# Ordinal patterns in clusters of subsequent extremes of regularly varying time series

**Marco Oesting**[1,2] **· Alexander Schnurr**[1]

## Abstract

In this paper, we investigate temporal clusters of extremes defined as subsequent exceedances of high thresholds in a stationary time series. Two meaningful features of these clusters are the probability distribution of the cluster size and the ordinal patterns giving the relative positions of the data points within a cluster. Since these patterns take only the ordinal structure of consecutive data points into account, the method is robust under monotone transformations and measurement errors. We verify the existence of the corresponding limit distributions in the framework of regularly varying time series, develop non-parametric estimators and show their asymptotic normality under appropriate mixing conditions. The performance of the estimators is demonstrated in a simulated example and a real data application to discharge data of the river Rhine.

## 1 Introduction

In time series data sets, extremes often do not occur at scattered instants of time, but tend to form clusters. Assigning a cluster of extremes to a single extreme event, such as a flood in the context of a hydrological time series or a stock market crash in the context of a financial data, the distribution of these clusters is crucial for risk assessment.

✉ Marco Oesting
oesting@mathematik.uni-siegen.de; marco.oesting@mathematik.uni-stuttgart.de

1  Department Mathematik, University of Siegen, Walter-Flex-Str. 3, 57068 Siegen, Germany

2  University of Stuttgart, Stuttgart Center for Simulation Science (SC SimTech) & Institute for Stochastics and Applications, Allmandring 5b, 70569 Stuttgart, Germany

In order to analyze the occurrence times of extremes defined as exceedances over some high threshold $u$, some profound theory has been built up since the 1970s. Within this framework, data $X_1, \ldots, X_n$ from a stationary time series $(X_t)_{t \in \mathbb{Z}}$ are typically divided into different blocks. Then, repeated extremes are said to form a cluster if they occur within the same temporal block. Due to the convergence of the process of exceedances to a Poisson point process under appropriate conditions as $u \to \infty$, the distribution of these clusters converges weakly provided that the block size increases at the right speed. The limit distribution is nicely linked to the well-known concept of the extremal index of the time series which can be interpreted as the reciprocal of the mean limiting cluster size (cf. Leadbetter et al. 1983; Embrechts and Klüppelberg 1997; Chavez-Demoulin and Davison 2012, for an overview). Besides the extremal index, several other cluster characteristics are of interest and can be estimated, such as the distribution of the cluster size (Robert 2009) or more general cluster functionals (Drees and Rootzén 2010). Convergence of clusters in appropriate sequence spaces preserving the order of observations can be shown within the framework of regular variation (Basrak et al. 2018).

Even though positive theoretical results exist, estimation of characteristics of clusters as defined above is difficult for finite samples. Here, besides the threshold $u$, also the block size or, equivalently, some cluster identification parameter giving the minimum distance between two separate clusters, needs to be chosen. Instead, in this paper, we will use a different definition of a cluster of extremes by restricting our attention to subsequent threshold exceedances, i.e. a realization of the $l$-dimensional vector $(X_i)_{i=t}^{t+l-1}$ will be called a $u$-exceedance cluster of size $l \in \mathbb{N}$ if and only if

$$X_{t-1} \leq u, X_t > u, \ldots, X_{t+l-1} > u \text{ and } X_{t+l} \leq u. \tag{1}$$

As any non-exceedance will separate two clusters, this definition is much stricter than the classical definition described above. An advantage of the definition of $u$-exceedance clusters is that it depends on one parameter, namely the threshold $u$, only. Such a cluster definition has already been employed in a series of papers by Markovich (2014, 2016, 2017) who analyzes the limit distribution of two cluster characteristics. First, she considers the number of inter-cluster times $T_1(u)$, i.e. the number of observations between two consecutive clusters, which is a random variable with the same distribution as

$$\min\{j \geq 1 : X_{j+1} > u\} \quad \text{conditional on } X_1 > u.$$

Note that this number of inter-cluster time also plays an important role in the estimation of the extremal index (Ferro and Segers 2003). Secondly, she studies the random variable $T_2(u)$ with the same distribution as

$$\min\{j \geq 1 : X_{j+1} \leq u\} \quad \text{conditional on } X_1 \leq u,$$

i.e. $T_2(u) - 1$ is the length of a $u$-exceedance cluster starting at some fixed time. Since we have $\lim_{u \to \infty} \mathbb{P}(X_2 \leq u) = 1$, the distribution of $T_2(u)$ is typically expected to converge weakly to a degenerate distribution, i.e. we have $\lim_{u \to \infty} \mathbb{P}(T_2(u) = 1) = 1$. In Markovich (2014, 2016), under appropriate mixing conditions, the rate of convergence is determined as a function of the extremal index. More precisely, for all $\varepsilon > 0$, there exist a threshold $u_0 = u_0(\varepsilon)$ a number $j_0 = j_0(\varepsilon)$ such that, for all

$u > u_0$ and $j > j_0$,

$$\left| \frac{1 - \mathbb{P}(X_0 > u)^\theta}{\theta^2(1 - \mathbb{P}(X_0 > u))} \cdot \frac{\mathbb{P}(T_2(u) = j)}{[1 - \mathbb{P}(X_0 > u)^\theta]\{\mathbb{P}(X_0 > u)^\theta\}^{j-1}} - 1 \right| < \varepsilon,$$

that is, for a sufficiently large threshold $u$, the tail of the distribution of $T_2(u)$ becomes approximately proportional to a geometric distribution with parameter $1 - \mathbb{P}(X_0 > u)^\theta$. Furthermore, Markovich (2014, 2016) provides results for the duration of clusters if the time between subsequent observations is random with regularly varying distribution.

In our paper, we will focus on the case of an equally spaced time series, i.e. the case when the terms of the duration and the size of cluster coincide. Here, instead of considering the probability that there is a cluster of a specific size at a certain time, we analyze the size of a a randomly chosen $u$-exceedance cluster or, equivalently, we examine the size of a cluster conditional on being a cluster of positive length. Thus, we first address the question:

How long does an extreme event in a time series last provided that it occurs?

Secondly, we analyze the distribution of so-called ordinal patterns which we find in the above mentioned clusters of extremes. Ordinal patterns keep the ordinal information of the data only and, thus, describe their 'up-and-down behavior'. Here, the relative position of the data points $x_0, \dots, x_{l-1}$ is encoded by a permutation $\pi$ on $\{0, \dots, l-1\}$ such that

$$x_{\pi(0)} \geq x_{\pi(1)} \geq \dots \geq x_{\pi(l-1)}.$$

Examples of ordinal patterns are shown in Fig. 1.

Note that this permutation is unique if the data points $x_0, \dots, x_{l-1}$ are pairwise distinct. The following precise definition also accounts for the ties by keeping the order of the indices in this case.

**Definition 1** For $l \in \mathbb{N}$, let $S_{l-1}$ be the set of permutations of $\{0, \dots, l-1\}$. The *l-ordinal pattern* is defined as the mapping $\Pi : \mathbb{R}^l \to S_{l-1}$ that maps a vector $(x_i)_{i=0}^{l-1}$ to the unique permutation $\pi$ satisfying

$$x_{\pi(0)} \geq x_{\pi(1)} \geq \dots \geq x_{\pi(l-1)}$$

and $\pi(\min\{i, j\}) < \pi(\max\{i, j\})$ if $x_i = x_j$ for $i \neq j$.
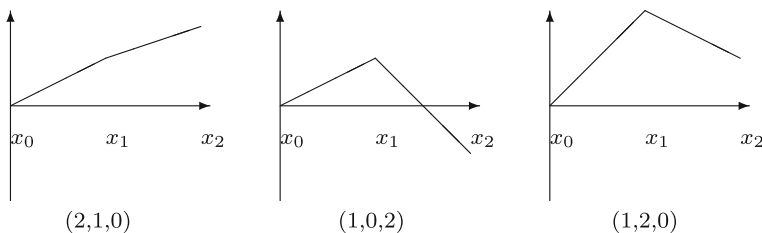


Fig. 1 Examples for ordinal patterns of size 3

Ordinal patterns have been introduced in order to analyze noisy data sets which appear in medicine, neuroscience and finance (cf. Bandt and Pompe 2002; Keller et al. 2007; Sinn et al. 2013). They have already been used successfully in the estimation of the Hurst parameter (Sinn and Keller 2011). Further applications include tests for structural breaks (Sinn et al. 2012) and the analysis of the Kolmogorov-Sinai entropy of dynamical systems (Keller et al. 2015).

Ordinal patterns can be used nicely to capture stylized facts as trends or inversions of the direction which might be used to characterize and classify certain events. To our knowledge the present paper is the first approach to analyze the ordinal behavior which can be observed in clusters of extremes of time series. The advantages of the proposed method include that the whole analysis is stable under monotone transformations of the state space. This will be useful in our analysis. Furthermore, the ordinal structure is typically not destroyed by small perturbations of the data or by measurement errors. There are fast algorithms to analyze the relative frequencies of ordinal patterns in given data sets (cf. Keller et al. 2007, Section 1.4).

In the future, ordinal patterns in clusters of extremes might be used to detect structural breaks in the extremes of the given time series (cf. Unakafov and Keller 2018). Dealing with 'correlated' time series, one could analyze the dependence between extreme events in a non-linear fashion as it has been developed in Schnurr (2014) and Schnurr and Dehling (2017). This might be advantageous in particular if the time series are on totally different scales. Finally, as we will point out in Section 6, ordinal patterns at the beginning of a cluster of extremes might be used in order to forecast the length of this cluster in an on-line analysis of data.

Our analysis is embedded in a different theoretical framework than the works of Markovich (2014, 2016, 2017), namely, we will assume that the stationary time series of interest, $(X_t)_{t \in \mathbb{Z}}$, is regularly varying. Note that this is a common assumption in extreme value theory allowing for convenient extrapolation to the tails of the distribution. More background on the theory of regularly varying time series will be provided in Section 2. In Section 3, we show that both the distribution of the size of $u$-exceedance clusters, as defined in Eq. (1), and the distribution of the ordinal pattern within a cluster converge to (typically non-degenerate) limit distributions in case of a regularly varying time series. Based on a sliding window approach, non-parametric empirical estimators for the limit distributions are introduced in Section 4. Under conditions, similar to those considered in Davis and Mikosch (2009) for the estimation of the extremogram, consistency (Proposition 1) and asymptotic normality (Corollary 1) of the estimators are established. In Section 5, we consider the example of max-stable time series and provide sufficient conditions in terms of extremal coefficients for Corollary 1 to hold. The conditions are verified for a Brown–Resnick time series which is then simulated to demonstrate the finite-sample behavior of the estimators. Finally, we apply the estimator to daily discharge data of the river Rhine at Cologne in Section 6. The proofs of our results can be found in the Appendix.

## 2 Background: Regular varying time series

Throughout this paper, we will assume that $X = (X_t)_{t \in \mathbb{Z}}$ is a stationary time series whose marginal distribution $F_0$, defined by $F_0(x) = \mathbb{P}(X_0 \leq x)$, is in the max-domain of attraction of an extreme value distribution, i.e., there exist constants $a_n > 0$, $b_n \in \mathbb{R}$, such that

$$F_0^n(a_n x + b_n) \overset{n \to \infty}{\longrightarrow} G_0(x), \quad x \in \mathbb{R},$$

for some non-degenerate distribution $G_0$. Without loss of generality, we may assume that $G_0$ is an $\alpha$-Fréchet distribution for some $\alpha > 0$,

$$G_0(x) = \Phi_\alpha(x) = \exp(-x^{-\alpha}), \qquad x > 0,$$

and that $F_0$ has a finite lower endpoint, $\inf\{x \in \mathbb{R} : F_0(x) > 0\} > -\infty$. Both properties can be achieved by applying strictly monotone marginal transformations to $(X_t)_{t \in \mathbb{Z}}$ provided that $F_0$ is continuous. As these transformations are the same for each $t \in \mathbb{Z}$ – remind that $X$ is stationary – they do not have any effect on ordinal structure of the data. In particular, ordinal patterns in extremes are invariant under these transformations.

A convenient framework for our further analysis will be provided by regular variation. Among several equivalent definitions of multivariate regular variation (cf. Resnick 2007, 2008, for instance), we will make use of the following convenient one (cf. Basrak et al. 2002, for instance): We say that the $d$-variate random vector $\mathbf{X} = (X_{t_1}, \ldots, X_{t_d})$, $t_1, \ldots, t_d \in \mathbb{Z}$, is *multivariate regularly varying* with index $\alpha > 0$ if, for some norm $\| \cdot \|$ on $\mathbb{R}^d$, there exists a probability measure $\sigma$ on the sphere $\mathbb{S}^{d-1} = \{\mathbf{x} \in [0, \infty)^d : \|\mathbf{x}\| = 1\}$ such that

$$\frac{\mathbb{P}(\|\mathbf{X}\| > rx, \ \mathbf{X}/\|\mathbf{X}\| \in \cdot)}{\mathbb{P}(\|\mathbf{X}\| > x)} \longrightarrow_w r^{-\alpha} \sigma(\cdot)$$

as $x \to \infty$, where $\to_w$ denotes weak convergence. The limit measure $\sigma$ is called spectral measure.

By Corollary 5.18 in Resnick (2008), multivariate regular variation of $\mathbf{X}$ with spectral measure $\sigma$ is equivalent to the fact that the distribution function $F$ of $\mathbf{X}$ is in the max-domain of attraction of a multivariate extreme value distribution $G$ with $\Phi_\alpha$ marginal distributions, i.e.

$$F^n(a_n x_1 + b_n, \ldots, a_n x_d + b_n) \overset{n \to \infty}{\longrightarrow} G(x_1, \ldots, x_d), \qquad x_1, \ldots, x_d > 0.$$

The limit distribution $G$ is of the form

$$G(x_1, \ldots, x_d) = \exp\left(-\mu\left\{[0, \infty)^d \setminus ([0, x_1] \times \ldots \times [0, x_d])\right\}\right),$$

for some Radon measure $\mu$ on $E = [0, \infty)^d \setminus \{\mathbf{0}\}$, the so-called *exponent measure $\mu$* of $G$. The exponent measure $\mu$ and the spectral measure $\sigma$ are related via

$$\mu(\{\mathbf{x} \in E : \|\mathbf{x}\| > r, \ \mathbf{x}/\|\mathbf{x}\| \in A\}) = r^{-\alpha} \sigma(A), \qquad r > 0, \ A \subset \mathbb{S}^{d-1}.$$

The time series $X$ is called *regularly varying* if all the finite-dimensional margins $(X_{t_1}, \ldots, X_{t_d})$, $t_1 < t_2 < \ldots < t_d \in \mathbb{Z}$, $d \in \mathbb{N}$, are multivariate regularly varying. By Basrak and Segers (2009), regular variation of $X$ is equivalent to the existence

of a process $Y = (Y_t)_{t \in \mathbb{Z}}$ with $\mathbb{P}(Y_0 > y) = y^{-\alpha}$ for $y \geq 1$ such that, for every $s < t \in \mathbb{Z}$,

$$\mathbb{P}\left(\left(\frac{X_s}{x}, \ldots, \frac{X_t}{x}\right) \in \cdot \,\middle|\, X_0 > x\right) \longrightarrow_d \mathbb{P}\left((Y_s, \ldots, Y_t) \in \cdot\right), \quad \text{as } x \to \infty. \quad (2)$$

The process $Y$ is called *tail process* of $X$, see Basrak and Segers (2009) and Planinić and Soulier (2018) for more details and further properties.

In the following, we will always assume that the time series $X$ is regularly varying with tail process $Y$. Furthermore, the probability measure induced by the random vector $(Y_i)_{i \in I}$ will be called $\mu_I$ for any index set $I \subset \mathbb{Z}$.

## 3 Distribution of clusters of extremes and ordinal patterns

In this section, we will analyze the limiting behavior of the size and ordinal pattern of $u$-exceedance clusters in the time series $X$ with $u$-exceedances being defined as in Eq. (1). Intuitively, the following expression gives a plausible definition of the distribution of the size $C_u$ of a $u$-exceedance cluster:

$$\mathbb{P}(C_u = l) = \lim_{n \to \infty} \frac{\#\{u\text{-exceedance clusters of size } l \text{ in } (X_t)_{t=-1}^n\}}{\#\{u\text{-exceedance clusters (of any size) in } (X_t)_{t=-1}^n\}}$$

$$= \lim_{n \to \infty} \frac{\sum_{k=0}^{n-l} \mathbf{1}_{\{X_{k-1} \leq u, X_k > u, \ldots, X_{k+l-1} > u, X_{k+l} \leq u\}}}{\sum_{l \in \mathbb{N}} \sum_{k=0}^n \mathbf{1}_{\{X_{k-1} \leq u, X_k > u, \ldots, X_{k+l-1} > u, X_{k+l} \leq u\}}}$$

$$= \lim_{n \to \infty} \frac{\sum_{k=0}^{n-l} \mathbf{1}_{\{X_{k-1} \leq u, X_k > u, \ldots, X_{k+l-1} > u, X_{k+l} \leq u\}}}{\sum_{k=0}^n \mathbf{1}_{\{X_{k-1} \leq u, X_k > u\}}}, \qquad l \in \mathbb{N}.$$

If $X$ is ergodic, we can apply the pointwise Birkhoff-Khinchin theorem and obtain that the above limit almost surely exists and equals

$$\mathbb{P}(C_u = l) = \frac{\mathbb{P}(X_{-1} \leq u, X_0 > u, \ldots, X_{l-1} > u, X_l \leq u)}{\mathbb{P}(X_{-1} \leq u, X_0 > u)}, \quad l \in \mathbb{N}. \quad (3)$$

Dividing both the enumerator and the denominator of Eq. (3) by $\mathbb{P}(X_0 > u)$, we can see from relation (2) that the distribution of $C_u$ eventually becomes independent from the threshold $u$ as $u \to \infty$ provided that

$$\mu_{\{-1,\ldots,l\}}\{(x_i)_{i=-1}^l \in [0, \infty)^{l+2} : x_i = 0 \text{ or } x_i = 1 \text{ for some } i = -1, \ldots, l\} = 0$$

for every $l \in \mathbb{N}$:

$$\lim_{u \to \infty} \mathbb{P}(C_u = l) = \lim_{u \to \infty} \frac{\mathbb{P}(X_{-1} \leq u, X_0 > u, \ldots, X_{l-1} > u, X_l \leq u \mid X_0 > u)}{\mathbb{P}(X_{-1} \leq u, X_0 > u \mid X_0 > u)}$$

$$= \frac{\mathbb{P}(Y_{-1} \leq 1, Y_0 > 1, \ldots, Y_{l-1} > 1, Y_l \leq 1)}{\mathbb{P}(Y_{-1} \leq 1, Y_0 > 1)}$$

$$= \frac{\mu_{\{-1,\ldots,l\}}([0, 1] \times (1, \infty)^l \times [0, 1])}{\mu_{\{-1,0\}}([0, 1] \times (1, \infty))}, \quad l \in \mathbb{N}. \quad (4)$$

*Example 1* We consider two examples to compare the limiting distribution of $C_u$ as $u \to \infty$, i.e. the distribution of the cluster size according to our definition, with the limiting cluster size distribution in the classical setting which has been studied extensively in the literature, cf. Robert (2009) and references therein. While both distributions coincide in the first examples, they significantly differ in the second one.

1. We consider a first order max-autoregressive model (cf. Davis and Resnick 1989)

$$X_t = \max\{a X_{t-1}, (1-a) Z_t\}, \quad t \in \mathbb{Z}, \tag{5}$$

where $a \in [0, 1]$ and $\{Z_t\}_{t \in \mathbb{Z}}$ is a unit Fréchet noise process. Equation (5) possesses a stationary solution with unit Fréchet margins that is regularly varying. From the fact that

$$\mathbb{P}(X_1 = a X_0, X_2 = a^2 X_0, \ldots, X_z = a^t X_t \mid X_0 > x)$$
$$= \mathbb{P}\left(X_0 \geq \max\left\{\tfrac{1-a}{a} Z_1, \tfrac{1-a}{a^2} Z_2, \ldots, \tfrac{1-a}{a^t} Z_t\right\} \,\middle|\, X_0 > x\right)$$
$$\geq \mathbb{P}\left(\max\left\{\tfrac{1-a}{a} Z_1, \tfrac{1-a}{a^2} Z_2, \ldots, \tfrac{1-a}{a^t} Z_t\right\} \leq x \,\middle|\, X_0 > x\right) \to 1,$$

for any $t > 0$ as $x \to \infty$, it can be seen that the tail process $\{Y_t\}_{t \in \mathbb{Z}}$ is given by

$$(Y_0, Y_1, \ldots, Y_t) =_d (Y, aY, \ldots, a^t Y)$$

where $Y$ is a standard Pareto random variable. Furthermore, we have

$$Y_{-1} = \begin{cases} a^{-1} Y & \text{with probability } a, \\ 0 & \text{with probability } 1 - a. \end{cases}$$

Thus, for $l \in \mathbb{N}$, we obtain

$$\lim_{u \to \infty} \mathbb{P}(C_u = l) = \frac{\mathbb{P}(Y_{-1} \leq 1, Y_0 > 1, \ldots, Y_{l-1} > 1, Y_l \leq 1)}{\mathbb{P}(Y_{-1} \leq 1, Y_0 > 1)}$$
$$= \frac{\mathbb{P}(Y_{-1} = 0, a^{1-l} < Y < a^{-l})}{\mathbb{P}(Y_{-1} = 0)} = a^{l-1}(1-a), \tag{6}$$

i.e. the limiting distribution is a geometric distribution with parameter $1 - a$. Even though, in principle, it should be possible to recover this limit distribution from results such as Proposition 1 in Markovich (2017), where formulae for $\mathbb{P}(T_2(u) = l + 1)$ are given, obtaining a closed-form expression seems to be much simpler using the tail process.

   Note that the limiting geometric distribution in Eq. (6) coincides with the limiting distribution for the size of clusters defined in the classical sense, cf. Perfekt (1994). This is due to the fact that, in the limit, exceedances over high thresholds always occur subsequently.

2. As second example, we consider a stationary moving maximum process (cf. Deheuvels 1983, for instance) defined by

$$X_t = \frac{1}{2} \max\{Z_t, Z_{t-2}\}, \quad t \in \mathbb{Z},$$

with $\{Z_t\}_{t\in\mathbb{Z}}$ being a unit Fréchet noise process. By definition, we have that $\mathbb{P}(X_t \leq u \mid X_0 > u) \to 1$ for all $t \in 2\mathbb{Z} + 1$. Consequently,

$$\lim_{u\to\infty} \mathbb{P}(C_u > 1) = \lim_{u\to\infty} \frac{\mathbb{P}(X_{-1} \leq u, X_0 > u, X_1 > u)}{\mathbb{P}(X_{-1} \leq u, X_0 > u)} = 0,$$

i.e. the distribution of $C_u$ converges weakly to a Dirac measure in 1.

In contrast, the limiting cluster size distribution according to the classical definition is obviously a Dirac measure in 2, that is, exceedances over high thresholds always occur in pairs. As the exceedances are separated by a non-exceedance, each pair is considered as two single clusters in our definition, while they belong to the same cluster according to the classical definition.

Similarly to their size, we can investigate ordinal patterns in $u$-exceedance clusters. Here, for fixed $l \in \mathbb{N}$, we are interested in the distribution of the $l$-ordinal pattern of a (randomly selected) $u$-exceedance cluster of size $l$:

$$\rho_{u,l}(\pi) = \mathbb{P}(\Pi((X_i)_{i=0}^{l-1}) = \pi \mid (X_i)_{i=0}^{l-1} \text{ is a } u\text{-exceedance cluster})$$
$$= \frac{\mathbb{P}(\Pi((X_i)_{i=0}^{l-1}) = \pi, X_{-1} \leq u, X_0 > u, \dots, X_{l-1} > u, X_l \leq u)}{\mathbb{P}(X_{-1} \leq u, X_0 > u, \dots, X_{l-1} > u, X_l \leq u)}$$

for each $\pi \in S_{l-1}$, i.e. $\rho_{u,l}$ defines a probability distribution on $S_{l-1}$. Again, this distribution converges as $u \to \infty$:

$$\lim_{u\to\infty} \rho_{u,l}(\pi) = \frac{\mathbb{P}(\Pi((Y_i)_{i=0}^{l-1}) = \pi, Y_{-1} \leq 1, Y_0 > 1, \dots, Y_{l-1} > 1, Y_l \leq 1)}{\mathbb{P}(Y_{-1} \leq 1, Y_0 > 1, \dots, Y_{l-1} > 1, Y_l \leq 1)}$$
$$= \frac{\mu_{\{-1,\dots,l\}}(\{(x, \mathbf{y}, z) \in [0, 1] \times (1, \infty)^l \times [0, 1] : \Pi(\mathbf{y}) = \pi\})}{\mu_{\{-1,\dots,l\}}([0, 1] \times (1, \infty)^l \times [0, 1])}. \quad (7)$$

## 4 Asymptotic results for empirical estimators

According to Eq. (4) and (7), both the limit distribution of clusters and the limit distribution of ordinal patterns within a cluster are given by a ratio of the type

$$\mu_{\{-1,\dots,t\}}(A)/\mu_{\{-1,\dots,t_0\}}(A_0)$$

for appropriate dimension $t, t_0 \in \mathbb{N}_0$ and sets $A \subset [0, \infty) \times (1, \infty) \times [0, \infty)^t$ and $A_0 \subset [0, \infty) \times (1, \infty) \times [0, \infty)^{t_0}$ with $\mu_{\{-1,\dots,t_0\}}(A_0) > 0$, i.e. a ratio of measures of two sets that are bounded from below by 1 in their second component.

More precisely, in case of the cluster size distribution in Eq. (4), we have $t = l$, $A = [0, 1] \times (1, \infty)^l \times [0, 1]$, $t_0 = 0$ and $A_0 = [0, 1] \times (1, \infty)$, while, in case of the distribution of the $l$-ordinal pattern in Eq. (7), we have $t = t_0 = l$, $A = \{(x, \mathbf{y}, z) \in [0, 1] \times (1, \infty)^l \times [0, 1] : \Pi(\mathbf{y}) = \pi\}$ and $A_0 = [0, 1] \times (1, \infty)^l \times [0, 1]$.

In the following, we will consider the general class of ratios of the above type, including both the limits in Eqs. (4) and (7) as special cases, and discuss their estimation from observations $X_{-1}, \dots, X_n$. It is worth noting that, making use of relation $\mu_{\{-1,\dots,t\}}(A) = \mu_{\{-1,\dots,t+1\}}(A \times [0, \infty))$, we may replace $t$ and $t_0$ by the maximum

of the two. Thus, in the following, without loss of generality, we will always assume that $t = t_0$.

Analogously to the calculations in Section 3, one can show that

$$\lim_{u \to \infty} \frac{\mathbb{P}((X_i)_{i=-1}^t \in uA)}{\mathbb{P}((X_i)_{i=-1}^t \in uA_0)} = \frac{\mu_{\{-1,\ldots,t\}}(A)}{\mu_{\{-1,\ldots,t\}}(A_0)}$$

provided that $\mu_{\{-1,\ldots,t\}}(\partial A) = \mu_{\{-1,\ldots,t\}}(\partial A_0) = 0$. This limit relation gives reason to set a high threshold $u$ and use a ratio estimator of the type

$$\widehat{R}_{n,u}(A, A_0) = \frac{\widehat{P}_{n,u}(A)}{\widehat{P}_{n,u}(A_0)} \tag{8}$$

where

$$\widehat{P}_{n,u}(A) = \frac{1}{n} \sum_{k=1}^{n-t} \mathbf{1}_{\{(X_i)_{i=k-1}^{k+t} \in uA\}} \tag{9}$$

is the empirical counterpart of the probability $\mathbb{P}((X_i)_{i=-1}^t \in uA)$.

We will show asymptotic properties of the ratio estimator $\widehat{R}_{n,u_n}(A, A_0)$ for some appropriate sequence of thresholds $(u_n)_{n \in \mathbb{N}}$ such that $u_n \to \infty$ and $n \, \mathbb{P}(X_0 > u_n) \to \infty$ as $n \to \infty$. To obtain consistency, we also need a mixing condition for the time series $(X_t)_{t \in \mathbb{Z}}$ expressed in terms of the $\alpha$-mixing coefficients

$$\alpha_h = \sup_{A, B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})} | \mathbb{P}((X_t)_{t \leq 0} \in A, \, (X_t)_{t \geq h} \in B)$$
$$- \mathbb{P}((X_t)_{t \leq 0} \in A) \, \mathbb{P}((X_t)_{t \geq h} \in B)|, \qquad h \in \mathbb{N}_0. \tag{10}$$

Here, we will assume that $(X_t)_{t \in \mathbb{Z}}$ is $\alpha$-mixing, i.e. $\alpha_h \to 0$ as $h \to \infty$, and that the coefficients $(\alpha_h)_{h=0}^{\infty}$ are summable. The proof is postponed to the Appendix.

**Proposition 1** *Let $(X_t)_{t \in \mathbb{Z}}$ be a regularly varying, strictly stationary time series with tail process $(Y_t)_{t \in \mathbb{Z}}$ whose finite-dimensional distributions are given by $(\mu_I)_{I \subset \mathbb{Z}}$. Assume that the corresponding $\alpha$-mixing coefficients satisfy*

$$\alpha_n \in \mathcal{O}(n^{-\delta}) \quad \text{for all } n \in \mathbb{N} \text{ and some } \delta > 1. \tag{11}$$

*Moreover, let $A \subset [0, \infty) \times (1, \infty) \times [0, \infty)^t$ and $A_0 \subset [0, \infty) \times (1, \infty) \times [0, \infty)^t$ be continuity sets w.r.t. $\mu_{\{-1,\ldots,t\}}$ and $\mu_{\{-1,\ldots,t\}}(A_0) > 0$. Then, for any sequence $(u_n)_{n \in \mathbb{N}}$ satisfying $u_n \to \infty$ and*

$$n^{\delta/(1+\delta)} \, \mathbb{P}(X_0 > u_n) \to \infty,$$

*we have*

$$\widehat{R}_{n,u_n}(A, A_0) \to_p \frac{\mu_{\{-1,\ldots,t\}}(A)}{\mu_{\{-1,\ldots,t\}}(A_0)}.$$

To further establish asymptotic normality, we will assume the following mixing condition which combines a condition on the $\alpha$-mixing coefficients and an anti-clustering condition.

*Condition* (M) There exist a sequence $\{u_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ of thresholds and an intermediate sequence $\{r_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$ with $\lim_{n \to \infty} u_n = \lim_{n \to \infty} r_n = \infty$, $\lim_{n \to \infty} n \, \mathbb{P}(X_0 >$

$u_n) = \infty$, $\lim_{n \to \infty} r_n \, \mathbb{P}(X_0 > u_n) = 0$ such that

$$\lim_{n \to \infty} \frac{1}{\mathbb{P}(X_0 > u_n)} \sum_{h=r_n}^{\infty} \alpha_h = 0 \tag{12}$$

and

$$\lim_{k \to \infty} \limsup_{n \to \infty} \sum_{h=k}^{r_n} \mathbb{P}(X_h > u_n \mid X_0 > u_n) = 0. \tag{13}$$

*Remark 1* Condition (M) is an adapted version of the condition in Davis and Mikosch (2009) who consider a ratio estimator of a similar type for the so-called extremogram. Similarly to the condition for consistency given in Eq. (11), also the mixing condition in Eq. (12) implies conditions on the decay of the sequence $\{\alpha_h\}_{h \in \mathbb{N}}$ which will be discussed below in more detail.

The anti-clustering condition in Eq. (13) is very similar to Condition (2.8) in Davis and Hsing (1995) and Condition 4.1 in Basrak and Segers (2009). By Proposition 4.2 in Basrak and Segers (2009), it implies that the tail process $\{Y_t\}_{t \in \mathbb{Z}}$ converges to 0 almost surely as $|t| \to \infty$ and thus ensures finite cluster size.

To prove the asymptotic normality of the ratio estimators, we first make use of the following auxiliary result in Davis and Mikosch (2009), adapted to our setting.

**Lemma 1** (Davis and Mikosch 2009, Thm. 3.1) *Let* $(X_t)_{t \in \mathbb{Z}}$ *be a regularly varying, strictly stationary time series with tail process* $(Y_t)_{t \in \mathbb{Z}}$ *whose finite-dimensional distributions are given by* $(\mu_I)_{I \subset \mathbb{Z}}$. *Moreover, let* $A \subset [0, \infty) \times (1, \infty) \times [0, \infty)^t$ *be a continuity set w.r.t.* $\mu_{\{-1,\dots,t\}}$. *If Condition (M) holds, then*

$$\lim_{n \to \infty} n \, \mathbb{P}(X_0 > u_n) \, \mathrm{Var}\left( \frac{\widehat{P}_{n,u_n}(A)}{\mathbb{P}(X_0 > u_n)} \right)$$

$$= \mu_{\{-1,\dots,t\}}(A) + 2 \sum_{h=1}^{\infty} \mathbb{P}((Y_i)_{i=-1}^{t} \in A, \, (Y_i)_{i=h-1}^{h+t} \in A) < \infty.$$

We further proceed by noting that Eq. (12) and $r_n \, \mathbb{P}(X > u_n) \to 0$ imply that

$$\lim_{n \to \infty} r_n \sum_{h=r_n}^{\infty} \alpha_h = 0.$$

Thus,

$$\liminf_{n \to \infty} n \sum_{h=n}^{\infty} \alpha_h = 0$$

and, consequently, $\liminf_{n \to \infty} n^2 \alpha_n = 0$. Imposing the existence of a finite limit superior, we may conclude that

$$\alpha_n \in \mathcal{O}(n^{-\delta}) \quad \text{for some } \delta \geq 2. \tag{14}$$

Using this slight strengthening of Condition (M), we can verify asymptotic normality of the estimators $\widehat{P}_{n,u_n} / \mathbb{P}(X_0 > u_n)$. The proof is postponed to the Appendix.

**Theorem 1** *Let* $(X_t)_{t \in \mathbb{Z}}$ *be a regularly varying, strictly stationary time series with tail process* $(Y_t)_{t \in \mathbb{Z}}$ *whose finite-dimensional distributions are denoted by* $(\mu_I)_{I \subset \mathbb{Z}}$.

*Moreover, let $A_0, \ldots, A_N \subset [0, \infty) \times (1, \infty) \times [0, \infty)^t$ be continuity sets w.r.t.* $\mu_{\{-1,\ldots,t\}}$. *We further assume that Condition (M) and* Eq. (14) *hold and*

$$\lim_{n \to \infty} n^{\delta/(4+\delta)} \, \mathbb{P}(X_0 > u_n) = \infty, \tag{15}$$

*if $\delta > 2$ in* Eq. (14) *or*

$$\lim_{n \to \infty} n^{1/2} \frac{\mathbb{P}(X_0 > u_n)^{3/2}}{|\log(\mathbb{P}(X_0 > u_n))|} = \infty \tag{16}$$

*if $\delta = 2$. Then,*

$$\sqrt{n \, \mathbb{P}(X_0 > u_n)} \left( \begin{array}{c} \frac{\widehat{P}_{n,u_n}(A_0)}{\mathbb{P}(X_0 > u_n)} - \mathbb{P}((X_i)_{i=-1}^t \in u_n A_0 \mid X_0 > u_n) \\ \vdots \\ \frac{\widehat{P}_{n,u_n}(A_N)}{\mathbb{P}(X_0 > u_n)} - \mathbb{P}((X_i)_{i=-1}^t \in u_n A_N \mid X_0 > u_n) \end{array} \right)$$
$$\longrightarrow_d \mathcal{N}(\mathbf{0}, \Sigma),$$

*as $n \to \infty$, where $\Sigma = (\sigma_{jl})_{0 \leq j,l \leq N}$ with*

$$\sigma_{j,l} = \mu_{\{-1,\ldots,t\}}(A_j \cap A_l) + \sum_{h=1}^{\infty} \mathbb{P}((Y_i)_{i=-1}^t \in A_j, \ (Y_i)_{i=h-1}^{h+t} \in A_l)$$
$$+ \sum_{h=1}^{\infty} \mathbb{P}((Y_i)_{i=-1}^t \in A_l, \ (Y_i)_{i=h-1}^{h+t} \in A_j).$$

*Remark 2* By regular variation,

$$\frac{\mathbb{E}\{\widehat{P}_{n,u_n}(A_j)\}}{\mathbb{P}(X_0 > u_n)} = \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j \mid X_0 > u_n)$$
$$\to \mathbb{P}((Y_i)_{i=-1}^t \in A_j) = \mu_{\{-1,\ldots,t\}}(A_j) \quad (n \to \infty),$$

i.e. the conditional distribution is approximately normal around the desired value even though the bias might be not negligible asymptotically. If the limit expression vanishes, i.e. $\mu_{\{-1,\ldots,t\}}(A_j) = 0$, we obtain the asymptotic variance $\sigma_{jj} = 0$, i.e. the limit distribution is degenerate.

The same arguments as in the proof of Corollary 3.3 in Davis and Mikosch (2009) allow us to go from asymptotic normality of the estimators $\widehat{P}_{n,u_n}(A_j)$, $j = 0, \ldots, N$, to asymptotic normality of the ratios $\widehat{R}_{n,u_n}(A_j, A_0)$, $j = 1, \ldots, N$, as stated in the following corollary.

**Corollary 1** *Under the assumptions of Theorem 1, we have that*

$$\sqrt{n \, \mathbb{P}(X_0 > u_n)} \left( \begin{array}{c} \widehat{R}_{n,u_n}(A_1, A_0) - \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_1)}{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_0)} \\ \vdots \\ \widehat{R}_{n,u_n}(A_N, A_0) - \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_N)}{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_0)} \end{array} \right)$$
$$\longrightarrow_d \mathcal{N}(\mathbf{0}, \mu_{\{-1,\ldots,t\}}(A_0)^{-4} F \Sigma F^\top),$$

*as $n \to \infty$, where*

$$F = \begin{pmatrix} -\mu_{\{-1,\ldots,t\}}(A_1) & \mu_{\{-1,\ldots,t\}}(A_0) & 0 & 0 & \ldots & 0 \\ -\mu_{\{-1,\ldots,t\}}(A_2) & 0 & \mu_{\{-1,\ldots,t\}}(A_0) & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\mu_{\{-1,\ldots,t\}}(A_N) & 0 & 0 & 0 & \ldots & \mu_{\{-1,\ldots,t\}}(A_0) \end{pmatrix}$$

*and $\Sigma$ is given in Theorem 1.*
*If, in addition,*

$$\lim_{n\to\infty} \sqrt{n\, \mathbb{P}(X_0 > u_n)} \left[ \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_i)}{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_0)} - \frac{\mu_{\{-1,\ldots,t\}}(A_i)}{\mu_{\{-1,\ldots,t\}}(A_0)} \right] = 0,$$

*for $i = 1, \ldots, N$, then*

$$\sqrt{n\, \mathbb{P}(X_0 > u_n)} \begin{pmatrix} \widehat{R}_{n,u_n}(A_1, A_0) - \frac{\mu_{\{-1,\ldots,t\}}(A_1)}{\mu_{\{-1,\ldots,t\}}(A_0)} \\ \vdots \\ \widehat{R}_{n,u_n}(A_N, A_0) - \frac{\mu_{\{-1,\ldots,t\}}(A_N)}{\mu_{\{-1,\ldots,t\}}(A_0)} \end{pmatrix}$$
$$\to \mathcal{N}(\mathbf{0}, \mu_{\{-1,\ldots,t\}}(A_0)^{-4} F \Sigma F^\top).$$

*Remark 3* Alternatively to our proofs, we could also show asymptotic normality of the vectors $(\widehat{P}_{n,u_n}(A_j))_{j=0}^N$ and $(\widehat{R}_{n,u_n}(A_j))_{j=1}^N$, respectively, using slightly adapted versions of Theorem 3.2 and Corollary 3.3 in Davis and Mikosch (2009). Therein, besides Condition (M), they also assume the conditions

$$\lim_{n\to\infty} n\, \mathbb{P}(X_0 > u_n) \cdot \alpha_{r_n} = 0 \tag{17}$$

and

$$\lim_{n\to\infty} n^{1/3}\, \mathbb{P}(X_0 > u_n) = \infty. \tag{18}$$

By using different techniques in the proof and extending Condition (M) by the slightly stronger assumption (14), we are able to drop condition (17). Furthermore, we replace condition (18) by conditions (15) and (16), respectively.

For $\delta > 2$, condition (15) is weaker than condition (18), which is the limiting case of condition (15) as $\delta \searrow 2$. If $\alpha_h$ even decays exponentially, i.e. if condition (15) holds for $\delta$ being arbitrarily large, the condition simplifies to $\lim_{n\to\infty} n^{1-\varepsilon}\, \mathbb{P}(X_0 > u_n) = \infty$ for some $\varepsilon$ which is close to the minimal assumption $\lim_{n\to\infty} n\, \mathbb{P}(X_0 > u_n) = \infty$ stated in Condition (M).

For $\delta = 2$, condition (16) is slightly stronger than condition (18). However, it is still weaker than

$$\lim_{n\to\infty} n^{1/3-\varepsilon}\, \mathbb{P}(X_0 > u_n) = \infty$$

for $\varepsilon > 0$.

Thus, even though our assumptions are not necessarily weaker than the assumptions in Davis and Mikosch (2009) due to the fact that we further assume (14), our results allow for a simplification of Eqs. (17) and (18).

In practical applications, the usability of central limit theorems in the flavor of Corollary 1 for uncertainty assessment of the resulting estimates is often limited by two obstacles:

- The rate of convergence includes the unknown threshold exceedance probability $\mathbb{P}(X_0 > u_n)$.
- The asymptotic (co-)variances are complex expressions including series expressions as given in Theorem 1.

In the case of Corollary 1, however, one can cope with both difficulties in the following way:

- Applying Lemma 1 to the set $A = [0, \infty) \times (1, \infty)$, we obtain that, under Condition (M),

$$\frac{\widehat{P}_{n,u_n}([0, \infty) \times (1, \infty))}{\mathbb{P}(X_0 > u_n)} = \frac{1}{n\,\mathbb{P}(X_0 > u_n)} \sum_{k=1}^{n} \mathbf{1}_{\{X_k > u_n\}} \to_p 1.$$

  Therefore, Corollary 1 still holds true if we replace the exceedance probability $\mathbb{P}(X_0 > u_n)$ by its empirical counterpart $n^{-1} \sum_{k=1}^{n} \mathbf{1}_{\{X_k > u_n\}}$.
- Similarly to the asymptotic (co-)variances for the extremogram estimators, the asymptotic (co-)variances arising in Corollary 1 can be estimated via various bootstrap techniques such as a stationary bootstrap (Davis et al. 2012) or a multiplier block bootstrap (Drees 2015). In Section 5, we will make use of the multiplier block bootstrap which has been demonstrated to provide more accurate and robust results than the stationary bootstrap in a simulation study in Davis et al. (2018).

## 5 Example: Max-stable time series

An important example of a stationary regularly varying times series $(X_t)_{t \in \mathbb{Z}}$ is a stationary max-stable time series with $\alpha$-Fréchet margins. According to de Haan (1984), such a time series can be represented as

$$X_t = \bigvee_{j=1}^{\infty} \left( \Gamma_j^{-1} W_t^{(j)} \right)^{1/\alpha}, \qquad t \in \mathbb{Z}, \tag{19}$$

where $\{\Gamma_j\}_{j \in \mathbb{N}}$ denote the arrival times of a unit rate Poisson process and $(W_t^{(j)})_{t \in \mathbb{Z}}$, $j \in \mathbb{N}$, are independent copies of a nonnegative time series $(W_t)_{t \in \mathbb{Z}}$ such that $\mathbb{E}\, W_t = 1$ for all $t \in \mathbb{Z}$. Then, $(X_t)_{t \in \mathbb{Z}}$ is regularly varying with index $\alpha$ and its tail process $(Y_t)_{t \in \mathbb{Z}}$ is of the form

$$Y_t = P \cdot \widetilde{W}_t^{\,1/\alpha}, \qquad t \in \mathbb{Z}, \tag{20}$$

where $P$ is an $\alpha$-Pareto random variable, i.e. $\mathbb{P}(P > x) = x^{-\alpha}$, $x > 1$, and $(\widetilde{W}_t)_{t \in \mathbb{Z}}$ is an independent time series with

$$\mathbb{P}\left(\widetilde{W} \in A\right) = \int_{[0,\infty)^{\mathbb{Z}}} w(0)\mathbf{1}_{w/w(0) \in A}\, \mathbb{P}(W \in \mathrm{d}w), \quad A \subset [0, \infty)^{\mathbb{Z}},$$

see Dombry and Ribatet (2015) for more details.

The dependence structure of a max-stable time series is often summarized by its extremal coefficient function, that is, a sequence $\{\theta(h)\}_{h\in\mathbb{Z}} \subset [1, 2]$ given by the relation

$$\mathbb{P}(X_h \leq x, X_0 \leq x) = \mathbb{P}(X_0 \leq x)^{\theta(h)}, \qquad x > 0, \ h \in \mathbb{Z}.$$

In particular, $X_h = X_0$ a.s. iff $\theta(h) = 1$ and $X_h$ and $X_0$ are (asymptotically) independent iff $\theta(h) = 2$.

By results of Dombry and Eyi-Minko (2012), an upper bound for the mixing coefficients $\alpha_h$ in Eq. (10) can be formulated in terms of the extremal coefficient function, see also Davis et al. (2013):

$$\alpha_h \leq 2 \sum_{s_1=-\infty}^{0} \sum_{s_2=0}^{\infty} (2 - \theta(s_2 - s_1 + h)) = 2 \sum_{s=0}^{\infty} (s + 1) (2 - \theta(s + h)). \tag{21}$$

This upper bound can be used to verify the validity of Eq. (14). Once $\delta \geq 2$ from Eq. (14) and the marginal distribution are known, it is straightforward to choose the sequence $\{u_n\}_{n\in\mathbb{N}}$ such that Eqs. (15) and (16), respectively, are satisfied. Furthermore, Eq. (21) can be used to obtain sufficient conditions for condition (M) given in the following proposition. The proof is postponed to the Appendix.

**Proposition 2** *Let $(X_t)_{t\in\mathbb{Z}}$ be a strictly stationary max-stable time series with $\alpha$-Fréchet margins and extremal coefficient function $\{\theta(h)\}_{h\in\mathbb{Z}}$ such that $h \mapsto \theta(h)$ is non-increasing on $\mathbb{N}_0$. If there exist a sequence $\{u_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ of thresholds and an intermediate sequence $\{r_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ with $\lim_{n\to\infty} u_n = \lim_{n\to\infty} r_n = \infty$, $\lim_{n\to\infty} n \mathbb{P}(X_0 > u_n) = \infty$, $\lim_{n\to\infty} r_n \mathbb{P}(X_0 > u_n) = 0$ such that*

$$\lim_{n\to\infty} u_n^\alpha \sum_{h=1}^{\infty} h^2 (2 - \theta(h + r_n)) = 0 \tag{22}$$

*and*

$$\lim_{k\to\infty} \sum_{h=k}^{r_n} (2 - \theta(h)) = 0, \tag{23}$$

*then Condition (M) holds.*

The additional assumption in the second part of Corollary 1 that ensures that the bias of the estimator is asymptotically negligible cannot be verified in this general setting. However, for the closely related extremogram (Davis and Mikosch 2009), we have, from Eq. (26) in the Appendix, that

$$|\mathbb{P}(X_h > u_n \mid X_0 > u_n) - (2 - \theta(h))| \in \mathcal{O}(u_n^{-\alpha})$$

(see also Buhl and Küppelberg 2018 Lemma A.1). Thus

$$\sqrt{n \mathbb{P}(X(0) > u_n)} \, |\mathbb{P}(X_h > u_n \mid X_0 > u_n) - (2 - \theta(h))| \to 0$$

holds if $nu_n^{-3\alpha} \to 0$. A similar behavior might be expected for the conditional probabilities in Corollary 1 which are of the same type.

In the following simulation study, we will focus on one of the most popular models for max-stable processes, namely the Brown–Resnick process, that is, a stationary max-stable time series with unit Fréchet margins and processes $(W_t)_{t\in\mathbb{Z}}$ and $(\widetilde{W}_t)_{t\in\mathbb{Z}}$ in Eqs. (19) and (20), respectively, of the form

$$W_t = \widetilde{W}_t = \exp(G_t - \text{Var}(G_t)/2), \quad t \in \mathbb{Z},$$

for a centered Gaussian time series $(G_t)_{t \in \mathbb{Z}}$ with $G_0 = 0$ a.s. and stationary increments. Thus, $(X_t)_{t \in \mathbb{Z}}$ is a stationary max-stable process and its law is uniquely determined by the semi-variogram

$$\gamma(h) = \tfrac{1}{2} \operatorname{Var}(G_h - G_0), \quad h \in \mathbb{Z},$$

see Kabluchko et al. (2009). In applications, often Brown–Resnick processes associated to a semi-variogram of power type

$$\gamma(h) = C|h|^{\beta}, \quad h \in \mathbb{Z},$$

for some $C > 0$ and $\beta \in (0, 2]$ are considered.

Similarly to Cho et al. (2016), Buhl and Klüppelberg (2018), Buhl et al. (2019), and Buhl and Klüppelberg (2019) for spatial and spatio-temporal Brown-Resnick processes, we can verify that such a Brown–Resnick process satisfies the assumptions of our central limit theorems (Theorem 1 and Corollary 1, respectively). As our assumptions are different from the ones in the papers mentioned above, we will verify them independently making use of our results for general max-stable processes in Proposition 2. The proof is postponed to the Appendix.

**Corollary 2** *Let $(X_t)_{t \in \mathbb{Z}}$ a Brown–Resnick time series associated to a semi-variogram $\gamma$ satisfying $\gamma(h) \geq C|h|^{\varepsilon}$ for all $h \in \mathbb{Z}$ and some $C, \varepsilon > 0$. Then, choosing $u_n \sim n^{\beta_1}$ for some $\beta_1 \in (0, 1)$ and $r_n \sim n^{\beta_2}$ for some $\beta_2 \in (0, \beta_1)$, the assumptions of Theorem 1 and, thus, of the first part of Corollary 1, are satisfied.*

We now simulate a Brown–Resnick process to demonstrate the performance of the estimators of the type $\widehat{R}_{n,u}(\cdot, \cdot)$ for the distribution of the cluster size and the ordinal patterns within a cluster. More precisely, we use the extremal functions approach (Dombry et al. 2016) to simulate a Brown-Resnick time series of length $1\,000\,000$ with unit Fréchet margins and associated to the semi-variogram $\gamma(h) = 0.1 \cdot |h|^{1.75}$. We then estimate

- the distribution of the cluster size,
- the distribution of ordinal patterns within clusters of size 2
- and the distribution of ordinal patterns within clusters of size 3

based on exceedances over the 95%-quantile of the unit Fréchet distribution using the ratio estimators according to Eq. (8). The results are displayed and compared to the exact limit distributions with probability masses of the form $\mu_{\{-1,\ldots,t\}}(A)/\mu_{\{-1,\ldots,t\}}(A_0)$ calculated via simulations from the tail process, in Fig. 2 and 3, respectively. The uncertainty of the estimators is assessed via the multiplier block bootstrap (Drees 2015; Davis et al. 2018) based on fixed blocks of size $1\,000$. The 95 % confidence intervals obtained from the bootstrap are compared to the theoretical confidence intervals according to the asymptotic distribution given in Corollary 1, i.e. intervals of the form

$$\left[ \; \widehat{R}_{n,u_n}(A, A_0) - \Phi^{-1}(0.975) \cdot \frac{1}{\sqrt{n \, \mathbb{P}(X_0 > u_n)}} \sqrt{\operatorname{AVar}(\widehat{R}_{n,u_n}(A, A_0))}, \right.$$
$$\left. \widehat{R}_{n,u_n}(A, A_0) + \Phi^{-1}(0.975) \cdot \frac{1}{\sqrt{n \, \mathbb{P}(X_0 > u_n)}} \sqrt{\operatorname{AVar}(\widehat{R}_{n,u_n}(A, A_0))} \right],$$
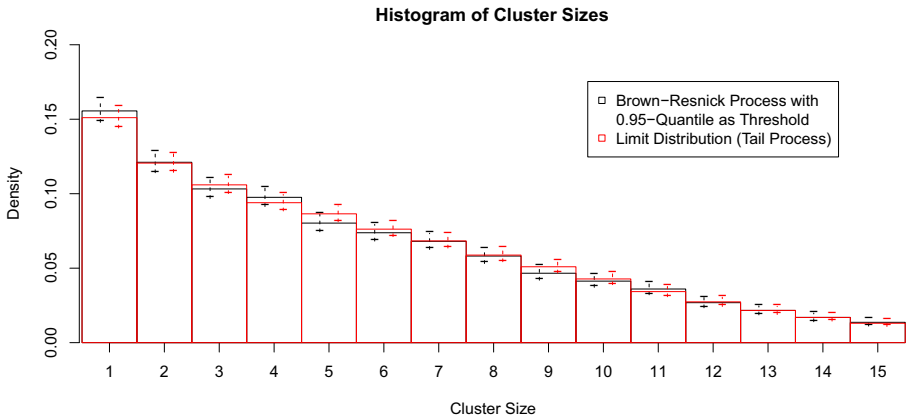
**Fig. 2** Histogram of the cluster size distribution for estimates from a simulated Brown–Resnick process (black) and the corresponding limit distribution (red). 95 % confidence intervals are added, obtained via a multiplier block bootstrap (black) and from the theoretical asymptotic distribution (red), respectively.

where $\mathrm{AVar}(\widehat{R}_{n,u_n}(A, A_0))$ denotes the estimated asymptotic variance of the ratio estimator given in Corollary 1 which is again calculated via simulations from the tail process. It can be seen that all the probabilities are estimated rather accurately and that the estimated uncertainty is close to the theoretical one, i.e. both types of confidence intervals have similar sizes.
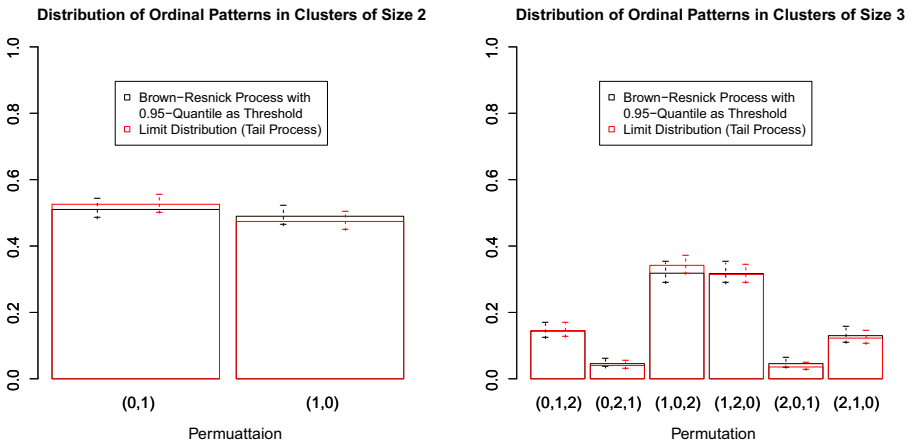


**Fig. 3** Bar charts of the distributions of 2- (left) and 3-ordinal patterns (right) within exceedance clusters estimated from a simulated Brown–Resnick process (black) and the corresponding limit distributions (red). The corresponding 95 % confidence intervals are obtained via a multiplier block bootstrap (black) and from the theoretical asymptotic distribution (red), respectively

## 6 Application: River discharge at cologne

As an application we consider a time series of daily discharge data of the river Rhine measured at Cologne. In many cases, river discharge data exhibit temporal clustering of extremes, which entails the use of declustering techniques for the statistical analysis of their tail behavior (cf. Kallache et al. 2011, Asadi et al. 2015, for instance). Here, we study the structure of these clusters making use of the estimators introduced above. We restrict ourselves to the analysis of floods in the extended winter season (DJFM), assuming stationarity of the time series within each winter period consisting of 121 days (and 122 days, respectively, in leak years).

The given data set, provided by The Global Runoff Data Centre, 56068 Koblenz, Germany, consists of data from 197 winter seasons from December 1816 to March 2013. In an exploratory analysis, we calculate the empirical version of the extremogram

$$\rho_{(1,\infty),(1,\infty)}(h) = \lim_{u \to \infty} \mathbb{P}(X_h > u \mid X_0 > u), \quad h \in \mathbb{Z},$$

based on exceedances the empirical 95 %-quantile according to Davis and Mikosch (2009). The result, displayed in Fig. 4, shows a decrease of extremal dependence as the temporal lag increases being close to asymptotic independence for lags larger than 40 days. Further analyses indicate that runoff data from different seasons may be assumed to be independent. These observations offer the applicability of the ratio estimator and the results on its asymptotic behavior from Section 4.

We choose two different thresholds for the empirical verification of the stability of different exceedance cluster characteristics. More precisely, we consider the empirical 95 %- and 97.5 %-quantiles as thresholds leading to 200 and 114 clusters, respectively. As the empirical distributions of cluster sizes are rather difficult to compare due to the large number of potential outcomes relatively to the small number of clusters, we focus on the distribution of the 2- and 3-ordinal patterns. The results are
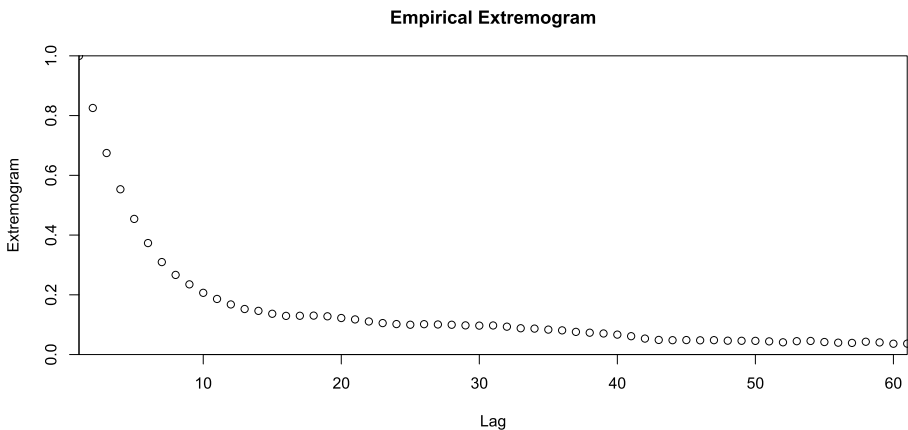


**Empirical Extremogram**

**Fig. 4** Empirical extremogram of the daily river discharge at Cologne based on the empirical 95 %–quantile
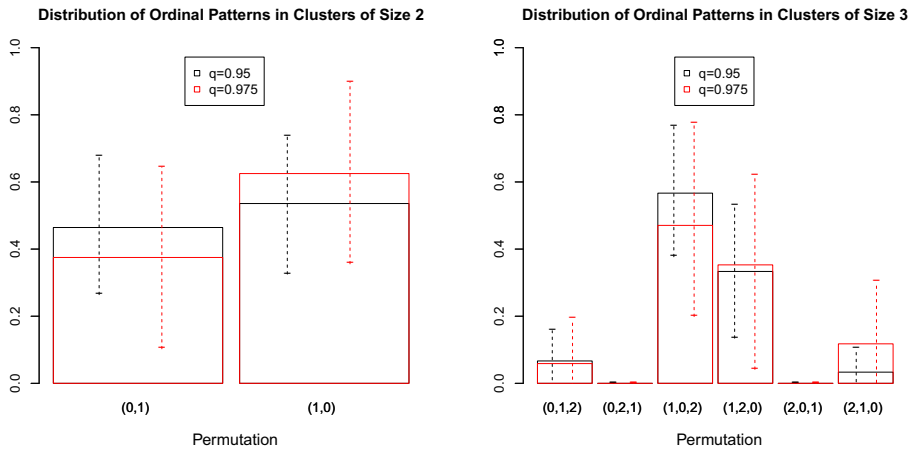
**Fig. 5** Bar charts of the distributions of ordinal patterns for the first 2 (left) and 3 (right) values of exceedance clusters of the daily river discharge at Cologne for two different thresholds (the empirical 95 %- and 97.5 %-quantiles)

displayed in Fig. 5 supplemented by 95 % confidence intervals obtained via a multiplier block bootstrap using each season as a fixed block. Even though the number of clusters is quite small, some interesting observations can be made: While both potential patterns of length two occur with almost the same frequency (in particular in case of the 95 %-quantile), for clusters of size 3, the patterns for which the second observation is the largest, i.e. $(1, 0, 2)$ and $(1, 2, 0)$, are clearly predominant. This means that extreme events that occur three time instants tend to show an "up-down" pattern. In contrast, patterns with a "down-up" behavior, i.e $(0, 2, 1)$ and $(2, 0, 1)$ do not occur at all.

In order to obtain more stable results based on a larger number of clusters, one could focus on patterns at the beginning of potentially longer clusters, i.e. one could consider the pattern for the first two instants within all clusters that are at least of size 2, for instance. Maybe one can even use these ordinal patterns at the beginning of clusters to predict the length of the clusters. Such an analysis, however, is beyond the scope of the present article.

## Appendix: Proofs

*Proof of Proposition 1* We first note that, by Eq. (11), there exists some $C > 0$ such that $\alpha_h \leq Cn^{-\delta}$ for all $n \in \mathbb{N}$. By a straightforward computation, for any intermediate sequence $(r_n)_{n\in\mathbb{N}}$ with $r_n \to \infty$ and $r_n/n \to 0$, we then obtain

$$
\text{Var}\left(\frac{\widehat{P}_{n,u_n}(A)}{\mathbb{P}(X_0 > u_n)}\right)
$$

$$
= \frac{1}{n^2\mathbb{P}(X_0 > u_n)^2}\sum_{k=1}^{n-t}\sum_{l=1}^{n-t}\text{Cov}\left(\mathbf{1}_{\{(X_i)_{i=k-1}^{k+t}\in u_n A\}}, \mathbf{1}_{\{(X_i)_{i=l-1}^{l+t}\in u_n A\}}\right)
$$

$$
\leq \frac{2}{n\mathbb{P}(X_0 > u_n)^2}\sum_{h=0}^{n-t-1}\left|\mathbb{P}((X_i)_{i=-1}^{t}\in u_n A, (X_i)_{i=h-1}^{h+t}\in u_n A)\right.
$$

$$
\left. -\mathbb{P}((X_i)_{i=-1}^{t}\in u_n A)\mathbb{P}((X_i)_{i=h-1}^{t}\in u_n A)\right|
$$

$$
\leq \frac{2}{n\mathbb{P}(X_0 > u_n)}\sum_{h=0}^{r_n}\left|\mathbb{P}((X_i)_{i=-1}^{t}\in u_n A, (X_i)_{i=h-1}^{h+t}\in u_n A \mid X_0 > u_n)\right.
$$

$$
\left. -\mathbb{P}((X_i)_{i=-1}^{t}\in u_n A \mid X_0 > u_n)\mathbb{P}((X_i)_{i=h-1}^{t}\in u_n A)\right|
$$

$$
+\frac{2}{n\mathbb{P}(X_0 > u_n)^2}\sum_{h=r_n+1}^{\infty}\alpha_{h-t-2}
$$

$$
\leq \frac{2(r_n+1)}{n\mathbb{P}(X_0 > u_n)} + \frac{2C}{n\mathbb{P}(X_0 > u_n)^2}\sum_{h=r_n-t-1}^{\infty}h^{-\delta}
$$

$$
\sim \frac{2(r_n+1)}{n\mathbb{P}(X_0 > u_n)} + \frac{2C(r_n-t-1)^{1-\delta}}{(\delta-1)n\mathbb{P}(X_0 > u_n)^2}.
$$

Setting $r_n = \mathbb{P}(X_0 > u_n)^{-1/\delta}$, the right-hand side is asymptotically equal to

$$
\frac{2(1+C/(\delta-1))+o(1)}{n\mathbb{P}(X_0 > u_n)^{1+1/\delta}} = \frac{2(1+C/(\delta-1))+o(1)}{[n^{\delta/(1+\delta)}\mathbb{P}(X_0 > u_n)]^{1+1/\delta}} \to 0.
$$

Thus, by Chebychev's inequality, this implies that

$$
\frac{\widehat{P}_{u_n,n}(A) - \mathbb{E}(\widehat{P}_{u_n,n}(A))}{\mathbb{P}(X_0 > u_n)} \to_p 0.
$$

Since, by regular variation,

$$
\frac{\mathbb{E}(\widehat{P}_{u_n,n}(A))}{\mathbb{P}(X_0 > u_n)} = \mathbb{P}((X_i)_{i=-1}^{t}\in u_n A \mid X_0 > u_n) \to \mu_{\{-1,...,t\}}(A), \qquad n \to \infty,
$$

we obtain that $\widehat{P}_{u_n,n}(A)/\mathbb{P}(X_0 > u_n) \to_p \mu_{\{-1,\dots,t\}}(A)$. The same result can be obtained if we replace $A$ by $A_0$. An application of the continuous mapping theorem for convergence in probability completes the proof.                                        $\square$

*Proof of Theorem 1* We prove the equivalent statement that all linear combinations of the form

$$\sum_{j=0}^{N} a_j \sqrt{n\mathbb{P}(X_0 > u_n)} \left( \frac{\widehat{P}_{n,u_n}(A_j)}{\mathbb{P}(X_0 > u_u)} - \mathbb{P}((X_i)_{i=-1}^{t} \in u_n A_j \mid X_0 > u_n) \right),$$

$a_0, \dots, a_N \in \mathbb{R}$, converge in distribution to a centered normal distribution with the corresponding variance $\sum_{j=0}^{N} \sum_{l=0}^{N} a_j a_l \sigma_{jl}$. To this end, define

$$Z_{n,k} = \frac{1}{\sqrt{n\mathbb{P}(X_0 > u_n)}} \sum_{j=0}^{N} a_j \left( \mathbf{1}_{\{(X_i)_{i=k-1}^{k+t} \in u_n A_j\}} - \mathbb{P}((X_i)_{i=-1}^{t} \in u_n A_j) \right).$$

We note that, for each $n \in \mathbb{N}$, the random variable $\sum_{k=1}^{l} Z_{n,k}$ is centered and that its variance converges

$$\begin{aligned}
&\mathrm{Var} \left( \sum_{k=1}^{n} Z_{n,k} \right) \\
&= \sum_{j=0}^{N} \sum_{l=0}^{N} a_j a_l \sum_{h=-(n-1)}^{n-1} \frac{n - |h|}{n} \cdot \frac{\mathrm{Cov} \left( \mathbf{1}_{\{(X_i)_{i=-1}^{t} \in u_n A_j\}}, \mathbf{1}_{\{(X_i)_{i=h-1}^{h+t} \in u_n A_l\}} \right)}{\mathbb{P}(X_0 > u_n)} \\
&\to \sum_{j=0}^{N} \sum_{l=0}^{N} a_j a_l \sigma_{jl}
\end{aligned}$$

to the desired quantity, which can be verified analogously to the proof of Lemma 1, i.e. the proof of Thm. 3.1 in Davis and Mikosch (2009).

It remains to show that the asymptotic distribution is normal. To this end, we verify that the triangular scheme $\{Z_{n,k}\}_{k=1,\dots,n}, n \in \mathbb{N}$, satisfies the conditions of Thm. 4.4 in Rio (2017):

– At first, all the variables $Z_{n,k}$ are required to be centered and have finite variance which holds true as they are bounded.
– Secondly, we need to verify that

$$\limsup_{n \to \infty} \max_{l=1,\dots,n} \mathrm{Var} \left( \sum_{k=1}^{l} Z_{n,k} \right) < \infty, \tag{24}$$

   which again can be shown analogously to the proof of Lemma 1.
– The third and last condition to be verified is

$$\lim_{x \to \infty} n \int_{0}^{1} \alpha^{-1}(x) Q_{n,a}^2(x) \min\{\alpha^{-1}(x) Q_{n,a}(x), 1\} \, dx = 0 \tag{25}$$

   where $\alpha^{-1}$ and $Q_{n,a}$ denote the inverse functions of $h \mapsto \alpha_h$ and

$$u \mapsto \mathbb{P}(|Z_{n,k}| > u),$$

respectively. Noting that $Q_{n,a}$ can be bounded via the relation

$$Q_{n,a}(x) \leq \sum_{j=1}^{N} |a_j| Q_{n,e_j}\left(\frac{x}{N}\right), \quad x \in [0, 1],$$

where $e_j$ is the $j$th standard basis vector in $\mathbb{R}^N$, $j = 1, \ldots, N$, it suffices to verify

$$\lim_{x \to \infty} n \int_0^1 \alpha^{-1}(x) Q_{n,e_j}^2\left(\frac{x}{N}\right) \min\left\{\alpha^{-1}(x) Q_{n,e_j}\left(\frac{x}{N}\right), 1\right\} dx = 0$$

for $j = 1, \ldots, N$. By Eq. (14), we have $\alpha^{-1}(x) \leq (x/C)^{-1/\delta}$ for some $C > 0$ and

$$Q_{n,e_j}\left(\frac{x}{N}\right) = \begin{cases} \frac{1 - \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\sqrt{n\mathbb{P}(X_0 > u_n)}}, & 0 \leq x < N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j), \\ \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\sqrt{n\mathbb{P}(X_0 > u_n)}}, & N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j) \leq x \leq 1, \end{cases}$$

for sufficiently large $n$. Thus, we have

$$n \int_0^1 \alpha^{-1}(x) Q_{n,e_j}^2\left(\frac{x}{N}\right) \min\left\{\alpha^{-1}(x) Q_{n,e_j}\left(\frac{x}{N}\right), 1\right\} dx$$

$$\leq n \int_0^{N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)} (x/C)^{-1/\delta} (n\mathbb{P}(X_0 > u_n))^{-1}$$

$$\min\{1, (x/C)^{-1/\delta}(n\mathbb{P}(X_0 > u_n)^{-1/2}\} dx$$

$$+ n \int_{N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}^1 (x/C)^{-2/\delta} \left(\frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\sqrt{n\mathbb{P}(X_0 > u_n)}}\right)^3 dx$$

$$=: I_1 + I_2.$$

For the assessment of the integral term $I_2$, we employ the upper bound

$$\left(\frac{x}{C}\right)^{-2/\delta} \leq (C/N)^{2/\delta} \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)^{-2/\delta}$$

and obtain

$$I_2 \leq n \int_0^1 \left(\frac{C}{N}\right)^{\frac{2}{\delta}} \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)^{-\frac{2}{\delta}} \left(\frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\sqrt{n\mathbb{P}(X_0 > u_n)}}\right)^3 dx$$

$$\leq \left(\frac{C}{N}\right)^{\frac{2}{\delta}} \frac{1}{\sqrt{n}} \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)^{\frac{3\delta - 4}{2\delta}} \left(\frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\mathbb{P}(X_0 > u_n)}\right)^{3/2} \longrightarrow 0$$

using that $\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)/\mathbb{P}(X_0 > u_n) \to \mu_{\{-1,\ldots,t\}}(A_j) < \infty$ and $\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j) \to 0$ as $n \to \infty$.

For the assessment of the integral $I_1$, we distinguish between the two cases $\delta = 2$ and $\delta > 2$.

In the case $\delta = 2$, we have

$$
\begin{aligned}
I_1 &\leq n \int_0^{N[n\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)]^{-1}} C^{1/2} x^{-1/2} (n\mathbb{P}(X_0 > u_n))^{-1} \, dx \\
&\quad + n \int_{N[n\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)]^{-1}}^{N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)} Cx^{-1} \sqrt{n\mathbb{P}(X_0 > u_n)}^{-3} \, dx \\
&= \sqrt{2CN} [n\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)]^{-1/2} \mathbb{P}(X_0 > u_n)^{-1} \\
&\quad + Cn^{-1/2}\mathbb{P}(X_0 > u_n)^{-3/2} \log(\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)) \\
&\quad + Cn^{-1/2}\mathbb{P}(X_0 > u_n)^{-3/2} \log(n\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)) \\
&= \sqrt{2CN} n^{-1/2}\mathbb{P}(X_0 > u_n)^{-3/2} \cdot \sqrt{\frac{\mathbb{P}(X_0 > u_n)}{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}} \\
&\quad + Cn^{-1/2}\mathbb{P}(X_0 > u_n)^{-3/2} \cdot \Big\{ -\log(\mathbb{P}(X_0 > u_n)) + \log(n\mathbb{P}(X_0 > u_n)^3) \\
&\quad + 2\log\left( \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\mathbb{P}(X_0 > u_n)} \right) \Big\}
\end{aligned}
$$

which vanishes as $n \to \infty$ because of Eq. (16) and $\lim_{x\to\infty} x^{-1/2} \log(x) = 0$. For $\delta > 2$, we obtain

$$
\begin{aligned}
I_1 &\leq n \int_0^{N\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)} C^{2/\delta} x^{-2/\delta} \sqrt{n\mathbb{P}(X_0 > u_n)}^{-3} \, dx \\
&= \frac{\delta}{\delta - 2} N^{1-\frac{2}{\delta}} C^{\frac{2}{\delta}} \frac{1}{\sqrt{n}} \mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)^{1-\frac{2}{\delta}} \mathbb{P}(X_0 > u_n)^{-3/2} \\
&= \frac{\delta}{\delta - 2} N^{1-\frac{2}{\delta}} C^{\frac{2}{\delta}} \frac{1}{\sqrt{n}} \mathbb{P}(X_0 > u_n)^{-\frac{1}{2}\left(\frac{4+\delta}{\delta}\right)} \left( \frac{\mathbb{P}((X_i)_{i=-1}^t \in u_n A_j)}{\mathbb{P}(X_0 > u_n)} \right)^{1-\frac{2}{\delta}}
\end{aligned}
$$

which vanishes as $n \to \infty$ because of Eq. (15). $\qquad\square$

*Proof of Proposition 2* We verify the validity of both Eqs. (12) and (13). By Eq. (21), the series considered in Eq. (12) can the bounded by

$$
\begin{aligned}
\frac{1}{\mathbb{P}(X_0 > u_n)} \sum_{h=r_n}^{\infty} \alpha_h &\leq \frac{2}{\mathbb{P}(X_0 > u_n)} \sum_{h=r_n}^{\infty} \sum_{s=0}^{\infty} (s+1)(2 - \theta(s+h)) \\
&= \frac{2}{\mathbb{P}(X_0 > u_n)} \sum_{t=0}^{\infty} \sum_{h'=0}^{t} (t-h'+1)(2-\theta(t+r_n)) \\
&= \frac{1}{\mathbb{P}(X_0 > u_n)} \sum_{t=0}^{\infty} (t+1)(t+2)(2-\theta(t+r_n)),
\end{aligned}
$$

where we set $h' = h - r_n$ and $t = s + h'$.

The asymptotic relation $\mathbb{P}(X_0 > u_n) = 1 - \exp(-u_n^{-\alpha}) \sim u_n^{-\alpha}$ then yields that Eq. (22) is a sufficient condition for Eq. (12).

In order to simplify condition (13), we first note that

$$\mathbb{P}(X_h > u_n \mid X_0 > u_n) = \frac{2(1 - \exp(-u_n^{-\alpha})) - (1 - \exp(-\theta(h)u_n^{-\alpha}))}{1 - \exp(-u_n^{-\alpha})}.$$

Using the inequality $x - x^2/2 \leq 1 - \exp(-x) \leq x$ for all $x \geq 0$, we obtain the bounds

$$2 - \theta(h) - \frac{u_n^{-\alpha}}{2} \leq \mathbb{P}(X_h > u_n \mid X_0 > u_n)$$

$$\leq \frac{2 - \theta(h) + \frac{\theta(h)^2 u_n^{-\alpha}}{2}}{1 - \frac{u_n^{-\alpha}}{2}} \sim 2 - \theta(h) + \frac{\theta(h)^2}{2} u_n^{-\alpha} \leq 2 - \theta(h) + 2u_n^{-\alpha}. \qquad (26)$$

As the sequences $r_n, u_n \to \infty$ are chosen such that $r_n \mathbb{P}(X_0 > u_n) \to 0$ as $n \to \infty$, we obtain that, for every $k \in \mathbb{N}$,

$$\sum_{h=k}^{r_n} u_n^{-\alpha} \leq \frac{r_n}{u_n^{\alpha}} = \frac{r_n \mathbb{P}(X_0 > u_n)}{u_n^{\alpha} \mathbb{P}(X_0 > u_n)} \to 0 \quad (n \to \infty),$$

and consequently, from the inequalities above, Eq. (13) holds if and only if Eq. (23) holds. $\qquad \square$

*Proof of Corollary 2* We first note that, for the Brown–Resnick model,

$$\theta(h) = 2\Phi\left(\sqrt{\gamma(h)/2}\right), \qquad h \in \mathbb{Z},$$

cf. Kabluchko et al. (2009). In particular, using $1 - \Phi(x) \sim x^{-1}\varphi(x)$ as $x \to \infty$, we obtain for large $h$ that

$$2 - \theta(h) \sim \frac{2}{\sqrt{\pi \gamma(h)}} \exp\left(-\frac{\gamma(h)}{4}\right).$$

Thus, similarly to Davis et al. (2013), we can employ (21) to bound the $\alpha$-mixing coefficients by

$$\alpha_h \leq \frac{4}{\sqrt{\pi C}} \sum_{s=0}^{\infty} \frac{s+1}{(s+h)^{\varepsilon/2}} \exp\left(-\frac{1}{4}C \cdot (s+h)^{\varepsilon}\right)$$

$$\leq \frac{4}{\sqrt{\pi C}} \sum_{s=0}^{\infty} \frac{s+1}{s^{\varepsilon/2}} \exp\left(-\frac{1}{8}Cs^{\varepsilon} - \frac{1}{8}Ch^{\varepsilon}\right) \leq C_{\alpha} \exp\left(-\frac{1}{8}Ch^{\varepsilon}\right)$$

for some appropriate constant $C_{\alpha} > 0$, that is, the $\alpha$-mixing coefficients $\alpha_h$ decay at an exponential rate. As Eq. (14) holds for every $\delta > 0$, Eq. (15) simplifies to $\lim_{n\to\infty} n^{1+\varepsilon^*} \mathbb{P}(X_0 > u_n) = \lim_{n\to\infty} n^{1-\varepsilon^*}/u_n = \infty$ for some $\varepsilon^* > 0$, which is true for $u_n \sim n^{\beta_1}$. Thus, the assumptions of Theory 1 reduce to Condition (M).

Choosing $u_n \sim n^{\beta_1}$ and $r_n \sim n^{\beta_2}$ (Buhl and Klüppelberg 2018, see also; Buhl et al. 2019), we have $u_n \to \infty$, $r_n \to \infty$, $n\mathbb{P}(X_0 > u_n) \sim n^{1-\beta_1} \to \infty$ and $r_n \mathbb{P}(X_0 > u_n) \sim n^{\beta_2-\beta_1} \to 0$ as $n \to \infty$. Furthermore, similarly to the assessment above,

$$u_n \sum_{h=1}^{\infty} h^2 \left(2 - \theta(h + r_n)\right)$$

$$\leq u_n \frac{4}{\sqrt{\pi C}} \exp\left(-\frac{1}{8}Cr_n^{\varepsilon}\right) \cdot \sum_{h=1}^{\infty} h^{2-\varepsilon/2} \exp\left(-\frac{1}{8}Ch^{\varepsilon}\right) \overset{n\to\infty}{\longrightarrow} 0,$$

i.e. Eq. (22) holds. We also obtain

$$\sum_{h=k}^{r_n} (2 - \theta(h)) \le \exp\left(-\frac{1}{8}C(k-1)^\varepsilon\right) \sum_{h=1}^{\infty} \frac{4}{\sqrt{\pi C h^\varepsilon}} \exp\left(-\frac{1}{8}C h^\varepsilon\right) \overset{k \to \infty}{\longrightarrow} 0$$

which implies Eq. (23). Consequently, the assertion of the corollary follows from Proposition 2. □

# References

Asadi, P., Davison, A.C., Engelke, S.: Extremes on river networks. Ann. Appl Stat. **9**(4), 2023–2050 (2015)

Bandt, C., Pompe, B.: Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. **88**(17), 174102 (2002)

Basrak, B., Segers, J.: Regularly varying multivariate time series. Stochastic Process Appl. **119**(4), 1055–1080 (2009)

Basrak, B., Davis, R.A., Mikosch, T.: A characterization of multivariate regular variation. Ann. Appl Probab. **12**(3), 908–920 (2002)

Basrak, B., Planinić, H., Soulier, P.: An invariance principle for sums and record times of regularly varying stationary sequences. Probab. Theory Rel. **172**(3-4), 869–914 (2018)

Buhl, S., Klüppelberg, C.: Limit theory for the empirical extremogram of random fields. Stochastic Process Appl. **128**(6), 2060–2082 (2018)

Buhl, S., Klüppelberg, C.: Generalised least squares estimation of regularly varying space-time processes based on flexible observation schemes. Extremes **22**, 223–269 (2019)

Buhl, S., Davis, R.A., Klüppelberg, C., Steinkohl, C.: Semiparametric estimation for isotropic max-stable space-time processes. Bernoulli **25**(4A), 2508–2537 (2019)

Chavez-Demoulin, V., Davison, A.: Modelling time series extremes. REVSTAT-Stat. J. **10**, 109–133 (2012)

Cho, Y.B., Davis, R.A., Ghosh, S.: Asymptotic properties of the empirical spatial extremogram. Scand. J Stat. **43**(3), 757–773 (2016)

Davis, R.A., Hsing, T.: Point process and partial sum convergence for weakly dependent random variables with infinite variance. Ann. Probab. **23**(2), 879–917 (1995)

Davis, R.A., Mikosch, T.: The extremogram: A correlogram for extreme events. Bernoulli **15**(4), 977–1009 (2009)

Davis, R.A., Resnick, S.I.: Basic properties and prediction of max-ARMA processes. Adv. Appl. Probab. **21**(4), 781–803 (1989)

Davis, R.A., Mikosch, T., Cribben, I.: Towards estimating extremal serial dependence via the bootstrapped extremogram. J. Econ. **170**(1), 142–152 (2012)

Davis, R.A., Mikosch, T., Zhao, Y.: Measures of serial extremal dependence and their estimation. Stochastic Process. Appl. **123**(7), 2575–2602 (2013)

Davis, R.A., Drees, H., Segers, J., Warchoł, M.: Inference on the tail process with application to financial time series modeling. J Econom. **205**(2), 508–525 (2018)

de Haan, L.: A spectral representation for max-stable processes. Ann. Probab. **12**(4), 1194–1204 (1984)

Deheuvels, P.: Point processes and multivariate extreme values. J. Multivar. Anal. **13**(2), 257–272 (1983)

Dombry, C., Eyi-Minko, F.: Strong mixing properties of max-infinitely divisible random fields. Stochastic Process Appl. **122**(11), 3790–3811 (2012)

Dombry, C., Ribatet, M.: Functional regular variations. Pareto processes and peaks over threshold. Stat. Interface **8**(1), 9–17 (2015)

Dombry, C., Engelke, S., Oesting, M.: Exact simulation of max-stable processes. Biometrika **103**(2), 303–317 (2016)

Drees, H.: Bootstrapping empirical processes of cluster functionals with application to extremograms. arXiv:1511.00420 (2015)

Drees, H., Rootzén, H.: Limit theorems for empirical processes of cluster functionals. Ann Stat. **38**(4), 2145–2186 (2010)

Embrechts, P., Klüppelberg, C.: T Mikosch Modelling extremal events for insurance and finance (1997)

Ferro, C.A., Segers, J.: Inference for clusters of extreme values. J Royal Stat. Soc. **65**(2), 545–556 (2003)

Kabluchko, Z., Schlather, M., de Haan, L.: Stationary max-stable fields associated to negative definite functions. Ann. Probab. **37**(5), 2042–2065 (2009)

Kallache, M., Rust, H.W., Lange, H., Kropp, J.P., Schellnhuber, H.: Extreme value analysis considering trends: Application to discharge data of the danube river basin. In: Kropp, J. (ed.) Extremis: Disruptive Events and Trends in Climate and Hydrology, pp. 166–183. Springer (2011)

Keller, K., Sinn, M., Emonds, J.: Time series from the ordinal viewpoint. Stochastics and Dynamics **7**(02), 247–272 (2007)

Keller, K., Maksymenko, S., Stolz, I.: Entropy determination based on the ordinal structure of a dynamical system. Discrete Contin. Dyn. Syst B **20**, 3507–3524 (2015)

Leadbetter, M., Lindgren, G., Rootzen, H.: Extremes and Related Properties of Random Sequences and Series. Springer, New York (1983)

Markovich, N.M.: Modeling clusters of extreme values. Extremes **17**(1), 97–125 (2014)

Markovich, N.M.: Erratum to: Modeling clusters of extreme values. Extremes **19**(1), 139–142 (2016)

Markovich, N.M.: Clusters of extremes: modeling and examples. Extremes **20**(3), 519–538 (2017)

Perfekt, R.: Extremal behaviour of stationary markov chains with applications. Ann. Appl. Probab. **4**(2), 529–548 (1994)

Planinić, H., Soulier, P.: The tail process revisited. Extremes **21**(4), 551–579 (2018)

Resnick, S.I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York (2007)

Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer, New York (2008)

Rio, E.: Asymptotic Theory of Weakly Dependent Random Processes. Springer (2017)

Robert, C.Y.: Inference for the limiting cluster size distribution of extreme values. Ann. Stat. **37**(1), 271–310 (2009)

Schnurr, A.: An ordinal pattern approach to detect and to model leverage effects and dependence structures between financial time series. Stat. Papers **55**(4), 919–931 (2014)

Schnurr, A., Dehling, H.: Testing for structural breaks via ordinal pattern dependence. J. Americ. Stat Assoc. **112**(518), 706–720 (2017)

Sinn, M., Keller, K.: Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments. Comp. Stat Data Anal. **55**(4), 1781–1790 (2011)

Sinn, M., Ghodsi, A., Keller, K.: Detecting change-points in time series by maximum mean discrepancy of ordinal pattern distributions. In: Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI), pp 786–794 (2012)

Sinn, M., Keller, K., Chen, B.: Segmentation and classification of time series using ordinal pattern distributions. Eur. Phys. J. Special Topics **222**(2), 587–598 (2013)

Unakafov, A., Keller, K.: Change-point detection using the conditional entropy of ordinal patterns. Entropy **20**(9), 709 (2018)