No. 2020-009

SPATIAL DEPENDENCE AND SPACE-TIME TRENDS IN EXTREME EVENTS

By

John H.J. Einmahl, Ana Ferreira, Laurens de Haan, Claudia Neves, Chen Zhou

23 March 2020

ISSN 0924-7815
ISSN 2213-9532
Spatial dependence and space-time trend in extreme events

John H.J. Einmahl
Tilburg University

Ana Ferreira
Instituto Superior Técnico, University of Lisbon

Laurens de Haan
University of Lisbon
Erasmus University Rotterdam

Cláudia Neves
University of Reading

Chen Zhou
Erasmus University Rotterdam
Bank of The Netherlands

Abstract

The statistical theory of extremes is extended to observations that are non-stationary and not independent. 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.

KEY WORDS AND PHRASES: Multivariate extreme value statistics, non-identical distributions, sequential tail empirical process, testing.

MSC2010 SUBJECT CLASSIFICATIONS: Primary 62G32, 62G30, 62G05, 62G10, 62G20; secondary 60F17, 60G70.

JEL CODES: C12, C13, C14.

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

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

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 temporal or a spatial context – with independent observations – a natural estimator of the 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 dependent observations as follows. Our leading example concerns daily rainfall in Northern Germany, with measurements taken at 49 stations over 84 years. All 49 time series of daily data are split into two seasons, summer and winter, each of which comprising at least 150 days so as to stay clear of transition periods. There is evidence that the extreme value index remains constant throughout the region and time span we chose to focus on. The distribution functions of the $49 \times 84 \times 150$ observations are assumed to be in one scedasis class. The distribution of the rainfall vector on any day is assumed to be in the domain of attraction of a multivariate extreme value distribution. We may assume independence in time. For the assumption of constant extreme value index and independence in time we refer to Buishand et al. (2008), Klein Tank et al. (2009), and the references therein.

Consider independent random vectors $(X_{i,1}, X_{i,2}, \ldots, X_{i,m})$, $i = 1, 2, \ldots, n$. In the rainfall context, “$m$” is the number of stations and “$i$” is time (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

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

(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 the scedasis $c(\cdot, j)$ a positive continuous function for each $j$. 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 $0 \leq t \leq 1$ and $j = 1, 2, \ldots, m$,

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

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 $G_\gamma$. As a consequence of (1.1), $\gamma$ is now the common extreme-value index: $F_{i,j} \in \mathcal{D}(G_\gamma)$, $i = 1, \ldots, n$; $j = 1, \ldots, m$. Consider independent random vectors $(X_{i,1}, X_{i,2}, \ldots, X_{i,m})$, $i = 1, 2, \ldots, n$. In the rainfall context, “$m$” is the
number of stations and “i” is time (in days).

Estimators $\hat{C}_j$ 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 $(X_{i,1}, X_{i,2}, \ldots, X_{i,m})$. Assume that the distribution function $F_{i,j}$ of $X_{i,j}$ is continuous and let $U_{i,j}(t) := F_{i,j}^{\uparrow}(1 - 1/t)$, where the arrow indicates the generalized inverse function. To model the spatial dependence, we further assume that $F_{i,j}(x_1, x_2, \ldots, x_m)$ 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 $j_2$:

$$R_{j_1,j_2}(x,y) = \lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[1 - F_{i,j_1}(X_{i,j_1}) \leq tx, 1 - F_{i,j_2}(X_{i,j_2}) \leq ty], \quad (x,y) \in [0,\infty)^2 \setminus \{(\infty, \infty)\}.$$  

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 $\{X_{i,j}\}_{i=1, j=1}^{n,m}$. We define the estimator

$$\hat{C}_j(t) := \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I}\{x_{i,j} > x_{N-k,N}\}$$ (1.2)

where $k$ is an intermediate sequence i.e. $k = k(n) \to \infty$, $k(n)/n \to 0$, as $n \to \infty$.

In this paper we make the following four contributions.

1. We establish the joint asymptotic behavior of $\{\hat{C}_j(t)\}_{t \in [0,1]}$, $j = 1, \ldots, m$.

2. We test $H_0 : C_j(1) = \frac{1}{m}$ for all $j = 1, \ldots, m$, i.e., the total scedasis is constant over the various locations. We perform the test by checking whether the limit vector (in distribution) of

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

has mean zero. This will be done via an adapted $\chi^2$-test.

3. We test $H_{0,j} : C_j(t) = tC_j(1)$ for $0 \leq t \leq 1$ and some given $j \in \{1, \ldots, m\}$, i.e., the scedasis $c(\cdot, j)$ is constant over time. Since, under $H_{0,j}$, the limit in distribution of the process $\{\sqrt{k}(\hat{C}_j(t) - C_j(t))\}$

4. We test $H_{0,j} : C_j(t) = tC_j(1)$ for $0 \leq t \leq 1$ and some given $j \in \{1, \ldots, m\}$, i.e., the scedasis $c(\cdot, j)$ is constant over time. Since, under $H_{0,j}$, the limit in distribution of the process $\{\sqrt{k}(\hat{C}_j(t) - C_j(t))\}$

5. We test $H_{0,j} : C_j(t) = tC_j(1)$ for $0 \leq t \leq 1$ and some given $j \in \{1, \ldots, m\}$, i.e., the scedasis $c(\cdot, j)$ is constant over time. Since, under $H_{0,j}$, the limit in distribution of the process $\{\sqrt{k}(\hat{C}_j(t) - C_j(t))\}$
Results

tC_j(1)_{0 \leq t \leq 1} is essentially 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 $\gamma$ based on all $n \times m$ observations.

Crucial for these results is a joint 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 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.

2 Results

Throughout the paper we assume the following conditions:

(i) Spatial dependence. Assume that $\tilde{F}(x_1, x_2, \ldots, x_m) := F_i(U_{i,1}(x_1), U_{i,2}(x_2), \ldots, U_{i,m}(x_m))$ 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 $\lim_{t \to 1} A_1(t) = 0$ such that

$$\sup \max_{i,j} \left| \frac{1 - F_{i,j}(x)}{1 - F_0(x)} - c \left( \frac{i}{n^j} \right) \right| = O \left( A_1 \left( \frac{1}{1 - F_0(x)} \right) \right), \text{ as } x \uparrow x^\ast. \quad (2.1)$$

(iii) Second order condition for $F_0$. Write $U_0(t) := F_0^+(1 - 1/t)$. Then 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 $x > 0$,

$$\lim_{t \to \infty} \frac{U_0(tx) - U_0(t)}{\tilde{a}_0(t)} - \frac{x^{\gamma-1}}{\gamma} = \Psi_{\gamma,\rho}(x), \quad (2.2)$$

where $\Psi_{\gamma,\rho}(x)$ is as in Corollary 2.3.5 of de Haan and Ferreira (2006).

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

$$k \to \infty, k/n \to 0,$$

$$\sqrt{k} A_1 \left( q \frac{n}{k} \right) \to 0, \text{ for all } q > 0,$$

$$\sqrt{k} A_0 \left( \frac{n}{k} \right) \to 0,$$

$$\sqrt{k} \sup_{|u-v| \leq \frac{1}{k}} |c(u,j) - c(v,j)| \to 0, \text{ for } j = 1, 2, \ldots, m.$$

We begin with presenting two fundamental approximations, which are the basis for the main results.
Results

(2.4) as given in Corollary 2.3.7 of de Haan and Ferreira (2006). Such inequalities are valid if we replace $U_0(t)$, $\tilde{a}_0$ and $\Psi$ with $b_0(t)$, $a_0$ and $\overline{W}$ therein.

Theorem 1. Assume conditions (i)-(iv) with $\rho < 0$. Let $x_0 > -1/\gamma_+$; set $x_1 := 1/(-\gamma_-)$.

a) Tail empirical distribution functions

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

$$
\max_{1 \leq j \leq m} \sup_{x_0 \leq x < x_1, 0 \leq t_j \leq 1} \left| \frac{1}{k} \sum_{i=1}^{n t_j} \mathbb{1}\left\{ \frac{x_{i,j} - b_0(\frac{k}{n})}{a_0(\frac{k}{n})} > t \right\} - (1 + \gamma x)^{-1/\gamma} C_j(t_j) \right| - W_j \left( (1 + \gamma x)^{-1/\gamma}, C_j(t_j) \right) \to 0,
$$

where $(W_1, \ldots, W_m)$ is a Gaussian vector of bivariate Wiener processes $W_j$ with a covariance matrix

$$
\Sigma = \Sigma(s_1, s_2, t_1, t_2)
$$

with entries

$$
\sigma_{j_1, j_2}(s_1, s_2, t_1, t_2) := \text{Cov} \left( W_{j_1}(s_1, C_{j_1}(t_1)), W_{j_2}(s_2, C_{j_2}(t_2)) \right) = \frac{1}{m^2} \int_0^{t_1 \wedge t_2} R_{j_1, j_2}(s_1 c(u, j_1), s_2 c(u, j_2)) du,
$$

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

b) Tail empirical quantile function

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

$$
\sup_{s \leq s \leq T} s^{-1/2+\varepsilon} \left| s^{1/2} \left\{ \mathbb{E}\left( \frac{X_{N-[ks]} - b_0(N)}{a_0(N/k)} - \frac{s^{-\gamma} - 1}{\gamma} \right) - \sum_{j=1}^{m} W_j(s, C_j(1)) \right\} \right| \to 0.
$$

Corollary 2. Under the conditions and in the setup of Theorem 1, with $\varepsilon > 0$,

$$
\sup_{1 \leq s \leq T} \left| \sqrt{k} \left( \frac{X_{N-[ks]} - X_{N-[k-1]}}{a_0(N/k)} - \frac{s^{-\gamma} - 1}{\gamma} \right) - \sum_{j=1}^{m} \left( s^{-\gamma-1} W_j(s, C_j(1)) - W_j(1, C_j(1)) \right) \right| \to 0.
$$

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

Theorem 3. Under the conditions and in the setup of Theorem 1, as $n \to \infty$,

$$
\max_{1 \leq j \leq m} \sup_{0 < r \leq 1} \left| \sqrt{k} \left( \hat{C}_j(t) - C_j(t) \right) - \left\{ W_j(1, C_j(t)) - C_j(t) \sum_{r=1}^{m} W_r(1, C_r(1)) \right\} \right| \to 0.
$$
Moreover, we have the uniform consistency of the estimator of the covariance matrix \( \Sigma \) as follows. For \( T > 0 \) and \( j_1 \neq j_2 \), as \( n \to \infty \),

\[
\sup_{0 \leq t_1 \leq t_2} \sup_{0 \leq s_1, s_2 \leq T} \left| \frac{1}{k} \sum_{n=1}^{mT} \mathbb{I} \{ X_{n+j_1} > X_{N-\lfloor k \frac{t_1}{T} \rfloor}, \ldots, X_{n+j_2} > X_{N-\lfloor k \frac{t_2}{T} \rfloor} \} - \frac{1}{m} \int_0^T R_{j_1,j_2}(s_1 c(u,j_1), s_2 c(u,j_2)) \right| \xrightarrow{P} 0.
\]

(2.7)

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

\[
H_0 : C_j(1) = \frac{1}{m}, \quad \text{for all } j = 1, \ldots, m,
\]

\[
H_1 : C_j(1) \neq \frac{1}{m}, \quad \text{for some } j = 1, \ldots, m.
\]

(2.8)

Let \( I_m \) be the \( m \)-unit vector, \( I_m \) the identity matrix of dimension \( m \) and define \( M := I_m - \frac{1}{m} I_m I_m' \).

From Theorem 3 and under \( H_0 \), \( D = \sqrt{k} (\hat{C}_1(1) - \frac{1}{m}, \ldots, \hat{C}_m(1) - \frac{1}{m})' \) is asymptotically an \( m \)-multivariate normal with zero mean vector and covariance matrix \( M \Sigma_1 M' \), where \( \Sigma_1 = \Sigma(1,1,1,1) \). Assume that \( \Sigma_1 \) is invertible. Then \( \text{rank}(M \Sigma_1 M') = \text{rank}(M) = m - 1 \). We therefore confine attention to the first \( m - 1 \) components of \( D \) denoted by \( D_{m-1} \), which has an asymptotic covariance matrix \( (M \Sigma_1 M')^{-1} \).

Here the notation \( A_{m-1} \) refers to the first \( m - 1 \) rows and \( m - 1 \) columns of an \( m \times m \) matrix \( A \), i.e. \( A_{m-1} = (I_{m-1}, 0_{m-1}) A (I_{m-1}, 0_{m-1})' \). Finally, we define the test statistic

\[
T_n := D_{m-1}' (M \hat{\Sigma}_1 M')^{-1} D_{m-1},
\]

with \( \hat{\Sigma}_1 \) estimated via the empirical counterpart given in (2.7). From Theorem 3 we immediately get the asymptotic behavior of \( T_n \) under \( H_0 \).

**Corollary 4.** Assume that \( \Sigma_1 \) is invertible. Then under \( H_0 \), \( T_n \xrightarrow{d} \chi_{m-1}^2 \), as \( n \to \infty \).

Next we consider, for \( j \in \{1, \ldots, m\} \), the testing problem \( H_{0,j} : C_j(t) = t C_j(1) \), \( 0 \leq t \leq 1 \), and \( H_{1,j} \) that this is not the case. We can use test statistics of the Kolmogorov-Smirnov-type or Cramér-von Mises-type based on the process \( \sqrt{k}(\hat{C}_j(t) - t \hat{C}_j(1)) / \sqrt{\hat{C}_j(1)} \), \( 0 \leq t \leq 1 \).

**Corollary 5.** Fix \( j \in \{1, \ldots, m\} \). Under the hypothesis that \( C_j(t) = t C_j(1) \), for \( 0 \leq t \leq 1 \),

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

with \( B \) a Brownian bridge.

Finally, we introduce the maximum likelihood estimator (MLE) of \((\gamma, a_0(\frac{T}{k}))\) based on dependent and non-identically distributed observations as described in Section 1. The estimator is based on the
sample \{X_{N-i+1:N} - X_{N-k:N}\}_{i=1}^k\ for which we have the result of Corollary 2. We highlight \(\gamma_0\) as the true unknown parameter value.

In particular, we know from (2.5) that an approximate model for \(\{X_{N-i+1:N} - X_{N-k:N}\}_{i=1}^k\) is the Generalized Pareto (GP\(_{\gamma,\sigma}\)), with well-known log-likelihood

\[
\ell(\gamma, \sigma, x) = -\log \sigma - \left(1 + \frac{1}{\gamma}\right) \log \left(1 + \gamma \frac{x}{\sigma}\right), \quad 0 < x < \frac{\sigma}{-\gamma},
\]

(for \(\gamma = 0\) the formula is interpreted as \(-\log \sigma - x/\sigma\)), and standard tail quantile function,

\[
(1 - GP_{\gamma,1})^{-1}(s) = \frac{s^{-\gamma} - 1}{\gamma}, \quad s \in (0, 1).
\]

The misspecified log-likelihood based on the above sample can be written as, with parameter space \((\gamma, \sigma) \in \mathbb{R} \times (0, \infty)\),

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

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

\[
\begin{align*}
\frac{\partial}{\partial \gamma} L_{N,k}(\gamma, \sigma_{N/k}) &= 0 \\
\frac{\partial}{\partial \sigma} L_{N,k}(\gamma, \sigma_{N/k}) &= 0.
\end{align*}
\]

**Theorem 6.** Under conditions (i)-(iv) with \(\gamma_0 > -1/2\), with probability tending to 1, there exists a unique sequence of estimators \((\hat{\gamma}_n, \hat{a}_0(N/k))\), maximizing (2.10), for which

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

where

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

and \(\Sigma_{\gamma_0}\) is the covariance matrix of the random vector

\[
\begin{bmatrix}
\sum_{j=1}^m \frac{1}{\gamma_0} \int_0^1 (s^{\gamma_0} - (1 + \gamma_0)s^{2\gamma_0}) \{s^{-\gamma_0-1}W_j(s, C_j(1)) - W_j(1, C_j(1))\} \, ds \\
\sum_{j=1}^m (1 + \gamma_0) \int_0^1 s^{2\gamma_0} \{s^{-\gamma_0-1}W_j(s, C_j(1)) - W_j(1, C_j(1))\} \, ds
\end{bmatrix}
\]
Application

with $(W_j)_{j=1}^m$ from Theorem 1.

Remark 1. The covariance matrix $\Sigma_{\gamma_0}$ can be calculated as follows: let $U$ be a $(1 \times 2m)$ vector with the first $i = 1, \ldots, m$, components as $\gamma_0^{-1} \int_0^1 (s^{\gamma_0} - (1 + \gamma_0)s^{2\gamma_0}) \{s^{-\gamma_0-1}W_i(s, C_i(1)) - W_i(1, C_i(1))\} ds$, and the remaining $i = m + 1, \ldots, 2m$ components as $(1 + \gamma_0)\int_0^1 s^{2\gamma_0} \{s^{-\gamma_0-1}W_i(s, C_i(1)) - W_i(1, C_i(1))\} ds$. The covariance matrix of $U$ has entries $\{\tau_{ij}\}_{i,j=1}^{2m}$ given by,

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

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

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

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

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

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

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

where

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

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

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

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. 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 procedure results in the declustered data set used for testing the presence of scedasis over time and/or across space.

First, we test whether the scedasis of extreme rainfall is constant across \( m = 49 \) stations by adopting the test statistic \( T_n \) in Corollary 4. We reject the null hypothesis of having constant 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 1, 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 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 $k$ depicted in the right panel of Figure 1.

Next, we investigate a possible temporal trend in the extreme rainfall process for each station mapped in Figure 2 by means of a Kolmogorov-Smirnov (KS) type test based on the left hand side of the limit relation in Corollary 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 2, 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 $\alpha^* = 5\%/49 = 0.1\%$. Since these low $p$-values do not breach the corrected nominal level, there is not enough evidence in the data to reject the null hypotheses of no local 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 $\gamma$ using all data from all stations in one season, using the maximum likelihood estimator in Theorem 6. Figure 3 shows the estimates against various values of $k$, for the winter and summer seasons respectively. We observe that for $k$ ranging between 950 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 grey line. The estimated extreme value indices are $\hat{\gamma} = 0.040$ and 0.078 for winter and summer seasons, respectively. We conclude that the magnitude of extreme rainfalls in the summer is higher than that in the winter.

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 $\hat{F}$. 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 1 a) Tail empirical distribution functions**

Consider one station $j$ for the time being and define

$$ C_{j,n}(t) := \frac{1}{N} \sum_{i=1}^{nt} c\left(\frac{i}{n}, j\right). $$

According to Proposition 1 of Einmahl et al. (2016) we have under a Skorokhod construction for any
Proofs

t_0 > 0 and 0 \leq \eta < 1/2, almost surely,

$$
\sup_{0 < v \leq t_0, 0 \leq t \leq 1} v^{-\eta} \left| \sqrt{K} \left\{ \frac{1}{K} \sum_{i=1}^{n_t} I \{ Y_{i,j} > \frac{m v C_{j,n}(1)}{C_{j,n}(1)} \} - v C_j(t) \right\} - W_j \left( v \frac{C_j(t)}{C_{j,n}(1)} \right) \right| \to 0,
$$
as \eta \to \infty. After some rearrangement we get, almost surely,

$$
\sup_{0 < v \leq t_0, 0 \leq t \leq 1} v^{-\eta} \left| \sqrt{K} \left\{ \frac{1}{K} \sum_{i=1}^{n_t} I \{ Y_{i,j} > \frac{n v C_{j,n}(1)}{C_{j,n}(1)} \} - m v C_j(t) \right\} - W_j \left( m v C_j(t) \right) \right| \to 0. \tag{4.1}
$$

We are going to transform this result in several steps. First replace \( v \) with \( \frac{n}{K c(x,j)} \left( 1 - F_{i,j} \left( U_0 \left( \frac{N}{K u} \right) \right) \right) \) for 0 < \( u \leq t_0 \). Note that by condition (ii), as \( \eta \to \infty \),

$$
\frac{n}{K c(x,j)} \left( 1 - F_{i,j} \left( U_0 \left( \frac{N}{K u} \right) \right) \right) = \frac{n}{K} \left( 1 - F_0 \left( U_0 \left( \frac{N}{K u} \right) \right) \right) \{ 1 + O \left( A_1 \left( \frac{N}{K u} \right) \right) \} = \frac{n}{m} \left[ 1 + O \left( A_1 \left( \frac{N}{K u} \right) \right) \right]
$$

and by condition (iv)

$$
\sqrt{K} A_1 \left( \frac{N}{K u} \right) \to 0 \text{ uniformly for } 0 < u \leq t_0.
$$

Hence we have, almost surely,

$$
\sup_{0 < u \leq t_0, 0 \leq t \leq 1} u^{-\eta} \left| \sqrt{K} \left\{ \frac{1}{K} \sum_{i=1}^{n_t} I \{ U_{i,j} \left( Y_{i,j} \right) > U_0 \left( \frac{N}{K} \right) \} - u C_j(t) \right\} - W_j \left( u, C_j(t) \right) \right| \to 0.
$$

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)

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

uniformly for \( x \geq x_0 \) and \( \sqrt{K} A_0 \left( \frac{N}{K} \right) \to 0 \) by condition (iv).

Recall \( X_{i,j} = U_{i,j} \left( Y_{i,j} \right) \). Hence we have for each \( j \) and \( x_0 \) larger than \( 1/(1-\gamma) \),

$$
\sup_{x_0 \leq x < x_1, 0 \leq t \leq 1} (1 + \gamma x)^{-\eta/(1-\gamma)} \left| \sqrt{K} \left\{ \frac{1}{K} \sum_{i=1}^{n_t} I \{ x_{i,j} > b_0 \left( \frac{N}{K} \right) + x \} \right\} - (1 + \gamma x)^{-1/(1-\gamma)} C_j(t) \right\} - W_j \left( (1 + \gamma x)^{-1/(1-\gamma)}, C_j(t) \right) \right| \to 0, \tag{4.2}
$$

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. For simplicity we do
that in the context of (4.1) (standard marginals), not (4.2). The general case follows similar to above. 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. \( \{(Y_{i,1}, Y_{i,2})\}_{i=1}^{n} \) and one point \((s_1, t_1)\) and \((s_2, t_2)\) at each dimension. According to Cramèr-Wold device, we look first at all linear combinations \( (x, y) \), \( s_1, t_1 > 0, 0 < t_1, t_2 \leq 1 \)

\[
\sqrt{k} \left[ \frac{1}{k} \sum_{i=1}^{n} \left( \mathbb{I}\{Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)}\} - P\left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)} \right\} \right) + y \frac{1}{k} \sum_{i=1}^{n} \left( \mathbb{I}\{Y_{i,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)}\} - P\left\{ Y_{i,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)} \right\} \right) \right].
\]  

(4.3)

For the application of Lyapunov’s theorem (cf. Billingsley, 1979, Theorem 27.3) it is sufficient to check the limit behavior of moments of orders 2 and 4. The variance of \( \sqrt{k} \sum_{i=1}^{n} \mathbb{I}\{Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)}\} \) is

\[
\frac{1}{n} \sum_{i=1}^{n} P\left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)} \right\} \left( 1 - P\left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)} \right\} \right) \to s_1 n C_1(t_1). \quad (4.4)
\]

Next we consider the covariance of the two terms which is

\[
\frac{1}{k} \sum_{i=1}^{n} \sum_{i_2=1}^{nt_2} E \left[ \left( \mathbb{I}\{Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)}\} - P\left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)} \right\} \right) \left( \mathbb{I}\{Y_{i_2,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)}\} - P\left\{ Y_{i_2,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)} \right\} \right) \right].
\]

Since all products are of lower order except one (cf. (4.4)), we only need to concentrate on

\[
\frac{1}{k} \sum_{i=1}^{n} \sum_{i_2=1}^{nt_2} E \left[ \mathbb{I}\{Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)}\} \mathbb{I}\{Y_{i_2,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)}\} \right]
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{k} \sum_{i_2=1}^{nt_2} P\left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{1}{n}, 1)} , \ Y_{i_2,2} > \frac{n}{k s_2 c(\frac{1}{n}, 2)} \right\}
\]

which, on the basis of assumption (i), Theorem 6.1.5 and page 222 of de Haan and Ferreira (2006), converges to \( \int_{0}^{t_1 \wedge t_2} R_{1,2}(s_1 c(u, j_1), s_2 c(u, j_2)) du \).

The fourth moments of the terms of (4.3) are similar. Again all terms of this expression are of lower
Proofs

order except

\[ Q_{n,i} := x^4 P \left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)} \right\} + y^4 P \left\{ Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \right\} + (4x^3y + 4xy^3 + 6x^2y^2) P \left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)}, Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \right\}. \]

Hence the sum of the fourth moments of (4.3) boils down, asymptotically, to

\[ \sum_{i=1}^{n(t_1 \land t_2)} \left( \sqrt{\frac{k}{n}} \right)^4 Q_{n,i} = \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \frac{n}{k} Q_{n,i} \]

which is of order \( k^{-1} \). Recall that the sum of the second order moments tends to a constant. Hence Lyapunov’s theorem applies and we have proved convergence of

\[ \sqrt{\frac{k}{n}} \left( \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \mathbb{I} \{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)} \} - P \left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)} \right\} \right) \]

\[ + \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \mathbb{I} \{ Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \} - P \left\{ Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \right\} \]

to \( (W_1(s_1, mc_1(t_1)), W_2(s_2, mc_2(t_2))) \) with the covariance given as \( \sigma_1,2(s_1, s_2, t_1, t_2) \).

Next note that by assumption (iv)

\[ x \frac{1}{n} \sum_{i=1}^{n(t_1 \land t_2)} P \left\{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)} \right\} + y \frac{1}{n} \sum_{i=1}^{n(t_1 \land t_2)} P \left\{ Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \right\} \]

\[ = x s_1 mC_1(t_1) + y s_2 mC_2(t_2) + o \left( \frac{1}{\sqrt{k}} \right). \]

It follows that

\[ \sqrt{\frac{k}{n}} \left( \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \mathbb{I} \{ Y_{i,1} > \frac{n}{k s_1 c(\frac{i}{n}, 1)} \} - s_1 mC_1(t_1) \right) \]

\[ + \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \mathbb{I} \{ Y_{i,2} > \frac{n}{k s_2 c(\frac{i}{n}, 2)} \} - s_2 mC_2(t_2) \] (4.5)

converges to \( (W_1(s_1, mc_1(t_1)), W_2(s_2, mc_2(t_2))) \).

Next we prove tightness of the process (4.5) in the space \( D([0, 1]^2 \times [0, T]^2) \) with index \( (t_1, t_2, s_1, s_2) \in [0, 1]^2 \times [0, T]^2 \) for \( T > 0 \). We know from Einmahl et al. (2016) that for \( 0 \leq \eta < 1/2 \) the sequence of processes

\[ s_j^{-\eta} \sqrt{\frac{k}{n}} \left( \frac{1}{k} \sum_{i=1}^{n(t_1 \land t_2)} \mathbb{I} \{ Y_{i,j} > \frac{n}{k s_j c(\frac{i}{n}, \eta)} \} - s_j mC_j(t_j) \right) \]

is tight in the space \( D([0, 1] \times [0, T]), T > 0, \) for \( j = 1 \) and \( j = 2 \). It then follows from Ferger and Vogel
(2015) that the joint process (4.5) is also tight.

Hence the weak convergence is established. A Skorokhod construction yields the result.

The following lemma is useful for the proof of Theorem 1 b).

**Lemma 7.** For every \( \delta > 0 \), there exists \( 0 < a < 1 \) such that for large \( n \),

\[
P \left\{ U_0 \left( \frac{aN}{ks} \right) \leq X_{N-[ks]:N} \leq U_0 \left( \frac{N}{aks} \right) \text{ for } \frac{1}{2k} \leq s \leq 1 \right\} > 1 - \delta.
\]

**Proof.** Condition (1.1) implies, by inversion, that there exists \( M > 1 \) and \( t_0 > 0 \) such that

\[
U_0 \left( \frac{t}{M} \right) \leq U_{ij}(t) \leq U_0(Mt)
\]

(4.6)
holds for all \( t \geq t_0 \).

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

\[
P \left\{ \frac{bn}{ks} \leq Y_{n-[ks]:n}^{(j)} \text{ for } \frac{1}{2k} \leq s \leq 1 \text{ and } Y_{n-[ks]:n}^{(j)} \leq \frac{n}{kbs} \text{ for } 0 \leq s \leq 1; j = 1, 2, \ldots, m \right\} > 1 - \delta/2. \quad (4.7)
\]

Note also that

\[
Y_{n-[ks]:n}^{(1)} \leq Y_{N-[ks]:N} \leq \max_{1 \leq j \leq m} Y_{n-[ks]:n}^{(j)} \quad (4.8)
\]

Further, using (4.6), with probability tending to 1,

\[
U_0 \left( \frac{1}{M} Y_{N-[ks]:N} \right) \leq X_{N-[ks]:N} \leq U_0(M Y_{N-[ks]:N}).
\]

The result follows combining (4.6), (4.7) and (4.8).

**Proof of Theorem 1 b) Tail empirical quantile function**

We start from (2.3) in Theorem 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,

\[
\sup_{x_0 \leq x < x_1} (1 + \gamma x)^{\eta/\gamma} \left| \sqrt{k} \left( P_n(x) - (1 + \gamma x)^{-1/\gamma} \right) - \sum_{j=1}^{m} W_j((1 + \gamma x)^{-1/\gamma}, C_j(1)) \right| \to 0, \quad (4.9)
\]
Proofs

where
\[ P_n(x) = \frac{1}{k} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbb{1}_{\left\{ \frac{x_{i,j} - x_0}{\frac{x_{i,j}}{y_0}} > x \right\}} \]

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

\[ P\left\{ \sup_{x_0 \leq x < x_1} (1 + \gamma x)^{\eta/\gamma} \left| \sqrt{k} \left( P_n(x) - (1 + \gamma x)^{-1/\gamma} \right) - \sum_{j=1}^{m} W_j((1 + \gamma x)^{-1/\gamma}, C_j(1)) \right| > \delta_1 \right\} \to 0, \quad (4.10) \]

We remark that the region \( x_0 \leq x < x_1 \) has different implications for \( \gamma \geq 0 \) and \( \gamma < 0 \). For \( \gamma \geq 0 \), 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 \( \infty \).

Then, split the range of \( s \) in two subintervals, \([1/(2k), s_0] \) and \([s_0, T]\), where \( s_0 \) is a sufficiently low but fixed constant. The upper bound of \( s_0 \) will be determined throughout the proof.

For the range \([s_0, T]\) we use (4.9) and Vervaat’s Lemma (cf. e.g. Appendix A of de Haan and Ferreira, 2006) with \( x_n(s) := P_n(s) \) and \( x_n^+(s) := (X_{N - \lfloor k \rfloor} N - b_0(N/k))/a_0(N/k) \). 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 \( \delta > 0 \) there exists an \( s_0 \), such that

\[ P\left\{ \sup_{0 < s \leq s_0} \left| W_0(s) \right| > \delta \right\} > 1 - \delta. \]

Now \( W_j(., C_j(1)) \overset{d}{=} \sqrt{C_j(1)} W_0 \) for all \( j \). Hence for a given \( \delta > 0 \), there exists an \( s_0(\delta) \), such that for all \( s_0 \leq s_0(\delta) \),

\[ P\left\{ \sup_{0 < s \leq s_0} \left| \sum_{j=1}^{m} W_j(s, C_j(1)) \right| < \delta \right\} > 1 - \delta. \quad (4.11) \]

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

\[ \sup_{\frac{1}{k} \leq s \leq s_0} s^{\gamma + \frac{\epsilon}{2} + \epsilon} \sqrt{k} \left( \frac{X_{N - \lfloor k \rfloor + N - b_0 \left( \frac{N}{k} \right)}}{a_0 \left( \frac{N}{k} \right)} - s^{-\gamma} - \frac{1}{\gamma} \right) \leq \delta, \quad (4.12) \]

and

\[ \inf_{\frac{1}{k} \leq s \leq s_0} s^{\gamma + \frac{\epsilon}{2} + \epsilon} \sqrt{k} \left( \frac{X_{N - \lfloor k \rfloor + N - b_0 \left( \frac{N}{k} \right)}}{a_0 \left( \frac{N}{k} \right)} - s^{-\gamma} - \frac{1}{\gamma} \right) \geq -\delta. \quad (4.13) \]

Again, we split the range of \( s \) in two subintervals, \([1/(2k), t_n] \) and \([t_n, s_0] \), where \( t_n \) depends only on the constant \( a \) in Lemma 7 (eventually depending on \( \delta \)) and a sufficiently small \( \xi > 0 \), although differently
for the upper and lower bounds in (4.12):
\[ k^{-1/2+\varepsilon} := \begin{cases} \delta \Delta_1^{-1} \text{ with } \Delta_1 := \frac{\gamma}{\gamma+1}(1 + \xi) > 0, \text{ for the upper bound,} \\
\delta \Delta_2^{-1} \text{ with } \Delta_2 := \frac{1}{\gamma+1}(1 + \xi) > 0, \text{ for the lower bound.} \end{cases} \]

(A) Upper bound and \( s \in [(2k)^{-1}, t_n] \): For simplicity, we assume that \( A_0 \) is eventually positive in the rest of the proof. Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 7 imply: for all \( \varepsilon, \delta, \theta > 0 \), there exists \( 0 < a < 1 \) such that for large \( n \), with probability at least 1 - \( \delta \),
\[
\frac{X_{N - [k\beta]N - b_0 \left( \frac{N}{k} \right)}}{a_0 \left( \frac{N}{k} \right)} - s^{-\gamma} - 1 \\
\leq \frac{(as)^{-\gamma} - s^{-\gamma}}{\gamma} + \overline{\Psi}_{\nu,\rho} \left( \frac{1}{as} \right) A_0 \left( \frac{N}{k} \right) + (as)^{-\gamma-\rho-\theta} A_0 \left( \frac{N}{k} \right)
\]
for some \( K > 0 \). Hence,

\[
s^{\gamma+1/2+\varepsilon} \sqrt{k} \left( \frac{X_{N - [k\beta]N - U_0 \left( \frac{N}{k} \right)}}{a_0 \left( \frac{N}{k} \right)} - s^{-\gamma} - 1 \right) \leq s^{1/2+\varepsilon} \sqrt{k} \left\{ \frac{a^{-\gamma} - 1}{\gamma} + s^{-\rho-\theta} K A_0 \left( \frac{N}{k} \right) \right\}
\]
\[
\leq \frac{\delta}{\Delta_1} \left\{ \frac{a^{-\gamma} - 1}{\gamma} + s^{-\rho-\theta} K A_0 \left( \frac{N}{k} \right) \right\} \leq \delta, \quad (4.14)
\]

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 \( n \),

\[
P_n \left( \frac{s^{-\gamma} - 1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma-1/2-\varepsilon} \right) \leq s, \quad \text{for all } s \in (t_n, s_0], \quad (4.15)
\]

which implies the upper bound in (4.12).

We start by giving some technical relations that provide the constants to determine an upper bound of \( s_0 \). Therefore note that for large enough \( \eta \) there exists \( \eta' \) with \( 1 - \eta < 1 - \eta' < 1/2 + \varepsilon \), such that the
Proofs

following hold:

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

\[
\sup_{s \in (t_n, s_0]} \frac{s^{-(1-\eta')}}{\sqrt{k}} \leq \frac{t_n^{-(1-\eta')}}{\sqrt{k}} = \frac{t_n^{-(1-\eta'+2\varepsilon)}}{\sqrt{k}t_n^{1/2+\varepsilon}} \to 0,
\]

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

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

where the last two inequalities follow from the inequalities \((1 + \gamma x)^{-1/\gamma} \leq 1 - c_I x \) and \((1 - \gamma x)^{-1/\gamma} \geq 1 + c_I x\) respectively, for \(0 < x < \min(1, 1/(\gamma - \varepsilon))\) and some \(0 < c_I, c_I \leq 1\). Take \(s_0 \leq (\delta c_I/(1 + \delta_1))^{(\eta' - \eta)^{-1}}\).

We intend to apply (4.10) 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 larger or equal to \(s_0^{-\gamma} \left( 1 + \delta \gamma k^{-1/2} s_0^{-1/2-\varepsilon} \right)\), consequently bounded away from zero. For \(\gamma < 0\) the inequality is reversed and the expression is bounded above. Hence,

\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma - 1/2-\varepsilon} \right) \leq s \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2-\varepsilon} \right)^{-1/\gamma} + \frac{1}{\sqrt{k}} \overline{W} \left( s \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2-\varepsilon} \right)^{-1/\gamma} \right) + \frac{\delta_1}{\sqrt{k}} s^{\eta} \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2-\varepsilon} \right)^{-\eta/\gamma}
\]

\[
\leq s \left( 1 - c_I \frac{\delta}{\sqrt{k}} s^{-(1-\eta')} \right) + \frac{1 + \delta_1}{\sqrt{k}} s^{\eta} = s - s_0^{\eta} \left( c_I \delta s^{\eta' - \eta} - (1 + \delta_1) \right)
\]

with \(\overline{W}(s) := \sum_{j=1}^{n} W_j(s, C_j(1))\) and where for the second inequality we have applied (4.16), \(\eta < 1/2\), (4.11) and the fact that

\[
\left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2-\varepsilon} \right)^{-1/\gamma} \leq 1.
\]

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

(C) Lower bound and \(s \in [(2k)^{-1}, t_n]\): Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 7
Proofs

imply: for all \( \varepsilon, \delta, \theta > 0 \), there exists \( 0 < a < 1 \) such that for large \( n \), with probability at least \( 1 - \delta \),

\[
\frac{X_{N-[k_s]} - b_0 \left( \frac{N}{k_s} \right)}{a_0 \left( \frac{N}{k_s} \right)} - \frac{s^{-\gamma} - 1}{\gamma} \geq \frac{(a/s)^\gamma - s^{-\gamma}}{\gamma} + \Psi_{\gamma, \rho} \left( \frac{a}{s} \right) A_0 \left( \frac{N}{k_s} \right) - \left( \frac{a}{s} \right)^{\gamma + \rho + \theta} A_0 \left( \frac{N}{k_s} \right)
\]

for some \( K > 0 \), hence,

\[
s^{\gamma + 1/2 + \varepsilon} \sqrt{k} \left( \frac{X_{N-[k_s]} - b_0 \left( \frac{N}{k_s} \right)}{a_0 \left( \frac{N}{k_s} \right)} - \frac{s^{-\gamma} - 1}{\gamma} \right) \geq s^{1/2 + \varepsilon} \sqrt{k} \left\{ \frac{a^\gamma - 1}{\gamma} - s^{-\rho - \theta} K A_0 \left( \frac{N}{k_s} \right) \right\} \\
\geq \frac{\delta}{\Delta_2} \left\{ \frac{a^\gamma - 1}{\gamma} - s^{-\rho - \theta} K A_0 \left( \frac{N}{k_s} \right) \right\} \geq -\delta,
\]

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

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

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

which implies the lower bound in (4.12). Similarly as in (B) apply (4.10) with \( x \) replaced by \( \gamma^{-1}(s^{-\gamma} - 1) - \delta k^{-1/2} s^{-\gamma - 1/2 - \varepsilon} \). Note that,

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

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

\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} - \frac{\delta}{\sqrt{k}} s^{-\gamma - 1/2 - \varepsilon} \right) \\
\geq s \left( 1 - \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-1/\gamma} + \frac{1}{\sqrt{k}} \left( 1 - \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-1/\gamma} - \frac{\delta_1}{\sqrt{k}} s^\eta \left( 1 - \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-\eta/\gamma} \\
\geq s \left( 1 + c_{II} \frac{\delta}{\sqrt{k}} s^{-1/2 - \varepsilon} \right) - s^\eta (1 + \xi) - s^\eta (1 + \delta_1) \frac{s^\eta}{\sqrt{k}} \\
= s + \frac{s^\eta}{\sqrt{k}} \left( c_{II} s^{\eta' - \eta} - s^\eta (1 + \xi) - s^\eta (1 + \delta_1) \right)
\]

where we have used in particular (4.16). It remains to check that the right-hand side of (4.19) is larger or equal to \( s + 1/k \) which is equivalent to \( s^\eta \sqrt{k} \left( c_{II} s^{\eta' - \eta} - s^\eta (1 + \xi) - s^\eta (1 + \delta_1) \right) \geq 1 \). This holds by
Proofs

the choice of $\eta' < \eta$ and choosing $s_0 \leq \delta c_{11}/((m/a)\eta(1+\xi)^{-\eta/\gamma}(1+\delta_1))^{(\eta-\eta')^{-1}}$.

The theorem is thus proved by choosing $s_0$ with combining all aforementioned upper bounds, i.e.
\[ s_0 \leq \min(s_0(\delta), (\delta c_{11}/(1+\delta_1))^{(\eta-\eta')^{-1}}, \delta c_{11}/((m/a)\eta(1+\xi)^{-\eta/\gamma}(1+\delta_1))^{(\eta-\eta')^{-1}}) \]

\[ \square \]

Proof of Theorem 3

Fix $j \in \{1, \ldots, m\}$. Replace $x$ in (2.3) 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
\[ \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I}\{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) - W_j \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) \}
\[ = o_p \left( 1 + \gamma \frac{X_{N-k,N} - b_0\left(\frac{N}{k} \right)}{a_0\left(\frac{N}{k} \right)} \right)^{-\eta/\gamma}. \]

Now by (2.4)
\[ \sqrt{k} \frac{X_{N-k,N} - b_0\left(\frac{N}{k} \right)}{a_0\left(\frac{N}{k} \right)} \xrightarrow{p} \sum_{j=1}^{m} W_j(1, C_j(1)) \]
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
\[ \sup_{0 \leq t \leq 1} \left| \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I}\{X_{i,j} > X_{N-k,N}\} - C_j(t) \right\} - \left( W_j(1, C_j(1)) - C_j(t) \sum_{r=1}^{m} W_r(1, C_r(1)) \right) \right| \xrightarrow{p} 0, \]
which yields (2.6).

To prove (2.7), we first prove the following limit relation: as $n \to \infty$,
\[ \frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{X_{i,j_1} - b_0\left(\frac{N}{k} \right)}{a_0\left(\frac{N}{k} \right)} > w_1 x - 1, \frac{X_{i,j_2} - b_0\left(\frac{N}{k} \right)}{a_0\left(\frac{N}{k} \right)} > w_2 x - 1 \right) \xrightarrow{p} \frac{1}{m} \int_{0}^{t} R_{j_1,j_2}(w_1 c(u,j_1), w_2 c(u,j_2)) du, \]
for fixed $(t, w_1, w_2) \in [0,1] \times [0,T]^2$. Since $X_{i,j} = U_{i,j}(Y_{i,j})$ with $Y_{i,j}$ standard Pareto distributed random
Proofs

variables, given any \( \varepsilon > 0 \), for sufficiently large \( n \),

\[
P \left( \frac{X_{i,j_1} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{i}^{-\gamma} - 1}{\gamma}, \frac{X_{i,j_2} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{j}^{-\gamma} - 1}{\gamma} \right)
\]

\[
= P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{N} \left( 1 - F_{i,j_1} \left( b_0 \left( \frac{N}{k} \right) + a_0 \left( \frac{N}{k} \right) \frac{w_{i}^{-\gamma} - 1}{\gamma} \right) \right), \frac{k}{n} Y_{1,j_2} > \frac{m}{N} \left( 1 - F_{i,j_2} \left( b_0 \left( \frac{N}{k} \right) + a_0 \left( \frac{N}{k} \right) \frac{w_{j}^{-\gamma} - 1}{\gamma} \right) \right) \right)
\]

\[
\leq P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{w_{1,c} \left( \frac{N}{k}, j_1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{1,j_2} > \frac{m}{w_{2,c} \left( \frac{N}{k}, j_2 \right) (1 + \varepsilon)} \right)
\]

The definition of \( R_{j_1,j_2} \) implies that, as \( n \to \infty \),

\[
\frac{n}{k} P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{w_{1,c} \left( \frac{N}{k}, j_1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{1,j_2} > \frac{m}{w_{2,c} \left( \frac{N}{k}, j_2 \right) (1 + \varepsilon)} \right) \to R_{j_1,j_2}(v_1, v_2),
\]

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

\[
\frac{n}{k} P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{w_{1,c} \left( \frac{N}{k}, j_1 \right) (1 + \varepsilon)} \right) \to 0, \quad (4.21)
\]

uniformly in \( i, w_1 \) and \( w_2 \). Hence, as \( n \to \infty \), uniformly in \( (t, w_1, w_2) \in [0, 1] \times [0, T]^2 \),

\[
\frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{X_{i,j_1} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{i}^{-\gamma} - 1}{\gamma}, \frac{X_{i,j_2} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{j}^{-\gamma} - 1}{\gamma} \right)
\]

\[
\leq \frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{w_{1,c} \left( \frac{N}{k}, j_1 \right) (1 + \varepsilon)} \right) \to 0, \quad (4.20)
\]

\[
= \int_0^t \frac{n}{k} P \left( \frac{k}{n} Y_{1,j_1} > \frac{m}{w_{1,c} \left( \frac{N}{k}, j_1 \right) (1 + \varepsilon)} \right) du.
\]

A lower bound is derived similarly. Hence, by the homogeneity of the function \( R \), (4.20) holds.

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

\[
\frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{X_{i,j_1} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{i}^{-\gamma} - 1}{\gamma}, \frac{X_{i,j_2} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{w_{j}^{-\gamma} - 1}{\gamma} \right) \to \frac{1}{m} \int_0^t R_{j_1,j_2}(w_{1,c}(u), w_{2,c}(u)) du. \quad (4.22)
\]

21
Proofs

We check the variance of the left hand side: using (4.20), as \( n \to \infty, \)

\[
Var \left( \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I} \left\{ \frac{x_{i,j_1} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \frac{x_{i,j_2} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \right\} \right) = \sum_{i=1}^{nt} \frac{1}{k^2} Var \left( \mathbb{I} \left\{ \frac{x_{i,j_1} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \frac{x_{i,j_2} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \right\} \right) \leq \sum_{i=1}^{nt} \frac{1}{k^2} P \left( \frac{X_{i,j_1} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \frac{X_{i,j_2} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \right) \leq \frac{1}{km} (1 + \varepsilon) \int_0^T R_{j_1,j_2}(w_1 c(u,j_1), w_2 c(u,j_2)) du \to 0.
\]

Hence, by Chebyshev inequality, as \( n \to \infty, \)

\[
\frac{1}{k} \sum_{i=1}^{nt} \mathbb{I} \left\{ \frac{x_{i,j_1} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \frac{x_{i,j_2} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} < \frac{1}{\gamma} \right\} \leq \frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{X_{i,j_1} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} > \frac{w_{1}^{1-\gamma} - 1}{\gamma}, \frac{X_{i,j_2} - b_0 \left( \frac{n}{k} \right)}{a_0 \left( \frac{n}{k} \right)} > \frac{w_{2}^{1-\gamma} - 1}{\gamma} \right) \Rightarrow 0. \tag{4.23}
\]

Then, (4.20) implies (4.22).

By the continuity of the right hand side of (4.22) and the monotonicity of both sides of (4.22) in \( t, w_1 \text{ and } w_2, \) 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 \( n \to \infty\)

\[
\left( 1 + \gamma \frac{X_{N-|s_1|} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} \right)^{-1/\gamma} \Rightarrow s_j,
\]

for \( j = 1, 2. \) We can then replace \( w_j \) in (4.22) with \( \left( 1 + \gamma \frac{X_{N-|s_1|} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} \right)^{-1/\gamma}, \) which yields (2.7) for fixed \((t, s_1, s_2) \in [0,1] \times [0,T]^2.\) The uniformity follows as before.

\[\blacksquare\]

**Proof of Corollary 5** Under the null hypothesis, using Theorem 3, under a Skorokhod construction,

\[
\sqrt{k}(\hat{C}_j(t) - t\hat{C}_j(1)) = \sqrt{k}(\hat{C}_j(t) - C_j(t)) - t\sqrt{k}(\hat{C}_j(1) - C_j(1))
\]
Proofs

converges uniformly \((0 \leq t \leq 1)\) to

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

\[
=W_j(1, tC_j(1)) - tW_j(1, C_j(1)) \overset{d}{=} \sqrt{C_j(1)} B(t).
\]

Then the result follows directly via Slutsky’s theorem.

The proof of Theorem 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 \(h = (h_1, h_2) \in \mathbb{R}^2:\)

\[
\begin{align*}
  h_1 &= \sqrt{k} \left( \gamma_{N/k} - \gamma_0 \right) \\
  h_2 &= \sqrt{k} \left( \sigma_{N/k}/a_0(N/k) - 1 \right)
\end{align*}
\]

\[
\L_{N,k}(h) = \int_0^1 \ell \left( \gamma_0 + \frac{h_1}{\sqrt{k}} a_0 \left( \frac{N}{k} \right) \left( 1 + \frac{h_2}{\sqrt{k}}, X_{N-|k|,N} - X_{N-k,N} \right) \right) ds
\]

\[
= k \int_0^1 \ell \left( \gamma_0 + \frac{h_1}{\sqrt{k}}, 1 + \frac{h_2}{\sqrt{k}} \frac{X_{N-|k|,N} - X_{N-k,N}}{a_0 \left( \frac{N}{k} \right)} \right) ds - k \log a_0 \left( \frac{N}{k} \right),
\]

and

\[
\begin{align*}
  \frac{\partial \L_{N,k}}{\partial h_1}(h_1, h_2) &= \frac{1}{\sqrt{k}} \frac{\partial \L_{N,k}}{\partial \gamma}(\gamma, \sigma_{N/k}) = 0 \\
  \frac{\partial \L_{N,k}}{\partial h_2}(h_1, h_2) &= \frac{1}{\sqrt{k}} \frac{\partial \L_{N,k}}{\partial \sigma}(\gamma, \sigma_{N/k}) = 0.
\end{align*}
\]

The main steps of the proof are as follows:

**a)** First prove that,

\[
\frac{\partial^2 \L_{N,k}(h)}{\partial \theta^2} = -I_{\gamma_0} + o_P(1), \quad I_{\gamma_0} = -\int_0^1 \frac{\partial^2 \ell}{\partial \theta \partial \theta^2} \left( \gamma_0, 1, \frac{s^{-\gamma_0} - 1}{\gamma_0} \right) ds, \quad \theta = (\gamma, \sigma) \in \mathbb{R} \times (0, \infty),
\]

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

\[
I_{\gamma_0} = \left( \begin{array}{cc}
\frac{2}{1 + 3\gamma_0 + 2\gamma_0^2} & \frac{1}{1 + 3\gamma_0 + 2\gamma_0^2} \\
\frac{1}{1 + 3\gamma_0 + 2\gamma_0^2} & \frac{1}{1 + 2\gamma_0}
\end{array} \right).
\]

\(I_{\gamma_0}\) 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):

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

(4.27)

where

$$
\frac{\partial L_{N,k}}{\partial h}(0) = \frac{1}{\sqrt{K}} \sum_{i=1}^K \frac{\partial \ell}{\partial \theta} \left( \gamma_{0,i}, 1, X_{N-i+1,N} - X_{N-k,N} \right) \xrightarrow{d} N(0, \Sigma_{\gamma_0}).
$$

(4.28)

**c)** Finally the Argmax Theorem (van der Vaart 1998, Corollary 5.58) provides the result: let

$$
M_n(h) = \bar{L}_{N,k}(h) - \bar{L}_{N,k}(0),
$$

$$
M(h) = h^T N(0, \Sigma_{\gamma_0}) - \frac{1}{2} h^T I_{\gamma_0} h, \quad h \in \mathbb{R}^2.
$$

Then,

$$
\hat{h}_n = \arg\max_{h \in H_n} M_n(h) \xrightarrow{d} h = \arg\max_{h \in \mathbb{R}^2} M(h) \xrightarrow{d} \Gamma_{\gamma_0}^{-1} N(0, \Sigma_{\gamma_0})
$$

provided $\hat{h}_n$ is tight which holds as in Dombrzy and Ferreira (2018). Finally, note that

$$
\bar{L}_{N,k}(\hat{h}_n) = L_{N,k} \left( \gamma_{0,i} + \frac{\hat{h}_n_{1}}{\sqrt{K}} a_0 \left( \frac{N}{K} \right) (1 + \frac{\hat{h}_n_{2}}{\sqrt{K}}) \right)
$$

and similarly for its derivatives.

**Acknowledgements**

John Einmahl holds the Arie Kapteyn Chair 2019-2022 and gratefully acknowledges the corresponding research support. Ana Ferreira was partially supported by FCT-Portugal: UID/Multi/04621/2019, 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/28640/2017. Cláudia Neves gratefully acknowledges support from EPSRC-UKRI Innovation Fellowship grant EP/S001263/1 and project FCT-UIDB/00006/2020.

**References**

REFERENCES


Supporting information

The detailed proof of Theorem 6, referenced in Section 4, as well as simulations showcasing finite sample performance of the proposed methods are available with this paper at https://bit.ly/3aJFM6B.