widetilde{W} is indistinguishable from W , in the sense of Definition 1.7 of [Revuz & Yor \(2005, p.19\)](#), because W is necessarily left-continuous as a uniform limit of left-continuous processes from (C.25), and so with probability 1 we can write $W = \widetilde{W}$. We do so in the sequel without further mention. We remark also that (C.25) entails, uniformly in $s \in [1/(2k), 1]$, $ks Z_{n-[ks],n}/n = O_{\mathbb{P}}(k^\varepsilon)$, and especially $Z_{n,n}/n = O_{\mathbb{P}}(k^\varepsilon)$, for any arbitrarily small $\varepsilon > 0$.

For a small fixed $\delta \in (0, 1/4)$, let $s_n(\delta) = k^{-1+2\delta}$. Then $s_n(\delta) \rightarrow 0$ and $[s_n(\delta)]^{-1/2-\delta}/\sqrt{k} = k^{-2\delta^2} \rightarrow 0$ as $n \rightarrow \infty$, which allows the use of Taylor expansions combined with (C.25) for $\varepsilon = \delta$ to get, uniformly in $s \in [s_n(\delta), 1]$,

$$\left(\frac{k}{n}Z_{n-[ks],n}\right)^b = s^{-b} \left(1 + \frac{b}{\sqrt{k}}s^{-1}W(s) + o_{\mathbb{P}}\left(\frac{s^{-1/2-\delta}}{\sqrt{k}}\right)\right) \text{ for any } b \neq 0 \quad (\text{C.26})$$

$$\text{and } \log\left(\frac{k}{n}Z_{n-[ks],n}\right) = -\log s + \frac{1}{\sqrt{k}}s^{-1}W(s) + o_{\mathbb{P}}\left(\frac{s^{-1/2-\delta}}{\sqrt{k}}\right). \quad (\text{C.27})$$

Recall now that $(Y_{n-[ks],n})_{s \in (0,1]} = (U(Z_{n-[ks],n}))_{s \in (0,1]}$ and $q_{1-k/n} = U(n/k)$. Combining (C.26), the uniform inequality on U stated in Theorem 2.3.9 of de Haan & Ferreira (2006, p.48) with (using the notation therein) $t = n/k$ and $x = kZ_{n-[ks],n}/n$ (so that $tx = Z_{n-[ks],n} \geq Z_{n-k,n} \rightarrow \infty$ in probability, by (C.25)), and the convergence $s^{-1/2+\delta}W(s) \rightarrow 0$ as $s \rightarrow 0$, we find, after straightforward calculations,

$$\frac{Y_{n-[ks],n}}{q_{1-k/n}} = s^{-\gamma} \left(1 + \frac{1}{\sqrt{k}}\gamma s^{-1}W(s) + \frac{s^{-\rho} - 1}{\rho}A(n/k) + o_{\mathbb{P}}\left(\frac{s^{-1/2-\delta}}{\sqrt{k}}\right)\right) \text{ uniformly in } s \in [s_n(\delta), 1]. \quad (\text{C.28})$$

After that, note that Condition C* yields

$$\forall x > 0, \lim_{s \rightarrow \infty} \frac{\log((sx)^{-\gamma}U(sx)) - \log(s^{-\gamma}U(s))}{A(s)} = \frac{x^{\rho} - 1}{\rho}. \quad (\text{C.29})$$

Combining (C.27), the uniform inequality on $s \mapsto \log(s^{-\gamma}U(s))$ resulting from Theorem B.2.18 in de Haan & Ferreira (2006, p.383) with (using the notation therein) $t = n/k$ and $x = kZ_{n-[ks],n}/n$ and carrying out the same kind of calculations, we get

$$\log \frac{Y_{n-[ks],n}}{q_{1-k/n}} = -\gamma \log s + \frac{1}{\sqrt{k}}\gamma s^{-1}W(s) + \frac{s^{-\rho} - 1}{\rho}A(n/k) + o_{\mathbb{P}}\left(\frac{s^{-1/2-\delta}}{\sqrt{k}}\right) \text{ uniformly in } s \in [s_n(\delta), 1]. \quad (\text{C.30})$$

To deal with the case $s \in (0, s_n(\delta))$, use the triangle inequality together with the convergence $s^{-1/2+2\delta}W(s) \rightarrow 0$ as $s \rightarrow 0$ and the assumption $\sqrt{k}A(n/k) = O(1)$ to write

$$\begin{aligned} & \sup_{0 < s < s_n(\delta)} s^{\gamma+1/2+2\delta} \left| \sqrt{k} \left[\frac{Y_{n-[ks],n}}{q_{1-k/n}} - s^{-\gamma} \right] - \gamma s^{-\gamma-1}W(s) - s^{-\gamma} \frac{s^{-\rho} - 1}{\rho} \sqrt{k}A(n/k) \right| \\ & \leq \sqrt{k} [s_n(\delta)]^{1/2+2\delta} \left(1 + \sup_{0 < s < s_n(\delta)} s^{\gamma} \frac{Y_{n-[ks],n}}{q_{1-k/n}} \right) + o_{\mathbb{P}}(1). \end{aligned} \quad (\text{C.31})$$

Write then

$$\begin{aligned} \sup_{0 < s < s_n(\delta)} s^\gamma \frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} &\stackrel{d}{=} \sup_{0 < s < s_n(\delta)} s^\gamma \frac{U(Z_{n-\lfloor ks \rfloor, n})}{U(n/k)} \\ &= O_{\mathbb{P}} \left(\sup_{0 < s < s_n(\delta)} \left(\frac{ks}{n} Z_{n-\lfloor ks \rfloor, n} \right)^\gamma \max \left[\left(\frac{k}{n} Z_{n-\lfloor ks \rfloor, n} \right)^{\delta'} , \left(\frac{k}{n} Z_{n-\lfloor ks \rfloor, n} \right)^{-\delta'} \right] \right) \end{aligned}$$

by Potter bounds on the regularly varying function U (see *e.g.* Proposition B.1.9.5 in [de Haan & Ferreira 2006](#)), for an arbitrarily small $\delta' > 0$. Now

$$\sup_{0 < s < s_n(\delta)} \max \left[\left(\frac{k}{n} Z_{n-\lfloor ks \rfloor, n} \right)^{\delta'} , \left(\frac{k}{n} Z_{n-\lfloor ks \rfloor, n} \right)^{-\delta'} \right] \leq \left(\frac{k}{n} Z_{n,n} \right)^{\delta'} + \left(\frac{k}{n} Z_{n-k,n} \right)^{-\delta'} = O_{\mathbb{P}}(k^{2\delta'})$$

and

$$\sup_{0 < s < s_n(\delta)} \left(\frac{ks}{n} Z_{n-\lfloor ks \rfloor, n} \right)^\gamma \leq \left(\frac{Z_{n,n}}{n} \right)^\gamma + \sup_{1/k \leq s \leq 1} \left(\frac{ks}{n} Z_{n-\lfloor ks \rfloor, n} \right)^\gamma = O_{\mathbb{P}}(k^{\delta'}).$$

Thus

$$\sup_{0 < s < s_n(\delta)} s^\gamma \frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} = O_{\mathbb{P}}(k^{3\delta'})$$

for an arbitrarily small $\delta' > 0$. Since $\sqrt{k} [s_n(\delta)]^{1/2+2\delta} = k^{-\delta(1-4\delta)} \rightarrow 0$, choosing $\delta' = \delta(1-4\delta)/4 > 0$ and using [\(C.31\)](#) yields

$$\sup_{0 < s < s_n(\delta)} s^{\gamma+1/2+2\delta} \left| \sqrt{k} \left[\frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} - s^{-\gamma} \right] - \gamma s^{-\gamma-1} W(s) - s^{-\gamma} \frac{s^{-\rho} - 1}{\rho} \sqrt{k} A(n/k) \right| = o_{\mathbb{P}}(1). \quad (\text{C.32})$$

Combining [\(C.28\)](#) and [\(C.32\)](#) results in

$$\begin{aligned} \frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} &= s^{-\gamma} \left(1 + \frac{1}{\sqrt{k}} \gamma s^{-1} W(s) + \frac{s^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{s^{-1/2-2\delta}}{\sqrt{k}} \right) \right) \\ &\text{uniformly in } s \in (0, 1]. \end{aligned} \quad (\text{C.33})$$

A similar argument based on a Potter bound for the (slowly varying) logarithm function gives

$$\sup_{0 < s < s_n(\delta)} s^{1/2+2\delta} \left| \sqrt{k} \left[\log \frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} + \gamma \log s \right] - \gamma s^{-1} W(s) - \frac{s^{-\rho} - 1}{\rho} \sqrt{k} A(n/k) \right| = o_{\mathbb{P}}(1). \quad (\text{C.34})$$

Combining [\(C.30\)](#) and [\(C.34\)](#) results in

$$\begin{aligned} \log \frac{Y_{n-\lfloor ks \rfloor, n}}{q_{1-k/n}} &= -\gamma \log s + \frac{1}{\sqrt{k}} \gamma s^{-1} W(s) + \frac{s^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{s^{-1/2-2\delta}}{\sqrt{k}} \right) \\ &\text{uniformly in } s \in (0, 1]. \end{aligned} \quad (\text{C.35})$$

Equations (C.33) and (C.35) constitute the desired result since $\delta > 0$ is arbitrarily small.

Step 2: the case of an arbitrary s_0 . We need only deal with the case when $s_0 \geq 2$ is a positive integer. Let $k_0 = s_0 k$ and apply the result obtained in Step 1 for $k = k_0$ to get, uniformly in $s \in [0, 1]$,

$$\begin{aligned} \frac{\widehat{q}_{1-k_0s/n}}{q_{1-k_0/n}} &= s^{-\gamma} \left(1 + \frac{1}{\sqrt{k_0}} \gamma s^{-1} W(s) + \frac{s^{-\rho} - 1}{\rho} A(n/k_0) + o_{\mathbb{P}} \left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k_0}} \right) \right) \\ \text{and } \log \frac{\widehat{q}_{1-k_0s/n}}{q_{1-k_0/n}} &= -\gamma \log s + \frac{1}{\sqrt{k_0}} \gamma s^{-1} W(s) + \frac{s^{-\rho} - 1}{\rho} A(n/k_0) + o_{\mathbb{P}} \left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k_0}} \right). \end{aligned}$$

Besides, as Gaussian random processes, $W(s) \stackrel{d}{=} W_0(s) := \sqrt{s_0} W(s/s_0)$, because the covariance structure of W is 1-homogeneous (since all the R_t are 1-homogeneous functions). Since $A(n/k_0)/A(n/k) \rightarrow s_0^{-\rho}$ as $n \rightarrow \infty$ by regular variation of $|A|$, we obtain

$$\frac{\widehat{q}_{1-k_0s/n}}{q_{1-k_0/n}} = s^{-\gamma} \left(1 + \frac{1}{\sqrt{k}} \gamma (s_0 s)^{-1} W_0(s_0 s) + \frac{(s_0 s)^{-\rho} - s_0^{-\rho}}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{(s_0 s)^{-1/2-\varepsilon}}{\sqrt{k}} \right) \right)$$

and

$$\log \frac{\widehat{q}_{1-k_0s/n}}{q_{1-k_0/n}} = -\gamma \log s + \frac{1}{\sqrt{k}} \gamma (s_0 s)^{-1} W_0(s_0 s) + \frac{(s_0 s)^{-\rho} - s_0^{-\rho}}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{(s_0 s)^{-1/2-\varepsilon}}{\sqrt{k}} \right).$$

Now, by Condition C* and its version (C.29) on $\log U$,

$$\begin{aligned} \frac{q_{1-k_0/n}}{q_{1-k/n}} &= s_0^{-\gamma} \left(1 + \frac{s_0^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right) \right) \\ \text{and } \log \frac{q_{1-k_0/n}}{q_{1-k/n}} &= -\gamma \log s_0 + \frac{s_0^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right). \end{aligned}$$

Conclude that

$$\frac{\widehat{q}_{1-k_0s/n}}{q_{1-k/n}} = (s_0 s)^{-\gamma} \left(1 + \frac{1}{\sqrt{k}} \gamma (s_0 s)^{-1} W_0(s_0 s) + \frac{(s_0 s)^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{(s_0 s)^{-1/2-\varepsilon}}{\sqrt{k}} \right) \right)$$

and

$$\log \frac{\widehat{q}_{1-k_0s/n}}{q_{1-k/n}} = -\gamma \log(s_0 s) + \frac{1}{\sqrt{k}} \gamma (s_0 s)^{-1} W_0(s_0 s) + \frac{(s_0 s)^{-\rho} - 1}{\rho} A(n/k) + o_{\mathbb{P}} \left(\frac{(s_0 s)^{-1/2-\varepsilon}}{\sqrt{k}} \right).$$

The proof is then completed by noting that $\widehat{q}_{1-k_0s/n} = \widehat{q}_{1-ks'/n}$ with $s' = s_0 s \in [0, s_0]$, and all the identities above, written in terms of s' , are uniform in $s' \in [0, s_0]$. \square

Proof of Corollary 3.4.* Remark first that

$$\begin{aligned} \widehat{\gamma}_n^H &= \int_0^1 \log \left(\frac{\widehat{q}_{1-\lfloor n(1-\tau_n) \rfloor s/n}}{q_{1-\lfloor n(1-\tau_n) \rfloor /n}} \right) ds - \log \left(\frac{\widehat{q}_{1-\lfloor n(1-\tau_n) \rfloor /n}}{q_{1-\lfloor n(1-\tau_n) \rfloor /n}} \right) \\ \text{and } \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} &= \frac{\widehat{q}_{1-\lfloor n(1-\tau_n) \rfloor /n}}{q_{1-\lfloor n(1-\tau_n) \rfloor /n}} \times \frac{q_{1-\lfloor n(1-\tau_n) \rfloor /n}}{q_{\tau_n}} = \frac{\widehat{q}_{1-\lfloor n(1-\tau_n) \rfloor /n}}{q_{1-\lfloor n(1-\tau_n) \rfloor /n}} \left(1 + o \left(\frac{1}{\sqrt{n(1-\tau_n)}} \right) \right). \end{aligned}$$

The latter holds by local uniformity of Condition C* and the regular variation property of $|A|$, see Theorem 2.3.9 in [de Haan & Ferreira \(2006\)](#). It is thus enough to tackle the case when $n(1 - \tau_n)$ is a sequence of integers. In this case, one has, from Theorem 3.3* and for appropriate versions of the processes involved,

$$\begin{aligned}\sqrt{n(1 - \tau_n)}(\widehat{\gamma}_n^H - \gamma) &= \int_0^1 \log\left(\frac{\widehat{q}_{1-(1-\tau_n)s}}{q_{\tau_n}}\right) ds - \log\left(\frac{\widehat{q}_{\tau_n}}{q_{\tau_n}}\right) \\ &= \gamma \int_0^1 [s^{-1}W(s) - W(1)] ds + \frac{\lambda_1}{1 - \rho} + o_{\mathbb{P}}(1)\end{aligned}$$

$$\text{and } \sqrt{n(1 - \tau_n)}\left(\frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} - 1\right) = \gamma W(1) + o_{\mathbb{P}}(1)$$

where W is a centred Gaussian process having covariance function $r(x, y) = \min(x, y) + \sum_{t=1}^{\infty} R_t(x, y) + R_t(y, x)$. It follows immediately that the pair

$$\sqrt{n(1 - \tau_n)}\left(\widehat{\gamma}_n^H - \gamma, \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} - 1\right)$$

is asymptotically Gaussian, with expectation $(\lambda_1/(1 - \rho), 0)^\top$ and covariance matrix

$$\gamma^2 \begin{pmatrix} \iint_{[0,1]^2} \left[\frac{r(s,t)}{st} - \frac{r(s,1)}{s} - \frac{r(1,t)}{t} + r(1,1) \right] ds dt & \int_0^1 \left[\frac{r(s,1)}{s} - r(1,1) \right] ds \\ \int_0^1 \left[\frac{r(s,1)}{s} - r(1,1) \right] ds & r(1,1) \end{pmatrix}.$$

Use the definition of the covariance function r and the identity

$$\iint_{[0,1]^2} \left[\frac{r(s,t)}{st} - \frac{r(s,1)}{s} - \frac{r(1,t)}{t} \right] ds dt = 0$$

valid because of the 1-homogeneity of each R_t (see the final stages of the proof of Theorem 4 in [Stupfler 2019](#)) to get the announced covariance matrix. The convergence of $\widehat{\xi}_{\tau_n}$ is then a direct consequence of Proposition [C.3](#). \square

The following lemma gives a joint convergence result on the intermediate empirical (LAWS) expectile and its intermediate empirical quantile counterpart, in our time series framework. It is the key ingredient in the proof of Theorem [C.5](#). To state this result, recall the notation $\widehat{F}_n(u) = n^{-1} \sum_{t=1}^n \mathbb{1}\{Y_t > u\}$ for the empirical survival function of the observations.

Lemma C.4. *Assume that Conditions A*, B* and C* are satisfied, and that there is $\delta > 0$ such that $\mathbb{E}|Y_-|^{2+\delta} < \infty$, $0 < \gamma < 1/(2 + \delta)$ and $\sum_{l \geq 1} [\beta(l)]^{\delta/(2+\delta)} < \infty$. Let $\tau_n \uparrow 1$ be such that $n(1 - \tau_n) \rightarrow \infty$,*

$r_n(1-\tau_n) \rightarrow 0$, $r_n(r_n/\sqrt{n(1-\tau_n)})^\delta \rightarrow 0$, $r_n \log^2(n(1-\tau_n))/\sqrt{n(1-\tau_n)} \rightarrow 0$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \rightarrow \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \rightarrow \lambda_2 \in \mathbb{R}$ as $n \rightarrow \infty$. Then

$$\sqrt{n(1-\tau_n)} \left(\frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} - 1, \frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1, \frac{\widehat{F}_n(\widetilde{\xi}_{\tau_n})}{1-\tau_n} - (\gamma^{-1} - 1) \right) \xrightarrow{d} (\Theta, \Phi, \Xi).$$

Here (Θ, Φ, Ξ) is a trivariate Gaussian random vector defined as

$$\Theta = \gamma W(1), \quad \Phi = \gamma^2(\gamma^{-1} - 1)^\gamma \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) ds$$

$$\text{and } \Xi = W(\gamma^{-1} - 1) - \gamma(\gamma^{-1} - 1)^{\gamma+1} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) ds - \lambda_1 \frac{(\gamma^{-1} - 1)^{1-\rho}}{\gamma(1-\gamma-\rho)} - \lambda_2(\gamma^{-1} - 1)^{\gamma+1} \mathbb{E}(Y)$$

where W is a centred Gaussian process with covariance function $r(x, y) = \min(x, y) + \sum_{t=1}^{\infty} R_t(x, y) + R_t(y, x)$.

Proof of Lemma C.4. Assume to start with that $k = n(1 - \tau_n)$ is a sequence of integers. The estimating equation for an expectile, obtained by differentiating the right-hand side of (1*) and setting it to 0, can also be written for its sample version: this results in

$$\xi_{1-k/n} - \mathbb{E}(Y) = \frac{n}{k} \left(1 - \frac{2k}{n} \right) \mathbb{E}((Y - \xi_{1-k/n}) \mathbb{1}\{Y > \xi_{1-k/n}\}) \quad (\text{C.36})$$

$$\text{and } \widetilde{\xi}_{1-k/n} - \bar{Y}_n = \frac{n}{k} \left(1 - \frac{2k}{n} \right) \times \frac{1}{n} \sum_{t=1}^n (Y_t - \widetilde{\xi}_{1-k/n}) \mathbb{1}\{Y_t > \widetilde{\xi}_{1-k/n}\}. \quad (\text{C.37})$$

[Here $\bar{Y}_n = n^{-1} \sum_{i=1}^n Y_i$ denotes the sample mean.] We now notice that

$$\begin{aligned} \mathbb{E}((Y - \xi_{1-k/n}) \mathbb{1}\{Y > \xi_{1-k/n}\}) &= \int_{\xi_{1-k/n}}^{\infty} \bar{F}(u) du \\ \text{and } \frac{1}{n} \sum_{t=1}^n (Y_t - \widetilde{\xi}_{1-k/n}) \mathbb{1}\{Y_t > \widetilde{\xi}_{1-k/n}\} &= \int_{1-\widehat{F}_n(\widetilde{\xi}_{1-k/n})}^1 \widehat{q}_w dw - \widetilde{\xi}_{1-k/n} \widehat{F}_n(\widetilde{\xi}_{1-k/n}). \end{aligned}$$

Subtracting (C.36) from (C.37) and dividing through by $\xi_{1-k/n}$ then produces

$$\begin{aligned} &\left(\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1 \right) - \frac{\bar{Y}_n - \mathbb{E}(Y)}{\xi_{1-k/n}} \\ &= \frac{n}{k} \left(1 - \frac{2k}{n} \right) \frac{1}{\xi_{1-k/n}} \left(\int_{1-\widehat{F}_n(\widetilde{\xi}_{1-k/n})}^1 \widehat{q}_w dw - \widetilde{\xi}_{1-k/n} \widehat{F}_n(\widetilde{\xi}_{1-k/n}) - \int_{\xi_{1-k/n}}^{\infty} \bar{F}(u) du \right). \end{aligned} \quad (\text{C.38})$$

Our goal will be to control or evaluate each term in Equation (C.38) except the first. As Y has a heavy-tailed distribution with tail index $\gamma < 1$,

$$1 - \frac{2k}{n} = 1 + o\left(\frac{1}{q_{1-k/n}}\right) = 1 + o\left(\frac{1}{\sqrt{k}}\right). \quad (\text{C.39})$$

Note also that, for any $n \geq 2$,

$$n \operatorname{Var}(\bar{Y}_n - \mathbb{E}(Y)) = \operatorname{Var}(Y) + 2 \sum_{j=1}^{n-1} \frac{n-j}{n} \mathbb{E}([Y_1 - \mathbb{E}(Y)][Y_{j+1} - \mathbb{E}(Y)]).$$

By Theorem 0 in the Introduction of [Bradley \(1985\)](#) applied to the sequence $(Y_t - \mathbb{E}(Y))$, the sum $\sum_{j \geq 1} |\mathbb{E}([Y_1 - \mathbb{E}(Y)][Y_{j+1} - \mathbb{E}(Y)])|$ converges. The dominated convergence theorem then entails $n \operatorname{Var}(\bar{Y}_n - \mathbb{E}(Y)) = O(1)$ and therefore, by the Chebyshev's inequality,

$$\frac{\bar{Y}_n - \mathbb{E}(Y)}{\xi_{1-k/n}} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) = o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right). \quad (\text{C.40})$$

A standard calculation based on Condition C* yields

$$\frac{1}{\xi_{1-k/n}} \int_{\xi_{1-k/n}}^{\infty} \bar{F}(u) \, du = \bar{F}(\xi_{1-k/n}) \left(\frac{\gamma}{1-\gamma} + A \left(\frac{1}{\bar{F}(\xi_{1-k/n})} \right) \frac{1+o(1)}{(1-\gamma)(1-\gamma-\rho)} \right);$$

see also (B.4) in the proof of Proposition 1 of [Daouia et al. \(2018\)](#). In fact, using this result together with the regular variation property of $|A|$, one finds

$$\frac{n}{k} \frac{1}{\xi_{1-k/n}} \int_{\xi_{1-k/n}}^{\infty} \bar{F}(u) \, du = 1 - \frac{\lambda_2}{\sqrt{k}} (\gamma^{-1} - 1)^{\gamma} \mathbb{E}(Y) + o\left(\frac{1}{\sqrt{k}}\right). \quad (\text{C.41})$$

To control the other two terms appearing on the right-hand side of Equation (C.38), we first note that for appropriate versions of the tail empirical quantile process $s \mapsto \hat{q}_{1-ks/n}$, Theorem 3.3* entails that

$$\sqrt{k} \left(\frac{\hat{q}_{1-ks/n}}{q_{1-k/n}} - s^{-\gamma} \right) \xrightarrow{\mathbb{P}} \gamma s^{-\gamma-1} W(s) + \lambda_1 s^{-\gamma} \frac{s^{-\rho} - 1}{\rho} \quad (\text{C.42})$$

uniformly on any compact interval $[s_0, s_1] \subset (0, \infty)$, where W is a continuous, centred Gaussian process W on $[0, 1]$ having covariance function $r(x, y) = \min(x, y) + \sum_{t=1}^{\infty} \{R_t(x, y) + R_t(y, x)\}$. By Vervaat's lemma (see *e.g.* Lemma A.0.2 on p.357 of [de Haan & Ferreira 2006](#)), we get

$$\sqrt{k} \left(\frac{n \hat{F}_n(x q_{1-k/n})}{k} - x^{-1/\gamma} \right) \xrightarrow{\mathbb{P}} W(x^{-1/\gamma}) + \lambda_1 x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\gamma \rho}$$

uniformly on any compact interval $[x_0, x_1] \subset (0, \infty)$. Now, by Theorem 3.1* and convergence (2*),

$$\frac{\tilde{\xi}_{1-k/n}}{q_{1-k/n}} = \frac{\tilde{\xi}_{1-k/n}}{\xi_{1-k/n}} \times \frac{\xi_{1-k/n}}{q_{1-k/n}} \xrightarrow{\mathbb{P}} (\gamma^{-1} - 1)^{-\gamma}.$$

Taking $x = \tilde{\xi}_{1-k/n}/q_{1-k/n}$ in the above convergence then yields

$$\sqrt{k} \left(\frac{n \hat{F}_n(\tilde{\xi}_{1-k/n})}{k} - \left[\frac{\tilde{\xi}_{1-k/n}}{q_{1-k/n}} \right]^{-1/\gamma} \right) \xrightarrow{\mathbb{P}} W(\gamma^{-1} - 1) + \lambda_1 (\gamma^{-1} - 1) \frac{(\gamma^{-1} - 1)^{-\rho} - 1}{\gamma \rho}. \quad (\text{C.43})$$

In particular, by Theorem 3.1* and Proposition 1(i) in [Daouia et al. \(2020\)](#), we have the expansion $(n/k) \times \widehat{F}_n(\widetilde{\xi}_{1-k/n}) = (\gamma^{-1} - 1) + \mathcal{O}_{\mathbb{P}}(1/\sqrt{k})$. By Theorem 3.1* again, this entails

$$\begin{aligned} \frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} \times \frac{n}{k} \widehat{F}_n(\widetilde{\xi}_{1-k/n}) &= (\gamma^{-1} - 1) + (\gamma^{-1} - 1) \left(\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1 \right) \\ &\quad + \left(\frac{n}{k} \widehat{F}_n(\widetilde{\xi}_{1-k/n}) - (\gamma^{-1} - 1) \right) + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right). \end{aligned} \quad (\text{C.44})$$

Lastly, by a change of variables,

$$\frac{n}{k} \frac{1}{\xi_{1-k/n}} \int_{1-\widehat{F}_n(\widetilde{\xi}_{1-k/n})}^1 \widehat{q}_w \, dw = \frac{q_{1-k/n}}{\xi_{1-k/n}} \left(\int_0^{\gamma^{-1}-1} \frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \, ds + \int_{\gamma^{-1}-1}^{(n/k)\widehat{F}_n(\widetilde{\xi}_{1-k/n})} \frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \, ds \right).$$

By Theorem 3.3*, the expansion $(n/k) \times \widehat{F}_n(\widetilde{\xi}_{1-k/n}) = (\gamma^{-1} - 1) + \mathcal{O}_{\mathbb{P}}(1/\sqrt{k})$, a Taylor expansion and the continuity of the process W ,

$$\begin{aligned} \int_{\gamma^{-1}-1}^{(n/k)\widehat{F}_n(\widetilde{\xi}_{1-k/n})} \frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \, ds &= \int_{\gamma^{-1}-1}^{(n/k)\widehat{F}_n(\widetilde{\xi}_{1-k/n})} s^{-\gamma} \, ds + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right) \\ &= (\gamma^{-1} - 1)^{-\gamma} \left(\frac{n}{k} \widehat{F}_n(\widetilde{\xi}_{1-k/n}) - (\gamma^{-1} - 1) \right) + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right). \end{aligned}$$

By Theorem 3.3*, moreover,

$$\begin{aligned} \int_0^{\gamma^{-1}-1} \frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \, ds &= \frac{(\gamma^{-1} - 1)^{-\gamma}}{\gamma} + \frac{\gamma}{\sqrt{k}} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) \, ds \\ &\quad + \frac{\lambda_1 (\gamma^{-1} - 1)^{-\gamma}}{\sqrt{k} \rho} \left(\frac{(\gamma^{-1} - 1)^{1-\rho}}{1 - \gamma - \rho} - \frac{1}{\gamma} \right) + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right). \end{aligned}$$

Another use of Proposition 1(i) in [Daouia et al. \(2020\)](#) and straightforward calculations give

$$\begin{aligned} \frac{n}{k} \frac{1}{\xi_{1-k/n}} \int_{1-\widehat{F}_n(\widetilde{\xi}_{1-k/n})}^1 \widehat{q}_w \, dw &= \frac{1}{\gamma} + \frac{\gamma(\gamma^{-1} - 1)^{\gamma}}{\sqrt{k}} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) \, ds \\ &\quad + \left(\frac{n}{k} \widehat{F}_n(\widetilde{\xi}_{1-k/n}) - (\gamma^{-1} - 1) \right) - \frac{\lambda_2}{\sqrt{k}} (\gamma^{-1} - 1)^{\gamma} \mathbb{E}(Y) \\ &\quad + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right). \end{aligned} \quad (\text{C.45})$$

Combining (C.38), (C.39), (C.40), (C.41), (C.44) and (C.45) thus yields

$$\sqrt{k} \left(\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1 \right) = \gamma^2 (\gamma^{-1} - 1)^{\gamma} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) \, ds + \mathcal{O}_{\mathbb{P}}(1). \quad (\text{C.46})$$

A combination of Equation (C.46) with Proposition 1(i) in Daouia et al. (2020) then gives

$$\begin{aligned} \sqrt{k} \left(\left[\frac{\tilde{\xi}_{1-k/n}}{q_{1-k/n}} \right]^{-1/\gamma} - (\gamma^{-1} - 1) \right) &= -\gamma(\gamma^{-1} - 1)^{\gamma+1} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) ds \\ &\quad - \lambda_1 \gamma^{-1} (\gamma^{-1} - 1) \left(\frac{(\gamma^{-1} - 1)^{-\rho}}{1 - \gamma - \rho} + \frac{(\gamma^{-1} - 1)^{-\rho} - 1}{\rho} \right) \\ &\quad - \lambda_2 (\gamma^{-1} - 1)^{\gamma+1} \mathbb{E}(Y) + o_{\mathbb{P}}(1). \end{aligned}$$

Reporting this identity in (C.43) gives

$$\begin{aligned} \sqrt{k} \left(\frac{n}{k} \widehat{F}_n(\tilde{\xi}_{1-k/n}) - (\gamma^{-1} - 1) \right) &= W(\gamma^{-1} - 1) - \gamma(\gamma^{-1} - 1)^{\gamma+1} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) ds \\ &\quad - \lambda_1 \frac{(\gamma^{-1} - 1)^{1-\rho}}{\gamma(1 - \gamma - \rho)} - \lambda_2 (\gamma^{-1} - 1)^{\gamma+1} \mathbb{E}(Y) + o_{\mathbb{P}}(1). \end{aligned} \quad (\text{C.47})$$

Combining Equations (C.46) and (C.47) with (C.42) for $s = 1$ completes the proof when $k = n(1 - \tau_n)$ is a sequence of integers. The general case follows because, by the monotonicity of quantiles and expectiles,

$$\begin{aligned} \frac{\widehat{q}_{1-k/n}}{q_{1-k/n}} - 1 &\leq \frac{\widehat{q}_{1-\lfloor k \rfloor/n}}{q_{1-\lfloor k \rfloor/n}} \left(\frac{q_{1-\lfloor k \rfloor/n}}{q_{1-k/n}} - 1 \right) + \frac{\widehat{q}_{1-\lfloor k \rfloor/n}}{q_{1-\lfloor k \rfloor/n}} - 1, \\ \frac{\tilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1 &\leq \frac{\tilde{\xi}_{1-\lfloor k \rfloor/n}}{\xi_{1-\lfloor k \rfloor/n}} \left(\frac{\xi_{1-\lfloor k \rfloor/n}}{\xi_{1-k/n}} - 1 \right) + \frac{\tilde{\xi}_{1-\lfloor k \rfloor/n}}{\xi_{1-\lfloor k \rfloor/n}} - 1, \end{aligned}$$

with similar lower bounds applying with the ceiling function in place of the floor function. Condition C* and Proposition 1(i) in Daouia et al. (2020) now yield

$$\frac{q_{1-\lfloor k \rfloor/n}}{q_{1-k/n}} - 1 = o\left(\frac{1}{\sqrt{k}}\right) \quad \text{and} \quad \frac{\xi_{1-\lfloor k \rfloor/n}}{\xi_{1-k/n}} - 1 = o\left(\frac{1}{\sqrt{k}}\right).$$

Retracing the steps of the proof when $k = n(1 - \tau_n)$ is a sequence of integers (in particular, applying Theorem 3.3* with $\lfloor k \rfloor$ in place of k to obtain analogues of (C.42) and (C.43)) and using the continuity of the process W , shows that the quantities

$$\frac{\widehat{q}_{1-\lfloor k \rfloor/n}}{q_{1-\lfloor k \rfloor/n}} - 1 \quad \text{and} \quad \frac{\widehat{q}_{1-\lceil k \rceil/n}}{q_{1-\lceil k \rceil/n}} - 1$$

on the one hand, and

$$\frac{\tilde{\xi}_{1-\lfloor k \rfloor/n}}{\xi_{1-\lfloor k \rfloor/n}} - 1 \quad \text{and} \quad \frac{\tilde{\xi}_{1-\lceil k \rceil/n}}{\xi_{1-\lceil k \rceil/n}} - 1$$

on the other hand, have the same limit in probability, and one can conclude the proof using the sandwich lemma. We omit the details. \square

Our next main result derives the joint weak convergence of the estimator $\widehat{\gamma}_n^E$ with \widehat{q}_{τ_n} .

Theorem C.5. *Assume that Conditions A^* , B^* and C^* are satisfied, and that there is $\delta > 0$ such that $\mathbb{E}|Y_-|^{2+\delta} < \infty$, $0 < \gamma < 1/(2 + \delta)$ and $\sum_{l \geq 1} [\beta(l)]^{\delta/(2+\delta)} < \infty$. Let $\tau_n \uparrow 1$ be such that $n(1 - \tau_n) \rightarrow \infty$, $r_n(1 - \tau_n) \rightarrow 0$, $r_n(r_n/\sqrt{n(1 - \tau_n)})^\delta \rightarrow 0$, $r_n \log^2(n(1 - \tau_n))/\sqrt{n(1 - \tau_n)} \rightarrow 0$, $\sqrt{n(1 - \tau_n)}A((1 - \tau_n)^{-1}) \rightarrow \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1 - \tau_n)}q_{\tau_n}^{-1} \rightarrow \lambda_2 \in \mathbb{R}$ as $n \rightarrow \infty$. Then*

$$\sqrt{n(1 - \tau_n)} \left(\hat{\gamma}_n^E - \gamma, \frac{\hat{q}_{\tau_n}}{q_{\tau_n}} - 1 \right) \xrightarrow{d} (\Gamma, \Theta).$$

Here (Γ, Θ) is a bivariate Gaussian random vector defined as

$$\Gamma = \gamma^3(\gamma^{-1} - 1)^{\gamma+1} \int_0^{\gamma^{-1}-1} s^{-\gamma-1} W(s) ds - \gamma^2 W(\gamma^{-1} - 1) + \lambda_1 \gamma \frac{(\gamma^{-1} - 1)^{1-\rho}}{1 - \gamma - \rho} + \lambda_2 \gamma^2 (\gamma^{-1} - 1)^{\gamma+1} \mathbb{E}(Y)$$

and $\Theta = \gamma W(1)$, where W is a centred Gaussian process with covariance function $r(x, y) = \min(x, y) + \sum_{t=1}^{\infty} R_t(x, y) + R_t(y, x)$.

Following Proposition C.3, the asymptotic Gaussian distribution of the indirect QB estimator with $\hat{\gamma}_n = \hat{\gamma}_n^E$ can also be derived. The expression of the covariance structure of the limiting random pair in Theorem C.5 is in general quite involved, however, so we do not pursue this here for the sake of simplicity. Let us mention though that in the independent case, the Gaussian process W is just a standard Brownian motion, and therefore the asymptotic variance of $\hat{\gamma}_n^E$ can be exactly calculated as $\gamma^3(1 - \gamma)/(1 - 2\gamma)$. As such, for $\gamma < (3 - \sqrt{5})/2 \approx 0.38$, this estimator will tend to be less variable than the Hill estimator, making it a valuable device in the construction of approximate confidence intervals.

Proof of Theorem C.5. Start by writing

$$\hat{\gamma}_n^E - \gamma = -\gamma \left(1 + \frac{\widehat{F}_n(\tilde{\xi}_{\tau_n})}{1 - \tau_n} \right)^{-1} \left(\frac{\widehat{F}_n(\tilde{\xi}_{\tau_n})}{1 - \tau_n} - (\gamma^{-1} - 1) \right).$$

It follows from Lemma C.4 that $\widehat{F}_n(\tilde{\xi}_{\tau_n})/(1 - \tau_n) \xrightarrow{\mathbb{P}} \gamma^{-1} - 1$. Consequently

$$\hat{\gamma}_n^E - \gamma = -\gamma^2 \left(\frac{\widehat{F}_n(\tilde{\xi}_{\tau_n})}{1 - \tau_n} - (\gamma^{-1} - 1) \right) (1 + o_{\mathbb{P}}(1)).$$

The result is now a straightforward consequence of Lemma C.4. □

Our next result gives high-level conditions for the convergence of $\bar{\xi}_{\tau_n}^*$. It is the key to the proof of Theorem 3.5*.

Proposition C.6. *Assume that $\mathbb{E}|Y_-| < \infty$ and Condition C* holds with $0 < \gamma < 1$ and $\rho < 0$. Assume further that $\tau_n, \tau'_n \uparrow 1$ with $n(1-\tau_n) \rightarrow \infty$, $n(1-\tau'_n) \rightarrow c \in [0, \infty)$ and $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau'_n)] \rightarrow \infty$ as $n \rightarrow \infty$. Suppose also that there are nondegenerate limiting random variables Γ, Δ such that*

$$\sqrt{n(1-\tau_n)}(\widehat{\gamma}_n - \gamma) \xrightarrow{d} \Gamma \quad \text{and} \quad \sqrt{n(1-\tau_n)} \left(\frac{\bar{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \xrightarrow{d} \Delta.$$

If moreover $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \rightarrow \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \rightarrow \lambda_2 \in \mathbb{R}$ as $n \rightarrow \infty$, then

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau'_n)]} \left(\frac{\bar{\xi}_{\tau'_n}^*}{\xi_{\tau'_n}} - 1 \right) \xrightarrow{d} \Gamma.$$

Proposition C.6 extends Theorem 3 in Daouia et al. (2018) by dropping the unnecessary assumption of an increasing function F and by being stated outside the i.i.d. case.

Proof of Proposition C.6. The key step is to write

$$\log \left(\frac{\bar{\xi}_{\tau'_n}^*}{\xi_{\tau'_n}} \right) = (\widehat{\gamma}_n - \gamma) \log \left(\frac{1-\tau_n}{1-\tau'_n} \right) + \log \left(\frac{\bar{\xi}_{\tau_n}}{\xi_{\tau_n}} \right) - \log \left(\left[\frac{1-\tau'_n}{1-\tau_n} \right]^\gamma \frac{\xi_{\tau'_n}}{\xi_{\tau_n}} \right).$$

One may then follow the steps of the proof of Theorem 5 in Daouia et al. (2020) with the appropriate straightforward changes, noting that the i.i.d. assumption is in fact never necessary. \square

Proof of Theorem 3.5.* This is a direct corollary of Proposition C.6. \square

The following lemma, of independent interest, is the key to the proof of Theorem 4.1*. It provides a general result on the convergence of array sums in a strong mixing framework.

Lemma C.7. *Let $(X_{n,i})_{1 \leq i \leq n}$ be a triangular, row-strictly stationary array of random variables. Let, for $1 \leq i < j \leq n$, $\mathcal{F}_{n,i,j} = \sigma(X_{n,i}, X_{n,i+1}, \dots, X_{n,j})$ and define a strong mixing coefficient of this array by*

$$\forall n \geq 2, \forall l < n, \alpha_n(l) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, A \in \mathcal{F}_{n,1,k}, B \in \mathcal{F}_{n,k+l,n}, 1 \leq k \leq n-l\}.$$

Suppose that there are sequences of integers (l_n) and (r_n) such that

$$l_n \rightarrow \infty, r_n \rightarrow \infty, \frac{l_n}{r_n} \rightarrow 0, \frac{r_n}{n} \rightarrow 0 \quad \text{and} \quad \frac{n \alpha_n(l_n)}{r_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Define also

$$S_n = \sum_{i=1}^n X_{n,i} = \sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j} + R'_n, \quad \text{where } Z_{n,j} = B_{n,j} + R_{n,j},$$

$$\text{with } B_{n,j} = \sum_{i=(j-1)r_n+1}^{jr_n-l_n} X_{n,i}, \quad R_{n,j} = \sum_{i=jr_n-l_n+1}^{jr_n} X_{n,i} \quad \text{and} \quad R'_n = \sum_{i=r_n \lfloor n/r_n \rfloor + 1}^n X_{n,i}.$$

(i) Assume here that $\sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j}^* \xrightarrow{\mathbb{P}} 0$ and $R'_n \xrightarrow{\mathbb{P}} 0$, where the $R_{n,j}^*$ are i.i.d. copies of $R_{n,1}$. Then S_n has the same limiting behaviour in distribution (if any) as $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^*$, where the $Z_{n,j}^*$ are i.i.d. copies of $Z_{n,1}$.

(ii) If one has $\mathbb{E}(X_{n,1}) = 0$ and

$$\begin{aligned} \frac{n}{r_n} \text{Var}(R_{n,1}) &= \frac{n}{r_n} \text{Var} \left(\sum_{i=1}^{\lfloor n/r_n \rfloor} X_{n,i} \right) \rightarrow 0, \quad \text{Var}(R'_n) = \text{Var} \left(\sum_{i=1}^{n-r_n \lfloor n/r_n \rfloor} X_{n,i} \right) \rightarrow 0, \\ \frac{n}{r_n} \text{Var}(Z_{n,1}) &= \frac{n}{r_n} \text{Var} \left(\sum_{i=1}^{r_n} X_{n,i} \right) \rightarrow \sigma^2 < \infty \quad \text{and} \quad \forall \varepsilon > 0, \quad \frac{n}{r_n} \mathbb{E}(Z_{n,1}^2 \mathbb{1}\{|Z_{n,1}| > \varepsilon\}) \rightarrow 0, \end{aligned}$$

then $S_n \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ (this should be read $S_n \xrightarrow{\mathbb{P}} 0$ if $\sigma^2 = 0$, in which case the final Lindeberg condition above is unnecessary).

Proof of Lemma C.7. (i) By assumption on R'_n , we need only show that $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}$ and $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^*$ have the same limiting behaviour in distribution. We then compare the Fourier transforms of $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}$ and $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^*$ using the big-block/small-block structure of $Z_{n,j} = B_{n,j} + R_{n,j}$ and the mixing assumption to show that the original $Z_{n,j}$ can be considered to be asymptotically independent. This argument and the result of the lemma are inspired by the proof of Lemma 2.1 of [Leadbetter & Rootzén \(1993\)](#).

First introduce the alternative strong mixing coefficient defined for any $n \geq 2$ and $l < n$ by

$$\begin{aligned} a_n(l) &= \sup\{|\text{Cov}(U, V)|, U \text{ is complex-valued and } \mathcal{F}_{n,1,k}\text{-measurable,} \\ &\quad V \text{ is complex-valued and } \mathcal{F}_{n,k+l,n}\text{-measurable, } |U|, |V| \leq 1, \text{ for } 1 \leq k \leq n-l\}. \end{aligned}$$

By Lemma 3 p.10 in [Doukhan \(1994\)](#) (see also Lemma 1.2 of [Ibragimov 1962](#)) there is a universal constant $C > 0$ such that $a_n(l) \leq C\alpha_n(l)$. Therefore, for any n , and using proof by induction,

$$|\mathbb{E}(U_1 U_2 \cdots U_{\lfloor n/r_n \rfloor}) - \mathbb{E}(U_1) \mathbb{E}(U_2) \cdots \mathbb{E}(U_{\lfloor n/r_n \rfloor})| \leq \left(\left\lfloor \frac{n}{r_n} \right\rfloor - 1 \right) a_n(l_n) \leq C \left\lfloor \frac{n}{r_n} \right\rfloor \alpha_n(l_n)$$

for all complex-valued random variables U_k such that $|U_k| \leq 1$ and which are respectively \mathcal{F}_{n,i_k,j_k} -measurable, with $i_1 < j_1 < i_2 < j_2 < \cdots < i_{\lfloor n/r_n \rfloor} < j_{\lfloor n/r_n \rfloor}$ and $i_k - j_{k-1} > l_n$. Conclude, first,

that for any $t \in \mathbb{R}$,

$$\begin{aligned}
& \left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j} \right] \right) - \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j}^* \right] \right) \right| \\
&= \left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j} \right] \right) - \prod_{j=1}^{\lfloor n/r_n \rfloor} \mathbb{E}(\exp[itR_{n,j}^*]) \right| \quad (\text{by independence}) \\
&= \left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j} \right] \right) - \prod_{j=1}^{\lfloor n/r_n \rfloor} \mathbb{E}(\exp[itR_{n,j}]) \right| \quad (\text{by strict stationarity}) \\
&\leq C \left\lfloor \frac{n}{r_n} \right\rfloor \alpha_n(l_n) \rightarrow 0
\end{aligned}$$

for n large enough, since each of the $R_{n,j}$ is $\mathcal{F}_{n,jr_n-l_n+1,jr_n}$ -measurable and $l_n/r_n \rightarrow 0$. Thus $\sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j} \xrightarrow{\mathbb{P}} 0$. It follows, by uniform integrability of the (bounded) sequence $\exp(it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j})$, that

$$\left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j} \right] \right) - \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j} \right] \right) \right| \leq \mathbb{E} \left| 1 - \exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j} \right] \right| \rightarrow 0.$$

Thus $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}$ has the same asymptotic behaviour in distribution as $\sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}$. Adapting the above argument comparing the Fourier transforms of the $R_{n,j}$ and $R_{n,j}^*$ shows that

$$\left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j} \right] \right) - \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}^* \right] \right) \right| \leq C \left\lfloor \frac{n}{r_n} \right\rfloor \alpha_n(l_n) \rightarrow 0,$$

where the $B_{n,j}^*$ are i.i.d. copies of $B_{n,1}$, since each of the $B_{n,j}$ is $\mathcal{F}_{n,(j-1)r_n+1,jr_n-l_n}$ -measurable, and therefore $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}$ has the same asymptotic behaviour in distribution as $\sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}^*$. Let finally $(X_{n,i}^*)_{1 \leq i \leq n}$ be a triangular array such that the $(X_{n,i}^*)_{(j-1)r_n+1 \leq i \leq jr_n}$ are independent copies of $(X_{n,1}, \dots, X_{n,r_n})$, for $1 \leq j \leq \lfloor n/r_n \rfloor$. Then

$$\begin{aligned}
\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^* &\stackrel{d}{=} \sum_{j=1}^{\lfloor n/r_n \rfloor} \sum_{i=(j-1)r_n+1}^{jr_n} X_{n,i}^*, \\
\sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}^* &\stackrel{d}{=} \sum_{j=1}^{\lfloor n/r_n \rfloor} \sum_{i=(j-1)r_n+1}^{jr_n-l_n} X_{n,i}^*, \\
\text{and } \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j}^* &\stackrel{d}{=} \sum_{j=1}^{\lfloor n/r_n \rfloor} \sum_{i=jr_n-l_n+1}^{jr_n} X_{n,i}^*.
\end{aligned}$$

As a consequence,

$$\begin{aligned} & \left| \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^* \right] \right) - \mathbb{E} \left(\exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}^* \right] \right) \right| \\ & \leq \mathbb{E} \left| 1 - \exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} \sum_{k=jr_n-l_n+1}^{jr_n} X_{n,k}^* \right] \right| = \mathbb{E} \left| 1 - \exp \left[it \sum_{j=1}^{\lfloor n/r_n \rfloor} R_{n,j}^* \right] \right| \rightarrow 0 \end{aligned}$$

by uniform integrability again. Conclude that $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^*$ has the same asymptotic behaviour in distribution as $\sum_{j=1}^{\lfloor n/r_n \rfloor} B_{n,j}^*$, and finally that $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}$ has the same asymptotic behaviour in distribution as $\sum_{j=1}^{\lfloor n/r_n \rfloor} Z_{n,j}^*$, as required.

(ii) This distributional convergence is obvious when $\sigma^2 = 0$. When $\sigma^2 > 0$, it is a simple corollary of (i) and the standard Lindeberg central limit theorem for i.i.d. sequences (see *e.g.* Theorem 27.2 on p.359 of [Billingsley 1995](#)). \square

The following lemma, a useful technical result on a marginal expected shortfall above a high level that will be used several times in the proof of Theorem 4.1*, generalises Proposition 2 in [Daouia et al. \(2018\)](#).

Lemma C.8. *Assume that (X, Y) is a random pair with heavy-tailed marginal distributions, having respective tail indices γ_X, γ_Y , continuous survival functions \bar{F}_X and \bar{F}_Y , and tail quantile functions U_X and U_Y . Suppose that there is a function $R_{(X,Y)}$ on $[0, \infty]^2 \setminus \{(\infty, \infty)\}$ such that*

$$\forall (x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}, \lim_{s \rightarrow \infty} s \mathbb{P} \left(\bar{F}_X(X) \leq \frac{x}{s}, \bar{F}_Y(Y) \leq \frac{y}{s} \right) = R_{(X,Y)}(x, y).$$

Let $z_n \rightarrow \infty$ as $n \rightarrow \infty$.

(i) Then, for any $a < 1/\gamma_X$,

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(X^a \mathbb{1}\{X > 0\} | Y > z_n)}{[U_X(1/\bar{F}_Y(z_n))]^a} = \int_0^\infty R_{(X,Y)}(x^{-1/a\gamma_X}, 1) dx < \infty.$$

(ii) Suppose moreover that X satisfies Condition C* with parameters (γ_X, ρ_X, A_X) , such that $\gamma_X < 1$ and $\rho_X < 0$, and that there exist $\beta > \gamma_X$ and $\kappa < 0$ such that

$$\sup_{x \in (0, \infty)} \left| \frac{s \mathbb{P}(\bar{F}_X(X) \leq x/s, \bar{F}_Y(Y) \leq y/s) - R_{(X,Y)}(x, y)}{\min(x^\beta, 1)} \right| = O(s^\kappa) \quad \text{as } s \rightarrow \infty$$

uniformly in y belonging to any compact subset of $(0, \infty)$. Then, for any $\varepsilon > 0$,

$$\frac{\mathbb{E}(X \mathbb{1}\{X > 0\} | Y > z_n)}{U_X(1/\bar{F}_Y(z_n))} - \int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx = O(|A_X(1/\bar{F}_Y(z_n))|^{\gamma_X/(1-\rho_X)-\varepsilon}) + O((\bar{F}_Y(z_n))^{-\kappa}).$$

Proof of Lemma C.8. In this proof we make, without further reference, extensive use of the following integration by parts calculation: for any $a > 0$ and any event B ,

$$\begin{aligned}\mathbb{E}(X^a \mathbb{1}\{X > 0\} \mathbb{1}\{B\}) &= \mathbb{E}\left(\int_0^\infty \mathbb{1}\{u < X^a\} \mathbb{1}\{X > 0\} du \mathbb{1}\{B\}\right) \\ &= \mathbb{E}\left(\int_0^\infty \mathbb{1}\{X > u^{1/a}\} \mathbb{1}\{X > 0\} du \mathbb{1}\{B\}\right) \\ &= \int_0^\infty \mathbb{P}(\{X > u^{1/a}\} \cap \{B\}) du.\end{aligned}$$

(i) By assumption and continuity of \bar{F}_X and \bar{F}_Y ,

$$\forall (x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}, \quad \lim_{s \rightarrow \infty} s \mathbb{P}(X > U_X(s/x), Y > U_Y(s/y)) = R_{(X,Y)}(x, y).$$

Remark now that

$$\mathbb{P}(Y > z_n) = \bar{F}_Y(z_n) = (\bar{F}_Y \circ U_Y)(1/\bar{F}_Y(z_n)) = \mathbb{P}(Y > U_Y(1/\bar{F}_Y(z_n))),$$

and use an integration by parts and the change of variables $u = x[U_X(1/\bar{F}_Y(z_n))]^a$ to find

$$\frac{\mathbb{E}(X^a \mathbb{1}\{X > 0\} | Y > z_n)}{[U_X(1/\bar{F}_Y(z_n))]^a} = \int_0^\infty \frac{1}{\bar{F}_Y(z_n)} \mathbb{P}(X > x^{1/a} U_X(1/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n))) dx. \quad (\text{C.48})$$

By Lemma C.1 then with $1 - \tau_n = \bar{F}_Y(z_n)$,

$$\forall x > 0, \quad \frac{1}{\bar{F}_Y(z_n)} \mathbb{P}(X > x^{1/a} U_X(1/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n))) \rightarrow R_{(X,Y)}(x^{-1/a\gamma_X}, 1) \quad \text{as } n \rightarrow \infty.$$

Besides, by continuity of \bar{F}_X and Potter bounds (see Proposition B.1.9.5 in [de Haan & Ferreira 2006](#)), we have, for any fixed $\varepsilon > 0$ such that $(1 + \varepsilon)a\gamma_X < 1$,

$$\begin{aligned}\frac{1}{\bar{F}_Y(z_n)} \mathbb{P}(X > x^{1/a} U_X(1/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n))) &\leq \mathbb{1}\{x \leq 1\} + \frac{\bar{F}_X(x^{1/a} U_X(1/\bar{F}_Y(z_n)))}{\bar{F}_X(U_X(1/\bar{F}_Y(z_n)))} \mathbb{1}\{x > 1\} \\ &\leq \mathbb{1}\{x \leq 1\} + C x^{-1/a\gamma_X + \varepsilon} \mathbb{1}\{x > 1\}\end{aligned}$$

for n large enough, where C is an unimportant positive constant. The upper bound defines an integrable function of $x > 0$, and therefore the dominated convergence theorem entails

$$\frac{\mathbb{E}(X^a \mathbb{1}\{X > 0\} | Y > z_n)}{[U_X(1/\bar{F}_Y(z_n))]^a} \rightarrow \int_0^\infty R_{(X,Y)}(x^{-1/a\gamma_X}, 1) dx \quad \text{as } n \rightarrow \infty,$$

as required.

(ii) We recall (C.48) for $a = 1$ and note that

$$\begin{aligned}
& \left| \int_0^\infty \frac{1}{\overline{F}_Y(z_n)} \mathbb{P}(X > x U_X(1/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n))) dx - \int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx \right| \\
& \leq \int_0^\infty \frac{|\mathbb{P}(X > x U_X(1/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n))) - \mathbb{P}(X > U_X(x^{1/\gamma_X}/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n)))|}{\overline{F}_Y(z_n)} dx \\
& + \int_0^\infty \frac{|\mathbb{P}(X > U_X(x^{1/\gamma_X}/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n))) - R_{(X,Y)}(x^{-1/\gamma_X}, 1)|}{\overline{F}_Y(z_n)} dx. \tag{C.49}
\end{aligned}$$

We now control each of the two terms on the right-hand side of (C.49); to do so, recall the elementary inequality

$$|\mathbb{P}(Z_1 > a, Z_2 > c) - \mathbb{P}(Z_1 > b, Z_2 > c)| \leq \min(|\mathbb{P}(Z_1 > a) - \mathbb{P}(Z_1 > b)|, \mathbb{P}(Z_2 > c)). \tag{C.50}$$

Controlling the first component of the right-hand side of (C.49) requires some additional preparation. Recall that, since X satisfies Condition C*, for any arbitrarily small $\eta > 0$, there is $s_0 > 0$ such that

$$s, sx \geq s_0 \Rightarrow \left| \frac{1}{B_X(s)} \left(\frac{\overline{F}_X(sx)}{\overline{F}_X(s)} - x^{-1/\gamma_X} \right) - x^{-1/\gamma_X} \frac{x^{\rho_X/\gamma_X} - 1}{\gamma_X \rho_X} \right| \leq x^{-(1-\rho_X)/\gamma_X} \max(x^\eta, x^{-\eta})$$

where $B_X(s)$ is asymptotically equivalent, as $s \rightarrow \infty$, to $A_X(1/\overline{F}_X(s))$. This is essentially a consequence of Theorem B.2.18 in [de Haan & Ferreira \(2006\)](#), and is stated as Lemma 5 in [Daouia et al. \(2020\)](#). Choose η so small that $(1 + \eta)\gamma_X < 1$ and

$$\frac{\gamma_X}{1 - \rho_X + \eta\gamma_X} > \frac{\gamma_X}{1 - \rho_X} - \varepsilon.$$

Split the integral defining the first term in (C.49) at $u_n = |A_X(1/\overline{F}_Y(z_n))|^{\gamma_X/(1-\rho_X+\eta\gamma_X)} \rightarrow 0$, remark that $U_X|A_X|^{\gamma_X/(1-\rho_X+\eta\gamma_X)}$ is regularly varying with index $\gamma_X(1 + \gamma_X\eta)/(1 - \rho_X + \gamma_X\eta) > 0$ and thus $u_n U_X(1/\overline{F}_Y(z_n)) \rightarrow \infty$, and use the continuity of \overline{F}_X and inequality (C.50) to find that for n large enough

$$\begin{aligned}
& \int_0^\infty \frac{|\mathbb{P}(X > x U_X(1/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n))) - \mathbb{P}(X > U_X(x^{1/\gamma_X}/\overline{F}_Y(z_n)), Y > U_Y(1/\overline{F}_Y(z_n)))|}{\overline{F}_Y(z_n)} dx \\
& \leq \int_0^{u_n} \frac{\mathbb{P}(Y > U_Y(1/\overline{F}_Y(z_n)))}{\overline{F}_Y(z_n)} dx + \int_{u_n}^\infty \frac{|\mathbb{P}(X > x U_X(1/\overline{F}_Y(z_n))) - \mathbb{P}(X > U_X(x^{1/\gamma_X}/\overline{F}_Y(z_n)))|}{\overline{F}_Y(z_n)} dx \\
& = \int_0^{u_n} dx + \int_{u_n}^\infty \left| \frac{\overline{F}_X(x U_X(1/\overline{F}_Y(z_n)))}{\overline{F}_X(U_X(1/\overline{F}_Y(z_n)))} - x^{-1/\gamma_X} \right| dx \\
& = O(u_n) + O\left(|A_X(1/\overline{F}_Y(z_n))| \int_{u_n}^\infty [x^{-1/\gamma_X} + x^{-(1-\rho_X)/\gamma_X} \max(x^\eta, x^{-\eta})] dx\right) \\
& = O(u_n + |A_X(1/\overline{F}_Y(z_n))| u_n^{-(1-\rho_X)/\gamma_X+1-\eta}) = O(|A_X(1/\overline{F}_Y(z_n))|^{\gamma_X/(1-\rho_X+\eta\gamma_X)}).
\end{aligned}$$

Consequently, by definition of η ,

$$\begin{aligned} & \int_0^\infty \frac{|\mathbb{P}(X > x U_X(1/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n))) - \mathbb{P}(X > U_X(x^{1/\gamma_X}/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n)))|}{\bar{F}_Y(z_n)} dx \\ &= O(|A_X(1/\bar{F}_Y(z_n))|^{\gamma_X/(1-\rho_X)-\varepsilon}). \end{aligned} \quad (\text{C.51})$$

Finally, note that by assumption the function $x \mapsto \min(x^{-\beta/\gamma_X}, 1)$ is integrable on $(0, \infty)$, so that the second term in (C.49) is immediately controlled as

$$\int_0^\infty \frac{|\mathbb{P}(X > U_X(x^{1/\gamma_X}/\bar{F}_Y(z_n)), Y > U_Y(1/\bar{F}_Y(z_n))) - R_{(X,Y)}(x^{-1/\gamma_X}, 1)|}{\bar{F}_Y(z_n)} dx = O((\bar{F}_Y(z_n))^{-\kappa}). \quad (\text{C.52})$$

Combine (C.48), (C.49), (C.51), and (C.52) to complete the proof. \square

Lemma C.9 contains a variance calculation which is crucial for the proof of Theorem 4.1*. It is an analogous of Lemma C.2 for the bivariate time series context.

Lemma C.9. *Assume that Condition D^* is satisfied. Let $r_n \rightarrow \infty$, $\tau_n \uparrow 1$ and $z_n \rightarrow \infty$ be such that $r_n(1 - \tau_n) \rightarrow 0$, $n(1 - \tau_n) \rightarrow \infty$ and $\bar{F}_Y(z_n)/(1 - \tau_n) \rightarrow z \in (0, \infty)$ as $n \rightarrow \infty$. Then, if $s_n \rightarrow \infty$ is a sequence of integers such that $s_n = O(r_n)$, we have, as $n \rightarrow \infty$,*

$$\text{Var} \left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} [\mathbb{1}\{Y_t > z_n\} - \mathbb{P}(Y_1 > z_n)] \right) \rightarrow z \left(1 + 2 \sum_{t=1}^{\infty} R_{Y,t}(1, 1) \right) < \infty.$$

Assume moreover that $R_{(X,Y)}(1, 1) > 0$ and $\gamma_X < 1/2$. Then

$$\begin{aligned} & \text{Var} \left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_n\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_n)} \right) \\ & \rightarrow z \left(\frac{\int_0^\infty R_{(X,Y)}(x^{-1/2\gamma_X}, 1) dx + 2 \int_0^\infty \int_0^\infty \sum_{t=1}^{\infty} r_t(x^{-1/\gamma_X}, y^{-1/\gamma_X}, 1, 1) dx dy}{[\int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx]^2} \right) < \infty. \end{aligned}$$

[Recall that, under Condition D^* , we define $R_{Y,t}(y_1, y_{t+1}) = r_t(\infty, \infty, y_1, y_{t+1})$ and $R_{(X,Y)} = r_1(x, \infty, y, \infty)$.]

Proof of Lemma C.9. The proof of the first convergence is exactly that of the first convergence in Lemma C.2, replacing ξ_{τ_n} by z_n and $\gamma^{-1} - 1 := \gamma_Y^{-1} - 1$ by z , so there is nothing to be proven here.

To show the second convergence, we note that by stationarity of (X_t, Y_t) ,

$$\begin{aligned} & \text{Var} \left(\frac{1}{\sqrt{s_n(1 - \tau_n)}} \sum_{t=1}^{s_n} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_n\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_n)} \right) \\ &= \frac{1}{[\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_n)]^2} \left[\frac{1}{1 - \tau_n} \text{Var}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\}) - \frac{s_n - 1}{1 - \tau_n} [\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\})]^2 \right. \\ & \quad \left. + \frac{2}{s_n(1 - \tau_n)} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E}(X_1 X_{t+1} \mathbb{1}\{X_1 > 0, X_{t+1} > 0, Y_1 > z_n, Y_{t+1} > z_n\}) \right]. \end{aligned} \quad (\text{C.53})$$

Apply Lemma C.8(i) with $a = 1$ and 2, and use the convergence $\bar{F}_Y(z_n)/(1 - \tau_n) \rightarrow z$ as $n \rightarrow \infty$ to get

$$\begin{aligned} & \frac{1}{[\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_n)]^2} \left[\frac{1}{1 - \tau_n} \text{Var}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\}) - \frac{s_n - 1}{1 - \tau_n} [\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_n\})]^2 \right] \\ & \rightarrow z \frac{\int_0^\infty R_{(X,Y)}(x^{-1/2\gamma_X}, 1) dx}{[\int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx]^2} \text{ as } n \rightarrow \infty. \end{aligned} \quad (\text{C.54})$$

We now concentrate on the sum in (C.53). An integration by parts and a change of variables provide

$$\begin{aligned} & \frac{2}{s_n(1 - \tau_n)} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E}(X_1 X_{t+1} \mathbb{1}\{X_1 > 0, X_{t+1} > 0, Y_1 > z_n, Y_{t+1} > z_n\}) \\ & = \frac{2}{s_n(1 - \tau_n)} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E} \left(\iint_{[0,\infty)^2} \mathbb{1}\{w_1 < X_1\} \mathbb{1}\{w_{t+1} < X_{t+1}\} dw_1 dw_{t+1} \mathbb{1}\{Y_1 > z_n, Y_{t+1} > z_n\} \right) \\ & = 2 q_{X,\tau_n}^2 \sum_{t=1}^{s_n-1} \frac{s_n - t}{s_n} \iint_{[0,\infty)^2} \frac{\mathbb{P}(X_1 > q_{X,\tau_n} u, X_{t+1} > q_{X,\tau_n} v, Y_1 > z_n, Y_{t+1} > z_n)}{1 - \tau_n} du dv. \end{aligned} \quad (\text{C.55})$$

[Here q_{X,τ_n} is the quantile of X_1 at level τ_n .] By Lemma C.1 and Condition D(iii)* we have

$$\frac{s_n - t}{s_n} \frac{1}{1 - \tau_n} \mathbb{P}(X_1 > q_{X,\tau_n} u, X_{t+1} > q_{X,\tau_n} v, Y_1 > z_n, Y_{t+1} > z_n) \mathbb{1}\{t < s_n\} \rightarrow r_t(u^{-1/\gamma_X}, v^{-1/\gamma_X}, z, z) \quad (\text{C.56})$$

as $n \rightarrow \infty$, pointwise in $u, v > 0$ for any t . To swap limit and sum of integrals in (C.55), we note that, by continuity of F_X and F_Y , Potter bounds (see Proposition B.1.9.5 in de Haan & Ferreira 2006), Condition D(iv)* and the convergence $\bar{F}_Y(z_n)/(1 - \tau_n) \rightarrow z$ as $n \rightarrow \infty$, we have, for any arbitrarily small $\varepsilon > 0$,

$$\begin{aligned} & \frac{s_n - t}{s_n} \times \frac{1}{1 - \tau_n} \mathbb{P}(X_1 > q_{X,\tau_n} u, X_{t+1} > q_{X,\tau_n} v, Y_1 > z_n, Y_{t+1} > z_n) \mathbb{1}\{t < s_n\} \\ & \leq \frac{1}{1 - \tau_n} \mathbb{P}(\bar{F}_X(X_1) \leq C u^{-1/\gamma_X + 2\varepsilon} (1 - \tau_n), \bar{F}_X(X_{t+1}) \leq C v^{-1/\gamma_X + 2\varepsilon} (1 - \tau_n), \\ & \quad \bar{F}_Y(Y_1) \leq (1 + \varepsilon) z (1 - \tau_n), \bar{F}_Y(Y_{t+1}) \leq (1 + \varepsilon) z (1 - \tau_n)) \mathbb{1}\{t < s_n\} \\ & \leq C \left(\rho(t) \sqrt{\min(u^{-1/\gamma_X + 2\varepsilon}, 1) \min(v^{-1/\gamma_X + 2\varepsilon}, 1)} \right. \\ & \quad \left. + (1 - \tau_n) \min(u^{-1/\gamma_X + 2\varepsilon}, 1) \min(v^{-1/\gamma_X + 2\varepsilon}, 1) \right) \mathbb{1}\{t < s_n\} \end{aligned} \quad (\text{C.57})$$

for n large enough, irrespective of the values of the positive integer t and $u, v > 0$; here C is an unimportant positive constant whose value may change from one line to the next. Choose $\varepsilon > 0$ such that $(1 + \varepsilon)2\gamma_X < 1$. We now remark that the upper bound, say $C \phi_n(t, u, v)$, satisfies

$$\phi_n(t, u, v) \rightarrow \rho(t) \sqrt{\min(u^{-1/\gamma_X + 2\varepsilon}, 1) \min(v^{-1/\gamma_X + 2\varepsilon}, 1)} \quad (\text{C.58})$$

as $n \rightarrow \infty$, pointwise in $t \geq 1$, $u, v > 0$. The functions $u \mapsto \sqrt{\min(u^{-1/\gamma_X+2\varepsilon}, 1)}$ and $u \mapsto \min(u^{-1/\gamma_X+2\varepsilon}, 1)$ are both integrable on $[0, \infty)$, from which it follows that

$$\begin{aligned} & \sum_{t=1}^{\infty} \iint_{[0, \infty)^2} \phi_n(t, u, v) \, du \, dv \mathbb{1}\{t < s_n\} \\ & \rightarrow \sum_{t=1}^{\infty} \iint_{[0, \infty)^2} \rho(t) \sqrt{\min(u^{-1/\gamma_X+2\varepsilon}, 1) \min(v^{-1/\gamma_X+2\varepsilon}, 1)} \, du \, dv < \infty, \end{aligned} \quad (\text{C.59})$$

as $n \rightarrow \infty$, by splitting the sum and since $s_n(1 - \tau_n) \rightarrow 0$. A combination of Theorem 1 in Pratt (1960) with (C.55), (C.56), (C.57), (C.58), (C.59), the Tonelli theorem and a change of variables now yields, as $n \rightarrow \infty$,

$$\begin{aligned} & \frac{2}{s_n(1 - \tau_n)q_{X, \tau_n}^2} \sum_{t=1}^{s_n-1} (s_n - t) \mathbb{E}(X_1 X_{t+1} \mathbb{1}\{X_1 > 0, X_{t+1} > 0, Y_1 > z_n, Y_{t+1} > z_n\}) \\ & \rightarrow 2 \sum_{t=1}^{\infty} \iint_{[0, \infty)^2} r_t(u^{-1/\gamma_X}, v^{-1/\gamma_X}, z, z) \, du \, dv \\ & = 2 z^{1-2\gamma_X} \iint_{[0, \infty)^2} \sum_{t=1}^{\infty} r_t(x^{-1/\gamma_X}, y^{-1/\gamma_X}, 1, 1) \, dx \, dy. \end{aligned} \quad (\text{C.60})$$

Noting that $U_X(1/\bar{F}_Y(z_n)) = z^{-\gamma_X} q_{X, \tau_n} (1 + o(1))$ (using the convergence $\bar{F}_Y(z_n)/(1 - \tau_n) \rightarrow z$ as $n \rightarrow \infty$ and the regular variation property of U_X) and combining Lemma C.8(i) with (C.53), (C.54) and (C.60) completes the proof. \square

*Proof of Theorem 4.1**. Consider $\text{MES}_{X, \tau}^{\pm} := \mathbb{E}(X \mathbb{1}\{X > 0\} | Y > z_{Y, \tau})$. Write

$$\frac{\overline{\text{MES}}_{X, \tau'_n}^*}{\text{MES}_{X, \tau'_n}} = \frac{\overline{\text{MES}}_{X, \tau'_n}^*}{\text{MES}_{X, \tau'_n}^+} \times \frac{\text{MES}_{X, \tau'_n}^+}{\text{MES}_{X, \tau'_n}}. \quad (\text{C.61})$$

Note that

$$\begin{aligned} \frac{\text{MES}_{X, \tau'_n}}{\text{MES}_{X, \tau'_n}^+} &= 1 + \frac{\mathbb{E}(X \mathbb{1}\{X < 0\} | Y > z_{Y, \tau'_n})}{\text{MES}_{X, \tau'_n}^+} \\ &= 1 + \frac{\mathbb{E}(X \mathbb{1}\{X < 0\} | Y > U_Y(1/\bar{F}_Y(z_{Y, \tau'_n})))}{\text{MES}_{X, \tau'_n}^+} \\ &= 1 + O([\bar{F}_Y(z_{Y, \tau'_n})]^{-\kappa(1-\gamma_X)}) \quad (\text{see Lemma C.8(i) and p.441 of Cai } et al., 2015) \\ &= 1 + O((1 - \tau'_n)^{-\kappa(1-\gamma_X)}) = 1 + o\left(\frac{1}{\sqrt{n(1 - \tau_n)}}\right). \end{aligned} \quad (\text{C.62})$$

Now let

$$\overline{\text{MES}}_{X, \tau_n} := \frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > \bar{z}_{Y, \tau_n}\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > \bar{z}_{Y, \tau_n}\}}$$

and write

$$\log \left(\frac{\overline{\text{MES}}_{X,\tau'_n}^*}{\text{MES}_{X,\tau'_n}^+} \right) = \log \left(\frac{\overline{\text{MES}}_{X,\tau_n}}{\text{MES}_{X,\tau_n}^+} \right) + \log \left(\frac{\text{MES}_{X,\tau_n}^+}{\text{MES}_{X,\tau'_n}^+} \left(\frac{1 - \tau'_n}{1 - \tau_n} \right)^{-\gamma_X} \right) + (\widehat{\gamma}_{X,n} - \gamma_X) \log \left(\frac{1 - \tau_n}{1 - \tau'_n} \right).$$

Because of (C.61) and (C.62), the proof will be complete if we can show that

$$\sqrt{n(1 - \tau_n)} \log \left(\frac{\overline{\text{MES}}_{X,\tau_n}}{\text{MES}_{X,\tau_n}^+} \right) = O_{\mathbb{P}}(1) \quad (\text{C.63})$$

and

$$\sqrt{n(1 - \tau_n)} \left(\frac{\text{MES}_{X,\tau_n}^+}{\text{MES}_{X,\tau'_n}^+} \left(\frac{1 - \tau'_n}{1 - \tau_n} \right)^{-\gamma_X} - 1 \right) = O(1). \quad (\text{C.64})$$

We first focus on showing (C.63) and write

$$\begin{aligned} \log \left(\frac{\overline{\text{MES}}_{X,\tau_n}}{\text{MES}_{X,\tau_n}^+} \right) &= \log \left(\frac{\mathbb{E}(\mathbb{1}\{Y_1 > z_{Y,\tau_n}\})}{\frac{1}{n} \sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} \right) + \log \left(\frac{\frac{1}{n} \sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})} \right) \\ &+ \log \left(\frac{\frac{1}{n} \sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}}{\frac{1}{n} \sum_{t=1}^n \mathbb{1}\{Y_t > \bar{z}_{Y,\tau_n}\}} \right) + \log \left(\frac{\frac{1}{n} \sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > \bar{z}_{Y,\tau_n}\}}{\frac{1}{n} \sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} \right). \end{aligned} \quad (\text{C.65})$$

Firstly, using the asymptotic proportionality relationship $\bar{F}_Y(z_{Y,\tau_n})/(1 - \tau_n) \rightarrow z$ as $n \rightarrow \infty$,

$$\sqrt{n(1 - \tau_n)} \left(\frac{\frac{1}{n} \sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}}{\mathbb{E}(\mathbb{1}\{Y_1 > z_{Y,\tau_n}\})} - 1 \right) = z^{-1}(1 + o(1)) \left\{ \frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^n [\mathbb{1}\{Y_t > z_{Y,\tau_n}\} - \mathbb{P}(Y_1 > z_{Y,\tau_n})] \right\}.$$

We check the assumptions of Corollary 4.2 of [Rootzén et al. \(1998\)](#) to control the quantity in curly brackets. By Lemma C.9, one has, as $n \rightarrow \infty$,

$$\begin{aligned} \frac{n}{r_n} \text{Var} \left(\frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^{l_n} [\mathbb{1}\{Y_t > z_{Y,\tau_n}\} - \mathbb{P}(Y_1 > z_{Y,\tau_n})] \right) &= O \left(\frac{l_n}{r_n} \right) \rightarrow 0, \\ \text{Var} \left(\frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^{n - r_n \lfloor n/r_n \rfloor} [\mathbb{1}\{Y_t > z_{Y,\tau_n}\} - \mathbb{P}(Y_1 > z_{Y,\tau_n})] \right) &= O \left(\frac{r_n}{n} \right) \rightarrow 0, \\ \text{and } \frac{n}{r_n} \text{Var} \left(\frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^{r_n} [\mathbb{1}\{Y_t > z_{Y,\tau_n}\} - \mathbb{P}(Y_1 > z_{Y,\tau_n})] \right) &\rightarrow z \left(1 + 2 \sum_{t=1}^{\infty} R_{Y,t}(1, 1) \right) < \infty \end{aligned}$$

where we recall that $R_{Y,t}(1, 1) = r_t(\infty, \infty, 1, 1)$. Setting

$$Z_{n,1} = \frac{1}{\sqrt{n(1 - \tau_n)}} \sum_{t=1}^{r_n} [\mathbb{1}\{Y_t > z_{Y,\tau_n}\} - \mathbb{P}(Y_1 > z_{Y,\tau_n})],$$

we also have, for any $\eta > 0$,

$$\frac{n}{r_n} \mathbb{E}(Z_{n,1}^2 \mathbb{1}\{|Z_{n,1}| > \eta\}) \leq \eta^{-\delta} \times \frac{n}{r_n} \mathbb{E}|Z_{n,1}|^{2+\delta} = O \left(r_n \left[\frac{r_n}{\sqrt{n(1 - \tau_n)}} \right]^{\delta} \right) \rightarrow 0$$

by Hölder's inequality and the asymptotic proportionality relationship $\bar{F}_Y(z_{Y,\tau_n})/(1-\tau_n) \rightarrow z$ as $n \rightarrow \infty$. [We can in fact take δ arbitrarily large here, but the more restrictive bound above appears again later in the proof.] Applying Corollary 4.2 in [Rootzén et al. \(1998\)](#) and the delta method, we find

$$\sqrt{n(1-\tau_n)} \left(\frac{\frac{1}{n} \sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}}{\mathbb{E}(\mathbb{1}\{Y_1 > z_{Y,\tau_n}\})} - 1 \right) = O_{\mathbb{P}}(1). \quad (\text{C.66})$$

Secondly, using again the asymptotic proportionality relationship $\bar{F}_Y(z_{Y,\tau_n})/(1-\tau_n) \rightarrow z$ as $n \rightarrow \infty$,

$$\begin{aligned} & \sqrt{n(1-\tau_n)} \left(\frac{\frac{1}{n} \sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})} - 1 \right) \\ &= z^{-1}(1 + o(1)) \left\{ \frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^n \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_{Y,\tau_n})} \right\}. \end{aligned}$$

This time, we check the assumptions of Lemma [C.7](#) to control the quantity between curly brackets above.

By Lemma [C.9](#), one has, as $n \rightarrow \infty$,

$$\begin{aligned} & \frac{n}{r_n} \text{Var} \left(\frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^{l_n} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_{Y,\tau_n})} \right) = O \left(\frac{l_n}{r_n} \right) \rightarrow 0, \\ & \text{Var} \left(\frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^{n-r_n \lfloor n/r_n \rfloor} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_{Y,\tau_n})} \right) = O \left(\frac{r_n}{n} \right) \rightarrow 0 \\ & \text{and } \lim_{n \rightarrow \infty} \frac{n}{r_n} \text{Var} \left(\frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^{r_n} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_{Y,\tau_n})} \right) < \infty. \end{aligned}$$

Setting

$$\mathcal{Z}_{n,1} = \frac{1}{\sqrt{n(1-\tau_n)}} \sum_{t=1}^{r_n} \frac{X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\} - \mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0\} | Y_1 > z_{Y,\tau_n})},$$

we also have, for any $\eta > 0$,

$$\frac{n}{r_n} \mathbb{E}(\mathcal{Z}_{n,1}^2 \mathbb{1}\{|\mathcal{Z}_{n,1}| > \eta\}) \leq \eta^{-\delta} \times \frac{n}{r_n} \mathbb{E}|\mathcal{Z}_{n,1}|^{2+\delta} = O \left(r_n \left[\frac{r_n}{\sqrt{n(1-\tau_n)}} \right]^\delta \right) \rightarrow 0$$

as $n \rightarrow \infty$, by Hölder's inequality, Lemma [C.8](#)(i) with $a = 2 + \delta$, and the asymptotic proportionality relationship $\bar{F}_Y(z_{Y,\tau_n})/(1-\tau_n) \rightarrow z$. By Lemma [C.7](#) then,

$$\sqrt{n(1-\tau_n)} \left(\frac{\frac{1}{n} \sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})} - 1 \right) = O_{\mathbb{P}}(1). \quad (\text{C.67})$$

Thirdly, by assumption, \bar{z}_{Y,τ_n} is a $\sqrt{n(1-\tau_n)}$ -relatively consistent estimator of z_{Y,τ_n} , so that for any $\eta > 0$, we may find $K > 0$ such that

$$\left| \frac{\bar{z}_{Y,\tau_n}}{z_{Y,\tau_n}} - 1 \right| \leq \frac{K}{\sqrt{n(1-\tau_n)}}$$

with probability larger than $1 - \eta$ eventually. Therefore

$$\left| \frac{\sum_{t=1}^n \mathbb{1}\{Y_t > \bar{z}_{Y,\tau_n}\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} - 1 \right| \leq \max \left(\left| \frac{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}(1 + K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} - 1 \right|, \right. \\ \left. \left| \frac{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}(1 - K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} - 1 \right| \right)$$

and likewise

$$\left| \frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > \bar{z}_{Y,\tau_n}\}}{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} - 1 \right| \leq \max \left(\left| \frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}(1 + K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} - 1 \right|, \right. \\ \left. \left| \frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}(1 - K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} - 1 \right| \right).$$

Recall (C.66) and (C.67): thanks to Lemma C.9, direct analogues of these equations can be shown with $z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})$ instead of z_{Y,τ_n} , and they imply

$$\sqrt{n(1-\tau_n)} \left(\frac{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} - 1 \right) \\ = O_{\mathbb{P}} \left(1 + \sqrt{n(1-\tau_n)} \left| \frac{\mathbb{P}(Y_1 > z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})\}}{\mathbb{P}(Y_1 > z_{Y,\tau_n})} - 1 \right| \right)$$

and

$$\sqrt{n(1-\tau_n)} \left(\frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})\}}{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} - 1 \right) \\ = O_{\mathbb{P}} \left(1 + \sqrt{n(1-\tau_n)} \left| \frac{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})\})}{\mathbb{E}(X_1 \mathbb{1}\{X_1 > 0, Y_1 > z_{Y,\tau_n}\})} - 1 \right| \right).$$

[Note the form of the $O_{\mathbb{P}}$ terms, which correct those of the proof of Theorem 4 in Daouia et al. (2018, p.14 of the Supplementary Material).] By Lemma C.8(ii) with $z_n = z_{Y,\tau_n}$ and $z_n = z_{Y,\tau_n}(1 \pm K/\sqrt{n(1-\tau_n)})$, local uniformity of second-order regular variation (see Theorem 2.3.9 in de Haan & Ferreira 2006), and the convergences $\sqrt{n(1-\tau_n)}A_X((1-\tau_n)^{-1}) \rightarrow 0$ (that follows from condition (iii) in the statement of the Theorem, since $\gamma_X < 1$) and $\sqrt{n(1-\tau_n)}A_Y((1-\tau_n)^{-1}) \rightarrow 0$ as $n \rightarrow \infty$, we therefore obtain

$$\sqrt{n(1-\tau_n)} \left(\frac{\sum_{t=1}^n \mathbb{1}\{Y_t > \bar{z}_{Y,\tau_n}\}}{\sum_{t=1}^n \mathbb{1}\{Y_t > z_{Y,\tau_n}\}} - 1 \right) = O_{\mathbb{P}}(1) \quad (\text{C.68})$$

and

$$\sqrt{n(1-\tau_n)} \left(\frac{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > \bar{z}_{Y,\tau_n}\}}{\sum_{t=1}^n X_t \mathbb{1}\{X_t > 0, Y_t > z_{Y,\tau_n}\}} - 1 \right) = O_{\mathbb{P}}(1). \quad (\text{C.69})$$

Combining (C.65), (C.66), (C.67), (C.68) and (C.69) concludes the proof of (C.63).

Now, to prove (C.64), apply Lemma C.8(ii) and use the regular variation properties of $|A_X|$ and \bar{F}_Y to get

$$\sqrt{n(1-\tau_n)} \left(\frac{\text{MES}_{X,\tau_n}^+}{U_X(1/\bar{F}_Y(z_{Y,\tau_n}))} - \int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx \right) = o(1) \quad (\text{C.70})$$

$$\text{and } \sqrt{n(1-\tau_n)} \left(\frac{\text{MES}_{X,\tau'_n}^+}{U_X(1/\bar{F}_Y(z_{Y,\tau'_n}))} - \int_0^\infty R_{(X,Y)}(x^{-1/\gamma_X}, 1) dx \right) = o(1). \quad (\text{C.71})$$

Using the assumptions

$$\sqrt{n(1-\tau_n)} \left(\frac{\bar{F}_Y(z_{Y,\tau_n})}{1-\tau_n} - z \right) = o(1) \quad \text{and} \quad \sqrt{n(1-\tau_n)} \left(\frac{\bar{F}_Y(z_{Y,\tau'_n})}{1-\tau'_n} - z \right) = o(1)$$

we get, by local uniformity of Condition C* (see Theorem 2.3.9 in de Haan & Ferreira 2006), the assumption $\rho_X < 0$ and convergence $\sqrt{n(1-\tau_n)}A_X((1-\tau_n)^{-1}) \rightarrow 0$ as $n \rightarrow \infty$ that

$$\sqrt{n(1-\tau_n)} \left(\frac{U_X(1/\bar{F}_Y(z_{Y,\tau_n}))}{U_X(1/\bar{F}_Y(z_{Y,\tau'_n}))} \left(\frac{1-\tau'_n}{1-\tau_n} \right)^{-\gamma_X} - 1 \right) = o(1). \quad (\text{C.72})$$

Combining (C.70), (C.71) and (C.72) completes the proof of (C.64) and thus the proof of the result itself. \square

Proof of Theorem 5.1.* We focus on the composite LAWS estimator, the proof being entirely similar for the composite QB estimator. Suppose generally that $\sqrt{n(1-\tau_n)}(\hat{\gamma}_n - \gamma) \xrightarrow{d} \Gamma$ where Γ is a nondegenerate distribution. Our first goal is to show that

$$\frac{1 - \hat{\tau}'_n(\alpha_n)}{1 - \tau'_n(\alpha_n)} = 1 + O_{\mathbb{P}} \left(\frac{1}{\sqrt{n(1-\tau_n)}} \right). \quad (\text{C.73})$$

To this end, we write

$$\frac{1 - \hat{\tau}'_n(\alpha_n)}{1 - \tau'_n(\alpha_n)} - 1 = \frac{\hat{\gamma}_n}{\gamma} \times \frac{1-\gamma}{1-\hat{\gamma}_n} \times \frac{(1-\alpha_n)\frac{\gamma}{1-\gamma}}{1-\tau'_n(\alpha_n)} - 1. \quad (\text{C.74})$$

The delta-method yields

$$\sqrt{n(1-\tau_n)} \left(\frac{\gamma}{\hat{\gamma}_n} \times \frac{1-\hat{\gamma}_n}{1-\gamma} - 1 \right) = O_{\mathbb{P}}(1). \quad (\text{C.75})$$

Writing $\mathbb{E}((Y-t)\mathbb{1}\{Y > t\}) = \int_t^\infty \bar{F}(x) dx$ for any t and using the heavy-tailed assumption entails, after straightforward calculations,

$$\frac{(1-\alpha_n)\frac{\gamma}{1-\gamma}}{\mathbb{E} \left[\left| \frac{Y}{q_{\alpha_n}} - 1 \right| \mathbb{1}\{Y > q_{\alpha_n}\} \right]} - 1 = O[A(1/\bar{F}(q_{\alpha_n}))] = o(1/\sqrt{n(1-\tau_n)}), \quad (\text{C.76})$$

because of the continuity of F and the regular variation property of $|A|$. Write further

$$\begin{aligned} \mathbb{E} \left| \frac{Y}{q_{\alpha_n}} - 1 \right| - 1 &= \mathbb{E} \left[\left| \frac{Y}{q_{\alpha_n}} - 1 \right| \mathbb{1}\{Y > q_{\alpha_n}\} \right] - \frac{\mathbb{E}(Y \mathbb{1}\{Y \leq q_{\alpha_n}\})}{q_{\alpha_n}} - \bar{F}(q_{\alpha_n}) \\ &= O(\max(1 - \alpha_n, 1/q_{\alpha_n})) = o(1/\sqrt{n(1 - \tau_n)}) \end{aligned} \quad (\text{C.77})$$

where we successively used (C.75), the dominated convergence theorem, the relationship $1 - \alpha_n = o(1/q_{\alpha_n})$ valid because $0 < \gamma < 1$, and the regular variation property of $t \mapsto q_{1-t}$. Combining (C.74), (C.75), (C.76) and (C.77) with the definition of $\tau'_n(\alpha_n)$ results in (C.73).

Write now

$$\tilde{\xi}_{\widehat{\tau}'_n(\alpha_n)}^* = \left(\frac{1 - \widehat{\tau}'_n(\alpha_n)}{1 - \tau_n} \right)^{-\widehat{\gamma}_n} \tilde{\xi}_{\tau_n} = \left(\frac{1 - \widehat{\tau}'_n(\alpha_n)}{1 - \tau'_n(\alpha_n)} \right)^{-\widehat{\gamma}_n} \times \left\{ \left(\frac{1 - \tau'_n(\alpha_n)}{1 - \tau_n} \right)^{-\widehat{\gamma}_n} \tilde{\xi}_{\tau_n} \right\}. \quad (\text{C.78})$$

By (C.73), the $\sqrt{n(1 - \tau_n)}$ -convergence of $\widehat{\gamma}_n$, and a Taylor expansion,

$$\left(\frac{1 - \widehat{\tau}'_n(\alpha_n)}{1 - \tau'_n(\alpha_n)} \right)^{-\widehat{\gamma}_n} = 1 + O_{\mathbb{P}} \left(\frac{1}{\sqrt{n(1 - \tau_n)}} \right). \quad (\text{C.79})$$

Furthermore

$$\left(\frac{1 - \tau'_n(\alpha_n)}{1 - \tau_n} \right)^{-\widehat{\gamma}_n} \tilde{\xi}_{\tau_n} = \tilde{\xi}_{\tau'_n(\alpha_n)}^*$$

by definition of the class of estimators $\tilde{\xi}^*$. From the convergence $(1 - \tau'_n(\alpha_n))/(1 - \alpha_n) \rightarrow \gamma/(1 - \gamma)$ as $n \rightarrow \infty$, we conclude that Proposition C.6 can be applied, and this gives

$$\frac{\sqrt{n(1 - \tau_n)}}{\log[(1 - \tau_n)/(1 - \tau'_n(\alpha_n))]} \left(\frac{\tilde{\xi}_{\tau'_n(\alpha_n)}^*}{\tilde{\xi}_{\tau'_n(\alpha_n)}} - 1 \right) \xrightarrow{d} \Gamma.$$

Finally

$$\log \left[\frac{1 - \tau_n}{1 - \tau'_n(\alpha_n)} \right] = \log \left[\frac{1 - \tau_n}{1 - \alpha_n} \right] + \log \left[\frac{1 - \alpha_n}{1 - \tau'_n(\alpha_n)} \right] = \log \left[\frac{1 - \tau_n}{1 - \alpha_n} \right] (1 + o(1)).$$

Together with the equality $\xi_{\tau'_n(\alpha_n)} = q_{\alpha_n}$ which is true by definition of $\tau'_n(\alpha_n)$, this entails

$$\frac{\sqrt{n(1 - \tau_n)}}{\log[(1 - \tau_n)/(1 - \alpha_n)]} \left(\frac{\tilde{\xi}_{\tau'_n(\alpha_n)}^*}{q_{\alpha_n}} - 1 \right) \xrightarrow{d} \Gamma. \quad (\text{C.80})$$

Combining (C.78), (C.79) and (C.80) completes the proof. \square

*Proof of Theorem 5.2**. This proof is similar to that of Theorem 5.1* and is thus omitted. \square

*Proof of Theorem 6.1**. Define

$$\begin{aligned}\chi_n(u) &:= \frac{1}{2\xi_{\tau_n}^2(\varepsilon)} \sum_{t=1}^n \left[\eta_{\tau_n} \left(\widehat{\varepsilon}_t^{(n)} - \xi_{\tau_n}(\varepsilon) - \frac{u\xi_{\tau_n}(\varepsilon)}{\sqrt{n(1-\tau_n)}} \right) - \eta_{\tau_n}(\widehat{\varepsilon}_t^{(n)} - \xi_{\tau_n}(\varepsilon)) \right], \\ \psi_n(u) &:= \frac{1}{2\xi_{\tau_n}^2(\varepsilon)} \sum_{t=1}^n \left[\eta_{\tau_n} \left(\varepsilon_t - \xi_{\tau_n}(\varepsilon) - \frac{u\xi_{\tau_n}(\varepsilon)}{\sqrt{n(1-\tau_n)}} \right) - \eta_{\tau_n}(\varepsilon_t - \xi_{\tau_n}(\varepsilon)) \right].\end{aligned}$$

The χ_n and ψ_n define sequences of convex and everywhere finite functions on \mathbb{R} . Moreover

$$\sqrt{n(1-\tau_n)} \left(\frac{\widetilde{\xi}_{\tau_n}(\varepsilon)}{\xi_{\tau_n}(\varepsilon)} - 1 \right) = \arg \min_{u \in \mathbb{R}} \chi_n(u)$$

and as $n \rightarrow \infty$,

$$\psi_n(u) \xrightarrow{d} -u \sqrt{\frac{2\gamma}{1-2\gamma} + 2(\gamma^{-1}-1) \iint_{[1,\infty)^2} \sum_{t=1}^{\infty} R_t(x^{-1/\gamma}, y^{-1/\gamma}) dx dy} \times Z + \frac{u^2}{2\gamma}$$

in the sense of finite-dimensional convergence (see the proof of Theorem 3.1*), so by Theorem 5 in [Knight \(1999\)](#), it is enough to prove that $\chi_n(u) - \psi_n(u) \xrightarrow{\mathbb{P}} 0$ for any $u \in \mathbb{R}$.

Define $I_n(u) = [0, |u|\xi_{\tau_n}(\varepsilon)/\sqrt{n(1-\tau_n)}]$. Follow the proof of Theorem 2.1 in [Girard et al. \(2021\)](#) to find that it is sufficient to prove $T_{1,n} \xrightarrow{\mathbb{P}} 0$ and $T_{2,n}(u) \xrightarrow{\mathbb{P}} 0$, where

$$\begin{aligned}T_{1,n} &:= \frac{\sqrt{1-\tau_n}}{\xi_{\tau_n}(\varepsilon)\sqrt{n}} \sum_{t=1}^n |\widehat{\varepsilon}_t^{(n)} - \varepsilon_t|, \\ T_{2,n}(u) &:= \frac{2}{\xi_{\tau_n}(\varepsilon)\sqrt{n(1-\tau_n)}} \sum_{t=1}^n \sup_{|s| \in I_n(u)} |\widehat{\varepsilon}_t^{(n)} - \varepsilon_t| \mathbb{1}\{\varepsilon_t - \xi_{\tau_n}(\varepsilon) - s > \min(\varepsilon_t - \widehat{\varepsilon}_t^{(n)}, 0)\}.\end{aligned}$$

Clearly, if $E_n := \max_{1 \leq t \leq n} \frac{|\widehat{\varepsilon}_t^{(n)} - \varepsilon_t|}{1 + |\varepsilon_t|}$, then $\sqrt{n(1-\tau_n)}E_n = o_{\mathbb{P}}(1)$ by Condition F*, and

$$\begin{aligned}T_{1,n} &\leq \left[\frac{\sqrt{n(1-\tau_n)}}{\xi_{\tau_n}(\varepsilon)} E_n \right] \times \frac{1}{n} \sum_{t=1}^n (1 + |\varepsilon_t|) \\ &= (1 + \mathbb{E}|\varepsilon|) \left[\frac{\sqrt{n(1-\tau_n)}}{\xi_{\tau_n}(\varepsilon)} E_n \right] \times \frac{1}{n} \sum_{t=1}^n \frac{1 + |\varepsilon_t|}{1 + \mathbb{E}|\varepsilon|}.\end{aligned}$$

The average on the right-hand side has unit expectation by stationarity of (ε_t) , so the Markov inequality applies to give

$$T_{1,n} = O_{\mathbb{P}} \left(\frac{\sqrt{n(1-\tau_n)}}{\xi_{\tau_n}(\varepsilon)} E_n \right) = o_{\mathbb{P}}(1).$$

To show $T_{2,n}(u) \xrightarrow{\mathbb{P}} 0$, we again follow the proof of Theorem 2.1 in [Girard et al. \(2021\)](#) to find that it is sufficient to prove that

$$T'_{2,n} := \frac{2}{\xi_{\tau_n}(\varepsilon)\sqrt{n(1-\tau_n)}} \sum_{t=1}^n |\widehat{\varepsilon}_t^{(n)} - \varepsilon_t| \mathbb{1} \left\{ \varepsilon_t > \frac{1}{6}\xi_{\tau_n}(\varepsilon) \right\} \xrightarrow{\mathbb{P}} 0.$$

Bound $T'_{2,n}$ as

$$\begin{aligned} T'_{2,n} &\leq 2 \left[\frac{\sqrt{n(1-\tau_n)}}{\xi_{\tau_n}(\varepsilon)} E_n \right] \times \frac{1}{n(1-\tau_n)} \sum_{t=1}^n (1 + |\varepsilon_t|) \mathbb{1} \left\{ \varepsilon_t > \frac{1}{6}\xi_{\tau_n}(\varepsilon) \right\} \\ &\leq 4 \frac{\mathbb{E}(\varepsilon \mathbb{1}\{\varepsilon > \xi_{\tau_n}(\varepsilon)/6\})}{\xi_{\tau_n}(\varepsilon)(1-\tau_n)} \left[\sqrt{n(1-\tau_n)} E_n \right] \times \frac{1}{n} \sum_{t=1}^n \frac{\varepsilon_t \mathbb{1}\{\varepsilon_t > \xi_{\tau_n}(\varepsilon)/6\}}{\mathbb{E}(\varepsilon \mathbb{1}\{\varepsilon > \xi_{\tau_n}(\varepsilon)/6\})} \end{aligned}$$

for n large enough. Using the Markov inequality again, we get

$$T'_{2,n} = O_{\mathbb{P}} \left(\frac{\mathbb{E}(\varepsilon \mathbb{1}\{\varepsilon > \xi_{\tau_n}(\varepsilon)/6\})}{\xi_{\tau_n}(\varepsilon)(1-\tau_n)} \left[\sqrt{n(1-\tau_n)} E_n \right] \right).$$

Straightforward calculations yield, for any $s_n \rightarrow \infty$, that

$$\frac{\mathbb{E}(\varepsilon \mathbb{1}\{\varepsilon > s_n\})}{s_n \overline{F}(s_n)} = 1 + \frac{\mathbb{E}((\varepsilon - s_n) \mathbb{1}\{\varepsilon > s_n\})}{s_n \overline{F}(s_n)} = 1 + \frac{1}{s_n} \int_{s_n}^{\infty} \frac{\overline{F}(x)}{\overline{F}(s_n)} dx \rightarrow \frac{1}{1-\gamma};$$

this follows from the heavy tail assumption on ε with $\gamma < 1$ and Proposition B.1.10 p.369 in [de Haan & Ferreira \(2006\)](#). Apply this with $s_n = \xi_{\tau_n}(\varepsilon)/6$ and use the heavy tail assumption and the asymptotic proportionality relationship $\overline{F}(\xi_{\tau_n}(\varepsilon))/(1-\tau_n) \rightarrow \gamma^{-1} - 1$ (see for example Proposition 1 in [Daouia et al. \(2019\)](#) with $p = 2$) to obtain

$$T'_{2,n} = O_{\mathbb{P}} \left(\sqrt{n(1-\tau_n)} E_n \right) = o_{\mathbb{P}}(1).$$

The proof is complete. □

Lemma C.10 is key to the asymptotic analysis of residual-based extreme value estimators under general dependence conditions.

Lemma C.10. *Let the time series $(\varepsilon_t, t = 1, 2, \dots)$ be stationary and ergodic, with a continuous one-dimensional marginal distribution function F having an infinite right endpoint. Let $k = k(n) \rightarrow \infty$ be a sequence of integers with $k/n \rightarrow 0$. Suppose further that the array of random variables $\widehat{\varepsilon}_t^{(n)}$, $1 \leq t \leq n$, satisfies*

$$E_n := \max_{1 \leq t \leq n} \frac{|\widehat{\varepsilon}_t^{(n)} - \varepsilon_t|}{1 + |\varepsilon_t|} \xrightarrow{\mathbb{P}} 0.$$

Then we have both

$$\sup_{0 < s \leq 1} \left| \frac{\widehat{\varepsilon}_{n-[ks],n}^{(n)}}{\varepsilon_{n-[ks],n}} - 1 \right| = O_{\mathbb{P}}(E_n) \quad \text{and} \quad \sup_{0 < s \leq 1} \left| \log \left(\frac{\widehat{\varepsilon}_{n-[ks],n}^{(n)}}{\varepsilon_{n-[ks],n}} \right) \right| = O_{\mathbb{P}}(E_n).$$

Proof of Lemma C.10. Inspecting the proof of Lemma A.3 in Girard et al. (2021) shows that it will carry over to the present case, and thus yield the desired result, provided $\varepsilon_{n-k,n} \xrightarrow{\mathbb{P}} +\infty$.

Suppose not. Then there exist $\delta, M > 0$ and a subsequence of integers $n = n_p \rightarrow \infty$ such that $\mathbb{P}(\varepsilon_{n_p-k(n_p),n_p} \leq M) > \delta$ for any p . Consequently,

$$\frac{1}{n_p} \sum_{i=1}^{n_p} \mathbb{1}\{\varepsilon_i \leq M\} \geq 1 - \frac{k(n_p)}{n_p} \rightarrow 1 \text{ as } p \rightarrow \infty$$

with probability at least δ . However, by ergodicity of the ε_i ,

$$\lim_{p \rightarrow \infty} \frac{1}{n_p} \sum_{i=1}^{n_p} \mathbb{1}\{\varepsilon_i \leq M\} = \mathbb{P}(\varepsilon \leq M) < 1$$

with probability 1, since the right endpoint of F is finite. This is an obvious contradiction. \square

Proof of Theorem 6.2.* Adapt the proof of Theorem 2.2 in Girard et al. (2021) by applying Lemma C.10 and Theorem 3.3* rather than Lemma A.3 in Girard et al. (2021) and Theorem 2.4.8 in de Haan & Ferreira (2006, p.52). \square

Proof of Corollary 6.3.* Adapt the proof of Corollary 2.1 in Girard et al. (2021) by applying our Theorem 6.2* rather than Theorem 2.2 in Girard et al. (2021). \square

Proof of Theorem 6.4.* Theorem 6.1* gives the $\sqrt{n(1-\tau_n)}$ -asymptotic normality of $\tilde{\xi}_{\tau_n}(\varepsilon)$, and Corollary 6.3* gives the $\sqrt{n(1-\tau_n)}$ -asymptotic normality of $\hat{\xi}_{\tau_n}(\varepsilon)$ and $\hat{\gamma}_n(\varepsilon) = \hat{\gamma}_n^H(\varepsilon)$. Apply Proposition C.6 to complete the proof of the asymptotic normality of $\bar{\xi}_{\tau_n}^*(\varepsilon)$ (note that the assumption that $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1}(\varepsilon) = O(1)$, rather than $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1}(\varepsilon) \rightarrow \lambda_2$, is sufficient because $\mathbb{E}(\varepsilon) = 0$). The proof of the asymptotic normality of $\bar{\xi}_{\tau_n}^*(Y_1 | \mathbf{X}_1 = \mathbf{x})$ follows immediately by a simple calculation. \square

Proof of Theorem 6.5.* The estimator $\hat{\gamma}_n^H(\varepsilon)$ is $\sqrt{n(1-\tau_n)}$ -asymptotically normal by Corollary 6.3*. The proof of Theorem 5.1* then carries over to the present context without change. \square

D Consistency of the estimator $\hat{w}_n(\gamma, R)$

We here sketch the proof of the consistency of the estimator $\hat{w}_n(\gamma, R)$ of the asymptotic variance $w(\gamma, R) = \gamma^2\{1 + 2\sum_{t=1}^{\infty} R_t(1, 1)\}$, see Section 5*. Recall the notation

$$\hat{Z}_j = \hat{Z}_{j,n} = \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{\hat{F}_n(Y_t) \geq \tau_n\}$$

for $j = 0, 1, \dots, m_n - 1$, where $m_n = \lfloor n/\ell_n \rfloor$, $\ell_n = r_n + l_n$, and \widehat{F}_n is the empirical distribution function calculated on the whole sample. Define

$$Z_{j,n} = \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{F(Y_t) \geq \tau_n\}$$

which would be the “ideal” version of the $\widehat{Z}_{j,n}$ that one would use instead if the distribution function F were known. By stationarity, $Z_{j,n} \stackrel{d}{=} \sum_{i=1}^{r_n} \mathbb{1}\{F(Y_i) \geq \tau_n\}$, whose variance is asymptotically equivalent to $\{1 + 2 \sum_{t=1}^{\infty} R_t(1, 1)\} r_n(1 - \tau_n)$ by Proposition 2.1 in Drees (2003). Up to using a suitable law of large numbers to handle the empirical variance of the $Z_{j,n}$ (see Lemma C.7), and since $\widehat{\gamma}_n^H$ is a consistent estimator of γ as a consequence of Theorem 3.3*, the main task in proving consistency of $\widehat{w}_n(\gamma, R)$ is to show that

$$\begin{aligned} & \frac{1}{r_n(1 - \tau_n)} \left| \Sigma_n - \frac{1}{m_n} \sum_{j=1}^{m_n} \left(Z_{j,n} - \frac{1}{m_n} \sum_{i=1}^{m_n} Z_{i,n} \right)^2 \right| \\ &= \frac{1}{r_n(1 - \tau_n)} \left| \frac{1}{m_n} \sum_{j=1}^{m_n} \left(\widehat{Z}_{j,n} - \frac{1}{m_n} \sum_{i=1}^{m_n} \widehat{Z}_{i,n} \right)^2 - \frac{1}{m_n} \sum_{j=1}^{m_n} \left(Z_{j,n} - \frac{1}{m_n} \sum_{i=1}^{m_n} Z_{i,n} \right)^2 \right| \xrightarrow{\mathbb{P}} 0, \end{aligned}$$

that is, the empirical variances of the $Z_{j,n}$ and $\widehat{Z}_{j,n}$ are asymptotically equivalent. It is then sufficient to show that

$$\frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n}^2 - Z_{j,n}^2| \xrightarrow{\mathbb{P}} 0$$

and

$$\frac{1}{r_n(1 - \tau_n)} \left| \left(\frac{1}{m_n} \sum_{j=1}^{m_n} \widehat{Z}_{j,n} \right)^2 - \left(\frac{1}{m_n} \sum_{j=1}^{m_n} Z_{j,n} \right)^2 \right| \xrightarrow{\mathbb{P}} 0.$$

Since, by Lemma C.7,

$$\frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} Z_{j,n} \xrightarrow{\mathbb{P}} 1,$$

straightforward algebra shows that it is in fact enough to prove that

$$\frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n} - Z_{j,n}| \xrightarrow{\mathbb{P}} 0 \text{ and } \frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n}^2 - Z_{j,n}^2| \xrightarrow{\mathbb{P}} 0. \quad (\text{D.81})$$

The proof of the first of these two convergences gives the idea for the proof of the second one. The key observation is that, for any distribution function G with associated quantile function Q , the inequalities $Q(\alpha) \leq x$ and $\alpha \leq G(x)$ (where $\alpha \in (0, 1)$ and $x \in \mathbb{R}$) are equivalent. Consequently,

$$\widehat{Z}_{j,n} = \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq \widehat{q}_{\tau_n}\} \text{ and } Z_{j,n} = \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}\}.$$

Set $\varepsilon_n = \{n(1 - \tau_n)\}^{-1/4}$ and apply Theorem 3.3* to find that the event $A_n := \{q_{\tau_n}(1 - \varepsilon_n) \leq \widehat{q}_{\tau_n} \leq q_{\tau_n}(1 + \varepsilon_n)\}$ has probability converging to 1 as $n \rightarrow \infty$, so that it suffices to work on A_n to prove the required convergences in probability, which we do from now on. We get

$$\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 + \varepsilon_n)\} \leq \widehat{Z}_{j,n} \leq \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 - \varepsilon_n)\}. \quad (\text{D.82})$$

It follows that

$$\begin{aligned} & \frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n} - Z_{j,n}| \\ & \leq \frac{1}{m_n} \sum_{j=1}^{m_n} \left(\frac{1}{r_n(1 - \tau_n)} \sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 - \varepsilon_n)\} - \mathbb{1}\{Y_t \geq q_{\tau_n}(1 + \varepsilon_n)\} \right). \end{aligned}$$

Taking into account that $\varepsilon_n \rightarrow 0$, and therefore that $\overline{F}(q_{\tau_n}(1 \pm \varepsilon_n)) = (1 - \tau_n)(1 + o(1))$ by local uniformity of the regular variation property, the expectation of the above upper bound is immediately found to be $o(1)$, so that

$$\frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n} - Z_{j,n}| \xrightarrow{\mathbb{P}} 0.$$

Then recall (D.82) to find

$$\left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 + \varepsilon_n)\} \right)^2 \leq \widehat{Z}_{j,n}^2 \leq \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 - \varepsilon_n)\} \right)^2$$

and therefore

$$\begin{aligned} & \frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n}^2 - Z_{j,n}^2| \\ & \leq \frac{1}{m_n} \sum_{j=1}^{m_n} \frac{1}{r_n(1 - \tau_n)} \left\{ \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 - \varepsilon_n)\} \right)^2 - \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 + \varepsilon_n)\} \right)^2 \right\}. \quad (\text{D.83}) \end{aligned}$$

Note finally that

$$\mathbb{E} \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 \pm \varepsilon_n)\} \right)^2 = \text{Var} \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 \pm \varepsilon_n)\} \right) + O(\{r_n(1 - \tau_n)\}^2).$$

Recall that $\overline{F}(q_{\tau_n}(1 \pm \varepsilon_n)) = (1 - \tau_n)(1 + o(1))$, so, by a straightforward extension of the first half of Lemma C.2,

$$\mathbb{E} \left(\sum_{t=1+j\ell_n}^{r_n+j\ell_n} \mathbb{1}\{Y_t \geq q_{\tau_n}(1 \pm \varepsilon_n)\} \right)^2 = r_n(1 - \tau_n) \left\{ 1 + 2 \sum_{t=1}^{\infty} R_t(1, 1) + o(1) \right\}.$$

Thus the expectation of the upper bound in (D.83) is $o(1)$, from which we obtain

$$\frac{1}{r_n(1 - \tau_n)} \times \frac{1}{m_n} \sum_{j=1}^{m_n} |\widehat{Z}_{j,n}^2 - Z_{j,n}^2| \xrightarrow{\mathbb{P}} 0.$$

This completes the proof of (D.81) as required.

References

- Billingsley, P. (1995), *Probability and Measure*, John Wiley & Sons, New York.
- Bradley, R. C. (1985), ‘On the central limit question under absolute regularity’, *Annals of Probability* **13**, 1314–1325.
- Bradley, R. C. (2005), ‘Basic properties of strong mixing conditions. A survey and some open questions’, *Probability Surveys* **2**, 107–144.
- Daouia, A., Girard, S. & Stupfler, G. (2018), ‘Estimation of tail risk based on extreme expectiles’, *Journal of the Royal Statistical Society: Series B* **80**, 263–292.
- Daouia, A., Girard, S. & Stupfler, G. (2019), ‘Extreme M-quantiles as risk measures: From L^1 to L^p optimization’, *Bernoulli* **25**, 264–309.
- Daouia, A., Girard, S. & Stupfler, G. (2020), ‘Tail expectile process and risk assessment’, *Bernoulli* **26**, 531–556.
- de Haan, L. & Ferreira, A. (2006), *Extreme Value Theory: An Introduction*, Springer.
- Doukhan, P. (1994), *Mixing: Properties and Examples*, Springer.
- Drees, H. (2002), Tail empirical processes under mixing conditions, *in* H. Dehling, T. Mikosch & M. Sørensen, eds, ‘Empirical Process Techniques for Dependent Data’, Birkhäuser, Boston.
- Drees, H. (2003), ‘Extreme quantile estimation for dependent data, with applications to finance’, *Bernoulli* **9**, 617–657.
- Drees, H. & Rootzén, H. (2010), ‘Limit theorems for empirical processes of cluster functionals’, *Annals of Statistics* **38**, 2145–2186.

- Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag.
- Girard, S., Stupfler, G. & Usseglio-Carleve, A. (2021), ‘Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models’, *Annals of Statistics* **49**, 3358–3382.
- Ibragimov, I. A. (1962), ‘Some limit theorems for stationary processes’, *Theory of Probability and its Applications* **7**, 349–382.
- Knight, K. (1999), Epi-convergence in distribution and stochastic equi-semicontinuity, Technical report, University of Toronto.
- Leadbetter, M. R. & Rootzén, H. (1993), On central limit theory for families of strongly mixing additive set functions, in S. Cambanis & P. Sen, eds, ‘In Festschrift in Honour of Gopinath Kallianpur’, Springer.
- Mikosch, T. & Stărică, C. (2000), ‘Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process’, *Annals of Statistics* **28**, 1427–1451.
- Pratt, J. W. (1960), ‘On interchanging limits and integrals’, *The Annals of Mathematical Statistics* **31**, 74–77.
- Revuz, D. & Yor, M. (2005), *Continuous Martingales and Brownian Motion*, Springer-Verlag.
- Rio, E. (2017), *Asymptotic Theory of Weakly Dependent Random Processes*, Springer-Verlag.
- Rootzén, H. (2009), ‘Non-parametric estimation of tail dependence’, *Stochastic Processes and their Applications* **119**, 468–490.
- Rootzén, H., Leadbetter, M. R. & de Haan, L. (1998), ‘On the distribution of tail array sums for strongly mixing stationary sequences’, *The Annals of Applied Probability* **8**, 868–885.
- Schmidt, R. & Stadtmüller, U. (2006), ‘Non-parametric estimation of tail dependence’, *Scandinavian Journal of Statistics* **33**, 307–335.
- Stupfler, G. (2019), ‘On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails’, *Extremes* **22**, 749–769.