Spatial dependence and space–time trend in extreme events

Article

Accepted Version

Einmahl, J. H. J., Ferreira, A., de Haan, L., Neves, C. ORCID: https://orcid.org/0000-0003-1201-5720 and Zhou, C. (2022) Spatial dependence and space–time trend in extreme events. Annals of Statistics, 50 (1). pp. 30-52. ISSN 2168-8966 doi: https://doi.org/10.1214/21-AOS2067 Available at https://centaur.reading.ac.uk/104770/

To link to this article DOI: http://dx.doi.org/10.1214/21-AOS2067

Publisher: Institute of Mathematical Statistics

SPATIAL DEPENDENCE AND SPACE-TIME TREND IN EXTREME EVENTS

BY JOHN H.J. EINMAHL[1], ANA FERREIRA[2], LAURENS DE HAAN[3], CLÁUDIA NEVES[4] AND CHEN ZHOU[5]

[1]Department of Econometrics and Operations Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands, j.h.j.einmahl@uvt.nl

[2]Instituto Superior Técnico (IST), University of Lisbon, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal, anafh@tecnico.ulisboa.pt

[3]Department of Statistics, University of Lisbon, , Bloco C6 Piso 4, Campo Grande, 1749-016 Lisboa, Portugal; Department of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands, ldehaan@ese.eur.nl

[4]Department of Mathematics and Statistics, Whiteknights Campus, University of Reading, Reading RG6 6AX, United Kingdom, c.neves@reading.ac.uk

[5]Econometric Institute, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands; Bank of The Netherlands, PO Box 98, 1000 AB Amsterdam, The Netherlands, zhou@ese.eur.nl

The statistical theory of extremes is extended to independent multivariate observations that are non-stationary both over time and across space. The non-stationarity over time and space is controlled via the scedasis (tail scale) in the marginal distributions. Spatial dependence stems from multivariate extreme value theory. We establish asymptotic theory for both the weighted sequential tail empirical process and the weighted tail quantile process based on all observations, taken over time and space. The results yield two statistical tests for homoscedasticity in the tail, one in space and one in time. Further, we show that the common extreme value index can be estimated via a pseudo-maximum likelihood procedure based on pooling all (non-stationary and dependent) observations. Our leading example and application is rainfall in Northern Germany.

1. Introduction. Within the domain of attraction of an extreme value distribution one can distinguish equivalence classes via the concept of scedasis (Einmahl et al., 2016; de Haan et al., 2015). The distribution function F has scedasis c (a positive, finite constant) with respect to the continuous distribution function F_0 if

$$\lim_{x \uparrow x^*} \frac{1-F(x)}{1-F_0(x)} = c,$$

where x^* is the right endpoint of F_0. The equivalence class consists of all probability distributions that have a scedasis with respect to the same distribution function F_0.

In a univariate context – with independent observations – a natural estimator of the integrated scedasis function and its asymptotic properties are known (Einmahl et al., 2016). The present paper sets out to extend the results to a situation with multivariate observations as follows. Our leading example concerns daily rainfall in Northern Germany, with measurements taken at 49 stations over 84 years. On any day, we have a 49-dimensional observation across all stations, which potentially possesses spatial dependence. For each station, the distributions of extreme rainfall over 84 years may vary according to a scedasis function.

MSC2020 subject classifications: Primary 62G32, 62G30, 62G05, 62G10, 62G20; secondary 60F17, 60G70.

Keywords and phrases: multivariate extreme value statistics, non-identical distributions, sequential tail empirical process, testing.

Consider *independent* random vectors $\big(X_{i,1}, X_{i,2}, \ldots, X_{i,m}\big)$, $i = 1, 2, \ldots, n$. In the rainfall context, "m" is the number of stations and "n" is the number of time points (in days). A key assumption is the existence of scedasis: for some continuous distribution function F_0 in the domain of attraction of an extreme value distribution

$$\lim_{x \uparrow x^*} \frac{1 - F_{i,j}(x)}{1 - F_0(x)} = c\Big(\frac{i}{n}, j\Big) \in (0, \infty), \tag{1.1}$$

holds for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, where $F_{i,j}$ is the distribution function of $X_{i,j}$ and where for each j the scedasis $c(\cdot, j)$ is a positive continuous function on [0,1]. In order to ensure that the function c is uniquely defined we impose the condition

$$\sum_{j=1}^{m} C_j(1) = 1,$$

where for $0 \le t \le 1$ and $j = 1, 2 \ldots, m$,

$$C_j(t) := \frac{1}{m} \int_0^t c(u, j)\, du;$$

C_j is called the integrated scedasis. The scedasis $c(i/n, j)$ can be interpreted as the relative frequency of extremes at time i and location j. We assume $F_0 \in \mathcal{D}(G_\gamma)$, $\gamma \in \mathbb{R}$, i.e., F_0 is in the max-domain of attraction of $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$, $1 + \gamma x > 0$. As a consequence of (1.1), γ is now the common extreme value index: $F_{i,j} \in \mathcal{D}(G_\gamma)$, $i = 1, \ldots, n$; $j = 1, \ldots, m$.

Estimators $\widehat{C}_j$ for C_j will be introduced and their joint asymptotic distribution derived. This will enable us to perform various tests. For each station j we test whether the scedasis is changing over time. We also test whether the $C_j(1)$ are different i.e. if there are real differences in extreme rainfall over space.

Let F_i be the distribution function of $\big(X_{i,1}, X_{i,2}, \ldots, X_{i,m}\big)$. Assume that the distribution function $F_{i,j}$ of $X_{i,j}$ is continuous and let $U_{i,j}(t) := F_{i,j}^{\leftarrow}\big(1 - 1/t\big)$, where the arrow indicates the generalized inverse function. To model the dependence, we further assume that

$$\widetilde{F}(x_1, x_2, \ldots, x_m) := F_i\big(U_{i,1}(x_1), U_{i,2}(x_2), \ldots, U_{i,m}(x_m)\big) \tag{1.2}$$

does not depend on i and is in the domain of attraction of a multivariate extreme value distribution (de Haan and Ferreira, 2006, Chapter 6). As a consequence, the multivariate tail dependence structure does not depend on i. Let R_{j_1,j_2} denote the tail copula of the components j_1 and j_2:

$$R_{j_1,j_2}(x, y) = \lim_{t \downarrow 0} \frac{1}{t} P\big(1 - F_{i,j_1}(X_{i,j_1}) \le tx,\, 1 - F_{i,j_2}(X_{i,j_2}) \le ty\big),$$

for $(x, y) \in [0, \infty]^2 \backslash \{(\infty, \infty)\}$. Note that a tail copula is continuous, non-decreasing, and homogeneous of order one.

As in Einmahl et al. (2016), the estimator of C_j could be the number of exceedances over a high empirical quantile at station j. But, since we want to compare the C_j's, we want to use the same threshold for all rain stations. Consequently the common threshold will be a high empirical quantile of all $N := n \times m$ observations taken together. Let $X_{N-k:N}$ be the $(N-k)$-th order statistic of the observations $\big\{X_{i,j}\big\}_{i=1,\, j=1}^{n \quad m}$. We define the estimator

$$\widehat{C}_j(t) := \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_{\left\{X_{i,j} > X_{N-k:N}\right\}}$$

where k is an intermediate sequence: $k = k(n) \to \infty$, $k(n)/n \to 0$, as $n \to \infty$; the sum is over all integers $1 \le i \le nt$.

In this paper we make the following four contributions.

1. We establish the joint asymptotic behavior of $\{\widehat{C}_j(t)\}_{t\in[0,1]}$, $j=1,\dots,m$.
2. We test $H_0: C_j(1)=\frac{1}{m}$ for all $j=1,\dots,m$, i.e., the total integrated scedasis is constant over the various locations. We perform the test by checking whether the limit vector (in distribution) of

$$\left(\sqrt{k}\left(\widehat{C}_1(1)-\frac{1}{m}\right),\sqrt{k}\left(\widehat{C}_2(1)-\frac{1}{m}\right),\dots,\sqrt{k}\left(\widehat{C}_m(1)-\frac{1}{m}\right)\right)$$

has mean zero. This will be done via an adapted χ^2-test.
3. We test $H_{0,j}: C_j(t)=tC_j(1)$ for $0\le t\le 1$ and some given $j\in\{1,\dots,m\}$, i.e., the scedasis $c(\cdot,j)$ is constant over time. Since, under $H_{0,j}$, the limit in distribution of the process $\{\sqrt{k}(\widehat{C}_j(t)-t\widehat{C}_j(1)\}_{0\le t\le 1}$ is essentially a Brownian bridge, we can use, e.g., a Kolmogorov-Smirnov-type statistic.
4. We establish the asymptotic behavior of the pseudo-maximum likelihood estimator of γ based on all $n\times m$ observations.

Crucial for these results is a joint weighted Gaussian approximation of the m sequential tail empirical processes as well as one for the tail quantile process based on all $n\times m$ observations, for general $\gamma\in\mathbb{R}$. The main challenge is to establish these results in the above setting of observations that are dependent and have different distributions. For the tail empirical processes this is achieved by disentangling these two complications, whereas for the tail quantile process, after aggregating the m tail empirical processes, a delicate proof is required to deal with the weight functions.

An early paper where in the univariate case a linear trend in the parameters of the limit distribution is studied is Davison and Smith (1990). Various models for spatial extremes, all quite different from the setup in the present paper, are reviewed in Davison et al. (2012), see also Coles and Tawn (1996) for a specific rainfall model. The assumption of constant tail dependence, cf. (1.2), is tested in Bücher et al. (2015).

The outline of the paper is as follows. Section 2 gives a detailed account of the conditions and the ensuing results. These results are applied to the mentioned rainfall data in Section 3. Proofs are collected in Section 4 and partly deferred to the Supplementary Material, along with a simulation study showing the performance of the proposed estimation and testing procedures and a validation of the assumptions for the data application.

2. Results.

Throughout the paper we assume the following conditions.

(i) Multivariate dependence: Assume that $\widetilde{F}$, defined in (1.2), does not depend on i and is in the domain of attraction of a multivariate extreme value distribution.

(ii) Sharpening of the scedasis condition (1.1): Assume that there exists an eventually decreasing function A_1 with $\lim_{t\to\infty}A_1(t)=0$ such that

$$\sup_{n\ge 1}\max_{i,j}\left|\frac{1-F_{i,j}(x)}{1-F_0(x)}-c\left(\frac{i}{n},j\right)\right|=O\left(A_1\left(\frac{1}{1-F_0(x)}\right)\right),\quad \text{as } x\uparrow x^*.$$

(iii) Second order condition for F_0: Write $U_0(t):=F_0^{\leftarrow}(1-1/t)$. There exists $\gamma\in\mathbb{R}$, $\rho<0$ and functions $\tilde{a}_0$, positive, and A_0 not changing sign eventually satisfying $\lim_{t\to\infty}A_0(t)=0$ such that for all $x>0$,

$$\lim_{t\to\infty}\frac{\frac{U_0(tx)-U_0(t)}{\tilde{a}_0(t)}-\frac{x^\gamma-1}{\gamma}}{A_0(t)}=\Psi_{\gamma,\rho}(x):=\begin{cases}\frac{x^{\gamma+\rho}-1}{\gamma+\rho}, & \gamma+\rho\neq 0,\\ \log x, & \gamma+\rho=0,\end{cases}\tag{2.1}$$

cf. Corollary 2.3.5 in de Haan and Ferreira (2006).

(iv) Conditions on the intermediate sequence k: Assume, as $n \to \infty$,

$$k \to \infty, \quad k/n \to 0,$$
$$\sqrt{k} A_1\Big(q\frac{n}{k}\Big) \to 0, \text{ for all } q > 0, \quad \sqrt{k} A_0\Big(\frac{n}{k}\Big) \to 0$$
$$\sqrt{k} \sup_{|u-v| \le \frac{1}{n}} \big|c(u,j) - c(v,j)\big| \to 0, \quad \text{for } j = 1, 2, \ldots, m.$$

Throughout the paper, quantities regarding $\gamma = 0$ should be read as the limit as γ tends to zero. In particular, in (2.1) if $\gamma = 0$, $(x^\gamma - 1)/\gamma$ should be read as $\log x$. Likewise, $(1+\gamma x)^{1/\gamma}$ is meant to be e^x. Write $\gamma_+ = \gamma \vee 0$ and $\gamma_- = \gamma \wedge 0$. If $\gamma \le 0$, $-1/\gamma_+$ means $-\infty$; if $\gamma \ge 0$, $1/(-\gamma_-)$ means ∞.

We begin with presenting two fundamental approximations, which are the basis for the main results (Theorem 2.3, Corollaries 2.4 and 2.5, Theorem 2.6), but they are also of independent interest. For the following theorem we need to bound the difference between the fraction to the left in (2.1) and its limit by a uniform inequality. For details, see the inequality (2.3.19) in Corollary 2.3.7, obtained from (B.3.19) in Theorem B.3.10 of de Haan and Ferreira (2006). Such an inequality holds provided that the functions U_0, $\tilde{a}_0$ and $\Psi_{\gamma,\rho}$ are replaced by appropriately chosen versions b_0, a_0 and $\overline{\Psi}_{\gamma,\rho}$ given in Corollary 2.3.7.

THEOREM 2.1. *Assume conditions (i)-(iv). Let $x_0 > -1/\gamma_+$; set $x_1 := 1/(-\gamma_-)$.*

a) Tail empirical distribution functions

Using a Skorokhod construction, for $0 \le \eta < 1/2$, as $n \to \infty$, it holds almost surely,

$$\max_{1 \le j \le m} \sup_{x_0 \le x_j < x_1} \sup_{0 \le t_j \le 1} (1+\gamma x_j)^{\eta/\gamma} \Bigg| \sqrt{k}\Bigg(\frac{1}{k} \sum_{i=1}^{nt_j} \mathbb{1}_{\Big\{ \frac{X_{i,j} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x_j \Big\}} - (1+\gamma x_j)^{-1/\gamma} C_j(t_j) \Bigg) - W_j\Big((1+\gamma x_j)^{-1/\gamma},\, C_j(t_j)\Big) \Bigg| \longrightarrow 0, \tag{2.2}$$

where $(W_1, \ldots, W_m)$ is a mean zero Gaussian vector of bivariate Wiener processes W_j. Its covariances follow from the covariance matrix $\Sigma = \Sigma(s_1, s_2, t_1, t_2)$ with entries

$$\begin{aligned} \sigma_{j_1,j_2}(s_1,s_2,t_1,t_2) :=& Cov\Big(W_{j_1}\big(s_1, C_{j_1}(t_1)\big),\, W_{j_2}\big(s_2, C_{j_2}(t_2)\big)\Big) \\ =& \frac{1}{m}\int_0^{t_1 \wedge t_2} R_{j_1,j_2}\Big(s_1 c(u,j_1),\, s_2 c(u,j_2)\Big)\, du, \end{aligned} \tag{2.3}$$

for $1 \le j_1, j_2 \le m$.

b) Tail empirical quantile function

With any $\varepsilon > 0$ and $X_{1:N} \le \ldots \le X_{N:N}$ the order statistics of the sample $\{X_{i,j}\}_{i,j}$ of all N observations, we have, for $T > 0$, as $n \to \infty$,

$$\sup_{\frac{1}{2k} \le s \le T} s^{-1/2+\varepsilon} \Bigg| s^{\gamma+1}\sqrt{k}\Bigg(\frac{X_{N-[ks]:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} - \frac{s^{-\gamma}-1}{\gamma} \Bigg) - \sum_{j=1}^m W_j\big(s, C_j(1)\big) \Bigg| \xrightarrow{P} 0. \tag{2.4}$$

We consider the two boundary cases for the dependence structure. In case of tail independence, that is $R_{j_1,j_2} \equiv 0$, for all pairs (j_1, j_2), $j_1 \ne j_2$, the bivariate Wiener processes $W_1, \ldots, W_m$ are independent, and (2.2) for $\gamma > 0$ specializes to a generalization of the results in Einmahl et al. (2016) where independence of the data is assumed. Further specializing the situation to having N i.i.d. data, Theorem 2.1 recovers Theorem 2.4.2 and generalizes Theorem 5.1.2 in de Haan and Ferreira (2006), on tail quantile and tail empirical processes, respectively. In case of complete tail dependence between two locations, that is $R_{j_1,j_2}(x,y) = x \wedge y$, we obtain $\sigma_{j_1,j_2}(s_1,s_2,t_1,t_2) = \frac{1}{m}\int_0^{t_1\wedge t_2}(s_1 c(u,j_1)) \wedge (s_2 c(u,j_2))du$. In case $c(\cdot,j_1) = c(\cdot,j_2)$, the covariance simplifies to $(s_1 \wedge s_2)C_{j_1}(t_1 \wedge t_2)$, which means that the processes W_{j_1} and W_{j_2} are the same.

COROLLARY 2.2. *Under the conditions and in the setup of Theorem 2.1, with $\varepsilon > 0$,*

$$\sup_{\frac{1}{2k}\le s\le T} \left(1 \wedge s^{\gamma+1/2+\varepsilon}\right) \Bigg|\sqrt{k}\Big(\frac{X_{N-[ks]:N} - X_{N-k:N}}{a_0(\frac{N}{k})} - \frac{s^{-\gamma}-1}{\gamma}\Big) - \sum_{j=1}^{m}\Big(s^{-\gamma-1}W_j\big(s, C_j(1)\big) - W_j\big(1, C_j(1)\big)\Big)\Bigg| \xrightarrow{P} 0. \tag{2.5}$$

As a result we get the joint asymptotic behavior of the $\widehat{C}_j$, $j = 1,\dots,m$.

THEOREM 2.3. *Under the conditions and in the setup of Theorem 2.1, as $n \to \infty$,*

$$\max_{1\le j\le m}\sup_{0<t\le 1}\left|\sqrt{k}\big(\widehat{C}_j(t) - C_j(t)\big) - \Big\{W_j\big(1, C_j(t)\big) - C_j(t)\sum_{r=1}^{m} W_r\big(1, C_r(1)\big)\Big\}\right| \xrightarrow{P} 0, \tag{2.6}$$

Moreover, we have the uniform consistency of the estimator of the covariance matrix Σ as follows. For $T > 0$ and $j_1 \ne j_2$, as $n \to \infty$,

$$\sup_{0\le t\le 1}\ \sup_{0\le s_1,s_2\le T}\Bigg|\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{X_{i,j_1}>X_{N-[ks_1]:N},\, X_{i,j_2}>X_{N-[ks_2]:N}\right\}} - \frac{1}{m}\int_0^t R_{j_1,j_2}\Big(s_1 c(u,j_1),\ s_2 c(u,j_2)\Big)\,du\Bigg| \xrightarrow{P} 0. \tag{2.7}$$

Now we proceed with the two aforementioned tests. First we discuss the testing problem

$$\begin{cases} H_0 : C_j(1) = \frac{1}{m}, & \text{for all } j = 1,\dots,m, \\ H_1 : C_j(1) \ne \frac{1}{m}, & \text{for some } j = 1,\dots,m. \end{cases}$$

Let $\mathbb{1}_m$ be the m-unit vector, I_m the identity matrix of dimension m and define $M := I_m - \frac{1}{m}\mathbb{1}_m\mathbb{1}_m'$. From Theorem 2.3 and under H_0, $D = \sqrt{k}\Big(\widehat{C}_1(1) - \frac{1}{m},\dots,\widehat{C}_m(1) - \frac{1}{m}\Big)'$ is asymptotically m-variate normal with zero mean vector and covariance matrix $M\Sigma_1 M'$, where $\Sigma_1 = \Sigma(1,1,1,1)$. Assume that Σ_1 is invertible. Then $\operatorname{rank}(M\Sigma_1 M') = \operatorname{rank}(M) = m-1$.

We therefore confine attention to the first $m-1$ components of D denoted by D_{m-1}, which has an asymptotic covariance matrix $(M\Sigma_1 M')_{m-1}$. Here for an $m\times m$ matrix A, the notation A_{m-1} refers to the matrix consisting of the first $m-1$ rows and $m-1$ columns of A. Finally, we define the test statistic

$$T_n := D_{m-1}'\left((M\hat{\Sigma}_1 M')_{m-1}\right)^{-1} D_{m-1},$$

with Σ_1 estimated via the empirical counterpart given in (2.7). From Theorem 2.3 we immediately get the asymptotic behavior of T_n under H_0.

COROLLARY 2.4. *Assume that Σ_1 is invertible. Then under H_0, $T_n \xrightarrow{d} \chi^2_{m-1}$, as $n \to \infty$.*

Next we consider, for $j \in \{1,\dots,m\}$, the testing problem $H_{0,j}: C_j(t) = tC_j(1)$, $0 \le t \le 1$, and $H_{1,j}$ that this is not the case. This null hypothesis means that the scedasis for station j, $c(\cdot,j)$, is constant over time. We can use test statistics of the Kolmogorov-Smirnov-type or Cramér-von Mises-type based on the process $\sqrt{k}\big(\widehat{C}_j(t) - t\widehat{C}_j(1)\big)/\sqrt{\widehat{C}_j(1)}$, $0 \le t \le 1$.

COROLLARY 2.5. *Fix $j \in \{1,\dots,m\}$. Under the hypothesis that $C_j(t) = tC_j(1)$, for $0 \le t \le 1$,*

$$\left\{\sqrt{k\,\widehat{C}_j(1)}\left(\frac{\widehat{C}_j(t)}{\widehat{C}_j(1)} - t\right)\right\}_{t\in[0,1]} \xrightarrow{d} \{B(t)\}_{t\in[0,1]},$$

with B a Brownian bridge.

Finally, we introduce the pseudo-maximum likelihood estimator (MLE) of $(\gamma, a_0(\frac{N}{k}))$. The estimator is based on the $N = n \times m$ dependent and non-identically distributed observations (as described in Section 1) and in particular on the order statistics of excesses $\{X_{N-i+1:N} - X_{N-k:N}\}_{i=1}^{k}$. The (pseudo) ML procedure is based on the assumption that these order statistics are taken from the (limiting) generalized Pareto distribution $1 - (1+\gamma x/\sigma)^{-1/\gamma}$, $\gamma \in \mathbb{R}$, $\sigma > 0$, where σ represents the scale component $a_0(\frac{N}{k})$; for more details see de Haan and Ferreira (2006), Section 3.4. This leads to the log-likelihood

$$\ell(\gamma,\sigma,x) = -\log\sigma - \left(1+\frac{1}{\gamma}\right)\log\left(1+\gamma\,\frac{x}{\sigma}\right), \quad 0 < x < \frac{\sigma}{-\gamma_-} \tag{2.8}$$

(for $\gamma = 0$ the formula is interpreted as $-\log\sigma - x/\sigma$). The pseudo log-likelihood based on the above sample can be written as, with parameter space $(\gamma,\sigma) \in \mathbb{R}\times(0,\infty)$,

$$\begin{aligned} L_{N,k}(\gamma,\sigma_{N/k}) &= \sum_{i=1}^{k} \ell(\gamma,\sigma_{N/k}, X_{N-i+1:N} - X_{N-k:N}) \\ &= k\int_0^1 \ell(\gamma,\sigma_{N/k}, X_{N-[ks]:N} - X_{N-k:N})\,ds \\ &= k\int_0^1 \ell\left(\gamma, \frac{\sigma_{N/k}}{a_0(N/k)}, \frac{X_{N-[ks]:N} - X_{N-k:N}}{a_0(N/k)}\right)ds - k\log a_0(N/k). \end{aligned} \tag{2.9}$$

Generally $(\widehat{\gamma}, \widehat{\sigma}_{N/k})$ is an MLE if it is a local maximizer of $L_{N,k}(\gamma,\sigma_{N/k})$ solving the score equations,

$$\begin{cases} \frac{\partial}{\partial\gamma} L_{N,k}(\gamma,\sigma_{N/k}) = 0 \\ \frac{\partial}{\partial\sigma} L_{N,k}(\gamma,\sigma_{N/k}) = 0. \end{cases}$$

In the following we highlight γ_0 as the true unknown parameter value.

THEOREM 2.6. *Under conditions (i)-(iv) with $\gamma_0 > -1/2$, with probability tending to 1, there exists a unique sequence of estimators $(\widehat{\gamma}_n, \widehat{a}_0(N/k))$, maximizing (2.9), for which*

$$\sqrt{k}\left(\widehat{\gamma}_n - \gamma_0, \frac{\widehat{a}_0(N/k)}{a_0(N/k)} - 1\right) \xrightarrow{d} N(0, I_{\gamma_0}^{-1}\Sigma_{\gamma_0} I_{\gamma_0}^{-1})$$

where

$$I_{\gamma_0}^{-1} = \begin{bmatrix} (\gamma_0+1)^2 & -(\gamma_0+1) \\ -(\gamma_0+1) & 2(\gamma_0+1) \end{bmatrix}$$

and Σ_{γ_0} is the covariance matrix of the random vector

$$\begin{bmatrix} \sum_{j=1}^{m}\int_0^1 \left(s^{\gamma_0}\gamma_0^{-1}(1-s^{\gamma_0})-s^{2\gamma_0}\right)\left\{s^{-\gamma_0-1}W_j(s,C_j(1))-W_j(1,C_j(1))\right\}ds \\ \sum_{j=1}^{m}(1+\gamma_0)\int_0^1 s^{2\gamma_0}\left\{s^{-\gamma_0-1}W_j(s,C_j(1))-W_j(1,C_j(1))\right\}ds \end{bmatrix}$$

with $\{W_j\}_{j=1}^m$ from Theorem 2.1.

REMARK 1. The covariance matrix Σ_{γ_0} can be calculated as follows: let U be a $(1\times 2m)$ vector with the first $i=1,\dots,m$, components as

$$\gamma_0^{-1}\int_0^1 \left(s^{\gamma_0}-(1+\gamma_0)s^{2\gamma_0}\right)\left\{s^{-\gamma_0-1}W_i(s,C_i(1))-W_i(1,C_i(1))\right\}ds,$$

and the remaining $i=m+1,\dots,2m$ components as

$$(1+\gamma_0)\int_0^1 s^{2\gamma_0}\left\{s^{-\gamma_0-1}W_i(s,C_i(1))-W_i(1,C_i(1))\right\}ds.$$

The covariance matrix of U has entries $\tau_{ij}, i=1,\dots,2m, j=1,\dots,2m$, given by,

$$\tau_{ii}=\frac{2+6\gamma_0+5\gamma_0^2}{(1+\gamma_0)^2(1+2\gamma_0)^2}C_i(1),\quad i=1,\dots,m;$$

$$\tau_{ii}=\left(\frac{1+\gamma_0}{1+2\gamma_0}\right)^2 C_{i-m}(1),\quad i=m+1,\dots,2m;$$

$$\tau_{i,i+m}=\tau_{i+m,i}=\frac{1+\gamma_0}{(1+2\gamma_0)^2}C_i(1),\quad i=1,\dots,m;$$

$$\tau_{ij}=\int_0^1\int_0^1 f(s)f(t)r_{ij}(s,t)-2f(s)g(t)r_{ij}(s,1)+g(s)g(t)r_{ij}(1,1)\,ds\,dt,$$
$$i\neq j, i,j=1,\dots,m;$$

$$\tau_{ij}=(1+\gamma_0)^2\left\{\int_0^1\int_0^1 s^{\gamma_0-1}t^{\gamma_0-1}r_{i-m,j-m}(s,t)-2s^{\gamma_0-1}t^{2\gamma_0}r_{i-m,j-m}(s,1)\right.$$
$$\left.+s^{2\gamma_0}t^{2\gamma_0}r_{i-m,j-m}(1,1)\,ds\,dt\right\},$$
$$i\neq j, i=m+1,\dots,2m, j=m+1,\dots,2m;$$

$$\tau_{ij}=(1+\gamma_0)\left\{\int_0^1\int_0^1 f(s)t^{\gamma_0-1}r_{i,j-m}(s,t)-f(s)t^{2\gamma_0}r_{i,j-m}(s,1)\right.$$
$$\left.-g(s)t^{\gamma_0-1}r_{i,j-m}(1,t)+g(s)t^{2\gamma_0}r_{i,j-m}(1,1)\,ds\,dt\right\},$$
$$j\neq i+m, i=1,\dots,m,\ j=m+1,\dots,2m;$$

$$\tau_{ij}=(1+\gamma_0)\left\{\int_0^1\int_0^1 f(s)t^{\gamma_0-1}r_{i-m,j}(s,t)-f(s)t^{2\gamma_0}r_{i-m,j}(s,1)\right.$$
$$\left.-g(s)t^{\gamma_0-1}r_{i-m,j}(1,t)+g(s)t^{2\gamma_0}r_{i-m,j}(1,1)\,ds\,dt\right\},$$
$$i\neq j+m, j=1,\dots,m,\ i=m+1,\dots,2m,$$

with

$$f(s)=s^{-1}\gamma_0^{-1}\left(1-(1+\gamma_0)s^{\gamma_0}\right),\ g(s)=s^{\gamma_0}\gamma_0^{-1}\left(1-s^{\gamma_0}\right)-s^{2\gamma_0}$$

$$r_{ij}(s,t) = \sigma_{i,j}(s,t,1,1) = EW_i(s,C_i(1))W_j(t,C_j(1)).$$

Then $\Sigma_{\gamma_0} = Cov(I \times U) = I\,Cov(U)\,I^T$ where $I = I_{2\times 2m} = \begin{bmatrix} 1 \cdots 1\, 0 \cdots 0 \\ 0 \cdots 0\, 1 \cdots 1 \end{bmatrix}$.

We consider the two boundary cases for the dependence structure again. In case of tail independence for all pairs (j_1, j_2), the matrix $I_{\gamma_0}^{-1}\Sigma_{\gamma_0}I_{\gamma_0}^{-1}$ reduces to

$$\Sigma_{iid} = \begin{bmatrix} (\gamma_0+1)^2 & -(\gamma_0+1) \\ -(\gamma_0+1) & 1+(\gamma_0+1)^2 \end{bmatrix},$$

which is the same covariance matrix as that for N i.i.d. data. In case of complete tail dependence for all pairs (j_1, j_2) and if all m scedases $c(\cdot, j)$ are the same, the matrix $I_{\gamma_0}^{-1}\Sigma_{\gamma_0}I_{\gamma_0}^{-1}$ becomes $m\Sigma_{iid}$, which is the same as that for n i.i.d. data with k replaced by k/m.

As an example, consider N data such that all m scedases $c(\cdot, j)$ are the same and the tail dependence structure is described by the well-known logistic stable tail dependence function

$$l_\theta(x_1,\ldots,x_m) = \left(\sum_{j=1}^m x_j^{1/\theta}\right)^\theta, \quad x_j \geq 0;$$

cf. Einmahl et al. (2012). Here the parameter $\theta \in (0,1]$ reflects the degree of tail dependence, with $\theta = 1$ corresponding to tail independence and the limiting case $\theta = 0$ corresponding to complete tail dependence. Then for all $j_1 \neq j_2$ the pairwise tail copula becomes

$$R_{j_1,j_2}(x,y) = x + y - \left(x^{1/\theta} + y^{1/\theta}\right)^\theta = mr_{j_1,j_2}(x,y), \quad x,y \geq 0.$$

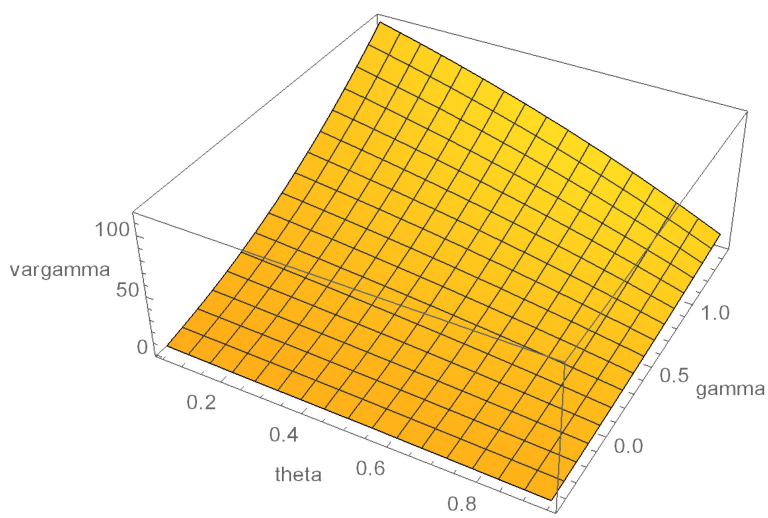

FIG 1. *Asymptotic variance of the MLE of γ_0 for the logistic stable tail dependence function in dimension $m = 20$, as a function of the dependence parameter θ and the extreme value index γ_0.*

With the latter formula we can numerically compute the τ_{ij} in Remark 1 and obtain $\Sigma_{\gamma_0}(\theta)$. Figure 1 shows, for $m = 20$, the asymptotic variance of $\sqrt{k}(\widehat{\gamma}_n - \gamma_0)$ as a function of θ and γ_0. The variance increases smoothly with dependence and with γ_0. For fixed γ_0, it increases from $(\gamma_0+1)^2$ to $20(\gamma_0+1)^2$, as θ decreases from 1 to 0. Clearly, tail independence yields the most accurate estimation of γ_0.

3. Application. This section is devoted to illustrating the testing methods for detecting a trend in extreme rainfalls, both across stations and over time. We use a subset of rainfall data from the German national meteorological service, which consists of daily rainfall amounts recorded in 49 stations ($m = 49$) in three regions of North-West Germany: Bremen, Niedersachsen and Hamburg. The data set comprises nearly complete time series records over 84 years (1931-2014). We divide the data into two seasons: winter from November to March, and summer from May to September, excluding the transitional months April and October.

Although the raw data set comprises a tally of $49 \times 84 \times 150$ rainfall amounts within each season, the actual number of observations we use at each station, n, will be determined by a declustering procedure. This has been designed to remove the effect of temporal dependence and is viewed as a key step to ensure that after pre-processing, the data set can be regarded as having no temporal dependence. The idea of our pre-processing procedure is to create gaps between consecutive observations by removing some days in the data set. The detailed procedure we have employed is outlined in the next paragraph.

The raw data set consists of daily rainfall amounts (in *mm*) at each gauging station $j = 1, \ldots, m$, including zero rainfall. From this data set we will use the daily maximum rainfall amount across the m stations, henceforth referred to as station-wise maxima, for eliciting potential serial dependence. We order all station-wise maxima from high to low. The declustering procedure is initiated by picking up the pair of calendar days with the largest and second largest station-wise maxima. If this second maximum was recorded within two consecutive days of the first station-wise maximum, then all m observations on its corresponding day are removed; otherwise both days are kept. This procedure then rolls out to the subsequent ordered station-wise maxima: for each station-wise maxima, we remove the corresponding day if it is recorded within two consecutive days of any of the previously kept days. This results in the declustered data sets with sample sizes $n = 3561$ and $n = 3552$ for winter and summer respectively. The two data sets are used for testing the presence of scedasis over time and/or across space, whilst accounting for the spatial dependence.[1] In Section 3.1 of the Supplementary Material, we find that the remaining serial dependence in the station-wise maxima after declustering is negligible.

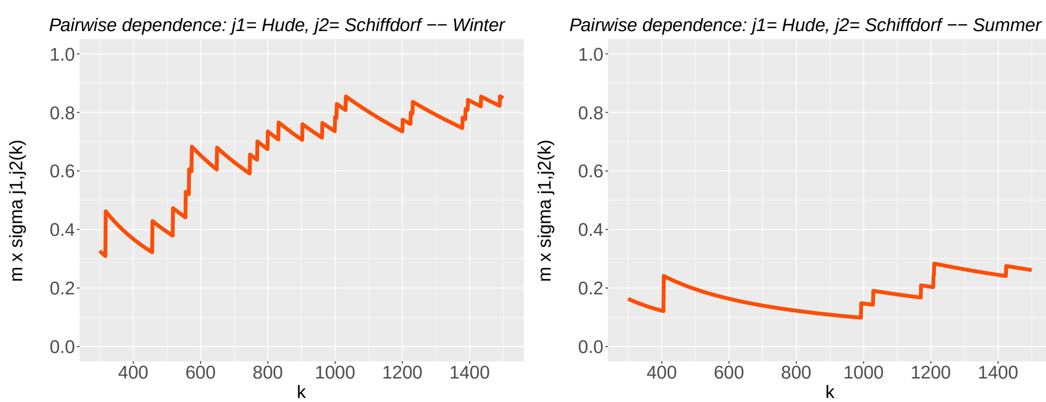

FIG 2. *Estimates of* $m\sigma_{j_1,j_2}(1,1,1,1)$ *for* $k = 300, 301, \ldots, 1500$.

In addition, we check the assumptions of having a constant extreme value index over time and over space (cf. Sections 3.2 and 3.3 in the Supplementary Material) and find no evidence

[1] We make the declustered data sets and the codes for application available at https://github.com/zhouchen0527/rainscedasis

against making such assumptions. In relation to the spatial dependence, we present in Figure 2 two plots displaying estimates for the pairwise spatial dependence through

$$m\hat{\sigma}_{j_1,j_2}(1,1,1,1) := \frac{m}{k}\sum_{i=1}^{n} \mathbb{1}_{\left\{X_{i,j_1}>X_{N-k:N},\, X_{i,j_2}>X_{N-k:N}\right\}},$$

as in (2.7), for $k = 300, \ldots, 1500$. Note that the choice of k in this estimation differs from that in the estimation of tail dependence coefficient by a factor of m, i.e. $k = 300$ corresponds to choosing $k = 6$ at each station. In addition, we choose to estimate $m\sigma_{j_1,j_2}(1,1,1,1)$, since by construction, a factor $1/m$ appears in the expression for $\sigma_{j_1,j_2}(1,1,1,1)$, see (2.3). The plots in Figure 2 illustrate estimation of pairwise spatial dependence for one single pair of stations for winter and for summer. These suggest that spatial dependence is stronger in the winter than in the summer, a finding consistent with both the convective nature of extreme rainfall in the summer and widespread persistent precipitation in the winter.

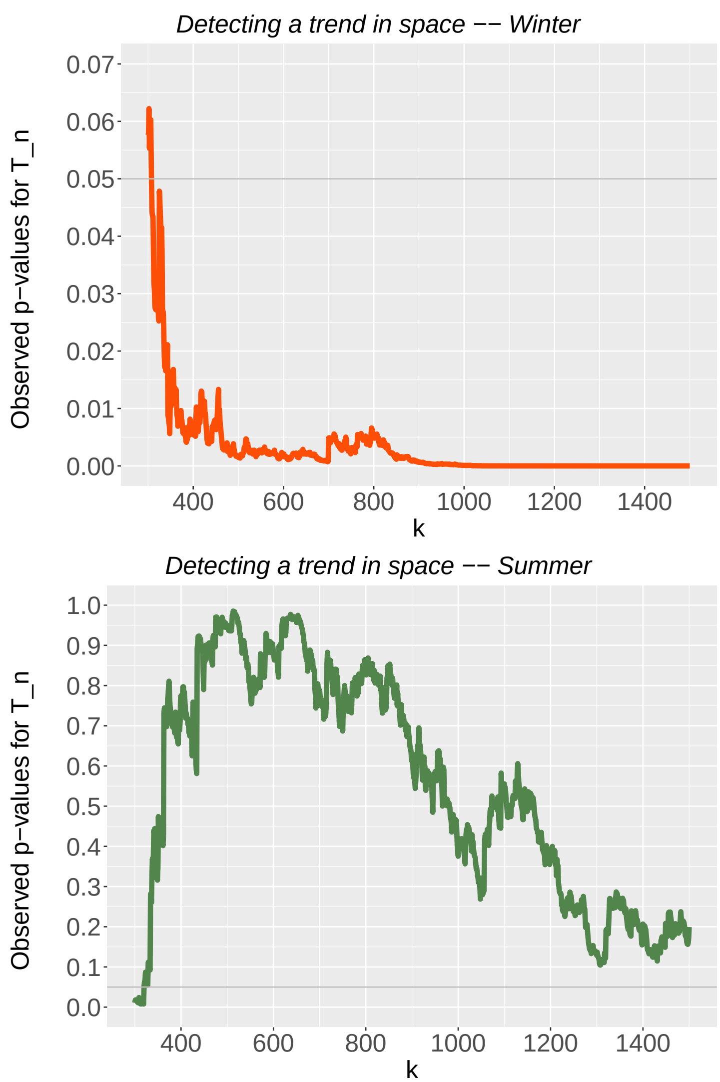

FIG 3. *Obtained p-values through the test T_n for homogeneity across space, all plotted as a function of $k = 300, \ldots, 1500$.*

First, we test whether the total integrated scedasis of extreme rainfall is constant across $m = 49$ stations by adopting the test statistic T_n in Corollary 2.4. We reject the null hypothesis of having constant total integrated scedasis across all stations for large values of T_n. We plot the p-values against k the number of upper observations used in the test in the two plots of Figure 3, for winter and summer seasons respectively. The p-values obtained for the winter season stay below 5% for all $k > 350$. Therefore, we conclude that for the winter season, the total integrated scedasis of extreme rainfall is not constant across stations. In other words, the frequencies of having extreme rainfall differ across stations. In contrast, there is no statistical evidence of a trend in the space-domain over the summer. This finding holds for almost all values of k in the lower panel of Figure 3.

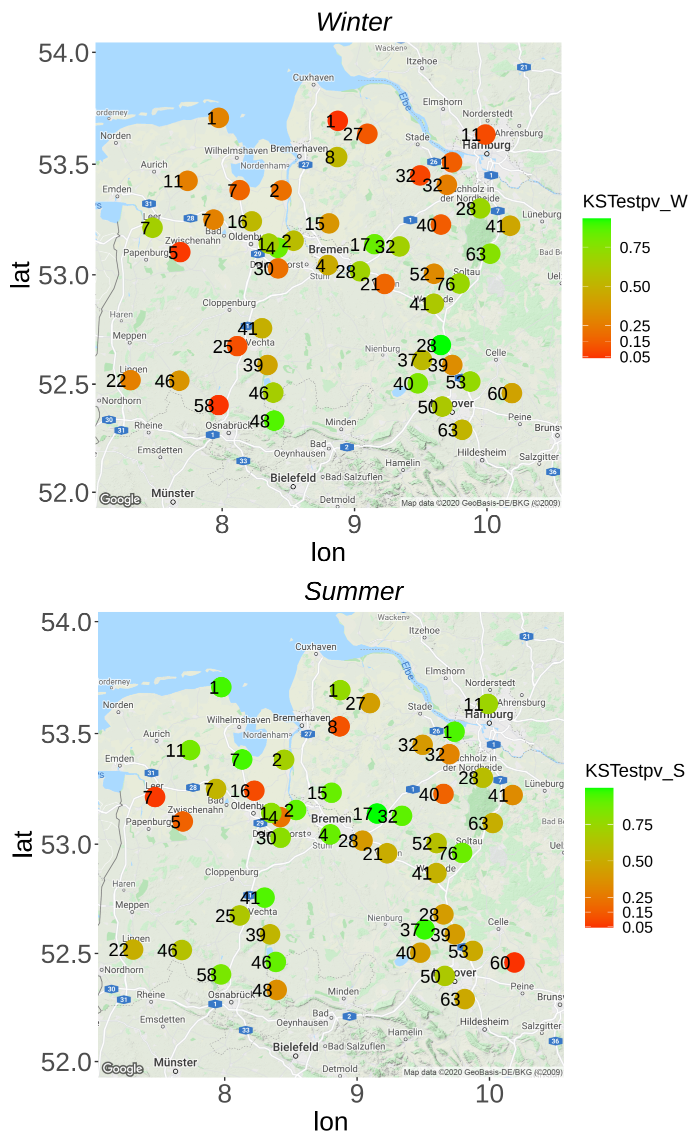

FIG 4. *Obtained p-values for the test of the null hypothesis $H_{0,j}$ of a non-existent time trend at each station $j = 1, 2, \ldots, 49$. The Kolmogorov-Smirnov statistic is applied with $k = 1000$ higher observations. The numbers next to the station-marks indicate elevation in meters; one station on the east side falls outside the map.*

Next, we investigate a possible temporal trend in the extreme rainfall process for each station mapped in Figure 4 by means of a Kolmogorov-Smirnov (KS) type test based on the left hand side of the limit relation in Corollary 2.5. For each season, we apply this test at each station j with $k=1000$, and plot the p-values of the test in the two plots in Figure 4, for winter and summer seasons respectively. The sharper the red in the renderings, the lower the estimated p-values, and the more evidence for rejecting the null hypothesis. The brighter the green marks, the higher the p-values. We find that the p-values vary widely across the selected region, and more so in the winter.

Overall, we find that p-values plunge in the winter but soar in the summer at many locations. In the winter season, the KS type test highlights two stations with p-values below the nominal level $\alpha = 5\%$: $p = 0.044$ for station *Steinau, Kr. Cuxhaven*, with elevation $1m$, and $p = 0.05$ for *Bramsche* at $58m$ high. Nevertheless, we need to interpret such p-values with caution. Given that these are the lower $p-$values across all 49 stations, we are encountering a potential multiple test problem. One potential solution is to consider the Bonferroni correction: the corrected nominal level is $\alpha^* = 5\%/49 \approx 0.1\%$. Since these low p-values do not breach the corrected nominal level, we find no temporal trend over the winter, at the usual significance levels. Similarly, for the summer, the KS type test identifies one significant individual p-value of 0.048 for station *Uetze*, standing at $60m$ of elevation. Again this individual p-value is not in the vicinity of the Bonferroni's corrected critical barrier $\alpha^* = 0.1\%$. To summarize, there seems to be no temporal trend in extreme rainfalls in the winter or in the summer.

Finally, we report the estimated extreme value index γ using *all* data from *all* stations in one season, using the maximum likelihood estimator (MLE) in Theorem 2.6. Figure 5 shows the estimates against various values of k, for the winter and summer seasons respectively. We observe that, for k ranging between 900 and 1100, both estimates paths seem to consolidate a plateau of stability. For the purpose of point estimation, we fix $k = 1000$, highlighted in both plots with a vertical gray line. The estimated extreme value indices are $\widehat{\gamma} = 0.041$ and 0.078 for winter and summer seasons, respectively, with corresponding 95% confidence intervals $(-0.0546,\, 0.1365)$ and $(0.0159,\, 0.1391)$. We find that the estimated standard deviation for the winter, based on the asymptotic distribution of the maximum likelihood estimator for γ, is greater than $(1+\widehat{\gamma})/\sqrt{k}$, the estimated standard deviation assuming tail independence. This indicates that ignoring spatial dependence results in underestimation of the asymptotic variance. The analogous results for the summer season lead to estimated standard deviations nearly matching the values $(1+\widehat{\gamma})/\sqrt{k}$, which suggests weaker spatial tail dependence during the summer months.

4. Proofs. Write for convenience $X_{i,j} = U_{i,j}(Y_{i,j})$, where $(Y_{i,1}, Y_{i,2}, \ldots, Y_{i,m})$ follows the distribution function $\widetilde{F}$ with standard Pareto marginals. Let $Y_{i:n}^{(j)}$ be the i-th order statistic from $Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j}$, for all j.

Proof of Theorem 2.1 a) Tail empirical distribution functions

Consider one station j for the time being and define $C_{j,n}(t) := \frac{1}{N}\sum_{i=1}^{nt} c\left(\frac{i}{n}, j\right)$. (Recall $N = nm$.) According to Proposition 1 in Einmahl et al. (2016) we have under a Skorokhod construction for any $t_0 > 0$ and $0 \le \eta < 1/2$, almost surely,

$$\sup_{0<v\le t_0,\, 0\le t\le 1} v^{-\eta}\left|\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{nt} \mathbb{1}_{\left\{Y_{i,j} > \frac{nmC_{j,n}(1)}{kv\, c(\frac{i}{n},j)}\right\}} - \frac{v\, C_j(t)}{C_{j,n}(1)}\right\} - W_j\left(v, \frac{C_j(t)}{C_{j,n}(1)}\right)\right| \longrightarrow 0,$$

as $n \to \infty$. After some rearrangement we get, almost surely,

$$(4.1)\quad \sup_{0<v\le t_0,\, 0\le t\le 1} v^{-\eta}\left|\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{nt} \mathbb{1}_{\left\{Y_{i,j} > \frac{n}{kv\, c(\frac{i}{n},j)}\right\}} - mvC_j(t)\right\} - W_j\big(mv, C_j(t)\big)\right| \longrightarrow 0.$$

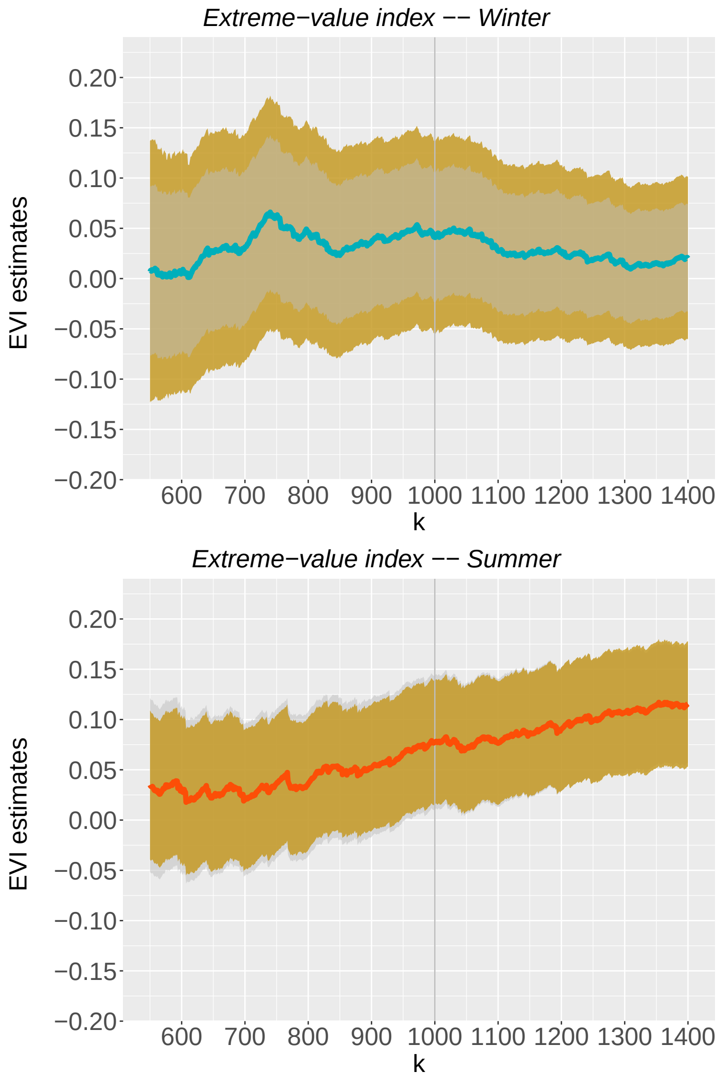

FIG 5. *The maximum likelihood estimates of the extreme value index with accompanying* 95% *confidence intervals, all plotted against* k*, the number of upper order statistics. The gray overlay gives the confidence intervals in case the tail independence setting were to be assumed.*

We are going to transform this result in several steps. First replace v with $\frac{n}{kc(\frac{i}{n},j)}\Big(1 - F_{i,j}\big(U_0(\frac{N}{ku})\big)\Big)$, $0 < u \le t_0$. Note that, as $n \to \infty$, by condition (ii),

$$\frac{n}{kc(\frac{i}{n},j)}\left(1 - F_{i,j}\left(U_0\left(\frac{N}{ku}\right)\right)\right) = \frac{n}{k}\left(1 - F_0\left(U_0\left(\frac{N}{ku}\right)\right)\right)\left\{1 + O\left(A_1\left(\frac{N}{ku}\right)\right)\right\}$$
$$= \frac{u}{m}\left\{1 + O\left(A_1\left(\frac{N}{ku}\right)\right)\right\},$$

and by condition (iv), $\sqrt{k}A_1\left(\frac{N}{ku}\right) \to 0$ uniformly for $0 < u \le t_0$. Hence we have, almost surely,

$$\sup_{0<u\le t_0,\,0\le t\le 1} u^{-\eta}\left|\sqrt{k}\Big\{\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{U_{i,j}(Y_{i,j})>U_0\left(\frac{N}{ku}\right)\right\}} - u\,C_j(t)\Big\} - W_j\big(u,\,C_j(t)\big)\right| \longrightarrow 0.$$

Next replace u with $\frac{N}{k}\left(1-F_0\left(b_0\left(\frac{N}{k}\right)+x\,a_0\left(\frac{N}{k}\right)\right)\right)$ and note that by condition (iii) and Proposition 3.2 (equation 3.2) of Drees et al. (2006)

$$\frac{N}{k}\left(1-F_0\Big(b_0\big(\frac{N}{k}\big)+x\,a_0\big(\frac{N}{k}\big)\Big)\right) = (1+\gamma x)^{-1/\gamma}\left\{1+O\left(A_0\Big(\frac{N}{k}\Big)\right)\right\}$$

uniformly for $x \ge x_0 > -1/\gamma_+$ and $\sqrt{k}A_0\big(\frac{N}{k}\big) \to 0$ by condition (iv). Recall $X_{i,j} = U_{i,j}(Y_{i,j})$. Hence we have for each j,

$$\text{(4.2)}\quad \sup_{x_0\le x_j<x_1,\,0\le t\le 1}(1+\gamma x_j)^{\eta/\gamma}\left|\sqrt{k}\Big\{\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{\frac{X_{i,j}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x_j\right\}} - (1+\gamma x_j)^{-1/\gamma}\,C_j(t)\Big\}\right.$$
$$\left. - W_j\Big((1+\gamma x_j)^{-1/\gamma},\,C_j(t)\Big)\right| \longrightarrow 0,$$

which yields the weak convergence of the weighted process on the left to the weighted Wiener process on the right.

It remains to prove the joint convergence of the processes at different locations. First we deal with convergence of the finite dimensional distributions. After that we consider tightness. For ease of writing we confine ourselves to the first two dimensions, i.e. $\big\{(X_{i,1},\,X_{i,2})\big\}_{i=1}^n$ and one point (s_1,t_1) and (s_2,t_2) at each dimension. According to the Cramèr-Wold device, we look first at all linear combinations $(u,v\in\mathbb{R},\,x_0\le x,y<x_1,\,0\le t_1,t_2\le 1)$

$$\text{(4.3)}\quad \sqrt{k}\left[u\,\frac{1}{k}\sum_{i=1}^{nt_1}\left(\mathbb{1}_{\left\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\right\}} - P\Big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\Big\}\right)\right.$$
$$\left. +\,v\,\frac{1}{k}\sum_{i=1}^{nt_2}\left(\mathbb{1}_{\left\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>y\right\}} - P\Big\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>y\Big\}\right)\right].$$

The Lindeberg-Feller central limit theorem applies to this expression since the summands are indicators and hence bounded. As a consequence we obtain that

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{nt_1}\mathbb{1}_{\left\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\right\}} - P\Big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\Big\},\right.$$
$$\left.\frac{1}{k}\sum_{i=1}^{nt_2}\mathbb{1}_{\left\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>y\right\}} - P\Big\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>y\Big\}\right)$$

converges in distribution to $\Big(W_1\big((1+\gamma x)^{-1/\gamma},C_1(t_1)\big),W_2\big((1+\gamma y)^{-1/\gamma},C_2(t_2)\big)\Big)$. Note that, e.g. for the first dimension, by assumptions (ii) and (iv),

$$\frac{1}{m\,n}\sum_{i=1}^{nt_1}\frac{N}{k}P\Big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\Big\} = \frac{1}{m\,n}\sum_{i=1}^{nt_1}(1+\gamma x)^{-1/\gamma}\,c\Big(\frac{i}{n},1\Big)+o\Big(\frac{1}{\sqrt{k}}\Big)$$
$$= (1+\gamma x)^{-1/\gamma}\,C_1(t_1)+o\Big(\frac{1}{\sqrt{k}}\Big).$$

It follows that

$$\sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{nt_1}\mathbb{1}_{\big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\big\}}-(1+\gamma x)^{-1/\gamma}C_1(t_1),\qquad (4.4)$$

$$\frac{1}{k}\sum_{i=1}^{nt_2}\mathbb{1}_{\big\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>y\big\}}-(1+\gamma y)^{-1/\gamma}C_2(t_2)\Big)$$

converges in distribution to $\Big(W_1\big((1+\gamma x)^{-1/\gamma},C_1(t_1)\big),\,W_2\big((1+\gamma y)^{-1/\gamma},C_2(t_2)\big)\Big)$.

Next we prove tightness of the process in (4.4) in the space $D\big([0,1]^2\times[x_0,x_1)^2\big)$ with index (t_1,t_2,x,y). We know from (4.2) that for $0\le\eta<1/2$ the sequence of processes

$$(1+\gamma x_j)^{\eta/\gamma}\sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{nt_j}\mathbb{1}_{\big\{\frac{X_{i,j}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x_j\big\}}-(1+\gamma x_j)^{-1/\gamma}C_j(t_j)\Big)$$

is tight in the space $D\big([0,1]\times[x_0,x_1)\big)$, for $j=1$ and $j=2$. It then follows from Ferger and Vogel (2015) that the joint process (4.4) is also tight. Hence the weak convergence is established. A Skorokhod construction yields the result.

For the proof of (2.3) consider the covariance of the two components in (4.4). It suffices to show that the leading term in this covariance

$$\frac{1}{k}\sum_{i=1}^{n(t_1\wedge t_2)}E\Big[\mathbb{1}_{\big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>\frac{s_1^{-\gamma}-1}{\gamma}\big\}}\mathbb{1}_{\big\{\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>\frac{s_2^{-\gamma}-1}{\gamma}\big\}}\Big]$$

$$=\frac{1}{m\,n}\sum_{i=1}^{n(t_1\wedge t_2)}\frac{N}{k}P\Big\{\frac{X_{i,1}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>\frac{s_1^{-\gamma}-1}{\gamma},\,\frac{X_{i,2}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>\frac{s_2^{-\gamma}-1}{\gamma}\Big\}$$

$$\to\frac{1}{m}\int_0^{t_1\wedge t_2}R_{1,2}(s_1c(u,1),s_2c(u,2))du,\quad n\to\infty,\qquad (4.5)$$

for fixed $(t_1,t_2,s_1,s_2)\in[0,1]^2\times[0,T]^2$, where $T=(1+\gamma x_0)^{-1/\gamma}>0$. Since $X_{i,j}=U_{i,j}(Y_{i,j})$ with $Y_{i,j}$ standard Pareto distributed random variables, given any $\varepsilon>0$, for sufficiently large n,

$$P\left(\frac{X_{i,1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{s_1^{-\gamma}-1}{\gamma},\frac{X_{i,2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{s_2^{-\gamma}-1}{\gamma}\right)$$

$$=P\left(\frac{k}{n}Y_{i,1}>\frac{m}{\frac{N}{k}\left(1-F_{i,1}\left(b_0\left(\frac{N}{k}\right)+a_0\left(\frac{N}{k}\right)\frac{s_1^{-\gamma}-1}{\gamma}\right)\right)},\right.$$

$$\left.\frac{k}{n}Y_{i,2}>\frac{m}{\frac{N}{k}\left(1-F_{i,2}\left(b_0\left(\frac{N}{k}\right)+a_0\left(\frac{N}{k}\right)\frac{s_2^{-\gamma}-1}{\gamma}\right)\right)}\right)$$

$$\le P\left(\frac{k}{n}Y_{i,1}>\frac{m}{s_1c\left(\frac{i}{n},1\right)(1+\varepsilon)},\frac{k}{n}Y_{i,2}>\frac{m}{s_2c\left(\frac{i}{n},2\right)(1+\varepsilon)}\right).$$

The definition of $R_{1,2}$ implies that, as $n\to\infty$,

$$\frac{n}{k}P\left(\frac{k}{n}Y_{i,1}>\frac{1}{v_1},\frac{k}{n}Y_{i,2}>\frac{1}{v_2}\right)\to R_{1,2}(v_1,v_2),$$

uniformly for $(v_1,v_2)\in[0,V]^2$ with any fixed $V>0$. By the continuity and boundedness of c and of $R_{1,2}$, as $n\to\infty$,

$$\frac{n}{k}P\left(\frac{k}{n}Y_{i,1}>\frac{m}{s_1c\left(\frac{i}{n},1\right)(1+\varepsilon)},\frac{k}{n}Y_{i,2}>\frac{m}{s_2c\left(\frac{i}{n},2\right)(1+\varepsilon)}\right)$$
$$-R_{1,2}\left(\frac{s_1c\left(\frac{i}{n},1\right)(1+\varepsilon)}{m},\frac{s_2c\left(\frac{i}{n},2\right)(1+\varepsilon)}{m}\right)\to 0,$$

uniformly in i, s_1 and s_2. Hence, uniformly in $(t,s_1,s_2)\in[0,1]\times[0,T]^2$,

$$\frac{1}{k}\sum_{i=1}^{nt}P\left(\frac{X_{i,1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{s_1^{-\gamma}-1}{\gamma},\frac{X_{i,2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{s_2^{-\gamma}-1}{\gamma}\right)$$
$$\le\frac{1}{k}\sum_{i=1}^{nt}P\left(\frac{k}{n}Y_{i,1}>\frac{m}{s_1c\left(\frac{i}{n},1\right)(1+\varepsilon)},\frac{k}{n}Y_{i,2}>\frac{m}{s_2c\left(\frac{i}{n},2\right)(1+\varepsilon)}\right)$$
$$=\int_{1/n}^{([nt]+1)/n}\frac{n}{k}P\left(\frac{k}{n}Y_{1,1}>\frac{m}{s_1c\left(\frac{[nu]}{n},1\right)(1+\varepsilon)},\frac{k}{n}Y_{1,2}>\frac{m}{s_2c\left(\frac{[nu]}{n},2\right)(1+\varepsilon)}\right)du$$
$$\to\int_0^t R_{1,2}\left(\frac{s_1c(u,1)(1+\varepsilon)}{m},\frac{s_2c(u,2)(1+\varepsilon)}{m}\right)du,\quad\text{as } n\to\infty.$$

The lower bound follows similarly by replacing $1+\varepsilon$ with $1-\varepsilon$ and reversing the inequality. Now by the homogeneity of R, (4.5) follows by letting $\varepsilon\to 0$ and taking $t=t_1\wedge t_2$. ❏

The following lemma is useful for the proof of Theorem 2.1 b).

LEMMA 4.1. *For all $\delta>0$, there exists $0<a<1$ such that for large n,*

$$P\Big\{U_0\big(\frac{aN}{ks}\big)\le X_{N-[ks]:N}\le U_0\big(\frac{N}{aks}\big)\ \textit{for}\ \frac{1}{2k}\le s\le 1\Big\}>1-\delta.$$

PROOF. Condition (1.1) implies, by inversion, that there exist $M>1$ and $t_0>0$ such that for all $t\ge t_0$,

$$U_0\Big(\frac{t}{M}\Big)\le U_{i,j}(t)\le U_0(Mt) \tag{4.6}$$

Note that from Inequality on page 419 of Shorack and Wellner (1986), we get that for every $\delta>0$ there exists $0<b<1$ such that

$$P\Big\{\frac{bn}{ks}\le Y^{(j)}_{n-[ks]:n}\ \text{for}\ \frac{1}{2k}\le s\le 1\ \text{and}\ Y^{(j)}_{n-[ks]:n}\le\frac{n}{kbs}$$
$$\text{for}\ 0\le s\le 1;\ j=1,2,\dots,m\Big\}>1-\delta/2. \tag{4.7}$$

Note also that

$$Y^{(1)}_{n-[ks]:n}\le Y_{N-[ks]:N}\le\max_{1\le j\le m}Y^{(j)}_{n-\left[\frac{ks}{m}\right]:n} \tag{4.8}$$

Next we show, using (4.6), that with probability tending to 1,

$$U_0\Big(\frac{1}{M}Y_{N-[ks]:N}\Big)\le X_{N-[ks]:N}\le U_0\big(MY_{N-[ks]:N}\big). \tag{4.9}$$

By definition $[ks]$ of the $Y_{i,j}$ are not less than $Y_{N-[ks]:N}$. Hence at least $[ks]$ of the $U_{i,j}(Y_{i,j})$ are not less than $U_0(\frac{1}{M}Y_{N-[ks]:N})$. The right-hand inequality is similar. The result follows combining (4.7), (4.8) and (4.9). □

Proof of Theorem 2.1 b) Tail empirical quantile function

We start from (2.2) in Theorem 2.1a). By taking $t_j = 1$ and aggregating over $1 \leq j \leq m$, we get that for any $0 \leq \eta < 1/2$, as $n \to \infty$, almost surely,

$$\sup_{x_0 \leq x < x_1} (1+\gamma x)^{\eta/\gamma} \left| \sqrt{k}\left(\mathbb{P}_n(x) - (1+\gamma x)^{-1/\gamma}\right) - \sum_{j=1}^{m} W_j((1+\gamma x)^{-1/\gamma}, C_j(1)) \right| \longrightarrow 0, \tag{4.10}$$

where

$$\mathbb{P}_n(x) = \frac{1}{k}\sum_{i=1}^{n}\sum_{j=1}^{m} \mathbb{1}_{\left\{\frac{X_{i,j}-b_0(\frac{N}{k})}{a_0(\frac{N}{k})}>x\right\}},$$

and $x_0 > -1/\gamma_+$ and $x_1 = 1/(-\gamma_-)$. As a consequence, for $\delta_1 > 0$,

$$P\left\{ \sup_{x_0 \leq x < x_1} (1+\gamma x)^{\eta/\gamma} \left| \sqrt{k}\left(\mathbb{P}_n(x) - (1+\gamma x)^{-1/\gamma}\right) - \sum_{j=1}^{m} W_j((1+\gamma x)^{-1/\gamma}, C_j(1)) \right| > \delta_1 \right\} \longrightarrow 0. \tag{4.11}$$

We remark that the region $x_0 \leq x < x_1$ has different implications for $\gamma > 0$, $\gamma < 0$, and $\gamma = 0$. For $\gamma > 0$, it implies that $1+\gamma x \geq 1+\gamma x_0 > 0$, i.e. $1+\gamma x$ is bounded away from zero. For $\gamma < 0$, $1+\gamma x > 1+\gamma x_1 = 0$. Hence $1+\gamma x > 0$ but not necessarily bounded away from zero. On the other hand, $1+\gamma x \leq 1+\gamma x_0$, i.e. $1+\gamma x$ is bounded away from ∞. For $\gamma = 0, 1+\gamma x = 1$.

Then, split the range of s in two subintervals, $[1/(2k), s_0]$ and $[s_0, T]$, where s_0 is a sufficiently small, but fixed, positive number.

For the range $[s_0, T]$ we use (4.10) and Vervaat's Lemma (cf. e.g. Appendix A of de Haan and Ferreira (2006)) with $x_n(s) := \mathbb{P}_n(s)$ and $x_n^{\leftarrow}(s) := \left(X_{N-[ks]:N} - b_0(N/k)\right)/a_0(N/k)$. We then obtain the statement in (2.4), with the 'sup' taken over $[s_0, T]$.

For $s \in [1/(2k), s_0]$, we first deal with the Gaussian processes term. Let W_0 be a univariate standard Wiener process. It is well-known (and follows from the law of the iterated logarithm) that for every $\tilde{\delta} > 0$ there exists an s_0, such that $P\{\sup_{0<s\leq s_0} |W_0(s)|/s^\eta < \tilde{\delta}\} > 1-\tilde{\delta}$. Now $W_j(\cdot, C_j(1)) \stackrel{d}{=} \sqrt{C_j(1)}W_0$ for all j. Hence for $\delta > 0$, there exists an $s_0(\delta)$, such that for all $s_0 \leq s_0(\delta)$,

$$P\left\{ \sup_{0<s\leq s_0} s^{-\eta} \left| \sum_{j=1}^{m} W_j(s, C_j(1)) \right| < \delta \right\} > 1-\delta. \tag{4.12}$$

Hence, we shall concentrate on proving that with probability larger than $1-\delta$, with a proper choice of s_0, for large n,

$$\sup_{\frac{1}{2k} \leq s \leq s_0} s^{\gamma+\frac{1}{2}+\varepsilon}\sqrt{k}\left(\frac{X_{N-[ks]:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma} \right) \leq \delta, \tag{4.13}$$

and

$$\inf_{\frac{1}{2k}\le s\le s_0} s^{\gamma+\frac{1}{2}+\varepsilon}\sqrt{k}\left(\frac{X_{N-[ks]:N}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}-\frac{s^{-\gamma}-1}{\gamma}\right)\ge -\delta. \tag{4.14}$$

In the following we further split the range of s, $[1/(2k), s_0]$, into two subintervals, $[1/(2k), t_n]$ and $(t_n, s_0]$, where t_n depends only on the constant a in Lemma 4.1 (eventually depending on δ) and a sufficiently small $\xi > 0$, although the choice of t_n is different for proving the upper bound in (4.13) and the lower bound in (4.14):

$$\sqrt{k}t_n^{1/2+\varepsilon} := \begin{cases} \delta\Delta_1^{-1} \text{ with } \Delta_1 := \frac{a^{-\gamma}-1}{\gamma}(1+\xi) > 0, \text{for the upper bound,} \\ \delta\Delta_2^{-1} \text{ with } \Delta_2 := \frac{1-a^{\gamma}}{\gamma}(1+\xi) > 0, \text{for the lower bound.} \end{cases}$$

Moreover the following technical relations provide the constants to determine an upper bound for s_0. Note that for large enough η there exists η' with $1-\eta < 1-\eta' < 1/2+\varepsilon$, such that the following hold:

$$\left(1+\frac{\delta\gamma}{\sqrt{k}}s^{-1/2-\varepsilon}\right)^{-1/\gamma} < \left(1+\frac{\delta\gamma}{\sqrt{k}}s^{-(1-\eta')}\right)^{-1/\gamma}, \tag{4.15}$$

$$\sup_{s>t_n}\frac{s^{-(1-\eta')}}{\sqrt{k}} \le \frac{t_n^{-(1-\eta')}}{\sqrt{k}} = \frac{t_n^{\eta'-1/2+\varepsilon}}{\sqrt{k}t_n^{1/2+\varepsilon}} \to 0,$$

$$\left(1+\frac{\delta\gamma}{\sqrt{k}}s^{-(1-\eta')}\right)^{-1/\gamma} \le 1-c_I\frac{\delta}{\sqrt{k}}s^{-(1-\eta')} \text{ for some } 0<c_I\le 1 \text{ and large } n,$$

$$\left(1-\frac{\delta\gamma}{\sqrt{k}}s^{-(1-\eta')}\right)^{-1/\gamma} \ge 1+c_{II}\frac{\delta}{\sqrt{k}}s^{-(1-\eta')} \text{ for some } 0<c_{II}\le 1 \text{ and large } n,$$

where the last two inequalities follow from the inequalities $(1+\gamma x)^{-1/\gamma} \le 1-c_I x$ and $(1-\gamma x)^{-1/\gamma} \ge 1+c_{II}x$ respectively, for $0 < x < \min(1, 1/(-\gamma_-))$ and some $0 < c_I, c_{II} \le 1$. Then, we should take

$$s_0 \le \min(s_0(\delta), (\delta c_I/(1+\delta_1))^{(\eta-\eta')^{-1}}, \delta c_{II}/((m/a)^\eta(1+\xi)^{-\eta/\gamma}(1+\delta_1))^{(\eta-\eta')^{-1}}).$$

(A) Upper bound and $s \in [(2k)^{-1}, t_n]$: Assume first that A_0 is eventually positive. Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 4.1 imply: for all $\varepsilon, \delta, \theta > 0$, there exists $0 < a < 1$ such that for large n, with probability at least $1-\delta$,

$$\begin{aligned}&\frac{X_{N-[ks]:N}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}-\frac{s^{-\gamma}-1}{\gamma}\\ \le&\frac{(as)^{-\gamma}-s^{-\gamma}}{\gamma}+\overline{\Psi}_{\gamma,\rho}\left(\frac{1}{as}\right)A_0\left(\frac{N}{k}\right)+(as)^{-\gamma-\rho-\theta}A_0\left(\frac{N}{k}\right)\\ \le&s^{-\gamma}\left\{\frac{a^{-\gamma}-1}{\gamma}+s^{-\rho-\theta}K\,A_0\left(\frac{N}{k}\right)\right\},\end{aligned}$$

for some $K > 0$. Hence,

$$\begin{aligned}&s^{\gamma+1/2+\varepsilon}\sqrt{k}\left(\frac{X_{N-[ks]:N}-U_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}-\frac{s^{-\gamma}-1}{\gamma}\right)\le s^{1/2+\varepsilon}\sqrt{k}\left\{\frac{a^{-\gamma}-1}{\gamma}+s^{-\rho-\theta}KA_0\left(\frac{N}{k}\right)\right\}\\ \le&\frac{\delta}{\Delta_1}\left\{\frac{a^{-\gamma}-1}{\gamma}+s^{-\rho-\theta}KA_0\left(\frac{N}{k}\right)\right\}\le\delta,\end{aligned}$$

for n large, uniformly in $s \in [(2k)^{-1}, t_n]$, since $\sup_{s\in[(2k)^{-1},t_n]} s^{-\rho-\theta} A_0\left(\frac{N}{k}\right) \to 0$ choosing $\theta < -\rho$.

(B) Upper bound and $s \in (t_n, s_0]$: We prove that, with probability at least $1-\delta$, for large n,

$$\mathbb{P}_n\left(\frac{s^{-\gamma}-1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) \le s, \quad \text{for all } s \in (t_n, s_0],$$

which implies the upper bound in (4.13).

We intend to apply (4.11) with x replaced by $\gamma^{-1}(s^{-\gamma}-1) + \delta k^{-1/2} s^{-\gamma-1/2-\varepsilon}$. For this, note that,

$$1+\gamma\left(\frac{s^{-\gamma}-1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) = s^{-\gamma}\left(1+\frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)$$

and, for $s \in (t_n, s_0]$ and $\gamma > 0$ the right-hand side is at least $s_0^{-\gamma}\left(1+\delta\gamma k^{-1/2} s_0^{-1/2-\varepsilon}\right)$, consequently bounded away from zero. For $\gamma < 0$ the inequality is reversed and the expression is bounded from above. Hence,

$$\begin{aligned}
&\mathbb{P}_n\left(\frac{s^{-\gamma}-1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) \le s\left(1+\frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-1/\gamma} \\
&\quad + \frac{1}{\sqrt{k}}\widetilde{W}\left(s\left(1+\frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-1/\gamma}\right) + \frac{\delta_1}{\sqrt{k}} s^{\eta}\left(1+\frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-\eta/\gamma} \\
&\le s\left(1-c_I\frac{\delta}{\sqrt{k}} s^{-(1-\eta')}\right) + \frac{1+\delta_1}{\sqrt{k}} s^{\eta} = s - \frac{s^{\eta}}{\sqrt{k}}\left(c_I\delta s^{\eta'-\eta} - (1+\delta_1)\right)
\end{aligned}$$

with $\widetilde{W}(s) := \sum_{j=1}^m W_j(s, C_j(1))$ and where for the second inequality we have applied (4.15), $\eta < 1/2$, (4.12) and $\left(1+\delta\gamma s^{-1/2-\varepsilon}/\sqrt{k}\right)^{-1/\gamma} \le 1$.

It remains to check that $c_I\delta s^{\eta'-\eta} - (1+\delta_1) \ge 0$ which holds by the choice of s_0.

(C) Lower bound and $s \in [(2k)^{-1}, t_n]$: As in (A) assume first that A_0 is eventually positive. Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 4.1 imply: for all $\varepsilon, \delta, \theta > 0$, there exists $0 < a < 1$ such that for large n, with probability at least $1-\delta$,

$$\begin{aligned}
&\frac{X_{N-[ks]:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma} \\
&\ge \frac{(a/s)^{\gamma} - s^{-\gamma}}{\gamma} + \overline{\Psi}_{\gamma,\rho}\left(\frac{a}{s}\right) A_0\left(\frac{N}{k}\right) - \left(\frac{a}{s}\right)^{\gamma+\rho+\theta} A_0\left(\frac{N}{k}\right) \\
&\ge s^{-\gamma}\left\{\frac{a^{\gamma}-1}{\gamma} - s^{-\rho-\theta} K\, A_0\left(\frac{N}{k}\right)\right\},
\end{aligned}$$

for some $K > 0$, hence,

$$\begin{aligned}
&s^{\gamma+1/2+\varepsilon}\sqrt{k}\left(\frac{X_{N-[ks]:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma}\right) \ge s^{1/2+\varepsilon}\sqrt{k}\left\{\frac{a^{\gamma}-1}{\gamma} - s^{-\rho-\theta} K A_0\left(\frac{N}{k}\right)\right\} \\
&\ge \frac{\delta}{\Delta_2}\left\{\frac{a^{\gamma}-1}{\gamma} - s^{-\rho-\theta} K A_0\left(\frac{N}{k}\right)\right\} \ge -\delta,
\end{aligned}$$

for large n, uniformly in $s \in [(2k)^{-1}, t_n]$, since $\sup_{s\in[(2k)^{-1},t_n]} s^{-\rho-\theta} A_0\left(\frac{N}{k}\right) \to 0$ choosing $\theta < -\rho$.

(D) Lower bound and $s \in (t_n, s_0]$: We prove that, with probability at least $1-\delta$, for large n,

$$\mathbb{P}_n\left(\frac{s^{-\gamma}-1}{\gamma} - \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) \geq s + \frac{1}{k}, \qquad \text{for all } s \in (t_n, s_0],$$

which implies the lower bound in (4.13). Similarly as in (B) apply (4.11) with x replaced by $\gamma^{-1}(s^{-\gamma}-1) - \delta k^{-1/2} s^{-\gamma-1/2-\varepsilon}$. Note that,

$$1 + \gamma\left(\frac{s^{-\gamma}-1}{\gamma} - \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) = s^{-\gamma}\left(1 - \frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)$$

and, for $s \in (t_n, s_0]$ and $\gamma > 0$ the right-hand side is at least $s_0^{-\gamma}(1-\gamma\Delta_2) > 0$ and consequently bounded away from zero. For $\gamma < 0$ the inequality is reversed and the expression is bounded above. Hence,

$$\begin{aligned}
&\mathbb{P}_n\left(\frac{s^{-\gamma}-1}{\gamma} - \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon}\right) \geq s\left(1 - \frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-1/\gamma} \\
&+ \frac{1}{\sqrt{k}} \widetilde{W}\left(s\left(1 - \frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-1/\gamma}\right) - \frac{\delta_1}{\sqrt{k}} s^{\eta}\left(1 - \frac{\delta\gamma}{\sqrt{k}} s^{-1/2-\varepsilon}\right)^{-\eta/\gamma} \\
&\geq s\left(1 + c_{II}\frac{\delta}{\sqrt{k}} s^{-(1-\eta')}\right) - a^{-\eta}(1+\xi)^{-\eta/\gamma}(1+\delta_1)\frac{s^{\eta}}{\sqrt{k}} \\
&= s + \frac{s^{\eta}}{\sqrt{k}}\left(c_{II}\delta s^{\eta'-\eta} - a^{-\eta}(1+\xi)^{-\eta/\gamma}(1+\delta_1)\right)
\end{aligned} \tag{4.16}$$

where we have used in particular (4.15). It remains to check that the right-hand side of (4.16) is at least $s + 1/k$ which is equivalent to
$s^{\eta}\sqrt{k}\left(c_{II}\delta s^{\eta'-\eta} - a^{-\eta}(1+\xi)^{-\eta/\gamma}(1+\delta_1)\right) \geq 1$. This holds by the choice of $\eta' < \eta$ and choosing $s_0 \leq \delta c_{II}/((m/a)^{\eta}(1+\xi)^{-\eta/\gamma}(1+\delta_1))^{(\eta-\eta')^{-1}}$.

Finally, if A_0 is eventually negative, the proofs are the same, except that the signs of the remaining terms in *(A)* and in *(C)*, $+(as)^{-\gamma-\rho-\theta}A_0\left(\frac{N}{k}\right)$ and $-(a/s)^{\gamma+\rho+\theta}A_0\left(\frac{N}{k}\right)$, should be interchanged. ❑

Proof of Theorem 2.3

Fix $j \in \{1,\ldots,m\}$. Replace x in (2.2) with $\left(X_{N-k:N} - b_0\left(\frac{N}{k}\right)\right)/a_0\left(\frac{N}{k}\right)$ and use condition (iii) jointly with Theorem 2.3.8 of de Haan and Ferreira (2006) to get

$$\begin{aligned}
&\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{nt} \mathbb{1}_{\left\{X_{i,j} > X_{N-k:N}\right\}} - \left(1 + \gamma\,\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-1/\gamma} C_j(t)\right. \\
&\left. - W_j\left(\left(1 + \gamma\,\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-1/\gamma}, C_j(t)\right)\right\} = o_p\left(\left(1 + \gamma\,\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-\eta/\gamma}\right).
\end{aligned}$$

Now by (2.4)

$$\sqrt{k}\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)} \xrightarrow{P} \sum_{j=1}^{m} W_j\left(1, C_j(1)\right)$$

and hence

$$\sqrt{k}\left\{1 - \left(1 + \gamma\,\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-1/\gamma}\right\} - \sqrt{k}\frac{X_{N-k:N} - b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)} \xrightarrow{P} 0.$$

Combining this we obtain

$$\sup_{0\le t\le 1}|\sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{X_{i,j}>X_{N-k:N}\right\}}-C_j(t)\Big)-\{W_j\big(1,C_j(t)\big)-C_j(t)\sum_{r=1}^{m}W_r\big(1,C_r(1)\big)\}|\xrightarrow{P}0,$$

which yields (2.6).

Next we prove for fixed $(t,w_1,w_2)\in[0,1]\times[0,T]^2$, as $n\to\infty$,

(4.17)
$$\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right\}}\xrightarrow{P}\frac{1}{m}\int_0^t R_{j_1,j_2}(w_1c(u,j_1),w_2c(u,j_2))du.$$

We check the variance of the left hand side: using (4.5), as $n\to\infty$,

$$\begin{aligned}
&Var\left(\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right\}}\right)\\
&=\sum_{i=1}^{nt}\frac{1}{k^2}Var\left(\mathbb{1}_{\left\{\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right\}}\right)\\
&\le\sum_{i=1}^{nt}\frac{1}{k^2}P\left(\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right)\\
&\le\frac{1}{km}(1+\varepsilon)\int_0^t R_{j_1,j_2}(w_1c(u,j_1),w_2c(u,j_2))du\to 0.
\end{aligned}$$

Hence, by Chebyshev inequality, as $n\to\infty$,

$$\begin{aligned}
&\frac{1}{k}\sum_{i=1}^{nt}\mathbb{1}_{\left\{\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right\}}\\
&\qquad-\frac{1}{k}\sum_{i=1}^{nt}P\left(\frac{X_{i,j_1}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_1^{-\gamma}-1}{\gamma},\frac{X_{i,j_2}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}>\frac{w_2^{-\gamma}-1}{\gamma}\right)\xrightarrow{P}0.
\end{aligned}$$

Then, (4.5) implies (4.17).

By the continuity of the right hand side of (4.17) and the monotonicity of both sides of (4.17) in t, w_1 and w_2, this result holds uniformly for $(t,w_1,w_2)\in[0,1]\times[0,T]^2$.

Finally, for fixed $(s_1,s_2)\in[0,T]^2$, by (2.4), as $n\to\infty$

$$\left(1+\gamma\frac{X_{N-[ks_j]:N}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-1/\gamma}\xrightarrow{P}s_j,$$

for $j=1,2$. We can then replace w_j in (4.17) with $\left(1+\gamma\frac{X_{N-[ks_j]:N}-b_0\left(\frac{N}{k}\right)}{a_0\left(\frac{N}{k}\right)}\right)^{-1/\gamma}$, which yields (2.7) for fixed $(t,s_1,s_2)\in[0,1]\times[0,T]^2$. The uniformity follows as before. ❏

Proof of Corollary 2.5 Under the null hypothesis, using Theorem 2.3, under a Skorokhod construction,

$$\sqrt{k}\big(\widehat{C}_j(t)-t\widehat{C}_j(1)\big)=\sqrt{k}\big(\widehat{C}_j(t)-C_j(t)\big)-t\sqrt{k}\big(\widehat{C}_j(1)-C_j(1)\big)$$

converges uniformly ($0 \le t \le 1$) to

$$W_j\big(1, C_j(t)\big) - C_j(t)\sum_{r=1}^{m} W_r\big(1, C_j(1)\big) - t\Big[W_j\big(1, C_j(1)\big) - C_j(1)\sum_{r=1}^{m} W_r\big(1, C_r(1)\big)\Big]$$

$$= W_j\big(1, tC_j(1)\big) - tW_j\big(1, C_j(1)\big) \stackrel{d}{=} \sqrt{C_j(1)}\, B(t).$$

Then the result follows directly via Slutsky's theorem. ❑

The proof of Theorem 2.6 is deferred to the Supplementary Material. Here we only present the main steps of the proof.

We use "local asymptotic normal theory", where the local log-likelihood and local score processes are fundamental, consisting of reparametrizations of the former with local parameter $h = (h_1, h_2) \in \mathbb{R}^2$:

$$\begin{cases} h_1 = \sqrt{k}\,\big(\gamma_{N/k} - \gamma_0\big) \\ h_2 = \sqrt{k}\,\big(\sigma_{N/k}/a_0(N/k) - 1\big) \end{cases} \quad \Leftrightarrow \quad \begin{cases} \gamma = \gamma_{N/k} = \gamma_0 + h_1/\sqrt{k} \\ \sigma = \sigma_{N/k} = a_0(\frac{N}{k})(1 + h_2/\sqrt{k}), \end{cases}$$

$$\begin{aligned} \widetilde{L}_{N,k}(h) =& k\int_0^1 \ell\left(\gamma_0 + \frac{h_1}{\sqrt{k}}, a_0(\frac{N}{k})(1 + \frac{h_2}{\sqrt{k}}), X_{N-[ks],N} - X_{N-k,N}\right) ds \\ =& k\int_0^1 \ell\left(\gamma_0 + \frac{h_1}{\sqrt{k}}, 1 + \frac{h_2}{\sqrt{k}}, \frac{X_{N-[ks],N} - X_{N-k,N}}{a_0(\frac{N}{k})}\right) ds - k\log a_0(\frac{N}{k}), \end{aligned}$$

and

$$\begin{cases} \frac{\partial \widetilde{L}_{N,k}}{\partial h_1}(h_1, h_2) = \frac{1}{\sqrt{k}}\frac{\partial L_{N,k}}{\partial \gamma}(\gamma, \sigma_{N/k}) = 0 \\ \frac{\partial \widetilde{L}_{N,k}}{\partial h_2}(h_1, h_2) = \frac{1}{\sqrt{k}}\frac{\partial L_{N,k}}{\partial \sigma}(\gamma, \sigma_{N/k}) = 0. \end{cases}$$

The main steps of the proof are as follows:

a) Denote $\theta = (\gamma, \sigma) \in \mathbb{R} \times (0, \infty)$. First prove that

$$\frac{\partial^2 \widetilde{L}_{N,k}(h)}{\partial h \partial h^T} = -I_{\gamma_0} + o_P(1), \quad I_{\gamma_0} = -\int_0^1 \frac{\partial^2 \ell}{\partial\theta\partial\theta^T}\left(\gamma_0, 1, \frac{s^{-\gamma_0} - 1}{\gamma_0}\right) ds,$$

uniformly in a large enough ball H_n to ensure that it covers the true solution; I_{γ_0} is the Fisher information matrix related to the approximate $GP_{\gamma_0,1}$ model

$$I_{\gamma_0} = \begin{pmatrix} \frac{2}{1+3\gamma_0+2\gamma_0^2} & \frac{1}{1+3\gamma_0+2\gamma_0^2} \\ \frac{1}{1+3\gamma_0+2\gamma_0^2} & \frac{1}{1+2\gamma_0} \end{pmatrix}.$$

I_{γ_0} is positive definite, which implies that the local log-likelihood process is eventually strictly concave on H_n with probability tending to 1.

b) Then, by integration one obtains an expansion for the local log-likelihood process (holding uniformly for h in compact sets):

$$\widetilde{L}_{N,k}(h) = \widetilde{L}_{N,k}(0) + h^T\frac{\partial \widetilde{L}_{N,k}}{\partial h}(0) - \frac{1}{2}h^T I_{\gamma_0} h + o_P(1)$$

where

$$\frac{\partial \widetilde{L}_{N,k}}{\partial h}(0) = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\frac{\partial \ell}{\partial\theta}\left(\gamma_0, 1, \frac{X_{N-i+1,N} - X_{N-k,N}}{a_0(\frac{N}{k})}\right) \stackrel{d}{\to} N(0, \Sigma_{\gamma_0}).$$

c) Finally the Argmax Theorem (van der Vaart, 2000, Corollary 5.58) provides the result: let

$$\begin{aligned} M_n(h) &= \widetilde{L}_{N,k}(h) - \widetilde{L}_{N,k}(0), \\ M(h) &= h^T N(0, \Sigma_{\gamma_0}) - \tfrac{1}{2} h^T I_{\gamma_0} h, \quad h \in \mathbb{R}^2. \end{aligned}$$

Then,

$$\widehat{h}_n = \underset{h \in H_n}{\operatorname{argmax}}\, M_n(h) \xrightarrow{d} h = \underset{h \in \mathbb{R}^2}{\operatorname{argmax}}\, M(h) \overset{d}{=} I_{\gamma_0}^{-1} N(0, \Sigma_{\gamma_0})$$

provided $\widehat{h}_n$ is tight which holds as in Dombry and Ferreira (2019). Finally, note that

$$\widetilde{L}_{N,k}(\widehat{h}_n) = L_{N,k}\left(\gamma_0 + \frac{\hat{h}_{n,1}}{\sqrt{k}}, a_0(\frac{N}{k})(1 + \frac{\hat{h}_{n,2}}{\sqrt{k}})\right)$$

and similarly for its derivatives.

Acknowledgements. We would like to thank the Editor, the Associate Editor and four anonymous referees for their detailed reviews, which helped to improve the paper substantially.

John Einmahl holds the Arie Kapteyn Chair 2019-2022 and gratefully acknowledges the corresponding research support. Ana Ferreira was partially supported by FCT-Portugal: UIDB/04621/2020, UIDB/00006/2020 and SFRH/BSAB/142912/2018; IST: P.5088. Laurens de Haan was financially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the projects UIDB/00006/2020 and PTDC/MAT-STA/28649/2017. Cláudia Neves gratefully acknowledges support from EPSRC-UKRI Innovation Fellowship grant EP/S001263/1 and project FCT-UIDB/00006/2020.

SUPPLEMENTARY MATERIAL

Supplementary Material for "Spatial dependence and space-time trend in extreme events" The supplementary material consists of three sections. Section 1 provides the detailed proof of Theorem 2.6. Section 2 shows a simulation study regarding the proposed estimation and testing procedures. Section 3 validates the assumptions for the data application.
(doi: TO BE COMPLETED BY THE TYPESETTER).

REFERENCES

Bücher, A., Jäschke, S., and Wied, D. (2015). Nonparametric tests for constant tail dependence with an application to energy and finance. *J. Econometrics*, 187(1):154–168.

Coles, S. G. and Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *J. R. Statist. Soc. B*, 58(2):329–347.

Davison, A. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *J. R. Statist. Soc. B*, 52:393–442.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.*, 27(2):161–186.

Dombry, C. and Ferreira, A. (2019). Maximum likelihood estimators based on the block maxima method. *Bernoulli*, 25(3):1690–1723.

Drees, H., de Haan, L., and Li, D. (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions. *Journal of Statistical Planning and Inference*, 136(10):3498 – 3538.

Einmahl, J. H. J., de Haan, L., and Zhou, C. (2016). Statistics of heteroscedastic extremes. *J. R. Statist. Soc. B*, 78:31–51.

Einmahl, J. H. J., Krajina, A., and Segers, J. (2012). An m-estimator for tail dependence in arbitrary dimensions. *Annals of Statistics*, 40(3):1764–1793.

Ferger, D. and Vogel, D. (2015). Weak convergence of the empirical process and the rescaled empirical distribution function in the Skorokhod product space. *Theory Probab. Appl.*, 4(54):609–625, 2010 (arXiv:1506.04324v1).

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

de Haan, L., Klein Tank, A., and Neves, C. (2015). On tail trend detection: modeling relative risk. *Extremes*, 18(2):141–178.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. John Wiley and Sons, New York.

van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.