Appendix A Proofs

A.1 Proof of Theorem 1

We collect some notation. Let $\mathbb{N}_0 := \{0, 1, \ldots\}$, and $K > 0$ a positive constant that may change from line to line. An empty sum $\sum_i^j$ $(i > j)$ is defined to be zero. All $o_{(P)}$- and $\mathcal{O}_{(P)}$-symbols are to be understood with respect to $n \to \infty$.

We first give a rough outline of the proof of Theorem 1. In the first step we show that

$$\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}_{\alpha,n}(t)}{x_{\alpha,n}} - 1\right) = \frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}^U_\alpha(t)}{x^U_\alpha} - 1\right)\big[1 + o_P(1)\big] + o_P(1) \tag{A.1}$$

uniformly in $t \in [t_0, 1]$. To do so, write

$$\begin{aligned}\frac{\widehat{x}_{\alpha,n}(t)}{x_{\alpha,n}} - 1 &= \frac{\widehat{\mu}_{n+1} + \widehat{\sigma}_{n+1}\widehat{x}^U_\alpha}{\mu_{n+1} + \sigma_{n+1}x^U_\alpha} - 1 \\ &= \frac{\frac{\widehat{\mu}_{n+1} - \mu_{n+1}}{\sigma_{n+1}x^U_\alpha} + \frac{\widehat{\sigma}_{n+1}}{\sigma_{n+1}}\left(\frac{\widehat{x}^U_\alpha(t)}{x^U_\alpha} - 1\right) + \left(\frac{\widehat{\sigma}_{n+1}}{\sigma_{n+1}} - 1\right)}{\frac{\mu_{n+1}}{\sigma_{n+1}x^U_\alpha} + 1}.\end{aligned} \tag{A.2}$$

Lemmas 1 and 2 below will show that $(\widehat{\mu}_{n+1} - \mu_{n+1})/(\sigma_{n+1}x^U_\alpha) = \mathcal{O}_P(n^{-1/2})$ and $\widehat{\sigma}_{n+1}/\sigma_{n+1} = 1 + \mathcal{O}_P(n^{-1/2})$, respectively. Combined with the stationarity of μ_{n+1}, $\sigma_{n+1} \geq \sqrt{\omega^\circ}$ and the fact that $x^U_\alpha \to \infty$ as $n \to \infty$ (see Eq. (A.6) below), (A.2) leads to the desired approximation in (A.1). In the second step, the expansion (see Eq. (A.7) in Hoga (2017b))

$$\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}^U_\alpha(t)}{x^U_\alpha} - 1\right) = \sqrt{k}\left(\widehat{\gamma}(t) - \gamma\right) + o_P(1) \qquad \text{uniformly in } t \in [t_0, 1] \tag{A.3}$$

is the motivation for deriving limit theory for $\widehat{\gamma}$, which relies on arguments in Hoga (2017a) and Hoga (2017b).

We begin with the first step.

Lemma 1. *Suppose Assumptions 1–5 hold and $\alpha = \alpha_n \to 0$ as $n \to \infty$. Then*

$$\frac{\widehat{\mu}_{n+1} - \mu_{n+1}}{\sigma_{n+1}x^U_\alpha} = \mathcal{O}_P\left(\frac{1}{\sqrt{n}}\right).$$

Proof: From the recursions in (7) and (8) we obtain, as in the proof of Chan et al. (2007, Thm. 2) and using $\sqrt{n}$-consistency of $\widehat{\boldsymbol{\theta}}$, that

$$\widehat{\mu}_{n+1}(\widehat{\boldsymbol{\theta}}) - \mu_{n+1} = \mathcal{O}_P\left(\frac{1}{\sqrt{n}}\right). \tag{A.4}$$

Furthermore, almost surely (a.s.),

$$\sigma_{n+1} \geq \sqrt{\omega^\circ} > 0 \tag{A.5}$$

and, by regular variation of U and, e.g., Haan and Ferreira (2006, Prop. B.1.9.1),

$$x_\alpha^U = U(1/\alpha) \underset{(n\to\infty)}{\longrightarrow} \infty. \tag{A.6}$$

Combining (A.4) – (A.6) gives the desired result. ■

The next lemma deals with the estimation of the conditional variance σ_{n+1}. For $\eta > 0$, let

$$N_n(\eta) := \left\{\boldsymbol{\theta} : \ |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \leq \frac{\eta}{\sqrt{n}}\right\}, \qquad N_n^-(\eta) := N_n(\eta) \setminus \{\boldsymbol{\theta}^\circ\}.$$

Define

$$\widetilde{\sigma}_i^2(\boldsymbol{\theta}) = \omega^\circ + \sum_{j=1}^{p} \psi_j^\circ \varepsilon_{i-j}^2(\boldsymbol{\theta}) + \sum_{j=1}^{q} \beta_j^\circ \widetilde{\sigma}_{i-j}^2(\boldsymbol{\theta}).$$

Lemma 2. *Suppose the assumptions of Theorem 1 hold. Then*

$$\frac{\widehat{\sigma}_{n+1}(\widehat{\boldsymbol{\theta}})}{\sigma_{n+1}} = 1 + \mathcal{O}_P\left(\frac{1}{\sqrt{n}}\right).$$

Proof: For $\boldsymbol{\theta} \in N_n(\eta)$ write

$$\begin{aligned}\frac{\widehat{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2} - 1 &= \frac{\widehat{\sigma}_{n+1}^2(\boldsymbol{\theta}) - \sigma_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2} + \frac{\sigma_{n+1}^2(\boldsymbol{\theta}) - \widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2} + \frac{\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta}) - \sigma_{n+1}^2}{\sigma_{n+1}^2} \\ &=: (I) + (II) + (III).\end{aligned} \tag{A.7}$$

For (I) we have from Lemma A 1 below that for some $r \in (0,1)$

$$\sup_{\boldsymbol{\theta} \in N_n(\eta)} \left|\widehat{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right| \leq r^i V_i \qquad \text{for all } i \in \mathbb{N} \text{ and } n \text{ large enough},$$

where $V_i \geq 0$ with $\sup_{i\in\mathbb{N}} \mathrm{E}\left[V_i^\delta\right] < \infty$ for some $\delta > 0$. Hence, for n large enough,

$$\begin{aligned} P\left\{\sqrt{n} \sup_{\boldsymbol{\theta}\in N_n(\eta)} \left|\widehat{\sigma}_{n+1}^2(\boldsymbol{\theta}) - \sigma_{n+1}^2(\boldsymbol{\theta})\right| > \varepsilon\right\} &\leq P\left\{r^{n+1} V_{n+1} > \varepsilon/\sqrt{n}\right\} \\ &\leq (\varepsilon/\sqrt{n})^{-\delta}\, \mathrm{E}\left[r^{n+1} V_{n+1}\right]^\delta \\ &\leq K n^{\delta/2} (r^\delta)^{n+1} \sup_{i\in\mathbb{N}} \mathrm{E}\left[V_i^\delta\right] \underset{(n\to\infty)}{\longrightarrow} 0, \end{aligned}$$

where we have used Markov's inequality in the second step. Since $\sigma_{n+1}^2 \geq \omega^\circ > 0$ a.s., we conclude

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} \frac{\left|\widehat{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right|}{\sigma_{n+1}^2} = \mathcal{O}_P\left(n^{-1/2}\right). \tag{A.8}$$

Now consider (II). Write

$$\frac{\sigma_{n+1}^2(\boldsymbol{\theta}) - \widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2} = \frac{\sigma_{n+1}^2(\boldsymbol{\theta}) - \widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2(\boldsymbol{\theta})} \cdot \frac{\sigma_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2}. \tag{A.9}$$

Observe that by Markov's inequality and Kim and Lee (2016, Lem. 18) we obtain

$$P\left\{\sqrt{n}\sup_{\boldsymbol{\theta}\in N_n^-(\eta)}\frac{|\sigma_{n+1}^2(\boldsymbol{\theta})-\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})|}{\sigma_{n+1}^2(\boldsymbol{\theta})}>K\right\}\leq K^{-1}\,\mathrm{E}\left[\sup_{\boldsymbol{\theta}\in N_n^-(\eta)}\frac{\left|\sigma_{n+1}^2(\boldsymbol{\theta})-\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})\right|}{|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}^\circ|\sigma_{n+1}^2(\boldsymbol{\theta})}\sqrt{n}|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}^\circ|\right]$$
$$=K^{-1}\mathcal{O}(1)\underset{(K\to\infty)}{\longrightarrow}0,$$

i.e.,

$$\sup_{\boldsymbol{\theta}\in N_n^-(\eta)}\frac{\left|\sigma_{n+1}^2(\boldsymbol{\theta})-\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})\right|}{\sigma_{n+1}^2(\boldsymbol{\theta})}=\mathcal{O}_P(n^{-1/2}).\tag{A.10}$$

From the displays below Eq. (65) in Kim and Lee (2016) we get

$$\sup_{\boldsymbol{\theta}\in N_n^-(\eta)}\left|\frac{\sigma_{n+1}^2}{\sigma_{n+1}^2(\boldsymbol{\theta})}-1\right|=o_P(1).\tag{A.11}$$

Now (A.9)–(A.11) yield

$$\sup_{\boldsymbol{\theta}\in N_n^-(\eta)}\frac{\left|\sigma_{n+1}^2(\boldsymbol{\theta})-\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})\right|}{\sigma_{n+1}^2}=\mathcal{O}_P(n^{-1/2}).\tag{A.12}$$

Finally, consider (III). Eq. (63) in Kim and Lee (2016) implies

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)}\left|\frac{\widetilde{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2}-1\right|=\mathcal{O}_P(n^{-1/2}).\tag{A.13}$$

Combine (A.7) with (A.8), (A.12) and (A.13) to obtain

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)}\left|\frac{\widehat{\sigma}_{n+1}^2(\boldsymbol{\theta})}{\sigma_{n+1}^2}-1\right|=\mathcal{O}_P(n^{-1/2}).\tag{A.14}$$

Since, by $\sqrt{n}$-consistency from Assumption 4, $\widehat{\boldsymbol{\theta}}\in N_n(\eta)$ with probability approaching 1 as $n\to\infty$ followed by $\eta\to\infty$, the conclusion follows from (A.14). ■

For the second step our first goal is to prove

Theorem 2. *Suppose the assumptions of Theorem 1 hold. Then, for any $\nu\in[0,1/2)$ under a Skorohod construction,*

$$\sup_{t\in[t_0,1]}\sup_{y\geq 1}y^{\nu/\gamma}\left|t\sqrt{k}\left[\widehat{F}_n(t,y)-y^{-1/\gamma}\right]-\left[W(t,y^{-1/\gamma})-y^{-1/\gamma}W(t,1)\right]\right|\underset{(n\to\infty)}{\overset{\text{a.s.}}{\longrightarrow}}0,$$

where $W(\cdot,\cdot)$ is a Brownian sheet.

The proof of Theorem 2 requires the following two propositions, which are proved in Subsection A.2 of this Appendix. Let $D([a,b]\times[c,d])$ be the space of two-parameter càdlàg functions defined on $[a,b]\times[c,d]$, which is equipped with the multivariate extension of the Skorohod metric; see Davidson (1994).

Proposition 1. *Suppose that $\{U_i\}$ is a sequence of i.i.d. r.v.s with distributions satisfying Assumption 5. Then for any $\iota > 0$ and $\nu \in [0,1/2)$*

$$M_n(t,y) := y^{-\nu}\sqrt{k}\left(\frac{1}{k}\sum_{i=m_n}^{\lfloor nt\rfloor} I_{\{U_i>y^{-\gamma}U(n/k)\}} - yt\right) \underset{(n\to\infty)}{\overset{\mathcal{D}}{\longrightarrow}} y^{-\nu}W(t,y) \quad \text{in } D([t_0,1]\times[0,1+\iota]), \tag{A.15}$$

where $W(\cdot,\cdot)$ is a Brownian sheet and $\frac{0}{0} := 0$.

The next proposition is a weighted version of Proposition 5 in Kim and Lee (2016).

Proposition 2. *Under the assumptions of Theorem 1 we have that for any $\iota > 0$ and $\nu \in [0,1/2)$*

$$\sup_{y\in(0,1+\iota]}\sup_{t\in[0,1]} y^{-\nu}\sqrt{k}\left|\frac{1}{k}\sum_{i=m_n}^{\lfloor nt\rfloor}\left(I_{\{U_i>y^{-\gamma}U(n/k)\}} - I_{\{\widehat{U}_i>y^{-\gamma}U(n/k)\}}\right)\right| = o_P(1).$$

Proof of Theorem 2: Combine Propositions 1 and 2 to get

$$y^{-\nu}\sqrt{k}\left(\frac{1}{k}\sum_{i=m_n}^{\lfloor nt\rfloor} I_{\{\widehat{U}_i>y^{-\gamma}U(n/k)\}} - yt\right) \underset{(n\to\infty)}{\overset{\mathcal{D}}{\longrightarrow}} y^{-\nu}W(t,y) \quad \text{in } D([t_0,1]\times[0,1+\iota]).$$

Then the conclusion follows as in the proof of Corollary 1 in Hoga (2017a). ■

Proof of Theorem 1: Theorem 2 implies

$$t\sqrt{k}\left(\widehat{\gamma}(t)-\gamma\right) \underset{(n\to\infty)}{\overset{\mathcal{D}}{\longrightarrow}} \sigma_{\widehat{\gamma},\gamma}W(t) \quad \text{in } D[t_0,1] \tag{A.16}$$

for a generic extreme value index estimator based on the sample $\widehat{U}_{m_n},\ldots,\widehat{U}_{\lfloor nt\rfloor}$. For instance, for the Hill (1975) estimator $\widehat{\gamma}(t) = \widehat{\gamma}_H(t)$ (moments ratio estimator $\widehat{\gamma}(t) = \widehat{\gamma}_{MR}(t)$) we have $\sigma_{\widehat{\gamma},\gamma} = \gamma$ ($\sigma_{\widehat{\gamma},\gamma} = \sqrt{2}\gamma$); see Examples 3 and 4 in Hoga (2017a). To see (A.16) for the Hill estimator (other cases can be dealt with similarly), recall (19) and write for $\nu \in (0,1/2)$ (under a Skorohod construction)

$$\begin{aligned}
t\sqrt{k}\left(\widehat{\gamma}(t)-\gamma\right) &= \int_1^\infty t\sqrt{k}\left(\widehat{F}_n(t,y) - y^{-1/\gamma}\right)\frac{\mathrm{d}y}{y}\\
&\overset{\text{a.s.}}{=} \int_1^\infty \left(W(t,y^{-1/\gamma}) - y^{-1/\gamma}W(t,1) + o(1)y^{-\nu/\gamma}\right)\frac{\mathrm{d}y}{y}\\
&= \gamma\int_0^1 \left(W(t,z) - zW(t,1) + o(1)z^{\nu}\right)\frac{\mathrm{d}z}{z}\\
&= \gamma W(t) + o(1),
\end{aligned} \tag{A.17}$$

where, by calculating the covariance function, $W(t) := \int_0^1 \left[W(t,z) - zW(t,1)\right]\frac{\mathrm{d}z}{z}$ can easily be identified as a Brownian motion. Note that the weighted convergence in Theorem 2 was required in the step leading to (A.17).

Hence, by (A.1), (A.3) and (A.16), we get

$$\frac{t\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}_{\alpha,n}(t)}{x_{\alpha,n}}-1\right)\underset{(n\to\infty)}{\overset{\mathcal{D}}{\longrightarrow}}\sigma_{\widehat{\gamma},\gamma}W(t)\quad\text{in } D[t_0,1]. \tag{A.18}$$

Using that $\log(1+x)\sim x$ as $x\to 0$, claim (23) for $z=x$ follows. The other claims, (24) and (25), follow from $\widehat{\gamma}(1)\to\gamma$ in probability from (A.16) and a suitable application of the continuous mapping theorem.

Note that in principle one could use (A.18) directly to construct confidence intervals. Yet, Drees (2003) and Gomes and Pestana (2007) found a log-transformation, as in (23), to produce a better agreement of finite-sample and asymptotic distribution for extreme quantile estimators.

For $z=S$ an exact analogue of (A.2) holds, i.e.,

$$\frac{\widehat{S}_{\alpha,n}(t)}{S_{\alpha,n}}-1=\frac{\frac{\widehat{\mu}_{n+1}-\mu_{n+1}}{\sigma_{n+1}S_\alpha^U}+\frac{\widehat{\sigma}_{n+1}}{\sigma_{n+1}}\left(\frac{\widehat{S}_\alpha^U(t)}{S_\alpha^U}-1\right)+\left(\frac{\widehat{\sigma}_{n+1}}{\sigma_{n+1}}-1\right)}{\frac{\mu_{n+1}}{\sigma_{n+1}S_\alpha^U}+1}. \tag{A.19}$$

Then, similarly as before, Lemmas 1 and 2 yield

$$\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{S}_{\alpha,n}(t)}{S_{\alpha,n}}-1\right)=\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{S}_\alpha^U(t)}{S_\alpha^U}-1\right)+o_P(1).$$

So it suffices to prove, by (A.3) and (A.16),

$$\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{S}_\alpha^U(t)}{S_\alpha^U}-1\right)=\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}_\alpha^U(t)}{x_\alpha^U}-1\right)+o_P(1). \tag{A.20}$$

To do so, write

$$\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{S}_\alpha^U(t)}{S_\alpha^U}-1\right)=\frac{\sqrt{k}}{\log(k/(n\alpha))}\left(\frac{\widehat{x}_\alpha^U(t)}{x_\alpha^U}\cdot\frac{1-\gamma}{1-\widehat{\gamma}(t)}\cdot\frac{\frac{x_\alpha^U}{1-\gamma}}{S_\alpha^U}-1\right). \tag{A.21}$$

Note that by (A.16)

$$\frac{1-\gamma}{1-\widehat{\gamma}(t)}=\frac{1-\gamma}{1-\gamma+\mathcal{O}_P(k^{-1/2})}=1+\mathcal{O}_P(k^{-1/2}) \tag{A.22}$$

uniformly in $t\in[t_0,1]$. Due to Assumption 5 and Haan and Ferreira (2006, Thm. 2.3.9), we may apply Pan, Leng, and Hu (2013, Thm. 4.2) to obtain

$$\lim_{n\to\infty}\frac{1}{A\left(1/(1-F(x_\alpha^U))\right)}\left(\frac{S_\alpha^U}{x_\alpha^U}-\frac{1}{1-\gamma}\right)=\frac{1}{(1/\gamma-1)(1/\gamma-1-\rho)}. \tag{A.23}$$

Now by basic properties of inverse functions, we have

$$F(x_\alpha^U)=F(F^{-1}(1-\alpha))\geq 1-\alpha,$$

so that for sufficiently large n

$$\frac{1}{1-F(x_\alpha^U)} \geq \frac{1}{\alpha} \overset{(22)}{\geq} \frac{n}{k}.$$

We conclude, by Assumption 5, that

$$A\left(\frac{1}{1-F(x_\alpha^U)}\right) = \mathcal{O}(A(n/k)) = o(1/\sqrt{k}),$$

which, together with (A.23), implies

$$\frac{S_\alpha^U}{\frac{x_\alpha^U}{1-\gamma}} = 1 + o(1/\sqrt{k}). \tag{A.24}$$

Combining (A.21)–(A.24) gives (A.20). ■

A.2 Proofs of Propositions 1 & 2

Proof of Proposition 1: Theorem 1 in Hoga (2017a) implies

$$y^{-\nu}\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{\lfloor nt\rfloor} I_{\{U_i>y^{-\gamma}U(n/k)\}} - yt\right) \underset{(n\to\infty)}{\overset{\mathcal{D}}{\longrightarrow}} y^{-\nu}W(t,y) \quad \text{in } D([t_0,1]\times[0,1+\iota]), \tag{A.25}$$

where $W(\cdot,\cdot)$ is a standard Brownian sheet. Hence it suffices to prove

$$\sup_{y\in[0,1+\iota]} y^{-\nu}\sqrt{k}\left|\frac{1}{k}\sum_{i=1}^{m_n-1} I_{\{U_i>y^{-\gamma}U(n/k)\}}\right| = o_P(1).$$

Setting $t_n = m_n/n$, the left hand-side can be bounded by

$$\sup_{\substack{t\in[0,t_n]\\ y\in[0,1+\iota]}} y^{-\nu}\sqrt{k}\left|\frac{1}{k}\sum_{i=1}^{\lfloor nt\rfloor} I_{\{U_i>y^{-\gamma}U(n/k)\}} - yt\right| + \sup_{\substack{t\in[0,t_n]\\ y\in[0,1+\iota]}} y^{-\nu}\sqrt{k}yt =: (I) + (II).$$

The term (I) is $o_P(1)$ by (A.25) and the uniform continuity of $y^{-\nu}W(t,y)$; see also Hoga (2017a, Prop. 2). That $(II) = o_P(1)$ follows easily from $m_n = o(\sqrt{k})$. The conclusion follows. ■

It remains to prove Proposition 2. To do so, we assume throughout that the assumptions of Proposition 2 hold. The proof relies on results in Kim and Lee (2016). For ease of reference and to make the proofs more self-contained, we include Lemma 17 and the slightly modified Lemma 19 from Kim and Lee (2016).

Define $\mathcal{U}_i = \sigma(U_i, U_{i-1}, \ldots)$ as the smallest σ-field generated by the $U_i, U_{i-1}, \ldots$ and write $\mathrm{E}_i[\cdot] = \mathrm{E}[\cdot|\mathcal{U}_i]$.

Lemma A 1 (cf. Kim and Lee (2016, Lem. 17))**.** *There exist $r_0\in(0,1)$ and $\mathcal{U}_{i-1}$-measurable r.v.s*

$V_i \geq 0$, *s.t.* $\sup_{i\in\mathbb{N}} \mathrm{E}\,|V_i^v| < \infty$ *for some* $v > 0$, *and for sufficiently large* n

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} \max\Big\{|\widehat{\varepsilon}_i(\boldsymbol{\theta}) - \varepsilon_i(\boldsymbol{\theta})|,\ |\widehat{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})|\Big\} \leq r^i V_i \quad \textit{for all } i \in \mathbb{N}.$$

Before stating our version of Kim and Lee (2016, Lem. 19), we have to introduce some additional notation. In the following $\Pi^*_{1,n,i}$, $\Pi^*_{2,n,i}$ and $\Pi_{3,n,i}$ denote non-negative, $\mathcal{U}_{i-1}$-measurable r.v.s with

$$\limsup_{n\to\infty} \mathrm{E}\,|\Pi^*_{1,n,i}|^{v_0} < \infty, \quad \limsup_{n\to\infty} \mathrm{E}\,|\Pi^*_{2,n,i}|^{v_0/2} < \infty, \quad \limsup_{n\to\infty} \mathrm{E}\,|\Pi_{3,n,i}|^{v_0} < \infty \quad \text{for some } v_0 > 2; \tag{A.26}$$

see p. 264 in Kim and Lee (2016). Note that due to our more stringent moment assumption on the GARCH errors $\{\varepsilon_i\}$ in Assumption 3, we easily obtain stronger moment bounds in (A.26) for $\Pi^*_{1,n,i}, \Pi^*_{2,n,i}, \Pi_{3,n,i}$ than Kim and Lee (2016, p. 264). These are required to show (A.28) in Lemma A 2. For $\eta_0 \in \mathbb{R}$ and $y > 0$, define

$$\begin{aligned}
A_i(y,\boldsymbol{\theta}) &= I_{\left\{\widehat{U}_i(\boldsymbol{\theta}) > y^{-\gamma}U(n/k)\right\}},\\
A_i(y) &= I_{\left\{U_i > y^{-\gamma}U(n/k)\right\}},\\
A_i(y,\eta,\eta_0) &= I_{\left\{U_i\left(1+r_0^i\eta_0V_i\right)\left(1+\frac{\eta_0\Pi^*_{1,n,i}}{\sqrt{n}}+\mathrm{sgn}(\eta_0)\frac{\eta_0^2\Pi^*_{2,n,i}}{n}\right)+\frac{\eta_0\Pi_{3,n,i}}{\sqrt{n}}+r_0^i\eta_0V_i > y^{-\gamma}U(n/k)\right\}}.
\end{aligned}$$

Lemma A 2 (cf. Kim and Lee (2016, Lem. 19))**.** *Let* $\eta > 0$, $\epsilon_0 \in (0, 1/10)$, $r_1 \in (r_0, 1)$ *and* $m_n \underset{(n\to\infty)}{\longrightarrow} \infty$, $m_n < n$. *Then, there exists* $\eta_0 > 0$ *such that when* $\boldsymbol{\theta} \in N_n(\eta)$

$$w_iA_i(y,\eta,-\eta_0) \leq w_iA_i(y,\boldsymbol{\theta}) \leq w_iA_i(y,\eta,\eta_0) \qquad \textit{for } i = m_n,\dots,n, \text{ and } y \in [0, 1+\iota],$$

where

$$w_i := I_{\left\{\max\left\{\frac{|\eta_0|\Pi^*_{1,n,i}}{\sqrt{n}},\frac{|\eta_0|^2\Pi^*_{2,n,i}}{n},\frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}U(n/k)}\right\}<\epsilon_0,\ r_0^i|\eta_0|V_i<r_1^i\right\}},$$

and for $y \in [0, 1+\iota]$

$$\frac{n}{k}w_i\left|\mathrm{E}_{i-1}\left[A_i(y,\eta,\eta_0) - A_i(y)\right]\right| \leq Kw_iy\max\left\{r_1^i, \frac{|\eta_0|\Pi^*_{1,n,i}}{\sqrt{n}}, \frac{|\eta_0|^2\Pi^*_{2,n,i}}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}U(n/k)}\right\}. \tag{A.27}$$

Furthermore,

$$\sum_{i=m_n}^{n}(1-w_i) = o_P(1), \tag{A.28}$$

i.e., $w_{m_n} = \ldots = w_n = 1$ *with probability approaching 1 as* $n \to \infty$.

Proof: We only show how statements (A.27) and (A.28), that differ slightly from the formulation of Kim and Lee (2016), can be proved. The insertion of y in the right-hand side of (A.27) follows as detailed next. Write

$$
\begin{aligned}
A_i(y,\eta,\eta_0) &= I_{\left\{U_i > \left(1 - y^{\gamma}\frac{\eta_0 \Pi_{3,n,i}}{\sqrt{n}U(n/k)} - y^{\gamma}\frac{r_0^i \eta_0 V_i}{U(n/k)}\right)\left(1 + r_0^i \eta_0 V_i\right)^{-1}\left(1 + \frac{\eta_0 \Pi^*_{1,n,i}}{\sqrt{n}} + \operatorname{sgn}(\eta_0)\frac{\eta_0^2 \Pi^*_{2,n,i}}{n}\right)^{-1} y^{-\gamma}U(n/k)\right\}} \\
&=: I_{\left\{U_i > C y^{-\gamma} U(n/k)\right\}}
\end{aligned}
$$

Note that for $\widetilde{C} > 0$ we get from Assumption 5 and arguments in Einmahl, Haan, and Zhou (2016, p. 46) that

$$
\frac{n}{k} P\left\{U_i > \widetilde{C} y^{-\gamma} U(n/k)\right\} = \widetilde{C}^{-1/\gamma} y + y\mathcal{O}(A(n/k)) = y\left[\widetilde{C}^{-1/\gamma} + o(1)\right] \tag{A.29}
$$

uniformly in $y \in (0, K]$ for any $K > 0$. Then, if $w_i = 1$, we may use (A.29) to obtain

$$
\begin{aligned}
\frac{n}{k}\left|\mathrm{E}_{i-1}\left[A_i(y,\eta,\eta_0 - A_i(y))\right]\right| &= \frac{n}{k}\left| P\left\{U_i > C y^{-\gamma} U(n/k) \middle| \mathcal{U}_{i-1}\right\} - P\left\{U_i > y^{-\gamma} U(n/k) \middle| \mathcal{U}_{i-1}\right\}\right| \\
&\le K y (C^{-1/\gamma} - 1).
\end{aligned}
$$

Now (A.27) follows using the mean value theorem and the definition of C.

To prove (A.28), use Markov's inequality to deduce

$$
\begin{aligned}
P\left\{\sum_{i=m_n}^{n}(1-w_i) \ge \varepsilon\right\} &\le \frac{1}{\varepsilon}\mathrm{E}\left|\sum_{i=m_n}^{n}(1-w_i)\right| \le K\sum_{i=m_n}^{n}\mathrm{E}\left|1-w_i\right| \\
&\le K\sum_{i=m_n}^{n}\Bigg[P\left\{\frac{|\eta_0|\Pi^*_{1,n,i}}{\sqrt{n}} \ge \epsilon_0\right\} + P\left\{\frac{|\eta_0|^2\Pi^*_{2,n,i}}{n} \ge \epsilon_0\right\} \\
&\qquad + P\left\{\frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}U(n/k)} \ge \epsilon_0\right\} + P\left\{(r_0/r_1)^i |\eta_0| V_i \ge 1\right\}\Bigg] \\
&\le K\sum_{i=m_n}^{n}\Big[n^{-v_0/2}\,\mathrm{E}\left|\Pi^*_{1,n,i}\right|^{v_0} + n^{-v_0/2}\,\mathrm{E}\left|\Pi^*_{2,n,i}\right|^{v_0} \\
&\qquad + n^{-v_0/2} U(n/k)^{-v_0}\,\mathrm{E}\left|\Pi_{3,n,i}\right|^{v_0} + \left(r_0/r_1\right)^{vi}\,\mathrm{E}\left|V_i\right|^{v}\Big] \\
&= o(1),
\end{aligned}
$$

because of $v_0 > 2$ and (A.26). ■

We shall need one additional lemma.

Lemma 3. *Let $\eta > 0$, $\eta_0 \in \mathbb{R}$ and $B_i(y,\eta,\eta_0) := A_i(y,\eta,\eta_0) - A_i(y)$. Then, for any $\iota > 0$ and $\nu \in [0, 1/2)$*

$$
\sup_{y\in(0,1+\iota]}\sup_{t\in[0,1]}\frac{1}{\sqrt{k}}\frac{1}{y^{\nu}}\left|\sum_{i=m_n}^{\lfloor nt\rfloor} B_i(y,\eta,\eta_0)\right| = o_P(1).
$$

Proof: We adapt the proof of Lemma 15 in Kim and Lee (2016). Let $R_n = \left\lfloor \sqrt{k}\log n \right\rfloor$ and write $B_i(y) = B_i(y, \eta, \eta_0)$ for short. There is no loss of generality in setting $\iota = 0$. Decompose

$$(0,1] = \sum_{j=0}^{\infty} \left(y_{j+1}, y_j\right], \qquad \text{where} \quad y_j := y_{j,n} := e^{-j/R_n}.$$

Then note that for $y \in (y_{j+1}, y_j]$

$$\begin{aligned}
\frac{1}{\sqrt{k}}\frac{1}{y^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor} B_i(y) &\leq \frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{A_i(y_j,\eta,\eta_0) - A_i(y_{j+1})\right\} \\
&= \frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{B_i(y_j) - \mathrm{E}_{i-1} B_i(y_j)\right\} + \frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\mathrm{E}_{i-1} B_i(y_j) \\
&\qquad + \frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{A_i(y_j) - A_i(y_{j+1})\right\}
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{\sqrt{k}}\frac{1}{y^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor} B_i(y) &\geq \frac{1}{\sqrt{k}}\frac{1}{y_{j}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{A_i(y_{j+1},\eta,\eta_0) - A_i(y_{j})\right\} \\
&= \frac{1}{\sqrt{k}}\frac{1}{y_{j}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{B_i(y_{j+1}) - \mathrm{E}_{i-1} B_i(y_{j+1})\right\} + \frac{1}{\sqrt{k}}\frac{1}{y_{j}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\mathrm{E}_{i-1} B_i(y_{j+1}) \\
&\qquad + \frac{1}{\sqrt{k}}\frac{1}{y_{j}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{A_i(y_{j+1}) - A_i(y_{j})\right\}.
\end{aligned}$$

Hence, it suffices to show that

$$\max_{j\in\mathbb{N}_0}\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{B_i(y_j) - \mathrm{E}_{i-1} B_i(y_j)\right\}\right| = o_P(1), \tag{A.30}$$

$$\max_{j\in\mathbb{N}_0}\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\mathrm{E}_{i-1} B_i(y_j)\right| = o_P(1), \tag{A.31}$$

$$\max_{j\in\mathbb{N}_0}\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor}\left\{A_i(y_j) - A_i(y_{j+1})\right\}\right| = o_P(1). \tag{A.32}$$

First, we verify (A.30). Note for this that $\left\{w_i\left\{B_i(y_j) - \mathrm{E}_{i-1} B_i(y_j)\right\},\ \mathcal{U}_i\right\}$ is a martingale difference sequence (m.d.s.). For some $r_1 \in (r_0, 1)$ we have

$$\begin{aligned}
&P\left\{\max_{j\in\mathbb{N}_0}\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor} w_i\left\{B_i(y_j) - \mathrm{E}_{i-1} B_i(y_j)\right\}\right| > \varepsilon\right\} \\
&\leq \sum_{j=0}^{\infty} P\left\{\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{k}}\frac{1}{y_{j+1}^{\nu}}\sum_{i=m_n}^{\lfloor nt\rfloor} w_i\left\{B_i(y_j) - \mathrm{E}_{i-1} B_i(y_j)\right\}\right| > \varepsilon\right\}
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{j=0}^{\infty} \frac{1}{\varepsilon^2 k} \frac{1}{y_{j+1}^{2\nu}} \operatorname{E}\left[\sup_{t\in[0,1]} \left| \sum_{i=m_n}^{\lfloor nt \rfloor} w_i \left\{B_i(y_j) - \operatorname{E}_{i-1} B_i(y_j)\right\} \right|^2\right] \\
&\leq \sum_{j=0}^{\infty} \frac{1}{\varepsilon^2 k} \frac{1}{y_{j+1}^{2\nu}} 4 \cdot \operatorname{E}\left[\sum_{i=m_n}^{n} w_i \left\{B_i(y_j) - \operatorname{E}_{i-1} B_i(y_j)\right\}\right]^2 \\
&= \sum_{j=0}^{\infty} \frac{1}{\varepsilon^2 k} \frac{1}{y_{j+1}^{2\nu}} 4 \cdot \sum_{i=m_n}^{n} \operatorname{E}\left[w_i \left\{B_i(y_j) - \operatorname{E}_{i-1} B_i(y_j)\right\}\right]^2 \\
&\leq K \sum_{j=0}^{\infty} \frac{1}{k} \frac{1}{y_{j+1}^{2\nu}} \sum_{i=m_n}^{n} \left|\operatorname{E}\left[w_i B_i(y_j)\right]\right| \\
&\leq K \sum_{j=0}^{\infty} \frac{y_j}{y_{j+1}^{2\nu}} \frac{1}{n} \sum_{i=m_n}^{n} \operatorname{E}\max\left\{r_1^i, \frac{|\eta_0|\Pi_{1,n,i}^*}{\sqrt{n}}, \frac{|\eta_0|^2\Pi_{2,n,i}^*}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}U(n/k)}\right\} \\
&\leq K \sum_{j=0}^{\infty} \frac{e^{-j/R_n}}{e^{-(j+1)2\nu/R_n}} \left[\frac{1}{n}\sum_{i=m_n}^{n} r_1^i + \frac{1}{n^{3/2}} \sum_{i=m_n}^{n} \operatorname{E}|\Pi_{1,n,i}^*| \right. \\
&\qquad\qquad \left. + \frac{1}{n^2} \sum_{i=m_n}^{n} \operatorname{E}|\Pi_{2,n,i}^*| + \frac{1}{n^{3/2}U(n/k)} \sum_{i=m_n}^{n} \operatorname{E}|\Pi_{3,n,i}|\right] \\
&= \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \sum_{j=0}^{\infty} \left(e^{(2\nu-1)/R_n}\right)^j = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \frac{1}{1-e^{(2\nu-1)/R_n}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \frac{1}{\frac{1-2\nu+o(1)}{R_n}} \\
&= \mathcal{O}\left(\frac{R_n}{\sqrt{n}}\right) = \mathcal{O}\left(\frac{\sqrt{k}\log n}{\sqrt{n}}\right) = o(1),
\end{aligned}
$$

where the first step follows by subadditivity, the second by Chebyshev's inequality, the third by Doob's inequality (e.g., Davidson, 1994, Thm. 15.15), the fourth by uncorrelatedness of the zero-mean m.d.s., the sixth by the law of iterated expectations and Lemma A 2 and the eighth by (A.26). Note for the fifth step that $B_i(y_j) \in \{0,1\}$ or $\{-1,0\}$ according as $\eta_0 > 0$ or < 0, so that

$$
\operatorname{E}\left[w_i \left\{B_i(y_j) - \operatorname{E}_{i-1} B_i(y_j)\right\}\right]^2 = \operatorname{E}\left[w_i B_i^2(y_j)\right] - \operatorname{E}\left[w_i \left\{\operatorname{E}_{i-1} B_i(y_j)\right\}^2\right] \leq \left|\operatorname{E}\left[w_i B_i(y_j)\right]\right|.
$$

The result follows, since $w_{m_n} = \ldots = w_n = 1$ with probability tending to 1 by Lemma A 2.

As for (A.31), observe that by Lemma A 2 the left-hand side can be bounded by

$$
K \frac{1}{\sqrt{k}} \sum_{i=m_n}^{n} \frac{k}{n} w_i \max\left\{r_1^i, \frac{|\eta_0|\Pi_{1,n,i}^*}{\sqrt{n}}, \frac{|\eta_0|^2\Pi_{2,n,i}^*}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}U(n/k)}\right\} = o_P(1).
$$

The result follows as before.

Now we show (A.32). Write

$$
\frac{1}{\sqrt{k}} \frac{1}{y_{j+1}^{\nu}} \sum_{i=m_n}^{\lfloor nt \rfloor} \left\{A_i(y_j) - A_i(y_{j+1})\right\} = \frac{\sqrt{k}}{y_{j+1}^{\nu}} \left\{ \left(\frac{1}{k} \sum_{i=m_n}^{\lfloor nt \rfloor} A_i(y_j) - y_j t \right) \right.
$$

$$
\left. - \left(\frac{1}{k} \sum_{i=m_n}^{\lfloor nt \rfloor} A_i(y_{j+1}) - y_{j+1}t \right) + (y_j - y_{j+1})t \right\}
$$

$$
= \left(\frac{y_j}{y_{j+1}} \right)^{\nu} M_n(t, y_j) - M_n(t, y_{j+1}) + \sqrt{k} \frac{y_j - y_{j+1}}{y_{j+1}^{\nu}} t
$$

$$
= M_n(t, y_j) - M_n(t, y_{j+1}) + o_P(1),
$$

because $(y_j/y_{j+1})^\nu = 1 + o(1)$ and $M_n(t, y) = \mathcal{O}_P(1)$ (by Proposition 1) uniformly, and $\sqrt{k}/R_n = o(1)$. Furthermore, for any $\delta > 0$

$$
P \left\{ \max_{j \in \mathbb{N}_0} \sup_{t \in [0,1]} \left| M_n(t, y_j) - M_n(t, y_{j+1}) \right| > \varepsilon \right\}
$$

$$
\leq P \left\{ \sup_{|w_1 - w_2| < \delta} \sup_{t \in [0,1]} \left| M_n(t, w_1) - M_n(t, w_2) \right| > \varepsilon \right\}
$$

$$
\underset{(n \to \infty)}{\longrightarrow} P \left\{ \sup_{|w_1 - w_2| < \delta} \sup_{t \in [0,1]} \left| w_1^{-\nu} W(t, w_1) - w_2^{-\nu} W(t, w_2) \right| > \varepsilon \right\} \underset{(\delta \downarrow 0)}{\longrightarrow} 0
$$

by continuity of the sample paths of the Brownian sheet $W(\cdot, \cdot)$; see also Hoga (2017a, Prop. 2). Hence, (A.32) follows and the proof is complete. ∎

Proof of Proposition 2: The proof resembles that of Proposition 5 in Kim and Lee (2016). We give it here to convey the main idea. Let $\eta > 0$ and $m_n \to \infty$ with $m_n = o(\sqrt{k})$ as $n \to \infty$. Then, due to Lemma A 2, there exists $\eta_0 > 0$ such that with probability approaching 1,

$$
\frac{1}{y^\nu} \left\{ A_i(y, \eta, -\eta_0) - A_i(y) \right\} \leq \frac{1}{y^\nu} \left\{ A_i(y, \boldsymbol{\theta}) - A_i(y) \right\} \leq \frac{1}{y^\nu} \left\{ A_i(y, \eta, \eta_0) - A_i(y) \right\}
$$

for $\boldsymbol{\theta} \in N_n(\eta)$, $i = m_n, \dots, n$. Whence from Lemma 3,

$$
\sup_{\boldsymbol{\theta} \in N_n(\eta)} \sup_{y \in (0, 1+\iota]} \sup_{t \in [0,1]} \frac{1}{y^\nu} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \left\{ A_i(y, \boldsymbol{\theta}) - A_i(y) \right\} \right| = o_P(1).
$$

Because by Assumption 4, $\widehat{\boldsymbol{\theta}} \in N_n(\eta)$ with probability tending to 1 as $n \to \infty$ followed by $\eta \to \infty$, the result follows. ∎

References

Chan, N.H. et al. (2007). "Interval Estimation of Value-at-Risk Based on GARCH Models with Heavy-Tailed Innovations". *Journal of Econometrics* 137, pp. 556–576.

Davidson, James (1994). *Stochastic Limit Theory*. Oxford: Oxford University Press.

Drees, H. (2003). "Extreme Quantile Estimation for Dependent Data, with Applications to Finance". *Bernoulli* 9, pp. 617–657.

Einmahl, J.H.J., L. de Haan, and C. Zhou (2016). “Statistics of Heteroscedastic Extremes”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 78, pp. 31–51.

Gomes, M.I. and D. Pestana (2007). “A Sturdy Reduced-Bias Extreme Quantile (VaR) Estimator”. *Journal of the American Statistical Association* 102, pp. 280–292.

Haan, L. de and A. Ferreira (2006). *Extreme Value Theory.* New York: Springer.

Hill, B. (1975). “A Simple General Approach to Inference About the Tail of a Distribution”. *The Annals of Statistics* 3, pp. 1163–1174.

Hoga, Y. (2017a). “Change Point Tests for the Tail Index of β-Mixing Random Variables”. *Econometric Theory* 33, pp. 915–954.

Hoga, Y. (2017b). “Testing for Changes in (Extreme) VaR”. *Econometrics Journal* 20, pp. 23–51.

Kim, L. and S. Lee (2016). “On the Tail Index Inference for Heavy-Tailed GARCH-Type Innovations”. *Annals of the Institute of Statistical Mathematics* 68, pp. 237–267.

Pan, X., X. Leng, and T. Hu (2013). “The Second-Order Version of Karamata’s Theorem with Applications”. *Statistics & Probability Letters* 83, pp. 1397–1403.