Statistics for Time Series Extremes

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Tobias Jennessen
aus Düsseldorf

Düsseldorf, Mai 2022

aus dem Mathematischen Institut
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Axel Bücher

2. Prof. Dr. Stanislav Volgushev

Tag der mündlichen Prüfung: 14. Juli 2022

Abstract

An understanding of the extremal behavior of time series is of importance in many applications. For stationary time series, the extremes typically occur in clusters. The extremal index θ, representing the reciprocal of the expected cluster size, and the limiting cluster size distribution π are important measures for analyzing the serial dependence of the extremes of stationary time series. In this thesis, new estimators for θ and π based on the blocks method are proposed. In contrast to many competing estimators from the literature, these estimators only depend on one tuning parameter, i.e., the block length. The introduced estimators are analyzed theoretically, establishing their asymptotic normality, and by means of a large-scale simulation study. Thereby, both disjoint and sliding blocks versions are considered. The sliding blocks estimators are shown to exhibit a smaller asymptotic variance than the corresponding disjoint blocks versions. Further, the sliding blocks estimators perform better with regard to their finite-sample behavior in the context of the simulation study. In specific scenarios, they are also found to be superior to recent competitors from the literature.

In various situations, time series data also exhibit non-stationary behavior, which needs to be accounted for in the statistical analysis. As an approach for modeling non-stationary time series extremes, the proportional tails model introduced by Einmahl, de Haan and Zhou (2016, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 78(1), 31– 51) is extended to allow for serially dependent observations. Here, the proportionality is described by the so-called *scedasis function* c, which can be interpreted as the frequency of extremes; the case where this frequency c is not constant is referred to as *heteroscedastic extremes*. Central limit theorems for estimators for the scedasis function and for the integrated scedasis function are provided. Moreover, different test procedures for assessing whether the extremes are heteroscedastic are developed that are based on a multiplier bootstrap-scheme and on the idea of self-normalization. These tests are examined theoretically, proving their consistency, and shown to perform well within a simulation study. Finally, an estimator for the extremal index of the underlying stationary time series, which governs the dynamics of the extremes, is proposed; its consistency is derived and it is investigated empirically.

Acknowledgements

First, I would like to thank Professor Axel Bücher for his invaluable support and guidance during the course of my doctoral studies. Thank you for numerous discussions and for often finding time for them spontaneously. Your confidence in me and your enthusiasm were very motivating and I have always enjoyed working together.

Further, I am also very grateful to Professor Stanislav Volgushev for agreeing to review this thesis and for affording me an exciting and educational research stay at the University of Toronto.

Moreover, I wish to thank my former and current colleagues at the Chair of Mathematical Statistics and Probability Theory for a very pleasant working atmosphere.

Finally, I owe special thanks to my family and in particular Rabea Gilleßen for their perpetual love and support.

Contents

1 Introduction

Extreme value theory is concerned with analyzing and modeling rare and extreme events. Such events may even be more extreme than any that have already been observed, which statistically means that they lie outside the range of the available data and cannot be adequately treated with traditional methods. However, the effects of extreme events on human societies and the environment are often severe, which is why a (statistical) understanding of extreme events is important to predict and potentially mitigate these effects. Extremes can be versatile and comprise heat waves, floods, earthquakes and large losses in the financial and insurance sector. For instance, a dike should be constructed in a way that it provides protection against water levels that have never been reached before. Or in financial applications, the possibility of extreme losses needs to be taken into account for general risk management. More details and examples in numerous domains of application like hydrology, meteorology, geology, finance and insurance are given in Beirlant et al. (2004) and Coles (2001). In classical extreme value theory, the extremal behavior of a series of independent and identically distributed random variables is analyzed. The probabilistic and statistical theory is well developed, see de Haan and Ferreira (2006); Beirlant et al. (2004) and Resnick (2007) for an overview. However, in many practical situations the assumption of independent variables is not reasonable and a more realistic model is given by considering an underlying stationary time series. The accompanying serial dependence can entail that extremes occur in clusters, rather than isolated as for independent sequences (Hsing et al., 1988). Indeed, it is often observed that the extremes of water levels, wind speeds, temperatures or financial times series cluster in time (Moloney et al., 2019). Such clustering of extremes is important to account for in risk assessment. For example, several days of heavy rainfall, and not just a single extreme rainfall, might be the cause of a flood or a landslide. In the case of temperature, a heat wave emerges which is connected with human health issues, agricultural losses or an increase in the number of forest fires (Scotto et al., 2011). Standard references to the literature on extreme value theory for dependent data are Hsing et al. (1988); Leadbetter (1983); Leadbetter et al. (1983); Leadbetter and Rootzén (1988) and O'Brien (1987).

The statistical analysis of the extremal behavior of a stationary time series typically consists of assessing the tail of the marginal law and assessing the serial dependence of the extremes, i.e., their tendency to occur in clusters. Associated statistical methods usually pursue one of two fundamental principles: the block maxima method and the peak-over-threshold (POT) method (Bücher and Zhou, 2018). The former approach, dating back to Gumbel (1958), consists of partitioning the observations into blocks and extracting the maximum within each block. In its simplest form, the distribution of the resulting block maxima is then approximated by the generalized extreme value (GEV) distribution, motivated by the Fisher-Tippet-Gnedenko Theorem (de Haan and Ferreira, 2006, Theorem 1.1.3). Important recent references to the literature, considering that the block maxima are only asymptotically GEV-distributed or may be serially dependent, include Dombry (2015); Ferreira and de Haan (2015); Dombry and Ferreira (2019); Bücher and Segers (2014) and Bücher and Segers (2018b). The POT method only considers observations that exceed a certain high threshold and the distribution of such exceedances is approximated by the generalized Pareto distribution (Balkema and de Haan, 1974; Pickands, 1975). The literature on the POT method is well developed, some exemplary references are de Haan and Ferreira (2006); Davison and Smith (1990); Hsing (1991b); Drees and Rootzén (2010); Resnick and Stărică (1998) and Drees and Knežević (2020). Heuristically, the POT approach may seem more efficient than the block maxima method since it takes all large observations into account, while the latter method may miss some extremes. However, the theoretical comparisons in Bücher and Zhou (2018) show that either method can be preferable, depending on the quantity of statistical interest such that in general neither can be considered superior.

Traditionally, in the block maxima method, the maxima are taken over disjoint blocks of observations. A relatively new idea, which goes back to Beirlant et al. (2004)(Chapter 10.3.4) and Robert et al. (2009), is to take maxima over sliding blocks. Mathematically, for random variables $X_1, \dots, X_n$, $n \in \mathbb{N}$, the disjoint blocks maxima sample for a block length $b \in \mathbb{N}$ consists of

$$\mathrm{M}_1^{\mathrm{dj}} = \max\{X_1, \dots, X_b\},\ \mathrm{M}_2^{\mathrm{dj}} = \max\{X_{b+1}, \dots, X_{2b}\}, \\ \dots,\ \mathrm{M}_k^{\mathrm{dj}} = \max\{X_{(k-1)b+1}, \dots, X_{kb}\},$$

where $k = \lfloor n/b \rfloor$ denotes the number of disjoint blocks, and the sliding blocks maxima are given by

$$\mathrm{M}_1^{\mathrm{sl}} = \max\{X_1, \dots, X_b\},\ \mathrm{M}_2^{\mathrm{sl}} = \max\{X_2, \dots, X_{b+1}\}, \\ \dots,\ \mathrm{M}_{n-b+1}^{\mathrm{sl}} = \max\{X_{n-b+1}, \dots, X_n\}.$$

It is worthwhile to mention that the maxima sample for sliding blocks is still stationary, provided this holds true for the underlying time series, but has stronger dependencies than for disjoint blocks, which makes the corresponding theoretical analysis more involved (Zou et al., 2021). Since the sample of sliding blocks maxima contains all disjoint blocks maxima and tends to lay more weight on the really large observations, it is plausible that it carries more information than the disjoint blocks sample and can thus be expected to lead to more accurate inference. In fact, the sliding blocks method has been shown to result in a smaller asymptotic variance in certain applications, while the bias is asymptotically the same (Robert et al., 2009; Berghaus and Bücher, 2018; Bücher and Segers, 2018a; Zou et al., 2021; Bücher and Zanger, 2021). On the other hand, the disjoint and sliding blocks variances can also be shown to be equal for estimators of cluster functionals within the POT framework (Cissokho and Kulik, 2020; Drees and Neblung, 2021). Altogether, the sliding blocks maxima method represents an alternative to the classical construction of block maxima that leads to an improvement in many situations.

This thesis is concerned with the statistical analysis of the extremal behavior of time series. New estimators for common measures of the extremal dependence of stationary time series are examined, which are based on the disjoint and sliding blocks method. Further, the extremal behavior is analyzed in a model that allows the underlying observations to be serially dependent and follow different distributions and thus allows for non-stationarities. A detailed description is given in the following.

A primary measure for capturing the serial dependence between the extremes of a stationary time series is provided by the extremal index $\theta \in [0,1]$. The extremal index was introduced in Leadbetter (1983) and is defined as follows. The real-valued stationary sequence $(X_n)_{n\in\mathbb{N}}$ with stationary cumulative distribution function F has an extremal index $\theta \in [0,1]$ if for any $\tau > 0$, there exists a sequence of thresholds $u_n = u_n(\tau)$ such that $\lim_{n\to\infty} n(1-F(u_n)) = \tau$ and

$$\lim_{n\to\infty} \mathrm{P}\Big(\max_{i=1,\dots,n} X_i \le u_n\Big) = e^{-\theta\tau}. \tag{1.1}$$

The first condition states that the expected number of exceedances among $X_1,\dots,X_n$ converges to τ, where every observation above the threshold u_n is called an exceedance. To illustrate the second condition, let $(\tilde{X}_n)_{n\in\mathbb{N}}$ be an independent sequence with cumulative distribution function F, and for $\tau > 0$ choose u_n to satisfy $\lim_{n\to\infty} n(1-F(u_n)) = \tau$ as above. Then, since

$$\lim_{n\to\infty} \mathrm{P}\Big(\max_{i=1,\dots,n} \tilde{X}_i \le u_n\Big) = \lim_{n\to\infty} F^n(u_n) = e^{-\tau},$$

condition (1.1) implies that $\mathrm{P}(\max_{i=1,\ldots,n} X_i \leq u_n) \approx \mathrm{P}(\max_{i=1,\ldots,n} \tilde{X}_i \leq u_n)^\theta$ for large n. Thus, the extremal index determines the (shrinking) effect the dependence of the extremes may have on the distribution of the maximum of the dependent sequence compared to its independent analogue (Leadbetter, 1983, Theorem 2.4). Obviously, if $(X_n)_{n\in\mathbb{N}}$ is an independent sequence, then $\theta = 1$; however, the case $\theta = 1$ can also be true for dependent sequences (Beirlant et al., 2004, page 378). From now on $\theta > 0$ is assumed; the case $\theta = 0$ is of little practical interest and commonly excluded (Leadbetter et al., 1983, page 72). The extremal index has several interpretations, among which the arguably most important one is its characterization as the reciprocal of the expected size of an extremal cluster. More precisely, split the observations into successive disjoint blocks of length b_n, where $b_n = o(n)$ and $b_n \to \infty$ as $n \to \infty$, and consider a threshold u_n as above. The set of all exceedances of the level u_n within a block is called a cluster. Then, under suitable conditions

$$\theta^{-1} = \lim_{n\to\infty} \mathrm{E}\left[\sum_{i=1}^{b_n} \mathbf{1}(X_i > u_n)\Big| \max_{i=1,\ldots,b_n} X_i > u_n\right] \tag{1.2}$$

such that θ^{-1} is the limiting mean number of exceedances in blocks with at least one exceedance (Hsing et al., 1988). Therefore, the extremal index measures how many extremes occur together on average. Another interpretation by O'Brien (1987) states that under suitable conditions

$$\theta = \lim_{n\to\infty} \mathrm{P}\Big(\max_{i=2,\ldots,b_n} X_i \leq u_n \big| X_1 > u_n\Big),$$

meaning that θ is the limiting probability that an exceedance is followed by a run of observations below the threshold. Further, the extremal index can be characterized in terms of the times between exceedances (Beirlant et al., 2004, Chapter 10.3.4).

As a consequence of these interpretations, the estimation of the extremal index can be an important part of the statistical analysis of the extremal dependence of a stationary time series. The statistical relevance of estimating θ is further emphasized by the fact that one risks underestimating the marginal quantiles and overestimating the return levels in many scenarios if the extremal index is neglected (Beirlant et al., 2004, page 381). Inference about θ has received a correspondingly large amount of attention in the literature. Common approaches are the blocks method, the runs method and the inter-exceedance times method. The first two methods typically depend on a threshold sequence and a cluster identification scheme (such as a block length), whereas estimators based on inter-exceedance times only depend on a threshold sequence. Respective references are Hsing (1993); Robert et al. (2009); Ferro and Segers (2003); Süveges (2007); Süveges and Davison (2010); Smith and Weissman (1994); Weissman and Novak (1998) and Laurini and Tawn (2003); see also Beirlant et al. (2004)(Chapter 10.3.4) for an overview of these methods. In many papers on the estimation of the extremal index, no or only incomplete asymptotic theory is given (Berghaus and Bücher, 2018, page 2308).

Chapter 2.1 of this thesis focuses on a class of method of moments estimators for the extremal index based on the blocks method, which shows an improvement over a recent (disjoint and sliding) blocks estimator proposed in Northrop (2015) and analyzed theoretically in Berghaus and Bücher (2018). These new estimators only require a block length parameter and rely on the construction of approximate samples from the exponential distribution with parameter θ. To this, the observations are partitioned into blocks and in each block a transformation of the block maximum is applied that asymptotically follows the exponential distribution with parameter θ by equation (1.1). This approximate sample of exponentially distributed observations is used to estimate θ via the method of moments, whereas in Northrop (2015) and Berghaus and Bücher (2018) the maximum likelihood estimator of the exponential distribution was used. Thereby, both disjoint and sliding blocks are considered. The asymptotic normality of the resulting estimators is established and the asymptotic variances in the sliding blocks case are shown to be smaller than their disjoint blocks counterparts. Further, the asymptotic variance can be seen to be smaller than the one of the estimator analyzed in Berghaus and Bücher (2018) in some scenarios. In a simulation study, all methods are compared with several estimators from the literature regarding their finite-sample properties.

Another important measure for describing the serial dependence of a stationary time series at extreme levels is the limiting cluster size distribution π. This object is a probability distribution on the positive integers, where $\pi(j)$ approximately represents the probability that extreme observations of a stationary sequence $(X_n)_{n\in\mathbb{N}}$ occur in a temporal cluster of size $j \in \mathbb{N}$. More formally, for appropriately chosen threshold sequence u_n and integer sequence b_n with $b_n = o(n)$ and $b_n \to \infty$ as $n \to \infty$,

$$\pi(j) = \lim_{n\to\infty} \mathrm{P}\Big(\sum_{i=1}^{b_n} \mathbf{1}(X_i > u_n) = j \;\Big|\; \max_{i=1,\dots,b_n} X_i > u_n \Big), \quad j \in \mathbb{N}, \tag{1.3}$$

see Hsing et al. (1988) for conditions under which this limit exists. By this definition, the distribution π is of natural interest, but beyond that it is further an appealing object to study since it shows up as one of two characterizing objects in the limiting distribution of the point process of exceedances; the other one being the extremal index. More precisely, the point process of exceedances is defined as $N_n(\cdot) = \sum_{i=1}^{n} \mathbf{1}(i/n \in \cdot, X_i > u_n)$ and is a commonly studied object in extreme value theory in general (Hsing et al., 1988; Beirlant et al., 2004, Chapter 10.3.1). Here, again the threshold sequence u_n is chosen in a way such that the expected number of exceedances remains finite, i.e., it satisfies $\lim_{n\to\infty} n(1 - F(u_n)) = \tau$ for some $\tau \in (0, \infty)$, where F is the stationary cumulative distribution function of $(X_n)_{n\in\mathbb{N}}$. This process counts the times, normalized by n, at which the threshold u_n is exceeded. In N_n, all the points making up a cluster from a block of length $b_n = o(n)$ converge to a single point such that, in the limit, the points in N_n represent the cluster positions. It turns out that if the extremal index and the limit in (1.3) exist and appropriate long range dependence conditions hold, the limiting point process is a compound Poisson process with intensity $\theta\tau$ and compounding distribution π (Hsing et al., 1988, Theorem 4.1 and 4.2). In particular, this means that, in the limit, the clusters occur randomly in the manner of a Poisson process, on average there are $\theta\tau$ clusters, and their sizes are independent and distributed according to π. Under mild additional assumptions the extremal index can be seen to satisfy $\theta^{-1} = \sum_{j\in\mathbb{N}} j\pi(j)$ as suggested by equation (1.2) (Hsing et al., 1988; Beirlant et al., 2004, Chapter 10.3.1).

There are only a few papers, which are concerned with estimating the limiting cluster size distribution (Robert, 2009b, page 273). Estimators for π have been studied in Hsing (1991a); Ferro (2003) and Robert (2009b). While the estimator in Ferro (2003) is based on the inter-exceedance times method, the estimators in the other two references are based on the (disjoint) blocks method. These two references also provide asymptotic theory, while the estimator from Ferro (2003) has been analyzed theoretically in Robert (2009a). Sliding blocks versions of peak-over-threshold estimators for a general class of cluster functionals including the limiting cluster size distribution are considered in Cissokho and Kulik (2020). It should be mentioned that a powerful framework for the asymptotic analysis of these methods is provided by results in Drees and Rootzén (2010) on empirical processes for cluster functionals. Further, at this point it is worthwhile to mention that a recently introduced alternative object for capturing the serial dependence of extremes is given by the tail process from Basrak and Segers (2009). Both θ and π can be seen to be functionals of this process (Cissokho and Kulik, 2020; Kulik and Soulier, 2020, Chapter 6.2). Statistical inference on the tail process (for selected functionals) has been considered in Davis et al. (2018); Drees and Knežević (2020); Drees et al. (2015) and Neblung (2021).

In Chapter 2.2, new estimators for π based on a disjoint and sliding blocks declustering scheme are proposed. These estimators are defined recursively and rest on the construction of approximate samples from the exponential distribution, as for the estimation of the extremal index, and on making use of the concrete form of the limiting distribution of the point process of exceedances. This approach is similar to the one in Robert (2009b), but, unlike in that reference, the resulting estimators only depend on one tuning parameter (i.e., the block length). The asymptotic normality of the estimators is derived for both disjoint and sliding blocks, under a side result on weak convergence of an empirical process associated with compounding probabilities and sliding blocks. The sliding blocks estimator can be seen to outperform the disjoint blocks version theoretically and both are shown, in the context of a simulation study, to exhibit good finite-sample properties compared to the estimators by Hsing (1991a); Ferro (2003) and Robert (2009b).

The assumption of stationarity of the underlying time series constitutes a realistic model in many situations. However, time series data may still exhibit non-stationary behavior in certain cases coming from numerous fields of application (Dahlhaus, 1997; Dahlhaus and Giraitis, 1998). In particular, there are suggestions in climatology that extreme weather events are becoming more frequent as a result of climate change (Klein Tank and Können, 2003; Zolina et al., 2009). Therefore, to account for non-stationarities and to investigate temporal trends for such extreme events is important. Chapter 2.3 is concerned with an extension of the proportional tails model introduced in Einmahl et al. (2016) to the case of dependent data as an approach for modeling non-stationary time series extremes. In this model, the observations $X_1^{(n)}, \dots, X_n^{(n)}$ exhibit serial dependence and are drawn from a distribution that changes as time progresses, more precisely, $X_i^{(n)}$ is assumed to have a continuous cumulative distribution function $F_{n,i}$, $i = 1, \dots, n$. Further, it is assumed that all these distribution functions share a common right endpoint x^* and that there is a continuous cumulative distribution function F with the same right endpoint x^* which is strictly increasing on its support, and a positive function c on $[0, 1]$ such that the following proportional tails condition holds:

$$\lim_{x \uparrow x^*} \frac{1 - F_{n,i}(x)}{1 - F(x)} = c\big(\tfrac{i}{n}\big).$$

The function c is called the *scedasis function* and assumed to be a continuous probability density function. Thereby, the scedasis function can be interpreted as the frequency of extremes, and the case where c is not constant equal to one is referred to as *heteroscedastic extremes*. It is worthwhile to mention that the above limit condition concerns the comparison of the distribution tails only and makes no assumption on the remaining parts of the distributions (Einmahl et al., 2016). In the latter reference, the observations were assumed to be independent. In Chapter 2.3, they may be serially dependent in such a way that, for each $n \in \mathbb{N}$, the marginal transformations $U_1^{(n)}, \dots, U_n^{(n)}$ where $U_i^{(n)} = F_{n,i}(X_i^{(n)})$ are an excerpt from a stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ whose distribution does not depend on n. The extremal behavior of this series will be governed by the extremal index and the concept of regular variation (Basrak and Segers, 2009). Altogether, this setting describes a non-parametric model which allows for serial dependence and different distribution tails and can be used for potential temporal trend detection of time series extremes.

There are other approaches in the literature that are concerned with non-identically distributed extremes. Smooth non-stationarity has often been captured by parametric regression models as in Davison and Smith (1990) and Coles (2001). They considered models for exceedances where a linear and log-linear trend is imposed on the parameters of the generalized Pareto distribution, respectively; in both, no asymptotic theory is provided. In Hall and Tajvidi (2000), non-parametric trends in parameters of the generalized Pareto and extreme value distribution are estimated and corresponding asymptotic results are developed, allowing for serial dependence. Further, parametric trends in a model similar to Einmahl et al. (2016) are considered in de Haan et al. (2015). Both provide asymptotic theory in the case of serially independent observations. Mefleh (2018) also imposed the proportional tails model by Einmahl et al. (2016), assuming that the scedasis function is of parametric form. Recently, de Haan and Zhou (2021) considered estimating a continuously changing extreme value index, and Einmahl et al. (2022) provided a multivariate extension of Einmahl et al. (2016) accounting for spatial dependence. A brief overview of other approaches is contained in de Haan et al. (2015).

Chapter 2.3 deals with an extension of the proportional tails model by Einmahl et al. (2016) to allow for serial dependence as described above. The asymptotic behavior of estimators for the scedasis function c and the integrated scedasis function $C(s) = \int_0^s c(x)\,\mathrm{d}x$, $s \in [0, 1]$, that were studied in Einmahl et al. (2016) in the independent case, is analyzed. Thereby, a pointwise and functional central limit theorem is provided for the estimation of c and C, respectively. The asymptotic variance and covariance functional turn out to be different than in the independent case. Moreover, different test procedures on the presence of heteroscedastic extremes are developed that are based on a multiplier bootstrap-scheme and on the idea of self-normalization. These tests are examined theoretically and shown to perform well in the context of a simulation study, with the bootstrap test being slightly more powerful but computationally more intensive. Finally, an estimator for the extremal index of the stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ is introduced, which constitutes a modification of the block maxima estimator from Berghaus and Bücher (2018). Its consistency is derived and it is further analyzed regarding its finite-sample performance within a simulation study.

This thesis is structured cumulatively and organized as follows. In Chapter 2, the articles in which the author of this thesis is involved are listed. Here, Chapter 2.1 contains the first article, which is concerned with the estimation of the extremal index of a stationary time series. The second article is included in Chapter 2.2 and deals with estimating the limiting cluster size distribution of a stationary time series. Finally, the third article is contained in Chapter 2.3 and concerns investigating the extremes of heteroscedastic time series in the proportional tails model described above. In Chapter 3, a brief outlook on a potential continuation of this work is presented along with some open research questions. Finally, an author contribution statement is deferred to the appendix, outlining the individual contributions of the authors to the articles included in Chapter 2.

2 Included articles

The following articles are included in this thesis. They are reprinted with the permission of the respective journals.

2.1) Bücher, A. and Jennessen, T. (2020). Method of moments estimators for the extremal index of a stationary time series. *Electronic Journal of Statistics*, 14(2):3103-3156. (DOI: 10.1214/20-EJS1734)

2.2) Bücher, A. and Jennessen, T. (2022). Statistical analysis for stationary time series at extreme levels: New estimators for the limiting cluster size distribution. *Stochastic Processes and their Applications*, 149:75-106. (DOI: 10.1016/j.spa.2022.03.004)

2.3) Bücher, A. and Jennessen, T. (2022). Statistics for Heteroscedastic Time Series Extremes.

Submitted for publication in *Bernoulli* and publicly available on *arXiv*:2204.09534. (DOI: 10.48550/arXiv.2204.09534)

Electronic Journal of Statistics
Vol. 14 (2020) 3103–3156
ISSN: 1935-7524
https://doi.org/10.1214/20-EJS1734

Method of moments estimators for the extremal index of a stationary time series

Axel Bücher and Tobias Jennessen

Heinrich-Heine-Universität Düsseldorf
Universitätsstr. 1
40225 Düsseldorf
e-mail: axel.buecher@hhu.de; tobias.jennessen@hhu.de

Abstract: The extremal index θ, a number in the interval $[0, 1]$, is known to be a measure of primal importance for analyzing the extremes of a stationary time series. New rank-based estimators for θ are proposed which rely on the construction of approximate samples from the exponential distribution with parameter θ that is then to be fitted via the method of moments. The new estimators are analyzed both theoretically as well as empirically through a large-scale simulation study. In specific scenarios, in particular for time series models with $\theta \approx 1$, they are found to be superior to recent competitors from the literature.

Received December 2019.

Contents

1. Introduction

The statistical analysis of the extremal behavior of a stationary time series is important in many fields of application, such as in hydrology, meteorology, finance or actuarial science [1]. Such an analysis typically consists of two steps: (1) assessing the tail of the marginal law and (2) assessing the serial dependence of the extremes, that is, the tendency that extreme observations occur in clusters. The present work is concerned with step (2). The most common and simplest mathematical object capturing the serial dependence between the extremes is provided by the extremal index $\theta \in [0,1]$. In a suitable asymptotic framework, the extremal index can be interpreted as the reciprocal of the expected size of a cluster of extreme observations. The underlying probabilistic theory was worked out in [18, 19, 23, 17, 20].

Estimating the extremal index based on a finite stretch of observations from the time series has been extensively studied in the literature. An early overview is provided in Section 10.3.4 in [1], where the estimators are classified into three groups: estimators based on the blocks method, the runs method or the inter-exceedance time method. Respective references are [16, 31, 13, 32, 27, 22, 12, 11, 5], among many others. The proposed estimators typically depend on two or, arguably preferable, one parameter to be chosen by the statistician. The present paper is on a class of method of moments estimators (based on the blocks method), which improves upon a recent estimator proposed by Paul Northrop in [22] and analyzed theoretically in [3].

Some notations and assumptions are necessary for the motivation of the new class of estimators. Throughout the paper, $X_1, X_2, \dots$ denotes a stationary sequence of real-valued random variables with continuous cumulative distribution function (c.d.f.) F. The sequence is assumed to have an extremal index $\theta \in (0,1]$, i.e., for any $\tau > 0$, there exists a sequence $u_b = u_b(\tau), b \in \mathbb{N}$, such that $\lim_{b\to\infty} b\bar F(u_b) = \tau$ and

$$\lim_{b\to\infty} \mathbb{P}(M_{1:b} \le u_b) = e^{-\theta\tau}, \tag{1.1}$$

where $\bar F = 1 - F$ and $M_{1:b} = \max\{X_1, \dots, X_b\}$. Next, define a sequence of standard uniform random variables by $U_s = F(X_s)$ and let

$$Y_{1:b} = -b\log(N_{1:b}), \qquad N_{1:b} = F(M_{1:b}) = \max\{U_1, \dots, U_b\}. \tag{1.2}$$

Since $b\bar{F}\{F^{\leftarrow}(e^{-x/b})\} = b(1-e^{-x/b}) \to x$ for $b \to \infty$, it follows from (1.1) that, for any $x > 0$,

$$\mathbb{P}(Y_{1:b} \ge x) = \mathbb{P}(M_{1:b} \le F^{\leftarrow}(e^{-x/b})) \to e^{-\theta x}, \tag{1.3}$$

where $F^{\leftarrow}(z) = \inf\{y \in \mathbb{R} : F(y) \ge z\}$ denotes the generalized inverse of F evaluated at $z \in \mathbb{R}$. In other words, for large block length b, $Y_{1:b}$ approximately follows an exponential distribution with parameter θ, denoted by $\mathrm{Exp}(\theta)$ throughout. This inspired [22] and [3] to estimate θ by the maximum likelihood estimator for the exponential distribution; see Section 2 below for details on how to arrive at an observable (rank-based) approximate sample from the $\mathrm{Exp}(\theta)$-distribution based on an observed stretch of length n from the time series $(X_s)_{s\in\mathbb{N}}$.

The idea of transforming observations into a sample of exponentially distributed observations is actually not new within extreme value statistics: it is also, among many others, the main motivation for the Pickands estimator in multivariate extremes [25, 14]. More precisely, if (X, Y) is a bivariate random vector from a multivariate extreme value distribution with Pickands function $A = (A(w))_{w\in[0,1]}$, then $\xi(w) = \min\{-\log F_X(X)/(1-w), -\log F_Y(Y)/w\}$ is exponentially distributed with parameter $A(w)$. Given a sample of size n from (X, Y), we may replace F_X and F_Y by their empirical counterparts and arrive at an approximate sample of size n from the $\mathrm{Exp}(A(w))$-distribution, to be, for instance, estimated by the maximum likelihood estimator.

The present paper is now motivated by the following observation: while the maximum likelihood estimator is asymptotically efficient in the ideal situation of observing an i.i.d. sample from the exponential distribution, it was shown in [14] for rank-based estimators of the Pickands function that it is in fact more efficient to consider alternative estimators based on the method of moments, such as a rank-based version of the CFG-estimator [6]. Given that Northrop's blocks estimator is also rank-based, the main motivation of this work is to consider CFG-type estimators for the extremal index θ. Alongside, we will also investigate other moment-based estimators, including one that is closely connected to the madogram estimator in [21]. We will show that, depending on the true value of θ, the new estimators may either exhibit a smaller or a larger asymptotic variance than Northrop's maximum likelihood estimator. In particular, we will show that the CFG-type estimator's variance is substantially smaller for θ close to one, i.e., for time series with little clustering of extremes.

The remaining parts of this paper are organized as follows: in Section 2, we collect some results about certain useful moments of the exponential distribution and use those to introduce the new estimators for θ. Regularity assumptions needed to prove asymptotic results are summarized and discussed in Section 3. The paper's main results are then presented in Section 4, alongside with a discussion of certain aspects of the derived asymptotic variance formulas. Section 5 is about a particular time series model, for which we show that all regularity conditions imposed in Section 3 are met. The finite-sample performance of the new estimators is investigated in a Monte-Carlo simulation study in Section 6. Finally, all proofs are postponed to Section A.

2. Definition of estimators

Recall the definition of $Y_{1:b}$ in (1.2), where $b \in \mathbb{N}$. Similarly, let

$$Z_{1:b} = b(1 - N_{1:b}), \qquad N_{1:b} = F(M_{1:b}) = \max\{U_1, \dots, U_b\},$$

and note that, as $b \to \infty$ and for any $x > 0$,

$$\mathbb{P}(Z_{1:b} \geq x) = \mathbb{P}(M_{1:b} \leq F^{\leftarrow}(1 - x/b)) \to e^{-\theta x} \tag{2.1}$$

by similar arguments as for $Y_{1:b}$. The convergence relations in (1.3) and (2.1) serve as a basis for the method of moments estimators defined below.

Subsequently, let $X_1, \dots, X_n$ denote a finite stretch of observations from the stationary sequence $(X_s)_{s \geq 1}$. Within Section 2.1 and 2.2, we start by using (1.3) and (2.1) to derive some observable, approximate samples from the Exp(θ)-distribution. In Section 2.3, we collect some moment equations for the exponential distribution, which will then be used to motivate new estimators for the extremal index in Section 2.4.

2.1. *Approximate* Exp(θ)*-samples based on disjoint blocks maxima*

Divide the sample $X_1, \dots, X_n$ into k_n successive blocks of size b_n, and for simplicity assume that $n = b_n k_n$ (otherwise, the last block of less than b_n observations should be deleted). For $i = 1, \dots, k_n$, let

$$M_{ni} = \max\{X_{(i-1)b_n+1}, \dots, X_{ib_n}\}$$

denote the maximum of the X_s in the ith block of observations and let

$$Y_{ni} = -b_n \log N_{ni}, \qquad Z_{ni} = b_n(1 - N_{ni}), \qquad N_{ni} = F(M_{ni}).$$

Due to relations (1.3) and (2.1), if the block size $b = b_n$ is sufficiently large, the (unobservable) random variables Y_{ni} and Z_{ni} are approximately exponentially distributed with parameter θ. Observable counterparts are obtained by replacing F by the (slightly adjusted) empirical c.d.f. $\hat{F}_n(x) = (n+1)^{-1} \sum_{s=1}^n \mathbb{1}(X_s \leq x)$, giving rise to the definitions

$$\hat{Y}_{ni} = -b_n \log \hat{N}_{ni}, \qquad \hat{Z}_{ni} = b_n(1 - \hat{N}_{ni}), \qquad \hat{N}_{ni} = \hat{F}_n(M_{ni}).$$

Both the samples $\mathcal{Y}_n^{\text{db}} = \{\hat{Y}_{ni} : i = 1, \dots, k_n\}$ and $\mathcal{Z}_n^{\text{db}} = \{\hat{Z}_{ni} : i = 1, \dots, k_n\}$ will be used later to define *disjoint blocks estimators for* θ (note that both samples are dependent over i due to the use of $\hat{F}_n$, which complicates the asymptotic analysis).

2.2. *Approximate* Exp(θ)*-samples based on sliding blocks maxima*

As in the previous paragraph, let n denote the sample size and b_n denote a block length parameter (the assumption that $k_n = n/b_n \in \mathbb{N}$ is not needed, no discarding is necessary). For $t = 1, \dots, n - b_n + 1$, let

$$M_{nt}^{\text{sb}} = M_{t:t+b_n-1} = \max\{X_t, \dots, X_{t+b_n-1}\}$$

denote the maximum of the X_s in a block of length b_n starting at observation t. Define

$$Y_{nt}^{\mathrm{sb}} = -b_n \log N_{nt}^{\mathrm{sb}}, \qquad Z_{nt}^{\mathrm{sb}} = b_n(1 - N_{nt}^{\mathrm{sb}}), \qquad N_{nt}^{\mathrm{sb}} = F(M_{nt}^{\mathrm{sb}}),$$
$$\hat{Y}_{nt}^{\mathrm{sb}} = -b_n \log \hat{N}_{nt}^{\mathrm{sb}}, \qquad \hat{Z}_{nt}^{\mathrm{sb}} = b_n(1 - \hat{N}_{nt}^{\mathrm{sb}}), \qquad \hat{N}_{nt}^{\mathrm{sb}} = \hat{F}_n(M_{nt}^{\mathrm{sb}}).$$

By the same heuristics as before, the observable samples $\mathcal{Y}_n^{\mathrm{sb}} = \{\hat{Y}_{nt}^{\mathrm{sb}} : t = 1, \dots, n - b_n + 1\}$ and $\mathcal{Z}_n^{\mathrm{sb}} = \{\hat{Z}_{nt}^{\mathrm{sb}} : t = 1, \dots, n - b_n + 1\}$ are approximate samples from the exponential distribution and will be used later to define *sliding blocks estimators for* θ (both samples are heavily dependent over i due to the use of $\hat{F}_n$ and the use of overlapping blocks).

2.3. Preliminaries on the exponential distribution

Some important moment equations, valid for a random variable ξ, which is $\mathrm{Exp}(\theta)$-distributed, are collected. First,

$$\mathrm{E}[\log \xi] = -\log \theta - \gamma =: \varphi_{(\mathrm{C})}(\theta), \tag{CFG}$$

where $\gamma = -\int_0^\infty \log(x) e^{-x} \,\mathrm{d}x \approx 0.577$ denotes the Euler-Mascheroni-constant. Equation (CFG) is the basis for motivating the CFG-estimator, see [6, 14] and the details in Section 1. Next, note that

$$\mathrm{E}[\exp(-\xi)] = \frac{\theta}{1+\theta} =: \varphi_{(\mathrm{M})}(\theta), \tag{MAD}$$

which serves as a basis for the madogram, see [21]. A further choice, including (CFG) as a limit, is provided by

$$\mathrm{E}[\xi^{1/p}] = \theta^{-1/p}\Gamma(1 + 1/p) =: \varphi_{(\mathrm{R}),p}(\theta), \tag{ROOT}$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \,\mathrm{d}t$ denotes the Gamma function and where $p > 0$. The moment estimator in case of $p = 1$ will turn out to coincide with Northrop's maximum likelihood estimator. Also note that the previous equation is equivalent to

$$\mathrm{E}\Big[\frac{\xi^{1/p} - 1}{1/p}\Big] = \frac{\theta^{-1/p}\Gamma(1 + 1/p) - 1}{1/p} =: \tilde{\varphi}_{(\mathrm{R}),\mathrm{p}}(\theta), \tag{2.2}$$

and taking the limits for $p \to \infty$ on both sides (interchanging the limit and the expectation on the left) exactly yields Equation (CFG).

2.4. Definition of the estimators

Let $\chi_m = \{\xi_1, \dots, \xi_m\}$ denote a generic sample (not necessarily independent) from the $\mathrm{Exp}(\theta)$-distribution. Replacing the moments in Equations (CFG), (MAD) and (ROOT) by their empirical counterparts and solving the equation for θ, we obtain the following three estimators for θ:

$$\hat\theta_{\mathrm{CFG}}(\chi_m) = e^{-\gamma}\exp\Big\{-\frac{1}{m}\sum_{i=1}^m \log(\xi_i)\Big\},$$
$$\hat\theta_{\mathrm{MAD}}(\chi_m) = \frac{\frac{1}{m}\sum_{i=1}^m \exp(-\xi_i)}{1-\frac{1}{m}\sum_{i=1}^m \exp(-\xi_i)},$$
$$\hat\theta_{\mathrm{R},p}(\chi_m) = \Gamma(1+1/p)^p\Big(\frac{1}{m}\sum_{i=1}^m \xi_i^{1/p}\Big)^{-p},$$

where $p>0$. It may be verified that $\lim_{p\to\infty}\hat\theta_{\mathrm{R},p}(\chi_m)=\hat\theta_{\mathrm{CFG}}(\chi_m)$, see also (2.2) for another relationship between the two estimators. Next, replacing χ_m by any of the four samples $\mathcal{Y}_n^{\mathrm{db}}, \mathcal{Z}_n^{\mathrm{db}}, \mathcal{Y}_n^{\mathrm{sb}}$ or $\mathcal{Z}_n^{\mathrm{sb}}$ defined in Sections 2.1 and 2.2, we finally arrive at 12 method of moments estimators for θ. We use the suggestive notations

$$\hat\theta^{\mathcal{Y}_n}_{\mathrm{db,CFG}} = \hat\theta_{\mathrm{CFG}}(\mathcal{Y}_n^{\mathrm{db}}), \qquad \hat\theta^{\mathcal{Z}_n}_{\mathrm{sb,MAD}} = \hat\theta_{\mathrm{MAD}}(\mathcal{Z}_n^{\mathrm{sb}})$$

to, e.g., denote the disjoint blocks CFG-estimator based on the $\hat Y_{ni}$ and the sliding blocks madogram-estimator based on the $\hat Z_{ni}$, respectively. Note that the four estimators of the form $\hat\theta^{\mathcal{Y}_n}_{\mathrm{m,R},1}, \hat\theta^{\mathcal{Z}_n}_{\mathrm{m,R},1}, \mathrm{m}\in\{\mathrm{db},\mathrm{sb}\}$, are the (pseudo) maximum likelihood (PML) estimators considered in [3].

3. Mathematical preliminaries

Further mathematical details are necessary before we can state asymptotic results about the estimators defined in the previous section. The asymptotic framework and the conditions are mostly similar as in Section 2 in [3], but will be repeated here for the sake of completeness.

The serial dependence of the time series $(X_s)_{s\in\mathbb{N}}$ will be controlled via mixing coefficients. For two sigma-fields $\mathcal{F}_1, \mathcal{F}_2$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, let

$$\alpha(\mathcal{F}_1,\mathcal{F}_2) = \sup_{A\in\mathcal{F}_1, B\in\mathcal{F}_2} |\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)|.$$

In time series extremes, one usually imposes assumptions on the decay of the mixing coefficients between sigma-fields generated by $\{X_s\mathbb{1}(X_s>F^{\leftarrow}(1-\varepsilon_n)) : s\le\ell\}$ and $\{X_s\mathbb{1}(X_s>F^{\leftarrow}(1-\varepsilon_n)) : s\ge \ell+k\}$, where $\varepsilon_n\to 0$ is some sequence reflecting the fact that only the dependence in the tail needs to be restricted (see, e.g., 29). As in [3], we need a slightly stronger condition, that also controls the dependence between the smallest of all block maxima. More precisely, for $-\infty\le p<q\le\infty$ and $\varepsilon\in(0,1]$, let $\mathcal{B}^\varepsilon_{p:q}$ denote the sigma algebra generated by $U_s^\varepsilon := U_s\mathbb{1}(U_s>1-\varepsilon)$ with $s\in\{p,\dots,q\}$ and define, for $\ell\ge 1$,

$$\alpha_\varepsilon(\ell) = \sup_{k\in\mathbb{N}} \alpha(\mathcal{B}^\varepsilon_{1:k}, \mathcal{B}^\varepsilon_{k+\ell:\infty}).$$

In Condition 3.1(iii) below, we will impose a condition on the decay of the mixing coefficients for small values of ε. Note that the coefficients are bounded by the standard alpha-mixing coefficients of the sequence U_s, which can be retrieved for $\varepsilon = 1$.

The extremes of a time series may be conveniently described by the point process of normalized exceedances. The latter is defined, for a Borel set $A \subset E := (0, 1]$ and a number $x \in [0, \infty)$, by

$$N_n^{(x)}(A) = \sum_{s=1}^{n} \mathbb{1}(s/n \in A, U_s > 1 - x/n).$$

Note that $N_n^{(x)}(E) = 0$ iff $N_{1:n} \le 1 - x/n$; the probability of that event converging to $e^{-\theta x}$ under the assumption of the existence of the extremal index θ.

Fix $m \ge 1$ and $x_1 > \cdots > x_m > 0$. For $1 \le p < q \le n$, let $\mathcal{F}_{p:q,n}^{(x_1,\ldots,x_m)}$ denote the sigma-algebra generated by the events $\{U_i > 1 - x_j/n\}$ for $p \le i \le q$ and $1 \le j \le m$. For $1 \le \ell \le n$, define

$$\alpha_{n,\ell}(x_1, \ldots, x_m) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{1:s,n}^{(x_1,\ldots,x_m)}, B \in \mathcal{F}_{s+\ell:n,n}^{(x_1,\ldots,x_m)}, 1 \le s \le n - \ell\}.$$

The condition $\Delta_n(\{u_n(x_j)\}_{1\le j\le m})$ is said to hold if there exists a sequence $(\ell_n)_n$ with $\ell_n = o(n)$ such that $\alpha_{n,\ell_n}(x_1, \ldots, x_m) = o(1)$ as $n \to \infty$. A sequence $(q_n)_n$ with $q_n = o(n)$ is said to be $\Delta_n(\{u_n(x_j)\}_{1\le j\le m})$-separating if there exists a sequence $(\ell_n)_n$ with $\ell_n = o(q_n)$ such that $nq_n^{-1}\alpha_{n,\ell_n}(x_1, \ldots, x_m) = o(1)$ as $n \to \infty$. If $\Delta_n(\{u_n(x_j)\}_{1\le j\le m})$ is met, then such a sequence always exists, simply take $q_n = \lfloor \max\{n\alpha_{n,\ell_n}^{1/2}, (n\ell_n)^{1/2}\}\rfloor$.

By Theorems 4.1 and 4.2 in [17], if the extremal index exists and the $\Delta(u_n(x))$-condition is met ($m = 1$), then a necessary and sufficient condition for weak convergence of $N_n^{(x)}$ is convergence of the conditional distribution of $N_n^{(x)}(B_n)$ with $B_n = (0, q_n/n]$ given that there is at least one exceedance of $1 - x/n$ in $\{1, \ldots, q_n\}$ to a probability distribution π on $\mathbb{N}$, that is,

$$\lim_{n\to\infty} \mathbb{P}(N_n^{(x)}(B_n) = j \mid N_n^{(x)}(B_n) > 0) = \pi(j) \qquad \forall j \ge 1,$$

where q_n is some $\Delta(u_n(x))$-separating sequence. Moreover, in that case, the convergence in the last display holds for any $\Delta(u_n(x))$-separating sequence q_n, and the weak limit of $N_n^{(x)}$ is a compound poisson process $\mathrm{CP}(\theta x, \pi)$. If the $\Delta(u_n(x))$-condition holds for any $x > 0$, then π does not depend on x (17, Theorem 5.1).

A multivariate version of the latter results is stated in [24], see also the summary in [27], page 278, and the thesis [15]. Suppose that the extremal index exists and that the $\Delta(u_n(x_1), u_n(x_2))$-condition is met for any $x_1 \ge x_2 \ge 0$, $x_1 \ne 0$. Moreover, assume that there exists a family of probability measures $\{\pi_2^{(\sigma)} : \sigma \in [0, 1]\}$ on $\mathcal{J} = \{(i, j) : i \ge j \ge 0, i \ge 1\}$, such that, for all $(i, j) \in \mathcal{J}$,

$$\lim_{n\to\infty} \mathbb{P}(N_n^{(x_1)}(B_n) = i, N_n^{(x_2)}(B_n) = j \mid N_n^{(x_1)}(B_n) > 0) = \pi_2^{(x_2/x_1)}(i, j),$$

where q_n is some $\Delta(u_n(x_1), u_n(x_2))$-separating sequence. In that case, the two-level point process $\boldsymbol{N}_n^{(x_1,x_2)} = (N_n^{(x_1)}, N_n^{(x_2)})$ converges in distribution to a point process with characterizing Laplace transform explicitly stated in [27] on top of page 278. Note that

$$\pi_2^{(1)}(i,j) = \pi(i)\mathbb{1}(i=j), \qquad \pi_2^{(0)}(i,j) = \pi(i)\mathbb{1}(j=0).$$

Finally, we will need the tail empirical pocess

$$e_n(x) = \frac{1}{\sqrt{k_n}} \sum_{s=1}^{n} \left\{ \mathbb{1}\left(U_s > 1 - \frac{x}{b_n}\right) - \frac{x}{b_n} \right\}, \qquad x \geq 0, \tag{3.1}$$

where $U_s = F(X_s)$, see, e.g., [10, 29].

The following set of conditions will be imposed to establish asymptotic normality of the estimators.

Condition 3.1.

(i) The stationary time series $(X_s)_{s\in\mathbb{N}}$ has an extremal index $\theta \in (0,1]$ and the above assumptions guaranteeing convergence of the one- and two-level point process of exceedances are satisfied.

(ii) There exists $\delta > 0$ such that, for any $m > 0$, there exists a constant $\tilde{C}_m$ such that, for all $0 \leq x_1 \leq x_2 \leq m, n \in \mathbb{N}$,

$$\mathrm{E}\left[|N_n^{(x_1)}(E) - N_n^{(x_2)}(E)|^{2+\delta}\right] \leq \tilde{C}_m(x_2 - x_1).$$

(iii) There exist constants $c_2 \in (0,1)$ and $C_2 > 0$ such that

$$\alpha_{c_2}(m) \leq C_2 m^{-\eta}$$

for some $\eta \geq 3(2+\delta)/(\delta-\mu) > 3$, where $0 < \mu < \min(\delta, 1/2)$ and $\delta > 0$ is from Condition (ii). The block size b_n converges to infinity and satisfies

$$k_n = o(b_n^2), \quad n \to \infty.$$

Further, there exists a sequence $\ell_n \to \infty$ with $\ell_n = o(b_n^{2/(2+\delta)})$ and $k_n \alpha_{c_2}(\ell_n) = o(1)$ as $n \to \infty$.

(iv) There exist constants $c_1 \in (0,1)$ and $C_1 > 0$ such that, for any $y \in (0, c_1)$ and $n \in \mathbb{N}$,

$$\mathrm{Var}\left\{\sum_{s=1}^{n} \mathbb{1}(U_s > 1 - y)\right\} \leq C_1(ny + n^2y^2).$$

(v) For any $c \in (0,1)$, one has

$$\lim_{n\to\infty} \mathbb{P}\Big(\min_{i=1,\dots,2k_n} N'_{ni} \leq c\Big) = 0,$$

where $N'_{ni} = \max\{U_s,\ s \in [(i-1)b_n/2 + 1, \dots, ib_n/2]\}$ for $i = 1, \dots, 2k_n$.

(vi) For any $x > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(N_{m:b_n} > 1 - \frac{x}{n} \,\Big|\, U_1 \geq 1 - \frac{x}{n}\right) = 0.$$

Condition 3.2 (Integrability)**.**

(i) With $\delta > 0$ from Condition 3.1(ii), one has

$$\limsup_{n\to\infty} \mathrm{E}\left[|\log(Z_{1:n})|^{2+\delta}\right] < \infty.$$

(ii) Fix $p > 0$. With $\delta > 0$ from Condition 3.1(ii), one has

$$\limsup_{n\to\infty} \mathrm{E}\left[Z_{1:n}^{(2+\delta)/p}\right] < \infty.$$

Condition 3.3 (Bias Condition)**.** Recall $\varphi_{(\mathrm{C})}, \varphi_{(\mathrm{M})}$ and $\varphi_{(\mathrm{R}),\mathrm{p}}$ defined in Equations (CFG), (MAD) and (ROOT), respectively.

(i) As $n \to \infty$, $\mathrm{E}[\log(Z_{1:b_n})] = \varphi_{(\mathrm{C})}(\theta) + o\big(k_n^{-1/2}\big)$.
(ii) As $n \to \infty$, $\mathrm{E}[\exp(-Z_{1:b_n})] = \varphi_{(\mathrm{M})}(\theta) + o\big(k_n^{-1/2}\big)$.
(iii) Fix $p > 0$. As $n \to \infty$, $\mathrm{E}\left[Z_{1:b_n}^{1/p}\right] = \varphi_{(\mathrm{R}),\mathrm{p}}(\theta) + o\big(k_n^{-1/2}\big)$.

Condition 3.4 (Technical Condition for the CFG-type estimator)**.**

(i) For some $q > 1/2$, we have $b_n = O(k_n^q)$ as $n \to \infty$.
(ii) For some $\tau \in (0, 1/2)$, we have, as $n \to \infty$,

$$\left\{\frac{e_n(x)}{x^\tau}\right\}_{x\in[0,1]} \xrightarrow{d} \left\{\frac{e(x)}{x^\tau}\right\}_{x\in[0,1]} \qquad \text{in } D([0,1]),$$

the càglàd space of functions on $[0, 1]$, where e_n denotes the tail empirical process defined in (3.1) and where e is a centered Gaussian process with continuous sample paths and covariance as given in Lemma B.1.

(iii) For any $c > 0$, we have, as $n \to \infty$,

$$\max_{Z_{ni}\geq c} \left|\frac{e_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}\right| = o_{\mathbb{P}}(1).$$

(iv) For any $c > 0$, there exists $\mu = \mu_c \in (1/2, 1/\{2(1-\tau)\})$ with τ from (ii) such that, as $n \to \infty$,

$$\mathbb{P}(Z_{n1} < ck_n^{-\mu}) - \mathbb{P}(\xi < ck_n^{-\mu}) = o\big(\log(n)^{-1}k_n^{-1/2}\big), \quad \text{where } \xi \sim \mathrm{Exp}(\theta).$$

The items of Condition 3.1 are the same as Condition 2.1(i)-(v) and (2.2) in [3] and are discussed in great detail in that reference. Condition 3.2 is needed for uniform integrability of the sequences $Z_{n1}^{2/p}$ and $\log^2 Z_{n1}$, respectively. It implies

$$\lim_{n\to\infty} \mathrm{Var}(Z_{n1}^{1/p}) = \mathrm{Var}(\xi^{1/p}), \qquad \lim_{n\to\infty} \mathrm{Var}(\log Z_{n1}) = \mathrm{Var}(\log \xi),$$

respectively, where ξ denotes an exponentially distributed random variable with parameter θ. Condition 3.3 is a bias condition requiring the approximation of the first moment of $f(Z_{n1})$ by $\mathrm{E}[f(\xi)]$ to be sufficiently accurate, where $f(x) \in \{x^{1/p}, \exp(-x), \log x\}$.

Condition 3.4 is a technical condition which is only needed for deriving the asymptotics of the CFG-estimator. The Condition 3.4(i) requires b to be not too large. Sufficient conditions for Condition 3.4(ii) in terms of beta mixing coefficients can be found in [10]. A sufficient condition for Condition 3.4(iii) is for instance strong mixing with polynomial rate $\alpha_1(n) = O(n^{-(1+\sqrt{2})-\varepsilon}), n \to \infty$, for some $\varepsilon > 0$, together with Condition 3.4(i) being met with $q < 1/(\sqrt{2}-1) \approx 2.41$. Indeed, for any $x \geq c$ and $\eta > 0$, one can write

$$\frac{e_n(x)}{x} = \frac{1}{\sqrt{k_n}} \sum_{s=1}^{n} \Big\{ \mathbb{1}\Big(U_s > 1 - \frac{x}{b_n}\Big) - \frac{x}{b_n} \Big\} \frac{1}{x} = -b_n^{1/2-\eta}\, \mathbb{U}_{n,\eta}\Big(1 - \frac{x}{b_n}\Big) \frac{1}{x^{1-\eta}},$$

where

$$\mathbb{U}_{n,\eta}(u) = \frac{\frac{1}{\sqrt{n}} \sum_{s=1}^{n} \{\mathbb{1}(U_s \leq u) - u\}}{(1-u)^{\eta}} \mathbb{1}_{(0,1)}(u).$$

By Theorem 2.2 in [30], we have $\sup_{x \geq 0} |\mathbb{U}_{n,\eta}(1 - x/b_n)| = O_{\mathbb{P}}(1)$ for all $\eta \leq 1 - 2^{-1/2} \approx 0.29$. Hence, by Condition 3.4(i),

$$\max_{Z_{ni} \geq c} \left| \frac{e_n(Z_{ni})}{Z_{ni}\sqrt{k_n}} \right| = O_{\mathbb{P}}\Big(\frac{b_n^{1/2-\eta}}{\sqrt{k_n}} \Big) = O_{\mathbb{P}}\Big(k_n^{q(1/2-\eta)-1/2} \Big).$$

The expression on the right-hand side is $o_{\mathbb{P}}(1)$ if we choose $\eta \in (1/2 - 1/\{2q\}, 1 - 2^{-1/2}]$; note that the latter interval is non-empty since $q < 1/(\sqrt{2}-1)$. Finally, Condition 3.4(iv) is another technical condition requiring the approximation of the law of Z_{n1} by the exponential distribution to be sufficiently accurate in the lower tail.

4. Asymptotic results

We present asymptotic results on all estimators defined in Section 2. For simplicity, all results are stated and proved for the $\hat{Z}_{ni}$-versions only. As in Theorem 3.1 in [3], it may be verified that the respective versions based on $\hat{Y}_{ni}$ show the same asymptotic behavior as the $\hat{Z}_{ni}$-versions. Throughout, for $z \in (0,1)$, let $(\xi_1^{(z)}, \xi_2^{(z)}) \sim \pi_2^{(z)}$.

Theorem 4.1. *Under Condition 3.1, 3.2(i), 3.3(i) and 3.4, we have*

$$\sqrt{k_n}(\hat{\theta}^{z_n}_{\mathrm{m,CFG}} - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{m,C}})$$

for $\mathrm{m} \in \{\mathrm{db}, \mathrm{sb}\}$ *and as* $n \to \infty$*, where*

$$\sigma^2_{\mathrm{db,C}} = 2\theta^3 \int_0^1 \frac{\theta\, \mathrm{E}[\xi_1^{(z)}\xi_2^{(z)}] - \mathrm{E}[\xi_1^{(z)} \mathbb{1}(\xi_2^{(z)} > 0)]}{z(1+z)} \,\mathrm{d}z + \{\pi^2/6 - 2\log(2)\}\theta^2,$$
$$\sigma^2_{\mathrm{sb,C}} = \sigma^2_{\mathrm{db,C}} - \{\pi^2/6 - 8\log(2) + 4\}\theta^2.$$

Theorem 4.2. *Under Condition 3.1 and 3.3(ii), we have*

$$\sqrt{k_n}(\hat\theta^{z_n}_{\mathrm{m,MAD}} - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{m,\mathrm{M}})$$

for $\mathrm{m} \in \{\mathrm{db}, \mathrm{sb}\}$ *and as* $n \to \infty$*, where*

$$\sigma^2_{\mathrm{db,M}} = 4\theta^2(1+\theta)\int_0^1 \frac{\theta\,\mathrm{E}[\xi_1^{(z)}\xi_2^{(z)}] - \mathrm{E}[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)} > 0)]}{(1+z)^3}\,\mathrm{d}z + \frac{\theta^2(1+\theta)}{2(2+\theta)}$$
$$\sigma^2_{\mathrm{sb,M}} = \sigma^2_{\mathrm{db,M}} - \frac{3\theta^2 + 4\theta - 4(1+\theta)(2+\theta)\log\{2(1+\theta)/(2+\theta)\}}{\theta(2+\theta)(1+\theta)^2}.$$

Theorem 4.3. *Fix* $p > 0$*. Under Condition 3.1, 3.2(ii) and 3.3(iii),*

$$\sqrt{k_n}(\hat\theta^{z_n}_{\mathrm{m,R},p} - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{m},p})$$

for $\mathrm{m} \in \{\mathrm{db}, \mathrm{sb}\}$ *and as* $n \to \infty$*, where*

$$\sigma^2_{\mathrm{db},p} = \frac{4p\theta^3}{\mathrm{B}(1/p,1/p)}\int_0^1 \frac{\theta\,\mathrm{E}[\xi_1^{(z)}\xi_2^{(z)}] + \mathrm{E}[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)} = 0)]z^{\frac{1}{p}-1}}{(1+z)^{1+\frac{2}{p}}}\,\mathrm{d}z + \Big\{\frac{2p^3}{\mathrm{B}(1/p,1/p)} - p^2 - 2p\Big\}\theta^2,$$
$$\sigma^2_{\mathrm{sb},p} = \sigma^2_{\mathrm{db},p} - \Big[p^2 + \frac{2p^3}{\mathrm{B}(1/p,1/p)} - \frac{4p}{\Gamma(1/p)^2}\int_0^\infty (1-e^{-z})z^{1/p-2}\Gamma(1/p,z)\,\mathrm{d}z\Big]\theta^2,$$

where $\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t$ *denotes the beta function and* $\Gamma(x,z) = \int_x^\infty t^{z-1}e^{-t}\,\mathrm{d}t$ *is the incomplete gamma function.*

It is worthwhile to mention that the imposed conditions in each theorem are exactly the same for the disjoint and the sliding blocks version. Furthermore, apart from the different bias conditions, the conditions regarding k_n are exactly the same in Theorem 4.2 and 4.3, and slightly stronger for Theorem 4.1 in that the additional technical Condition 3.4 is imposed.

The proofs are provided in Section A and bear some similarities with the one of Theorem 3.2 in [3]. In particular, they rely on the delta method, Wichura's theorem and empirical process theory to adequately handle the asymptotic contribution of the rank transformation. The most sophisticated proof is the one of Theorem 4.1, which is essentially due to the fact that $\mathrm{E}[\log\xi] = \int_0^\infty \log(t)\theta e^{-\theta t}\,\mathrm{d}t$ is an improper integral both at zero and at infinity (see also [14] for similar technical difficulties with the CFG-estimator for the Pickands dependence function in multivariate extremes).

It is worth to mention that the difference

$$\mathrm{AsyVar}(\sqrt{k_n}\hat\theta^{z_n}_{\mathrm{db,CFG}}/\theta) - \mathrm{AsyVar}(\sqrt{k_n}\hat\theta^{z_n}_{\mathrm{sb,CFG}}/\theta) = (\sigma^2_{\mathrm{db,C}} - \sigma^2_{\mathrm{sb,C}})/\theta^2 \approx 0.0977$$

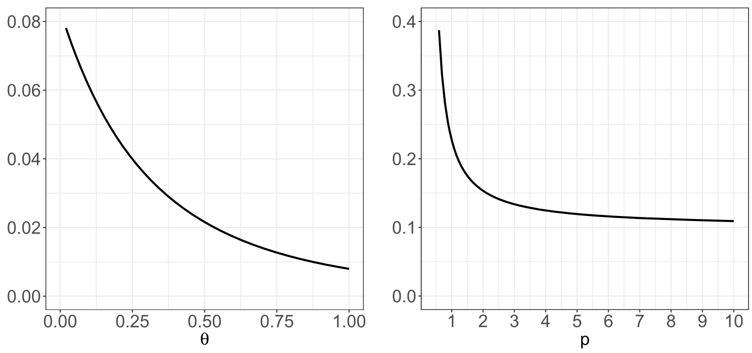

FIG 1. *Graph of the functions* $\theta \mapsto (\sigma^2_{\mathrm{db,M}} - \sigma^2_{\mathrm{sb,M}})/\theta^2$ *(left) and* $p \mapsto (\sigma^2_{\mathrm{db,p}} - \sigma^2_{\mathrm{sb,p}})/\theta^2$ *(right).*

is a universal constant independent of any properties of the observed time series. The same holds true for the Root-estimator with a constant depending in a complicated way on the parameter p (the graph of $p \mapsto (\sigma^2_{\mathrm{db},p} - \sigma^2_{\mathrm{sb},p})/\theta^2$ is depicted in Figure 1, with a value of approximately 0.2274 for the PML-estimator). For the Madogram-estimator, this difference depends on θ (see Figure 1 for the graph of $\theta \mapsto (\sigma^2_{\mathrm{db,M}} - \sigma^2_{\mathrm{sb,M}})/\theta^2$); it is non-negative and decreasing with value $1/12 \approx 0.083$ for $\theta \to 0$ and approximately 0.0079 for $\theta = 1$. In that regard, the use of sliding blocks over disjoint blocks is least beneficial for the Madogram-estimator.

Example 4.4. In the case that the time series is serially independent, the cluster size distributions are given by $\pi(i) = \mathbb{1}(i = 1)$ and $\pi_2^{(z)}(i, j) = (1 - z)\mathbb{1}(i = 1, j = 0) + z\mathbb{1}(i = 1, j = 1)$, which implies

$$\theta = 1, \quad E[\xi_1^{(z)}\xi_2^{(z)}] = z \quad \text{and} \quad E[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)} = 0)] = 1 - z.$$

It can be seen that these formulas hold true whenever $\theta = 1$. Consequently, the limiting variances in Theorem 4.1 and 4.2 are equal to

$$\sigma^2_{\mathrm{db,C}} = \frac{\pi^2}{6} - 2\log(2) \approx 0.2586, \qquad \sigma^2_{\mathrm{sb,C}} = 6\log(2) - 4 \approx 0.1588,$$
$$\sigma^2_{\mathrm{db,M}} = 1/3, \qquad \sigma^2_{\mathrm{sb,M}} \approx 0.32536.$$

It is remarkable that the asymptotic variances are substantially smaller than those of the maximum likelihood estimator, see Example 3.1 in [3], which are equal to 1/2 and 0.2726 for the disjoint and sliding blocks version, respectively.

The limiting variance in the case of the Root-estimator is given by

$$\sigma^2_{\mathrm{db},p} = \frac{2p}{\mathrm{B}(\frac{1}{p}, \frac{1}{p})}\Big[p^2 + 2^{-2/p}p\Big] - p^2 - p,$$
$$\sigma^2_{\mathrm{sb},p} = \sigma^2_{\mathrm{db},p} - \Big[p^2 + \frac{2p^3}{\mathrm{B}(\frac{1}{p}, \frac{1}{p})} - \frac{4p}{\Gamma(\frac{1}{p})^2}\int_0^\infty (1 - e^{-z})z^{1/p-2}\Gamma(\tfrac{1}{p}, z)\,\mathrm{d}z\Big].$$

Some values are

$$\sigma^2_{\mathrm{db},1/2} = \frac{15}{16}, \qquad \sigma^2_{\mathrm{db},1} = \frac{1}{2}, \qquad \sigma^2_{\mathrm{db},2} \approx 0.3662,$$
$$\sigma^2_{\mathrm{sb},1/2} = \frac{7}{16}, \qquad \sigma^2_{\mathrm{sb},1} \approx 0.2726, \qquad \sigma^2_{\mathrm{sb},2} \approx 0.212909.$$

It can further be shown that $\lim_{p\to\infty} \sigma^2_{\mathrm{m},p} = \sigma^2_{\mathrm{m,C}}$ for $\mathrm{m} \in \{\mathrm{db}, \mathrm{sb}\}$.

Remark 4.5. Instead of working with $\hat{F}_n$ in the definition of $\hat{Z}_{ni} = b_n\{1 - \hat{F}_n(M_{ni})\}$, one may alternatively use the empirical c.d.f. of $(X_s)_{s\notin I_i}$ multiplied by $(n - b_n)/(n - b_n + 1)$ for $I_i = \{(i-1)b_n + 1, \ldots, ib_n\}$, denoted by $\hat{F}_{n,-i}$, and define $\tilde{Z}_{ni} = b_n\{1 - \hat{F}_{n,-i}(M_{ni})\}$ and $\tilde{\theta} = \hat{\theta}(\tilde{Z}_{n1}, \ldots, \tilde{Z}_{nk_n})$. This modification has been motivated as a bias reduction scheme in [22]. Since

$$\tilde{Z}_{ni} = b_n\{1 - \hat{F}_{n,-i}(M_{ni})\} = b_n\{1 - \hat{F}_n(M_{ni})\}\frac{n+1}{n - b_n + 1} = \hat{Z}_{ni}\frac{n+1}{n - b_n + 1},$$

some simple calculations show that, for instance for the CFG-estimator,

$$e^{-\gamma} \exp\Big\{ - \frac{1}{k_n}\sum_{i=1}^{k_n} \log(\tilde{Z}_{ni})\Big\} = \frac{n - b_n + 1}{n+1}\hat{\theta}^{z_n}_{\mathrm{db,CFG}},$$

showing that the modification is asymptotically negligible. It is however beneficial in finite-sample situations, whence it has been applied throughout the finite-sample situations considered in Section 6. Obviously, similar adaptions can be applied to the sliding blocks version and the other moment based estimators.

5. Example: max-autoregressive process

In this section, we exemplarily discuss the new estimators when applied to a max-autoregressive process, defined by the recursion

$$X_s = \max\{\alpha X_{s-1}, (1-\alpha)Z_s\}, \quad s \in \mathbb{Z},$$

where $\alpha \in [0, 1)$ and where $(Z_s)_{s\in\mathbb{Z}}$ is an i.i.d. sequence of Fréchet(1)-distributed random variables. A stationary solution of the above recursion is

$$X_s = \max_{j\geq 0} (1-\alpha)\alpha^j Z_{s-j},$$

such that the stationary solution is again Fréchet(1)-distributed. Note that a model with an arbitrary stationary c.d.f. F may be obtained by considering $\tilde{X}_s = F^{\leftarrow}\{\exp(-1/X_s)\}$ and that all subsequent results are also valid for $(\tilde{X}_s)_s$.

We start by explicitly calculating the asymptotic variances of the estimators in Section 5.1, and then show in Section 5.2 that all regularity conditions from Section 3 are met.

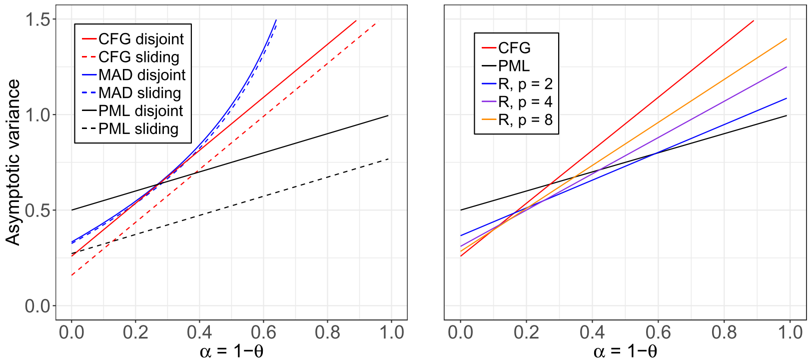

FIG 2. *Asymptotic variance of $\sqrt{k_n}(\hat\theta_n/\theta - 1)$ in the ARMAX-model. The estimators in the right figure rely on disjoint blocks.*

5.1. Asymptotic variances for the ARMAX-model

Recall that the ARMAX-model has extremal index $\theta = 1 - \alpha$ and that the corresponding cluster size distribution is geometric, that is, $\pi(j) = \alpha^{j-1}(1 - \alpha), j \geq 1$, see, e.g., Chapter 10 in [1]. From Example 6.1 in [3], one further has

$$\begin{aligned} \mathrm{E}[\xi_1^{(z)}\xi_2^{(z)}] &= \frac{\alpha^{w+1} + z + zw(1-\alpha)}{(1-\alpha)^2}, \\ \mathrm{E}[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)} = 0)] &= \frac{1 - \alpha^{w+1}}{1-\alpha} - z(w+1), \end{aligned}$$

where $w = \lfloor \log(z)/\log(\alpha) \rfloor$ and $(\xi_1^{(z)}, \xi_2^{(z)}) \sim \pi_2^{(z)}$. This allows to calculate the limiting variances in Theorem 4.1–4.3 explicitly. For the CFG-type estimator, some tedious but straightforward calculations imply

$$\frac{\sigma^2_{\mathrm{db,C}}}{\theta^2} = \frac{\pi^2}{6} + 2\log(2)(\alpha - 1) \quad \text{and} \quad \frac{\sigma^2_{\mathrm{sb,C}}}{\theta^2} = 2\log(2)(3+\alpha) - 4,$$

see also Figure 2 for a picture of the graph of these functions. Next, we compare these variances with the disjoint and sliding blocks variances of the PML-estimator in [3], which are given by $\sigma^2_{\mathrm{db},1}$ and $\sigma^2_{\mathrm{sb},1}$ and satisfy

$$\frac{\sigma^2_{\mathrm{db},1}}{\theta^2} = \frac{1}{2}(1+\alpha) \quad \text{and} \quad \frac{\sigma^2_{\mathrm{sb},1}}{\theta^2} = \frac{8\log(2) - 5 + \alpha}{2},$$

respectively. Thus, $\sigma^2_{\mathrm{db,C}} \leq \sigma^2_{\mathrm{db},1}$ iff $\alpha \leq \{1 + 4\log(2) - \pi^2/3\}/\{4\log(2) - 1\} \approx 0.2723$ and $\sigma^2_{\mathrm{sb,C}} \leq \sigma^2_{\mathrm{sb},1}$ iff $\alpha \leq \{3 - 4\log(2)\}/\{4\log(2) - 1\} \approx 0.128$.

Further comparisons can be drawn from Figure 2, where the asymptotic variances of $\sqrt{k_n}(\hat\theta_n/\theta - 1)$ are additionally illustrated for the Madogram- and the Root-estimators.

5.2. Regularity conditions for the ARMAX-model

Recall that X_s is Fréchet(1)-distributed, i.e., the stationary c.d.f. F is given by $F(x) = \exp(-1/x), x > 0$, with inverse $F^{-1}(x) = -\log(x)^{-1}$.

The assumptions in Condition 3.1 are satisfied as shown in [3], page 2322, provided b_n and k_n are chosen to satisfy the conditions in item (iii). Next, by induction,

$$\mathbb{P}\big(\max_{s=1,\dots,b} X_s \le x\big) = F(x)^{1+\theta(b-1)},$$

which implies that the c.d.f. of $Z_{1:b} = b\{1 - F(M_{1:b})\}$ is given by

$$\mathbb{P}(Z_{1:b} \le x) = \begin{cases} 1, & x \ge b, \\ 1 - \big(1 - \frac{x}{b}\big)^{1+\theta(b-1)}, & x \in [0, b], \\ 0, & b \le 0. \end{cases} \tag{5.1}$$

A tedious but straightforward calculation then shows that the assumptions in Condition 3.2 and 3.3 are met, provided $k_n/b_n^2 = o(1)$, cf. Condition 3.1(iii). Condition 3.4(i) is a condition on the choice of b_n, that is under the control of the statistician. Conditions 3.4(ii) and 3.4(iii) are consequences of mixing properties of $(X_s)_s$ as argued at the end of Section 3. It remains to show that Condition 3.4(iv) is satisfied. By (5.1) and with $\xi \sim \text{Exp}(\theta)$, we have

$$\begin{aligned} \mathbb{P}(Z_{n1} < ck_n^{-\mu}) - \mathbb{P}(\xi < ck_n^{-\mu}) &= \exp(-\theta ck_n^{-\mu}) - \Big(1 - \frac{ck_n^{-\mu}}{b_n}\Big)^{1+\theta(b_n-1)} \\ &= o(k_n^{-1/2}(\log n)^{-1}), \qquad n \to \infty, \end{aligned}$$

for any $\mu > 1/2$, where the final estimate follows from Taylor's theorem and Condition 3.4(i).

6. Finite-sample results

A Monte-Carlo simulation study was performed to assess the finite-sample performance of the introduced estimators and to compare them with competing estimators from the literature. The data is simulated from the following four time series models that were also investigated in [3]:

- The **ARMAX-model**:

  $$X_s = \max\{\alpha X_{s-1}, (1-\alpha)Z_s\}, \qquad s \in \mathbb{Z},$$

  where $\alpha \in [0, 1)$ and where $(Z_s)_s$ is an i.i.d. sequence of standard Fréchet random variables. We consider $\alpha = 0, 0.25, 0.5, 0.75$ resulting in $\theta = 1, 0.75, 0.5, 0.25$.
- The **squared ARCH-model**:

  $$X_s = (2 \times 10^{-5} + \lambda X_{s-1})Z_s^2, \qquad s \in \mathbb{Z},$$

where $\lambda \in (0,1)$ and where $(Z_s)_s$ denotes an i.i.d. sequence of standard normal random variables. We consider $\lambda = 0.1, 0.5, 0.9, 0.99$ for which the simulated values $\theta = 0.997, 0.727, 0.460, 0.422$ were obtained, respectively; see Table 3.1 in [8].

- The **ARCH-model**:

$$X_s = (2 \times 10^{-5} + \lambda X_{s-1}^2)^{1/2} Z_s, \qquad s \in \mathbb{Z},$$

where $\lambda \in (0,1)$ and where $(Z_s)_s$ denotes an i.i.d. sequence of standard normal random variables. We consider $\lambda = 0.1, 0.5, 0.7, 0.99$ for which the simulated values $\theta = 0.999, 0.835, 0.721, 0.571$ were obtained, respectively; see Table 3.2 in [8].

- The **Markovian Copula-model** ([7]):

$$X_s = F^{\leftarrow}(U_s), \quad (U_s, U_{s-1}) \sim C_\vartheta, \qquad s \in \mathbb{Z}.$$

Here, $F^{\leftarrow}$ is the left-continuous quantile function of some arbitrary continuous c.d.f. F, $(U_s)_s$ is a stationary Markovian time series of order 1 and C_ϑ denotes the Survival Clayton Copula with parameter $\vartheta > 0$. We consider choices $\vartheta = 0.23, 0.41, 0.68, 1.06, 1.90$ such that (approximately) $\theta = 0.95, 0.8, 0.6, 0.4, 0.2$ [3] and fix F as the standard uniform c.d.f. (the results are independent of this choice, as the estimators are rank-based). Algorithm 2 in [26] allows to simulate from this model.

In each case, the sample size is fixed to $n = 2^{13} = 8192$ and the block size is chosen from $b = b_n \in \{2^2, \dots, 2^9\}$. The performance is assessed based on $N = 3000$ simulation runs each.

6.1. *Comparison of the introduced estimators*

We start by comparing the finite-sample properties of the proposed sliding blocks estimators $\hat\theta^x_{\mathrm{m,CFG}}$, $\hat\theta^x_{\mathrm{m,MAD}}$ and $\hat\theta^x_{\mathrm{m,R,p}}$ for $p \in \{0.5, 0.75, 1, 2, 4, 8, 16\}$ for $x \in \{z_n, y_n\}$ and for m $\in$\{sb, db\}.

As the simulation results are, to a large extent, similar among the different models and estimators, they are only partially reported, with a particular view on highlighting selected interesting qualitative features. We begin by a detailed investigation of the variance, the squared bias and the mean squared error (MSE) as a function of the block size parameter b. In Figure 3, we present results for the disjoint and sliding blocks version of the CFG- and the PML-estimator in a representative ARMAX-model with $\theta = 0.75$. Similarly as in [3] and as to be expected from the asymptotic results, the bias of the disjoint and the sliding blocks version are almost identical, while the variance is uniformly smaller for the sliding blocks version (in particular for large values of b_n). Since this qualitative behavior holds uniformly over all models and estimators, we omit the disjoint blocks estimator from the subsequent discussions and write $\hat\theta^x_{\mathrm{CFG}} = \hat\theta^x_{\mathrm{sb,CFG}}$ etc. for simplicity.

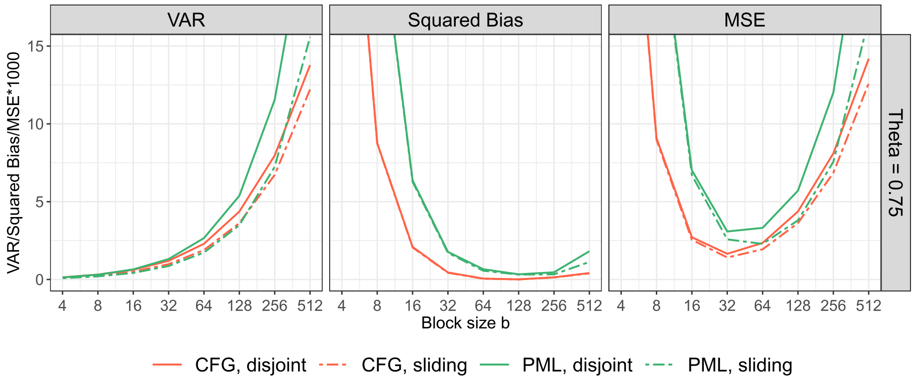

FIG 3. *Comparison of variance, squared bias and MSE, multiplied by* 10^3*, of the disjoint and sliding blocks CFG- and PML-estimator in the ARMAX-model.*

Next, we compare the different moment estimators. For illustrative purposes, we begin by restricting the presentation to the z_n-versions and the ARCH-model. The corresponding results are depicted in Figure 4 (for the CFG-, the Madogram- and three selected Root-estimators). In general, as to be expected from the underlying theory, the variance curves are increasing in b, while the squared bias curves are (mostly) decreasing in b, resulting in a typical U-shape for the MSE curves. The hierarchy of the estimators with regard to the considered performance measures is similar among the considered values of θ. In terms of the MSE, up to an intermediate block size, the CFG- and Madogram-estimator are superior to the other estimators (especially to the PML-estimator), while for large block sizes the Madogram-estimator has a relatively high MSE, but the CFG-estimator partly remains superior. The Root-estimators are, as expected, ordered in p and located between the PML- and CFG-estimator.

Next, a comparison between the z_n- and y_n-versions of the estimators is drawn in Figure 5; for illustrative purposes, attention is restricted to six different models and two estimators. Remarkably, there are many models, especially for smaller values of θ, in which the MSE-curves of the y_n-versions lie uniformly below the ones of the z_n-versions. In the remaining models, neither version can be said to be strictly preferable. Furthermore, it is remarkable that, for θ close to one, the MSE-curves of the y_n-versions are often no longer U-shaped, but increasing in the block size instead. The latter behavior may be explained by the proximity to the i.i.d. case, since in that case, we have

$$\mathbb{P}(Y_{1:b} \geq y) = \mathbb{P}(N_{1:b} \leq e^{-y/b}) = \mathbb{P}(U_1 \leq e^{-y/b})^b = e^{-y}$$

for all $b \in \mathbb{N}$, such that there is real equality in relation (1.3), resulting in a vanishing bias.

Next, we investigate the dependence of the performance of the Root-estimators on the parameter p; recall that $p = 1$ yields the PML-estimator, while '$p = \infty$' yields the CFG-estimator. In Figure 6, the MSE-curves are depicted

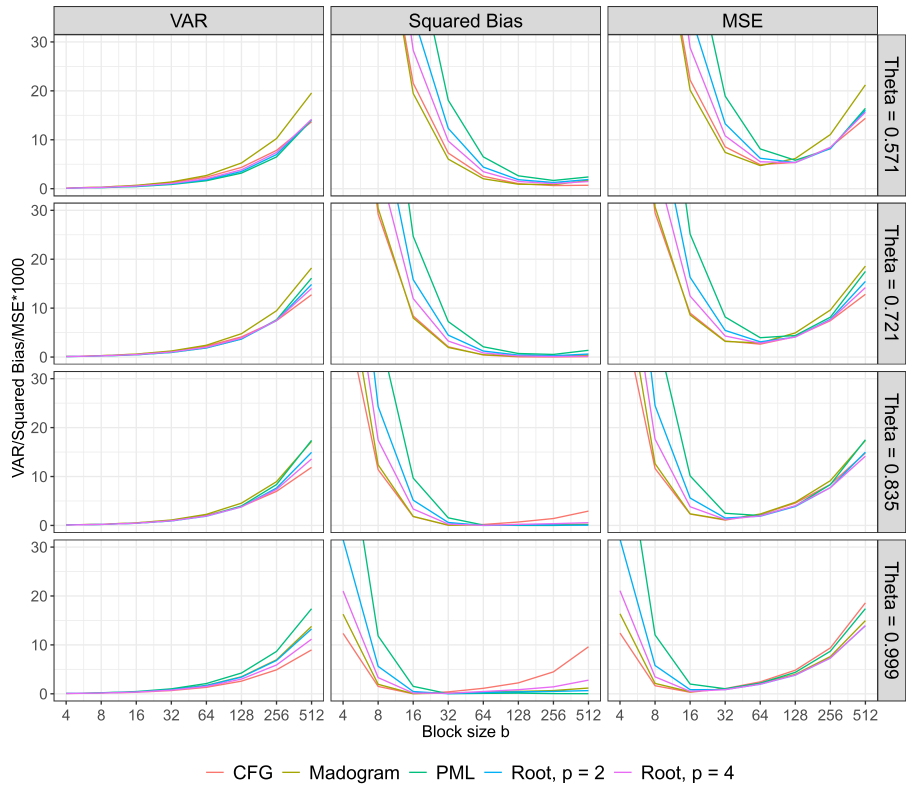

FIG 4. *Variance, squared bias and MSE, multiplied by 10^3, for the estimation of θ within the ARCH-model for four values of θ.*

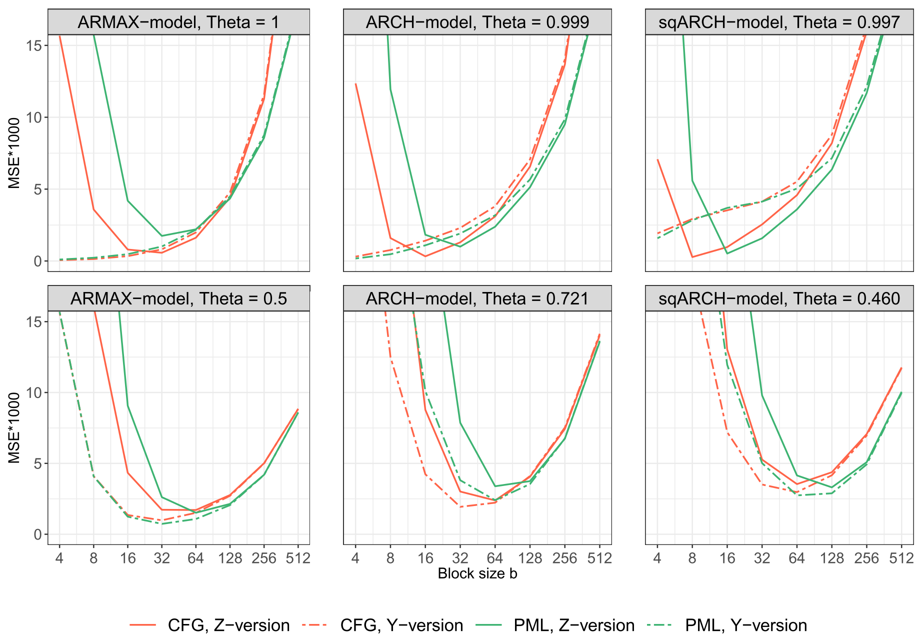

FIG 5. *Comparison of the MSE multiplied by 10^3 of the z_n- and y_n-versions of the estimators.*

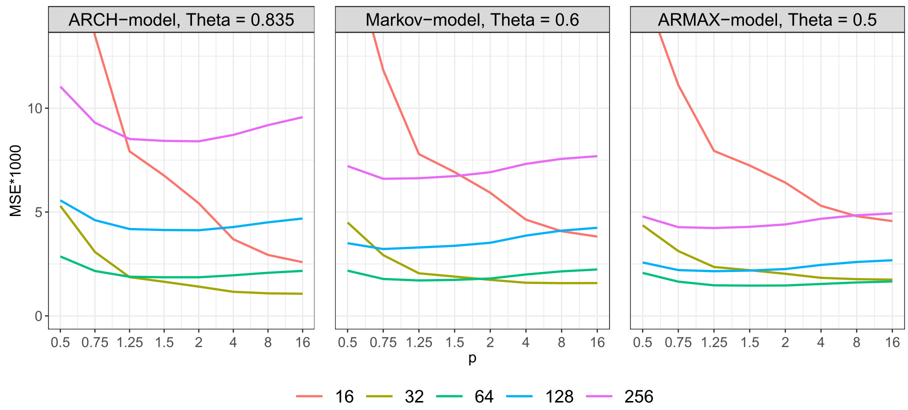

FIG 6. *Mean Squared Error multiplied by* 10^3 *of the Root-estimators as a function of the parameter* p *for block sizes* $b \in \{16, 32, 64, 128, 256\}$ *and three different models.*

TABLE 1
Identification of the Root-estimator p *with the minimum MSE for the ARCH- and ARMAX-model and every considered block size* b*. The* p *with the minimum MSE over all blocksizes is presented in the last line.*

Model	ARCH				ARMAX			
Theta	0.999	0.835	0.721	0.571	1	0.75	0.5	0.25
$b = 4$	∞	∞	∞	∞	∞	∞	∞	∞
8	∞	∞	∞	∞	∞	∞	∞	∞
16	∞	∞	∞	∞	∞	∞	∞	∞
32	2	∞	∞	∞	∞	∞	∞	∞
64	2	2	∞	∞	16	8	1.5	2
128	2	1.5	4	4	8	4	1	1
256	2	4	∞	1.25	4	8	1	0.75
512	2	8	∞	∞	4	∞	1	0.75
$\min_b$	∞	∞	∞	∞	∞	∞	1.5	1

as a function of p for various fixed block sizes and for three selected models. It can be seen that choices of $p < 1$ lead to a poor behavior of the corresponding estimators. At the same time, the results do not allow to identify some 'optimal' choice of $p \geq 1$ which is valid uniformly over all models. A similar conclusion can be drawn from Table 1, which presents, for the ARCH- and ARMAX-model and every block size b, the value of p for which the Root-estimator attains the minimal MSE ($p = \infty$ corresponds to the CFG-estimator). One can see that most values of p are represented, with $p = \infty$ appearing most often, but that there is no optimal choice of p universally over all models.

6.2. Comparison with other estimators for the extremal index

In this section, we compare the performance of the introduced new estimators with the following estimators: the bias-reduced sliding blocks estimator from [28] (with a data-driven choice of the threshold as outlined in Section 7.1 of that paper), the integrated version of the blocks estimator from [27], the intervals

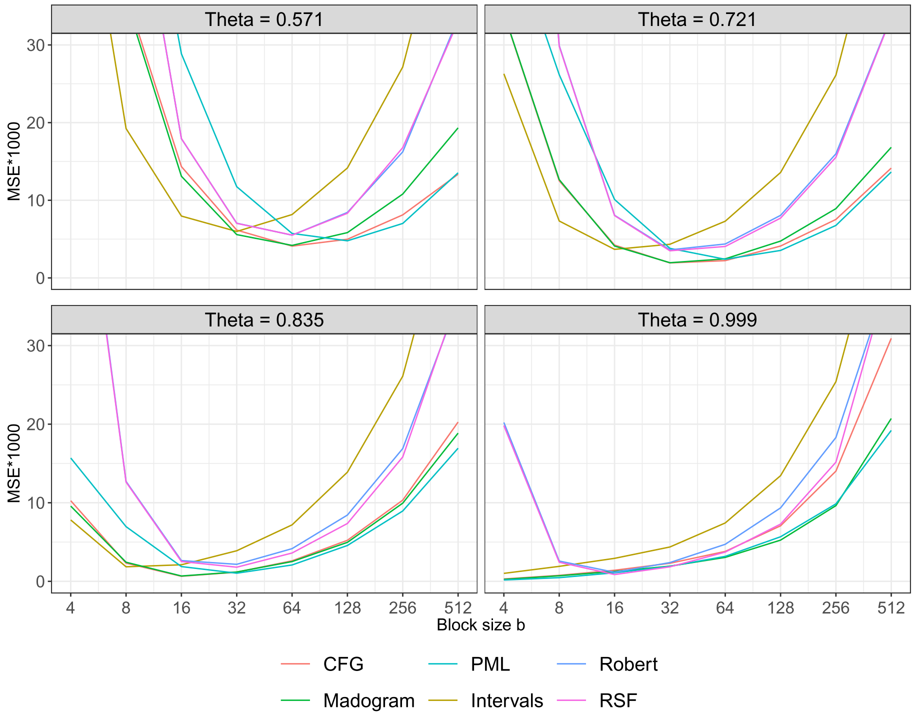

Fig 7. *Mean Squared Error multiplied by* 10^3 *for the estimation of* θ *within the ARCH-model for four values of* θ*.*

estimator from [13] and the ML-estimator from [32]. The parameters σ and ϕ for the Robert-estimator (cf. page 276 of 27) are chosen as $\sigma = 0.7$ and $\phi = 1.3$. In the case of the intervals- and Süveges-estimator, the choice of a threshold u is required, which is here chosen as the $1 - 1/b_n$ empirical quantile of the observed data. With regard to our estimators, we present results for the sliding-blocks, bias-reduced and z_n-versions, if not indicated otherwise.

In Figure 7, we depict the MSE as a function of the block size b. For most models, the MSE-curves of the estimators from the literature are again U-shaped due to the bias-variance tradeoff already described in Section 6.1. It can further be seen that no estimator is uniformly best in any model under consideration. The method-of-moment estimators do however compare quite well to the competitors.

The minimum values of the MSE-curves in Figure 7 are of particular interest. Due to the large amount of estimators and models under consideration (in total 26 estimators and 17 models) we try to simplify possible comparisons by the following aggregation, summarized in Table 2. First, in the first four columns of the table, we calculate for each time series model and each estimator under consideration, the sum (sum over all values of θ considered for the specific model) of the minimum MSE-values (minimum over b). Second, in the last four columns of the table, we present the sum of the minimum MSE-values (minimum over b) over all models, for which the extremal index θ lies in the in-

TABLE 2
Sum of minimal Mean Squared Error multiplied by 10^3 *over different models and* $\theta_1 \in (0, 0.3], \theta_2 \in (0.3, 0.6], \theta_3 \in (0.6, 0.8]$ *and* $\theta_4 \in (0.8, 1]$*. The three smallest values per column are in boldface.*

Estimator	armax	arch	arch^2	markov	(0, .3]	(.3, .6]	(.6, .8]	(.8, 1]
CFG, Z	4.80	8.54	8.46	11.19	5.84	19.08	5.46	**2.61**
CFG, Y	2.56	**6.98**	8.41	12.63	5.08	15.45	**3.56**	6.49
Madogram, Z	5.17	8.87	7.92	10.77	5.66	18.12	5.68	3.27
Madogram, Y	3.00	**7.08**	8.62	12.65	5.10	15.72	**3.59**	6.94
PML, Z	6.18	11.74	7.99	10.89	4.44	18.62	7.37	6.38
PML, Y	**1.96**	8.40	7.45	10.99	**3.73**	14.83	4.04	6.21
R, p = 0.5, Z	9.64	17.37	11.57	12.11	4.90	24.18	11.25	10.35
R, p = 0.5, Y	2.33	11.99	8.49	10.94	3.90	18.14	5.66	6.05
R, p = 0.75, Z	7.08	13.33	8.83	10.99	4.44	19.80	8.79	7.20
R, p = 0.75, Y	2.03	9.26	7.63	10.74	**3.66**	15.53	4.41	6.06
R, p = 1.25, Z	5.77	11.02	7.82	10.80	4.56	18.33	6.61	5.89
R, p = 1.25, Y	**1.96**	8.06	**7.37**	11.04	3.74	**14.47**	3.90	6.32
R, p = 1.5, Z	5.54	10.48	7.86	**10.47**	4.72	18.38	6.21	5.04
R, p = 1.5, Y	1.98	7.93	**7.32**	11.10	3.76	**14.32**	3.84	6.40
R, p = 2, Z	5.22	9.82	8.11	**10.22**	4.84	18.67	5.76	4.10
R, p = 2, Y	2.03	7.88	**7.34**	11.16	3.84	**14.34**	3.72	6.51
R, p = 4, Z	4.84	9.10	8.40	**10.14**	5.07	18.81	5.39	3.20
R, p = 4, Y	2.20	7.52	7.64	11.58	4.21	14.53	3.67	6.52
R, p = 8, Z	4.76	8.88	8.42	10.48	5.37	18.96	5.36	**2.85**
R, p = 8, Y	2.35	7.31	7.95	12.02	4.56	14.91	3.68	6.48
R, p = 16, Z	4.76	8.69	8.41	10.78	5.58	18.99	5.39	**2.68**
R, p = 16, Y	2.45	**7.14**	8.16	12.32	4.80	15.18	**3.61**	6.47
Intervals	3.49	12.53	11.72	21.86	**3.60**	15.55	11.46	18.98
ML Süveges	**1.90**	22.67	8.70	25.20	14.93	30.46	4.95	8.13
Robert	8.54	12.45	9.97	13.61	6.46	22.42	8.34	7.34
RSF	8.09	11.68	9.77	15.85	7.28	23.52	7.52	7.06

terval $(0, 0.3], (0.3, 0.6], (0.6, 0.8]$ or $(0.8, 1]$, respectively. It can be seen that the CFG-estimator wins thrice, the Madogram- and PML-estimator wins twice, the Süveges and the Intervals-estimator wins once, and that the remaining smallest values are covered by a version of the Root-estimator. Also note that for large values of $\theta \in (0.8, 1]$ (last column), the CFG-estimator and the Root-estimator for $p \in \{8, 16\}$ are the best performing estimators. As a final interesting observation, note that the y-versions of the moment estimators mostly outperform the z-version, except for the column corresponding to $\theta \in (0.8, 1]$ and some entries in the columns 'Markov' and 'sqARCH'. A more refined analysis showed that these differences were almost exclusively attributable to the two specific models 'Markov($\theta = 0.95$)' and 'sqARCH($\theta = 0.997$)', which appear to be rather difficult to estimate for all estimators under consideration.

7. Conclusion

Estimating the extremal index is a classical problem in extreme value analysis for univariate stationary time series, with many ad-hoc solutions based on diverse motivations. This paper considers a new approach that is based on certain rescaled samples of ranks of block maxima and the method of moment principle. The underlying samples have also been used by [22] and [3] to define explicit (pseudo) maximum likelihood estimators for the extremal index. Using the method of moment principle instead results in a large variety of alternative estimators. Studying their properties was initially motivated by the fact that a similar approach in multivariate extremes (the rank-based CFG-estimator for the Pickands function) was found to yield a more efficient estimator than the (pseudo) maximum likelihood method [14].

The method of moment principle being a rather universal principle, the present paper goes far beyond only considering a CFG-type estimator. In fact, based on natural moment equations for the exponential distribution (see Section 2.3), three classes of method of moment estimators were considered, which may each be based on (1) either disjoint or sliding block maxima, and (2) on certain y- or z-transformations of the block maxima. The sliding blocks version was always found to be more efficient than the disjoint blocks version. The y- and z-version share a similar behavior in terms of their asymptotic variances, but their bias may differ substantially depending on the underlying data generating process. The initial conjecture derived from [14] was partially confirmed: for θ in an explicit neighbourhood of 1, the asymptotic variance of the CFG-type estimator is always smaller than the one of the ML-type estimator. A comparison between the various method of moment estimators is more cumbersome, with no universal answer, neither theoretically nor in terms of simulated finite sample results. If one were to come up with a single proposal, then the simulation study overall suggests to use the sliding blocks y-version of the root-estimator with an intermediate choice of p, say, $p = 1.25$.

In comparison with many other estimators for the extremal index, the proposed estimators have the advantage of being based on only one parameter to be chosen by the statistician, namely the block size b. Moreover, the estimators perform equally well or even better in some typical finite sample situations.

Finally, this work leaves some interesting questions for future research: (1) what is the minimal asymptotic variance that can be achieved by estimators based on the considered rank-based samples? (2) More generally, are there estimators for the extremal index that are semiparametrically efficient? (3) Can the sliding blocks method be used to derive more efficient estimators for the cluster size distribution, for instance by generalizing the disjoint blocks versions in [27]?

Appendix A: Proofs of Theorems 4.1–4.3

The proofs of Theorems 4.1–4.3 are actually quite similar in that each proof will be decomposed into a sequence of similar intermediate lemmas. Occasionally, those lemmas will be hardest to prove for Theorem 4.1 and easiest to prove for Theorem 4.2; this is also reflected by the larger number of conditions required for the proof of Theorem 4.1. The proof of Theorem 4.3 in turn is quite similar to the one in [3], and of intermediate difficulty. For the above reasons, we will carry out the proof of Theorem 4.1 in great detail (Section A.1), and skip parts of the technical arguments needed for Theorem 4.2 and 4.3 where possible (Sections A.2 and A.3). Intermediate, but less central results for the proof of Theorem 4.1 are given in Sections B.1, B.2 and B.3.

All convergences are for $n \to \infty$ if not stated otherwise.

A.1. Proof of Theorem 4.1

The following notations will be used throughout:

$$\hat{S}_n = \frac{1}{k_n}\sum_{i=1}^{k_n}\log(\hat{Z}_{ni}), \qquad S_n = \frac{1}{k_n}\sum_{i=1}^{k_n}\log(Z_{ni}),$$

$$\hat{S}_n^{\mathrm{sb}} = \frac{1}{n-b_n+1}\sum_{i=1}^{n-b_n+1}\log(\hat{Z}_{ni}^{\mathrm{sb}}), \qquad S_n^{\mathrm{sb}} = \frac{1}{n-b_n+1}\sum_{i=1}^{n-b_n+1}\log(Z_{ni}^{\mathrm{sb}}).$$

Note that $\hat{\theta}^{z_n}_{\mathrm{db,CFG}} = \varphi_{(\mathrm{C})}^{-1}(\hat{S}_n)$ and $\hat{\theta}^{z_n}_{\mathrm{sb,CFG}} = \varphi_{(\mathrm{C})}^{-1}(\hat{S}_n^{\mathrm{sb}})$, where $\varphi_{(\mathrm{C})}^{-1}(x) = \exp\{-(x+\gamma)\}$. Observing that $(\varphi_{(\mathrm{C})}^{-1})'\{\varphi_{(\mathrm{C})}(\theta)\} = \theta$, the two assertions of the theorem are a consequence of the delta-method and Proposition A.1 and Proposition A.2, respectively. □

Proposition A.1. *Under Condition 3.1, 3.2(i), 3.3(i) and 3.4, we have*

$$\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{C})}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{db,C}}/\theta^2) \qquad \text{as } n \to \infty.$$

Proof. We may decompose

$$\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{C})}(\theta)\} = A_n + B_n + C_n,$$

where

$$A_n = \sqrt{k_n}\{\hat{S}_n - S_n\}, \quad B_n = \sqrt{k_n}\{S_n - \mathrm{E}(S_n)\}, \quad C_n = \sqrt{k_n}\{\mathrm{E}(S_n) - \varphi_{(\mathrm{C})}(\theta)\}.$$

We have $C_n = o(1)$ by Condition 3.3(i). For the treatment of A_n, recall the tail empirical process defined in (3.1). Further, let $\tilde{N}_{ni} = (n+1)/n \times \hat{N}_{ni}$, and note that

$$\begin{aligned}
1 - \tilde{N}_{ni} &= \frac{1}{n}\sum_{s=1}^{n}\mathbb{1}(X_s > M_{ni}) \\
&= \frac{1}{n}\sum_{s=1}^{n}\mathbb{1}\Big(U_s > 1 - \frac{Z_{ni}}{b_n}\Big) \\
&= \frac{\sqrt{k_n}}{n}\frac{1}{\sqrt{k_n}}\sum_{s=1}^{n}\Big\{\mathbb{1}\Big(U_s > 1 - \frac{Z_{ni}}{b_n}\Big) - \frac{Z_{ni}}{b_n}\Big\} + \frac{Z_{ni}}{b_n} \\
&= \frac{\sqrt{k_n}}{n}e_n(Z_{ni}) + \frac{Z_{ni}}{b_n}. \qquad \text{(A.1)}
\end{aligned}$$

Finally, let

$$\hat{H}_{k_n}(x) := \frac{1}{k_n}\sum_{i=1}^{k_n} \mathbb{1}(Z_{ni} \le x) \tag{A.2}$$

denote the empirical c.d.f. of $Z_{n1}, \dots, Z_{nk_n}$. By Equation (A.1), we obtain

$$\begin{aligned} A_n &= \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n} \log(1-\hat{N}_{ni}) - \log\left(Z_{ni}b_n^{-1}\right) \\ &= \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n} \log\left\{\frac{n}{n+1}\left(\frac{1}{n}+1-\tilde{N}_{ni}\right)\right\} - \log\left(\frac{Z_{ni}}{b_n}\right) \\ &= \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n} \left[\log\left\{\frac{1}{n}+\frac{\sqrt{k_n}}{n}e_n(Z_{ni})+\frac{Z_{ni}}{b_n}\right\} - \log\left(\frac{Z_{ni}}{b_n}\right) + \log\left(\frac{n}{n+1}\right)\right] \\ &= \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n} \log\left\{1+\frac{\sqrt{k_n}b_n}{n}\cdot\frac{e_n(Z_{ni})}{Z_{ni}}+\frac{b_n}{nZ_{ni}}\right\} + \sqrt{k_n}\log\left(\frac{n}{n+1}\right) \\ &= \int_0^\infty W_n(x)\,\mathrm{d}\hat{H}_{k_n}(x) + o(1), \end{aligned} \tag{A.3}$$

where

$$W_n(x) = \sqrt{k_n}\log\left\{1+\frac{1}{\sqrt{k_n}}\left(\frac{e_n(x)}{x}+\frac{1}{\sqrt{k_n}x}\right)\right\}.$$

Heuristically, $\hat{H}_{k_n}(x) \approx 1-\exp(-\theta x)$ and $W_n(x) \approx e(x)/x$ (where e denotes the limit of the tail empirical process), whence the tentative limit of A_n should be

$$A = \int_0^\infty \frac{e(x)}{x}\theta e^{-\theta x}\,\mathrm{d}x.$$

For a rigorous treatment of $A_n + B_n$, let

$$E_n = \int_0^\infty W_n(x)\,\mathrm{d}\hat{H}_{k_n}(x), \qquad E_{n,m} = \int_{1/m}^m W_n(x)\,\mathrm{d}\hat{H}_{k_n}(x),$$

$$E'_m = \int_{1/m}^m \frac{e(x)}{x}\,\theta e^{-\theta x}\,\mathrm{d}x$$

and let B be defined as in Lemma B.1 below. As shown above, $A_n = E_n + o(1)$. The proposition is hence a consequence of Wichura's theorem ([4], Theorem 25.5) and the following items:

(i) For all $m \in \mathbb{N}$: $E_{n,m} + B_n \xrightarrow{d} E'_m + B$ as $n \to \infty$.

(ii) $E'_m + B \xrightarrow{d} A + B \sim \mathcal{N}(0, \sigma^2_{\mathrm{db,C}}/\theta^2)$ as $m \to \infty$.

(iii) For all $\delta > 0$: $\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(|E_n - E_{n,m}| > \delta) = 0$.

The assertion in (i) is proven in Lemma B.4. The assertion in (ii) follows from the fact that $E'_m + B$ is normally distributed with variance τ_m^2 as specified in Lemma B.4, and the fact that $\tau_m^2 \to \sigma^2_{\mathrm{db,C}}/\theta^2$ as $m \to \infty$ by Lemma B.5. Finally, Lemma B.6 proves (iii). □

Proposition A.2. *Under Condition 3.1, 3.2(i), 3.3(i) and 3.4, we have*

$$\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - \varphi_{(\mathrm{C})}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{sb,C}}/\theta^2) \qquad \text{as } n \to \infty.$$

Proof. The proof is very similar to the proof of Proposition A.1. Decompose

$$\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - g(\theta)\} = A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}} + C_n^{\mathrm{sb}},$$

where

$$A_n^{\mathrm{sb}} := \sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - S_n^{\mathrm{sb}}\}, \qquad B_n^{\mathrm{sb}} := \sqrt{k_n}\{S_n^{\mathrm{sb}} - E[S_n^{\mathrm{sb}}]\},$$
$$C_n^{\mathrm{sb}} := \sqrt{k_n}\{E[S_n^{\mathrm{sb}}] - \varphi_{(\mathrm{C})}(\theta)\}.$$

Again, we have $C_n^{\mathrm{sb}} = o(1)$ by Condition 3.3(i). A similar calculation as in (A.3) in the case of the disjoint blocks shows that A_n^{sb} can be written in the following way

$$A_n^{\mathrm{sb}} = \int_0^\infty W_n(x)\, \mathrm{d}\hat{H}_n^{\mathrm{sb}}(x) + o(1),$$

where

$$\hat{H}_n^{\mathrm{sb}}(x) = \frac{1}{n - b_n + 1} \sum_{t=1}^{n-b_n+1} \mathbb{1}(Z_{nt}^{\mathrm{sb}} \le x)$$

denotes the empirical c.d.f. of $Z_{n1}^{\mathrm{sb}}, \dots, Z_{n,n-b_n+1}^{\mathrm{sb}}$. We may now treat $A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}}$ exactly as $A_n + B_n$ in the proof of Proposition A.1, with $E_n, E_{n,m}$ and Lemma B.4, B.5 and B.6 replaced by

$$E_n^{\mathrm{sb}} = \int_0^\infty W_n(x)\, \mathrm{d}\hat{H}_n^{\mathrm{sb}}(x), \qquad E_{n,m}^{\mathrm{sb}} = \int_{1/m}^m W_n(x)\, \mathrm{d}\hat{H}_n^{\mathrm{sb}}(x),$$

and Lemma B.10, B.11 and B.12, respectively. □

A.2. Proof of Theorem 4.2

The following notation will be used throughout:

$$\hat{S}_n = \frac{1}{k_n}\sum_{i=1}^{k_n} \exp(-\hat{Z}_{ni}), \qquad S_n = \frac{1}{k_n}\sum_{i=1}^{k_n} \exp(-Z_{ni}),$$
$$\hat{S}_n^{\mathrm{sb}} = \frac{1}{n - b_n + 1}\sum_{i=1}^{n-b_n+1} \exp(-\hat{Z}_{ni}^{\mathrm{sb}}), \quad S_n^{\mathrm{sb}} = \frac{1}{n - b_n + 1}\sum_{i=1}^{n-b_n+1} \exp(-Z_{ni}^{\mathrm{sb}}).$$

Note that $\hat{\theta}_{\mathrm{db,MAD}}^{z_n} = \varphi_{(\mathrm{M})}^{-1}(\hat{S}_n)$ and $\hat{\theta}_{\mathrm{sb,MAD}}^{z_n} = \varphi_{(\mathrm{M})}^{-1}(\hat{S}_n^{\mathrm{sb}})$, where $\varphi_{(\mathrm{M})}(x) = x/(1+x)$. The assertion follows from the delta-method and Proposition A.3 and A.5. □

Proposition A.3. *Under Condition 3.1 and 3.3(ii), we have*

$$\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{M})}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{db,M}}/(1+\theta)^4) \quad \text{as } n \to \infty.$$

Proof. Write $\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{M})}(\theta)\} = A_n + B_n + C_n$, where

$$A_n = \sqrt{k_n}\{\hat{S}_n - S_n\}, \; B_n = \sqrt{k_n}\{S_n - E[S_n]\}, \; C_n = \sqrt{k_n}\{E[S_n] - \varphi_{(\mathrm{M})}(\theta)\}.$$

The term C_n is asymptotically negligible by Condition 3.3(ii). A straightforward calculation shows that the summand A_n can be written in terms of the tail empirical process e_n as

$$A_n = \int_0^\infty W_n(x)\, \mathrm{d}\hat{H}_{k_n}(x), \qquad W_n(x) = \sqrt{k_n} e^{-x}\left[\exp(-e_n(x)k_n^{-1/2}) - 1\right],$$

where $\hat{H}_{k_n}$ is the empirical c.d.f. of $Z_{n1}, \dots, Z_{nk_n}$, see (A.2). The asymptotic normality of $A_n + B_n$ can now be shown as in the proof of Proposition A.1. The corresponding key result is given by Lemma A.4; whose proof is similar (but easier) as for the CFG-estimator (Lemma B.1) and is omitted for the sake of brevity. □

Lemma A.4. *(a) For any $x_1, \dots, x_m \in [0, \infty)$, as $n \to \infty$,*

$$(e_n(x_1), \dots, e_n(x_m), B_n) \xrightarrow{d} (e(x_1), \dots, e(x_m), B) \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1}),$$

with

$$\Sigma_{m+1} = \begin{pmatrix} r(x_1,x_1) & \dots & r(x_1,x_m) & f(x_1) \\ \vdots & \ddots & \vdots & \vdots \\ r(x_m,x_1) & \dots & r(x_m,x_m) & f(x_m) \\ f(x_1) & \dots & f(x_m) & \frac{\theta}{\theta+2} - \frac{\theta^2}{(\theta+1)^2} \end{pmatrix},$$

where the covariance function r is given as in Lemma B.1 and

$$f(x) = \sum_{i=1}^\infty i \int_0^1 p^{(x)}(i) - p_2^{(x,-\log(y))}(i,0)\mathbb{1}(x \geq -\log(y))\, \mathrm{d}y - x\varphi_{(\mathrm{M})}(\theta).$$

(b) For any $x_1, \dots, x_m \in [0, \infty)$, as $n \to \infty$,

$$(W_n(x_1), \dots, W_n(x_m), B_n) \xrightarrow{d} (-e^{-x_1}e(x_1), \dots, -e^{-x_m}e(x_m), B).$$

Proposition A.5. *Under Condition 3.1 and 3.3(ii), we have*

$$\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - \varphi_{(\mathrm{M})}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{sb,M}}/(1+\theta)^4) \quad \text{as } n \to \infty.$$

Proof. The proof is similar to the proof of Proposition A.3. We may decompose $\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - \varphi_{(\mathrm{M})}(\theta)\} = A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}} + C_n^{\mathrm{sb}}$, where

$$A_n^{\mathrm{sb}} = \sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - S_n^{\mathrm{sb}}\}, \qquad B_n^{\mathrm{sb}} = \sqrt{k_n}\{S_n^{\mathrm{sb}} - E[S_n^{\mathrm{sb}}]\},$$

$$C_n^{\mathrm{sb}} = \sqrt{k_n}\{E[S_n^{\mathrm{sb}}] - \varphi_{(\mathrm{M})}(\theta)\}.$$

Again, we have $C_n^{\mathrm{sb}} = o(1)$ by Condition 3.3(ii) and

$$A_n^{\mathrm{sb}} = \int_0^\infty W_n(x)\,\mathrm{d}\hat{H}_n^{\mathrm{sb}}(x),$$

where $\hat{H}_n^{\mathrm{sb}}$ denotes the empirical c.d.f. of $Z_{n1}^{\mathrm{sb}}, \dots, Z_{n,n-b_n+1}^{\mathrm{sb}}$. The sum $A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}}$ can now be treated as in proof of Proposition A.2. The corresponding key result, Lemma B.7, needs to be replaced by Lemma A.6; whose proof is again omitted for the sake of brevity. □

Lemma A.6. *(a) For any $x_1, \dots, x_m \in [0, \infty)$, as $n \to \infty$,*

$$(e_n(x_1), \dots, e_n(x_m), B_n^{\mathrm{sb}}) \xrightarrow{d} (e(x_1), \dots, e(x_m), B^{\mathrm{sb}}) \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1}^{\mathrm{sb}}),$$

where all entries of $\Sigma_{m+1}^{\mathrm{sb}}$ are the same as those of Σ_{m+1} in Lemma A.4 except for the entry at position $(m+1, m+1)$, which needs to be replaced by

$$v(\theta) = 2 - \frac{4}{\theta+1} + 4\,\frac{\log(\theta+1) - \log(\theta+2) + \log(2)}{\theta(\theta+1)} - \frac{2\theta^2}{(\theta+1)^2}.$$

(b) For any $x_1, \dots, x_m \in [0, \infty)$, as $n \to \infty$,

$$(W_n(x_1), \dots, W_n(x_m), B_n^{\mathrm{sb}}) \xrightarrow{d} (-e^{-x_1} e(x_1), \dots, -e^{-x_m} e(x_m), B^{\mathrm{sb}}).$$

A.3. Proof of Theorem 4.3

For fixed $p > 0$, define

$$\hat{S}_n = \frac{1}{k_n}\sum_{i=1}^{k_n} \hat{Z}_{ni}^{1/p}, \qquad S_n = \frac{1}{k_n}\sum_{i=1}^{k_n} Z_{ni}^{1/p},$$

$$\hat{S}_n^{\mathrm{sb}} = \frac{1}{n-b_n+1}\sum_{i=1}^{n-b_n+1} \hat{Z}_{ni}^{1/p}, \qquad S_n^{\mathrm{sb}} = \frac{1}{n-b_n+1}\sum_{i=1}^{n-b_n+1} Z_{ni}^{1/p}.$$

Note that $\hat{\theta}_{\mathrm{db,R},p}^{z_n} = \varphi_{(\mathrm{R}),\mathrm{p}}^{-1}(\hat{S}_n)$ and $\hat{\theta}_{\mathrm{sb,R},p}^{z_n} = \varphi_{(\mathrm{R}),\mathrm{p}}^{-1}(\hat{S}_n^{\mathrm{sb}})$, where $\varphi_{(\mathrm{R}),\mathrm{p}}(x) = x^{-1/p}\Gamma(1+1/p)$. By the delta-method, the assertion follows from Proposition A.7 and A.9. □

Proposition A.7. *Under Condition 3.1, 3.2(ii) and 3.3(iii), we have*

$$\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma_{\mathrm{db},p}^2 \psi_p(\theta)) \text{ as } n \to \infty,$$

where $\psi_p(\theta) = \Gamma(1+1/p)^2 p^{-2} \theta^{-(2+2/p)}$.

Proof. Decompose $\sqrt{k_n}\{\hat{S}_n - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\} = A_n + B_n + C_n$, where

$$A_n = \sqrt{k_n}\{\hat{S}_n - S_n\}, \qquad B_n = \sqrt{k_n}\{S_n - E[S_n]\},$$
$$C_n = \sqrt{k_n}\{E[S_n] - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\}.$$

By Condition 3.3(iii), the term C_n converges to zero. A straightforward calculation shows that the term A_n can be written as

$$A_n = \int_0^\infty W_n(x)\,\mathrm{d}\hat{H}_{k_n}(x), \qquad W_n(x) = \sqrt{k_n}\left\{\left[\frac{e_n(x)}{\sqrt{k_n}} + x\right]^{1/p} - x^{1/p}\right\}.$$

The asymptotic normality of $A_n + B_n$ can be shown as in the proof of Proposition A.1 by an application of Wichura's theorem. Here, Lemma B.1 needs to be replaced by Lemma A.8, whose proof is similar but easier and therefore omitted for the sake of brevity. □

Lemma A.8. *(a) For any $x_1, \dots, x_m \in (0,\infty)$, as $n \to \infty$,*

$$(e_n(x_1), \dots, e_n(x_m), B_n) \xrightarrow{d} (e(x_1), \dots, e(x_m), B) \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1})$$

with

$$\Sigma_{m+1} = \begin{pmatrix} r(x_1,x_1) & \dots & r(x_1,x_m) & f_p(x_1) \\ \vdots & \ddots & \vdots & \vdots \\ r(x_m,x_1) & \dots & r(x_m,x_m) & f_p(x_m) \\ f_p(x_1) & \dots & f_p(x_m) & v_p(\theta) \end{pmatrix},$$

where the covariance function r is defined as in Lemma B.1 and

$$f_p(x) = \sum_{i=1}^\infty i \int_0^\infty p_2^{(x,y^p)}(i,0)\mathbb{1}(x \ge y^p)\,\mathrm{d}y - x\varphi_{(\mathrm{R}),\mathrm{p}}(\theta),$$
$$v_p(\theta) = \theta^{\frac{-2}{p}}\left\{\Gamma(1+2/p) - \Gamma(1+1/p)^2\right\}.$$

(b) For any $x_1, \dots, x_m \in (0,\infty)$, as $n \to \infty$,

$$(W_n(x_1), \dots, W_n(x_m), B_n) \xrightarrow{d} \big(e(x_1)x_1^{\frac{1}{p}-1}p^{-1}, \dots, e(x_m)x_m^{\frac{1}{p}-1}p^{-1}, B\big).$$

Proposition A.9. *Under Condition 3.1, 3.2(ii) and 3.3(iii), we have*

$$\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathrm{sb},p}\psi_p(\theta)) \text{ as } n \to \infty,$$

where $\psi_p(\theta) = \Gamma(1+1/p)^2 p^{-2}\theta^{-(2+2/p)}$.

Proof. The proof is similar to the proof of Proposition A.7. Write $\sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\} = A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}} + C_n^{\mathrm{sb}}$, where

$$A_n^{\mathrm{sb}} = \sqrt{k_n}\{\hat{S}_n^{\mathrm{sb}} - S_n^{\mathrm{sb}}\}, \qquad B_n^{\mathrm{sb}} = \sqrt{k_n}\{S_n^{\mathrm{sb}} - E[S_n^{\mathrm{sb}}]\},$$

$$C_n^{\mathrm{sb}} = \sqrt{k_n}\{E[S_n^{\mathrm{sb}}] - \varphi_{(\mathrm{R}),\mathrm{p}}(\theta)\}.$$

By Condition 3.3(iii), $C_n^{\mathrm{sb}} = o(1)$, and a straightforward calculation yields

$$A_n^{\mathrm{sb}} = \int_0^\infty W_n(x)\, \mathrm{d}\hat{H}_n^{\mathrm{sb}}(x),$$

where $\hat{H}_n^{\mathrm{sb}}$ denotes the empirical c.d.f. of $Z_{n1}^{\mathrm{sb}}, \dots, Z_{n,n-b_n+1}^{\mathrm{sb}}$. The sum $A_n^{\mathrm{sb}} + B_n^{\mathrm{sb}}$ can be treated as in the proof of Proposition A.2, where the main result, Lemma B.7, needs to be replaced by Lemma A.10, whose proof is omitted for the sake of brevity. □

Lemma A.10. *(a) For any $x_1, \dots, x_m \in (0, \infty)$, as $n \to \infty$,*

$$(e_n(x_1), \dots, e_n(x_m), B_n^{\mathrm{sb}}) \xrightarrow{d} (e(x_1), \dots, e(x_m), B^{\mathrm{sb}}) \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1}^{\mathrm{sb}}),$$

where all entries of $\Sigma_{m+1}^{\mathrm{sb}}$ are the same as those of Σ_{m+1} in Lemma A.8 except for the entry at position $(m+1, m+1)$, which needs to be replaced by

$$v_p^{\mathrm{sb}}(\theta) = 4p^{-2}\theta^{-2/p} \int_0^\infty (1 - e^{-z}) z^{1/p-2} \Gamma(1/p, z)\, \mathrm{d}z - 2\theta^{-2/p}\Gamma(1 + 1/p)^2.$$

(b) For any $x_1, \dots, x_m \in (0, \infty)$, as $n \to \infty$,

$$(W_n(x_1), \dots, W_n(x_m), B_n^{\mathrm{sb}}) \xrightarrow{d} \big(e(x_1) x_1^{\frac{1}{p}-1} p^{-1}, \dots, e(x_m) x_m^{\frac{1}{p}-1} p^{-1}, B^{\mathrm{sb}}\big).$$

Appendix B: Auxiliary results for the proof of Theorem 4.1

B.1. Auxiliary lemmas – disjoint blocks

Throughout this section, we assume that Condition 3.1, 3.2(i) and 3.3(i) are met.

Lemma B.1. *For any $x_1, \dots, x_m \in [0, \infty)$ and $m \in \mathbb{N}$, we have*

$$(e_n(x_1), \dots, e_n(x_m), B_n)' \xrightarrow{d} (e(x_1), \dots, e(x_m), B)',$$

where $(e(x_1), \dots, e(x_m), B)' \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1})$ with

$$\Sigma_{m+1} = \begin{pmatrix} r(x_1, x_1) & \dots & r(x_1, x_m) & f(x_1) \\ \vdots & \ddots & \vdots & \vdots \\ r(x_m, x_1) & \dots & r(x_m, x_m) & f(x_m) \\ f(x_1) & \dots & f(x_m) & \pi^2/6 \end{pmatrix}.$$

Here, $r(0,0) = 0$ and, for $x \ge y \ge 0$ with $x \ne 0$,

$$r(x, y) = \theta x \sum_{i=1}^\infty \sum_{j=0}^i ij\pi_2^{(y/x)}(i, j), \qquad f(x) = h(x) - x\varphi_{(\mathrm{C})}(\theta),$$

$$h(x) = \sum_{i=1}^{\infty} i \left[\int_0^{\infty} \mathbb{1}(e^y \le x) p_2^{(x,e^y)}(i,0) \, \mathrm{d}y \right.$$
$$\left. - \int_{-\infty}^{0} p^{(x)}(i) - \mathbb{1}(e^y \le x) p_2^{(x,e^y)}(i,0) \, \mathrm{d}y \right]$$

and where, for $i \ge j \ge 0, \ i \ge 1,$

$$p_2^{(x,y)}(i,j) = \mathbb{P}\big(\boldsymbol{N}_E^{(x,y)} = (i,j)\big), \quad \boldsymbol{N}_E^{(x,y)} = \sum_{i=1}^{\eta} \big(\xi_{i1}^{(y/x)}, \xi_{i2}^{(y/x)}\big)$$

with $\eta \sim \mathrm{Poisson}(\theta x)$ *independent of i.i.d. random vectors* $\big(\xi_{i1}^{(y/x)}, \xi_{i2}^{(y/x)}\big) \sim \pi_2^{(y/x)}, i \in \mathbb{N}$ *and*

$$p^{(x)}(i) = \mathbb{P}(N_E^{(x)} = i), \quad N_E^{(x)} = \sum_{i=1}^{\eta_2} \xi_i$$

with $\eta_2 \sim \mathrm{Poisson}(\theta x)$ *independent of i.i.d. random variables* $\xi_i \sim \pi, \ i \in \mathbb{N}$.

Lemma B.2. *For any* $m \in \mathbb{N}$*, we have*

$$\{(W_n(x), B_n)'\}_{x\in[1/m,m]} \xrightarrow{d} \left\{ \left(\frac{e(x)}{x}, B \right)' \right\}_{x\in[1/m,m]} \quad \text{in } D([1/m,m]) \times \mathbb{R},$$

where $(e, B)'$ *is a centered Gaussian process with continuous sample paths and with covariance functional as specified in Lemma B.1.*

Lemma B.3. *For any* $m \in \mathbb{N}$*, we have*

$$E_{n,m} = E'_{n,m} + o_{\mathbb{P}}(1) \quad \text{as } n \to \infty,$$

where $E'_{n,m} = \int_{1/m}^{m} W_n(x)\theta e^{-\theta x} \, \mathrm{d}x$.

Lemma B.4. *For any* $m \in \mathbb{N}$*, we have*

$$E_{n,m} + B_n \xrightarrow{d} E'_m + B \sim \mathcal{N}(0, \tau_m^2) \quad \text{as } n \to \infty,$$

where, with r *and* f *defined as in Lemma B.1,*

$$\tau_m^2 = \theta^2 \int_{1/m}^{m} \int_{1/m}^{m} r(x,y) \frac{1}{xy} e^{-\theta(x+y)} \, \mathrm{d}x\mathrm{d}y + 2\theta \int_{1/m}^{m} f(x) \frac{1}{x} e^{-\theta x} \, \mathrm{d}x + \frac{\pi^2}{6}.$$

Lemma B.5. *As* $m \to \infty$*,* $\tau_m^2 \to \sigma_{\mathrm{db,(C)}}^2/\theta^2$*, where* $\sigma_{\mathrm{db,(C)}}^2$ *is specified in Theorem 4.1.*

Lemma B.6. *If, in addition to Condition 3.1, 3.2(i) and 3.3(i), Condition 3.4 holds, then, for all* $\delta > 0$*,*

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(|E_{n,m} - E_n| > \delta) = 0.$$

Proof of Lemma B.1. We proceed similarly as in the proof of Lemma 9.3 in [3]. Weak convergence of the $(e_n(x_1), \dots, e_n(x_m))'$ is a consequence of Theorem 4.1 in [27]. For the treatment of the joint convergence with B_n, we only consider the case $m = 1$ and set $x_1 = x$; the general case can be treated analogously. For $i = 1, \dots, k_n$, we decompose a block $I_i = \{(i-1)b_n + 1, \dots, ib_n\}$ into a big block I_i^+ and a small block I_i^-, where, recalling ℓ_n from Condition 3.1(iii),

$$I_i^+ = \{(i-1)b_n + 1, \dots, ib_n - \ell_n\}, \qquad I_i^- = \{ib_n - \ell_n + 1, \dots, ib_n\},$$

and set

$$e_n^+(x) = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \sum_{s \in I_i^+} \Big\{ \mathbb{1}\Big(U_s > 1 - \frac{x}{b_n}\Big) - \frac{x}{b_n} \Big\},$$

$$B_n^+ = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \Big\{ \log(Z_{ni}^+) - \mathrm{E}[\log(Z_{ni}^+)] \Big\},$$

where $Z_{ni}^+ = b_n(1 - N_{ni}^+), N_{ni}^+ = \max_{s \in I_i^+} U_s$. Next, according to Lemma 6.6 in [27],

$$e_n^-(x) := e_n(x) - e_n^+(x) = o_{\mathbb{P}}(1).$$

It can further be shown by the same arguments as in the proof of Lemma 9.3 in [3] that

$$B_n^- := B_n - B_n^+ = o_{\mathbb{P}}(1).$$

Finally, for $\varepsilon \in (0, c_1 \wedge c_2)$, define $A_n^+ = \{\min_{i=1,\dots,k_n} N_{ni}^+ > 1 - \varepsilon\}$, and note that $\mathbb{P}(A_n^+) \to 1$ by Condition 3.1(v). As a consequence of the previous three statements, it suffices to show that, using the Cramér-Wold device,

$$\{\lambda_1 e_n^+(x) + \lambda_2 B_n^+\} \mathbb{1}_{A_n^+} \xrightarrow{d} \lambda_1 e(x) + \lambda_2 B, \tag{B.1}$$

for any $\lambda_1, \lambda_2 \in \mathbb{R}$.

Now, the left-hand side of (B.1) can be written as

$$\{\lambda_1 e_n^+(x) + \lambda_2 B_n^+\} \mathbb{1}_{A_n^+} = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \tilde{g}_{i,n} + o_{\mathbb{P}}(1),$$

where $\tilde{g}_{i,n} = g_{i,n} \mathbb{1}(Z_{ni}^+ < \varepsilon b_n)$ and where

$$g_{i,n} = \lambda_1 \sum\nolimits_{s \in I_i^+} \Big\{ \mathbb{1}\Big(U_s > 1 - \frac{x}{b_n}\Big) - \frac{x}{b_n} \Big\} + \lambda_2 \big\{ \log(Z_{ni}^+) - \mathrm{E}[\log(Z_{ni}^+)] \big\}.$$

Note, that $\tilde{g}_{i,n}$ only depends on the block I_i^+ and is $\mathcal{B}^{\varepsilon}_{(i-1)b_n+1:ib_n-\ell_n}$-measurable. In particular, the $(\tilde{g}_{i,n})_{i=1,\dots,k_n}$ are each separated by a small block of length ℓ_n. A standard argument based on characteristic functions and the assumption on alpha mixing may then be used to show that the weak limit of $k_n^{-1/2} \sum_{i=1}^{k_n} \tilde{g}_{i,n}$ is the same as if the $\tilde{g}_{i,n}$ were independent.

Next, we show that Ljapunov's condition ([4], Theorem 27.3) is satisfied. By Minkowski's inequality, for any $p \in (2, 2+\delta)$, we have $C_\infty = \sup_{n\in\mathbb{N}} \mathrm{E}[|\tilde g_{1,n}|^p] < \infty$ by Condition 3.1(ii) and 3.2(i). Further, by stationarity and independence, we get

$$\frac{\sum_{i=1}^{k_n} \mathrm{E}[|\tilde g_{i,n}|^p]}{\mathrm{Var}\left(\sum_{i=1}^{k_n} \tilde g_{i,n}\right)^{p/2}} = k_n^{1-p/2} \frac{\mathrm{E}[|\tilde g_{1,n}|^p]}{\left(\mathrm{E}[\tilde g_{1,n}^2]\right)^{p/2}} \le C_\infty \times k_n^{1-p/2}\, \mathrm{E}[\tilde g_{1,n}^2]^{-p/2}.$$

Hence, provided $\lim_{n\to\infty} \mathrm{E}[\tilde g_{1,n}^2]$ exists, the last expression converges to 0 and hence Ljapunov's condition is met. As a consequence, $k_n^{-1/2}\sum_{i=1}^{k_n} \tilde g_{i,n}$ weakly converges to a centered normal distribution with variance $\lim_{n\to\infty} \mathrm{E}[\tilde g_{1,n}^2]$.

Finally, since $\lim_{n\to\infty} \mathrm{E}[\tilde g_{1,n}^2] = \lim_{n\to\infty} \mathrm{E}[g_{1,n}^2]$, it remains to be shown that

$$\lim_{n\to\infty} \mathrm{E}[g_{1,n}^2] = \lambda_1^2 r(x,x) + 2\lambda_1\lambda_2 h(x) + \lambda_2^2 \pi^2/6.$$

Since similar arguments as in the proof of $B_n^- = o_\mathbb{P}(1)$ and $e_n^- = o_\mathbb{P}(1)$ allow us to replace I_1^+ by I_1 and then b_n by n, this in turn is a consequence of

$$\lim_{n\to\infty} \mathrm{Var}\left(N_n^{(x)}(E)\right) = r(x,x), \tag{B.2}$$

$$\lim_{n\to\infty} \mathrm{Cov}\left\{N_n^{(x)}(E), \log(Z_{1:n})\right\} = f(x), \tag{B.3}$$

$$\lim_{n\to\infty} \mathrm{Var}\{\log(Z_{1:n})\} = \pi^2/6. \tag{B.4}$$

The assertion in (B.2) follows from Theorem 4.1 in [27]. Further, since $Z_{1:n} \xrightarrow{d} \xi \sim \mathrm{Exp}(\theta)$ and since since $|\log(Z_{1:n})|^2$ is uniformly integrable by Condition 3.2(i), we have

$$\lim_{n\to\infty} \mathrm{Var}\{\log(Z_{1:n})\} = \mathrm{Var}\{\log(\xi)\} = \frac{\pi^2}{6},$$

which is (B.4). Finally, note that $\mathrm{E}[N_n^{(x)}(E)] = x$ and $\mathrm{E}[\log(Z_{1:n})] \to \varphi_{(\mathrm{C})}(\theta)$ by similar arguments as given above. As a consequence, (B.3) follows from $\lim_{n\to\infty} \mathrm{E}\left[N_n^{(x)}(E)\log(Z_{1:n})\right] = h(x)$. The latter in turn can be seen as follows: first,

$$\mathrm{E}\left[N_n^{(x)}(E)\log(Z_{1:n})\right] = \sum_{i=1}^n i\, \mathrm{E}\left[\mathbb{1}(N_n^{(x)}(E) = i)\log(Z_{1:n})\right]. \tag{B.5}$$

The expected value on the right-hand side can be written as

$$\begin{aligned} &\int_0^\infty \mathbb{P}\big(\mathbb{1}(N_n^{(x)}(E) = i)\log(Z_{1:n}) > y\big)\,\mathrm{d}y \\ &\qquad\qquad - \int_{-\infty}^0 1 - \mathbb{P}(\mathbb{1}(N_n^{(x)}(E) = i)\log(Z_{1:n}) > y)\,\mathrm{d}y \\ &= \int_0^\infty \mathbb{P}(N_n^{(x)}(E) = i, Z_{1:n} > e^y)\,\mathrm{d}y \end{aligned}$$

$$- \int_{-\infty}^{0} \mathbb{P}(N_n^{(x)}(E) = i) - \mathbb{P}(N_n^{(x)}(E) = i, Z_{1:n} > e^y) \, \mathrm{d}y.$$

Now,

$$\mathbb{P}(N_n^{(x)}(E) = i,\ Z_{1:n} > e^y) = \mathbb{P}(N_n^{(x)}(E) = i,\ N_n^{(e^y)}(E) = 0)$$
$$\to \begin{cases} p_2^{(x,e^y)}(i,0) & , \ x \geq e^y \geq 0, \\ 0 & , e^y > x \geq 0 \end{cases}$$

and $\mathbb{P}(N_n^{(x)}(E) = i) \to p^{(x)}(i)$, see [24] and [27]. By uniform integrability we obtain that the expected value on the right-hand side of (B.5) converges to $h(x)$. The proof is finished. ☐

Proof of Lemma B.2. For fixed $x > 0$, consider the function

$$f_n : \mathbb{R} \to \mathbb{R},\ f_n(z) = \sqrt{k_n} \log \Big\{ 1 + \frac{1}{\sqrt{k_n}} \Big(\frac{z}{x} + \frac{1}{\sqrt{k_n} x} \Big) \Big\}.$$

For $z_n \to z$, one has $f_n(z_n) \to e(z)/z$. Hence, since $(e_n(x_1), \ldots, e_n(x_m), B_n)'$ converges in distribution to $(e(x_1), \ldots, e(x_m), B)'$ for any $x_1, \ldots, x_m > 0$ and $m \in \mathbb{N}$ by Lemma B.1, we can apply the extended continuous mapping theorem (Theorem 18.11 in 33) to obtain $(W_n(x_1), \ldots, W_n(x_m), B_n)' \to (e(x_1)/x_1, \ldots, e(x_m)/x_m, B)'$ in distribution. This is the fidi-convergence needed to prove Lemma B.2.

Asymptotic tightness of the tail empirical process e_n follows from Theorem 4.1 in [27]. Asymptotic tightness of B_n follows from its weak convergence. This implies asymptotic tightness of the vector (e_n, B_n), for instance by a simple adaptation of Lemma 1.4.3 in [34]. ☐

Proof of Lemma B.3. Let $H(x) = 1 - e^{-\theta x}$ be the cdf of the Exp(θ)-distribution. From the proof of Lemma 9.2 in [3], we have, for any $m \in \mathbb{N}$,

$$\sup_{x \in [1/m, m]} |\hat{H}_{k_n}(x) - H(x)| = o_{\mathbb{P}}(1), \quad n \to \infty.$$

Since

$$E_{n,m} - E'_{n,m} = \int_{1/m}^{m} W_n(x) \, \mathrm{d}(\hat{H}_{k_n} - H)(x),$$

the assertion follows from Lemma B.2, Lemma C.8 in [2] and the continuous mapping theorem. ☐

Proof of Lemma B.4. As a consequence of Lemma B.3, Lemma B.2 and the continuous mapping theorem, we have

$$E_{n,m} + B_n = \int_{1/m}^{m} W_n(x) \ \theta e^{-\theta x} \ \mathrm{d}x + B_n + o_{\mathbb{P}}(1)$$
$$\xrightarrow{d} \int_{1/m}^{m} \frac{e(x)}{x} \theta e^{-\theta x} \ \mathrm{d}x + B = E'_m + B \sim \mathcal{N}(0, \tau_m^2),$$

where the variance τ_m^2 is given by

$$\tau_m^2 = \operatorname{Var}\Big\{ \int_{1/m}^m e(x)\frac{1}{x}\theta e^{-\theta x}\,\mathrm{d}x\Big\} + 2\operatorname{Cov}\Big\{ \int_{1/m}^m e(x)\frac{1}{x}\theta e^{-\theta x}\,\mathrm{d}x, B\Big\} + \operatorname{Var}(B)$$
$$= \theta^2 \int_{1/m}^m\int_{1/m}^m r(x,y)\frac{1}{xy}e^{-\theta(x+y)}\,\mathrm{d}x\,\mathrm{d}y + 2\theta\int_{1/m}^m f(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x + \frac{\pi^2}{6}$$

as asserted. □

Proof of Lemma B.5. By the definition of τ_m^2 in Lemma B.4

$$\lim_{m\to\infty}\tau_m^2 = \theta^2\int_0^\infty\int_0^\infty r(x,y)\frac{1}{xy}e^{-\theta(x+y)}\,\mathrm{d}x\,\mathrm{d}y + 2\theta\int_0^\infty f(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x + \frac{\pi^2}{6}. \tag{B.6}$$

For $x>y$, we have $r(x,y) = \theta x\,\mathrm{E}\big[\xi_1^{(y/x)}\xi_2^{(y/x)}\big]$ with $(\xi_1^{(y/x)},\xi_2^{(y/x)}) \sim \pi_2^{(y/x)}$. Hence, applying the transformation $z=y/x$, the first summand on the right-hand side of (B.6) can be written as

$$\begin{aligned}\theta^2\int_0^\infty\int_0^\infty \frac{r(x,y)}{xy}e^{-\theta(x+y)}\,\mathrm{d}x\,\mathrm{d}y &= 2\theta^3\int_0^\infty\int_0^x \frac{\mathrm{E}\big[\xi_1^{(y/x)}\xi_2^{(y/x)}\big]}{y}e^{-\theta(x+y)}\,\mathrm{d}y\,\mathrm{d}x\\ &= 2\theta^3\int_0^\infty\int_0^1 \frac{\mathrm{E}\big[\xi_1^{(z)}\xi_2^{(z)}\big]}{z}e^{-\theta x(1+z)}\,\mathrm{d}z\,\mathrm{d}x\\ &= 2\theta^2\int_0^1 \frac{\mathrm{E}\big[\xi_1^{(z)}\xi_2^{(z)}\big]}{z(z+1)}\,\mathrm{d}z. \end{aligned}\tag{B.7}$$

For the second summand on the right-hand side of (B.6), note that

$$\sum_{i=1}^\infty i p_2^{(x,e^y)}(i,0) = \mathrm{E}\big[\xi_1^{(e^y/x)}\mathbb{1}(\xi_2^{(e^y/x)}=0)\big]\theta x e^{-\theta e^y}, \tag{B.8}$$

see Formula (A.7) in the proof of Lemma 9.6 in [3] and $\sum_{i=1}^\infty ip^{(x)}(i) = \mathrm{E}[N_E^{(x)}] = x$, see [27]. Therefore, we can rewrite h from Lemma B.1 as follows

$$\begin{aligned}h(x) &= \int_0^\infty \mathbb{1}(e^y\le x)\,\mathrm{E}\big[\xi_1^{(e^y/x)}\mathbb{1}(\xi_2^{(e^y/x)}=0)\big]\theta x e^{-\theta e^y}\,\mathrm{d}y\\ &\qquad - \int_{-\infty}^0 x - \mathbb{1}(e^y\le x)\,\mathrm{E}\big[\xi_1^{(e^y/x)}\mathbb{1}(\xi_2^{(e^y/x)}=0)\big]\theta x e^{-\theta e^y}\,\mathrm{d}y.\\ &= x\int_{1/x}^\infty \mathbb{1}(z\le 1)\,\mathrm{E}\big[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\big]\theta\frac{e^{-\theta zx}}{z}\,\mathrm{d}z\\ &\qquad - x\int_0^{1/x}\frac{1}{z} - \mathbb{1}(z\le 1)\,\mathrm{E}\big[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\big]\theta\frac{e^{-\theta zx}}{z}\,\mathrm{d}z,\end{aligned}$$

where we have used the transformation $z = e^y/x$. For $0 < x \le 1$, the first integral is zero and we obtain

$$\begin{aligned} h(x) &= -x\int_0^1 \frac{1}{z} - \mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\right]\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z - x\int_1^{1/x}\frac{1}{z}\,\mathrm{d}z \\ &= -x\int_0^1 \frac{1}{z} - \mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\right]\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z + x\log(x), \end{aligned}$$

while for $x > 1$,

$$\begin{aligned} h(x) = x\int_{1/x}^1 &\mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\right]\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z \\ &- x\int_0^{1/x} \frac{1}{z} - \mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\right]\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z. \end{aligned}$$

As a consequence, writing $g(z) = \mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}=0)\right]$, we obtain

$$\begin{aligned} \int_0^\infty h(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x = \int_0^1 \log(x)e^{-\theta x}\,\mathrm{d}x &- \int_0^1 e^{-\theta x}\int_0^1 \frac{1}{z} - g(z)\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z\,\mathrm{d}x \\ &+ \int_1^\infty e^{-\theta x}\int_{1/x}^1 g(z)\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z\,\mathrm{d}x \\ &- \int_1^\infty e^{-\theta x}\int_0^{1/x}\frac{1}{z} - g(z)\theta\frac{e^{-\theta z x}}{z}\,\mathrm{d}z\,\mathrm{d}x. \end{aligned}$$

Next, some tedious calculations based on Fubini's theorem allow to rewrite the sum of the last three double integrals as

$$s = \int_0^1 \frac{e^{-\theta/z}-1}{\theta z} + \frac{g(z)}{z(1+z)}\,\mathrm{d}z.$$

Using the fact that $g(z) = \frac{1}{\theta} - \mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}>0)\right]$, we thus obtain

$$\begin{aligned} &\int_0^\infty h(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x \\ &= \int_0^1 \log(z)e^{-\theta z} + \frac{e^{-\theta/z}-1}{\theta z} + \frac{1}{\theta z(1+z)} - \frac{\mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}>0)\right]}{z(1+z)}\,\mathrm{d}z \\ &= \int_0^1 \log(z)e^{-\theta z} + \frac{e^{-\theta/z}}{\theta z} - \frac{1}{\theta(1+z)} - \frac{\mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)}>0)\right]}{z(1+z)}\,\mathrm{d}z. \end{aligned}$$

Finally, one can show

$$\int_0^1 \log(z)e^{-\theta z} + \frac{e^{-\theta/z}}{\theta z}\,\mathrm{d}z = -(\log\theta + \gamma)/\theta = \varphi_{(\mathrm{C})}(\theta)/\theta,$$

such that, assembling terms and recalling $f(x) = h(x) - x\varphi_{(\mathrm{C})}(\theta)$,

$$\int_0^1 f(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x = \int_0^1 h(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x - \varphi_{(\mathrm{C})}(\theta)\int_0^\infty e^{-\theta x}\,\mathrm{d}x$$
$$= -\log(2)/\theta - \int_0^1 \frac{\mathrm{E}\left[\xi_1^{(z)}\mathbb{1}(\xi_2^{(z)} > 0)\right]}{z(1+z)}\,\mathrm{d}z. \tag{B.9}$$

The lemma is now an immediate consequence of (B.6), (B.7) and (B.9). □

Proof of Lemma B.6. By Lemma B.3, it suffices to show the assertion with $E_{n,m}$ replaced by $E'_{n,m}$. Define $\tilde{e}_n(x) := e_n(x) + k_n^{-1/2}$, such that, by Condition 3.4(iii), we have

$$\max_{Z_{ni}\ge c}\left|\frac{\tilde{e}_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}\right| = o_{\mathbb{P}}(1)$$

for any constant $c > 0$. Fix $m \in \mathbb{N}$. By the previous display, for any $\varepsilon > 0$, the event

$$B_n = B_n(m,\varepsilon) = \left\{\max_{Z_{ni}\ge m}\left|\frac{\tilde{e}_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}\right| \le \varepsilon\right\}$$

satisfies $\mathbb{P}(B_n) \to 1$. Next,

$$|E_{n,m} - E_n| \le \left|\int_0^\infty \log\left(1 + \frac{\tilde{e}_n(x)}{x\sqrt{k_n}}\right)\sqrt{k_n}\,\mathbb{1}_{(0,1/m]}(x)\,\mathrm{d}\hat{H}_{k_n}(x)\right|$$
$$+ \left|\int_0^\infty \log\left(1 + \frac{\tilde{e}_n(x)}{x\sqrt{k_n}}\right)\sqrt{k_n}\,\mathbb{1}_{[m,\infty)}(x)\,\mathrm{d}\hat{H}_{k_n}(x)\right|$$
$$=: |V_{n1}| + |V_{n2}|,$$

such that

$$|E_{n,m} - E_n| = |E_{n,m} - E_n|\mathbb{1}_{B_n} + o_{\mathbb{P}}(1) \le |V_{n1}| + |V_{n2}|\mathbb{1}_{B_n} + o_{\mathbb{P}}(1). \tag{B.10}$$

We begin by treating the term $|V_{n2}|\mathbb{1}_{B_n}$. Since $\log(1+x) = \int_0^1 x/(1+sx)\,\mathrm{d}s$ for any $x > -1$, we have

$$V_{n2}\mathbb{1}_{B_n} = \int_0^\infty \frac{\tilde{e}_n(x)}{x}\int_0^1 \frac{1}{1 + s\frac{\tilde{e}_n(x)}{x\sqrt{k_n}}}\,\mathrm{d}s\;\;\mathbb{1}(x \ge m)\,\mathrm{d}\hat{H}_{k_n}(x)\,\mathbb{1}_{B_n}$$
$$= \frac{1}{k_n}\sum_{i=1}^{k_n}\frac{\tilde{e}_n(Z_{ni})}{Z_{ni}}\mathbb{1}(Z_{ni} \ge m)\int_0^1 \frac{1}{1 + s\frac{\tilde{e}_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}}\,\mathrm{d}s\,\mathbb{1}_{B_n}$$
$$= k_n^{-3/2}\sum_{i=1}^{k_n}\frac{\mathbb{1}(Z_{ni} \ge m)}{Z_{ni}}\int_0^1 \frac{1}{1 + s\frac{\tilde{e}_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}}\,\mathrm{d}s\,\left\{\sum_{t=1}^n f(U_t, Z_{ni}) + 1\right\}\mathbb{1}_{B_n},$$

where

$$f(U_t, Z_{ni}) = \mathbb{1}(U_t > 1 - Z_{ni}/b_n) - Z_{ni}/b_n. \tag{B.11}$$

For given $\varepsilon \in (0, c_1 \wedge c_2)$ with c_j as in Condition 3.1, let $C_n = C_n(\varepsilon)$ denote the event $\{\min_{i=1,\dots,k_n} N_{ni} > 1-\varepsilon/2\} = \{\max_{i=1,\dots,k_n} Z_{ni} < \varepsilon b_n/2\}$, which satisfies $\mathbb{P}(C_n) \to 1$ by Condition 3.1(v). Hence, we can write $V_{n2}\mathbb{1}_{B_n} = \bar V_{n2}\mathbb{1}_{C_n} + o_{\mathbb{P}}(1)$, where

$$\bar V_{n2} = k_n^{-3/2} \sum_{i=1}^{k_n} \frac{1}{Z_{ni}} \mathbb{1}(\varepsilon b_n/2 > Z_{ni} \ge m) \int_0^1 \frac{1}{1 + s\frac{\tilde e_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}} \,\mathrm{d}s \times \Big\{ \sum_{t=1}^n f(U_t, Z_{ni}) + 1 \Big\} \mathbb{1}_{B_n}.$$

We obtain

$$|\bar V_{n2}| \le \frac{1}{m} k_n^{-3/2} \sum_{i=1}^{k_n} \mathbb{1}(\varepsilon b_n/2 > Z_{ni} \ge m) \int_0^1 \frac{1}{\Big|1 + s\frac{\tilde e_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}\Big|} \,\mathrm{d}s \times \Big\{ \Big|\sum_{t=1}^n f(U_t, Z_{ni})\Big| + 1 \Big\} \mathbb{1}_{B_n}.$$

On the event B_n the integral over s can be bounded as follows

$$\int_0^1 \frac{1}{\Big|1 + s\frac{\tilde e_n(Z_{ni})}{Z_{ni}\sqrt{k_n}}\Big|} \,\mathrm{d}s\ \mathbb{1}_{B_n} \le \int_0^1 \frac{1}{1 - s\varepsilon} \,\mathrm{d}s\ \mathbb{1}_{B_n} \le \frac{1}{1-\varepsilon}.$$

The previous two displays imply that $|\bar V_{n2}|$ is bounded by

$$\begin{aligned} &\frac{1}{m}\frac{1}{1-\varepsilon}\ k_n^{-3/2} \sum_{i=1}^{k_n} \mathbb{1}(\varepsilon b_n/2 > Z_{ni} \ge m) \Big\{ \Big|\sum_{t=1}^n f(U_t, Z_{ni})\Big| + 1 \Big\} \\ =&\frac{1}{m}\frac{1}{1-\varepsilon}\ k_n^{-3/2} \sum_{i=1}^{k_n} \mathbb{1}(\varepsilon b_n/2 > Z_{ni} \ge m) \Big|\sum_{t=1}^n f(U_t, Z_{ni})\Big| + O_{\mathbb{P}}(k_n^{-1/2}). \end{aligned}$$

The upper bound can now be treated exactly as in the proof of Lemma 9.1 in [3], finally yielding

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(|V_{n2}\mathbb{1}_{B_n}| > \delta) = 0. \tag{B.12}$$

It remains to treat $|V_{n1}|$. Write

$$\begin{aligned} |V_{n1}| &\le T_n(0, dk_n^{-1}) + T_n(dk_n^{-1}, dk_n^{-\mu}) + T_n(dk_n^{-\mu}, 1/m) \\ &=: T_{n1} + T_{n2} + T_{n3}, \end{aligned} \tag{B.13}$$

where, for some constant $d > 0$ and $\mu = \mu_d$ determined below,

$$T_n(a,b) = \sqrt{k_n} \int_0^\infty \mathbb{1}(x \in (a,b]) \Big| \log\Big(1 + \frac{\tilde e_n(x)}{x\sqrt{k_n}}\Big)\Big| \,\mathrm{d}\hat H_{k_n}(x).$$

We start by covering the term $T_{n1} = T_n(0, dk_n^{-1})$ and determining the constants d and μ. Note that for the event $J_n = \{\min_{i=1,\dots,k_n} Z_{ni} > dk_n^{-1}\}$ one has

$$\mathbb{P}(J_n) = \mathbb{P}\big(k_n \min_{i=1,\dots,k_n} Z_{ni} > d\big) = \mathbb{P}\big(n\ (1 - \max_{i=1,\dots,k_n} N_{ni}) > d\big) \\ = \mathbb{P}(Z_{1:n} > d) \to e^{-d\theta}.$$

Then,

$$\begin{aligned}\mathbb{P}(T_{n1} > \delta) &= \mathbb{P}(T_{n1}\mathbb{1}_{J_n} + T_{n1}\mathbb{1}_{J_n^c} > \delta) \\ &\le \mathbb{P}(T_{n1}\mathbb{1}_{J_n} > \delta/2) + \mathbb{P}(T_{n1}\mathbb{1}_{J_n^c} > \delta/2) \\ &\le \mathbb{P}(J_n^c) \to 1 - \exp(-d\theta).\end{aligned}$$

Hence, for any given $\varepsilon > 0$ we can choose $d = d(\varepsilon) < -\log(1-\varepsilon)/\theta$, such that

$$\limsup_{n\to\infty} \mathbb{P}(T_{n1} > \delta) \le \limsup_{n\to\infty} \mathbb{P}(J_n^c) = 1 - \exp(-d\theta) < \varepsilon. \tag{B.14}$$

Now, choose $\mu = \mu_d \in (1/2, 1/\{2(1-\tau)\})$ from Condition 3.4(iv), where $\tau \in (0, 1/2)$ is from Condition 3.4(ii). Next, consider $T_{n3} = T_n(dk_n^{-\mu}, 1/m)$ and note that, for $x \in (dk_n^{-\mu}, 1/m]$, we have

$$\left|\frac{\tilde e_n(x)}{x\sqrt{k_n}}\right| = \left|\frac{\tilde e_n(x)}{x^\tau}\right| \frac{1}{x^{1-\tau}\sqrt{k_n}} \le \frac{1}{d^{1-\tau}} \left|\frac{\tilde e_n(x)}{x^\tau}\right| k_n^{\mu(1-\tau)-1/2} = o_{\mathbb{P}}(1)$$

uniformly in x, by Condition 3.4(ii). As a consequence, the event

$$D_n = \left\{\left|\frac{\tilde e_n(x)}{x\sqrt{k_n}}\right| \le \frac12\right\}$$

satisfies $\mathbb{1}_{D_n^c} = o_{\mathbb{P}}(1)$, whence, recalling that $x/(1+x) \le \log(1+x) \le x$ for any $x > -1$, we have

$$\begin{aligned}T_{n3} &= \sqrt{k_n} \int_{(dk_n^{-\mu}, 1/m]} \left|\log\left(1 + \frac{\tilde e_n(x)}{x\sqrt{k_n}}\right)\right| \mathbb{1}_{D_n}\ \mathrm{d}\hat H_{k_n}(x) + o_{\mathbb{P}}(1) \\ &\le \int_{(dk_n^{-\mu}, 1/m]} \max\left\{\left|\frac{\tilde e_n(x)}{x}\right|, \left|\frac{\tilde e_n(x)}{x}\right|\left(1 + \frac{\tilde e_n(x)}{x\sqrt{k_n}}\right)^{-1}\right\}\mathbb{1}_{D_n}\ \mathrm{d}\hat H_{k_n}(x) + o_{\mathbb{P}}(1) \\ &\le 2\int_{(dk_n^{-\mu}, 1/m]} \left|\frac{\tilde e_n(x)}{x^\tau}\right| \frac{1}{x^{1-\tau}}\mathbb{1}_{D_n}\ \mathrm{d}\hat H_{k_n}(x) + o_{\mathbb{P}}(1).\end{aligned}$$

By Lemma B.15, Condition 3.4(ii) and the continuous mapping theorem, the last expression converges weakly to

$$T_3(m) = 2\int_0^{1/m} \left|\frac{e(x)}{x^\tau}\right| \frac{1}{x^{1-\tau}}\ \mathrm{d}H(x).$$

As a consequence,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(T_{n3} \ge \delta) \le \lim_{m\to\infty} \mathbb{P}(T_3(m) > \delta) = 0. \tag{B.15}$$

Finally, regarding T_{n2}, note that, for $x \in (dk_n^{-1}, dk_n^{-\mu})$,

$$\frac{\tilde{e}_n(x)}{x\sqrt{k_n}} \leq \frac{1}{x}\Big(\frac{1}{k_n} + \frac{1}{k_n}\sum_{i=1}^n \mathbb{1}(U_i > 1 - x/b_n)\Big) \leq \frac{n+1}{d},$$
$$\frac{\tilde{e}_n(x)}{x\sqrt{k_n}} \geq \frac{1}{x}\Big(\frac{1}{k_n} - \frac{1}{k_n}\sum_{i=1}^n x/b_n\Big) \geq \frac{1}{dk_n^{1-\mu}} - 1,$$

which implies

$$\begin{aligned}
&\Big|\log\Big(1 + \frac{\tilde{e}_n(x)}{x\sqrt{k_n}}\Big)\Big| \\
&= \log\Big(1 + \frac{\tilde{e}_n(x)}{x\sqrt{k_n}}\Big)\mathbb{1}\Big(\frac{\tilde{e}_n(x)}{x\sqrt{k_n}} > 0\Big) - \log\Big(1 + \frac{\tilde{e}_n(x)}{x\sqrt{k_n}}\Big)\mathbb{1}\Big(\frac{\tilde{e}_n(x)}{x\sqrt{k_n}} < 0\Big) \\
&\leq \log\big((n+1)d^{-1} + 1\big) + \log\big(dk_n^{1-\mu}\big) \\
&\lesssim \log(n).
\end{aligned}$$

As a consequence, the term $T_{n2} = T_n(dk_n^{-1}, dk_n^{-\mu})$ can be bounded as follows

$$T_{n2} \lesssim \log(n)\ \sqrt{k_n}\int_{(dk_n^{-1}, dk_n^{-\mu}]} \mathrm{d}\hat{H}_{k_n}(x) = \frac{\log(n)}{\sqrt{k_n}}\sum_{i=1}^{k_n}\mathbb{1}\big(Z_{ni} \in (dk_n^{-1}, dk_n^{-\mu}]\big).$$

Hence, by Condition 3.4(iv),

$$\begin{aligned}
E\left[T_{n2}\right] &\lesssim \log(n)\ \sqrt{k_n}\mathbb{P}(Z_{n1} < dk_n^{-\mu}) \\
&= \log(n)\ \sqrt{k_n}\{1 - \exp(-\theta dk_n^{-\mu})\} + o(1) \\
&= \theta d\log(n)\ k_n^{1/2-\mu}\{1 + o(1)\} + o(1) \\
&= O(\log(k_n)k_n^{1/2-\mu}) = o(1), \qquad \text{(B.16)}
\end{aligned}$$

where the last line follows from $\log n = \log k_n + \log b_n \lesssim (1+q)\log k_n$ by Condition 3.4(i).

The assertion follows from (B.10), combined with (B.12), (B.13), (B.14), (B.15) and (B.16). □

B.2. Auxiliary lemmas – sliding blocks

Throughout this section, we assume that Condition 3.1, 3.2(i) and 3.3(i) are met.

Lemma B.7. *For any $x_1, \dots, x_m \in [0, \infty)$ and $m \in \mathbb{N}$, we have*

$$(e_n(x_1), \dots, e_n(x_m), B_n^{\mathrm{sb}})' \xrightarrow{d} (e(x_1), \dots, e(x_m), B^{\mathrm{sb}})',$$

where $(e(x_1), \dots, e(x_m), B^{\mathrm{sb}})' \sim \mathcal{N}_{m+1}(0, \Sigma_{m+1}^{\mathrm{sb}})$ *with*

$$\Sigma_{m+1}^{\mathrm{sb}} = \begin{pmatrix} r(x_1, x_1) & \dots & r(x_1, x_m) & f(x_1) \\ \vdots & \ddots & \vdots & \vdots \\ r(x_m, x_1) & \dots & r(x_m, x_m) & f(x_m) \\ f(x_1) & \dots & f(x_m) & 8\log(2) - 4 \end{pmatrix}.$$

Here, the functions r *and* f *are defined as in Lemma B.1.*

Lemma B.8. *For any* $m \in \mathbb{N}$*, we have*

$$\left\{(W_n(x), B_n^{\mathrm{sb}})'\right\}_{x\in[1/m,m]} \xrightarrow{d} \left\{\left(\frac{e(x)}{x}, B^{\mathrm{sb}}\right)'\right\}_{x\in[1/m,m]} \quad \text{in } D([1/m,m]) \times \mathbb{R},$$

where $(e, B^{\mathrm{sb}})'$ *is a centered Gaussian process with continuous sample paths and with covariance functional as specified in Lemma B.7.*

Lemma B.9. *For any* $m \in \mathbb{N}$*, we have*

$$E_{n,m}^{\mathrm{sb}} = E'_{n,m} + o_{\mathbb{P}}(1) \qquad \text{as } n \to \infty,$$

where $E'_{n,m} = \int_{1/m}^m W_n(x)\theta e^{-\theta x}\,\mathrm{d}x$ *is as in Lemma B.3.*

Lemma B.10. *For any* $m \in \mathbb{N}$*, we have*

$$E_{n,m}^{\mathrm{sb}} + B_n^{\mathrm{sb}} \xrightarrow{d} E'_m + B^{\mathrm{sb}} \sim \mathcal{N}(0, \tau_{\mathrm{sb},m}^2) \quad \text{as } n \to \infty,$$

where, with r *and* f *defined as in Lemma B.1,*

$$\tau_{\mathrm{sb},m}^2 = \theta^2 \int_{1/m}^m \int_{1/m}^m r(x,y)\frac{1}{xy}e^{-\theta(x+y)}\,\mathrm{d}x\mathrm{d}y \\ + 2\theta \int_{1/m}^m f(x)\frac{1}{x}e^{-\theta x}\,\mathrm{d}x + 8\log(2) - 4.$$

Lemma B.11. *As* $m \to \infty$*,* $\tau_{\mathrm{sb},m}^2 \to \sigma_{\mathrm{sb,(C)}}^2/\theta^2$*, where* $\sigma_{\mathrm{sb,(C)}}^2$ *is specified in Theorem 4.1.*

Lemma B.12. *If in addition, Condition 3.4 holds, then, for all* $\delta > 0$*,*

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(|E_{n,m}^{\mathrm{sb}} - E_n^{\mathrm{sb}}| > \delta\right) = 0.$$

Proof of Lemma B.7. As in the proof of Lemma B.1 we only show joint weak convergence of $(e_n(x), B_n^{\mathrm{sb}})$ for some fixed $x > 0$; the general case can be shown analogously. For given $\varepsilon \in (0, c_1 \wedge c_2)$ let $A'_n = \{\min_{t=1,\dots,n-b_n+1} N_{nt}^{\mathrm{sb}} > 1 - \varepsilon\}$, such that $\mathbb{P}(A_n) \to 1$ by Condition 3.1(v). By the Cramér-Wold device, it suffices to prove weak convergence of

$$\lambda_1 e_n(x) + \lambda_2 B_n^{\mathrm{sb}} = \sum_{j=1}^{k_n-1} \sum_{s\in I_j} \Big[\frac{\lambda_1}{\sqrt{k_n}}\Big\{\mathbb{1}\Big(U_s > 1-\frac{x}{b_n}\Big) - \frac{x}{b_n}\Big\} + \frac{\lambda_2\sqrt{k_n}}{n-b_n+1}\Big\{\log(Z_{ns}^{\mathrm{sb}}) - E[\log(Z_{ns}^{\mathrm{sb}})]\Big\}\Big] + o_{\mathbb{P}}(1),$$

for some arbitrary $\lambda_1, \lambda_2 \in \mathbb{R}$, where the negligible term stems from omitting a negligible number of summands.

We are going to apply a big block-small block argument, based on a suitable 'blocking of blocks' to take care of the serial dependence introduced through the use of sliding blocks. For that purpose, let $k_n^* < k_n$ be an integer sequence with $k_n^* \to \infty$ and $k_n^* = o(k_n^{\delta/(2(1+\delta))})$, where δ is from Condition 3.1(ii). For $q_n^* = \lfloor k_n/(k_n^*+2)\rfloor$ and $j = 1, \dots, q_n^*$, define

$$J_j^+ = \bigcup_{i=(j-1)(k_n^*+2)+1}^{j(k_n^*+2)-2} I_i \quad \text{and} \quad J_j^- = I_{j(k_n^*+2)-1} \cup I_{j(k_n^*+2)}.$$

Thus we have q_n^* big blocks J_j^+ of size $k_n^* b_n$, which are separated by a small block J_j^- of size $2b_n$, just as in the construction in the proof of Lemma 10.3 in [3]. Consequently, we have $\lambda_1 e_n(x) + \lambda_2 B_n^{\mathrm{sb}} = L_n^+ + L_n^- + o_{\mathbb{P}}(1)$, where

$$L_n^{\pm} = \frac{1}{\sqrt{q_n^*}} \sum_{j=1}^{q_n^*} W_{nj}^{\pm}$$

with

$$W_{nj}^{\pm} = \sqrt{\frac{q_n^*}{k_n}} \sum_{s\in J_j^{\pm}} \lambda_1\Big\{\mathbb{1}\Big(U_s > 1-\frac{x}{b_n}\Big) - \frac{x}{b_n}\Big\} + \frac{\lambda_2 n}{n-b_n+1}\frac{1}{b_n}\Big\{\log(Z_{ns}^{\mathrm{sb}}) - E[\log(Z_{ns}^{\mathrm{sb}})]\Big\}$$

for $j = 1, \dots, q_n^*$. In the following, we show that, on the one hand, $L_n^- \mathbb{1}_{A_n'} = o_{\mathbb{P}}(1)$ and that, on the other hand, $L_n^+ \mathbb{1}_{A_n'}$ converges to the claimed normal distribution. First, we cover $L_n^- \mathbb{1}_{A_n'}$. As in the proof of Lemma B.1, we have

$$Z_{ns}^{\mathrm{sb}} = b_n\Big(1 - \max_{t=s,\dots,s+b_n-1} U_t\Big) = b_n\Big(1 - \max_{t=s,\dots,s+b_n-1} U_t^{\varepsilon}\Big) =: Z_{ns}^{\varepsilon,\mathrm{sb}}$$

on the event A_n', where $U_t^{\varepsilon} = U_t \mathbb{1}(U_t > 1-\varepsilon)$. Hence, we can write $L_n^- \mathbb{1}_{A_n'} = \tilde{L}_n^- \mathbb{1}_{A_n'} + o_{\mathbb{P}}(1) = \tilde{L}_n^- + o_{\mathbb{P}}(1)$ with

$$\tilde{L}_n^- = \frac{1}{\sqrt{q_n^*}} \sum_{j=1}^{q_n^*} W_{nj}^{\varepsilon-},$$

where

$$W_{nj}^{\varepsilon-} = \sqrt{\frac{q_n^*}{k_n}} \sum_{s \in J_j^-} \lambda_1 \Big\{ \mathbb{1}\big(U_s^\varepsilon > 1 - \frac{x}{b_n}\big) - \frac{x}{b_n} \Big\} + \frac{\lambda_2 n}{n - b_n + 1} \frac{1}{b_n} \Big\{ \log(Z_{ns}^{\varepsilon,\mathrm{sb}}) - E[\log(Z_{ns}^{\varepsilon,\mathrm{sb}})] \Big\}.$$

We proceed by showing that $\mathrm{Var}[\tilde{L}_n^-] = o(1)$. By stationarity, one has

$$\mathrm{Var}[\tilde{L}_n^-] = \mathrm{Var}[W_{n1}^{\varepsilon-}] + \frac{2}{q_n^*} \sum_{j=1}^{q_n^*} (q_n^* - j) \,\mathrm{Cov}\left(W_{n1}^{\varepsilon-}, W_{n,j+1}^{\varepsilon-}\right),$$

which is bounded by $3\,\mathrm{Var}[W_{n1}^{\varepsilon-}] + 2\sum_{j=2}^{q_n^*} |\,\mathrm{Cov}\left(W_{n1}^{\varepsilon-}, W_{n,j+1}^{\varepsilon-}\right)|$ in absolute value. First, we show $\mathrm{Var}[W_{n1}^{\varepsilon-}] = o(1)$, for which it suffices to show that $||W_{n1}^{\varepsilon-}||_p = o(1)$ for some $p \in (2, 2+\delta)$. By Minkowski's inequality, one has

$$||W_{n1}^{\varepsilon-}||_p \le 2\sqrt{\frac{q_n^*}{k_n}} \left[|\lambda_1| \; ||N_{b_n}^{(x)}(E)||_p + |\lambda_2| \; ||\log(Z_{n1}^{\varepsilon,\mathrm{sb}}) - \mathrm{E}[\log(Z_{n1}^{\varepsilon,\mathrm{sb}})]||_p \right] \tag{B.17}$$
$$= O(\sqrt{q_n^*/k_n}) = o(1)$$

by Condition 3.1(ii) and 3.2(i). It remains to treat the sum over the covariances. Since $W_{nj}^{\varepsilon-}$ is $\mathcal{B}^{\varepsilon}_{\{(j(k_n^*+2)-2)b_n+1\}:\{j(k_n^*+2)b_n\}}$- measurable, we may apply Lemma 3.11 in [9] to obtain

$$|\,\mathrm{Cov}(W_{n1}^{\varepsilon-}, W_{n,j+1}^{\varepsilon-})| \le 10 \; ||W_{n1}^{\varepsilon-}||_p^2 \; \alpha_{c_2}(jk_n^* b_n)^{1-2/p}.$$

By Condition 3.1(iii), the sum $\sum_{j=2}^{q_n^*} \alpha_{c_2}(jk_n^* b_n)^{1-2/p}$ converges to zero, hence $||W_{n1}^{\varepsilon-}||_p = o(1)$ as asserted.

Let us now treat the term $L_n^+ \mathbb{1}_{A_n'}$ and show weak convergence to the asserted normal distribution. One can write

$$L_n^+ \mathbb{1}_{A_n'} = \frac{1}{\sqrt{q_n^*}} \sum_{j=1}^{q_n^*} \tilde{W}_{nj}^+ + o_{\mathbb{P}}(1), \qquad \tilde{W}_{nj}^+ = W_{nj}^+ \mathbb{1}\big(\max_{t \in J_j^+} Z_{nt}^{\mathrm{sb}} < \varepsilon b_n\big).$$

A standard argument based on characteristic functions shows that the weak limit of ${q_n^*}^{-1/2} \sum_{j=1}^{q_n^*} \tilde{W}_{nj}^+$ is the same as if the summands were independent. By arguments as before, we may also pass back to an independent sample W_{nj}^+, $j = 1, \dots, q_n^*$. The assertion then follows from Ljapunov's central limit theorem, once we have shown the Ljapunov condition.

For that purpose, note that $||W_{nj}^+||_{2+\delta} = O(\sqrt{q_n^* k_n}) = O(\sqrt{k_n^*})$ by similar arguments as in (B.17) such that $E[|W_{nj}^+|^{2+\delta}] = O({k_n^*}^{(2+\delta)/2})$. As a consequence,

$$\frac{\sum_{j=1}^{q_n^*} E[|W_{nj}^+|^{2+\delta}]}{\mathrm{Var}\left[\sum_{j=1}^{q_n^*} W_{nj}^+\right]^{\frac{2+\delta}{2}}} = {q_n^*}^{-\frac{\delta}{2}} \frac{E[|W_{n1}^+|^{2+\delta}]}{E[|W_{n1}^+|^2]^{\frac{2+\delta}{2}}} = O(k_n^{-\delta/2} {k_n^*}^{1+\delta}) = o(1),$$

since $k_n^* = o(k_n^{\delta/(2(1+\delta))})$ by construction and provided that the limit of $\mathrm{E}[|W_{n1}^+|^2]$ exists. If it does, we can conclude that $L_n^+ \xrightarrow{d} \mathcal{N}(0, \lim_{n\to\infty} E[|W_{n1}^+|^2])$. and it suffices to show that

$$\lim_{n\to\infty} E[|W_{n1}^+|^2] = \lambda_1^2 r(x,x) + 2\lambda_1\lambda_2 f(x) + \lambda_2^2\{8\log(2) - 4\}.$$

To this, note that $W_{n1}^+ = \lambda_1 e_{n^*}(x) + \lambda_2 B_{n^*}^{\mathrm{sb}} + o_{\mathbb{P}}(1)$, where e_{n^*} and $B_{n^*}^{\mathrm{sb}}$ are defined as e_n and B_n^{sb} with n replaced by $n^* = k_n^* b_n$ and k_n by k_n^*; and our general conditions still hold with this replacement. The result follows from Lemma B.13 and Lemma B.14 and the proof of Theorem 4.1 in [27]. □

Proof of Lemma B.8. Up to notation, the proof is exactly the same as the one of Lemma B.2 in the disjoint blocks case. □

Proof of Lemma B.9. The result follows immediately from the argument in the proof of Lemma B.3 and the proof of Lemma 10.2 in [3]. □

Proof of Lemma B.10. Up to notation, the proof is exactly the same as the one of Lemma B.4 in the disjoint blocks case. □

Proof of Lemma B.11. By the definition of τ_m^2 and $\tau_{\mathrm{sb},m}^2$ in Lemma B.4 and B.10, we have

$$\tau_{\mathrm{sb},m}^2 = \tau_m^2 - \pi^2/6 + 8\log(2) - 4.$$

Hence, by the proof of Lemma B.5 and the definition of $\sigma_{\mathrm{sb,C}}^2$ in Theorem 4.1,

$$\lim_{m\to\infty} \tau_{\mathrm{sb},m}^2 = \sigma_{\mathrm{db,C}}^2/\theta^2 - \pi^2/6 + 8\log 2 - 4 = \sigma_{\mathrm{sb,C}}^2/\theta^2. \qquad \square$$

Proof of Lemma B.12. The proof is similar to the one of Lemma B.6, which is why we keep it short. Write $|E_{n,m}^{\mathrm{sb}} - E_n^{\mathrm{sb}}| \le |V_{n1}| + |V_{n2}|$ with

$$V_{n1} = \int_0^\infty \log\left(1 + \frac{\tilde e_n(x)}{x\sqrt{k_n}}\right)\sqrt{k_n}\,\mathbb{1}_{(0,1/m]}(x)\,\mathrm{d}\hat H_n^{\mathrm{sb}}(x),$$
$$V_{n2} = \int_0^\infty \log\left(1 + \frac{\tilde e_n(x)}{x\sqrt{k_n}}\right)\sqrt{k_n}\,\mathbb{1}_{[m,\infty)}(x)\,\mathrm{d}\hat H_n^{\mathrm{sb}}(x),$$

where $\tilde e_n(x) = e_n(x) + k_n^{-1/2}$. For some $\varepsilon > 0$ define the event

$$B_n = \Big\{ \max_{Z_{ni}^{\mathrm{sb}}\ge m} \Big| \frac{\tilde e_n(Z_{ni}^{\mathrm{sb}})}{Z_{ni}^{\mathrm{sb}}\sqrt{k_n}} \Big| \le \varepsilon \Big\},$$

such that $\mathbb{P}(B_n) \to 1$ by Condition 3.4(iii). As in the proof of Lemma B.6, with f defined in (B.11), we can write

$$V_{n2}\mathbb{1}_{B_n} = k_n^{-3/2} \sum_{i=1}^{k_n-1} \sum_{w\in I_i} \frac{1}{Z_{nw}^{\mathrm{sb}}} \mathbb{1}(Z_{nw}^{\mathrm{sb}} \ge m) \int_0^1 \frac{1}{1 + s\,\frac{\tilde e_n(Z_{nw}^{\mathrm{sb}})}{Z_{nw}^{\mathrm{sb}}\sqrt{k_n}}}\,\mathrm{d}s$$

$$\times\, b_n^{-1}\Big\{\sum_{j=1}^{k_n}\sum_{t\in I_j} f(U_t, Z_{nw}^{\mathrm{sb}}) + 1\Big\}\mathbb{1}_{B_n} + o_{\mathbb{P}}(1).$$

By Condition 3.1(v), $\mathbb{P}(C_n)\to 1$ where $C_n = \big\{\min_{i=1,\dots,n-b_n+1} N_{ni}^{\mathrm{sb}} > 1-\varepsilon/2\big\}$. Hence, $V_{n2}\mathbb{1}_{B_n} = \bar V_{n2}\mathbb{1}_{B_n}\mathbb{1}_{C_n} + o_{\mathbb{P}}(1)$, where

$$\bar V_{n2} = k_n^{-3/2}\sum_{i=1}^{k_n-1}\sum_{w\in I_i}\frac{1}{Z_{nw}^{\mathrm{sb}}}\mathbb{1}(\varepsilon b_n/2 > Z_{nw}^{\mathrm{sb}} \ge m)\int_0^1 \frac{1}{1+s\,\frac{\tilde e_n(Z_{nw}^{\mathrm{sb}})}{Z_{nw}^{\mathrm{sb}}\sqrt{k_n}}}\,\mathrm{d}s \\ \times\, b_n^{-1}\Big\{\sum_{j=1}^{k_n}\sum_{t\in I_j} f(U_t, Z_{nw}^{\mathrm{sb}}) + 1\Big\},$$

such that $\bar V_{n2}$ can be bounded as in the proof of Lemma B.6 as follows

$$|\bar V_{n2}\mathbb{1}_{B_n}| \le \frac{1}{m}\frac{1}{1-\varepsilon}k_n^{-3/2}\sum_{i=1}^{k_n-1}\sum_{w\in I_i}\mathbb{1}(\varepsilon b_n/2 > Z_{nw}^{\mathrm{sb}} \ge m) \\ \times\, b_n^{-1}\Big|\sum_{j=1}^{k_n}\sum_{t\in I_j} f(U_t, Z_{nw}^{\mathrm{sb}})\Big| + o_{\mathbb{P}}(1).$$

This expression can be handled as in the proof of Lemma 10.1 in [3], such that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(|\bar V_{n2}\mathbb{1}_{B_n}\mathbb{1}_{C_n}| > \delta) = 0.$$

The remaining term $|V_{n1}|$ can be treated analogously to the eponymous term in the proof of Lemma B.6. □

Lemma B.13. *(a) For $x \ge 0$, as $n\to\infty$,*

$$\mathrm{Cov}(e_n(x), B_n^{\mathrm{sb}}) \to 2\int_0^1 h_{\mathrm{sb},x}(\xi)\,\mathrm{d}\xi - 2x\varphi_{(\mathrm{C})}(\theta),$$

where

$$h_{\mathrm{sb},x}(\xi) = \sum_{i=1}^{\infty} i\int_0^{\infty}\mathbb{1}(y\le\log(x))\sum_{l=0}^{i} p^{(\xi x)}(l)\,p_2^{((1-\xi)x,(1-\xi)e^y)}(i-l,0)\,e^{-\theta\xi e^y} \\ + \mathbb{1}(y>\log(x))\;p^{(\xi x)}(i)\;e^{-\theta e^y}\;\mathrm{d}y \\ -\sum_{i=1}^{\infty} i\int_{-\infty}^{0} p^{(x)}(i) - \mathbb{1}(y\le\log(x))\sum_{l=0}^{i} p^{(\xi x)}(l)\,p_2^{((1-\xi)x,(1-\xi)e^y)}(i-l,0) \\ \times\, e^{-\theta\xi e^y} \\ - \mathbb{1}(y>\log(x))\;p^{(\xi x)}(i)\;e^{-\theta e^y}\;\mathrm{d}y.$$

(b) We have

$$2\int_0^1 h_{\mathrm{sb},x}(\xi)\;\mathrm{d}\xi = h(x) + x\varphi_{(\mathrm{C})}(\theta),$$

where h is defined in Lemma B.1.

Proof. (a) We assume that both U_s and Z_{nt}^{sb} are measurable with respect to the appropriate $\mathcal{B}_{:,:}^{\varepsilon}$ sigma-algebra; the general case can be treated by multiplying with suitable indicator functions as in the proof of Lemma B.7. Let $A_j = \sum_{s\in I_j} \mathbb{1}\big(U_s > 1 - \frac{x}{b_n}\big)$ and $D_j = \sum_{s\in I_j} \log(Z_{ns}^{\mathrm{sb}})$. Then

$$\operatorname{Cov}(e_n(x), B_n^{\mathrm{sb}}) = \frac{1}{n-b_n+1}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n-1}\operatorname{Cov}(A_i, D_j) + \frac{1}{n-b_n+1}\sum_{i=1}^{k_n}\operatorname{Cov}(A_i, \log(Z_{n,n-b_n+1}^{\mathrm{sb}})).$$

The second sum is asymptotically negligible, since $||A_j||_2 = ||N_{b_n}^{(x)}(E)||_2 = O(1)$ and $||\log(Z_{n,n-b_n+1}^{\mathrm{sb}})||_2 = O(1)$ by Condition 3.1(ii) and 3.2(i). Next, following the argument in the proof of Lemma B.1 in [3], we may write

$$\begin{aligned}\operatorname{Cov}(e_n(x), B_n^{\mathrm{sb}}) &= \frac{1}{b_n}\operatorname{Cov}(A_2, D_1 + D_2) + o(1)\\ &= \frac{1}{b_n}\sum_{t=1}^{2b_n}\operatorname{Cov}\Big\{\sum_{s\in I_2}\mathbb{1}\big(U_s > 1-\frac{x}{b_n}\big), \log(Z_{nt}^{\mathrm{sb}})\Big\} + o(1)\\ &= \int_0^1 f_n(\xi) + g_n(\xi)\,\mathrm{d}\xi - 2x\,\mathrm{E}\left[\log(Z_{n1}^{\mathrm{sb}})\right] + o(1),\end{aligned}$$

where

$$f_n(\xi) = \sum_{t=1}^{b_n}\mathrm{E}\Big[\sum_{s\in I_2}\mathbb{1}\big(U_s > 1-\tfrac{x}{b_n}\big)\log(Z_{nt}^{\mathrm{sb}})\Big]\mathbb{1}\big\{\xi\in[\tfrac{t-1}{b_n},\tfrac{t}{b_n})\big\},$$

$$g_n(\xi) = \sum_{t=b_n+1}^{2b_n}\mathrm{E}\Big[\sum_{s\in I_2}\mathbb{1}\big(U_s > 1-\tfrac{x}{b_n}\big)\log(Z_{nt}^{\mathrm{sb}})\Big]\mathbb{1}\big\{\xi\in[\tfrac{t-b_n-1}{b_n},\tfrac{t-b_n}{b_n})\big\}.$$

Note that $\lim_{n\to\infty}\mathrm{E}[\log(Z_{n1}^{\mathrm{sb}})] = \varphi_{(\mathrm{C})}(\theta)$ by uniform integrability of $\log(Z_{1:n})$, and that $\sup_{n\in\mathbb{N}}||f_n||_\infty + ||g_n||_\infty < \infty$ as a consequence of $\|\sum_{s\in I_1}\mathbb{1}\big(U_s > 1-\frac{x}{b_n}\big)\|_2 \times \|\log(Z_{n1}^{\mathrm{sb}})\|_2 < \infty$ by Condition 3.1(ii) and 3.2(i). Hence, the lemma is proven if we show that, for any $\xi\in(0,1)$,

$$\lim_{n\to\infty} f_n(1-\xi) = \lim_{n\to\infty} g_n(\xi) = h_{\mathrm{sb},x}(\xi).$$

Since the proof for $f_n(1-\xi)$ is similar, we only treat $g_n(\xi)$, which can be written as

$$g_n(\xi) = \mathrm{E}\Big[\sum_{s\in I_2}\mathbb{1}\big(U_s > 1-\tfrac{x}{b_n}\big)\log(Z_{n,\lfloor(1+\xi)b_n\rfloor+1}^{\mathrm{sb}})\Big].$$

Let us proceed by showing joint weak convergence of $\sum_{s\in I_2} \mathbb{1}(U_s > 1-\frac{x}{b_n})$ and $\log(Z^{\mathrm{sb}}_{n,\lfloor(1+\xi)b_n\rfloor+1})$. For that purpose, note that

$$G_n(i,y) := \mathbb{P}\Big(\sum_{s\in I_2} \mathbb{1}\big(U_s > 1-\tfrac{x}{b_n}\big) = i,\ \log(Z^{\mathrm{sb}}_{n,\lfloor(1+\xi)b_n\rfloor+1}) \ge y\Big)$$
$$= \mathbb{P}\Big(\sum_{s\in I_2} \mathbb{1}\big(U_s > 1-\tfrac{x}{b_n}\big) = i,\ Z^{\mathrm{sb}}_{n,\lfloor(1+\xi)b_n\rfloor+1} \ge e^y\Big),$$

coincides with $F_n(i,e^y)$ in the proof of Lemma B.1 in [3]. Hence, by that proof, we have

$$\lim_{n\to\infty} G_n(i,y) = \sum_{l=0}^{i} p^{(\xi x)}(l) p_2^{((1-\xi)x,(1-\xi)e^y)}(i-l,0)e^{-\theta\xi e^y}$$

for $y \le \log x$ and

$$\lim_{n\to\infty} G_n(i,y) = p^{(\xi x)}(i)\ e^{-\theta e^y}$$

for $y > \log x$. Further, note that

$$\lim_{n\to\infty} \mathbb{P}\big(N_{b_n}^{(x)}(E) = i\big) = p^{(x)}(i).$$

As a consequence of the previous three displays, and since weak convergence and uniform integrability implie convergence of moments, we have

$$g_n(\xi) = \sum_{i=1}^{\infty} i \int_0^{\infty} \mathbb{P}\Big(\sum_{s=b_n+1}^{2b_n} \mathbb{1}\big(U_s > 1-\frac{x}{b_n} = i\big), \log(Z^{\mathrm{sb}}_{n,\lfloor(1+\xi)b_n\rfloor+1}) \ge y\Big)\ \mathrm{d}y$$
$$- i\int_{-\infty}^{0} \mathbb{P}\Big(\sum_{s=b_n+1}^{2b_n} \mathbb{1}\big(U_s > 1-\frac{x}{b_n}\big) = i, \log(Z^{\mathrm{sb}}_{n,\lfloor(1+\xi)b_n\rfloor+1}) \le y\Big)\ \mathrm{d}y$$
$$= \sum_{i=1}^{\infty} i\int_0^{\infty} G_n(i,y)\ \mathrm{d}y - i\int_{-\infty}^{0} \mathbb{P}\big(N_{b_n}^{(x)}(E) = i\big) - G_n(i,y)\ \mathrm{d}y$$
$$\to h_{\mathrm{sb},x}(\xi)$$

as asserted, which implies part (a) of the lemma.

(b) In the proof of Lemma B.3 in [3] it is shown that, for $y \le \log(x)$,

$$S(x,y,\xi) = e^{-\theta\xi e^y} \sum_{i=1}^{\infty} i \sum_{l=0}^{i} p^{(\xi x)}(l)\ p_2^{((1-\xi)x,(1-\xi)e^y)}(i-l,0)$$
$$= \xi x e^{-\theta e^y} + \mathrm{E}\left[\xi_{11}^{(e^y/x)} \mathbb{1}(\xi_{12}^{(e^y/x)} = 0)\right] \theta(1-\xi)xe^{-\theta e^y},$$

where $(\xi_{11}^{(y/x)}, \xi_{12}^{(y/x)}) \sim \pi_2^{(y/x)}$. Equation (B.8) then allows to rewrite

$$S(x,y,\xi) = \xi x e^{-\theta e^y} + (1-\xi)\sum_{i=1}^{\infty} i p_2^{(x,e^y)}(i,0) \equiv \xi x e^{-\theta e^y} + (1-\xi)T(x,y).$$

As a consequence, further noting that $\sum_{i=1}^\infty i\, p^{(\xi x)}(i) = \xi x$, we obtain

$$\begin{aligned} h_{\mathrm{sb},x}(\xi) = \int_0^\infty \xi x e^{-\theta e^y} + \mathbb{1}(y \le \log(x))(1-\xi) T(x,y)\,\mathrm{d}y \\ - \int_{-\infty}^0 x - \xi x e^{-\theta e^y} - \mathbb{1}(y \le \log(x))(1-\xi)T(x,y)\,\mathrm{d}y. \end{aligned}$$

Then, by Fubinbi's theorem,

$$\begin{aligned} 2\int_0^1 h_{\mathrm{sb},x}(\xi)\ \mathrm{d}\xi = \int_0^\infty x e^{-\theta e^y} + \mathbb{1}(y \le \log x) T(x,y)\,\mathrm{d}y \\ - \int_{-\infty}^0 x(1-e^{-\theta e^y}) + x - \mathbb{1}(y \le \log(x))T(x,y)\,\mathrm{d}y. \end{aligned}$$

The assertion now follows from the fact that

$$\int_0^\infty e^{-\theta e^y}\,\mathrm{d}y = \int_\theta^\infty \frac{e^{-z}}{z}\,\mathrm{d}z = -\operatorname{Ei}(-\theta)$$

and

$$\begin{aligned} \int_{-\infty}^0 1 - e^{-\theta e^y}\,\mathrm{d}y &= \int_0^\theta \frac{1-e^{-z}}{z}\,\mathrm{d}z = (1-e^{-z})\log(z)\big|_0^\theta - \int_0^\theta e^{-z}\log(z)\,\mathrm{d}z \\ &= \log(\theta) - e^{-\theta}\log(\theta) - \Big\{\gamma - \int_\theta^\infty e^{-z}\log(z)\,\mathrm{d}z\Big\} \\ &= \log(\theta) - e^{-\theta}\log(\theta) - \gamma + \Big\{ - e^{-z}\log(z)\big|_\theta^\infty + \int_\theta^\infty \frac{e^{-z}}{z}\,\mathrm{d}z\Big\} \\ &= \log(\theta) + \gamma - \operatorname{Ei}(-\theta) = -\varphi_{(\mathrm{C})}(\theta) - \operatorname{Ei}(-\theta) \end{aligned}$$

after assembling terms, where $\operatorname{Ei}(x) = -\int_{-x}^\infty e^{-t}/t\ \mathrm{d}t$ for $x > 0$ is the exponential integral. □

Lemma B.14. *One has*

$$\lim_{n\to\infty} \operatorname{Var}(B_n^{\mathrm{sb}}) = 8\log(2) - 4 \approx 1.545.$$

Proof. As in the proof of Lemma B.13, we assume that the Z_{nt}^{sb} are measurable with respect to the appropriate $\mathcal{B}^{\varepsilon}_{\cdot,\cdot}$ sigma-algebra. We may then argue as in that proof to obtain

$$\begin{aligned} \operatorname{Var}(B_n^{\mathrm{sb}}) &= \frac{2}{b_n}\sum_{t=1}^{b_n} \mathrm{E}\left[\log(Z_{n1}^{\mathrm{sb}})\log(Z_{n,1+t}^{\mathrm{sb}})\right] - 2\,\mathrm{E}[\log(Z_{n1}^{\mathrm{sb}})]^2 + o(1) \\ &= 2\int_0^1 f_n(\xi)\ \mathrm{d}\xi - 2\,\mathrm{E}[\log(Z_{n1}^{\mathrm{sb}})]^2 + o(1), \end{aligned} \tag{B.18}$$

where $f_n : [0,1] \to \mathbb{R}$ is defined as

$$f_n(\xi) = \sum_{t=1}^{b_n} \mathrm{E}[\log(Z_{n1}^{\mathrm{sb}}) \log(Z_{n,1+t}^{\mathrm{sb}})]\, \mathbb{1}\big(\xi \in \big[\tfrac{t-1}{b_n}, \tfrac{t}{b_n}\big)\big)$$
$$= \mathrm{E}[\log(Z_{n1}^{\mathrm{sb}}) \log(Z_{n,\lfloor b_n\xi\rfloor+1}^{\mathrm{sb}})].$$

By Condition 3.2(i), we have $\mathrm{E}[\log(Z_{n1}^{\mathrm{sb}})] \to \varphi_{(\mathrm{C})}(\theta)$. Further,

$$\sup_{n\in\mathbb{N}} ||f_n||_\infty \le \sup_{n\in\mathbb{N}} E[\log(Z_{n1}^{\mathrm{sb}})^2] < \infty,$$

whence convergence of the integral over f_n in (B.18) may be concluded from the dominated convergence theorem, once we have shown pointwise convergence of f_n. To this end we show that, for any fixed $\xi \in (0,1)$,

$$\big(\log(Z_{n1}^{\mathrm{sb}}),\ \log(Z_{n,\lfloor b_n\xi\rfloor+1}^{\mathrm{sb}}\big) \xrightarrow{d} \big(X^{(\xi)}, Y^{(\xi)}\big) \tag{B.19}$$

for some random vector $\big(X^{(\xi)}, Y^{(\xi)}\big)$. This in turn will imply

$$\lim_{n\to\infty} f_n(\xi) = \lim_{n\to\infty} \mathrm{E}[\log(Z_{n1}^{\mathrm{sb}})\ \log(Z_{n,\lfloor b_n\xi\rfloor+1}^{\mathrm{sb}})] = \mathrm{E}[X^{(\xi)}Y^{(\xi)}]$$

by Condition 3.2(i) and therefore

$$\lim_{n\to\infty} \mathrm{Var}(B_n^{\mathrm{sb}}) = 2\int_0^1 \mathrm{E}[X^{(\xi)}Y^{(\xi)}]\ \mathrm{d}\xi - 2\varphi_{(\mathrm{C})}(\theta)^2 = 2\int_0^1 \mathrm{Cov}(X^{(\xi)}, Y^{(\xi)})\ \mathrm{d}\xi. \tag{B.20}$$

For the proof of (B.19), define, for $x, y \in \mathbb{R}$,

$$G_{n,\xi}(x,y) = \mathbb{P}\big(\log(Z_{n1}^{\mathrm{sb}}) > x,\ \log(Z_{n,\lfloor b_n\xi\rfloor+1}^{\mathrm{sb}}) > y\big)$$
$$= \mathbb{P}\big(Z_{n1}^{\mathrm{sb}} > e^x,\ Z_{n,\lfloor b_n\xi\rfloor+1}^{\mathrm{sb}} > e^y\big),$$

which converges to

$$G_\xi(x,y) = \exp\big(-\theta\big[\xi(e^x \wedge e^y) + (e^x \vee e^y)\big]\big)$$

by the proof of Lemma B.2 in [3]. Hence, (B.19), where the random vector $(X^{(\xi)}, Y^{(\xi)})$ has joint c.d.f.

$$F_\xi(x,y) = \mathbb{P}\big(X^{(\xi)} \le x,\ Y^{(\xi)} \le y\big)$$
$$= 1 - \mathbb{P}\big(X^{(\xi)} > x\big) - \mathbb{P}\big(Y^{(\xi)} > y\big) + G_\xi(x,y),$$
$$= 1 - \exp(-\theta e^x) - \exp(-\theta e^y) + G_\xi(x,y).$$

We are left with calculating the right-hand side of (B.20). By Lemma B.16, we have

$$
\begin{aligned}
V &\equiv \int_0^1 \mathrm{Cov}(X^{(\xi)}, Y^{(\xi)})\,\mathrm{d}\xi \\
&= \int_0^1 \int_0^\infty \int_0^\infty G_\xi(x,y) - e^{-\theta e^x} e^{-\theta e^y}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\xi \\
&\qquad + \int_0^1 \int_{-\infty}^0 \int_{-\infty}^0 F_\xi(x,y) - (1-e^{-\theta e^x})(1-e^{-\theta e^y})\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\xi \\
&\qquad - 2\int_0^1 \int_{-\infty}^0 \int_0^\infty \mathbb{P}(X^{(\xi)} > x, Y^{(\xi)} \le y) - e^{-\theta e^x}(1-e^{-\theta e^y})\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\xi, \\
&\equiv A + B - 2\cdot C. \qquad \text{(B.21)}
\end{aligned}
$$

We start with the first summand A. Recall the exponential integral $\mathrm{Ei}(x) = -\int_{-x}^\infty e^{-t}/t\,\mathrm{d}t$ for $x > 0$, and note that $\int_y^\infty e^{-\theta e^x}\,\mathrm{d}x = -\mathrm{Ei}(-\theta e^y)$ for $y \in \mathbb{R}$ and $\int_0^1 e^{-a\xi}\,\mathrm{d}\xi = (1-e^{-a})/a$ for $a > 0$. Fubini's theorem allows to rewrite A as

$$
\begin{aligned}
&\int_0^\infty \int_0^y e^{-\theta e^y}\Big\{\int_0^1 e^{-\theta\xi e^x}\,\mathrm{d}\xi - e^{-\theta e^x}\Big\}\,\mathrm{d}x \\
&\qquad\qquad + \int_y^\infty e^{-\theta e^x}\Big\{\int_0^1 e^{-\theta\xi e^y}\,\mathrm{d}\xi - e^{-\theta e^y}\Big\}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty e^{-\theta e^y}\int_0^y \frac{1-e^{-\theta e^x}}{\theta e^x} - e^{-\theta e^x}\,\mathrm{d}x \\
&\qquad\qquad + \int_y^\infty e^{-\theta e^x}\,\mathrm{d}x\Big\{\frac{1-e^{-\theta e^y}}{\theta e^y} - e^{-\theta e^y}\Big\}\,\mathrm{d}y \\
&= \int_0^\infty e^{-\theta e^y}\Big\{\frac{e^{-\theta e^y}-1}{\theta e^y} - \frac{e^{-\theta}-1}{\theta}\Big\} + \{-\mathrm{Ei}(-\theta e^y)\}\Big\{\frac{1-e^{-\theta e^y}}{\theta e^y} - e^{-\theta e^y}\Big\}\,\mathrm{d}y.
\end{aligned}
$$

Next, invoke the substitution $z = \theta e^y$ to obtain that

$$
A = \int_\theta^\infty \Big\{\frac{e^{-z}}{z} - \frac{1}{z} + \frac{1-e^{-\theta}}{\theta}\Big\}\frac{e^{-z}}{z} - \mathrm{Ei}(-z)\Big\{\frac{1}{z} - \frac{e^{-z}}{z} - e^{-z}\Big\}\frac{1}{z}\,\mathrm{d}z. \qquad \text{(B.22)}
$$

A similar calculation allows to rewrite

$$
\begin{aligned}
B &= \int_0^1 \int_{-\infty}^0 \int_{-\infty}^0 G_\xi(x,y) - e^{-\theta e^x}e^{-\theta e^y}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\xi \\
&= \int_{-\infty}^0 \int_{-\infty}^y e^{-\theta e^y}\Big\{\int_0^1 e^{-\theta\xi e^x}\,\mathrm{d}\xi - e^{-\theta e^x}\Big\}\,\mathrm{d}x \\
&\qquad\qquad + \int_y^0 e^{-\theta e^x}\Big\{\int_0^1 e^{-\theta\xi e^y}\,\mathrm{d}\xi - e^{-\theta e^y}\Big\}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^0 e^{-\theta e^y}\int_{-\infty}^y \frac{1-e^{-\theta e^x}}{\theta e^x} - e^{-\theta e^x}\,\mathrm{d}x \\
&\qquad\qquad + \int_y^0 e^{-\theta e^x}\,\mathrm{d}x\Big\{\frac{1-e^{-\theta e^y}}{\theta e^y} - e^{-\theta e^y}\Big\}\,\mathrm{d}y
\end{aligned}
$$

$$= \int_{-\infty}^{0} e^{-\theta e^y} \Big\{ \frac{e^{-\theta e^y} - 1}{\theta e^y} + 1 \Big\} + \Big\{ \mathrm{Ei}(-\theta) - \mathrm{Ei}(-\theta e^y) \Big\} \Big\{ \frac{1 - e^{-\theta e^y}}{\theta e^y} - e^{-\theta e^y} \Big\} \,\mathrm{d}y,$$

and the substitution $z = \theta e^y$ yields

$$B = \int_0^\theta \Big\{ \frac{e^{-z}}{z} - \frac{1}{z} + 1 \Big\} \frac{e^{-z}}{z} + \Big\{ \mathrm{Ei}(-\theta) - \mathrm{Ei}(-z) \Big\} \Big\{ \frac{1}{z} - \frac{e^{-z}}{z} - e^{-z} \Big\} \frac{1}{z} \,\mathrm{d}z. \tag{B.23}$$

Finally, regarding the term C, we have

$$\begin{aligned} C &= \int_0^1 \int_{-\infty}^0 \int_0^\infty e^{-\theta e^x} e^{-\theta e^y} - G_\xi(x,y) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}\xi \\ &= \int_{-\infty}^0 \int_0^\infty e^{-\theta e^x} \Big\{ e^{-\theta e^y} - \int_0^1 e^{-\theta \xi e^y} \,\mathrm{d}\xi \Big\} \,\mathrm{d}x \,\mathrm{d}y \\ &= \{-\mathrm{Ei}(-\theta)\} \int_{-\infty}^0 e^{-\theta e^y} - \frac{1 - e^{-\theta e^y}}{\theta e^y} \,\mathrm{d}y \\ &= \mathrm{Ei}(-\theta) \int_0^\theta \Big\{ \frac{1}{z} - \frac{e^{-z}}{z} - e^{-z} \Big\} \frac{1}{z} \,\mathrm{d}z. \end{aligned} \tag{B.24}$$

Next, the expressions in (B.22), (B.23) and (B.24) may be plugged-into (B.21). Using the notations

$$g(z) = \Big\{ \frac{1}{z} - \frac{e^{-z}}{z} - e^{-z} \Big\} \frac{1}{z}, \qquad h(z) = \Big\{ \frac{e^{-z}}{z} - \frac{1}{z} + 1 \Big\} \frac{e^{-z}}{z},$$

we obtain that

$$\begin{aligned} V &= \int_\theta^\infty \Big\{ \frac{1 - e^{-\theta}}{\theta} - 1 \Big\} \frac{e^{-z}}{z} + h(z) + \{-\mathrm{Ei}(-z)\} g(z) \,\mathrm{d}z \\ &\qquad + \int_0^\theta \Big\{ \mathrm{Ei}(-\theta) - \mathrm{Ei}(-z) \Big\} g(z) + h(z) - 2\,\mathrm{Ei}(-\theta) g(z) \,\mathrm{d}z \\ &= \int_0^\infty h(z) + \{-\mathrm{Ei}(-z)\} g(z) \,\mathrm{d}z + \frac{1 - e^{-\theta} - \theta}{\theta} \{-\mathrm{Ei}(-\theta)\} \\ &\qquad - \mathrm{Ei}(-\theta) \int_0^\theta g(z) \,\mathrm{d}z \end{aligned}$$

The first integral is independent of θ, and can be seen to be equal to $4 \log 2 - 2$. Further, $\int_0^\theta g(z) \,\mathrm{d}z = (e^{-\theta} - 1 + \theta)/\theta$, whence the last two summands cancel out. This proves the lemma. □

B.3. Further auxiliary lemmas

Lemma B.15. *Let A be a continuous function on $[0,1]$ with $\lim_{x \to 0} A(x)/x^\eta = 0$ for some $\eta \in (0, 1/2)$. Further, let H_n and H be monotone and non-negative* *functions on* $[0,1]$ *with*

$$\limsup_{n\to\infty} \int_{[0,1]} \frac{1}{x^{1-\eta}} \, \mathrm{d}H_n(x) < \infty \quad \textit{and} \quad \int_{[0,1]} \frac{1}{x^{1-\eta}} \, \mathrm{d}H(x) < \infty.$$

If $\lim_{n\to\infty} \sup_{x\in[0,1]} |B_n(x)| = 0$*, where* $B_n := H_n - H$*, and if there is a sequence of measurable functions* A_n *such that*

$$\lim_{n\to\infty} \sup_{x\in[0,1]} \left| \frac{A_n(x) - A(x)}{x^\eta} \right| = 0,$$

then we have

$$\lim_{n\to\infty} \int_{[0,1]} \frac{A_n(x)}{x} \, \mathrm{d}B_n(x) = 0.$$

Proof. For $r \in \mathbb{N}$ define the piecewise constant function

$$\tilde{A}_r(x) := \sum_{k=1}^{r} \mathbb{1}_{\left(\frac{k-1}{r}, \frac{k}{r}\right]}(x) \, \frac{A\big(k/r\big)}{k/r}$$

as an approximation of $A(x)/x$. We write $\int_{[0,1]} A_n(x)/x \, \mathrm{d}B_n(x) = I_{n1} + I_{n2} + I_{n3}$, where

$$I_{n1} = \int_{[0,1]} \frac{A_n(x) - A(x)}{x} \, \mathrm{d}B_n(x), \qquad I_{n2} = \int_{[0,1]} \frac{A(x)}{x} - \tilde{A}_r(x) \, \mathrm{d}B_n(x),$$
$$I_{n3} = \int_{[0,1]} \tilde{A}_r(x) \, \mathrm{d}B_n(x).$$

The first integral is bounded by

$$\int_{[0,1]} \left| \frac{A_n(x) - A(x)}{x} \right| \mathrm{d}(H_n + H)(x)$$
$$\leq \sup_{x\in[0,1]} \left| \frac{A_n(x) - A(x)}{x^\eta} \right| \int_{[0,1]} \frac{1}{x^{1-\eta}} \, \mathrm{d}(H_n + H)(x),$$

which converges to zero by assumption. Regarding I_{n2}, we obtain

$$|I_{n2}| = \Big| \int_{[0,1]} \frac{A(x) - \tilde{A}_r(x)x}{x^\eta} \, \frac{1}{x^{1-\eta}} \, \mathrm{d}B_n(x) \Big|$$
$$\leq \sup_{x\in[0,1]} \left| \frac{A(x) - \tilde{A}_r(x)x}{x^\eta} \right| \int_{[0,1]} \frac{1}{x^{1-\eta}} \, \mathrm{d}(H_n + H)(x). \qquad \text{(B.25)}$$

By uniform continuity of $x \mapsto A(x)/x^\eta$ on $[0,1]$, we have

$$\sup_{x\in[0,1]} \left| \frac{A(x) - \tilde{A}_r(x)x}{x^\eta} \right| \to 0 \quad \text{for } r \to \infty.$$

Thus, the limes superior (for $n \to \infty$) of the expression on the right-hand side of (B.25) can be made arbitrarily small by increasing r. Finally, we can bound $|I_{n3}|$ as follows

$$\begin{aligned} |I_{n3}| &\le \sum_{k=1}^{r} \frac{|A(k/r)|}{k/r} \Big| \int_{[0,1]} \mathbb{1}_{\left(\frac{k-1}{r}, \frac{k}{r}\right]}(x) \, \mathrm{d}B_n(x) \Big| \\ &= \sum_{k=1}^{r} \frac{|A(k/r)|}{k/r} \Big| B_n\Big(\frac{k}{r}\Big) - B_n\Big(\frac{k-1}{r}\Big) \Big| \\ &\le 2r^2 \sup_{x\in[0,1]} |A(x)| \sup_{x\in[0,1]} |B_n(x)|, \end{aligned}$$

which converges to zero by assumption. □

Lemma B.16. *Let X and Y be real-valued random variables such that XY is integrable. Then,*

$$\begin{aligned} \mathrm{E}[XY] = &\int_0^\infty \int_0^\infty \mathbb{P}(X > x, Y > y) \, \mathrm{d}x \, \mathrm{d}y + \int_{-\infty}^0 \int_{-\infty}^0 \mathbb{P}(X \le x, Y \le y) \, \mathrm{d}x \, \mathrm{d}y \\ &- \int_{-\infty}^0 \int_0^\infty \mathbb{P}(X > x, Y \le y) \, \mathrm{d}x \, \mathrm{d}y - \int_0^\infty \int_{-\infty}^0 \mathbb{P}(X \le x, Y > y) \, \mathrm{d}x \, \mathrm{d}y. \end{aligned}$$

Proof. This is a standard calculation based on Fubini's theorem. □

Acknowledgements

This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation, which is gratefully acknowledged. The authors are grateful to two referees and an associate editor for their constructive comments on an earlier version of this article which lead to a substantial improvements. The authors are also grateful to Shi-Yuan Zhou for her support with the simulation study.

References

[1] Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004). *Statistics of extremes: Theory and Applications*. Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd. MR2108013

[2] Berghaus, B. and A. Bücher (2017, 004). Goodness-of-fit tests for multivariate copula-based time series models. *Econometric Theory 33*(2), 292–330. MR3600045

[3] Berghaus, B. and A. Bücher (2018, 10). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Ann. Statist. 46*(5), 2307–2335. MR3845019

[4] Billingsley, P. (1979). *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. New York-Chichester-Brisbane: John Wiley & Sons. MR0534323
[5] Cai, J. J. (2019). A nonparametric estimator of the extremal index. *ArXiv e-prints*.
[6] Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika 84*(3), 567–577. MR1603985
[7] Darsow, W. F., B. Nguyen, and E. T. Olsen (1992). Copulas and Markov processes. *Illinois J. Math. 36*(4), 600–642. MR1215798
[8] de Haan, L., S. I. Resnick, H. Rootzén, and C. G. de Vries (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl. 32*(2), 213–224. MR1014450
[9] Dehling, H. and W. Philipp (2002). Empirical process techniques for dependent data. In *Empirical Process Techniques for Dependent Data*, pp. 3–113. Boston, MA: Birkhäuser Boston. MR1958777
[10] Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl. Probab. 10*(4), 1274–1301. MR1810875
[11] Ferreira, H. and M. Ferreira (2018). Estimating the extremal index through local dependence. *Annales de l'Institut Henri Poincaré, Probabilités et Statistique 54*(2), 587–605. MR3795060
[12] Ferreira, M. (2018). Analysis of estimation methods for the extremal index. *Electronic Journal of Applied Statistical Analysis 11*(1), 296–306. MR3807534
[13] Ferro, C. A. T. and J. Segers (2003). Inference for clusters of extreme values. *J. R. Stat. Soc. Ser. B Stat. Methodol. 65*(2), 545–556. MR1983763
[14] Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist. 37*(5B), 2990–3022. MR2541453
[15] Hsing, T. (1984). *Point Processes Associated with Extreme Value Theory*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–The University of North Carolina at Chapel Hill. MR2634031
[16] Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist. 21*(4), 2043–2071. MR1245780
[17] Hsing, T., J. Hüsler, and M. R. Leadbetter (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields 78*(1), 97–112. MR0940870
[18] Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete 65*(2), 291–306. MR0722133
[19] Leadbetter, M. R., G. Lindgren, and H. Rootzén (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. New York-Berlin: Springer-Verlag. MR0691492
[20] Leadbetter, M. R. and H. Rootzén (1988). Extremal theory for stochastic processes. *Ann. Probab. 16*(2), 431–478. MR0929071
[21] Naveau, P., A. Guillou, D. Cooley, and J. Diebolt (2009). Modelling pairwise dependence of maxima in space. *Biometrika 96*(1), 1–17. MR2482131
[22] Northrop, P. J. (2015). An efficient semiparametric maxima estimator of

the extremal index. *Extremes 18*(4), 585–603. MR3418769

[23] O'Brien, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab. 15*(1), 281–291. MR0877604

[24] Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications. *Ann. Appl. Probab. 4*(2), 529–548. MR1272738

[25] Pickands, III, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, Volume 49, pp. 859–878, 894–902. With a discussion. MR0820979

[26] Rémillard, B., N. Papageorgiou, and F. Soustra (2012). Copula-based semiparametric models for multivariate time series. *J. Multivariate Anal. 110*, 30–42. MR2927508

[27] Robert, C. Y. (2009). Inference for the limiting cluster size distribution of extreme values. *Ann. Statist. 37*(1), 271–310. MR2488352

[28] Robert, C. Y., J. Segers, and C. A. T. Ferro (2009). A sliding blocks estimator for the extremal index. *Electron. J. Stat. 3*, 993–1020. MR2540849

[29] Rootzén, H. (2009). Weak convergence of the tail empirical process for dependent sequences. *Stochastic Process. Appl. 119*(2), 468–490. MR2494000

[30] Shao, Q.-M. and H. Yu (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab. 24*(4), 2098–2127. MR1415243

[31] Smith, R. L. and I. Weissman (1994). Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B 56*(3), 515–528. MR1278224

[32] Süveges, M. (2007). Likelihood estimation of the extremal index. *Extremes 10*(1-2), 41–55. MR2407640

[33] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics, Vol. 3. Cambridge: Cambridge University Press. MR1652247

[34] van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer Series in Statistics. New York: Springer. MR1385671

Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 149 (2022) 75–106
stochastic processes and their applications
www.elsevier.com/locate/spa

Statistical analysis for stationary time series at extreme levels: New estimators for the limiting cluster size distribution

Axel Bücher*, Tobias Jennessen

Heinrich-Heine-Universität Düsseldorf, Mathematisches Institut, Universitätsstr. 1, 40225 Düsseldorf, Germany

Received 9 November 2020; received in revised form 18 February 2022; accepted 7 March 2022
Available online 18 March 2022

Abstract

The serial dependence of a stationary time series at extreme levels may be captured by the limiting cluster size distribution. New estimators based on a blocks declustering scheme are proposed and analyzed both theoretically and by means of a large-scale simulation study. A sliding blocks version of the estimators is shown to outperform a disjoint blocks version. In contrast to some competitors from the literature, the estimators only depend on one tuning parameter to be chosen by the statistician.

MSC: 62G32; 60G70

Keywords: Asymptotic theory; Block maxima; Clusters of extremes; Mixing coefficients

1. Introduction

The serial dependence of a stationary time series $(X_t)_{t\in\mathbb{Z}}$ at extreme levels may be described by various, partially interrelated limiting objects. The most traditional approach consists of studying the point process of exceedances and its weak convergence (see [18], or Section 10.3 in [2]). Two characterizing objects show up in the limit: the extremal index $\theta \in [0, 1]$ and the limiting cluster size distribution π, a probability distribution on the positive integers with $\pi(m)$ approximately representing the probability that extreme observations occur in a temporal cluster of size m. Under mild additional assumptions, the extremal index is in fact the reciprocal of the expectation of the limiting cluster size distribution [21].

* Corresponding author.
E-mail addresses: axel.buecher@hhu.de (A. Bücher), tobias.jennessen@hhu.de (T. Jennessen).

https://doi.org/10.1016/j.spa.2022.03.004

A recently introduced alternative object for assessing the serial dependence is given by the tail process $(Y_t)_{t\in\mathbb{Z}}$ (or the spectral process $(\Theta_t)_{t\in\mathbb{Z}}$) that may be associated with a suitably standardized version of $(X_t)_{t\in\mathbb{Z}}$ [1]. Heuristically, the law of those processes on $\mathbb{R}^{\mathbb{Z}}$ provides a more detailed description of the serial dependence. In fact, relying on results from [19], it can be shown that the limiting cluster size distribution π may be expressed as a functional of the tail process under mild additional conditions, see Remark 2.2.

Estimating the above mentioned objects based on a finite stretch of observations has received a lot of attention in recent years. For instance, estimators for the extremal index have been studied in [5,13,17,23,28,29], among many others. Estimators for π have been studied in [12,16,25,26]. To the best of our knowledge, inference on the law of the tail process has only been studied for selected functionals (note that the above mentioned contributions fall into this category as well). For instance, Drees et al. [11], Davis et al. [7] and Drees and Knežević [9] investigate estimators for the c.d.f. of Y_t, at a fixed lag t, which are based on making sophisticated use of the time change formula. Cissokho and Kulik [6] consider sliding blocks versions of peak-over-threshold estimators for a general class of functionals, including the extremal index and the limiting cluster size distribution. It is worthwhile to mention that asymptotic theory for many of the afore-mentioned estimators may be (non-trivially) derived from high level results in [10] on empirical processes for cluster functionals, see also [20].

The present paper is motivated by the apparently little amount of well-studied estimators for the limiting cluster size distribution π. Inspired by recent contributions on the estimation of the extremal index, we study an estimator that is based on a (disjoint or sliding) blocks declustering method. The sliding blocks estimator is shown to be more efficient than the disjoint blocks version. Moreover, by extensive Monte Carlo simulations, they are shown to exhibit very good finite-sample behavior in comparison to the competitors from [12,16,26].

The remaining parts of this paper are organized as follows: mathematical preliminaries, including precise definitions of the limiting objects described above, are provided in Section 2. In that section, we also define the new estimators. Regularity conditions needed to derive asymptotic normality are collected in Section 3, with the respective theoretical results given in Section 4. Section 5 contains results from a large scale Monte Carlo simulation study. The main arguments for the proofs are collected in Section 6, with an interesting side result on weak convergence of an empirical process associated with compound probabilities presented in Section 7 and proven in Section 8. Finally, all remaining proofs as well as additional simulation results are collected in a supplementary material.

2. Mathematical preliminaries and definition of estimators

Throughout the paper, $(X_t)_{t\in\mathbb{Z}}$ denotes a stationary time series with marginal cumulative distribution function (c.d.f.) F. The sequence is assumed to have an *extremal index* $\theta \in (0, 1]$, i.e., we assume that, for any $\tau > 0$, there exists a sequence $(u_n(\tau))_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} n\bar{F}(u_n(\tau)) = \tau$ and

$$\lim_{n\to\infty} \mathbb{P}(M_{1:n} \le u_n(\tau)) = e^{-\theta\tau}, \tag{2.1}$$

where $\bar{F} = 1 - F$ and $M_{1:n} = \max\{X_1, \dots, X_n\}$. Some thoughts reveal that, if the extremal index exists, then the convergence in (2.1) holds for any sequence $u_n(\tau)$ such that $\lim_{n\to\infty} n\bar{F}(u_n(\tau)) = \tau$ (see, e.g., the beginning of Section 5 in [18]) and that we may always choose $u_n(\tau) = F^{\leftarrow}(1 - \tau/n)$ (see the proof of Theorem 1.7.13 in [22]). Subsequently, the latter definition is tacitly employed, where $F^{\leftarrow}(p) = \inf\{x \in \mathbb{R} : F(x) \ge p\}$ denotes the (left-continuous) generalized inverse of F.

The point process of exceedances is defined as

$$N_n^{(\tau)}(B) = \sum_{t=1}^{n} \mathbb{1}(t/n \in B, X_t > u_n(\tau)),$$

for any Borel set $B \subset E := (0, 1]$ and $\tau \geq 0$. If the time series is serially independent, then it is well-known that $N_n^{(\tau)}$ converges in distribution to a homogeneous Poisson process on E with intensity τ. In the serial dependent case, if the extremal index exists and a certain mixing condition is met, then a necessary and sufficient condition for weak convergence of $N_n^{(\tau)}$ is as follows, see Theorems 4.1 and 4.2 in [18]: there exists a $\Delta(u_n(\tau))$-separating sequence $(q_n)_n$ (see Section 3 for a definition) such that the following limit exists for all $m \in \mathbb{N}_{\geq 1}$:

$$\pi(m) = \lim_{n\to\infty} \pi_n(m), \quad \pi_n(m) = \mathbb{P}(N_n^{(\tau)}(B_n) = m \mid N_n^{(\tau)}(B_n) > 0), \tag{2.2}$$

where $B_n = (0, q_n/n]$. In that case, the convergence in the last display holds for any $\Delta(u_n(\tau))$-separating sequence $(q_n)_n$ and the weak limit of $N_n^{(\tau)}$, say $N^{(\tau)}$, is a compound Poisson process with intensity $\theta\tau$ and compounding distribution π, notionally $N^{(\tau)} \sim \mathrm{CPP}(\theta\tau, \pi)$. If the $\Delta(u_n(\tau))$-condition holds for all $\tau > 0$, then π does not depend on τ ([18], Theorem 5.1), which will be tacitly assumed throughout. Motivated by (2.2), the distribution π is commonly referred to as the *(limiting) cluster size distribution*. Remark 2.2 provides a theoretical connection to the *tail process* introduced in [1].

Let $N_E^{(\tau)}$ denote the distributional limit of $N_n^{(\tau)}(E)$. Since the distribution of $N^{(\tau)}$ is $\mathrm{CPP}(\theta\tau, \pi)$, we have the stochastic representation

$$N_E^{(\tau)} \stackrel{d}{=} \sum_{i=1}^{\eta(\theta\tau)} \xi_i$$

for independent random variables $\eta(\theta\tau) \sim \mathrm{Poisson}(\theta\tau)$ and $\xi_i \sim \pi$. As a consequence, we have

$$p^{(\tau)}(0) = \mathbb{P}(N_E^{(\tau)} = 0) = e^{-\theta\tau},$$
$$p^{(\tau)}(m) = \mathbb{P}(N_E^{(\tau)} = m) = \sum_{j=1}^{m} \frac{e^{-\theta\tau}(\theta\tau)^j}{j!} \pi^{*j}(m), \quad m \in \mathbb{N}_{\geq 1},$$

where π^{*j} is the jth convolution of π. As explicitly written down in Equation (1.5) in [26], the previous equations allow to obtain, for any $\tau > 0$, a recursion expressing $\pi(m)$ as a function of θ, $p^{(\tau)}(1), \ldots, p^{(\tau)}(m)$ and $\pi(1), \ldots, \pi(m-1)$. This recursion then allows for estimation of $\pi(m)$ based on estimation of θ, $p^{(\tau)}(1), \ldots, p^{(\tau)}(m)$, which is precisely the approach followed in [26].

It may be argued that this approach suffers from the fact that the obtained recursion is depending on τ, which ultimately implies that the final estimator depends on τ as well. Hence, the statistician has either to make a choice, or to apply a suitable aggregation scheme. Within the present paper, we propose to instead consider a different recursion based on

$$\bar{p}(m) = \int_0^\infty p^{(\tau)}(m)\theta e^{-\theta\tau} \, \mathrm{d}\tau = \mathrm{E}[p^{(Z)}(m)],$$

where $Z \sim \mathrm{Exponential}(\theta)$. Perhaps surprisingly, and unlike for $p^{(\tau)}(m)$ above, the respective recursion does not even depend on θ, which allows for even simpler estimation. More precisely, a simple calculation shows that $\bar{p}(0) = \int_0^\infty \theta e^{-2\theta\tau}\, \mathrm{d}\tau = 1/2$ and

$$\bar{p}(m) = \sum_{j=1}^{m} \pi^{*j}(m) \frac{\theta}{j!} \int_0^\infty (\theta\tau)^j e^{-2\theta\tau}\, \mathrm{d}\tau = \sum_{j=1}^{m} \frac{1}{2^{j+1}} \pi^{*j}(m)$$

for $m \in \mathbb{N}_{\geq 1}$. As a consequence,

$$\begin{aligned}\bar{p}(m) &= \frac{1}{4}\pi(m) + \sum_{j=2}^{m} \frac{1}{2^{j+1}} \sum_{k=j-1}^{m-1} \pi^{*(j-1)}(k)\pi(m-k) \\ &= \frac{1}{4}\pi(m) + \sum_{k=1}^{m-1} \pi(m-k) \sum_{j=2}^{k+1} \pi^{*(j-1)}(k) \frac{1}{2^{j+1}} \\ &= \frac{1}{4}\pi(m) + \frac{1}{2}\sum_{k=1}^{m-1} \pi(m-k)\bar{p}(k),\end{aligned}$$

which in turn implies

$$\pi(m) = 4\bar{p}(m) - 2\sum_{k=1}^{m-1} \pi(m-k)\bar{p}(k), \quad m \in \mathbb{N}_{\geq 1}. \tag{2.3}$$

Obviously, Eq. (2.3) allows to recursively derive $(\pi(1), \ldots, \pi(m))$ from $(\bar{p}(1), \ldots, \bar{p}(m))$. The plug-in principle hence allows to estimate the former vector based on suitable estimators for the latter vector.

For the estimation of $(\bar{p}(1), \ldots, \bar{p}(m))$, a transformation extensively used in [4,5] comes in handy: the random variable

$$Z_{1:n} = n\{1 - F(M_{1:n})\}$$

is asymptotically exponentially distributed with parameter θ, for $n \to \infty$. Indeed, since $v_n(\tau) = F^{\to}(1-\tau/n)$ with the right-continuous generalized inverse $F^{\to}$ satisfies $\lim_{n\to\infty} \bar{F}(v_n(\tau)) = \tau$, whence

$$\mathbb{P}(Z_{1:n} \geq \tau) = \mathbb{P}(M_{1:n} \leq v_n(\tau)) \to e^{-\theta\tau} \tag{2.4}$$

for $n \to \infty$ by (2.1). Next, for motivating our estimator it is instructive to consider, for two independent copies $(X_t)_{t\in\mathbb{Z}}, (\tilde{X}_t)_{t\in\mathbb{Z}}$, the random variable

$$N_n^{(\tilde{Z}_{1:n})}(E) = \sum_{t=1}^{n} \mathbb{1}(X_t > u_n(\tilde{Z}_{1:n})) \overset{a.s.}{=} \sum_{t=1}^{n} \mathbb{1}(X_t > \tilde{M}_{1:n}),$$

where $\tilde{Z}_{1:n} = n\{1 - F(\tilde{M}_{1:n})\}$. Then, conditional on $\tilde{Z}_{1:n}$, the random variable $N_n^{(\tilde{Z}_{1:n})}(E)$ approximately follows a compound Poisson distribution with intensity $\theta\tilde{Z}_{1:n}$ and compounding distribution π, for sufficiently large n. As a consequence,

$$\mathbb{P}\Big(N_n^{(\tilde{Z}_{1:n})}(E) = m \mid \tilde{Z}_{1:n}\Big) \approx p^{(\tilde{Z}_{1:n})}(m),$$

which readily implies

$$\mathbb{P}\Big(N_n^{(\tilde{Z}_{1:n})}(E) = m\Big) \approx \mathrm{E}\Big[p^{(\tilde{Z}_{1:n})}(m)\Big] \approx \bar{p}(m), \tag{2.5}$$

where the second approximation is due to (2.4). The latter display allows for estimation of $\bar{p}(m)$ based on the method of moments.

More precisely, suppose we observe a finite-stretch of observation from the time series, say $X_1, \ldots, X_n$. Divide the observation period into non-overlapping successive blocks of size $b = b_n$, that is, into blocks $I_i = I_i^{\mathrm{db}}$ (db for 'disjoint blocks'),

$$I_1 = \{1, \ldots, b\}, \quad I_2 = \{b+1, \ldots, 2b\}, \quad \ldots \quad, I_k = \{(k-1)b+1, \ldots, kb\},$$

where $k = k_n = \lfloor n/b_n \rfloor$. A possible remainder block $I_{k+1}^\circ = \{kb+1, \ldots, n\}$ of cardinality $|I_{k+1}| < b_n$ will have a negligible influence on the subsequent estimators and will hence be discarded. Asymptotically, $(b_n)_n$ needs to be an intermediate sequence satisfying $b = b_n \to \infty$ and $b_n = o(n)$. Now, by well-known heuristics, cluster functionals (i.e., statistics that depend only the 'large observations' within a specific block I_j) calculated based on disjoint blocks of observations may be considered asymptotically independent, whence (2.5) suggests to estimate $\bar{p}(m)$ by

$$\hat{\bar{p}}_n(m) = \hat{\bar{p}}_n^{\mathrm{db}}(m) = \frac{1}{k_n(k_n-1)} \sum_{\substack{i,i'=1 \\ i \neq i'}}^{k_n} \mathbb{1}\Big\{ \sum_{s \in I_{i'}} \mathbb{1}\big(X_s > M_{ni}^{\mathrm{db}}\big) = m \Big\},$$

where the upper index 'db' refers to the fact that the underlying blocks are disjoint and where $M_{ni}^{\mathrm{db}} = \max\{X_t : t \in I_i\}$. Following Berghaus and Bücher [4] and Bücher and Jennessen [5], a possibly more efficient version that is based on sliding/overlapping blocks instead of disjoint blocks is given by

$$\hat{\bar{p}}_n^{\mathrm{sb}}(m) = \frac{1}{|D_n|} \sum_{(i,i') \in D_n} \mathbb{1}\Big\{ \sum_{s \in I_{i'}^{\mathrm{sb}}} \mathbb{1}\Big(X_s > M_{ni}^{\mathrm{sb}}\Big) = m \Big\},$$

where $I_i^{\mathrm{sb}} = \{i, \ldots, i+b_n-1\}$, $M_{ni}^{\mathrm{sb}} = \max\{X_t : t \in I_i^{\mathrm{sb}}\}$ and where D_n is the set of all pairs $(i,i') \in \{1, \ldots, n-b_n+1\}^2$ such that $I_i^{\mathrm{sb}} \cap I_{i'}^{\mathrm{sb}} = \varnothing$. Obviously, since $I_i^{\mathrm{sb}} \cap I_{i'}^{\mathrm{sb}} = \varnothing$, the same heuristics as in the disjoint blocks case applies: the expectation of each summand is approximately equal to $\bar{p}(m)$.

Based on the recursion (2.3), the final (disjoint and sliding blocks) estimators for $\pi(m)$, $m \in \mathbb{N}_{\geq 1}$, are defined, for mb $\in$ {db, sb}, by

$$\hat{\pi}_n^{\mathrm{mb}}(m) = 4\hat{\bar{p}}_n^{\mathrm{mb}}(m) - 2\sum_{k=1}^{m-1} \hat{\pi}_n^{\mathrm{mb}}(m-k)\hat{\bar{p}}_n^{\mathrm{mb}}(k). \tag{2.6}$$

Remark 2.1. The estimators $\hat{\bar{p}}_n^{\mathrm{db}}$ and $\hat{\bar{p}}_n^{\mathrm{sb}}(m)$ can be interpreted as U-statistics, which we exemplarily illustrate for $\hat{\bar{p}}_n^{\mathrm{db}}$. Indeed, we may write

$$\hat{\bar{p}}_n^{\mathrm{db}}(m) = \binom{k_n}{2}^{-1} \sum_{\substack{i,i'=1 \\ i < i'}}^{k_n} h_m\big((X_s)_{s \in I_{i'}}, (X_s)_{s \in I_i}\big)$$

with h_m defined by

$$h_m\big((X_s)_{s \in I_{i'}}, (X_s)_{s \in I_i}\big) = \frac{1}{2}\Big[\, \mathbb{1}\Big\{ \sum_{s \in I_{i'}} \mathbb{1}\Big(X_s > M_{ni}^{\mathrm{db}}\Big) = m \Big\} + \mathbb{1}\Big\{ \sum_{s \in I_i} \mathbb{1}\Big(X_s > M_{ni'}^{\mathrm{db}}\Big) = m \Big\} \,\Big].$$

Since h_m is a symmetric kernel function with

$$\mathrm{E}\Big[h_m\big((X_s)_{s\in I_1},(X_s)_{s\in I_2}\big)\Big] = \mathbb{P}\Big(\sum_{s\in I_1}\mathbb{1}(X_s > M_{n2}^{\mathrm{db}}) = m\Big) \approx \bar{p}(m),$$

the estimator $\hat{\bar{p}}_n(m)$ can be considered an (approximate) U-statistic for $\bar{p}(m)$. The same applies for $\hat{\bar{p}}_n^{\mathrm{sb}}(m)$. The U-statistics representation suggests an alternative approach to proving our central results based on, e.g., suitable adaptations of the Hoeffding decomposition. The approach will be investigated in a future research project from a higher level.

Remark 2.2. The limiting cluster size distribution is closely connected to the tail process introduced in [1], see also the monograph [19]. Since the tail process may only be defined for heavy tailed stationary time series, a standardization is necessary first. For simplicity, we assume that F is continuous. In that case, for any $t \in \mathbb{Z}$, $Z_t = 1/\{1 - F(X_t)\}$ is standard Pareto-distributed and the event $X_t > u_n(\tau)$ is (almost surely) equivalent to $Z_t > n/\tau$. Under the assumption that $(Z_t)_{t\in\mathbb{Z}}$ is regularly varying (i.e., all vectors of the form $(Z_k, \dots, Z_\ell)$ are multivariate regularly varying), there exists a process $(Y_t)_{t\in\mathbb{Z}}$, the *tail process of* $(Z_t)_{t\in\mathbb{Z}}$, such that, for every $s, t \in \mathbb{Z}$ with $s \le t$,

$$\mathbb{P}\big(x^{-1}(Z_s, \dots, Z_t) \in \cdot \mid Z_0 > x\big) \xrightarrow{w} \mathbb{P}\big((Y_s, \dots, Y_t) \in \cdot\big) \quad (x \to \infty),$$

see Theorem 2.1 in [1]. If we additionally assume that, for the sequence $(q_n)_n$ from (2.2) and for all $x, y > 0$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(\max_{m\le|t|\le q_n} Z_t > nx \mid Z_0 > ny\Big) = 0, \tag{2.7}$$

then π may be expressed through the tail process, see Example 6.2.9 in [19]:

$$\begin{aligned}\pi(m) &= \lim_{n\to\infty}\mathbb{P}\Big(\sum_{1\le t\le q_n}\mathbb{1}(X_t > u_n(\tau)) = m \,\Big|\, \max_{1\le t\le q_n} X_t > u_n(\tau)\Big)\\ &= \mathbb{P}\Big(\sum_{t\ge 0}\mathbb{1}(Y_t > 1) = m \,\Big|\, \max_{t\le -1} Y_t \le 1\Big), \quad m \in \mathbb{N}_{\ge 1}.\end{aligned}$$

In other words, $\pi(m)$ is the conditional probability that the 'number of time points where the tail process exceeds the value 1' equals m, conditional on the event that the tail process does not exceed 1 until $t = -1$. It is worthwhile to mention that (2.7) is for instance satisfied for geometrically ergodic Markov chains, short-memory linear or max-stable processes and m-dependent sequences; see [6], page 7, and [19], page 151.

3. Regularity conditions

This section summarizes technical regularity conditions which are imposed to derive asymptotic properties for the estimators from the previous section. First of all, the serial dependence will be controlled via alpha- and beta-mixing coefficients. For two sigma-fields $\mathcal{F}_1, \mathcal{F}_2$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let

$$\alpha(\mathcal{F}_1, \mathcal{F}_2) = \sup_{A\in\mathcal{F}_1, B\in\mathcal{F}_2}|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

$$\beta(\mathcal{F}_1, \mathcal{F}_2) = \frac{1}{2}\sum_{i\in I}\sum_{j\in J}\sup|\mathbb{P}(A_i\cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|,$$

where the last supremum is over all finite partitions $(A_i)_{i\in I} \subset \mathcal{F}_1$ and $(B_j)_{j\in J} \subset \mathcal{F}_2$ of Ω. For $-\infty \leq p < q \leq \infty$ and $\varepsilon \in (0,1]$, let $\mathcal{B}^\varepsilon_{p:q}$ denote the sigma algebra generated by $U_s^\varepsilon := U_s \mathbb{1}(U_s > 1-\varepsilon)$ with $s \in \{p,\dots,q\}$; here, $U_s = F(X_s)$. Finally, for $\ell \geq 1$, let

$$\alpha_\varepsilon(\ell) = \sup_{k\in\mathbb{N}} \alpha(\mathcal{B}^\varepsilon_{1:k}, \mathcal{B}^\varepsilon_{k+\ell:\infty}), \qquad \beta_\varepsilon(\ell) = \sup_{k\in\mathbb{N}} \beta(\mathcal{B}^\varepsilon_{1:k}, \mathcal{B}^\varepsilon_{k+\ell:\infty}).$$

Conditions on the decay of the mixing coefficients will be imposed below.

Formally introducing conditions that connect the existence of the cluster size distribution to weak convergence of the exceedance process requires yet another weaker version of the alpha mixing coefficient. Fix $m \geq 1$ and $\tau_1 > \cdots > \tau_m > 0$. For $1 \leq p < q \leq n$, let $\mathcal{F}^{(\tau_1,\dots,\tau_m)}_{p:q,n}$ denote the sigma-algebra generated by the events $\{X_s > u_n(\tau_j)\}$ for $s \in \{p,\dots,q\}$ and $j \in \{1,\dots m\}$. For $\ell \in \{1,\dots,n\}$, define

$$\alpha_{n,\ell}(\tau_1,\dots,\tau_m) = \sup\{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| : \\ A \in \mathcal{F}^{(\tau_1,\dots,\tau_m)}_{1:s,n}, B \in \mathcal{F}^{(\tau_1,\dots,\tau_m)}_{s+\ell:n,n}, 1 \leq s \leq n-\ell\}.$$

The condition $\Delta_n(\{u_n(\tau_j)\}_{1\leq j\leq m})$ is said to hold if there exists a sequence $(\ell_n)_n$ with $\ell_n = o(n)$ such that $\alpha_{n,\ell_n}(\tau_1,\dots,\tau_m) = o(1)$ as $n \to \infty$. A sequence $(q_n)_n$ with $q_n = o(n)$ is said to be $\Delta_n(\{u_n(\tau_j)\}_{1\leq j\leq m})$-separating if there exists a sequence $(\ell_n)_n$ with $\ell_n = o(q_n)$ such that $\alpha_{n,\ell_n}(\tau_1,\dots,\tau_m) = o(q_n/n)$ as $n\to\infty$. If $\Delta_n(\{u_n(\tau_j)\}_{1\leq j\leq m})$ is met, then such a sequence always exists, simply take $q_n = \lfloor \max\{n\alpha^{1/2}_{n,\ell_n}, (n\ell_n)^{1/2}\}\rfloor$.

As already stated in Section 2, by Theorems 4.1 and 4.2 in [18], if the extremal index exists and the $\Delta(u_n(\tau))$-condition is met ($m = 1$), then a necessary and sufficient condition for weak convergence of $N_n^{(\tau)}$ is the convergence in (2.2) for some $\Delta(u_n(\tau))$-separating sequence $(q_n)_n$. Moreover, in that case, the convergence in (2.2) holds for any $\Delta(u_n(\tau))$-separating sequence $(q_n)_n$, and the weak limit of $N_n^{(\tau)}$, say $N^{(\tau)}$, is a compound Poisson process CPP$(\theta\tau,\pi)$. If the $\Delta(u_n(\tau))$-condition holds for any $\tau > 0$, then π does not depend on τ ([18], Theorem 5.1).

A multivariate version of the latter results is stated in [24], see also the summary in [26], page 278, and the thesis [15]. Suppose that the extremal index exists and that the $\Delta(u_n(\tau_1), u_n(\tau_2))$-condition is met for any $\tau_1 \geq \tau_2 \geq 0, \tau_1 \neq 0$. Moreover, assume that there exists a family of probability measures $\{\pi_2^{(\sigma)} : \sigma \in [0,1]\}$ on $\mathcal{J} = \{(i,j) \in \mathbb{N}^2_{\geq 0} : i \geq j \geq 0, i \geq 1\}$, such that, for all $(i,j) \in \mathcal{J}$,

$$\lim_{n\to\infty} \mathbb{P}(N_n^{(\tau_1)}(B_n) = i, N_n^{(\tau_2)}(B_n) = j \mid N_n^{(\tau_1)}(B_n) > 0) = \pi_2^{(\tau_2/\tau_1)}(i,j),$$

where $B_n = (0, q_n/n]$ and q_n is some $\Delta(u_n(\tau_1), u_n(\tau_2))$-separating sequence. In that case, the two-level point process $\boldsymbol{N}_n^{(\tau_1,\tau_2)} = (N_n^{(\tau_1)}, N_n^{(\tau_2)})$ converges in distribution to a point process $\boldsymbol{N}^{(\tau_1,\tau_2)} = (N_1^{(\tau_1,\tau_2)}, N_2^{(\tau_1,\tau_2)})$ with characterizing Laplace transform explicitly stated in [26] on top of page 278. It can further be shown that, under the above mixing assumptions and if the extremal index exists, the existence of the limit in the latter display is even necessary for the distributional convergence of $\boldsymbol{N}_n^{(\tau_1,\tau_2)}$; the limit $\pi_2^{(\tau_2/\tau_1)}$ then necessarily depends on τ_1 and τ_2 only through τ_2/τ_1, see Theorem 2.5 and page 535 in [24]. Throughout, let

$$\boldsymbol{N}_E^{(\tau_1,\tau_2)} = (N_{E,1}^{(\tau_1,\tau_2)}, N_{E,2}^{(\tau_1,\tau_2)}) = \boldsymbol{N}^{(\tau_1,\tau_2)}(E),$$

whose marginal distributions are equal to $N_E^{(\tau_1)}$ and $N_E^{(\tau_2)}$ and which further allows for the stochastic representation

$$\boldsymbol{N}_E^{(\tau_1,\tau_2)} \stackrel{d}{=} \sum_{i=1}^{\eta(\theta\tau_1)} (\xi_{i,1}^{(\tau_2/\tau_1)}, \xi_{i,2}^{(\tau_2/\tau_1)}),$$

where $\eta(\theta\tau_1) \sim \text{Poisson}(\theta\tau_1)$ is independent of the bivariate i.i.d. sequence $(\xi_{i,1}^{(\tau_2/\tau_1)}, \xi_{i,2}^{(\tau_2/\tau_1)}) \sim \pi_2^{(\tau_2/\tau_1)}$. As a consequence, the distribution of $\boldsymbol{N}_E^{(\tau_1,\tau_2)}$ on $\mathbb{N}_{\geq 0}^2$, say

$$p_2^{(\tau_1,\tau_2)}(i,j) = \mathbb{P}(\boldsymbol{N}_E^{(\tau_1,\tau_2)} = (i,j)),$$

is given by $p_2^{(\tau_1,\tau_2)}(0,0) = e^{-\theta\tau_1}$, $p_2^{(\tau_1,\tau_2)}(i,j) = 0$ for $i < j$ and

$$p_2^{(\tau_1,\tau_2)}(i,j) = e^{-\theta\tau_1} \sum_{k=1}^{i} \frac{(\theta\tau_1)^k}{k!} \pi_2^{(\tau_2/\tau_1),*k}(i,j), \quad i \geq j \geq 0,\ i \geq 1,$$

where $\pi_2^{(\tau_2/\tau_1),*k}$ is the kth convolution of $\pi_2^{(\tau_2/\tau_1)}$.

The assumptions needed to derive asymptotic properties for $\hat{p}_n^{\mathrm{mb}}(m)$ and $\hat{\pi}_n^{\mathrm{mb}}(m)$ are collected in the following condition.

Condition 3.1.

(i) The stationary time series $(X_s)_{s\in\mathbb{N}}$ has an extremal index $\theta \in (0,1]$ and the two-level point process of exceedances $\boldsymbol{N}_n^{(\tau_1,\tau_2)}$ converges weakly to $\boldsymbol{N}^{(\tau_1,\tau_2)}$.

(ii) There exist constants $\varepsilon_1 \in (0,1)$, $\eta > 3$ and $C > 0$ such that

$$\alpha_{\varepsilon_1}(n) \leq Cn^{-\eta} \qquad \forall\ n \in \mathbb{N}.$$

The block size b_n converges to infinity and satisfies

$$k_n = o(b_n^\eta),\ n \to \infty,$$

(i.e., a slow decrease of the mixing coefficients requires large block sizes). Further, there exists a sequence $\ell_n \to \infty$ with $\ell_n = o(b_n)$ and $k_n\alpha_{\varepsilon_1}(\ell_n) = o(1)$ as $n \to \infty$.

(iii) For some $c > 1 - \varepsilon_1$ with ε_1 from (ii), one has

$$\mathbb{P}(N'_{n1} \leq c) = o(k_n^{-1}),$$

where $N'_{n1} = \max\{U_s : s \in \{1,\ldots,\lfloor b_n/2\rfloor\}\}$ and $U_s = F(X_s)$.

(iv) (Bias.) For any $j \in \mathbb{N}_{\geq 1}$, as $n \to \infty$,

$$\mathrm{E}\big[\varphi_{n,j}(Z_{1:b_n})\big] = \bar{p}(j) + o\big(k_n^{-1/2}\big),$$

where $\varphi_{n,j}(z) = \mathbb{P}(N_{b_n}^{(z)} = j)$ and $N_{b_n}^{(z)} = \sum_{s=1}^{b_n} \mathbb{1}(U_s > 1 - z/b_n)$.

Remark 3.2. Under Condition 3.1(i)–(ii), Condition 3.1(iii) is equivalent to the following condition: For some $c > 1 - \varepsilon_1$ with ε_1 from 3.1(ii), one has

$$\lim_{n\to\infty} \mathbb{P}\Big(\min_{i=1,\ldots,2k_n} N'_{ni} \leq c\Big) = 0, \tag{3.1}$$

where $N'_{ni} = \max\{U_s : s \in [(i-1)b_n/2 + 1, ib_n/2] \cap \mathbb{N}\}$ for $i \in \{1,\ldots,2k_n\}$ (note that (3.1) corresponds to Condition 2.1(v) in [4]). This can be seen as follows. First,

$$\Big|\mathbb{P}\Big(\min_{i=2,4,\ldots,2k_n} N'_{ni} > c\Big) - \prod_{i=2,4,\ldots,2k_n} \mathbb{P}\big(N'_{ni} > c\big)\Big| \leq k_n\alpha_{\varepsilon_1}(\lfloor b_n/2\rfloor + 1) \leq Ck_nb_n^{-\eta},$$

which converges to zero by Condition 3.1(ii). Next, by stationarity,

$$\prod_{i=2,4,\ldots,2k_n} \mathbb{P}\big(N'_{ni} > c\big) = \mathbb{P}\big(N'_{n1} > c\big)^{k_n} = \Big(1 - \frac{k_n\mathbb{P}\big(N'_{n1} \leq c\big)}{k_n}\Big)^{k_n},$$

which converges to 1 iff Condition 3.1(iii) holds. The previous two displays imply

$$\lim_{n\to\infty} \mathbb{P}\Big(\min_{i=2,4,\dots,2k_n} N'_{ni} > c\Big) = 1 \qquad \Longleftrightarrow \qquad \mathbb{P}(N'_{n1} \le c) = o(k_n^{-1}).$$

The same can be shown for the minimum over the odd indices, which shows that (3.1) follows from Condition 3.1(iii). Along with this equivalence the other implication is trivial since $\mathbb{P}(\min_{i=2,4,\dots,2k_n} N'_{ni} \le c) \le \mathbb{P}(\min_{i=1,\dots,2k_n} N'_{ni} \le c)$, which also holds for odd indices i.

The conditions are weaker versions of the conditions imposed in [4], which in turn are mostly based on [26]. In contrast to those papers, no moment condition on the increments of $\tau \mapsto N_n^{(\tau)}(E)$ is needed, which may be explained by the fact that the cluster functionals showing up in the definition of $\hat{\bar{p}}_n(m)$ are bounded by 1. This also allows for a great simplification of the α-mixing condition in comparison to the last-named references. For the treatment of the sliding blocks estimator, we will additionally impose a beta-mixing condition below, which is used for proving tightness of the scaled estimation error of empirical compound probabilities, see Section 7.

An exemplary time series model meeting the above conditions is given by the max-autoregressive process of order 1, ARMAX in short, defined by the recursion

$$X_s = \max\{\alpha X_{s-1}, (1-\alpha)Z_s\}, \qquad s \in \mathbb{Z},$$

where $\alpha \in [0, 1)$ and $(Z_s)_s$ is an i.i.d. sequence of standard Fréchet random variables. A stationary solution is given by $X_s = \max_{j\ge 0}(1-\alpha)\alpha^j Z_{s-j}$ such that the stationary distribution is standard Fréchet as well. The extremal index is $\theta = 1-\alpha$ and the cluster size distribution is geometric, i.e., $\pi(j) = \alpha^{j-1}(1-\alpha)$ for $j \ge 1$, see Chapter 10 in [2]. Conditions (i)–(iii) were shown to be satisfied in [4], page 2322. Regarding Condition (iv), extensive simulations have shown that the bias $\mathrm{E}[\varphi_{n,j}(Z_{1:b_n})] - \bar{p}(j)$ is of the order b_n^{-1}, such that Condition (iv) is met provided $k_n = o(b_n^2)$.

Further discussions of the conditions in general and details on models defined by stochastic difference equations fulfilling slight adaptations of conditions (i) and (ii) and (iv) are provided in [4,26].

4. Main results

In this section we derive asymptotic normality of both the disjoint and sliding blocks estimators from Section 2. A comparison of the asymptotic variances shows that the sliding blocks version exhibits a smaller asymptotic variance than the disjoint blocks version. Subsequently, for mb $\in$ {db, sb}, let

$$\begin{aligned} s_{n,j}^{\mathrm{mb}} &= \sqrt{k_n}\big\{\hat{\bar{p}}_n^{\mathrm{mb}}(j) - \bar{p}(j)\big\}, \quad j \in \mathbb{N}_{\ge 1}, \\ v_{n,j}^{\mathrm{mb}} &= \sqrt{k_n}\big\{\hat{\pi}_n^{\mathrm{mb}}(j) - \pi(j)\big\}, \quad j \in \mathbb{N}_{\ge 1}. \end{aligned} \tag{4.1}$$

For simplicity, we will further assume that F is continuous.

Theorem 4.1. *Assume that Condition 3.1 is met. Then, for any $m \in \mathbb{N}_{\ge 1}$,*

$$(s_{n,1}^{\mathrm{db}}, \dots, s_{n,m}^{\mathrm{db}}) \xrightarrow{d} (s_1^{\mathrm{db}}, \dots, s_m^{\mathrm{db}}) \sim \mathcal{N}_m(0, \Sigma_m^{\mathrm{db}})$$

as $n \to \infty$, where the covariance matrix $\Sigma_m^{\mathrm{db}} = (d_{j,j'}^{\mathrm{db}})_{1\le j,j'\le m}$ is given by

$$d_{j,j'}^{\mathrm{db}} = \int_0^\infty \int_0^\infty \mathrm{Cov}\Big(\mathbb{1}(N_E^{(\tau)} = j) + p^{(Z)}(j),$$

$$\mathbb{1}(N_E^{(\tau')} = j') + p^{(Z)}(j') \Big) \, \mathrm{d}H(\tau) \, \mathrm{d}H(\tau'). \tag{4.2}$$

Here, H denotes the c.d.f. of the $\mathrm{Exp}(\theta)$*-distribution, and $N_E^{(\tau)} \sim p^{(\tau)}$ and $Z \sim \mathrm{Exp}(\theta)$ are such that*

$$\mathbb{P}(N_E^{(\tau)} = j, Z > \mu) = \begin{cases} p_2^{(\tau,\mu)}(j,0) & ,\tau \geq \mu \\ e^{-\theta\mu}\mathbb{1}(j=0) & ,\tau < \mu \end{cases} \qquad (j \in \mathbb{N}_{\geq 0}, \mu > 0)$$

and

$$\mathbb{P}(N_E^{(\tau)} = j, N_E^{(\tau')} = j') = \begin{cases} p_2^{(\tau,\tau')}(j,j') & ,\tau \geq \tau' \\ p_2^{(\tau',\tau)}(j',j) & ,\tau < \tau' \end{cases} \qquad (j, j' \in \mathbb{N}_{\geq 0}).$$

Theorem 4.2. *In addition to Condition 3.1 assume that $\sqrt{k_n}\beta_{\varepsilon_2}(b_n) = o(1)$ for some $\varepsilon_2 > 0$. Then, for any $m \in \mathbb{N}_{\geq 1}$,*

$$(s_{n,1}^{\mathrm{sb}}, \ldots, s_{n,m}^{\mathrm{sb}}) \xrightarrow{d} (s_1^{\mathrm{sb}}, \ldots, s_m^{\mathrm{sb}}) \sim \mathcal{N}_m(0, \Sigma_m^{\mathrm{sb}})$$

as $n \to \infty$, where the covariance matrix $\Sigma_m^{\mathrm{sb}} = (d_{j,j'}^{\mathrm{sb}})_{1 \leq j,j' \leq m}$ is given by

$$\begin{aligned} d_{j,j'}^{\mathrm{sb}} = 2\int_0^1 \Big\{ \int_0^\infty \int_0^\infty & \mathrm{Cov}\big(\mathbb{1}(X_{1,\xi}^{(\tau)} = j), \mathbb{1}(Y_{1,\xi}^{(\tau')} = j')\big) \, \mathrm{d}H(\tau)\mathrm{d}H(\tau') \\ & + \int_0^\infty \mathrm{Cov}\big(\mathbb{1}(X_{3,\xi}^{(\tau)} = j), p^{(Y_{3,\xi})}(j')\big) \, \mathrm{d}H(\tau) \\ & + \int_0^\infty \mathrm{Cov}\big(\mathbb{1}(X_{3,\xi}^{(\tau)} = j'), p^{(Y_{3,\xi})}(j)\big) \, \mathrm{d}H(\tau) \\ & + \mathrm{Cov}\big(p^{(X_{2,\xi})}(j), p^{(Y_{2,\xi})}(j')\big) \Big\} \, \mathrm{d}\xi, \end{aligned} \tag{4.3}$$

where for $0 \leq \tau \leq \tau'$ and $x, y > 0$,

$$\begin{aligned} \mathbb{P}\big(X_{1,\xi}^{(\tau)} = j, Y_{1,\xi}^{(\tau')} = j'\big) &= \sum_{l=0}^{j} \sum_{r=j-l}^{j'} p^{(\xi\tau)}(l) p^{(\xi\tau')}(j'-r) \\ &\qquad \times p_2^{((1-\xi)\tau',(1-\xi)\tau)}(r, j-l), \\ \mathbb{P}\big(X_{2,\xi} > x, Y_{2,\xi} > y\big) &= \exp\big(-\theta\{(x \wedge y)\xi + (x \vee y)\}\big), \\ \mathbb{P}\big(X_{3,\xi}^{(\tau)} = j, Y_{3,\xi} > x\big) &= e^{-\theta\xi x} \sum_{l=0}^{j} p^{(\xi\tau)}(l) p_2^{((1-\xi)\tau,(1-\xi)x)}(j-l,0)\mathbb{1}(x \leq \tau) \\ &\qquad + e^{-\theta x} p^{(\tau\xi)}(j)\mathbb{1}(x > \tau). \end{aligned}$$

It is worthwhile to mention that $X_{1,\xi}^{(\tau)}, Y_{1,\xi}^{(\tau)}, X_{3,\xi}^{(\tau)}$ are equal in distribution to $N_E^{(\tau)}$ and that $X_{2,\xi}, Y_{2,\xi}, Y_{3,\xi}$ are exponentially distributed with parameter θ.

Regarding the estimator $\hat{\pi}_n^{\mathrm{mb}}(j)$ from (2.6), recall the definition of $\upsilon_{n,j}^{\mathrm{mb}}$ in (4.1) and of $(s_1^{\mathrm{mb}}, \ldots, s_m^{\mathrm{mb}})$ and Σ_m^{mb} in Theorem 4.1 (mb = db) or Theorem 4.2 (mb = sb).

Corollary 4.3. *Let* mb $\in$ {db, sb}. *Under the conditions of Theorem 4.1* (mb = db) *or Theorem 4.2* (mb = sb) *we have, for any $m \in \mathbb{N}_{\geq 1}$ and as $n \to \infty$,*

$$(\upsilon_{n,1}^{\mathrm{mb}}, \ldots, \upsilon_{n,m}^{\mathrm{mb}}) \xrightarrow{d} (\upsilon_1^{\mathrm{mb}}, \ldots, \upsilon_m^{\mathrm{mb}}) \sim \mathcal{N}_m(0, \Gamma_m^{\mathrm{mb}}),$$

where $v_1^{\mathrm{mb}} = 4s_1^{\mathrm{mb}}$ *and*

$$v_j^{\mathrm{mb}} = 4s_j^{\mathrm{mb}} - 2\sum_{k=1}^{j-1}\pi(j-k)s_k^{\mathrm{mb}} - 2\sum_{k=1}^{j-1}\bar{p}(j-k)v_k^{\mathrm{mb}}, \quad j \geq 2.$$

This recursion allows to write $(v_1^{\mathrm{mb}}, \dots, v_m^{\mathrm{mb}})^\top = A_m(s_1^{\mathrm{mb}}, \dots, s_m^{\mathrm{mb}})^\top$ *for some matrix* $A_m \in \mathbb{R}^{m\times m}$*, such that the covariance matrix* Γ_m^{mb} *may be written as* $\Gamma_m^{\mathrm{mb}} = A_m \Sigma_m^{\mathrm{mb}} A_m^\top$.

In the next theorem it will be shown that the asymptotic variances of the sliding blocks estimators are not larger than the asymptotic variances of their disjoint blocks counterparts. As a consequence, the sliding blocks estimators can be considered at least as efficient and should usually be preferred in practice.

Theorem 4.4. *Under the conditions of Theorem 4.2 we have, for any* $m \in \mathbb{N}$,

$$\Sigma_m^{\mathrm{sb}} \leq_L \Sigma_m^{\mathrm{db}} \quad \text{and} \quad \Gamma_m^{\mathrm{sb}} \leq_L \Gamma_m^{\mathrm{db}},$$

where $\leq_L$ *denotes the Loewner-ordering between symmetric matrices. In particular,* $\mathrm{Var}(s_j^{\mathrm{sb}}) \leq \mathrm{Var}(s_j^{\mathrm{db}})$ *and* $\mathrm{Var}(v_j^{\mathrm{sb}}) \leq \mathrm{Var}(v_j^{\mathrm{db}})$ *for any* $j \in \mathbb{N}_{\geq 1}$.

Example 4.5. In the case that the time series is serially independent, a simple calculation yields $\pi(i) = \mathbb{1}(i = 1)$ and $\pi_2^{(\sigma)}(i, j) = (1-\sigma)\mathbb{1}(i = 1, j = 0) + \sigma\mathbb{1}(i = 1, j = 1)$, which implies

$$p^{(\tau)}(1) = \tau e^{-\tau}, \quad p_2^{(\tau',\tau)}(1, 0) = (\tau' - \tau)e^{-\tau'}, \quad p_2^{(\tau',\tau)}(1, 1) = \tau e^{-\tau'}$$

for $\tau' \geq \tau \geq 0, \tau' \neq 0$. Lengthy computations show that $d_{1,1}^{\mathrm{db}} = 5/108$, such that $\sigma^{2,\mathrm{db}} = \mathrm{Var}(v_1^{\mathrm{db}}) = 20/27 \approx 0.7407$. Likewise, $\sigma^{2,\mathrm{sb}} = \mathrm{Var}(v_1^{\mathrm{sb}}) \approx 0.3790$. The competing blocks estimator $\hat{\pi}_n^{(\tau),\mathrm{Rob}}$ from [26] is known to satisfy

$$\sqrt{k_n}\{\hat{\pi}_n^{(\tau),\mathrm{Rob}}(1) - \pi(1)\} \xrightarrow{d} \mathcal{N}(0, \mu^2(\tau)), \quad \mu^2(\tau) = e^\tau(\tau + (1-\tau)^2 - e^{-\tau}).$$

see Corollary 4.2 in that reference or p. 3300 in [25]. It is worth to mention that μ^2 is strictly increasing with $\sigma^{2,\mathrm{db}} < \mu^2(\tau)$ iff $\tau > 0.7573$.

Recall that $\theta = \{\sum_{j=1}^\infty j\pi(j)\}^{-1}$. As a consequence, following Hsing [16],Robert [26], the extremal index θ may be estimated by

$$\hat{\theta}_n^{\mathrm{mb}}(m) = \left\{\sum_{j=1}^m j\hat{\pi}_n^{\mathrm{mb}}(j)\right\}^{-1}, \quad \mathrm{mb} \in \{\mathrm{db}, \mathrm{sb}\}, \tag{4.4}$$

for sufficiently large m. More precisely, $\hat{\theta}_n^{\mathrm{mb}}(m)$ should be considered an estimator for the partial sum approximation $\theta(m) = \{\sum_{j=1}^m j\pi(j)\}^{-1}$. The following result is an immediate consequence of Corollary 4.3, see also Corollary 4.2 in [26].

Corollary 4.6. *Under the conditions of Theorem 4.1* (mb = db) *or Theorem 4.2* (mb = sb) *we have, for any* $m \in \mathbb{N}$ *and as* $n \to \infty$,

$$\sqrt{k_n}\{\hat{\theta}_n^{\mathrm{mb}}(m) - \theta(m)\} \xrightarrow{d} -\left\{\sum_{j=1}^m j\pi(j)\right\}^{-2} \sum_{j=1}^m j v_j^{\mathrm{mb}} \sim \mathcal{N}(0, \sigma_{\mathrm{mb}}^2(m)),$$

where $\sigma_{\mathrm{mb}}^2(m) = \left\{\sum_{j=1}^m j\pi(j)\right\}^{-4} (1, \dots, m)\Gamma_m^{\mathrm{mb}}(1, \dots, m)^\top$.

5. Finite-sample results

A simulation study was carried out to analyze the finite-sample performance of the introduced estimators and to compare them with estimators from the literature. Results are presented for the following three time series models which were also considered in [26] (with a slightly different ARMAX-model).

- **ARMAX-model**:

$$X_s = \max\{\alpha X_{s-1}, (1-\alpha)Z_s\}, \qquad s \in \mathbb{Z},$$

where $\alpha \in [0,1)$ and $(Z_s)_s$ is an i.i.d. sequence of standard Fréchet random variables. We consider $\alpha = 0.5$ resulting in $\theta = 0.5$ and $\pi(1) = 0.5$, $\pi(2) = 0.25$, $\pi(3) = 0.125$, $\pi(4) = 0.0625$ and $\pi(5) = 0.03125$ by Perfekt [24].

- **Squared ARCH-model**:

$$X_s = (2 \times 10^{-5} + \lambda X_{s-1})Z_s^2, \qquad s \in \mathbb{Z},$$

where $\lambda \in (0,1)$ and where $(Z_s)_s$ denotes an i.i.d. sequence of standard normal random variables. We consider $\lambda = 0.5$, for which the simulated values $\theta = 0.727$ and $\pi(1) = 0.751$, $\pi(2) = 0.168$, $\pi(3) = 0.055$, $\pi(4) = 0.014$ and $\pi(5) = 0.008$ were obtained in [14].

- **AR-model**:

$$X_s = r^{-1}X_{s-1} + Z_s, \qquad s \in \mathbb{Z},$$

where $(Z_s)_s$ is an i.i.d. sequence of random variables that are uniformly distributed on $\{0, 1/r, \ldots, (r-1)/r\}$. We consider $r = 4$, for which the simulated values $\theta = 0.75$ and $\pi(1) = 0.75$, $\pi(2) = 0.1875$, $\pi(3) = 0.0469$, $\pi(4) = 0.0117$ and $\pi(5) = 0.0029$ were obtained in [24].

In all scenarios the sample size was fixed to $n = 2\,000$, attention was restricted to $\pi(m)$ for $m \le 5$, and the block size b was chosen from the set $\{6, 8, \ldots, 36, 38\}$. Note that it is not sensible to use block sizes smaller than m, as the summands making up $\hat{\hat{p}}_n(m)$ are necessarily zero in such a case, which eventually results in a large bias. All results are based on $N = 500$ simulation runs each.

For completeness, and inspired by Northrop [23], a slight modification of the estimators from Section 2 has been considered as well. For its motivation, note that $X_s > M_{ni}^{\mathrm{mb}}$ iff $\hat{F}_n(X_s) > 1 - \hat{Z}_{ni}^{\mathrm{mb}}/b_n$ (a.s.), where $\hat{Z}_{ni}^{\mathrm{mb}} = b_n\{1 - \hat{F}_n(M_{ni}^{\mathrm{mb}})\}$ with the empirical c.d.f. $\hat{F}_n(x) = n^{-1}\sum_{i=1}^n \mathbb{1}(X_i \le x)$. For large block size b_n, we further have $\hat{Z}_{ni}^{\mathrm{mb}} \approx \hat{Y}_{ni}^{\mathrm{mb}} = -b_n \log \hat{F}_n(M_{ni}^{\mathrm{mb}})$, which suggests to define

$$\hat{\hat{p}}_n^{y,\mathrm{db}}(m) = \frac{1}{k_n(k_n-1)} \sum_{\substack{i,i'=1 \\ i \ne i'}}^{k_n} \mathbb{1}\Big\{\sum_{s \in I_{i'}} \mathbb{1}\big(\hat{F}_n(X_s) > 1 - \hat{Y}_{ni}^{\mathrm{db}}/b_n\big) = m\Big\},$$

$$\hat{\hat{p}}_n^{y,\mathrm{sb}}(m) = \frac{1}{|D_n|} \sum_{(i,i') \in D_n} \mathbb{1}\Big\{\sum_{s \in I_{i'}^{\mathrm{sb}}} \mathbb{1}\Big(\hat{F}_n(X_s) > 1 - \hat{Y}_{ni}^{\mathrm{sb}}/b_n\Big) = m\Big\}.$$

Finally, let $\hat{\pi}_n^{y,\mathrm{mb}}$ be defined in terms of $\hat{\hat{p}}_n^{y,\mathrm{mb}}$ as in (2.6). For the ease of a unified notation, the estimators from Section 2 will subsequently be denoted by $\hat{\hat{p}}_n^{z,\mathrm{mb}}$ and $\hat{\pi}_n^{z,\mathrm{mb}}$.

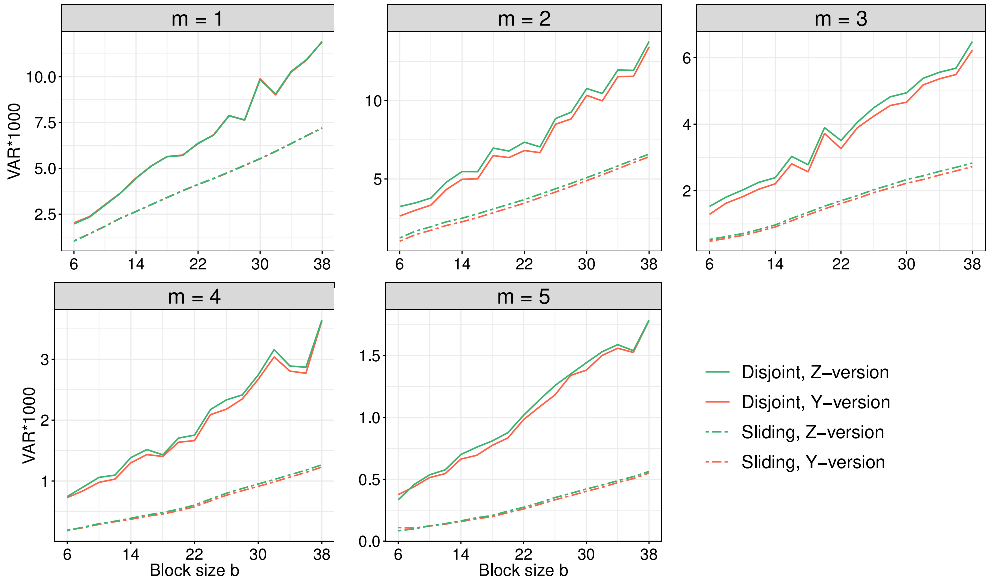

Fig. 1. Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \ldots, 5$.

5.1. Comparison of the introduced estimators for π

In this section we compare the finite-sample performance of the four introduced estimators $\hat{\pi}_n^{z,\mathrm{db}}$, $\hat{\pi}_n^{z,\mathrm{sb}}$, $\hat{\pi}_n^{y,\mathrm{db}}$ and $\hat{\pi}_n^{y,\mathrm{sb}}$.

We start with a detailed analysis of the variance, bias and mean squared error (MSE) as a function of the block size parameter b. Results are only reported for the squared ARCH-model; the corresponding figures for the ARMAX- and AR-model show roughly the same qualitative behavior and can be found in Appendix D in the supplementary material. The variance is depicted in Fig. 1, which can be seen to be increasing in the block size for all estimators. It is further apparent that the Z- and Y-versions behave nearly identical (the curves of the Y-versions are barely visible for $m = 1$ as they are covered by the curves of the Z-versions), whereas the variance of the sliding blocks estimators is considerably smaller than for the disjoint blocks estimators, uniformly over all block sizes. For the Z-version, this is in accordance with the theoretical result from Theorem 4.4.

The bias is presented in Fig. 2 and can be seen to be either increasing or decreasing in b. The largest absolute value of the bias is mostly decreasing in b and attained for small block sizes, which may be explained by the fact that the approximation to the exponential distribution in (2.4) becomes better. The bias curves for the sliding blocks estimators are smoother than for the disjoint blocks versions, which may be explained by the fact that no observations have to be discarded when b is not a divisor of n. One can further see that the Y-versions exhibit a substantially smaller absolute bias for small block sizes (except for $m = 2$); an observation that has also been made in [5]. However, we observe that neither of our estimators can be said to be overall superior with regard to the smallest bias.

The mean squared error is outlined in Fig. 3. In many cases, the MSE-curves show a similar behavior as the variance-curves for large block sizes, since there the variance is dominating over the squared bias. Likewise, the large squared bias for small block sizes can

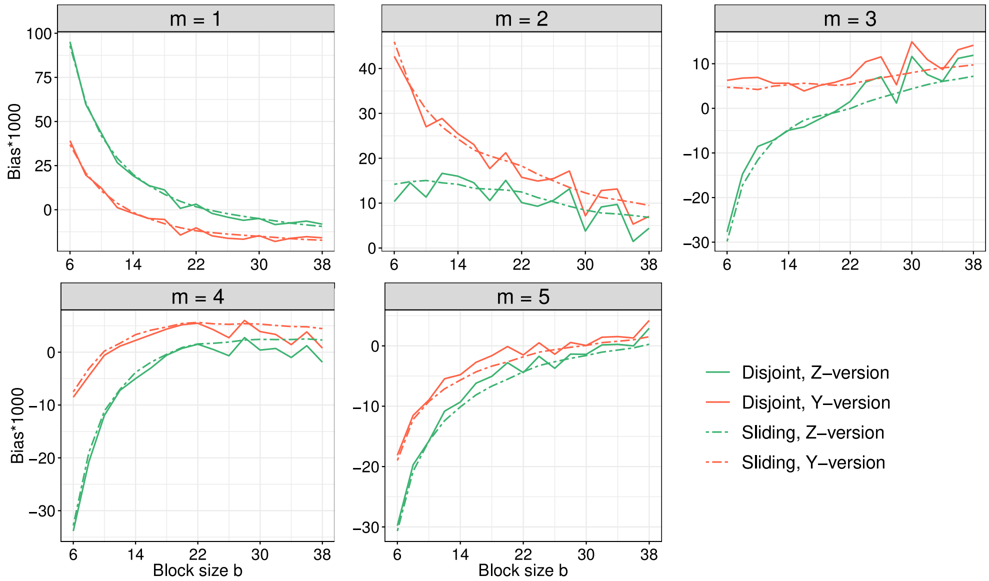

Fig. 2. Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \dots, 5$.

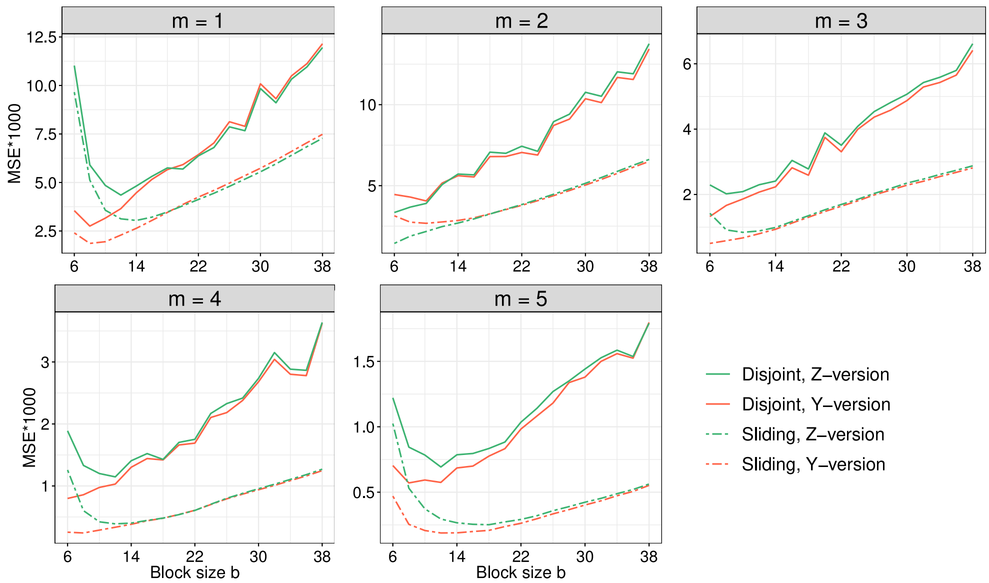

Fig. 3. Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \dots, 5$.

be identified in the MSE-curves as well, eventually resulting in a typical u-shape. Again, the Y-versions perform better for small block sizes (except for $m = 2$). Moreover, the sliding blocks estimators outperform the disjoint blocks estimators with regard to the MSE. Since this qualitative behavior holds uniformly over all models under consideration, we omit the disjoint blocks estimators in the subsequent discussion.

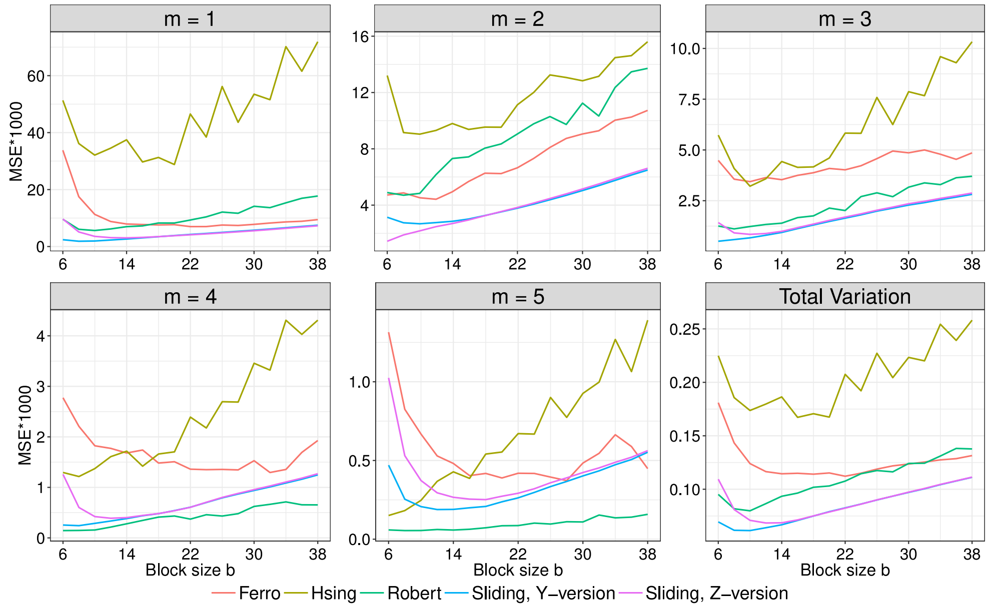

Fig. 4. Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \ldots, 5$ and the total variation between $\hat{\pi}$ and π.

We finally remark on the choice of the block size parameter in practical applications, which is a difficult problem in general, with no universal (yet optimal) solution. In the specific context of estimating $\pi(j)$ up to $j \leq m$, one may apply standard eye-based approaches, like identifying plateaus in the plot $b \mapsto \hat{\pi}_n(j)$, possibly for each value of j separately, or, for convenience, just for $j = m$ (see also [4], page 2328).

5.2. *Comparison with competing estimators for π*

In this section, we compare the performance of our sliding blocks estimators for $\pi(m)$ with the following competitors from the literature: the integrated version of the blocks estimator from [26] with parameters $\sigma = 0.7$ and $\phi = 1.3$ (page 276 in that reference), the blocks estimator from [16] with $v_n = X_{n-\lfloor n/s_n \rfloor :n}$, where $s_n = 2(b_n - 3)$ (see (1.4) in [16] and (1.2) in [26], where a similar same choice has been made), and the inter-exceedance times estimator from [12] with $N = 3k_n$ (see equation (4.12) in that reference).

In Fig. 4, the MSE is plotted as a function of the blocksize in the squared ARCH-model (see Appendix D in the supplementary material for other models and the bias- and variance-curves). In addition, in order to evaluate the overall accuracy of the estimators, Fig. 4 also presents results on a version of the total variation distance between the cluster size distribution and its estimator defined by

$$d_{\mathrm{TV}(5)}(\hat{\pi}, \pi) := \frac{1}{2} \sum_{m=1}^{5} |\hat{\pi}_n(m) - \pi(m)|.$$

We can see that the MSE is mostly decreasing for small blocksizes and tends to increase from an intermediate blocksize onwards, which is due to the common bias–variance-tradeoff. The MSE-curves of our sliding blocks estimators are very smooth compared to the competing

Table 1
Minimal mean squared error multiplied by 10^3 for the AR-model, the maxAR-model and the squared ARCH-model. The estimator with the row-wise smallest MSE is in boldface.

Model	m	$\pi(m)$	Sliding, Z	Sliding, Y	Robert	Hsing	Ferro
AR	1	0.750	8.255	6. 746	12.951	20.094	**4.007**
	2	0.188	2.374	**1.636**	8.732	7.352	3.683
	3	0.047	1.679	1.301	**1.090**	2.864	1.423
	4	0.012	0.861	0.497	**0.113**	0.277	0.236
	5	0.003	0.159	0.088	**0.008**	0.017	0.035
ARMAX	1	0.500	2.642	**1.650**	6.819	5.318	5.343
	2	0.250	0.495	**0.434**	2.177	1.586	3.460
	3	0.125	**0.186**	0.311	1.763	1.816	2.118
	4	0.062	0.252	**0.179**	1.144	1.011	2.454
	5	0.031	0.206	**0.086**	0.474	0.390	2.350
sqARCH	1	0.751	3.044	**1.860**	5.631	28.795	7.001
	2	0.168	**1.436**	2.677	4.706	9.043	4.418
	3	0.055	0.842	**0.503**	1.111	3.214	3.439
	4	0.014	0.389	0.242	**0.145**	1.215	1.294
	5	0.008	0.251	0.188	**0.055**	0.150	0.372

estimators and lie uniformly below their MSE-curves in many cases. Generally, the estimator by Robert and our sliding blocks estimators outperform the estimators by Ferro and Hsing in almost all scenarios under consideration. With regard to the total variation distance, one can see that our sliding blocks estimators outperform the competitors almost uniformly over all blocksizes.

The minimum values of the mean squared error (minimum over b) are of particular interest. They are presented for all models under consideration in Table 1. The estimator $\hat{\pi}_n^{z,\mathrm{sb}}$ wins twice, $\hat{\pi}_n^{y,\mathrm{sb}}$ wins seven times and Robert's estimator five times, while the estimators by Ferro wins once. It is worth to mention that the sliding blocks estimators cover all minimum values within the ARMAX-model, and Robert's estimator seems to perform especially well for estimating $\pi(m)$ in case that value is very close to zero. The latter may be explained by the fact that the estimator is forced to be non-negative by definition (which is not the case for our estimators), which results in a high proportion of zero estimates if $\pi(m) \approx 0$ and hence a small estimation variance.

Remark 5.1. The performance of the extremal index estimator $\hat{\theta}_n^{\mathrm{mb}}(m)$ from (4.4) was also investigated in the above setting. More precisely, attention was restricted to the z-version of $\hat{\theta}_n^{\mathrm{sb}}(8)$, which was then compared with other estimators for the extremal index from the literature: the bias-reduced sliding blocks estimator from [27], the integrated version of the blocks estimator from [26], the intervals estimator from [13], the ML-estimator from [29], the pseudo ML-estimator from [4] and the CFG-estimator from [5] (based on sliding blocks). While our estimator showed a similar qualitative behavior (see Fig. D.19 in the supplementary material), it was found to be mostly inferior to its competitors, whence we cannot recommend it for further use.

6. Proofs of the main results

Throughout the paper, we use the notation $a_n \lesssim b_n$ if there exists a constant C not depending on n such that $a_n \leq C b_n$. We start by arguing that we may slightly redefine the estimators, which will greatly simplify the notational complexity. For $m \in \mathbb{N}_{\ge 0}$, let

$$\tilde{\hat{p}}_n^{\mathrm{db}}(m) = \frac{1}{k_n^2} \sum_{i,i'=1}^{k_n} \mathbb{1}\Big\{ \sum_{s \in I_{i'}^{\mathrm{db}}} \mathbb{1}\big(X_s > M_{ni}^{\mathrm{db}}\big) = m \Big\},$$

$$\tilde{\hat{p}}_n^{\mathrm{sb}}(m) = \frac{1}{(n-b_n+1)^2} \sum_{i,i'=1}^{n-b_n+1} \mathbb{1}\Big\{ \sum_{s \in I_{i'}^{\mathrm{sb}}} \mathbb{1}\big(X_s > M_{ni}^{\mathrm{sb}}\big) = m \Big\}.$$

Since $\sum_{s \in I_i^{\mathrm{db}}} \mathbb{1}\big(X_s > M_{ni}^{\mathrm{db}}\big) = \mathbb{1}(m=0)$ and $|\tilde{\hat{p}}_n^{\mathrm{db}}| \le 1$, we have, for $m \ge 1$,

$$\tilde{\hat{p}}_n^{\mathrm{db}}(m) - \hat{\hat{p}}_n^{\mathrm{db}}(m) = \Big(1 - \frac{k_n}{k_n - 1}\Big) \tilde{\hat{p}}_n^{\mathrm{db}}(m) = o_{\mathbb{P}}(k_n^{-1/2}).$$

As a consequence, throughout the proof, we may redefine $\hat{\hat{p}}_n^{\mathrm{db}}(m) = \tilde{\hat{p}}_n^{\mathrm{db}}(m)$. A similar argument holds for the sliding blocks version, whence we subsequently set $\hat{\hat{p}}_n^{\mathrm{sb}}(m) = \tilde{\hat{p}}_n^{\mathrm{sb}}(m)$.

Next, we will introduce some additional notation. For $s \in \mathbb{Z}$, let $U_s = F(X_s)$. For $\tau > 0$ and $m \in \mathbb{N}_{\ge 0}$, let

$$p_n^{(\tau),\mathrm{db}}(m) = \frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}\big(N_{b_n,i}^{(\tau),\mathrm{db}} = m\big),$$

$$p_n^{(\tau),\mathrm{sb}}(m) = \frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbb{1}\big(N_{b_n,i}^{(\tau),\mathrm{sb}} = m\big),$$

where, for mb $\in$ {db, sb},

$$N_{b_n,i}^{(\tau),\mathrm{mb}} = \sum_{s \in I_i^{\mathrm{mb}}} \mathbb{1}\Big(U_s > 1 - \frac{\tau}{b_n}\Big).$$

Denote the rescaled estimation error by

$$e_{n,m}^{\mathrm{mb}}(\tau) = \sqrt{k_n}\Big\{ p_n^{(\tau),\mathrm{mb}}(m) - \varphi_{n,m}(\tau) \Big\}, \tag{6.1}$$

where $\varphi_{n,m}$ is defined in Condition 3.1(iv). Note that the disjoint blocks version $e_{n,m}^{\mathrm{db}}$ has been extensively studied in [26]. Next, let $Z_{ni}^{\mathrm{mb}} = b_n\{1 - F(M_{ni}^{\mathrm{mb}})\}$ and, for $x > 0$, let

$$\hat{H}_n^{\mathrm{db}}(x) = \frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}(Z_{ni}^{\mathrm{db}} \le x), \qquad \hat{H}_n^{\mathrm{sb}}(x) = \frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbb{1}(Z_{ni}^{\mathrm{sb}} \le x),$$

denote the empirical c.d.f. of $Z_{n1}^{\mathrm{db}}, \dots, Z_{nk_n}^{\mathrm{db}}$ and $Z_{n1}^{\mathrm{sb}}, \dots, Z_{n,n-b_n+1}^{\mathrm{sb}}$, respectively. Finally, recall $H(x) = (1 - e^{-\theta x})\mathbb{1}(x \ge 0)$, the c.d.f. of the exponential distribution with parameter θ.

Proof of Theorem 4.1. By continuity of F, we have $U_s > 1 - Z_{ni}^{\mathrm{db}}/b_n$ iff $X_s > M_{ni}^{\mathrm{db}}$ almost surely, whence we may write, for $j \in \mathbb{N}_{\ge 1}$,

$$\hat{\hat{p}}_n^{\mathrm{db}}(j) \stackrel{a.s.}{=} k_n^{-1} \sum_{i=1}^{k_n} p_n^{(Z_{ni}^{\mathrm{db}}),\mathrm{db}}(j).$$

We may thus decompose

$$s_{n,j}^{\mathrm{db}} = \sqrt{k_n}\{\hat{\hat{p}}_n^{\mathrm{db}}(j) - \bar{p}(j)\} \stackrel{a.s.}{=} A_{n1} + A_{n2} + A_{n3}, \tag{6.2}$$

where

$$A_{n1} = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \bigg\{ \int_0^\infty \mathbb{1}\big(N_{b_n,i}^{(\tau),\mathrm{db}} = j\big) - \varphi_{n,j}(\tau)\, \mathrm{d}H(\tau) + \varphi_{n,j}(Z_{ni}^{\mathrm{db}}) - \mathrm{E}[\varphi_{n,j}(Z_{n1}^{\mathrm{db}})] \bigg\},$$

$$A_{n2} = \int_0^\infty e_{n,j}^{\mathrm{db}}(\tau)\, \mathrm{d}(\hat{H}_n^{\mathrm{db}} - H)(\tau), \qquad A_{n3} = \sqrt{k_n}\big\{\mathrm{E}[\varphi_{n,j}(Z_{n1}^{\mathrm{db}})] - \bar{p}(j)\big\}.$$

We have $A_{n3} = o(1)$ by Condition 3.1(iv) and $A_{n2} = o_{\mathbb{P}}(1)$ by Lemma A.1 in the supplementary material. Hence, setting

$$W_{n,i}^{\mathrm{db}}(j) = \int_0^\infty \mathbb{1}\big(N_{b_n,i}^{(\tau),\mathrm{db}} = j\big) - \varphi_{n,j}(\tau)\,\mathrm{d}H(\tau) + \varphi_{n,j}(Z_{ni}^{\mathrm{db}}) - \mathrm{E}[\varphi_{n,j}(Z_{ni}^{\mathrm{db}})], \tag{6.3}$$

we have $s_{n,j}^{\mathrm{db}} = k_n^{-1/2} \sum_{i=1}^{k_n} W_{n,i}^{\mathrm{db}}(j) + o_{\mathbb{P}}(1)$. The assertion then follows from

$$\frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \big(W_{n,i}^{\mathrm{db}}(1), \ldots, W_{n,i}^{\mathrm{db}}(m)\big) \xrightarrow{d} \mathcal{N}_m(0, \Sigma_m^{\mathrm{db}})$$

as a consequence of Lemma A.2 in the supplementary material. □

Proof of Theorem 4.2. As in the proof of Theorem 4.1, we have

$$\hat{\bar{p}}_n^{\mathrm{sb}}(j) \overset{a.s.}{=} \frac{1}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} p_n^{(Z_{ni}^{\mathrm{sb}}),\mathrm{sb}}(j).$$

Similarly as in (6.2) and by using the bias Condition 3.1(vi), we can thus write

$$\begin{aligned} s_{n,j}^{\mathrm{sb}} &= \sqrt{k_n}\big\{\hat{\bar{p}}_n^{\mathrm{sb}}(j) - \bar{p}(j)\big\} \\ &\overset{a.s.}{=} \frac{\sqrt{k_n}}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} W_{n,i}^{\mathrm{sb}}(j) + \int_0^\infty e_{n,j}^{\mathrm{sb}}(\tau)\, \mathrm{d}(\hat{H}_n^{\mathrm{sb}} - H)(\tau) + o(1), \end{aligned}$$

where $W_{n,i}^{\mathrm{sb}}$ is defined as in (6.3), but with 'db' replaced by 'sb' everywhere. The assertion then follows from $\int_0^\infty e_{n,j}^{\mathrm{sb}}\, \mathrm{d}(\hat{H}_n^{\mathrm{sb}} - H) = o_{\mathbb{P}}(1)$ by Lemma B.1 in the supplementary material and

$$\frac{\sqrt{k_n}}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} \big(W_{n,i}^{\mathrm{sb}}(1), \ldots, W_{n,i}^{\mathrm{sb}}(m)\big) \xrightarrow{d} \mathcal{N}_m(0, \Sigma_m^{\mathrm{sb}})$$

by Lemma B.2 in the supplementary material. □

Proof of Corollary 4.3. Throughout, we omit the index mb $\in \{\mathrm{db}, \mathrm{sb}\}$. For $j \in \mathbb{N}_{\geq 1}$, set $\varphi_j : \mathbb{R}^{2j-1} \to \mathbb{R}$, $\varphi_j(x) = 4x_j - 2\sum_{k=1}^{j-1} x_{2j-k} x_k$, such that

$$\begin{aligned} \hat{\pi}_n(j) &= \varphi_j(\hat{\bar{p}}_n(1), \ldots, \hat{\bar{p}}_n(j), \hat{\pi}_n(1), \ldots, \hat{\pi}_n(j-1)), \\ \pi(j) &= \varphi_j(\bar{p}(1), \ldots, \bar{p}(j), \pi(1), \ldots, \pi(j-1)). \end{aligned}$$

By Theorems 4.1 and 4.2, we know that $(s_{n,1},\ldots,s_{n,m}) \xrightarrow{d} (s_1,\ldots,s_m) \sim \mathcal{N}_m(0,\Sigma_m)$. To prove the theorem, we use this result and apply induction over m. First,

$$v_{n,1} = \sqrt{k_n}\{\hat{\pi}_n(1) - \pi(1)\} = 4\sqrt{k_n}\{\hat{\bar{p}}_n(1) - \bar{p}(1)\} = 4s_{n,1},$$

such that $(s_{n,1}, s_{2,n}, v_{n,1}) \xrightarrow{d} (s_1, s_2, 4s_1) = (s_1, s_2, v_1)$. Second, assume we have

$$(s_{n,1},\ldots,s_{n,m},v_{n,1},\ldots,v_{n,m-1}) \xrightarrow{d} (s_1,\ldots,s_m,v_1,\ldots,v_{m-1})$$

for $m \geq 2$. Then, the delta-method implies

$$\begin{aligned} v_{n,m} &= \sqrt{k_n}\{\varphi_m(\hat{\bar{p}}_n(1),\ldots,\hat{\bar{p}}_n(m),\hat{\pi}_n(1),\ldots,\hat{\pi}_n(m-1)) - \\ &\qquad \varphi_m(\bar{p}(1),\ldots,\bar{p}(m),\pi(1),\ldots,\pi(m-1))\} \\ &= \varphi_m'(\bar{p}(1),\ldots,\bar{p}(m),\pi(1),\ldots,\pi(m-1)) \\ &\qquad \cdot (s_{n,1},\ldots,s_{n,m},v_{n,1},\ldots,v_{n,m-1})^\top + o_{\mathbb{P}}(1) \\ &\xrightarrow{d} 4s_m - 2\sum_{k=1}^{m-1}\pi(m-k)s_k - 2\sum_{k=1}^{m-1}\bar{p}(m-k)v_k =: v_m, \end{aligned}$$

where φ_m' denotes the gradient of φ_m. We obtain that

$$(s_{n,1},\ldots,s_{n,m},v_{n,1},\ldots,v_{n,m}) \xrightarrow{d} (s_1,\ldots,s_m,v_1,\ldots,v_m).$$

Since every v_j is a linear function of $(s_1,\ldots,s_m) \sim \mathcal{N}_m(0,\Sigma_m)$, the vector $(v_1,\ldots,v_m)$ follows an m-dimensional normal distribution as well. □

Proof of Theorem 4.4. We only need to prove $\Sigma_m^{\mathrm{sb}} \leq_L \Sigma_m^{\mathrm{db}}$; the assertion regarding Γ_m^{mb} is an immediate consequence.

In the following, we assume for simplicity that U_s and Z_{ni}^{sb} are measurable with respect to the $\mathcal{B}_{::}^{\varepsilon_1}$-sigma fields with ε_1 from Condition 3.1(ii); the general case can be treated by multiplication with suitable indicator functions as in the proofs in the appendices. Now, $\Sigma_m^{\mathrm{sb}} \leq_L \Sigma_m^{\mathrm{db}}$ is equivalent to

$$\mathrm{Var}\Big(\sum_{j=1}^m a_j s_j^{\mathrm{sb}}\Big) \leq \mathrm{Var}\Big(\sum_{j=1}^m a_j s_j^{\mathrm{db}}\Big) \tag{6.4}$$

for any $a = (a_1,\ldots,a_m)^\top \in \mathbb{R}^m$. To prove the latter, we are going to apply Lemma A.10 in [31]. For $j \in \{1,\ldots,m\}$ and $i \in \mathbb{N}_{\geq 1}$, let $S_{n,i} = \sum_{j=1}^m a_j V_{n,i}(j)$, where

$$V_{n,i}(j) = \int_0^\infty \mathbb{1}\Big(\sum_{s\in J_i}\mathbb{1}\Big(U_s > 1 - \frac{\tau}{b_n}\Big) = j\Big)\,\mathrm{d}H(\tau) + \varphi_{n,j}\Big(b_n(1 - \max_{s\in J_i} U_s)\Big)$$

and where $J_i = \{i, i+1, \ldots, i+b_n-1\}$. Note that $I_i^{\mathrm{db}} = J_{(i-1)b_n+1}$ for $i \in \{1,\ldots,k_n\}$ and that $I_i^{\mathrm{sb}} = J_i$ for $i \in \{1,\ldots,n-b_n+1\}$. By the proofs of Theorems 4.1 and 4.2 we can write

$$\mathrm{Var}\Big(\sum_{j=1}^m a_j s_j^{\mathrm{sb}}\Big) = \lim_{n\to\infty}\mathrm{Var}\Big(\sqrt{\frac{n}{b_n}}\frac{1}{n}\sum_{i=1}^n S_{n,i}\Big),$$

$$\mathrm{Var}\Big(\sum_{j=1}^m a_j s_j^{\mathrm{db}}\Big) = \lim_{n\to\infty}\mathrm{Var}\Big(\sqrt{\frac{n}{b_n}}\frac{b_n}{n}\sum_{i=1}^{\lfloor n/b_n\rfloor} S_{n,(i-1)b_n+1}\Big).$$

For $h \in \mathbb{N}_{\geq 0}$, set $\gamma_n(h) = \mathrm{Cov}(S_{n,1}, S_{n,h+1})$; note that $S_{n,1}, \ldots, S_{n,n-b_n+1}$ is stationary. Since $0 \leq V_{n,i}(j) \leq 2$ we obtain

$$|\gamma_n(h)| \leq \sum_{j,j'=1}^{m} |a_j a_{j'}| \big| \mathrm{Cov}(V_{n,1}(j), V_{n,h+1}(j')) \big| \leq 8 \sum_{j,j'=1}^{m} |a_j a_{j'}|,$$

such that $\sup_{n \in \mathbb{N}, h \in \mathbb{N}_{\geq 0}} |\gamma_n(h)| < \infty$. Further, by Lemma 3.9 in [8] we have

$$\begin{aligned} |\gamma_n(h + b_n)| &\leq 4 \sum_{j,j'=1}^{m} |a_j a_{j'}| \; \|V_{n,1}(j)\|_\infty \; \|V_{n,h+b_n+1}(j')\|_\infty \; \alpha_{\varepsilon_1}(1 + h) \\ &\leq C' \times \alpha_{\varepsilon_1}(h) \end{aligned}$$

for some constant C' depending on $a_1, \ldots, a_m$ only. This implies

$$\sum_{h=1}^{\infty} |\gamma_n(h + b_n)| \leq C' \sum_{h=1}^{\infty} \alpha_{\varepsilon_1}(h) < \infty$$

by Condition 3.1(ii). Relation (6.4) then follows from Lemma A.10 in [31]. □

7. On sliding blocks estimators for compound probabilities

Throughout this section, we derive an extension of Theorem 4.1 in [26] from the disjoint blocks process $e_{n,m}^{\mathrm{db}}$ in (6.1) to the sliding blocks version $e_{n,m}^{\mathrm{sb}}$. The result is used for proving Theorem 4.2, but might in fact be of general interest for statistics for time series extremes based on sliding blocks. For $m \in \mathbb{N}_{\geq 0}$ and $\tau \geq 0$, let

$$E_{n,m}^{\mathrm{sb}}(\tau) = \big(e_{n,0}^{\mathrm{sb}}(\tau), \ldots, e_{n,m}^{\mathrm{sb}}(\tau)\big).$$

For simplicity, we impose the same mixing conditions as needed for the results in Section 4. We denote by $D([0, \infty))$ the space of real-valued càdlàg functions on $[0, \infty)$, equipped with the metric $d(f, g) = \sum_{k=1}^{\infty} 2^{-k} \min(\sup_{x \in [0,k]} |f(x) - g(x)|, 1)$.

Theorem 7.1. *Suppose that Condition 3.1(i)–(ii) is met and that, additionally, $\sqrt{k_n}\beta_{\varepsilon_2}(b_n) = o(1)$ for some $\varepsilon_2 > 0$. Then, for any $m \in \mathbb{N}_{\geq 1}$,*

$$E_{n,m}^{\mathrm{sb}} \xrightarrow{d} E_m^{\mathrm{sb}} \quad \text{in} \quad D([0, \infty))^{m+1},$$

where $E_m^{\mathrm{sb}}(\cdot) = \big(e_0^{\mathrm{sb}}(\cdot), \ldots, e_m^{\mathrm{sb}}(\cdot)\big)$ is a centered Gaussian process with continuous sample paths, almost surely, and with covariance functional given by, for $0 \leq \tau \leq \tau'$ and $j, j' \in \{0, \ldots, m\}$,

$$\begin{aligned} \mathrm{Cov}\big(e_j^{\mathrm{sb}}(\tau), e_{j'}^{\mathrm{sb}}(\tau')\big) &= 2 \int_0^1 \mathrm{Cov}\big(\mathbb{1}(X_\xi^{(\tau)} = j), \mathbb{1}(Y_\xi^{(\tau')} = j')\big) \, \mathrm{d}\xi \\ &= 2 \int_0^1 H_{j,j'}^{(\tau,\tau')}(\xi) \, \mathrm{d}\xi - 2 p^{(\tau)}(j) p^{(\tau')}(j'), \end{aligned}$$

where $X_\xi^{(\tau)} = Y_\xi^{(\tau)} = N_E^{(\tau)}$ in distribution with joint probability mass function

$$\begin{aligned} H_{j,j'}^{(\tau,\tau')}(\xi) &= \mathbb{P}\big(X_\xi^{(\tau)} = j, Y_\xi^{(\tau')} = j'\big) \\ &= \sum_{l=0}^{j} \sum_{r=j-l}^{j'} p^{(\xi\tau)}(l) p^{(\xi\tau')}(j' - r) p_2^{((1-\xi)\tau',(1-\xi)\tau)}(r, j - l). \end{aligned}$$

Proof. The result is a consequence of the next two lemmas. □

It is worthwhile to mention that one may add the classical tail empirical process $\bar{e}_n$ as an $(m+2)$th-coordinate to $E^{\mathrm{sb}}_{n,m}$ (just as in Theorem 4.2 in [26]). Additional conditions as in that reference would be necessary then, including a moment bound on the increments of $\tau \mapsto N_n^{(\tau)}$ and adapted mixing conditions. Details are omitted for the sake of brevity.

Further, it is interesting to note that in specific cases the asymptotic variance of the sliding blocks process can be seen to be smaller than that of its disjoint blocks counterpart. For instance, some tedious but straightforward calculations show that in the i.i.d. model, for $\tau = 1$,

$$\mathrm{Var}(e_1^{\mathrm{sb}}(\tau)) = 2e^{-2}(2e-5) \approx 0.1182, \quad \mathrm{Var}(e_2^{\mathrm{sb}}(\tau)) = e^{-2}(5e-13) \approx 0.0800,$$

which are substantially smaller than

$$\mathrm{Var}(e_1^{\mathrm{db}}(\tau)) = e^{-1} - e^{-2} \approx 0.2325, \quad \mathrm{Var}(e_2^{\mathrm{db}}(\tau)) = \frac{1}{2e} - \frac{1}{4e^2} \approx 0.1501,$$

where e_j^{db} denotes the disjoint blocks limit from Theorem 4.1 in [26].

Lemma 7.2 (*Tightness.*)**.** *Under the conditions of Theorem 7.1, and for any $0 < \phi < \infty$ and $m \in \mathbb{N}_{\geq 0}$, the process $(E^{\mathrm{sb}}_{n,m})_{n\in\mathbb{N}}$ is asymptotically tight in $D([0,\phi])^{m+1}$.*

Lemma 7.3 (*Fidis-Convergence.*)**.** *Suppose that Condition 3.1(i)–(ii) are met. Then, for $m \in N_{\geq 0}$ and $\tau_1, \dots, \tau_r \geq 0, r \in \mathbb{N}_{\geq 1}$, we have*

$$\big(E^{\mathrm{sb}}_{n,m}(\tau_1), \dots, E^{\mathrm{sb}}_{n,m}(\tau_r)\big) \xrightarrow{d} \big(E^{\mathrm{sb}}_{m}(\tau_1), \dots, E^{\mathrm{sb}}_{m}(\tau_r)\big).$$

8. Proofs for Section 7

Proof of Lemma 7.2. Since marginal asymptotic tightness implies joint asymptotic tightness, it is sufficient to show asymptotic tightness of $e^{\mathrm{sb}}_{n,j}$ for fixed $j \in \mathbb{N}_{\geq 0}$. Subsequently, we omit the upper index sb.

For sufficiently large n, the summands making up $(e_{n,j}(\tau))_{\tau\in[0,\phi]}$ are only depending on $U_s^{\varepsilon_2} = U_s \mathbb{1}(U_s > 1-\varepsilon_2)$, whence the beta-mixing coefficients based on the $\mathcal{B}^{\varepsilon_2}_{:,:}$-sigma fields become available; in particular, we may use that $\sqrt{k_n}\beta_{\varepsilon_2}(b_n) = o(1)$.

Let $b_n' = 2b_n$ and $\mathcal{K}_n = (n - b_n + 1)/(2b_n') = O(n/b_n)$. For simplicity we assume that $\mathcal{K}_n$ is an integer (otherwise, a potential remainder block can be shown to be asymptotically negligible). For $k \in \{1, \dots, \mathcal{K}_n\}$, define

$$A_k = \{2(k-1)b_n' + 1, \dots, 2(k-1)b_n' + b_n'\},$$
$$B_k = \{(2k-1)b_n' + 1, \dots, (2k-1)b_n' + b_n'\}$$

such that $|A_k| = |B_k| = b_n'$ and $A_1 \cup B_1 \cup \dots \cup A_{\mathcal{K}_n} \cup B_{\mathcal{K}_n} = \{1, \dots, n - b_n + 1\}$. Next, to simplify the notation, define

$$N_i^{(\tau)} = N_{b_n,i}^{(\tau),\mathrm{sb}} = \sum_{s\in I_i^{\mathrm{sb}}} \mathbb{1}\big(U_s > 1 - \tau/b_n\big)$$

Write $e_{n,j}(\tau) = A_{n,j}(\tau) + B_{n,j}(\tau)$, where

$$A_{n,j}(\tau) = \frac{1}{\sqrt{\mathcal{K}_n}} \sum_{k=1}^{\mathcal{K}_n} \big\{\bar{A}_{n,j,k}(\tau) - \mathrm{E}[\bar{A}_{n,j,k}(\tau)]\big\},$$

with

$$\bar{A}_{n,j,k}(\tau) = \frac{\sqrt{k_n \mathcal{K}_n}}{n - b_n + 1} \sum_{i \in A_k} \mathbb{1}(N_i^{(\tau)} = j),$$

and where $B_{n,j}$ is defined analogously, but with A_k replaced by B_k. Since finite sums of asymptotically tight processes are asymptotically tight, it is sufficient to show tightness of $A_{n,j}$ and $B_{n,j}$. We only treat $A_{n,j}$. Note that $(U_t : t \in I_i^{\text{sb}})_{i \in A_k}$ only depends on

$$U_n^{(k)} := \big(U_{2(k-1)b_n'+1}, \dots, U_{2(k-1)b_n'+b_n'+b_n-1}\big) \in \mathbb{R}^{3b_n-1}$$

by the definition of I_i^{sb}. Further, write

$$\bar{A}_{n,j,k}(\tau) = h_{n,j}^{(\tau)}(U_n^{(k)}) - h_{n,j-1}^{(\tau)}(U_n^{(k)}),$$

where

$$h_{n,j}^{(\tau)} : \mathbb{R}^{3b_n-1} \to \mathbb{R},\ u \mapsto \frac{\sqrt{k_n \mathcal{K}_n}}{n - b_n + 1} \sum_{i=1}^{2b_n} \mathbb{1}\Big(\sum_{t \in I_i^{\text{sb}}} \mathbb{1}\big(u_t > 1 - \tau/b_n\big) \le j\Big),$$

As a consequence, we may write $A_{n,j}(\tau) = C_{n,j}(\tau) - C_{n,j-1}(\tau)$, where

$$C_{n,j}(\tau) = \frac{1}{\sqrt{\mathcal{K}_n}} \sum_{k=1}^{\mathcal{K}_n} \big\{h_{n,j}^{(\tau)}(U_n^{(k)}) - \mathrm{E}[h_{n,j}^{(\tau)}(U_n^{(k)})]\big\}$$

for $j \ge 0$ and $C_{n,-1} = 0$. It is hence sufficient to show asymptotic tightness of $C_{n,j}$ for fixed $j \in \mathbb{N}_{\ge 0}$. By the coupling lemma in [3] (see Lemma C.2 in the supplementary material), we can inductively construct an array $\big\{(\tilde{U}_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\}_{k=1,\dots,\mathcal{K}_n}$ such that

$$\begin{aligned}
&\text{(i)} \quad \forall\, k \in \{1, \dots, \mathcal{K}_n\} : \big\{(\tilde{U}_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\} \stackrel{D}{=} \big\{(U_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\},\\
&\text{(ii)} \quad \forall\, k \in \{1, \dots, \mathcal{K}_n\} :\\
&\qquad\qquad \mathbb{P}\Big(\big\{(\tilde{U}_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\} \ne \big\{(U_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\}\Big) \le \beta_{\varepsilon_2}(b_n),\\
&\text{(iii)} \quad \big\{(\tilde{U}_s)_{s \in I_i^{\text{sb}}} : i \in A_k\big\}_{k=1,\dots,\mathcal{K}_n} \ \text{ is (row-wise) independent.}
\end{aligned} \tag{8.1}$$

Set

$$\tilde{N}_i^{(\tau)} = \sum_{s \in I_i^{\text{sb}}} \mathbb{1}\big(\tilde{U}_s > 1 - \tau/b_n\big)$$

and let $\tilde{C}_{n,j}(\tau)$ be defined as $C_{n,j}(\tau)$ but with $U_n^{(k)}$ substituted by

$$\tilde{U}_n^{(k)} := \big(\tilde{U}_{2(k-1)b_n'+1}, \dots, \tilde{U}_{2(k-1)b_n'+b_n'+b_n-1}\big) \in \mathbb{R}^{3b_n-1}.$$

We begin by showing that

$$\sup_{\tau \in [0,\phi]} |C_{n,j}(\tau) - \tilde{C}_{n,j}(\tau)| = o_{\mathbb{P}}(1), \tag{8.2}$$

for which we write

$$C_{n,j}(\tau) - \tilde{C}_{n,j}(\tau) = \frac{\sqrt{k_n}}{n - b_n + 1} \sum_{k=1}^{\mathcal{K}_n} \sum_{i \in A_k} \big\{\mathbb{1}(N_i^{(\tau)} \le j) - \mathbb{1}(\tilde{N}_i^{(\tau)} \le j)\big\}.$$

For fixed $k \in \{1, \ldots, \mathcal{K}_n\}$, we obtain

$$\begin{aligned}
&\Big|\sum_{i \in A_k} \mathbb{1}(N_i^{(\tau)} \le j) - \mathbb{1}(\tilde N_i^{(\tau)} \le j)\Big| \\
&\le 2b_n \times \mathbb{1}\big(\{N_i^{(\tau)} : i \in A_k\} \ne \{\tilde N_i^{(\tau)} : i \in A_k\}\big) \\
&\le 2b_n \times \mathbb{1}\big(\{(U_s)_{s \in I_i^{\mathrm{sb}}} : i \in A_k\} \ne \{(\tilde U_s)_{s \in I_i^{\mathrm{sb}}} : i \in A_k\}\big).
\end{aligned}$$

Hence, by Item (ii) in (8.1), we obtain

$$\mathrm{E}\Big[\sup_{\tau \in [0,\phi]} |C_{n,j}(\tau) - \tilde C_{n,j}(\tau)|\Big] \le \frac{\sqrt{k_n}\mathcal{K}_n}{n - b_n + 1} 4 b_n \beta_{\varepsilon_2}(b_n) = 2\sqrt{k_n}\beta_{\varepsilon_2}(b_n),$$

which converges to zero by assumption. Markov's inequality implies (8.2). As a consequence, it suffices to show that the process $(\tilde C_{n,j})_{n\in\mathbb{N}}$ is tight.

Note that by the items (i) and (iii) in (8.1), $(\tilde U_n^{(k)})_{k=1,\ldots,\mathcal{K}_n}$ is a row-wise i.i.d. triangular array. Let $(\mathcal{F}, \rho) = ([0,\phi], |\cdot|)$ and

$$Z_{nk}(\tau) = \mathcal{K}_n^{-1/2} h_{n,j}^{(\tau)}(\tilde U_n^{(k)}), \quad \tau \in \mathcal{F},\ k = 1, \ldots, \mathcal{K}_n,$$

such that

$$\tilde C_{n,j}(\tau) = \sum_{k=1}^{\mathcal{K}_n} Z_{nk}(\tau) - \mathrm{E}[Z_{nk}(\tau)].$$

In the following, we apply Theorem 2.11.9 in [30]. First, note that

$$\begin{aligned}
&\sup_{\tau \in \mathcal{F}, u \in \mathbb{R}^{3b_n - 1}} |h_{n,j}^{(\tau)}(u)| \\
&= \sup_{\tau \in \mathcal{F}, u \in \mathbb{R}^{3b_n - 1}} \Big|\frac{\sqrt{k_n \mathcal{K}_n}}{n - b_n + 1} \sum_{i=1}^{2b_n} \mathbb{1}\Big(\sum_{t \in I_i^{\mathrm{sb}}} \mathbb{1}\big(u_t > 1 - \tau/b_n\big) \le j\Big)\Big| \\
&\le \frac{\sqrt{k_n \mathcal{K}_n}}{n - b_n + 1} 2b_n = \sqrt{\frac{n}{n - b_n + 1}} \le 2
\end{aligned}$$

since $n - b_n + 1 \ge n/2$ for sufficiently large n. Consequently, $\|Z_{nk}\|_{\mathcal{F}} := \sup_{\tau\in[0,\phi]} |Z_{nk}(\tau)| \le 2\mathcal{K}_n^{-1/2}$, such that the first condition in Theorem 2.11.9 in [30] is satisfied. Next, let $\|\cdot\|_{n,2}$ be the norm $\|f\|_{n,2} = \mathrm{E}[f(\tilde U_n^{(1)})^2]^{1/2}$. We prove the subsequent inequality: for any $\tau, \tau' \in [0, \phi + 1]$,

$$\|h_{n,j}^{(\tau)} - h_{n,j}^{(\tau')}\|_{n,2} \le 2|\tau - \tau'|^{1/2}. \tag{8.3}$$

Indeed, by Jensen's inequality

$$\begin{aligned}
\|h_{n,j}^{(\tau)} - h_{n,j}^{(\tau')}\|_{n,2}^2 &= \frac{k_n \mathcal{K}_n}{(n - b_n + 1)^2} \mathrm{E}\Big[\Big(\sum_{i=1}^{2b_n} \{\mathbb{1}(\tilde N_i^{(\tau)} \le j) - \mathbb{1}(\tilde N_i^{(\tau')} \le j)\}\Big)^2\Big] \\
&\le \frac{k_n \mathcal{K}_n (2b_n)^2}{(n - b_n + 1)^2} \mathrm{E}\big[\{\mathbb{1}(\tilde N_1^{(\tau)} \le j) - \mathbb{1}(\tilde N_1^{(\tau')} \le j)\}^2\big] \\
&\le 4\,\mathrm{E}\big[\{\mathbb{1}(\tilde N_1^{(\tau)} \le j) - \mathbb{1}(\tilde N_1^{(\tau')} \le j)\}^2\big] \qquad (8.4)
\end{aligned}$$

for sufficiently large n. Without loss of generality, let $\tau \le \tau'$. Since $z \mapsto \mathbb{1}(\tilde N_1^{(z)} \le j)$ is monotonically decreasing, one has

$$\{\mathbb{1}(\tilde N_1^{(\tau)} \le j) - \mathbb{1}(\tilde N_1^{(\tau')} \le j)\}^2 = \mathbb{1}(\tilde N_1^{(\tau)} \le j) - \mathbb{1}(\tilde N_1^{(\tau')} \le j)$$

$$= \mathbb{1}(\tilde N_1^{(\tau)} \le j < \tilde N_1^{(\tau')})$$
$$\le \mathbb{1}(\tilde N_1^{(\tau')} - \tilde N_1^{(\tau)} \ge 1).$$

Hence, by (i) in (8.1), the expression on the right-hand side of (8.4) can be bounded by

$$4 \cdot \mathbb{P}(N_1^{(\tau')} - N_1^{(\tau)} \ge 1) \le 4 \cdot \mathrm{E}[N_1^{(\tau')} - N_1^{(\tau)}] = 4(\tau' - \tau)$$

as asserted in (8.3). Therefore, we obtain

$$\sup_{|\tau-\tau'|<\delta_n} \sum_{k=1}^{\mathcal{K}_n} \mathrm{E}\big[(Z_{nk}(\tau) - Z_{nk}(\tau'))^2\big] = \sup_{|\tau-\tau'|<\delta_n} \|h_{n,j}^{(\tau)} - h_{n,j}^{(\tau')}\|_{n,2}^2 \le 4\delta_n,$$

which converges to 0 for every $\delta_n \to 0$. It remains to bound the bracketing number $N_{[]}$ as given on page 211 in [30]. First, we construct a cover of $\mathcal{F}$. For $\varepsilon \in (0,1)$ and $a \in \mathbb{N}_{\ge 1}$ let $D_{\varepsilon,a} = [(a-1)\varepsilon^2/4, a\varepsilon^2/4]$. Then

$$[0,\phi] \subset \bigcup_{a\in\{1,2,\dots,M_\varepsilon\}} D_{\varepsilon,a} \subset [0,\phi+1], \qquad M_\varepsilon = \lfloor 4(\phi+1)/\varepsilon^2 \rfloor.$$

Now, since $\tau \mapsto h_{n,j}^{(\tau)}$ is monotonically decreasing, we may choose, for any $\tau \in [0,\phi]$, an integer $a \in \{1,\dots,M_\varepsilon\}$ such that

$$h_{n,j}^{(a\varepsilon^2/4)} \le h_{n,j}^{(\tau)} \le h_{n,j}^{((a-1)\varepsilon^2/4)}.$$

Therefore, using (8.3), we get that, for any $a \in \{1,\dots,M_\varepsilon\}$,

$$\begin{aligned}
&\sum_{k=1}^{\mathcal{K}_n} \mathrm{E}\Big[\sup_{\tau,\tau'\in D_{\varepsilon,a}} |Z_{nk}(\tau) - Z_{nk}(\tau')|^2\Big] \\
&\le \mathcal{K}_n^{-1} \sum_{k=1}^{\mathcal{K}_n} \mathrm{E}\Big[|h_{n,j}^{(a\varepsilon^2/4)}(\tilde U_n^{(k)}) - h_{n,j}^{((a-1)\varepsilon^2/4)}(\tilde U_n^{(k)})|^2\Big] \\
&= \|h_{n,j}^{(a\varepsilon^2/4)} - h_{n,j}^{((a-1)\varepsilon^2/4)}\|_{n,2}^2 \le \varepsilon^2.
\end{aligned}$$

Hence, the bracketing number as on page 211 in [30] is obviously bounded by M_ε, such that the last condition in Theorem 2.11.9 in that reference is satisfied, and the proof is finished. □

Proof of Lemma 7.3. By the Cramér–Wold device it suffices to show that

$$D_n = \sum_{l=1}^{r}\sum_{j=0}^{m} \lambda_{l,j}\, e_{n,j}^{\mathrm{sb}}(\tau_l) \xrightarrow{d} \sum_{l=1}^{r}\sum_{j=0}^{m} \lambda_{l,j}\, e_j^{\mathrm{sb}}(\tau_l) = D \tag{8.5}$$

for any $\lambda_{l,j} \in \mathbb{R}$. Throughout the proof, let $I_i = I_i^{\mathrm{db}}$ and write

$$D_n = \sum_{j=1}^{k_n-1}\sum_{s\in I_j}\sum_{l=1}^{r}\sum_{j=0}^{m} \lambda_{l,j} \frac{\sqrt{k_n}}{n-b_n+1}\big\{\mathbb{1}(N_{b_n,s}^{(\tau_l),\mathrm{sb}} = j) - \varphi_{n,j}(\tau)\big\} + o_{\mathbb{P}}(1).$$

Let $k_n^* < k_n$ be an integer sequence with $k_n^* \to \infty$ and $k_n^* = o(k_n^{1/4})$. For $q_n^* = \lfloor k_n/(k_n^*+2) \rfloor \to \infty$ and $p = 1,\dots,q_n^*$, define

$$J_p^+ = \bigcup_{i=(p-1)(k_n^*+2)+1}^{p(k_n^*+2)-2} I_i, \qquad J_p^- = I_{p(k_n^*+2)-1} \cup I_{p(k_n^*+2)}.$$

Thus, we have decomposed the observation period into q_n^* 'big blocks' J_p^+ of size $k_n^* b_n$, which are separated by 'small blocks' J_p^- of size $2b_n$. We may hence rewrite $D_n = V_n^+ + V_n^- + o_{\mathbb{P}}(1)$, where

$$V_n^\pm = \frac{1}{\sqrt{q_n^*}} \sum_{p=1}^{q_n^*} T_{np}^\pm$$

and, for $p \in \{1, \dots, q_n^*\}$,

$$T_{np}^\pm = \sqrt{\frac{q_n^*}{k_n}} \sum_{s \in J_p^\pm} \sum_{l=1}^{r} \sum_{j=0}^{m} \lambda_{l,j} \frac{n}{n - b_n + 1} \frac{1}{b_n} \big\{ \mathbb{1}(N_{b_n,s}^{(\tau_l),\mathrm{sb}} = j) - \varphi_{n,j}(\tau_l) \big\}.$$

Let us show that $V_n^- = o_{\mathbb{P}}(1)$. For that purpose, take $\varepsilon_1 \in (0,1)$ from Condition 3.1. Observe that, for sufficiently large n, T_{np}^- only depends on $U_s^{\varepsilon_1} = U_s \mathbb{1}(U_s > 1 - \varepsilon_1)$ with $s \in \{(p(k_n^* + 2) - 2)b_n + 1, \dots, p(k_n^* + 2)b_n + b_n - 1\}$, whence, in particular, the alpha-mixing coefficients based on the $\mathcal{B}_{:,:}^{\varepsilon_1}$-sigma fields become available. Now, since $\mathrm{E}[V_n^-] = 0$, it is enough to prove $\mathrm{Var}(V_n^-) = o(1)$. By stationarity,

$$\mathrm{Var}(V_n^-) \le 3\,\mathrm{Var}(T_{n1}^-) + 2 \sum_{p=2}^{q_n^*} \big| \mathrm{Cov}(T_{n1}^-, T_{n,p+1}^-) \big|. \tag{8.6}$$

Observing that $|J_1^-| = 2b_n$ and $n/(n - b_n + 1) \le 2$ for sufficiently large n, we have

$$|T_{n1}^-| \le 4 \sqrt{\frac{q_n^*}{k_n}} \sum_{l=1}^{r} \sum_{j=0}^{m} |\lambda_{l,j}| = O\Big(\sqrt{\frac{q_n^*}{k_n}}\Big) = O\Big(\frac{1}{\sqrt{k_n^*}}\Big) = o(1), \tag{8.7}$$

which implies $\mathrm{Var}(T_{n1}^-) = o(1)$ as well. Next, by Lemma 3.9 in [8], Condition 3.1(ii) and since $T_{n,p}^-$ is bounded (and since by construction the observations making up T_{n1}^- and $T_{n,p+1}^-$ are separated by $pk_n^* b_n$ observations), we obtain

$$\sum_{p=2}^{q_n^*} \big| \mathrm{Cov}(T_{n1}^-, T_{n,p+1}^-) \big| \le 4 \|T_{n1}^-\|_\infty \sum_{p=2}^{q_n^*} \alpha_{\varepsilon_1}(pk_n^* b_n)$$
$$\lesssim o(1) \sum_{p=2}^{q_n^*} (pk_n^* b_n)^{-\eta} = o(1),$$

such that altogether $\mathrm{Var}(V_n^-) = o(1)$ by (8.6).

It remains to show that V_n^+ converges in distribution to D from (8.5). Note that T_{np}^+ and $T_{np'}^+$ are based on $U_s^{\varepsilon_1}$-observations that are at least b_n observations apart for $p \ne p'$. This allows to apply an argument based on characteristic functions to reason that $(T_{np}^+)_p$ may be considered independent. Indeed, let $(\tilde T_{np}^+)_p$ denote iid random variables with $\tilde T_{np}^+ =_d T_{np}^+$. Recursively applying Lemma 3.11 in [8], we have, for any $t \in \mathbb{R}$,

$$\Big| \mathrm{E}\Big[\exp\Big(\frac{\mathbf{i}t}{\sqrt{q_n^*}} \sum_{p=1}^{q_n^*} T_{np}^+\Big)\Big] - \mathrm{E}\Big[\exp\Big(\frac{\mathbf{i}t}{\sqrt{q_n^*}} \sum_{p=1}^{q_n^*} \tilde T_{np}^+\Big)\Big] \Big|$$
$$= \Big| \mathrm{E}\Big[\exp\Big(\frac{\mathbf{i}t}{\sqrt{q_n^*}} \sum_{p=1}^{q_n^*} T_{np}^+\Big)\Big] - \prod_{p=1}^{q_n^*} \mathrm{E}\Big[\exp\Big(\frac{\mathbf{i}t}{\sqrt{q_n^*}} T_{np}^+\Big)\Big] \Big| \lesssim q_n^* \alpha_{\varepsilon_1}(b_n),$$

where $\mathbf{i}$ denotes the imaginary unit. The upper bound satisfies $q_n^* \alpha_{\varepsilon_1}(b_n) \le k_n \alpha_{\varepsilon_1}(b_n) \lesssim k_n b_n^{-\eta} = o(1)$ by Condition 3.1(ii), whence, by Lévy's continuity theorem, the weak limits of $(q_n^*)^{-1/2} \sum_{p=1}^{q_n^*} T_{np}^+$ and $(q_n^*)^{-1/2} \sum_{p=1}^{q_n^*} \tilde{T}_{np}^+$ coincide, provided one of the limits exists. This implies that $(T_{np}^+)_{p=1,\dots,q_n^*}$ may be considered independent, which is assumed from now on.

As in (8.7), we obtain that $|T_{np}^+| = O(\sqrt{k_n^*})$, whence

$$\frac{\sum_{p=1}^{q_n^*} \mathrm{E}\big[|T_{np}^+|^3\big]}{\big\{\sum_{p=1}^{q_n^*} \mathrm{Var}(T_{np}^+)\big\}^{3/2}} = O\big(k_n^{-1/2}(k_n^*)^2\big) = o(1),$$

provided that $\lim_{n\to\infty} \mathrm{Var}(T_{n1}^+)$ exists. In this case, the Lyapunov condition is satisfied and the central limit theorem implies that V_n^+ converges in distribution to a centered normal distribution with variance $\lim_{n\to\infty} \mathrm{Var}(T_{n1}^+)$. Note that

$$T_{n1}^+ = \sum_{l=1}^{r} \sum_{j=0}^{m} \lambda_{l,j}\; e_{n^*,j}^{\mathrm{sb}}(\tau_l) + R_n,$$

where $R_n \to 0$ in $L_2(\mathbb{P})$, with $n^* = k_n^* b_n$ and that our assumptions in Condition 3.1 still hold if n and k_n are substituted by n^* and k_n^*. The limiting variance of the above expression is calculated in Lemma 8.1 and is seen to be of the required form. □

Lemma 8.1. *Suppose that Condition 3.1(i)–(ii) are met. Then, for $0 \le \tau \le \tau'$ and $j, j' \in \mathbb{N}_{\ge 0}$, we have*

$$\lim_{n\to\infty} \mathrm{Cov}(e_{n,j}^{\mathrm{sb}}(\tau), e_{n,j'}^{\mathrm{sb}}(\tau')) = 2\int_0^1 \mathrm{Cov}\big(\mathbb{1}(X_\xi^{(\tau)} = j), \mathbb{1}(Y_\xi^{(\tau')} = j')\big)\,\mathrm{d}\xi$$
$$= 2\int_0^1 H_{j,j'}^{(\tau,\tau')}(\xi)\,\mathrm{d}\xi - 2p^{(\tau)}(j)p^{(\tau')}(j'),$$

where $X_\xi^{(\tau)} = Y_\xi^{(\tau)} = N_E^{(\tau)}$ in distribution with joint probability mass function

$$H_{j,j'}^{(\tau,\tau')}(\xi) = \mathbb{P}\big(X_\xi^{(\tau)} = j, Y_\xi^{(\tau')} = j'\big)$$
$$= \sum_{l=0}^{j} \sum_{r=j-l}^{j'} p^{(\xi\tau)}(l)p^{(\xi\tau')}(j'-r)p_2^{((1-\xi)\tau',(1-\xi)\tau)}(r, j-l). \tag{8.8}$$

Proof of Lemma 8.1. Fix $0 \le \tau \le \tau'$ and $j, j' \in \mathbb{N}_{\ge 0}$. Note that we may replace U_s by $U_s^{\varepsilon 1} = U_s \mathbb{1}(U_s > 1 - \varepsilon_1)$ for n large enough, where $\varepsilon = \varepsilon_1$ is from Condition 3.1(ii). Write

$$r_n(\tau,\tau') \equiv \mathrm{Cov}(e_{n,j}^{\mathrm{sb}}(\tau), e_{n,j'}^{\mathrm{sb}}(\tau'))$$
$$= \frac{k_n}{(n-b_n+1)^2} \sum_{s,t=1}^{n-b_n+1} \mathrm{Cov}(\mathbb{1}(N_{b_n,s}^{(\tau),\mathrm{sb}} = j), \mathbb{1}(N_{b_n,t}^{(\tau'),\mathrm{sb}} = j'))$$
$$= \frac{k_n}{(n-b_n+1)^2} \sum_{i,i'=1}^{k_n-1} \sum_{s\in I_i} \sum_{t\in I_{i'}} \mathrm{Cov}(A_s, B_t) + o(1),$$

where $A_s = \mathbb{1}(N_{b_n,s}^{(\tau),\mathrm{sb}} = j)$, $B_t = \mathbb{1}(N_{b_n,t}^{(\tau'),\mathrm{sb}} = j')$ and $I_i = I_i^{\mathrm{db}}$. By stationarity, we may further write

$$\begin{aligned} r_n(\tau,\tau') &= \frac{k_n(k_n-1)}{(n-b_n+1)^2}\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_1} B_t\Big) \\ &\quad + \frac{k_n}{(n-b_n+1)^2}\sum_{i=2}^{k_n-1}(k_n-i)\Big\{\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_i} B_t\Big) \\ &\qquad + \operatorname{Cov}\Big(\sum_{s\in I_q} A_s, \sum_{t\in I_1} B_t\Big)\Big\} + o(1) \\ &= T_{n1} + T_{n2} + T_{n3} + T_{n4} + o(1), \end{aligned} \tag{8.9}$$

where

$$\begin{aligned} T_{n1} &= \frac{k_n(k_n-1)}{(n-b_n+1)^2}\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_1} B_t\Big) \\ T_{n2} &= \frac{k_n(k_n-2)}{(n-b_n+1)^2}\Big\{\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_2} B_t\Big) + \operatorname{Cov}\Big(\sum_{s\in I_2} A_s, \sum_{t\in I_1} B_t\Big)\Big\} \\ T_{n3} &= \frac{k_n(k_n-3)}{(n-b_n+1)^2}\Big\{\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_3} B_t\Big) + \operatorname{Cov}\Big(\sum_{s\in I_3} A_s, \sum_{t\in I_1} B_t\Big)\Big\} \\ T_{n4} &= \frac{k_n}{(n-b_n+1)^2}\sum_{i=4}^{k_n-1}(k_n-i)\Big\{\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_i} B_t\Big) \\ &\qquad + \operatorname{Cov}\Big(\sum_{s\in I_i} A_s, \sum_{t\in I_1} B_t\Big)\Big\}. \end{aligned}$$

Next, we show that

$$T_{n3} = o(1), \quad T_{n4} = o(1). \tag{8.10}$$

For that purpose note that $\sum_{s\in I_1} A_s$ and $\sum_{t\in I_i} B_s$ are at least $(i-3)b_n$ observations apart. By Lemma 3.9 in [8] we obtain

$$\Big|\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_i} B_t\Big)\Big| \le 4\, b_n^2 \alpha_{\varepsilon_1}((i-3)b_n),$$

such that

$$|T_{n4}| \le \frac{8\, k_n^2 b_n^2}{(n-b_n+1)^2}\sum_{i=4}^{k_n-1}\alpha_{\varepsilon_1}((i-3)b_n) \lesssim \frac{n^2 b_n^{-\eta}}{(n-b_n+1)^2}\sum_{i=1}^{k_n-4} i^{-\eta} = o(1)$$

since $\eta > 1$ by Condition 3.1(ii). Regarding T_{n3}, note that

$$\begin{aligned} \Big|\operatorname{Cov}\Big(\sum_{s\in I_1} A_s, \sum_{t\in I_3} B_t\Big)\Big| &\le \sum_{t=2b_n+1}^{3b_n}\Big|\operatorname{Cov}\Big(\sum_{s=1}^{b_n} A_s, B_t\Big)\Big| \\ &\le 4\, b_n \sum_{t=2b_n+1}^{3b_n}\alpha_{\varepsilon_1}(t-2b_n) = 4\, b_n \sum_{t=1}^{b_n}\alpha_{\varepsilon_1}(t) \end{aligned}$$

by Lemma 3.9 in [8], which implies

$$|T_{n3}| \le 8\frac{k_n(k_n-3)b_n}{(n-b_n+1)^2}\sum_{t=1}^{b_n}\alpha_{\varepsilon_1}(t) \lesssim \frac{k_n^2 b_n}{(n-b_n+1)^2}\sum_{t=1}^{b_n} t^{-\eta} = O(b_n^{-1}).$$

Hence, (8.10) is shown.

Next, consider T_{n1}. Since $k_n(k_n-1)/(n-b_n+1)^2 = 1/b_n^2 + o(1)$ and $\mathrm{E}[A_s] \to p^{(\tau)}(j)$ and $\mathrm{E}[B_t] \to p^{(\tau')}(j')$, we may write

$$T_{n1} = \frac{1}{b_n^2}\sum_{s,t=1}^{b_n}\mathrm{E}[A_s B_t] - p^{(\tau)}(j)p^{(\tau)}(j') + o(1),$$

Next, we have $b_n^{-2}\sum_{s,t=1}^{b_n}\mathrm{E}[A_s B_t] = \int_0^1 f_n(\xi)\,\mathrm{d}\xi$, where, for $\xi \in (0,1)$,

$$\begin{aligned} f_n(\xi) &= \frac{1}{b_n}\sum_{s,t=1}^{b_n}\mathrm{E}[A_s B_t]\mathbb{1}\Big(\xi \in \Big[\frac{t-1}{b_n},\frac{t}{b_n}\Big)\Big) \\ &= \frac{1}{b_n}\sum_{s=1}^{b_n}\mathrm{E}[A_s B_{\lfloor b_n\xi\rfloor+1}] = \int_0^1 \varphi_n(\xi,z)\,\mathrm{d}z, \end{aligned}$$

where, for $z \in (0,1)$,

$$\begin{aligned} \varphi_n(\xi,z) &= \sum_{s=1}^{b_n}\mathrm{E}[A_s B_{\lfloor b_n\xi\rfloor+1}]\mathbb{1}\Big(z \in \Big[\frac{s-1}{b_n},\frac{s}{b_n}\Big)\Big) \\ &= \mathrm{E}\big[A_{\lfloor b_n z\rfloor+1}B_{\lfloor b_n\xi\rfloor+1}\big] = \mathbb{P}\big(N^{(\tau),\mathrm{sb}}_{b_n,\lfloor b_n z\rfloor+1} = j,\, N^{(\tau').\mathrm{sb}}_{b_n,\lfloor b_n\xi\rfloor+1} = j'\big) \end{aligned} \tag{8.11}$$

For $0 < z \le \xi < 1$, we may rewrite

$$\begin{aligned} \varphi_n(\xi,z) = \sum_{l=0}^{j}\sum_{r=0}^{j'}\mathbb{P}\Big(& N_{\lfloor b_n z\rfloor+1:\lfloor b_n\xi\rfloor}(\tau) = l,\, N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n z\rfloor+b_n}(\tau) = j-l, \\ & N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n z\rfloor+b_n}(\tau') = r,\, N_{\lfloor b_n z\rfloor+b_n+1:\lfloor b_n\xi\rfloor+b_n}(\tau') = j'-r\Big). \end{aligned} \tag{8.12}$$

where, for $s, s' \in \mathbb{N}_{\ge 1}$ with $s \le s'$ and $\tau \ge 0$,

$$N_{s:s'}(\tau) = \sum_{t=s}^{s'}\mathbb{1}\Big(U_t > 1 - \frac{\tau}{b_n}\Big).$$

We will next argue that the first, the intersection of the second and the third and the fourth of the four events in each summand in (8.12) may be considered independent. Indeed, for any fixed $y > 0$ and any integer sequence q_n converging to infinity with $q_n = o(b_n)$, we have

$$\mathbb{P}\big(N_{1:q_n}(y) = 0\big) \ge 1 - q_n\mathbb{P}\big(U_1 > 1 - \frac{y}{b_n}\big) = 1 - \frac{yq_n}{b_n} \to 1$$

As a consequence, we may intersect the events inside the sum in (8.12) with

$$\{N_{\lfloor b_n\xi\rfloor-q_n:\lfloor b_n\xi\rfloor}(\tau) = 0,\, N_{\lfloor b_n z\rfloor+b_n+1:\lfloor b_n z\rfloor+b_n+1+q_n}(\tau') = 0\}. \tag{8.13}$$

at the expense of a $O(q_n/b_n)$-term. On the intersected event, we must then have $N_{\lfloor b_n z\rfloor+1:\lfloor b_n\xi\rfloor-q_n}(\tau) = l$ and $N_{\lfloor b_n z\rfloor+b_n+q_n:\lfloor b_n\xi\rfloor+b_n}(\tau) = j'-r$. After discarding the events in (8.13) again, we are left with an intersection of three events that are based on observations that are at least q_n observations apart. As a consequence, at the expense of an $\alpha_{\varepsilon_1}(q_n)$-error, they may be considered independent. Finally, we may sneak in the omitted observations once again at the expense of an additional $O(q_n/b_n)$-term, and we arrive at

$$\begin{aligned}\varphi_n(\xi, z) = \sum_{l=0}^{j}\sum_{r=0}^{j'} &\mathbb{P}\Big(N_{\lfloor b_n z\rfloor+1:\lfloor b_n\xi\rfloor}(\tau) = l\Big)\\ &\times \mathbb{P}\Big(N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n z\rfloor+b_n}(\tau) = j-l, N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n z\rfloor+b_n}(\tau') = r\Big)\\ &\times \mathbb{P}\Big(N_{\lfloor b_n z\rfloor+b_n+1:\lfloor b_n\xi\rfloor+b_n}(\tau') = j'-r\Big)\\ &+ O(\alpha_{\varepsilon_1}(q_n)) + O(q_n/b_n) \end{aligned} \tag{8.14}$$

which converges to

$$\begin{aligned} H(\xi - z) &= H_{j,j'}^{(\tau,\tau')}(\xi - z)\\ &= \sum_{l=0}^{j}\sum_{r=j-l}^{j'} p^{((\xi-z)\tau)}(l)p^{((\xi-z)\tau')}(j'-r)p_2^{((1-\xi+z)\tau',(1-\xi+z)\tau)}(r, j-l) \end{aligned}$$

by Condition 3.1(i), where $H_{j,j'}^{(\tau,\tau')}$ is defined in (8.8). Changing the roles of z and ξ, we obtain

$$\varphi_n(\xi, z) \to H(\xi - z)\mathbb{1}(z \le \xi) + H(z - \xi)\mathbb{1}(z > \xi).$$

For fixed $\xi \in (0, 1)$, $\sup_{n\in\mathbb{N}} \|\varphi_n(\xi, \cdot)\|_\infty \le 1$, such that the dominated convergence theorem implies

$$f_n(\xi) = \int_0^1 \varphi_n(\xi, z)\, \mathrm{d}z \to \int_0^\xi H(\xi - z)\, \mathrm{d}z + \int_\xi^1 H(z - \xi)\, \mathrm{d}z.$$

Moreover, since $\|f_n\|_\infty \le 1$, dominated convergence also implies that

$$\begin{aligned}\lim_{n\to\infty} T_{n1} &= \int_0^1 \int_0^\xi H(\xi - z)\,\mathrm{d}z + \int_\xi^1 H(z - \xi)\,\mathrm{d}z\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j')\\ &= 2\int_0^1\int_0^\xi H(\xi - z)\,\mathrm{d}z\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j')\\ &= 2\int_0^1 (1-\xi)H(\xi)\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j'), \end{aligned} \tag{8.15}$$

where the last step is due to Fubini's theorem.

It remains to treat T_{n2} in (8.9), which consists of two summands, say $T_{n2,1}$ and $T_{n2,2}$. By similar arguments as for T_{n1}, the first summand $T_{n2,1}$ can be written as

$$\begin{aligned} T_{n2,1} &= \frac{1}{b_n^2}\sum_{s=1}^{b_n}\sum_{t=b_n+1}^{2b_n} \mathrm{E}[A_s B_t] - p^{(\tau)}(j)p^{(\tau')}(j') + o(1)\\ &= \int_0^1\int_0^1 \psi_n(\xi, z)\,\mathrm{d}z\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j') + o(1) \end{aligned}$$

where

$$\psi_n(\xi, z) = \mathrm{E}\big[A_{\lfloor b_n\xi\rfloor+1} B_{\lfloor (z+1)b_n\rfloor+1}\big]$$

$$= \mathbb{P}\Big(N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n\xi\rfloor+b_n}(\tau) = j, N_{\lfloor b_n(z+1)\rfloor+1:\lfloor b_n(z+1)\rfloor+b_n}(\tau') = j'\Big).$$

If $\xi \le z$, then $\lfloor b_n\xi\rfloor + b_n \le \lfloor b_n(1+z)\rfloor + 1$ and we can manipulate the above probability as in (8.14), such that it equals

$$\psi_n(\xi, z) = \mathbb{P}\Big(N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n\xi\rfloor+b_n}(\tau) = j\Big)\mathbb{P}\Big(N_{\lfloor b_n(z+1)\rfloor+1:\lfloor b_n(z+1)\rfloor+b_n}(\tau') = j'\Big) + O(\alpha_{\varepsilon_1}(q_n)) + O(q_n/b_n),$$

which converges to $p^{(\tau)}(j)p^{(\tau')}(j')$. In the case $z \le \xi$, we again need to separate the sums as in (8.14) and obtain that $\psi_n(\xi, z)$ equals

$$\sum_{l=0}^{j}\sum_{r=0}^{j'} \mathbb{P}\Big(N_{\lfloor b_n\xi\rfloor+1:\lfloor b_n(z+1)\rfloor}(\tau) = l\Big) \times \mathbb{P}\Big(N_{\lfloor b_n(z+1)\rfloor+1:\lfloor b_n\xi\rfloor+b_n}(\tau) = j-l, N_{\lfloor b_n(z+1)\rfloor+1:\lfloor b_n\xi\rfloor+b_n}(\tau') = r\Big) \times \mathbb{P}\Big(N_{\lfloor b_n\xi\rfloor+b_n+1:\lfloor b_n(z+1)\rfloor+b_n}(\tau') = j'-r\Big) + O(\alpha_{\varepsilon_1}(q_n)) + O(q_n/b_n)$$

which converges to

$$H(1-(\xi-z)) = H_{j,j'}^{(\tau,\tau')}(1-(\xi-z)) = \sum_{l=0}^{j}\sum_{r=j-l}^{j'} p^{((1-\xi+z)\tau)}(l)p^{((1-\xi+z)\tau')}(j'-r)p_2^{((\xi-z)\tau',(\xi-z)\tau)}(r, j-l).$$

Since $\|\psi_n\|_\infty \le 1$, dominated convergence implies

$$\lim_{n\to\infty} T_{n2,1} = \int_0^1\int_0^\xi H(1-(\xi-z))\,\mathrm{d}z + \int_\xi^1 p^{(\tau)}(j)p^{(\tau')}(j')\,\mathrm{d}z\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j') = \int_0^1 \xi H(\xi)\,\mathrm{d}\xi - \frac{1}{2}p^{(\tau)}(j)p^{(\tau')}(j')$$

as $n \to \infty$. By symmetry, the second summand in T_{n2} has the same limit, such that

$$\lim_{n\to\infty} T_{n2} = 2\int_0^1 \xi H(\xi)\,\mathrm{d}\xi - p^{(\tau)}(j)p^{(\tau')}(j'), \tag{8.16}$$

where the last equation follows as in (8.15). Altogether, by (8.10), (8.15) and (8.16), we have

$$\lim_{n\to\infty} \mathrm{Cov}\big(e_{n,j}^{\mathrm{sb}}(\tau), e_{n,j'}^{\mathrm{sb}}(\tau')\big) = 2\int_0^1 H_{j,j'}^{(\tau,\tau')}(\xi)\,\mathrm{d}\xi - 2p^{(\tau)}(j)p^{(\tau')}(j')$$

as asserted. □

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation, which is gratefully acknowledged. The authors would like to thank two anonymous referees and an Associate Editor for their constructive comments on an earlier version of this manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spa.2022.03.004.

References

[1] B. Basrak, J. Segers, Regularly varying multivariate time series, Stochastic Process. Appl. 119 (4) (2009) 1055–1080, http://dx.doi.org/10.1016/j.spa.2008.05.004.

[2] J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, Statistics of Extremes: Theory and Applications, in: Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester, 2004.

[3] H.C.P. Berbee, Random Walks with Stationary Increments and Renewal Theory, in: Mathematical Centre Tracts, vol. 112, Mathematisch Centrum, Amsterdam, 1979, p. iii+223.

[4] B. Berghaus, A. Bücher, Weak convergence of a pseudo maximum likelihood estimator for the extremal index, Ann. Statist. 46 (5) (2018) 2307–2335, http://dx.doi.org/10.1214/17-AOS1621.

[5] A. Bücher, T. Jennessen, Method of moments estimators for the extremal index of a stationary time series, Electron. J. Stat. 14 (2) (2020) 3103–3156, http://dx.doi.org/10.1214/20-EJS1734.

[6] Y. Cissokho, R. Kulik, Estimation of cluster functionals for regularly varying time series: sliding blocks estimators, 2020, arXiv:2005.11378.

[7] R.A. Davis, H. Drees, J. Segers, M. Warchoł, Inference on the tail process with application to financial time series modeling, J. Econometrics 205 (2) (2018) 508–525, http://dx.doi.org/10.1016/j.jeconom.2018.01.009.

[8] H. Dehling, W. Philipp, Empirical process techniques for dependent data, in: Empirical Process Techniques for Dependent Data, Birkhäuser Boston, Boston, MA, 2002, pp. 3–113, http://dx.doi.org/10.1007/978-1-4612-0099-4_1.

[9] H. Drees, M. Knežević, Peak-over-threshold estimators for spectral tail processes: random vs deterministic thresholds, Extremes 23 (3) (2020) 465–491, http://dx.doi.org/10.1007/s10687-019-00367-x.

[10] H. Drees, H. Rootzén, Limit theorems for empirical processes of cluster functionals, Ann. Statist. 38 (4) (2010) 2145–2186, http://dx.doi.org/10.1214/09-AOS788.

[11] H. Drees, J. Segers, M. Warchoł, Statistics for tail processes of Markov chains, Extremes 18 (3) (2015) 369–402, http://dx.doi.org/10.1007/s10687-015-0217-1.

[12] C.A.T. Ferro, Statistical Methods for Clusters of Extreme Values (Ph.D. Thesis), Lancaster University, 2003.

[13] C.A.T. Ferro, J. Segers, Inference for clusters of extreme values, J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2) (2003) 545–556, http://dx.doi.org/10.1111/1467-9868.00401.

[14] L. de Haan, S.I. Resnick, H. Rootzén, C.G. de Vries, Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes, Stochastic Process. Appl. 32 (2) (1989) 213–224, http://dx.doi.org/10.1016/0304-4149(89)90076-8.

[15] T. Hsing, Point Processes Associated with Extreme Value Theory, ProQuest LLC, Ann Arbor, MI, 1984, p. 94, URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:8508594. Thesis (Ph.D.)–The University of North Carolina at Chapel Hill.

[16] T. Hsing, Estimating the parameters of rare events, Stochastic Process. Appl. 37 (1) (1991) 117–139, http://dx.doi.org/10.1016/0304-4149(91)90064-J.

[17] T. Hsing, Extremal index estimation for a weakly dependent stationary sequence, Ann. Statist. 21 (4) (1993) 2043–2071, http://dx.doi.org/10.1214/aos/1176349409.

[18] T. Hsing, J. Hüsler, M.R. Leadbetter, On the exceedance point process for a stationary sequence, Probab. Theory Related Fields 78 (1) (1988) 97–112, http://dx.doi.org/10.1007/BF00718038.

[19] R. Kulik, P. Soulier, Heavy-Tailed Time Series, in: Springer Series in Operations Research and Financial Engineering, Springer New York, 2020, URL: https://books.google.de/books?id=6-HuDwAAQBAJ.

[20] R. Kulik, P. Soulier, O. Wintenberger, The tail empirical process of regularly varying functions of geometrically ergodic Markov chains, Stochastic Process. Appl. 129 (11) (2019) 4209–4238, http://dx.doi.org/10.1016/j.spa.2018.11.014.

[21] M.R. Leadbetter, Extremes and local dependence in stationary sequences, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 65 (2) (1983) 291–306, http://dx.doi.org/10.1007/BF00532484.

[22] M.R. Leadbetter, G. Lindgren, H. Rootzén, Extremes and Related Properties of Random Sequences and Processes, in: Springer Series in Statistics, Springer-Verlag, New York-Berlin, 1983, p. xii+336.

[23] P.J. Northrop, An efficient semiparametric maxima estimator of the extremal index, Extremes 18 (4) (2015) 585–603, http://dx.doi.org/10.1007/s10687-015-0221-5.

[24] R. Perfekt, Extremal behaviour of stationary Markov chains with applications, Ann. Appl. Probab. 4 (2) (1994) 529–548, URL: http://links.jstor.org/sici?sici=1050-5164(199405)4:2<529:EBOSMC>2.0.CO;2-X&origin=MSN.

[25] C.Y. Robert, Asymptotic distributions for the intervals estimators of the extremal index and the cluster-size probabilities, J. Statist. Plann. Inference 139 (9) (2009) 3288–3309, http://dx.doi.org/10.1016/j.jspi.2009.03.010.

[26] C.Y. Robert, Inference for the limiting cluster size distribution of extreme values, Ann. Statist. 37 (1) (2009) 271–310, http://dx.doi.org/10.1214/07-AOS551.

[27] C.Y. Robert, J. Segers, C.A.T. Ferro, A sliding blocks estimator for the extremal index, Electron. J. Stat. 3 (2009) 993–1020, http://dx.doi.org/10.1214/08-EJS345.

[28] R.L. Smith, I. Weissman, Estimating the extremal index, J. R. Stat. Soc. Ser. B Stat. Methodol. 56 (3) (1994) 515–528, URL: http://links.jstor.org/sici?sici=0035-9246(1994)56:3<515:ETEI>2.0.CO;2-5&origin=MSN.

[29] M. Süveges, Likelihood estimation of the extremal index, Extremes 10 (1–2) (2007) 41–55, http://dx.doi.org/10.1007/s10687-007-0034-2.

[30] A.W. van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes - Springer Series in Statistics, Springer, New York, 1996.

[31] N. Zou, S. Volgushev, A. Bücher, Multiple block sizes and overlapping blocks for multivariate time series extremes, Ann. Statist. 49 (1) (2021) 295–320, URL: https://doi.org/10.1214/20-AOS1957.

Supplement to the paper: Statistical analysis for stationary time series at extreme levels: new estimators for the limiting cluster size distribution

Axel Bücher*, Tobias Jennessen

Heinrich-Heine-Universität Düsseldorf, Mathematisches Institut, Universitätsstr. 1, 40225 Düsseldorf, Germany

---

Abstract

This supplement contains auxiliary results needed for the proofs in the main paper (Sections A-C) as well as additional simulation results (Section D).

A. Auxiliary lemmas - Disjoint blocks

Throughout, assume that Condition 3.1 is met. All convergences are for $n \to \infty$ if not stated otherwise.

Lemma A.1. *For any $j \in \mathbb{N}_{\geq 1}$,*

$$\int_0^\infty e_{n,j}^{\mathrm{db}}(\tau) \ \mathrm{d}(\hat{H}_n^{\mathrm{db}} - H)(\tau) = o_{\mathbb{P}}(1).$$

Proof of Lemma A.1. Throughout the proof, we omit the upper index db at all instances of $\hat{H}_n^{\mathrm{db}}, e_{n,j}^{\mathrm{db}}$ and Z_{ni}^{db}. For any $\delta > 0$ and $\ell \in \mathbb{N}_{\geq 1}$, we have

$$\mathbb{P}\Big(\Big|\int_0^\infty e_{n,j}(\tau) \ \mathrm{d}(\hat{H}_n - H)(\tau)\Big| > 3\delta\Big) \leq \mathbb{P}(|A_{n,\ell}| > \delta) + \mathbb{P}(|B_{n,\ell,1}| > \delta) + \mathbb{P}(|B_{n,\ell,2}| > \delta),$$

where

$$A_{n,\ell} = \int_0^\ell e_{n,j}(\tau) \ \mathrm{d}(\hat{H}_n - H)(\tau) \tag{A.1}$$

*Corresponding author

Email addresses: axel.buecher@hhu.de (Axel Bücher), tobias.jennessen@hhu.de (Tobias Jennessen)

March 25, 2022

and

$$B_{n,\ell,1} = \int_\ell^\infty e_{n,j}(\tau)\, \mathrm{d}\hat{H}_n(\tau), \quad B_{n,\ell,2} = \int_\ell^\infty e_{n,j}(\tau)\, \mathrm{d}H(\tau). \tag{A.2}$$

The proof is finished once we have shown that

$$\forall\, \ell \in \mathbb{N}_{\geq 1}: \quad A_{n,\ell} = o_{\mathbb{P}}(1), \tag{A.3}$$

and that, for $v \in \{1, 2\}$,

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|B_{n,\ell,v}| > \delta\big) = 0. \tag{A.4}$$

We start by showing (A.3). Fix $\ell \in \mathbb{N}_{\geq 1}$. Let us first show that

$$\sup_{\tau\in[0,\ell]} \big|\hat{H}_n(\tau) - H(\tau)\big| = o_{\mathbb{P}}(1). \tag{A.5}$$

This result follows from the pointwise convergence (in probability) of $\hat{H}_n$ to H by a standard Glivenko-Cantelli-type argument. For the pointwise convergence, note that $\mathrm{E}[\hat{H}_n(\tau)] = \mathbb{P}(Z_{n1} \leq \tau)$, which converges to $H(\tau)$ by (2.1), such that it suffices to show $\lim_{n\to\infty} \mathrm{Var}(\hat{H}_n(\tau)) = 0$. In the following we prove

$$\lim_{n\to\infty} k_n \,\mathrm{Var}(\hat{H}_n(\tau)) = e^{-\theta\tau}(1 - e^{-\theta\tau})$$

for any $\tau \geq 0$. Set $\tilde{M}_{ni} = \max\{U_s : s \in I_i\}$. For any $\tau \geq 0$, we have

$$k_n \,\mathrm{Var}\left(\hat{H}_n(\tau)\right) = \mathbb{P}(\tilde{M}_{n1} > 1 - \tau/b_n)\big(1 - \mathbb{P}(\tilde{M}_{n1} > 1 - \tau/b_n)\big) + R_n,$$

where

$$R_n = \frac{2}{k_n} \sum_{1\leq i<j\leq k_n} \mathrm{Cov}\left(\mathbb{1}(\tilde{M}_{ni} > 1 - \tau/b_n), \mathbb{1}(\tilde{M}_{nj} > 1 - \tau/b_n)\right).$$

By definition of the extremal index, the first term converges to $e^{-\theta\tau}(1 - e^{-\theta\tau})$, and it remains to show that $R_n = o(1)$. By stationarity

$$\begin{aligned} R_n = {} & \frac{2(k_n - 1)}{k_n} \mathrm{Cov}\left(\mathbb{1}(\tilde{M}_{n1} > 1 - \tau/b_n), \mathbb{1}(\tilde{M}_{n2} > 1 - \tau/b_n)\right) \\ & + \frac{2}{k_n} \sum_{s=3}^{k_n-1} (k_n - s)\, \mathrm{Cov}\left(\mathbb{1}(\tilde{M}_{n1} > 1 - \tau/b_n), \mathbb{1}(\tilde{M}_{ns} > 1 - \tau/b_n)\right), \end{aligned}$$

which in absolute value is bounded by

$$\begin{aligned} & 2\, |\,\mathrm{Cov}\left(\mathbb{1}(\tilde{M}_{n1} > 1 - \tau/b_n), \mathbb{1}(\tilde{M}_{n2} > 1 - \tau/b_n)\right)| \\ & + 2 \sum_{s=3}^{k_n-1} |\,\mathrm{Cov}\left(\mathbb{1}(\tilde{M}_{n1} > 1 - \tau/b_n), \mathbb{1}(\tilde{M}_{ns} > 1 - \tau/b_n)\right)|. \end{aligned} \tag{A.6}$$

For the first term note that

$$
\begin{aligned}
&\operatorname{Cov}\big(\mathbb{1}(\tilde{M}_{n1} > 1-\tau/b_n), \mathbb{1}(\tilde{M}_{n2} > 1-\tau/b_n)\big) \\
&= \operatorname{Cov}\big(\mathbb{1}(\tilde{M}_{n1} < 1-\tau/b_n), \mathbb{1}(\tilde{M}_{n2} < 1-\tau/b_n)\big) \\
&= \mathbb{P}(\max\{U_s : s \in I_1 \cup I_2\} < 1-\tau/b_n) - \mathbb{P}(\tilde{M}_{n1} < 1-\tau/b_n)^2 \\
&= e^{-2\theta\tau} - e^{-2\theta\tau} + o(1) = o(1)
\end{aligned}
$$

by definition of the extremal index. Further, by Lemma 3.9 in Dehling & Philipp (2002) and Condition 3.1(ii) the second term in (A.6) can be bounded by

$$
8\sum_{s=3}^{k_n-1} \alpha_{\varepsilon_1}((s-1)b_n) \le 8C\sum_{s=3}^{k_n-1}((s-1)b_n)^{-\eta} \lesssim 8Cb_n^{-\eta},
$$

which converges to 0 by Condition 3.1(ii). Altogether, this proves equation (A.5). Further, by Lemma A.3 we know that

$$
\{e_{n,j}(\tau)\}_{\tau\in[0,\ell]} \xrightarrow{d} \{e_j(\tau)\}_{\tau\in[0,\ell]}
$$

in $D([0,\ell])$, for some centered Gaussian process e_j. Then, combining this result with the convergence in (A.5), we readily obtain (A.3) by Lemma C.8 in Berghaus & Bücher (2017).

Next, consider (A.4) with $v = 1$. We have

$$
\begin{aligned}
B_{n,\ell,1} &= k_n^{-3/2}\sum_{i,i'=1}^{k_n}\Big\{\mathbb{1}\big(N_{b_n,i'}^{(Z_{ni})} = j\big) - \varphi_{n,j}(Z_{ni})\Big\}\mathbb{1}(Z_{ni} \ge \ell) \\
&= T_{n,\ell} + S_{n,\ell,1} + S_{n,\ell,2},
\end{aligned}
$$

where

$$
\begin{aligned}
T_{n,\ell} &= k_n^{-3/2}\sum_{i=1}^{k_n}\sum_{i'\in\{i-1,i,i+1\}}\Big\{\mathbb{1}\big(N_{b_n,i'}^{(Z_{ni})} = j\big) - \varphi_{n,j}(Z_{ni})\Big\}\mathbb{1}(Z_{ni} \ge \ell), \\
S_{n,\ell,1} &= k_n^{-3/2}\sum_{i=3}^{k_n}\sum_{i'=1}^{i-2}\Big\{\mathbb{1}\big(N_{b_n,i'}^{(Z_{ni})} = j\big) - \varphi_{n,j}(Z_{ni})\Big\}\mathbb{1}(Z_{ni} \ge \ell), \\
S_{n,\ell,2} &= k_n^{-3/2}\sum_{i=1}^{k_n-2}\sum_{i'=i+2}^{k_n}\Big\{\mathbb{1}\big(N_{b_n,i'}^{(Z_{ni})} = j\big) - \varphi_{n,j}(Z_{ni})\Big\}\mathbb{1}(Z_{ni} \ge \ell).
\end{aligned}
$$

Clearly, $|T_{n,\ell}| \le 3k_n^{-1/2} = o(1)$. Next, write $\varepsilon = \varepsilon_1 \in (0,1)$ and $c > 1-\varepsilon$ from Condition 3.1(iii) as $c = 1-\kappa\varepsilon$ for some $\kappa \in (0,1)$, and let

$$
C_n = C_n(\varepsilon) = \big\{\max_{i=1,\dots,k_n} Z_{ni} < \kappa\varepsilon b_n\big\} = \big\{\min_{i=1,\dots,k_n} N_{ni} > 1-\kappa\varepsilon\big\},
$$

where $N_{ni} = \max\{U_s : s \in I_i^{\mathrm{db}}\}$. We obtain $\mathbb{P}(C_n) \to 1$ as $n \to \infty$ by Remark 3.2. As a consequence, (A.4) with $v = 1$ follows once we have shown that

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|S_{n,\ell,w} \mathbb{1}_{C_n}| > \delta\big) = 0, \quad w \in \{1,2\}. \tag{A.7}$$

We only prove this for the term $S_{n,\ell,1}$, as $S_{n,\ell,2}$ can be treated analogously. Define $N_{b_n,j,\varepsilon}^{(\tau)}$ as $N_{b_n,j}^{(\tau)}$ and Z_{ni}^{ε} as Z_{ni}, but with U_s substituted by $U_s^{\varepsilon} = U_s \mathbb{1}(U_s > 1-\varepsilon)$, respectively. Then, $Z_{ni} < \varepsilon\kappa b_n$ iff $Z_{ni}^{\varepsilon\kappa} < \varepsilon\kappa b_n$, and in that case we have

(1) $Z_{ni} = Z_{ni}^{\varepsilon\kappa}$,

(2) $U_s > 1 - Z_{ni}^{\varepsilon\kappa}/b_n$ iff $U_s^{\varepsilon} > 1 - Z_{ni}^{\varepsilon\kappa}/b_n$.

As a consequence, $S_{n,\ell,1}\mathbb{1}_{C_n} = S_{n,\ell,1}^{\varepsilon}\mathbb{1}_{C_n}$, where

$$S_{n,\ell,1}^{\varepsilon} = \frac{1}{k_n} \sum_{i=3}^{k_n} f_{n,i-2}(Z_{ni}^{\varepsilon\kappa})\, \mathbb{1}\big(\varepsilon\kappa b_n > Z_{ni}^{\varepsilon\kappa} \geq \ell\big)$$

and where

$$f_{n,i-2}(\tau) = k_n^{-1/2} \sum_{i'=1}^{i-2} \Big\{ \mathbb{1}\big(N_{b_n,i',\varepsilon}^{(\tau)} = j\big) - \varphi_{n,j}(\tau) \Big\}. \tag{A.8}$$

We may further write $f_{n,i-2}(\tau) = h_{n,i-2,j}(\tau) - h_{n,i-2,j-1}(\tau)$, where

$$h_{n,i-2,p}(\tau) = k_n^{-1/2} \sum_{i'=1}^{i-2} \mathbb{1}\big(N_{b_n,i',\varepsilon}^{(\tau)} \leq p\big) - \mathbb{P}\big(N_{b_n,i'}^{(\tau)} \leq p\big), \quad p \in \mathbb{N}_{\geq 0}. \tag{A.9}$$

Next, we apply Bradley's coupling lemma (see Lemma C.1 in the appendix) with $X = (U_s^{\varepsilon})_{s \in I_1 \cup \dots \cup I_{i-2}}, Y = Z_{ni}^{\varepsilon\kappa}$ and $q = q_n = ||Z_{n1}^{\varepsilon\kappa}||_{\gamma}/(\sqrt{k_n} b_n)$ for some $\gamma > 0$. We obtain the existence of a random variable $Y^* = Z_{ni}^{*\varepsilon\kappa}$, which is independent of $(U_s^{\varepsilon})_{s \in I_1 \cup \dots \cup I_{i-2}}$, has the same distribution as $Z_{ni}^{\varepsilon\kappa}$ and satisfies

$$\mathbb{P}(|Z_{ni}^{\varepsilon\kappa} - Z_{ni}^{*\varepsilon\kappa}| > q) \leq 18\, (\sqrt{k_n} b_n)^{\frac{\gamma}{2\gamma+1}} \alpha_{\varepsilon}(b_n)^{\frac{2\gamma}{2\gamma+1}}.$$

Thus, we obtain the bound

$$\begin{aligned} \mathrm{E}\left[|S_{n,\ell,1}^{\varepsilon}|\right] \leq \frac{1}{k_n} \sum_{i=3}^{k_n} \sum_{p \in \{j-1,j\}} \mathrm{E}\Big[&|h_{n,i-2,p}(Z_{ni}^{\varepsilon\kappa})| \mathbb{1}\big(\varepsilon b_n \kappa > Z_{ni}^{\varepsilon\kappa} \geq \ell\big) \\ &\times \mathbb{1}\big(|Z_{ni}^{\varepsilon\kappa} - Z_{ni}^{*\varepsilon\kappa}| < q\big)\Big] \\ &+ 36\, \frac{1}{k_n} \sum_{i=3}^{k_n} k_n^{-1/2} i (\sqrt{k_n} b_n)^{\frac{\gamma}{2\gamma+1}} \alpha_{\varepsilon}(b_n)^{\frac{2\gamma}{2\gamma+1}}, \end{aligned} \tag{A.10}$$

where the second sum is of the order

$$O\Big(k_n^{\frac{1}{2}+\frac{\gamma}{4\gamma+2}} b_n^{\frac{\gamma(1-2\eta)}{2\gamma+1}}\Big) = O\Big(\big(k_n b_n^{-\frac{2\gamma(2\eta-1)}{3\gamma+1}}\big)^{\frac{3\gamma+1}{4\gamma+2}}\Big) = o(1)$$

by Condition 3.1(ii), choosing $\gamma = \eta/(\eta-2) > 0$. To bound the first sum, note that for all $x, y \geq 0$ with $y - a \leq x \leq y + a$ for some $a > 0$, we have, for any $p \in \mathbb{N}_{\geq 0}$,

$$|h_{n,i,p}(x)| \leq \max\big\{|h_{n,i,p}(y+a)|, |h_{n,i,p}((y-a)_+)|\big\} + 2a\sqrt{k_n} \tag{A.11}$$

where $z_+ = \max(z, 0)$, which follows from monotonicity arguments. Indeed, $\tau \leq \tau'$ implies $N_{b_n,1}^{(\tau)} \leq N_{b_n,1}^{(\tau')}$, whence, for $y + a \geq x \geq y - a \geq 0$,

$$\begin{aligned}
0 < h_{n,i,p}(x) &\leq h_{n,i,p}(y-a) + \sqrt{k_n}\mathbb{P}(N_{b_n,1}^{(y-a)} \leq p < N_{b_n,1}^{(y+a)}) \\
&\leq h_{n,i,p}(y-a) + \sqrt{k_n}\mathbb{P}(N_{b_n,1}^{(y+a)} - N_{b_n,1}^{(y-a)} \geq 1) \\
&\leq h_{n,i,p}(y-a) + \sqrt{k_n}\,\mathrm{E}[N_{b_n,1}^{(y+a)} - N_{b_n,1}^{(y-a)}] \\
&= h_{n,i,p}(y-a) + 2a\sqrt{k_n},
\end{aligned}$$

where we have used the facts that $N_{b_n,1}^{(\tau)}$ is integer-valued. A similar inequality to the bottom implies (A.11). As a consequence of (A.11), we may bound the first sum on the right-hand side of (A.10) by

$$\begin{aligned}
\frac{1}{k_n}\sum_{i=3}^{k_n}\sum_{p\in\{j-1,j\}} \mathrm{E}\Big[\Big\{&|h_{n,i-2,p}(Z_{ni}^{*\varepsilon\kappa} + q_n)| + |h_{n,i-2,p}((Z_{ni}^{*\varepsilon\kappa} - q_n)_+)| \\
&+ 2\|Z_{n1}^{\varepsilon\kappa}\|_\gamma/b_n\Big\}\mathbb{1}\big(\varepsilon b_n\kappa + q_n > Z_{ni}^{*\varepsilon\kappa} \geq \ell - q_n\big)\Big].
\end{aligned}$$

Now, since $Z_{n1}^{\varepsilon}/b_n \leq 1$ and $q_n \to 0$, we have

$$\limsup_{n\to\infty} \|Z_{n1}^{\varepsilon\kappa}\|_\gamma/b_n\ \mathbb{P}\big(\varepsilon b_n\kappa > Z_{ni}^{*\varepsilon\kappa} \geq \ell - q_n\big) \leq \limsup_{n\to\infty}\mathbb{P}\big(Z_{ni}^{\varepsilon\kappa} \geq \ell\kappa\big)$$

which converges to 0 as $\ell \to \infty$. Hence, for proving (A.7) with $w = 1$, it remains to treat, for $p \in \{j-1, j\}$,

$$\frac{1}{k_n}\sum_{i=3}^{k_n} \mathrm{E}\Big[\big\{|h_{n,i-2,p}((Z_{ni}^{*\varepsilon\kappa} \pm q_n)_+)|\mathbb{1}\big(\varepsilon b_n\kappa + q_n > Z_{ni}^{*\varepsilon\kappa} \geq \ell - q_n\big)\Big]. \tag{A.12}$$

We only consider the case with the plus sign. After conditioning on $Z_{ni}^{*\varepsilon\kappa}$ we need to bound $\mathrm{E}[|h_{n,i-2,p}(x)|]$ for $\ell \leq x \leq \varepsilon b_n$ (note that $Z_{ni}^{*\varepsilon\kappa} + q_n \leq \varepsilon b_n\kappa + 2q_n \leq \varepsilon b_n$ for large n, since q_n converges to zero). Write $h_{n,i-2,p} = h_{n,i-2,p}^{even} + h_{n,i-2,p}^{odd}$, where $h_{n,i-2,p}^{even}$ and $h_{n,i-2,p}^{odd}$ correspond to the sum over the even and odd blocks in (A.9), respectively. Fix $x \in [\ell, \varepsilon b_n]$. Set

$$V_j = \big\{\mathbb{1}\big(N_{b_n,2j,\varepsilon}^{(x)} \leq p\big) - \mathbb{P}\big(N_{b_n,2j}^{(x)} \leq p\big)\big\},$$

such that $h_{n,i-2,v}^{even}(x) = k_n^{-1/2}\sum_{j=1}^{\lfloor i/2\rfloor-1} V_j$. Note that V_j is centered. Recursive application of Bradley's coupling lemma (see Lemma C.1) with some $\gamma > 0, V_1^* = V_1$ and, in the j-th step, $X = (V_1^*, \ldots, V_j^*), Y = V_{j+1}$ and

$q' = q_n' = 1/\sqrt{k_n}$ (note that $\alpha(\sigma(V_j), \sigma(V_{j+1}) \le \alpha_\varepsilon(b_n))$ in combination with Theorem 5.1 in Bradley (2005) lets us construct an i.i.d. sequence $(V_j^*)_{j\ge1}$, such that V_j^* has the same distribution as V_j and

$$\mathbb{P}(|V_j - V_j^*| \ge q_n') \le 18\ k_n^{\frac{\gamma}{4\gamma+2}} \alpha_\varepsilon(b_n)^{\frac{2\gamma}{2\gamma+1}}.$$

Note that the i.i.d. sequence $(V_j^*)_{j\ge1}$ is centered with $|V_j^*| \le 1$. As a consequence, by Condition 3.1(ii),

$$\begin{aligned}
\mathrm{E}[|h^{even}_{n,i-2,v}(x)|] &\le k_n^{-1/2}\,\mathrm{E}\left[\left|\textstyle\sum_{j=1}^{\lfloor i/2\rfloor-1} V_j^*\right|\right] + ik_n^{-1/2}\,\mathrm{E}\left[|V_1 - V_1^*|\right] \\
&\le (i/k_n)^{1/2} + ik_n^{-1/2}\big\{q_n' + 36\, k_n^{\frac{\gamma}{4\gamma+2}} \alpha_\varepsilon(b_n)^{\frac{2\gamma}{2\gamma+1}}\big\} \\
&\le (i/k_n)^{1/2} + ik_n^{-1} + 36C^{\frac{2\gamma}{2\gamma+1}}\, k_n^{\frac12+\frac{\gamma}{4\gamma+2}} b_n^{-\eta\frac{2\gamma}{2\gamma+1}}. \qquad \text{(A.13)}
\end{aligned}$$

A similar bound can be obtained for the sum over the odd blocks. Assembling terms, the expression in (A.12) can be bounded by

$$\begin{aligned}
&\mathbb{P}\big(Z_{n1}^{*\varepsilon\kappa} \ge \ell - q_n\big)\frac{1}{k_n}\sum_{i=3}^{k_n}\Big[(i/k_n)^{1/2} + ik_n^{-1} + 36C^{\frac{2\gamma}{2\gamma+1}}\, k_n^{\frac12+\frac{\gamma}{4\gamma+2}} b_n^{-\eta\frac{2\gamma}{2\gamma+1}}\Big] \\
&\lesssim \mathbb{P}\big(Z_{n1} \ge \ell/2\big)\Big\{1 + k_n^{\frac12+\frac{\gamma}{4\gamma+2}} b_n^{-\eta\frac{2\gamma}{2\gamma+1}}\Big\},
\end{aligned}$$

where

$$k_n^{\frac12+\frac{\gamma}{4\gamma+2}} b_n^{-\eta\frac{2\gamma}{2\gamma+1}} = \big(k_n b_n^{-\frac{4\eta\gamma}{3\gamma+1}}\big)^{\frac{3\gamma+1}{4\gamma+2}} = o(1)$$

by Condition 3.1(ii), after setting $\gamma = 1$. Hence, since $\lim_{n\to\infty}\mathbb{P}(Z_{n1} \ge \ell/2) = e^{-\theta\ell/2} \to 0$ for $\ell \to \infty$, we obtain (A.7) and hence (A.4) with $v = 1$. Next, consider (A.4) with $v = 2$. By Markov's inequality

$$\mathbb{P}(|B_{n,\ell,2}| > \delta) \le \delta^{-1}\int_\ell^\infty \mathrm{E}[|e_{n,j}(\tau)|]\,\mathrm{d}H(\tau).$$

Split the integral on the right-hand side into two integrals over $[\ell, \varepsilon b_n]$ and $(\varepsilon b_n, \infty)$. For $\tau \in [\ell, \varepsilon b_n]$, we have $e_{n,j}(\tau) = f_{n,k_n}(\tau)$, with f_{n,k_n} from (A.8). Hence, similar as for the treatment of (A.12), see in particular relation (A.13), we have $\mathrm{E}[|f_{n,k_n}(\tau)|] \lesssim 1 + o(1)$, where the upper bound is uniform in τ. As a consequence, the integral on the right-hand side of the previous display can be bounded by

$$(1 + o(1))\int_\ell^{\varepsilon b_n} \mathrm{d}H(\tau) + \sqrt{k_n}\int_{\varepsilon b_n}^\infty \mathrm{d}H(\tau),$$

which converges to zero for $n \to \infty$ followed by $\ell \to \infty$. This proves (A.4) with $v = 2$. □

Lemma A.2. *For any $m \in \mathbb{N}_{\ge1}$,*

$$\frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n}\big(W_{n,i}^{\mathrm{db}}(1), \dots, W_{n,i}^{\mathrm{db}}(m)\big) \xrightarrow{d} (s_1^{\mathrm{db}}, \dots, s_m^{\mathrm{db}}) \sim \mathcal{N}_m(0, \Sigma_m^{\mathrm{db}}),$$

where $W_{n,i}^{\mathrm{db}}(j)$ *and* $\Sigma_m^{\mathrm{db}} = (d_{j,j'}^{\mathrm{db}})_{1 \le j,j' \le m}$ *are defined in (6.3) and (4.2), respectively.*

Proof of Lemma A.2. Throughout the proof, we omit the upper index db at all instances of $\hat H_n^{\mathrm{db}}, e_{n,j}^{\mathrm{db}}$ and Z_{ni}^{db}. Define

$$B_{n,j} = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \Big\{ \varphi_{n,j}(Z_{ni}) - \mathrm{E}[\varphi_{n,j}(Z_{ni})] \Big\}, \quad j \in \mathbb{N}_{\ge 1}.$$

Decompose each block $I_i = I_i^{\mathrm{db}} = I_i^+ \cup I_i^-$, $i = 1, \dots, k_n$, into a big block $I_i^+ = \{(i-1)b_n + 1, \dots, ib_n - \ell_n\}$ and a small one $I_i^- = \{ib_n - \ell_n + 1, \dots, ib_n\}$, where ℓ_n is from Condition 3.1(ii), and define $Z_{ni}^+ = b_n(1 - N_{ni}^+)$ with $N_{ni}^+ = \max\{U_s : s \in I_i^+\}$. Set

$$B_{n,j}^+ = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \varphi_{n,j}(Z_{ni}^+) - \mathrm{E}[\varphi_{n,j}(Z_{ni}^+)], \quad j \in \mathbb{N}_{\ge 1},$$

and write

$$B_{n,j}^- = B_{n,j} - B_{n,j}^+ = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} Y_{ni}^- - \mathrm{E}[Y_{ni}^-],$$

where $Y_{ni}^- = \varphi_{n,j}(Z_{ni}) - \varphi_{n,j}(Z_{ni}^+)$.

Let us start by showing that

$$B_{n,j}^- = o_{\mathbb{P}}(1), \tag{A.14}$$

for which we may proceed similar as in the proof of Lemma 9.3 in Berghaus & Bücher (2018): denote $G_n^- = G_n - G_n^+$, $Z_{ni}^- = Z_{ni} - Z_{ni}^+$.

For $\varepsilon = 1 - c$ with c from Condition 3.1(iii), let $A_n^+ = \{\min_{i=1}^{k_n} N_{ni}^+ > 1 - \varepsilon\}$ and note that $\mathbb{P}(A_n^+) \to 1$ by Remark 3.2. We can write $B_{n,j}^- = B_{n,j}^- \mathbb{1}_{A_n^+} + o_{\mathbb{P}}(1) = \tilde B_{n,j}^- \mathbb{1}_{A_n^+} = \tilde B_{n,j}^- + o_{\mathbb{P}}(1)$, where

$$\tilde B_{n,j}^- = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} \{Y_{ni}^- - \mathrm{E}[Y_{ni}^-]\} \mathbb{1}(N_{ni}^+ > 1 - \varepsilon)$$

It suffices to show that $\tilde B_{n,j}^- = o_{\mathbb{P}}(1)$. For that purpose, note that $|Y_{ni}^-| \le 1$, such that, by stationarity and Minkowski's inequality,

$$\mathrm{E}\left[|B_{n,j}^- - \tilde B_{n,j}^-|^2 \right] \le k_n \mathbb{P}(N_{n1}^+ \le 1 - \varepsilon) \le k_n \mathbb{P}(N_{n1}' \le 1 - \varepsilon),$$

which converges to zero by Condition 3.1(iii). As a consequence, $\mathrm{E}[\tilde B_{n,j}^-] = \mathrm{E}[\tilde B_{n,j}^- - B_{n,j}^-] + \mathrm{E}[B_{n,j}^-] = o(1)$.

Next, we show that $\mathrm{Var}(\tilde B_{n,j}^-) = o(1)$. Now $N_{ni}^+ > 1 - \varepsilon$ implies that $Z_{ni}^+ = Z_{ni}^{\varepsilon +}$ and $Z_{ni} = Z_{ni}^{\varepsilon}$, where the variables with an upper index ε are defined in terms of the U_i^ε instead of the U_i. Hence, $\tilde B_{n,j}^- = k_n^{-1/2} \sum_{i=1}^{k_n} S_{ni}^\varepsilon$, where

$$S_{ni}^\varepsilon = \{Y_{ni}^{\varepsilon -} - \mathrm{E}[Y_{ni}^-]\} \mathbb{1}(N_{ni}^{\varepsilon +} > 1 - \varepsilon)$$

with

$$Y_{ni}^{\varepsilon-} = \varphi_{n,j}(Z_{ni}^{\varepsilon}) - \varphi_{n,j}(Z_{ni}^{\varepsilon+}).$$

As a consequence, by stationarity

$$\begin{aligned}\operatorname{Var}(\tilde{B}_{n,j}^{-}) &= \operatorname{Var}(S_{n1}^{\varepsilon}) + \frac{2}{k_n}\sum_{i=1}^{k_n}(k_n - i)\operatorname{Cov}(S_{n1}^{\varepsilon}, S_{n,1+i}^{\varepsilon}) \\ &\leq 3\operatorname{Var}(S_{n1}^{\varepsilon}) + \frac{2}{k_n}\sum_{i=2}^{k_n}(k_n - i)\operatorname{Cov}(S_{n1}^{\varepsilon}, S_{n,1+i}^{\varepsilon}) \qquad \text{(A.15)}\end{aligned}$$

Let us first show that $\operatorname{Var}(S_{n1}^{\varepsilon}) = o(1)$ as $n \to \infty$, which would follow, in view of the boundedness of $|\varphi_{n,j}| \leq 1$, from $|Y_{n1}^{\varepsilon-}| = o_{\mathbb{P}}(1)$ and $|Y_{n1}^{-}| = o_{\mathbb{P}}(1)$. To this end, let $Z_{n1}^{\varepsilon-} = Z_{n1}^{\varepsilon} - Z_{n1}^{\varepsilon+}$ and note that $|Z_{n1}^{\varepsilon-}| \leq |Z_{n1}|$ (by studying the cases $N_{ni}^{\varepsilon+} > 1-\varepsilon$ and $N_{ni}^{\varepsilon+} \leq 1-\varepsilon$). Therefore, since $\ell_n = o(b_n)$,

$$\begin{aligned}\mathbb{P}(Y_{n1}^{\varepsilon-} \neq 0) &\leq \mathbb{P}(Z_{n1}^{\varepsilon} \neq Z_{n1}^{\varepsilon+}) \\ &= \mathbb{P}(Z_{n1}^{\varepsilon-} \neq 0), \\ &\leq \mathbb{P}(Z_{n1}^{-} \neq 0), \\ &= \mathbb{P}\Big(\max_{s\in I_1} U_s > \max_{s\in I_1^+} U_s\Big) \\ &\leq \mathbb{P}\Big(\max_{s=1}^{b_n-\ell_n} U_s \leq 1 - y/b_n\Big) + \mathbb{P}\Big(\max_{s=1}^{\ell_n} U_s > 1 - y/b_n\Big) \\ &\leq \mathbb{P}\Big(Z_{1:b_n-\ell_n} \geq y(b_n-\ell_n)/b_n\Big) + \ell_n y/b_n \to \exp(-\theta y),\end{aligned}$$

which can be made arbitrary small by increasing y. This implies $|Y_{n1}^{\varepsilon-}| = o_{\mathbb{P}}(1)$, and the same arguments can be used for showing that $|Y_{n1}^{-}| = o_{\mathbb{P}}(1)$.

It remains to treat the sum over the covariances on the right-hand side of (A.15). Note that S_{ni}^{ε} is $\mathcal{B}_{\{(i-1)b_n+1\}:(ib_n)}^{\varepsilon}$-measurable (defined on page 8). By Lemma 3.9 in Dehling & Philipp (2002),

$$|\operatorname{Cov}(S_{n1}^{\varepsilon}, S_{n,1+i}^{\varepsilon})| \leq 4\alpha_{\varepsilon_1}((i-1)b_n)$$

Now, for $i \geq 2$, $\alpha_{\varepsilon_1}((i-1)b_n) \leq Cb_n^{-\eta}(i-1)^{-\eta}$ by Condition 3.1(ii). The sum over the covariances in (A.15) can thus be bounded by a multiple of $Cb_n^{-\eta}\sum_{i=2}^{k_n}(i-1)^{-\eta} \leq Cb_n^{-\eta}\sum_{i=2}^{\infty} i^{-\eta} = o(1)$. Overall, we obtain $\tilde{B}_{n,j}^{-} = o_{\mathbb{P}}(1)$ as required.

Next, let us show that

$$\int_0^\infty e_{n,j}(\tau)\,\mathrm{d}H(\tau) = \int_0^\infty e_{n,j}^{+}(\tau)\,\mathrm{d}H(\tau) + o_{\mathbb{P}}(1), \qquad \text{(A.16)}$$

where

$$e_{n,j}^{+}(\tau) = \sqrt{k_n}\{p_n^{(\tau),+}(j) - \mathbb{P}(N_{b_n,1}^{(\tau),+} = j)\}, \quad j \in \mathbb{N}_{\geq 1},$$

and where

$$p_n^{(\tau),+}(j) = \frac{1}{k_n}\sum_{i=1}^{k_n}\mathbb{1}(N_{b_n,i}^{(\tau),+} = j), \quad N_{b_n,i}^{(\tau),+} = \sum\nolimits_{s\in I_i^+}\mathbb{1}(U_s > 1 - \tau/b_n).$$

For that purpose, write, for $\tau > 0$,

$$e_{n,j}(\tau) - e^+_{n,j}(\tau) = \frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} X_{n,i}(\tau) - Y_{n,i}(\tau) - \mathrm{E}[X_{n,i}(\tau) - Y_{n,i}(\tau)],$$

where

$$X_{n,i}(\tau) = \mathbb{1}\Big(N^{(\tau)}_{b_n,i} = j, N^{(\tau),-}_{b_n,i} > 0\Big), \quad Y_{n,i}(\tau) = \mathbb{1}\Big(N^{(\tau),+}_{b_n,i} = j, N^{(\tau),-}_{b_n,i} > 0\Big)$$

with $N^{(\tau),-}_{b_n,i}$ defined as $N^{(\tau),+}_{b_n,i}$ but with the sum ranging over I_i^- instead of I_i^+. We obtain

$$\mathrm{E}[|e_{n,j}(\tau) - e^+_{n,j}(\tau)|^2] \le \frac{2}{k_n} \operatorname{Var}\Big(\sum_{i=1}^{k_n} X_{n,i}(\tau)\Big) + \frac{2}{k_n} \operatorname{Var}\Big(\sum_{i=1}^{k_n} Y_{n,i}(\tau)\Big).$$

By stationarity and Lemma 3.11 in Dehling & Philipp (2002) (with $t = 2, s = r = 4$) we have

$$\begin{aligned}
&\frac{1}{k_n} \operatorname{Var}\Big(\sum_{i=1}^{k_n} X_{n,i}(\tau)\Big) \\
=&\frac{1}{k_n} \sum_{i=1}^{k_n} \operatorname{Var}(X_{n,i}(\tau)) + \frac{2}{k_n} \sum_{i=1}^{k_n} (k_n - i) \operatorname{Cov}(X_{n,1}(\tau), X_{n,i+1}(\tau)) \\
\le& \; 20 \; \mathbb{P}(N^{(\tau),-}_{b_n,1} > 0)^{1/2} \Big(3 + \sum_{i=2}^{k_n} \alpha_{\varepsilon_1}((i-1)b_n)^{1/2}\Big) \\
\le& \; 20 \Big(\frac{\tau \ell_n}{b_n}\Big)^{1/2} \Big(3 + C^{1/2} \sum_{i=1}^{\infty} i^{-\eta/2}\Big),
\end{aligned}$$

where we used Condition 3.1(ii) in the last step. Since the series in the last display is finite and the variance over the $Y_{n,i}(\tau)$ can be treated analogously, we obtain, for any $\tau > 0$, $\mathrm{E}[|e_{n,j}(\tau) - e^+_{n,j}(\tau)|^2] \lesssim (\tau \ell_n / b_n)^{1/2}$ such that

$$\mathrm{E}\Big[\Big| \int_0^\infty e_{n,j}(\tau) - e^+_{n,j}(\tau) \,\mathrm{d}H(\tau) \Big|^2\Big] \lesssim (\ell_n/b_n)^{1/2} \int_0^\infty \tau^{1/2} \,\mathrm{d}H(\tau),$$

which converges to 0 by Condition 3.1(ii). This readily implies (A.16).

As a consequence of (A.14) and (A.16), we have

$$k_n^{-1/2} \sum_{i=1}^{k_n} W_{n,i}(j) = \int_0^\infty e^+_{n,j}(\tau) \,\mathrm{d}H(\tau) + B^+_{n,j} + o_{\mathbb{P}}(1). \tag{A.17}$$

Next, define $A_n^+ = \big\{ \min_{i=1,\dots,k_n} N^+_{ni} > 1 - \varepsilon \big\}$ with $\varepsilon = \varepsilon_1$ from Condition 3.1(ii), such that $\lim_{n\to\infty} \mathbb{P}(A_n^+) = 1$ by Remark 3.2. Hence, by (A.17) and the Cramér-Wold-device, the lemma is shown once we prove that

$$\sum_{j=1}^m \lambda_j \Big\{ \int_0^\infty e^+_{n,j}(\tau) \,\mathrm{d}H(\tau) + B^+_{n,j} \Big\} \mathbb{1}_{A_n^+} \xrightarrow{d} \sum_{j=0}^m \lambda_j s_j \tag{A.18}$$

for arbitrary $\lambda_j \in \mathbb{R}$. For that purpose, rewrite the left-hand side of (A.18) as $k_n^{-1/2} \sum_{i=1}^{k_n} f_{i,n} \mathbb{1}_{A_n^+}$, where

$$f_{i,n} = \sum_{j=0}^{m} \lambda_j \bigg\{ \int_0^\infty \mathbb{1}(N_{b_n,i}^{(\tau),+} = j) - \mathbb{P}(N_{b_n,1}^{(\tau),+} = j) \, \mathrm{d}H(\tau) \\ + \varphi_{n,j}(Z_{ni}^+) - \mathrm{E}[\varphi_{n,j}(Z_{ni}^+)] \bigg\}.$$

By the definition of A_n^+, we have

$$k_n^{-1/2} \sum_{i=1}^{k_n} f_{i,n} \mathbb{1}_{A_n^+} = k_n^{-1/2} \sum_{i=1}^{k_n} \tilde f_{i,n} + o_{\mathbb{P}}(1),$$

where $\tilde f_{i,n} = f_{i,n} \mathbb{1}(Z_{ni}^+ < \varepsilon b_n)$. Observing that $\tilde f_{i,n}$ is $\mathcal{B}^{\varepsilon}_{\{(i-1)b_n+1\}:\{ib_n-\ell_n\}}$-measurable (and that $\tilde f_{i,n}$ and $\tilde f_{j,n}$ are at least ℓ_n observations apart for $i \neq j$) and recursively applying Lemma 3.11 in Dehling & Philipp (2002), we obtain that, for any $t \in \mathbb{R}$, the characteristic functions satisfy

$$\left| \mathrm{E}\left[\exp\left(\frac{\mathbf{i}t}{\sqrt{k_n}} \sum_{i=1}^{k_n} \tilde f_{i,n} \right) \right] - \prod_{i=1}^{k_n} \mathrm{E}\left[\exp\left(\frac{\mathbf{i}t}{\sqrt{k_n}} \tilde f_{i,n} \right) \right] \right| \lesssim k_n \alpha_{\varepsilon_1}(\ell_n).$$

The upper bound converges to 0 by Condition 3.1(ii). Therefore, $\{\tilde f_{i,n} : i = 1, \dots, k_n\}$ may be considered independent in the remaining part of this proof (see also the argumentation in the proof of Lemma 7.3). To obtain asymptotic normality, we apply Lyapunov's central limit theorem. First, note that $|\tilde f_{1,n}| \le 2\sum_{j=1}^m |\lambda_j| < \infty$. This implies, by stationarity, for any $p > 2$,

$$\frac{\sum_{i=1}^{k_n} \mathrm{E}[|\tilde f_{i,n}|^p]}{\big\{ \sum_{i=1}^{k_n} \mathrm{Var}(\tilde f_{i,n}) \big\}^{p/2}} = k_n^{1-p/2} \frac{\mathrm{E}[|\tilde f_{1,n}|^p]}{\mathrm{E}[|\tilde f_{1,n}|^2]^{p/2}} \lesssim k_n^{1-p/2} \, \mathrm{E}[\tilde f_{1,n}^2]^{-p/2},$$

which converges to zero provided that $\lim_{n\to\infty} \mathrm{E}[\tilde f_{1,n}^2]$ exists. The central limit theorem then implies that $k_n^{-1/2} \sum_{i=1}^{k_n} \tilde f_{i,n}$ converges in distribution to a centered normal distribution with variance $\lim_{n\to\infty} \mathrm{E}[\tilde f_{1,n}^2]$, whence it remains to calculate the latter limit.

For that purpose, note that $\lim_{n\to\infty} \mathrm{E}[\tilde f_{1,n}^2] = \lim_{n\to\infty} \mathrm{E}[f_{1,n}^2]$. Set

$$C_{n,j} = \int_0^\infty \mathbb{1}(N_{b_n,1}^{(\tau),+} = j) - \mathbb{P}(N_{b_n,1}^{(\tau),+} = j) \, \mathrm{d}H(\tau),$$
$$D_{n,j} = \varphi_{n,j}(Z_{n1}^+) - \mathrm{E}[\varphi_{n,j}(Z_{n1}^+)],$$

and note that

$$\mathrm{E}[f_{1,n}^2] = \sum_{j,j'=1}^{m} \lambda_j \lambda_{j'} \, \mathrm{E}[(C_{n,j} + D_{n,j})(C_{n,j'} + D_{n,j'})],$$

which we need to show to converge to $\sum_{j,j'=1}^m \lambda_j \lambda_{j'} \operatorname{E}[s_j s_{j'}]$. Similar arguments as in the proof of (A.14) and (A.16) allow us to replace I_1^+ by I_1.

We start by considering the product of the $C_{n,j}$-terms. Invoking the dominated convergence theorem and

$$\mathbb{P}(N_{b_n,1}^{(\tau)} = j, N_{b_n,1}^{(\tau')} = j') \to \mathbb{P}(N_E^{(\tau)} = j, N_E^{(\tau')} = j')$$

with $(N_E^{(\tau)}, N_E^{(\tau')})$ as defined in Theorem 4.1 (following from Condition 3.1(i)), we obtain that

$$\lim_{n\to\infty} \operatorname{E}[C_{n,j} C_{n,j'}] = \int_0^\infty \int_0^\infty \operatorname{Cov}[\mathbb{1}(N_E^{(\tau)} = j), \mathbb{1}(N_E^{(\tau')} = j')] \, \mathrm{d}H(\tau)\mathrm{d}H(\tau').$$

Second, we consider the product of the $D_{n,j}$-terms. For this purpose, we first show that $\varphi_{n,j}(Z_{n1})$ converges weakly to $p^{(Z)}(j)$, for $Z \sim \operatorname{Exp}(\theta)$, which in turn is a consequence of weak convergence of Z_{n1} to Z and the extended continuous mapping theorem. For the latter, one needs to prove that $\varphi_{n,j}(x_n) \to p^{(x)}(j)$ for any $x_n \to x$, which follows from

$$\begin{aligned} |\varphi_{n,j}(x_n) - \varphi_{n,j}(x)| &\le \operatorname{E}\left[|\mathbb{1}(N_{b_n,1}^{(x_n)} = j) - \mathbb{1}(N_{b_n,1}^{(x)} = j)|\right] \\ &\le \operatorname{E}\left[\mathbb{1}(|N_{b_n,1}^{(x_n)} - N_{b_n,1}^{(x)}| \ge 1)\right] \le \operatorname{E}\left[|N_{b_n,1}^{(x_n)} - N_{b_n,1}^{(x)}|\right] \\ &= \operatorname{E}\left[N_{b_n,1}^{(x_n \vee x)} - N_{b_n,1}^{(x_n \wedge x)}\right] = |x_n - x|. \end{aligned}$$

Likewise, $\varphi_{n,j}(Z_{n1})\varphi_{j',n}(Z_{n1})$ weakly converges to $p^{(Z)}(j)p^{(Z)}(j')$. Since $|\varphi_{n,j}| \le 1$, Theorem 2.20 in van der Vaart (1998) implies convergence of the corresponding moments, i.e.,

$$\operatorname{E}[D_{n,j} D_{n,j'}] = \operatorname{Cov}\left(\varphi_{n,j}(Z_{n1}), \varphi_{n,j'}(Z_{n1})\right) = \operatorname{Cov}\left(p^{(Z)}(j), p^{(Z)}(j')\right) + o(1).$$

With regard to the mixed $C_{n,j}$- and $D_{n,j'}$-terms, note that, for $j \in \mathbb{N}_{\ge 0}$ and $\mu \ge 0$,

$$\begin{aligned} \mathbb{P}(N_{b_m,1}^{(\tau)} = j, Z_{n1} > \mu) &= \mathbb{P}(N_{b_n,1}^{(\tau)} = j, N_{b_n,1}^{(\mu)} = 0) \\ &\to \begin{cases} p_2^{(\tau,\mu)}(j, 0) & , \tau \ge \mu \ge 0 \\ e^{-\theta\mu}\mathbb{1}(j = 0) & , \mu > \tau \ge 0, \end{cases} \end{aligned}$$

such that $(N_{b_n,1}^{(\tau)}, Z_{1:n}) \xrightarrow{d} (N_E^{(\tau)}, Z)$ with $(N_E^{(\tau)}, Z)$ as specified in Theorem 4.1. The extended continuous mapping theorem and boundedness, $|\varphi_{n,j}| \le 1$, implies

$$\begin{aligned} \operatorname{E}[C_{n,j} D_{n,j'}] &= \int_0^\infty \operatorname{Cov}\left\{\mathbb{1}(N_{b_n,1}^{(\tau)} = j), \varphi_{n,j'}(Z_{n1})\right\} \mathrm{d}H(\tau) \\ &= \int_0^\infty \operatorname{Cov}\left\{\mathbb{1}(N_E^{(\tau)} = j), p^{(Z)}(j')\right\} \mathrm{d}H(\tau) + o(1). \end{aligned}$$

The last three paragraphs imply

$$\lim_{n\to\infty} \operatorname{E}[(C_{n,j} + D_{n,j})(C_{n,j'} + D_{n,j'})] = d_{j,j'}$$

with $d_{j,j'} = d_{j,j'}^{\mathrm{db}}$ from (4.2), which finalizes the proof. □

Lemma A.3. *For $j \in \mathbb{N}_{\geq 1}$, we have*

$$e_{n,j} \xrightarrow{d} e_j \quad in \quad D([0,\infty))$$

for some centered Gaussian process e_j, whose covariance function is given in Theorem 4.1 in Robert (2009).

Proof. The result follows by suitable adaptations of the proof of Theorem 4.1 in Robert (2009) (see also Theorem 7.1 for a similar result for the sliding blocks version $e^{\mathrm{sb}}_{n,j}$ under slightly different mixing conditions). In the following we explain why Lemma 6.7 in Robert (2009) regarding the finite-dimensional convergence of the above process is applicable under our set of conditions, and give a detailed proof for the tightness of that process that substitutes his Lemma 6.8.

Note that Robert's Lemma 6.7 yields the finite-dimensional convergence of the vector of functions $(e_{n,0},\dots,e_{n,m},\bar e_n)$ in $D([\sigma,\ell])^{m+2}$, with $\bar e_n$ denoting the tail empirical process, and with arbitrary fixed $0<\sigma<\ell$. Extending the result for the margin $e_{n,j}$ to $\sigma=0$ is straightforward. It remains to argue why the marginal convergence is valid under our weaker conditions; note that our assumptions are the same as in Robert (2009) except that we do not impose Conditions (C0.b) and (C2.a) and that we impose a slightly different mixing condition (see Condition 3.1(ii)) than in (C2.b) and (C2.c) in that reference. First, a close look at Robert's proof reveals that Conditions (C0.b) and (C2.a) are only needed for weak convergence of the last component $\bar e_n$ and not for weak convergence of the component $e_{n,j}$. The argumentation regarding (C2.b) and (C2.c) is more involved, and requires referring to specific pages and arguments in Robert's paper. First of all, the assumption $\ell_n = o(r_n^{2/r})$ is used on page 302 only, where it is used for showing that the small-blocks version of $\bar e_n$ is asymptotically negligible. The corresponding result for $e_{n,j}$, however, only requires $\ell_n = o(r_n)$, which is exactly $\ell_n = o(b_n)$ in our notation as imposed in Condition 3.1(ii). Next, his condition $\lim_{n\to\infty} nr_n^{-1}\alpha_{l_n}=0$ from (C2.c) (used on page 303 only) is actually stronger than needed, and can be replaced by the weaker condition $\lim_{n\to\infty} nr_n^{-1}\alpha_{r_n,l_n}(\tau_1,\dots,\tau_r)=0$. The latter however is a simple consequence of our condition $k_n\alpha_{\varepsilon_1}(\ell_n)=o(1)$ in Condition 3.1(ii). Note that $\eta>3$ from Condition 3.1(ii) also suffices for the convergence of the series appearing on page 302 in the upper bound of I_1.

It remains to show tightness of $e_{n,j}$ on $[0,\ell]$ for any $\ell\in\mathbb{N}$, for which we give a self-contained proof. Write $e_{n,j} = \tilde e_{n,j} - \tilde e_{n,j-1}$ (set $\tilde e_{n,-1} := 0$), where

$$\tilde e_{n,j}(\tau) = \sqrt{k_n}\Big\{\frac{1}{k_n}\sum_{i=1}^{k_n} \mathbb{1}\Big(N^{(\tau)}_{b_n,i}\leq j\Big) - \mathbb{P}\Big(N^{(\tau)}_{b_n,i}\leq j\Big)\Big\}.$$

It suffices to show tightness of $\tilde e_{n,j}$. By Theorem 15.5 and Theorem 8.3 in Billingsley (1968) and the finite-dimensional convergence of $e_{n,j}$ it is sufficient to show that: for any (sufficiently small) $\varepsilon>0$ and $\nu>0$ there exist some $\delta>0$ and $n_0\in\mathbb{N}$, such that

$$\mathbb{P}\Big(\sup_{\tau_2\leq\tau_1\leq\tau_2+\delta}|\tilde e_{n,j}(\tau_1)-\tilde e_{n,j}(\tau_2)|>\varepsilon\Big)\leq\delta\nu,$$

for all $\tau_2 \in [0, \ell - \delta]$ and all $n \geq n_0$. Fix $0 < \varepsilon < 1/4$ and $\nu > 0$. Choose $2 < v < p < r < \infty$ such that, with η from Condition 3.1(ii), $\eta > v/(v-2)$ and $\eta \geq (p-1)r/(r-p)$ and $p/2 > 1 + \varepsilon$ (decrease ε is necessary); see (A.20) below that such choices are possible. Let $0 \leq \tau_2 < \tau_1 \leq \ell$. By Theorem 4.1 in Shao & Yu (1996), there exists some constant $K < \infty$, such that

$$
\begin{aligned}
&\mathrm{E}\Big[\big|\sqrt{k_n}\big(\tilde{e}_{n,j}(\tau_1) - \tilde{e}_{n,j}(\tau_2)\big)\big|^p\Big] \\
&\leq K\big(k_n^{p/2}\mathbb{P}\big(N_{b_n,1}^{(\tau_2)} \leq j < N_{b_n,1}^{(\tau_1)}\big)^{p/v} + k_n^{1+\varepsilon}\mathbb{P}\big(N_{b_n,1}^{(\tau_2)} \leq j < N_{b_n,1}^{(\tau_1)}\big)^{p/r}\big) \\
&\leq K\big(k_n^{p/2}(\tau_1 - \tau_2)^{p/v} + k_n^{1+\varepsilon}(\tau_1 - \tau_2)^{p/r}\big),
\end{aligned}
$$

where the last inequality follows as in the proof of Lemma A.2. Recall that $p/2 > 1 + \varepsilon$, and suppose that $\tau_2 < \tau_1$ and n satisfy $\varepsilon \leq k_n^{p/2-(1+\varepsilon)}(\tau_1 - \tau_2)^{p/v-p/r}$. Then, the above inequality implies

$$
\mathrm{E}\Big[\big|\tilde{e}_{n,j}(\tau_1) - \tilde{e}_{n,j}(\tau_2)\big|^p\Big] \leq \frac{2K}{\varepsilon}(\tau_1 - \tau_2)^{p/v}.
$$

Now let $\kappa = \kappa_n \in \mathbb{N}$ and $\mu = \mu_n > 0$ such that $\mu \geq (\varepsilon k_n^{-(p/2-(1+\varepsilon))})^{1/(p/v-p/r)}$ and $\delta := \kappa\mu$ is independent of n. We obtain, for all $\tau \in [0, \ell - \delta]$,

$$
\mathrm{E}\Big[\big|\tilde{e}_{n,j}(\tau + i\mu) - \tilde{e}_{n,j}(\tau + (i-1)\mu)\big|^p\Big] \leq \frac{2K}{\varepsilon}\mu^{p/v}, \quad i \in \{1, \ldots, \kappa\},
$$

and, by Theorem 12.2 in Billingsley (1968),

$$
\mathbb{P}\Big(\max_{1 \leq i \leq \kappa}\big|\tilde{e}_{n,j}(\tau + i\mu) - \tilde{e}_{n,j}(\tau)\big| > \varepsilon\Big) \leq \frac{2KK'}{\varepsilon^{p+1}}(\kappa\mu)^{p/v} = C\delta^{p/v} \tag{A.19}
$$

for some constants $K', C < \infty$. Further, note that $\tau \mapsto \mathbb{1}\big(N_{b_n,i}^{(\tau)} \leq j\big)$ is monotonically decreasing, which implies, for any $\delta' > 0$,

$$
\sup_{\tau_2 \leq \tau_1 \leq \tau_2 + \delta'}\big|\tilde{e}_{n,j}(\tau_1) - \tilde{e}_{n,j}(\tau_2)\big| \leq \big|\tilde{e}_{n,j}(\tau_2 + \delta') - \tilde{e}_{n,j}(\tau_2)\big| + \sqrt{k_n}\delta',
$$

yielding

$$
\sup_{\tau_2 \leq \tau_1 \leq \tau_2 + \kappa\mu}\big|\tilde{e}_{n,j}(\tau_1) - \tilde{e}_{n,j}(\tau_2)\big| \leq 3\max_{1 \leq i \leq \kappa}\big|\tilde{e}_{n,j}(\tau_2 + i\mu) - \tilde{e}_{n,j}(\tau_2)\big| + \sqrt{k_n}\mu
$$

by a similar reasoning as for the proof of (A.11). Let $\mu \leq \varepsilon/\sqrt{k_n}$ (see below that this is a valid choice). Then

$$
\begin{aligned}
&\mathbb{P}\Big(\sup_{\tau_2 \leq \tau_1 \leq \tau_2 + \delta}|\tilde{e}_{n,j}(\tau_1) - \tilde{e}_{n,j}(\tau_2)| > 4\varepsilon\Big) \\
&\leq \mathbb{P}\Big(\max_{1 \leq i \leq \ell}\big|\tilde{e}_{n,j}(\tau_2 + i\mu) - \tilde{e}_{n,j}(\tau_2)\big| > \varepsilon\Big) \\
&\leq C\delta^{p/v} < \delta\nu
\end{aligned}
$$

by inequality (A.19) and choosing δ such that $C\delta^{p/v-1} < \nu$. The proof is finished once we have shown that our choice of parameters is valid and in accordance with our Condition 3.1(ii). Above, we required

$$\left(\frac{\varepsilon}{k_n^{p/2-(1+\varepsilon)}}\right)^{\frac{1}{p/v-p/r}} \le \mu \le \frac{\varepsilon}{\sqrt{k_n}}, \qquad \eta > v/(v-2),$$
$$\eta \ge (p-1)r/(r-p), \qquad p/2 > 1+\varepsilon. \tag{A.20}$$

Some straightforward calculation similar to those on page 305f. in Robert (2009) yield that all four conditions are satisfied if $p = v(1+\varepsilon)$, $v = (3+\varepsilon)r/(r+1+\varepsilon)$, $\varepsilon < ((r-2) \wedge 1/2)/4$ and

$$\eta \ge \frac{3r}{r-2(1+2\varepsilon)}.$$

Since $\eta > 3$ by assumption, the latter can be guaranteed by increasing $r \to \infty$. □

B. Auxiliary lemmas - Sliding blocks

Throughout, we assume that Condition 3.1 is met and that, additionally, $\sqrt{k_n}\beta_{\varepsilon_2}(b_n) = o(1)$ for some $\varepsilon_2 > 0$. All convergences are for $n \to \infty$ if not stated otherwise. We will also occasionally omit the upper index sb at $\hat{H}_n^{\mathrm{sb}}, e_{n,j}^{\mathrm{sb}}$ and Z_{ni}^{sb}.

Lemma B.1. *For any* $j \in \mathbb{N}_{\ge 1}$,

$$\int_0^\infty e_{n,j}^{\mathrm{sb}}(\tau)\, \mathrm{d}(\hat{H}_n^{\mathrm{sb}} - H)(\tau) = o_{\mathbb{P}}(1).$$

Proof of Lemma B.1. The proof is very similar to the one of Lemma A.1. In fact, we need to show that (A.3) and (A.4) is met (for $v = 1, 2$), where $A_{n,\ell}$ and $B_{n,\ell,v}$ are defined as in (A.1) and (A.2), but with $\hat{H}_n = \hat{H}_n^{\mathrm{sb}}$ and $e_{n,j} = e_{n,j}^{\mathrm{sb}}$.

Invoking Theorem 7.1 instead of Lemma A.3, the proof of (A.3) is the same as in the proof Lemma A.1. (Note that we still have $\mathrm{E}[\hat{H}_n^{\mathrm{sb}}] = H$, and similar to $\mathrm{Var}(\hat{H}_n) = o(1)$ in the proof of Lemma A.1), it can be shown that $\mathrm{Var}(\hat{H}_n^{\mathrm{sb}}) = o(1)$.)

Regarding (A.4) with $v = 1$, write

$$\begin{aligned} B_{n,\ell,1} &= \frac{\sqrt{k_n}}{(n-b_n+1)^2} \sum_{i,i'=1}^{n-b_n+1} \Big\{\mathbb{1}\big(N_{b_n,i'}^{(Z_{ni})} = j\big) - \varphi_{n,j}(Z_{ni})\Big\}\mathbb{1}(Z_{ni} \ge \ell) \\ &= k_n^{-3/2} \sum_{i,i'=1}^{k_n-1} b_n^{-2} \sum_{s\in I_i}\sum_{s'\in I_{i'}} \Big\{\mathbb{1}\big(N_{b_n,s'}^{(Z_{ns})} = j\big) - \varphi_{n,j}(Z_{ns})\Big\} \\ &\qquad \times \mathbb{1}(Z_{ns} \ge \ell) + o(1) \\ &= V_{n,\ell,1} + V_{n,\ell,2} + o(1), \end{aligned}$$

where $V_{n,\ell,w}$ is made up from the same summands as in the line before, with the only difference that for $w = 1$ the sum over i' ranges from 1 to $i - 3$, and for $w = 2$ it goes from $i + 3$ to $k_n - 1$ (c.f. the proof of Lemma A.1). Now, for $\varepsilon = \varepsilon_1$ from Condition 3.1(ii), write c from Condition 3.1(iii) as $c = 1 - \varepsilon\kappa$ with $\kappa \in (0, 1)$, and let $C_n = \{\min_{i=1,...,n-b_n+1} N_{ni} > 1 - \varepsilon\kappa\}$, such that $\mathbb{P}(C_n) \to 1$ as $n \to \infty$ by Remark 3.2. Consequently, (A.4) with $v = 1$ follows if we show

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|V_{n,\ell,w}\mathbb{1}_{C_n}| > \delta\big) = 0$$

for $w \in \{1, 2\}$. In the following, we consider the case $w = 1$; the case $w = 2$ can be treated analogously. By the same reasoning and using the same notation as on page 38, we can write $V_{n,\ell,1}\mathbb{1}_{C_n} = V^{\varepsilon}_{n,\ell,1}\mathbb{1}_{C_n}$, where

$$V^{\varepsilon}_{n,\ell,1} = \frac{1}{k_n}\sum_{i=4}^{k_n-1}\frac{1}{b_n}\sum_{s\in I_i} f_{n,i-3}\big(Z^{\varepsilon\kappa}_{ns}\big)\mathbb{1}(\varepsilon\kappa b_n \ge Z^{\varepsilon\kappa}_{ns} \ge \ell)$$

and $f_{n,i-3}$ is given by

$$f_{n,i-3}(\tau) = k_n^{-1/2}\sum_{i'=1}^{i-3}\frac{1}{b_n}\sum_{s'\in I_{i'}}\big\{\mathbb{1}\big(N^{(\tau)}_{b_n,s',\varepsilon} = j\big) - \varphi_{n,j}(\tau)\big\}.$$

Note that, by construction, the observations making up $f_{n,i-3}$ are separated by at least one block of size b_n from the observations occurring in $Z^{\varepsilon\kappa}_{ns}$ for any $s \in I_4 \cup \ldots \cup I_{k_n-1}$, just as in the disjoint blocks case in the proof of Lemma A.1. As a matter of fact, following the proof of this lemma from page 38 onwards, one can show that

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|V^{\varepsilon}_{n,\ell,1}\mathbb{1}_{C_n}| > \delta\big) = 0,$$

overall proving (A.4) with $v = 1$.

Likewise, as for the process $e^{\mathrm{db}}_{n,j}$ in the disjoint blocks setting, we obtain the bound $\mathrm{E}[|e^{\mathrm{sb}}_{n,j}(\tau)|] = O(1)$ uniformly in $\tau \in [\ell, \varepsilon b_n]$. Consequently, by Markov's inequality, (A.4) with $v = 2$ follows from

$$\mathbb{P}(|B_{n,\ell,2}| > \delta) \lesssim \delta^{-1}\int_{\ell}^{\varepsilon b_n} \mathrm{d}H(\tau) + \delta^{-1}\sqrt{k_n}\int_{\varepsilon b_n}^{\infty} \mathrm{d}H(\tau),$$

which which converges to zero for $n \to \infty$ followed by $\ell \to \infty$. This concludes the proof. □

Lemma B.2. *For any $m \in \mathbb{N}_{\ge 1}$,*

$$\frac{\sqrt{k_n}}{n - b_n + 1}\sum_{i=1}^{n-b_n+1}\big(W^{\mathrm{sb}}_{n,i}(1), \ldots, W^{\mathrm{sb}}_{n,i}(m)\big) \xrightarrow{d} (s^{\mathrm{sb}}_1, \ldots, s^{\mathrm{sb}}_m) \sim \mathcal{N}_m(0, \Sigma^{\mathrm{sb}}_m),$$

where $W^{\mathrm{sb}}_{n,i}(j)$ is defined as in (6.3) but with db *replaced by* sb, *and where $\Sigma^{\mathrm{sb}}_m = (d^{\mathrm{sb}}_{j,j'})_{1\le j,j'\le m}$ is defined in (4.3).*

Proof of Lemma B.2. By the Cramér-Wold device it suffices to show that

$$\sum_{j=1}^{m} \lambda_j \frac{\sqrt{k_n}}{n-b_n+1} \sum_{i=1}^{n-b_n+1} W_{n,i}^{\mathrm{sb}}(j) \xrightarrow{d} \sum_{j=1}^{m} \lambda_j s_j^{\mathrm{sb}}$$

for arbitrary $\lambda_j \in \mathbb{R}$. Write the right-hand side as

$$\sum_{i=1}^{k_n-1} \sum_{s\in I_i} \sum_{j=1}^{m} \lambda_j \frac{\sqrt{k_n}}{n-b_n+1} \Big\{ \int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j) - \varphi_{n,j}(\tau)\,\mathrm{d}H(\tau)$$
$$+ \varphi_{n,j}(Z_{ns}) - \mathrm{E}\left[\varphi_{n,j}(Z_{ns})\right\} + o_{\mathbb{P}}(1),$$

where the small $o_{\mathbb{P}}(1)$ term is due to the fact that a negligible number of summands has been omitted. To take care of the serial dependence of the sliding blocks, we apply a similar construction as in the proof of Lemma 7.3. Using the same notation as in that proof, write $V_n^{\pm} = (q_n^*)^{-1/2} \sum_{i=1}^{q_n^*} T_{ni}^{\pm}$ with

$$T_{ni}^{\pm} = \sqrt{\frac{q_n^*}{k_n}} \sum_{s\in J_i^{\pm}} \sum_{j=1}^{m} \lambda_j \frac{n}{n-b_n+1} \frac{1}{b_n} \Big\{ \int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j) - \varphi_{n,j}(\tau)\,\mathrm{d}H(\tau)$$
$$+ \varphi_{n,j}(Z_{ns}) - \mathrm{E}\left[\varphi_{n,j}(Z_{ns})\right]\Big\}.$$

Since

$$\Big| \int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j) - \varphi_{n,j}(\tau)\,\mathrm{d}H(\tau) \Big| + \left|\varphi_{n,j}(Z_{ns}) - \mathrm{E}\left[\varphi_{n,j}(Z_{ns})\right]\right| \le 2,$$

we still obtain the upper bound in (8.7). Note that from relation (8.7) forward the proof of Lemma 7.3 actually does not depend on the concrete form of the $T_{ni}^{\pm}$ but only makes use of the block structure and mixing conditions, which is why the remaining proof is the same as in Lemma 7.3. In particular, note that

$$T_{n1}^{+} = \sum_{j=1}^{m} \lambda_j \frac{\sqrt{k_n^*}}{n^*-b_n+1} \sum_{i=1}^{n^*-b_n+1} W_{n^*,i}(j) + R_n,$$

where $R_n \to 0$ in $L_2(\mathbb{P})$ and $n^* = k_n^* b_n$, and that our assumptions in Condition 3.1 still hold if n and k_n are substituted by n^* and k_n^*. The assertion then follows from Lemma B.3 below. □

Lemma B.3. *For any $j, j' \in \mathbb{N}_{\ge 1}$, we have*

$$\lim_{n\to\infty} \mathrm{Cov}\Big(\frac{\sqrt{k_n}}{n-b_n+1} \sum_{i=1}^{n-b_n+1} W_{n,i}^{\mathrm{sb}}(j), \frac{\sqrt{k_n}}{n-b_n+1} \sum_{i=1}^{n-b_n+1} W_{n,i}^{\mathrm{sb}}(j') \Big) = d_{j,j'}^{\mathrm{sb}},$$

where $d_{j,j'}^{\mathrm{sb}}$ is defined in (4.3).

Proof of Lemma B.3. Assume that all U_s are $\mathcal{B}^{\varepsilon}_{s:s}$-measurable with $\varepsilon = \varepsilon_1$ from Condition 3.1; the general case can be treated by multiplying with suitable indicator functions as in the previous proofs. Write

$$\begin{aligned}
&\operatorname{Cov}\Big(\frac{\sqrt{k_n}}{n-b_n+1}\sum_{i=1}^{n-b_n+1} W^{\mathrm{sb}}_{n,i}(j), \frac{\sqrt{k_n}}{n-b_n+1}\sum_{i=1}^{n-b_n+1} W^{\mathrm{sb}}_{n,i}(j')\Big)\\
&= C_{n1} + C_{n2} + C_{n3} + C_{n4}, \qquad \text{(B.1)}
\end{aligned}$$

where

$$\begin{aligned}
C_{n1} &= \operatorname{Cov}\Big(\int_0^\infty e_{n,j}(\tau)\,\mathrm{d}H(\tau), \int_0^\infty e_{n,j'}(\tau)\,\mathrm{d}H(\tau)\Big)\\
C_{n2} &= \frac{k_n}{(n-b_n+1)^2}\sum_{i,i'=1}^{n-b_n+1} \operatorname{Cov}\big(\varphi_{n,j}(Z_{ni}), \varphi_{n,j'}(Z_{ni'})\big)\\
C_{n3} &= \frac{k_n}{(n-b_n+1)^2}\sum_{i,i'=1}^{n-b_n+1} \operatorname{Cov}\Big(\int_0^\infty \mathbb{1}(N^{(\tau)}_{b_n,i} = j)\,\mathrm{d}H(\tau), \varphi_{n,j'}(Z_{ni'})\Big)\\
C_{n4} &= \frac{k_n}{(n-b_n+1)^2}\sum_{i,i'=1}^{n-b_n+1} \operatorname{Cov}\Big(\int_0^\infty \mathbb{1}(N^{(\tau)}_{b_n,i} = j')\,\mathrm{d}H(\tau), \varphi_{n,j}(Z_{ni'})\Big).
\end{aligned}$$

By Lemma 8.1, the first term C_{n1} satisfies

$$\begin{aligned}
\lim_{n\to\infty} C_{n1} = 2\int_0^1\int_0^\infty\int_0^\infty \operatorname{Cov}\big(&\mathbb{1}(X^{(\tau)}_{1,\xi} = j),\\
&\mathbb{1}(Y^{(\tau')}_{1,\xi} = j')\big)\,\mathrm{d}H(\tau)\mathrm{d}H(\tau') \qquad \text{(B.2)}
\end{aligned}$$

As at the beginning of the proof of Lemma 8.1, the second term can be shown to satisfy $C_{n2} = T_{n1} + T_{n2} + o(1)$, where

$$\begin{aligned}
T_{n1} &= \frac{1}{b_n^2}\sum_{s,t\in I_1} \operatorname{Cov}\big(\varphi_{n,j}(Z_{ns}), \varphi_{n,j'}(Z_{nt})\big),\\
T_{n2} &= \frac{1}{b_n^2}\sum_{s\in I_1}\sum_{t\in I_2} \operatorname{Cov}\big(\varphi_{n,j}(Z_{ns}), \varphi_{n,j'}(Z_{nt})\big)\\
&\qquad + \frac{1}{b_n^2}\sum_{s\in I_2}\sum_{t\in I_1} \operatorname{Cov}\big(\varphi_{n,j}(Z_{ns}), \varphi_{n,j'}(Z_{nt})\big).
\end{aligned}$$

Let us start with T_{n1}. We know that $\mathrm{E}[\varphi_{n,j'}(Z_{nt})] \to \bar p(j')$ by the proof of Lemma A.2, which implies

$$T_{n1} = \frac{1}{b_n^2}\sum_{s,t=1}^{b_n} \mathrm{E}[\varphi_{n,j}(Z_{ns})\varphi_{n,j'}(Z_{nt})] - \bar p(j)\bar p(j') + o(1).$$

As in the proof of Lemma 8.1 we can write

$$\frac{1}{b_n^2}\sum_{s,t=1}^{b_n} \mathrm{E}[\varphi_{n,j}(Z_{ns})\varphi_{n,j'}(Z_{nt})] = \int_0^1\int_0^1 g_n(\xi,z)\,\mathrm{d}z\mathrm{d}\xi,$$

where

$$g_n(\xi,z) = \mathrm{E}[\varphi_{n,j}(Z_{n,\lfloor b_n z\rfloor+1})\varphi_{n,j'}(Z_{n,\lfloor b_n\xi\rfloor+1})].$$

Let $z\le\xi$. Set $r_n=\lfloor b_n\xi\rfloor-\lfloor b_n z\rfloor$. For $x,y>0$ consider

$$\begin{aligned}
&\mathbb{P}\big(Z_{n,\lfloor b_n z\rfloor+1}>x, Z_{n,\lfloor b_n\xi\rfloor+1}>y\big)\\
&=\mathbb{P}\big(N_{1:b_n}<1-\tfrac{x}{b_n},\ N_{r_n+1:r_n+b_n}<1-\tfrac{y}{b_n}\big)\\
&=\mathbb{P}\big(N_{1:r_n}<1-\tfrac{x}{b_n},\ N_{r_n+1:b_n}<1-\tfrac{x\vee y}{b_n},\ N_{b_n+1:r_n+b_n}<1-\tfrac{y}{b_n}\big)
\end{aligned}$$

where $N_{s:t}=\max(U_s,\dots,U_t)$ for $s,t\in\mathbb{N}_{\ge1}$ with $s\le t$. Note that $\mathbb{P}(N_{1:q_n}>1-z/b_n)\le zq_n/b_n\to0$ for any integer sequence $q_n=o(b_n)$ that is converging to infinity. Similar as in the step (8.14) in the proof of Lemma 8.1, this implies that the expression in the previous display equals

$$\begin{aligned}
\mathbb{P}\big(N_{1:r_n}<1-\tfrac{x}{b_n}\big)\,\mathbb{P}\big(N_{r_n+1:b_n}<1-\tfrac{x\vee y}{b_n}\big)\,\mathbb{P}\big(N_{b_n+1:r_n+b_n}<1-\tfrac{y}{b_n}\big)&\\
+O(\alpha_\varepsilon(q_n))+O((x\vee y)q_n/b_n),&
\end{aligned}$$

which by (2.4) converges to

$$H_{\xi-z}(x,y):=\exp\big(-\theta\{(x\wedge y)(\xi-z)+(x\vee y)\}\big).$$

As a consequence, by the definition of $(X_{2,\xi-z},Y_{2,\xi-z})$ in Theorem 4.2,

$$(Z^{\mathrm{sb}}_{n,\lfloor b_n z\rfloor+1},Z^{\mathrm{sb}}_{n,\lfloor b_n\xi\rfloor+1})\xrightarrow{d}(X_{2,\xi-z},Y_{2,\xi-z}),$$

As in the proof of Lemma A.2, the extended continuous mapping theorem and Theorem 2.20 in van der Vaart (1998) imply

$$\lim_{n\to\infty}g_n(z,\xi)=\mathrm{E}[p^{(X_{2,\xi-z})}(j)p^{(Y_{2,\xi-z})}(j')]$$

for $z\le\xi$. By symmetry, for $z>\xi$,

$$\lim_{n\to\infty}g_n(z,\xi)=\mathrm{E}[p^{(X_{2,z-\xi})}(j)p^{(Y_{2,z-\xi})}(j')]$$

A simple calculation then shows that

$$\lim_{n\to\infty}\int_0^1\int_0^1 g_n(\xi,z)\,\mathrm{d}z\mathrm{d}\xi = 2\int_0^1(1-\xi)\,\mathrm{E}[p^{(X_{2,\xi})}(j)p^{(Y_{2,\xi})}(j')]\,\mathrm{d}\xi,$$

Altogether, we have that

$$\lim_{n\to\infty}T_{n1}=2\int_0^1(1-\xi)\,\mathrm{E}\left[p^{(X_{2,\xi})}(j)p^{(Y_{2,\xi})}(j')\right]\mathrm{d}\xi-\bar p(j)\bar p(j').$$

Analogously, the term T_{n2} can be seen to satisfy

$$\lim_{n\to\infty} T_{n2} = 2\int_0^1 \xi \,\mathrm{E}\left[p^{(X_{2,\xi})}(j)p^{(Y_{2,\xi})}(j')\right] \mathrm{d}\xi - \bar{p}(j)\bar{p}(j'),$$

such that

$$\begin{aligned} C_{n2} = T_{n1} + T_{n2} + o(1) &\to 2\int_0^1 \mathrm{E}\left[p^{(X_{2,\xi})}(j)p^{(Y_{2,\xi})}(j')\right] \mathrm{d}\xi - 2\bar{p}(j)\bar{p}(j') \\ &= 2\int_0^1 \mathrm{Cov}\left(p^{(X_{2,\xi})}(j), p^{(Y_{2,\xi})}(j')\right) \mathrm{d}\xi. \end{aligned} \tag{B.3}$$

Next, consider C_{n3} in (B.1), which may be written as $C_{n3} = S_{n1}+S_{n2}+o(1)$, where

$$\begin{aligned} S_{n1} &= \frac{1}{b_n^2}\sum_{s,t\in I_1} \mathrm{Cov}\left(\int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j)\,\mathrm{d}H(\tau), \varphi_{n,j'}(Z_{nt})\right) \\ S_{n2} &= \frac{1}{b_n^2}\bigg\{\sum_{s\in I_1}\sum_{t\in I_2} \mathrm{Cov}\left(\int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j)\,\mathrm{d}H(\tau), \varphi_{n,j'}(Z_{nt})\right) \\ &\qquad + \sum_{s\in I_2}\sum_{t\in I_1} \mathrm{Cov}\left(\int_0^\infty \mathbb{1}(N_{b_n,s}^{(\tau)} = j)\,\mathrm{d}H(\tau), \varphi_{n,j'}(Z_{nt})\right)\bigg\}. \end{aligned}$$

By similar arguments as before, we obtain

$$\begin{aligned} S_{n1} = \int_0^1\int_0^1\int_0^\infty \mathrm{E}\left[\mathbb{1}\left(N_{b_n,\lfloor b_n z\rfloor+1}^{(\tau)} = j\right)\varphi_{n,j'}(Z_{n,\lfloor b_n\xi\rfloor+1})\right] \mathrm{d}H(\tau)\,\mathrm{d}z\,\mathrm{d}\xi \\ - \bar{p}(j)\bar{p}(j') + o(1). \end{aligned}$$

To analyze the convergence of the product moment in the previous display we start by showing that

$$\left(N_{b_n,\lfloor b_n z\rfloor+1}^{(\tau)}, Z_{n,\lfloor b_n\xi\rfloor+1}\right) \xrightarrow{d} \left(X_{3,|\xi-z|}^{(\tau)}, Y_{3,|\xi-z|}\right)$$

where $(X_{3,\zeta}^{(\tau)}, Y_{3,\zeta})$ is defined in Theorem 4.2. For $x > 0, j \in \mathbb{N}_{\geq 0}$ and $0 \leq z \leq \xi \leq 1$, write

$$\mathbb{P}\left(N_{b_n,\lfloor b_n z\rfloor+1}^{(\tau)} = j, Z_{n,\lfloor b_n\xi\rfloor+1} > x\right) = \mathbb{P}\left(N_{b_n,\lfloor b_n z\rfloor+1}^{(\tau)} = j, N_{b_n,\lfloor b_n\xi\rfloor+1}^{(x)} = 0\right)$$

which is exactly of the form of φ_n in (8.11) and hence converges to

$$\begin{aligned} &H_{0,j}^{(x,\tau)}(\xi - z)\mathbb{1}(x \leq \tau) + H_{j,0}^{(\tau,x)}(\xi - z)\mathbb{1}(x > \tau) \\ &= \sum_{l=0}^{j} p^{(\tau(\xi-z))}(l)p^{(x(\xi-z))}(0)p_2^{((1-\xi+z)\tau,(1-\xi+z)x)}(j-l,0)\mathbb{1}(x \leq \tau) \\ &\qquad + p^{(\tau(\xi-z))}(j)e^{-\theta x}\mathbb{1}(x > \tau) \end{aligned}$$

$$= \mathbb{P}\big(X^{(\tau)}_{3,\xi-z} = j, Y_{3,\xi-z} > x\big)$$

by the proof of Lemma 8.1, where the last equation follows from the definition of $(X^{(\tau)}_{3,\zeta}, Y_{3,\zeta})$ in Theorem 4.2. The same arguments as before in combination with the dominated convergence theorem implies

$$\lim_{n\to\infty} \mathrm{E}\left[\mathbb{1}\big(N^{(\tau)}_{b_n,\lfloor b_n z\rfloor+1} = j\big)\varphi_{n,j'}(Z_{n,\lfloor b_n\xi\rfloor+1})\right] = \mathrm{E}\left[\mathbb{1}\big(X^{(\tau)}_{3,\xi-z} = j\big)p^{(Y_{3,\xi-z})}(j')\right].$$

For the case $z > \xi$, one obtains the same limiting expression, but with z and ξ interchanged. As a consequence, by similar arguments as for C_{n2},

$$\begin{aligned}
&\lim_{n\to\infty} S_{n1} \\
&= \int_0^1\int_0^1\int_0^\infty \mathrm{E}\left[\mathbb{1}\big(X^{(\tau)}_{3,\xi-z} = j\big)p^{(Y_{3,\xi-z})}(j')\right] \mathrm{d}H(\tau)\,\mathrm{d}z\,\mathrm{d}\xi - \bar{p}(j)\bar{p}(j') \\
&= 2\int_0^1\int_0^\infty (1-\xi)\,\mathrm{E}\left[\mathbb{1}\big(X^{(\tau)}_{3\xi} = j\big)p^{(Y_{3\xi})}(j')\right] \mathrm{d}H(\tau)\,\mathrm{d}\xi - \bar{p}(j)\bar{p}(j').
\end{aligned}$$

A similar argumentation for S_{n2} finally implies

$$\begin{aligned}
C_{n3} &= S_{n1} + S_{n2} + o(1) \\
&\to 2\int_0^1\int_0^\infty \mathrm{E}\left[\mathbb{1}\big(X^{(\tau)}_{3\xi} = j\big)p^{(Y_{3\xi})}(j')\right] \mathrm{d}H(\tau)\,\mathrm{d}\xi - \bar{p}(j)\bar{p}(j') \\
&= 2\int_0^1\int_0^\infty \mathrm{Cov}\left(\mathbb{1}\big(X^{(\tau)}_{3,\xi} = j\big), p^{(Y_{3,\xi})}(j')\right) \mathrm{d}H(\tau)\,\mathrm{d}\xi, \qquad \text{(B.4)}
\end{aligned}$$

where we have used that $X_{3,\xi} \sim p^{(\tau)}$ and $Y_{3,\xi} \sim \mathrm{Exponential}(\theta)$. The assertion is a consequence of (B.1) and (B.2), (B.3), (B.4), and the fact that C_{n4} has the same limit as C_{n3}, but with interchanged roles of j and j'. □

C. Further auxiliary results

Lemma C.1 (Bradley, 1983)**.** *If X and Y are two random variables in some Borel space S and $\mathbb{R}$, respectively, if U is uniform on $[0,1]$ and independent of (X,Y) and if $q > 0$ and $\gamma > 0$ are such that $q \le ||Y||_\gamma = \mathrm{E}[|Y|^\gamma]^{1/\gamma}$, then there exists a measurable function f such that $Y^* = f(X,Y,U)$ has the same distribution as Y, is independent of X and satisfies*

$$\mathbb{P}(|Y - Y^*| \ge q) \le 18(||Y||_\gamma/q)^{\gamma/(2\gamma+1)}\alpha(\sigma(X),\sigma(Y))^{2\gamma/(2\gamma+1)}.$$

Lemma C.2 (Berbee, 1979)**.** *If X and Y are two random variables in some Borel spaces S_1 and S_2, respectively, then there exists a random variable U independent of (X,Y) and a measurable function f such that $Y^* = f(X,Y,U)$ has the same distribution as Y, is independent of X and satisfies $\mathbb{P}(Y \ne Y^*) = \beta(\sigma(X),\sigma(Y))$.*

D. Further simulation results

This section contains additional simulation results for the ARCH,- ARMAX- and AR-model described in Section 5, see Figure D.1-D.14.

Acknowledgements

This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation, which is gratefully acknowledged. The authors would like to thank two anonymous referees and an Associate Editor for their constructive comments on an earlier version of this manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Berbee, H. C. P. (1979). *Random walks with stationary increments and renewal theory* volume 112 of *Mathematical Centre Tracts*. Amsterdam: Mathematisch Centrum.

Berghaus, B., & Bücher, A. (2017). Goodness-of-fit tests for multivariate copula-based time series models. *Econometric Theory*, *33*, 292–330.

Berghaus, B., & Bücher, A. (2018). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Ann. Statist.*, *46*, 2307–2335. URL: https://doi.org/10.1214/17-AOS1621. doi:10.1214/17-AOS1621.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Bradley, R. C. (1983). Approximation theorems for strongly mixing random variables. *Michigan Math. J.*, *30*, 69–81. URL: http://dx.doi.org/10.1307/mmj/1029002789. doi:10.1307/mmj/1029002789.

Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, *2*, 107–144. URL: http://dx.doi.org/10.1214/154957805100000104. doi:10.1214/154957805100000104. Update of, and a supplement to, the 1986 original.

Dehling, H., & Philipp, W. (2002). Empirical process techniques for dependent data. In *Empirical process techniques for dependent data* (pp. 3–113). Boston, MA: Birkhäuser Boston. URL: http://dx.doi.org/10.1007/978-1-4612-0099-4_1. doi:10.1007/978-1-4612-0099-4_1.

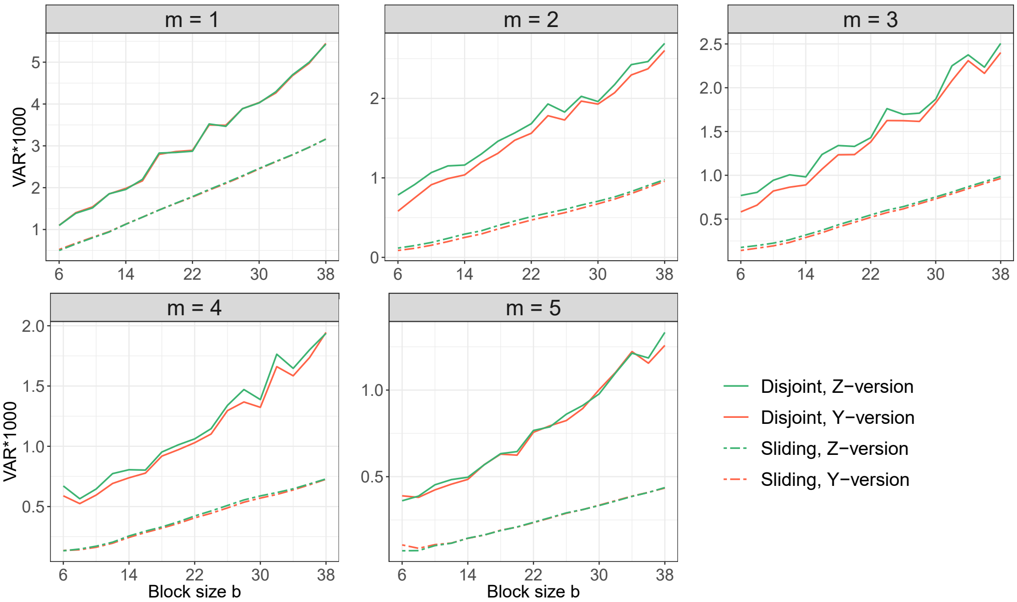

Figure D.1: Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \dots, 5$.

Robert, C. Y. (2009). Inference for the limiting cluster size distribution of extreme values. *Ann. Statist.*, *37*, 271–310. URL: `http://dx.doi.org/10.1214/07-AOS551`. doi:`10.1214/07-AOS551`.

Shao, Q.-M., & Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.*, *24*, 2098–2127. URL: `http://dx.doi.org/10.1214/aop/1041903220`. doi:`10.1214/aop/1041903220`.

van der Vaart, A. W. (1998). *Asymptotic statistics* volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press.

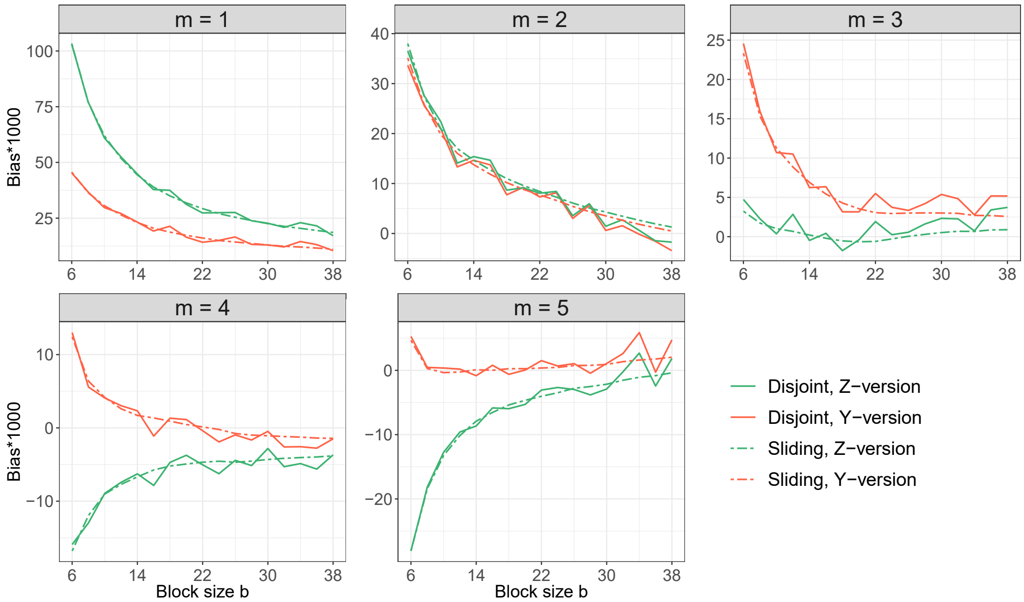

Figure D.2: Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \ldots, 5$.

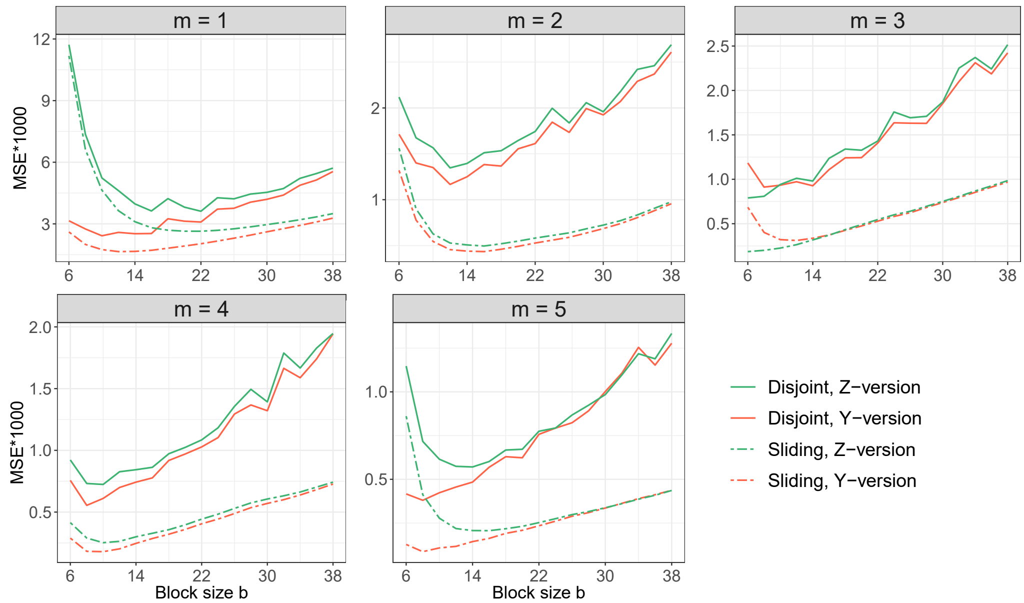

Figure D.3: Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \ldots, 5$.

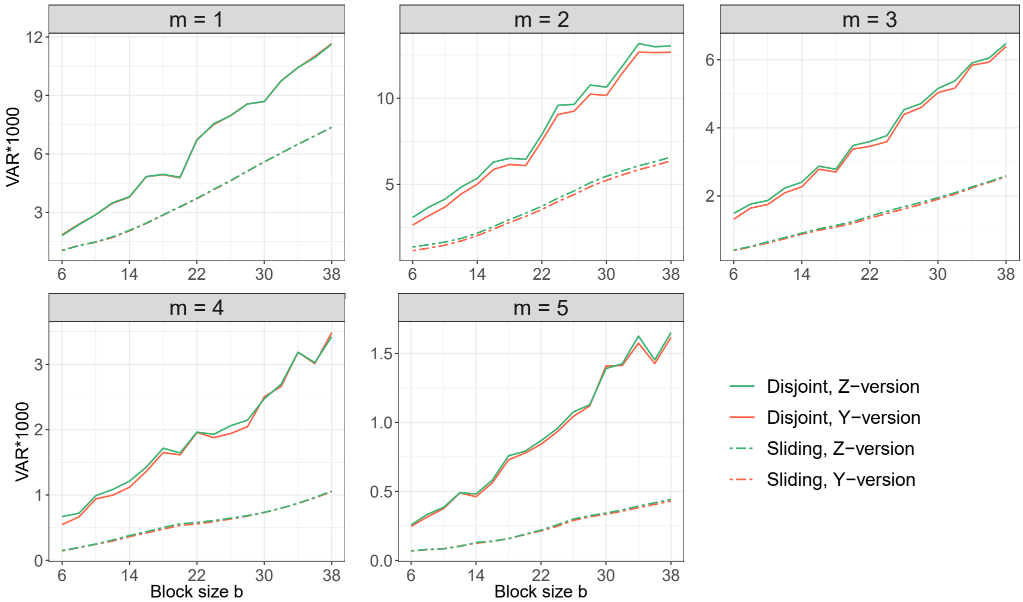

Figure D.4: Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \dots, 5$.

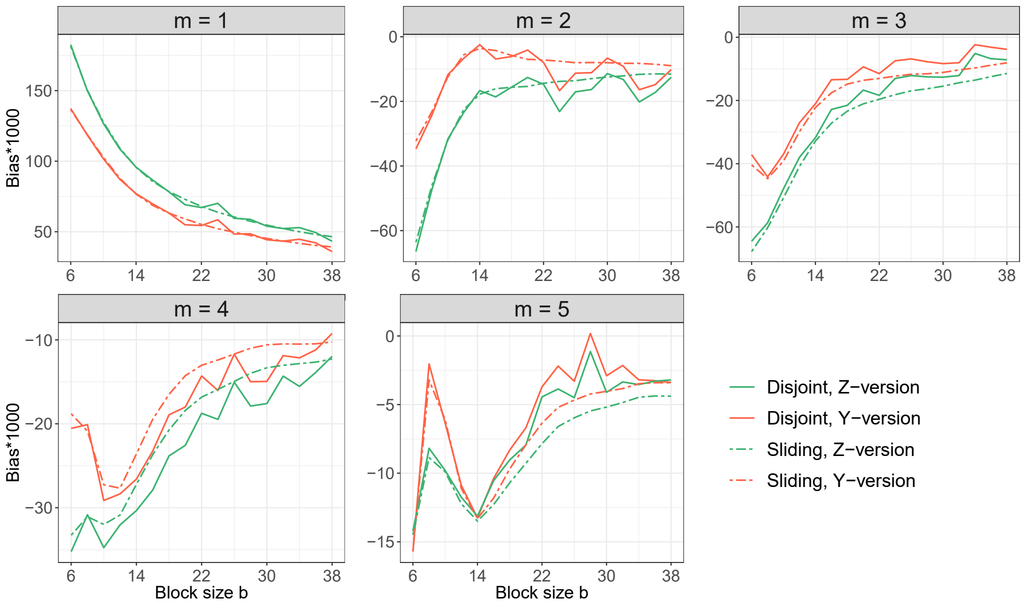

Figure D.5: Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \dots, 5$.

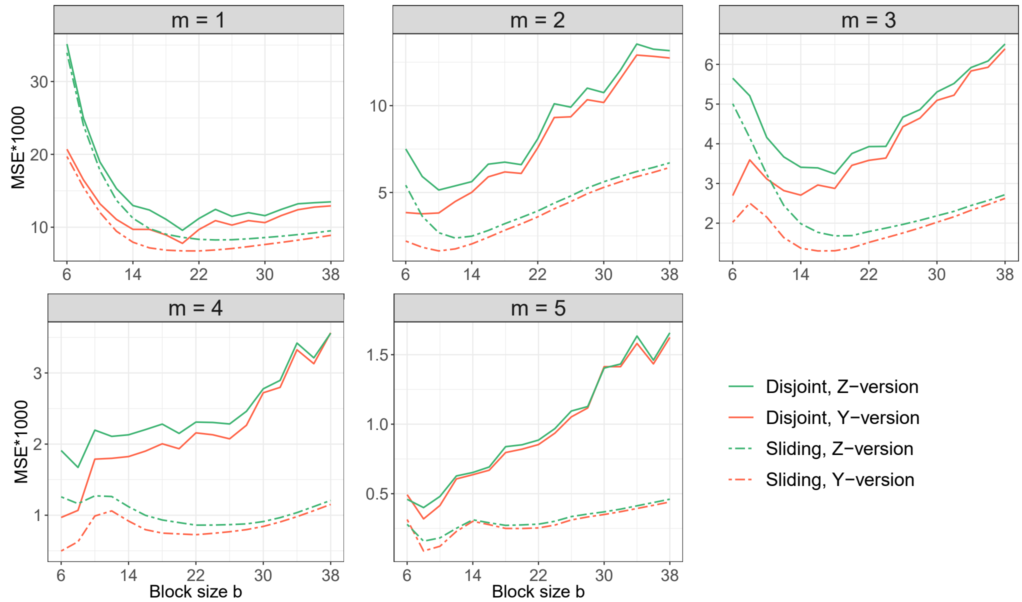

Figure D.6: Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \dots, 5$.

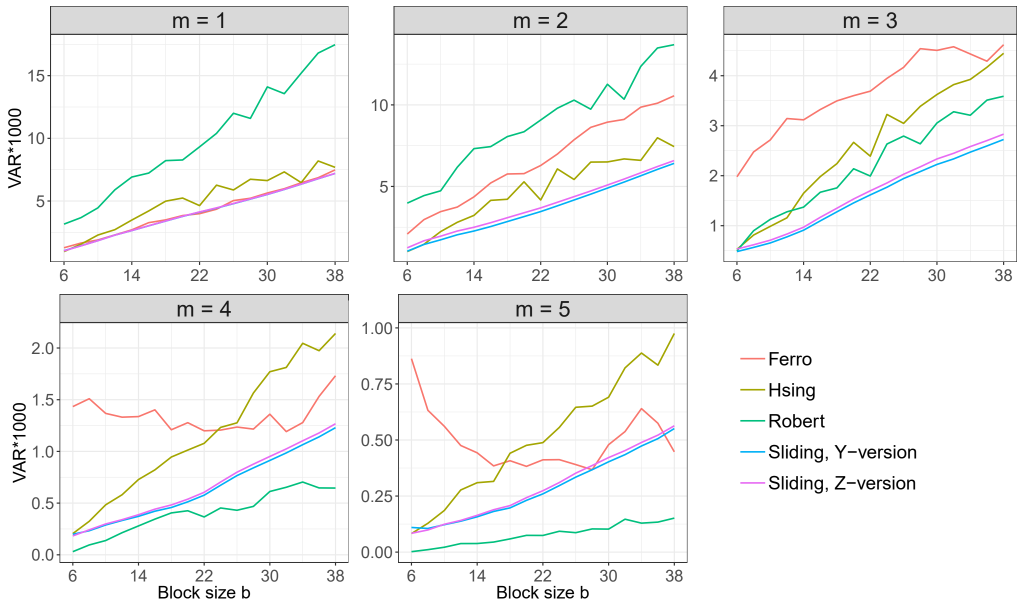

Figure D.7: Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \dots, 5$.

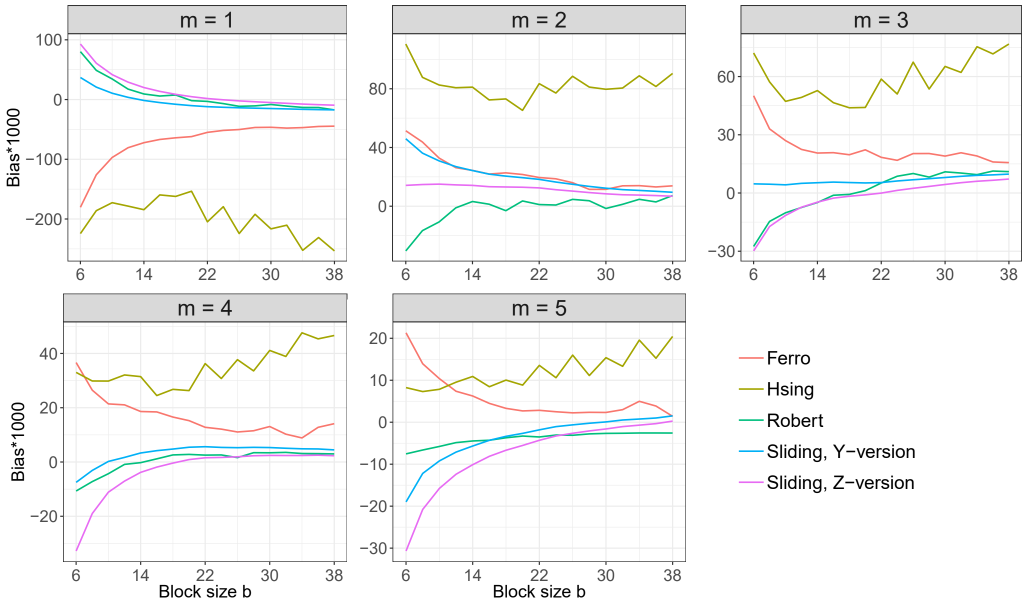

Figure D.8: Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the squared ARCH-model for $m = 1, \ldots, 5$.

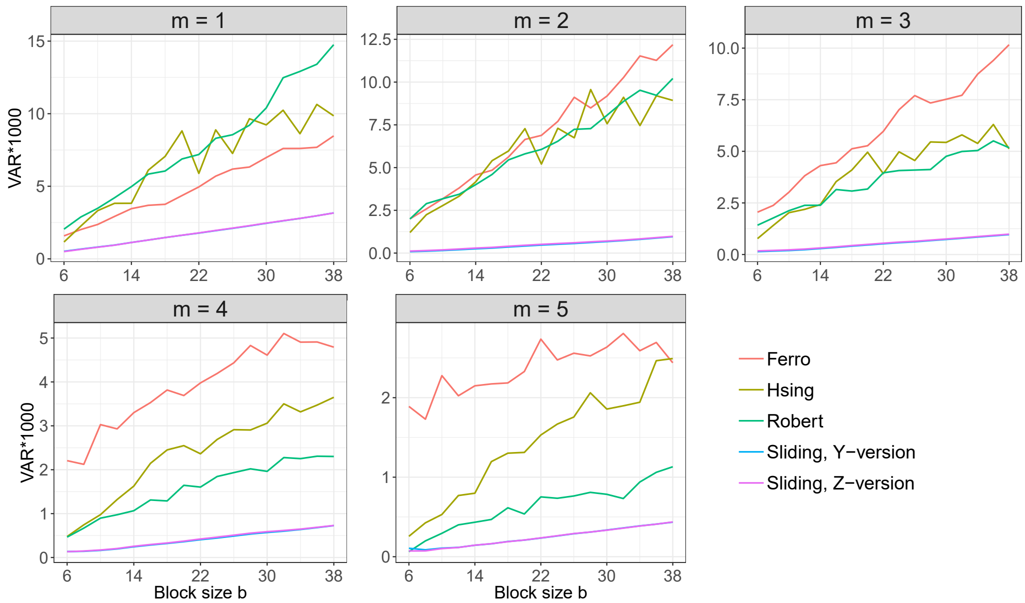

Figure D.9: Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \ldots, 5$.

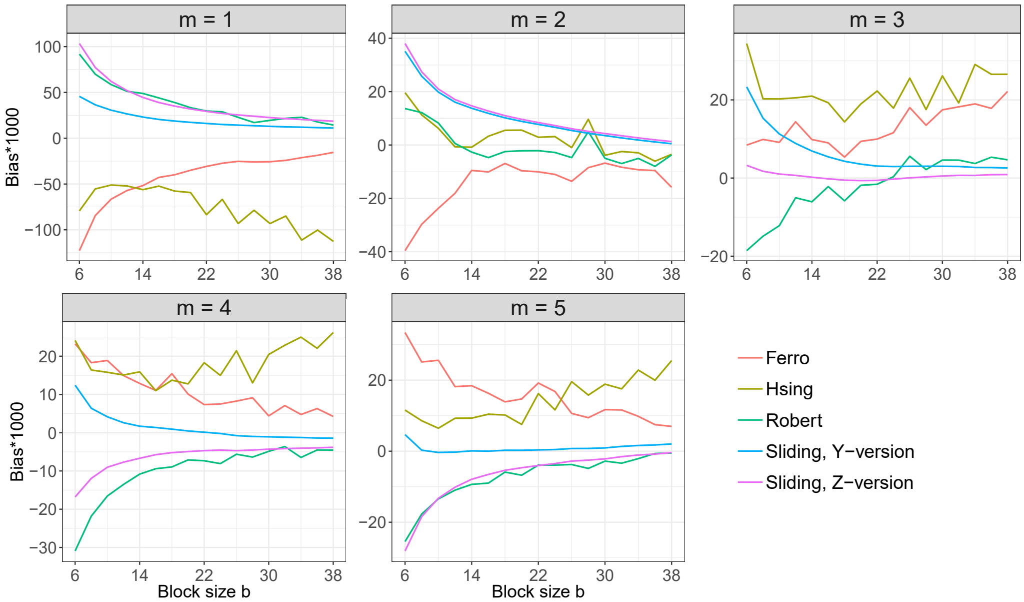

Figure D.10: Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \ldots, 5$.

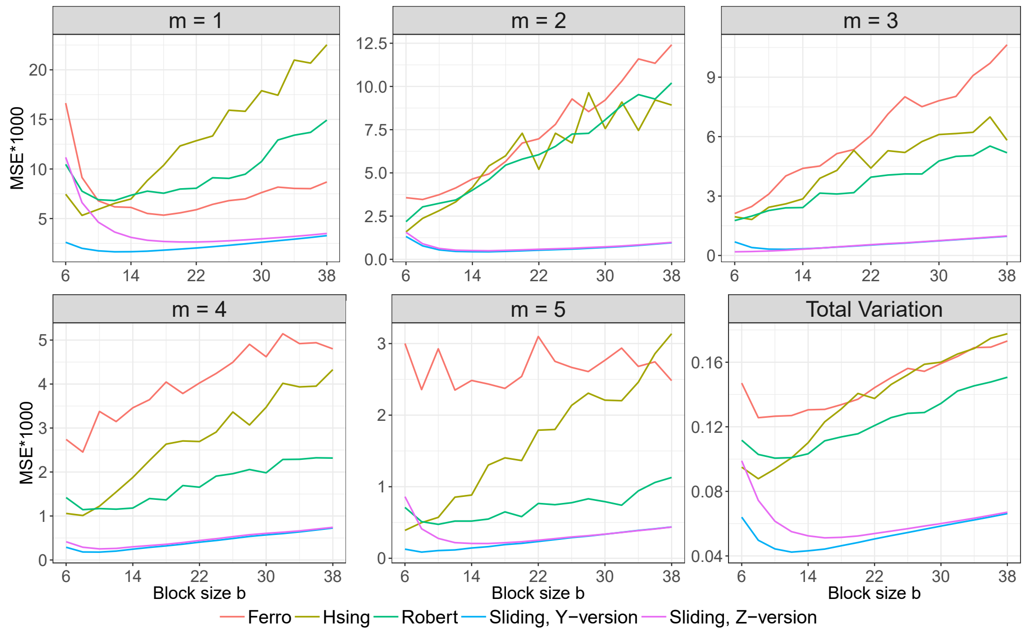

Figure D.11: Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the ARMAX-model for $m = 1, \ldots, 5$ and the total variation between $\hat{\pi}$ and π..

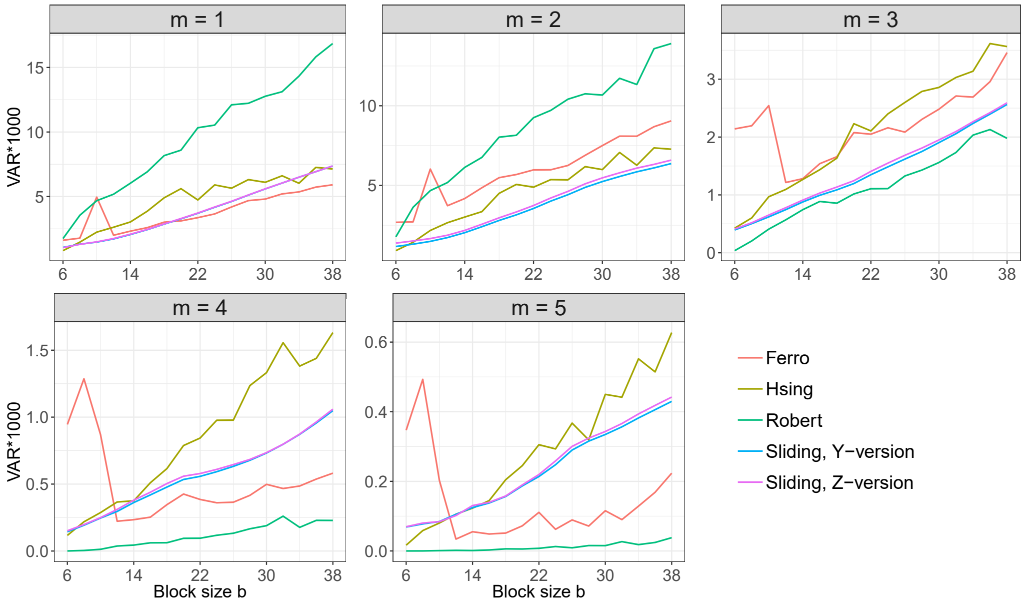

Figure D.12: Variance multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \ldots, 5$.

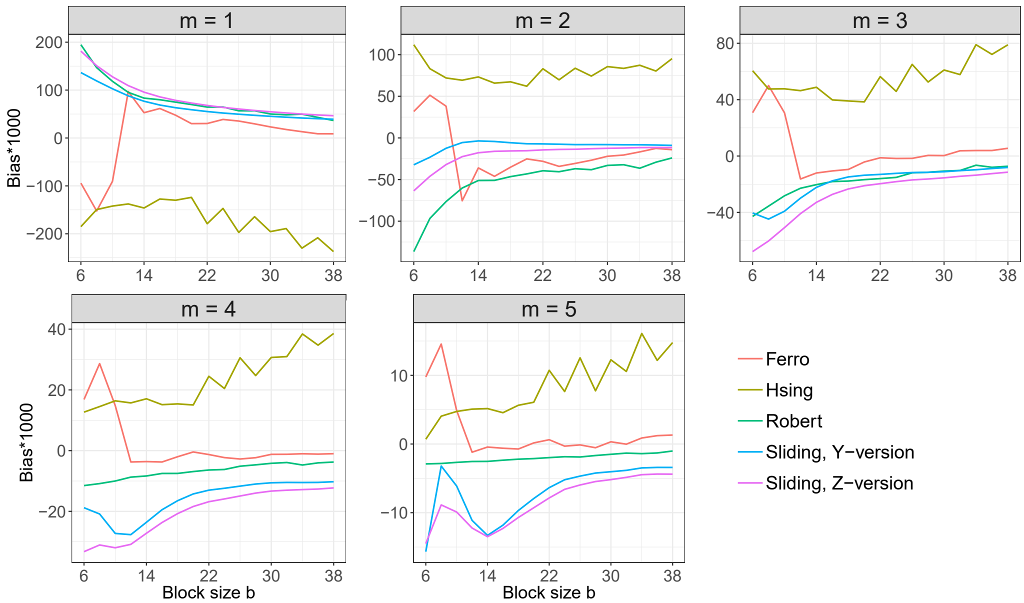

Figure D.13: Bias multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \ldots, 5$.

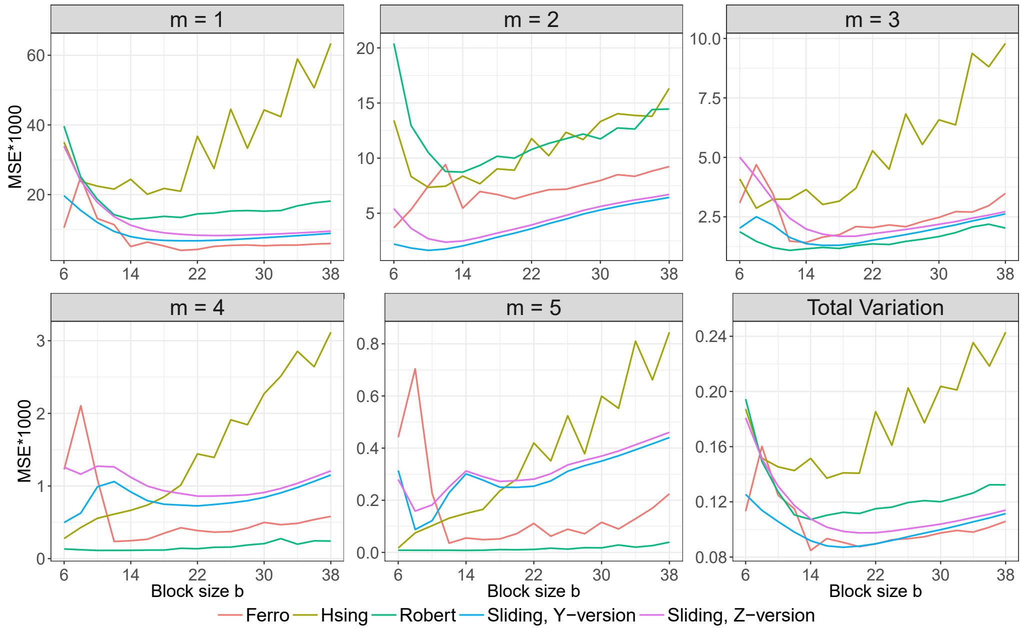

Figure D.14: Mean squared error multiplied by 10^3 for the estimation of $\pi(m)$ within the AR-model for $m = 1, \ldots, 5$ and the total variation between $\hat{\pi}$ and π.

Figure D.15: Mean squared error multiplied by 10^3 for the estimation of θ within the AR-, ARMAX- and squared ARCH-model.

STATISTICS FOR HETEROSCEDASTIC TIME SERIES EXTREMES

AXEL BÜCHER, TOBIAS JENNESSEN

Abstract. Einmahl, de Haan and Zhou (2016, Journal of the Royal Statistical Society: Series B, 78(1), 31– 51) recently introduced a stochastic model that allows for heteroscedasticity of extremes. The model is extended to the situation where the observations are serially dependent, which is crucial for many practical applications. We prove a local limit theorem for a kernel estimator for the scedasis function, and a functional limit theorem for an estimator for the integrated scedasis function. We further prove consistency of a bootstrap scheme that allows to test for the null hypothesis that the extremes are homoscedastic. Finally, we propose an estimator for the extremal index governing the dynamics of the extremes and prove its consistency. All results are illustrated by Monte Carlo simulations. An important intermediate result concerns the sequential tail empirical process under serial dependence.

Key words. Extremal Index; Kernel Estimator; Multiplier Bootstrap; Non-Stationary Extremes; Regular Varying Time Series.

Contents

Date: April 21, 2022.

1. Introduction

Classical extreme value statistics is concerned with analyzing the extremal behavior of a series of independent and identically distributed (i.i.d.) random variables. However, in many practical applications, the latter assumption is not justifiable, since the data typically consist of observations collected on one or more variables over time. The observations may then both exhibit serial dependence and they may be drawn from a distribution that changes smoothly (or even abruptly) as time progresses. The latter is particularly the case in many applications from environmental statistics (e.g., due to climate change), while the former is also omnipresent in typical applications from finance.

While an abundance of methods has been proposed for tackling the resulting challenges concerning the bulk of the data (see, e.g., Brockwell and Davis, 1991 for a classical account on time series analysis; Dahlhaus, 2012 for an overview on locally stationary processes that allow for nonparametric smooth changes over time; or Aue and Horváth, 2013 for an overview on results for change point analysis involving abrupt changes), respective results concerning extreme value analysis are much less developed, in particular for the situation exhibiting both serial dependence and non-stationarity.

Theoretical results on extreme value analysis for stationary time series build on corresponding probabilistic theory summarized in Leadbetter et al. (1983), see also Chapter 10 in Beirlant et al. (2004) for an overview or Kulik and Soulier (2020) for a modern account in the heavy tailed case. Respective asymptotic results on a large class of estimators for the tail index can be found in Drees (2000), with some substantial extensions on important intermediate results in Drees and Rootzén (2010). Results regarding the time series dynamics for the heavy tailed case can be found in Kulik and Soulier (2020) and the references therein. Smooth non-stationarity has often been approached by parametric regression models, see, e.g., Davison and Smith (1990); Coles (2001), where, however, no asymptotic theory is provided. Nonparametric approaches that were supported by asymptotic results can be found in Hall and Tajvidi (2000); these authors also explicitly allow for serial dependence. de Haan et al. (2015) consider a situation in which the smooth non-stationarity was formulated in a parametric way on the level of the domain of attraction condition rather than the limit situation. A nonparametric version of that model was investigated in Einmahl et al. (2016) (see below for details). Under the assumption of serial independence, these authors also provide asymptotic theory, which was recently extended in de Haan and Zhou (2021) to trends in the tail index and in Einmahl et al. (2022) to multivariate, spatial applications. Finally, change point tests for the tail index and the extremal dependence (i.e., abrupt changes in the tail behavior) can be found in Kojadinovic and Naveau (2017); Bücher et al. (2017); Hoga (2017, 2018), with the latter two references explicitly allowing for serially dependent observations.

The above literature review reveals a crucial gap which motivates the present paper: the models initiated by Einmahl et al. (2016) (subsequently referred to as EdHZ) have never been investigated under the assumption that the observations are serially dependent. Throughout the paper, we therefore work under the following model adapted from EdHZ: for sample size n and at time points $i \in \{1, \dots, n\}$, we observe possibly dependent random variables $X_1^{(n)}, \dots, X_n^{(n)}$ with continuous cumulative distribution functions (c.d.f.s) $F_{n,1}, \dots, F_{n,n}$, i.e., $X_i^{(n)} \sim F_{n,i}$. We assume that all these distribution functions share a common right endpoint $x^* = \sup\{x \in \mathbb{R} : F_{n,i}(x) < 1\}$, and that there exists some continuous reference c.d.f. F with the same right endpoint x^* that is strictly increasing on its support and some positive function c on $[0,1]$ such that

$$\lim_{x \uparrow x^*} \frac{1-F_{n,i}(x)}{1-F(x)} = c\Big(\frac{i}{n}\Big). \tag{1.1}$$

As in EdHZ, we refer to c as the *scedasis function*, which we additionally assume to be a bounded and continuous probability density function. The case where $c \equiv 1$ corresponds to *homogeneous extremes*, while the opposite is referred to as *heteroscedastic extremes*. The integrated scedasis function shall be denoted by

$$C(s) := \int_0^s c(x)\,\mathrm{d}x, \quad s \in [0,1].$$

We allow for serial dependence in the following sense: for each $n \in \mathbb{N}$, the unobservable sample $U_1^{(n)}, \dots, U_n^{(n)}$ with $U_i^{(n)} = F_{n,i}(X_i^{(n)})$ is assumed to be an excerpt from a strictly stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ whose distribution does not depend on n. The dynamics of the extremes of the latter series will later be captured by the concept of regular variation (Basrak and Segers, 2009), see Condition (B1) below for details, and by the extremal index θ (Leadbetter, 1983), see Condition (B8). Recall that the reciprocal of the extremal index may be interpreted as the mean cluster size of subsequent extreme observations.

Our contributions within the above model are as follows: first, we provide a (pointwise) central limit theorem on the kernel estimator for the scedasis function that was studied in EdHZ for the independent case. Notably, the serial dependence will only show up in the asymptotic estimation variance. Second, we study an empirical version of the integrated scedasis function from EdHZ and provide a respective functional central limit theorem; again, the asymptotic covariance functional will be different from that in the serially independent case. The latter is a major nuisance for testing the null hypothesis of homoscedastic extremes, i.e., $H_0 : c \equiv 1$, where standard approaches based on functionals of the law of the Brownian bridge as proposed in EdHZ do not work any more. As a circumvent, we develop a suitable multiplier bootstrap scheme and show its consistency; for this, we need to extent results from Drees (2015) and Section 12 in Kulik and Soulier (2020) to the non-stationary case. The bootstrap scheme is then used to define a classical bootstrap test as well as a test based on self-normalization, the latter being computationally much more efficient but slightly less powerful. Finally, we also propose an estimator for the extremal index θ of the underlying stationary time series that governs the dynamics of the extremes and show its consistency. For that purpose, we use a suitable modification of the block-maxima estimator from Northrop (2015); Berghaus and Bücher (2018) to the current non-stationary setting. On a theoretical level, a crucial tool for most of the afore-mentioned asymptotics is a functional central limit theorem for the sequential tail empirical process (STEP), which may also be of interest for other statistical problems not tackled in this paper.

The remaining parts of this paper are organized as follows: the assumptions needed to prove the asymptotic results are summarized and discussed in Section 2, where we also introduce a location-scale model meeting these assumptions. Section 3 is concerned with the estimation of the scedasis function and the integrated scedasis function. Section 4 is about testing for the null hypothesis that the extremes are homoscedastic. The assessment of the serial dependence is dealt with in Section 5, where we also extend the discussion on the location-scale model. A functional central limit theorem for the sequential tail empirical process under serial dependence is presented in Section 6. The finite-sample behavior of the introduced methods is investigated in a Monte Carlo simulation study in Section 7. The proofs for Section 6 are given in Section 8, with some auxiliary lemmas collected in Section 9. Finally, all other proofs are deferred to a supplementary material.

Throughout, all convergences are for $n \to \infty$ if not mentioned otherwise. Weak convergence is denoted by $\rightsquigarrow$. The left-continuous generalized inverse of some increasing function H is denoted by $H^{-1}(p) = \inf\{x \in \mathbb{R} : H(x) > p\}$. The sup-norm of some real-valued function f defined on some domain T is denoted by $\|f\|_\infty$.

2. Mathematical Preliminaries

Let $k = k_n$ be an increasing integer sequence satisfying $k \to \infty$ and $k = o(n)$ as $n \to \infty$; the STEP and our estimators for the scedasis function will be defined in terms of k, which essentially determines the threshold for declaring an observation as extreme. Let $L \in \mathbb{N}$ be some arbitrary but fixed constant (later determining, on which set the STEP will be defined; most often, we need $L = 1$ or $L = 2$). We impose the following set of assumptions:

(B0) **Basic assumptions.** The conditions on the model in Section 1 are met.

(B1) **Multivariate regular variation.** For each $n \in \mathbb{N}$, $U_1^{(n)}, \dots, U_n^{(n)}$ is an excerpt from a strictly stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ whose marginal stationary distribution is necessarily standard uniform on $(0, 1)$. The processes $(U_t^{(n)})_{t\in\mathbb{Z}}$ are all equal in law; denote a generic version by $(U_t)_{t\in\mathbb{Z}}$. The process $Z_t = 1/(1 - U_t)$ (note that Z_t is standard Pareto) is stationary and regularly varying, necessarily with index $\alpha = 1$ (Basrak and Segers, 2009).

(B2) **Regularity of c.** The function c is Hölder-continuous of order 1/2, that is, there exists $K_c > 0$ such that

$$|c(s) - c(s')| \le K_c|s - s'|^{1/2} \qquad \forall\ s, s' \in [0, 1].$$

(B3) **Blocking sequences and Beta-mixing.** There exist integer sequences $1 < \ell_n < r = r_n < n$, both converging to infinity as $n \to \infty$ and satisfying $\ell_n = o(r), r = o(\sqrt{k} \vee \frac{n}{k})$, such that the beta-mixing coefficients of $(U_t)_{t\in\mathbb{Z}}$ satisfy $\beta(n) = o(1)$ and $\frac{n}{r}\beta(\ell_n) = o(1)$.

(B4) **Moment bound on the number of extreme observations.** Let $c_\infty = c_\infty(L) = 1 + L\|c\|_\infty$, where $\|\cdot\|_\infty$ denotes the sup norm of a real-valued function. There exists $\delta > 0$ such that

$$\mathrm{E}\Big[\Big\{\sum_{s=1}^{r} \mathbf{1}(U_s > 1 - \tfrac{k}{n}c_\infty(L))\Big\}^{2+\delta}\Big] = O(r\tfrac{k}{n}).$$

(B5) **Moment bound on extreme increments.** There exists a non-decreasing, continuous function $h : [0, c_\infty(L)] \to [0, \infty)$, positive on $(0, c_\infty(L)]$ and with $h(0) = 0$, such that, for all sufficiently large n,

$$\mathrm{E}\Big[\Big\{\sum_{s=1}^{r} \mathbf{1}(1 - \tfrac{k}{n}x \ge U_s > 1 - \tfrac{k}{n}y)\Big\}^2\Big] \le r\frac{k}{n} \times h(y - x)$$

for all $0 \le x \le y \le c_\infty(L)$ with $c_\infty(L)$ from (B4).

(B6) **Second order condition.** There exists a positive, eventually decreasing function A with $\lim_{t\to\infty} A(t) = 0$ such that, as $x \uparrow x^*$,

$$\sup_{n\in\mathbb{N}} \max_{1\le i\le n} \left|\frac{1 - F_{n,i}(x)}{1 - F(x)} - c\Big(\frac{i}{n}\Big)\right| = O\Big(A\Big(\frac{1}{1 - F(x)}\Big)\Big).$$

Condition (B1) allows to control the serial dependence within the observed time series via *tail processes* (Basrak and Segers, 2009). More precisely, by Theorem 2.1 in Basrak and Segers (2009), regular variation of $(Z_t)_{t\in\mathbb{Z}}$ is equivalent to the fact that there exists a process $(Y_t)_{t\in\mathbb{N}_0}$ (the *tail process*) with Y_0 standard Pareto such that, for every $\ell \in \mathbb{N}$ and as $x \to \infty$,

$$\mathrm{P}(x^{-1}(Z_0, \dots, Z_\ell) \in \cdot \mid Z_0 > x) \rightsquigarrow \mathrm{P}((Y_0, \dots, Y_\ell) \in \cdot), \tag{2.1}$$

where, necessarily, $Y_j \ge 0$ for $j \ge 1$. Further, by Theorem 2 and its subsequent discussion in Segers (2003), Y_j is absolutely continuous on $(0, \infty)$ and may have an atom at 0.

Condition (B2) has also been imposed in Einmahl et al. (2016). Since $k = o(n)$, it implies that

$$\lim_{n\to\infty} \sup_{s\in[0,1]} \sqrt{k} \Big| \frac{1}{n} \sum_{i=1}^{\lfloor ns \rfloor} c(\tfrac{i}{n}) - C(s) \Big| = 0,$$

which will imply that there is no asymptotic bias in our main result below. The condition will however also be needed to prove (8.14) below.

The conditions in (B3), (B4), (B5) are essentially conditions imposed in Example 3.8 in Drees and Rootzén (2010) for deriving weak convergence of the standard non-sequential univariate tail empirical process under stationarity. Condition (B5) has mostly been shown with $h(z) = Kz$, for some $K > 0$, see, e.g., Drees (2000) for solutions of stochastic recurrence equations. Condition (B6) is a second-order condition on the speed of convergence in (1.1); it was also used in Einmahl et al. (2016). It is worth noting that Conditions (B4)-(B5) (and only these) depend on the constant $L \in \mathbb{N}$. The sequence ℓ_n in (B3) plays the role of a small-block length in a big-block-small-block technique, while $r - \ell_n$ is the length of a corresponding big block.

Example 2.1. Let us consider the following location-scale model, for which the above conditions can be shown to hold. Let

$$X_i^{(n)} = \sigma\big(\tfrac{i}{n}\big) W_i + \mu\big(\tfrac{i}{n}\big), \quad i = 1, \dots, n,$$

where $(W_t)_{t\in\mathbb{Z}}$ is a strictly stationary time series (see below for an explicit example) with c.d.f. F and where $\sigma : [0,1] \to (0,\infty)$ and $\mu : [0,1] \to \mathbb{R}$ are sufficiently smooth functions. Then, we obtain

$$F_{n,i}(x) = F\Big(\frac{x - \mu(\frac{i}{n})}{\sigma(\frac{i}{n})}\Big), \quad x \in \mathbb{R},$$

and $U_i^{(n)} = F_{n,i}(X_i^{(n)}) = F(W_i)$, $i = 1, \dots, n$, such that $U_1^{(n)}, \dots, U_n^{(n)}$ is an excerpt from a strictly stationary time series, with marginal distribution given by the uniform distribution on $[0, 1]$.

Next, as a special case, consider $(W_t)_{t\in\mathbb{Z}}$ to be a max-autoregressive process (ARMAX), defined by the recursion

$$W_t = \max\{\lambda W_{t-1}, (1-\lambda)V_t\}, \quad t \in \mathbb{Z}, \tag{2.2}$$

where $\lambda \in [0, 1)$ and $(V_t)_{t\in\mathbb{Z}}$ is an i.i.d. sequence of Fréchet(1)-distributed random variables. A stationary solution of the above recursion is given by $W_t = \max_{j\ge0}(1-\lambda)\lambda^j V_{t-j}$, such that the stationary solution is again Fréchet(1)-distributed, i.e.,

$F(x) = \exp(-1/x)$. Then, the scedasis function c can be easily calculated via

$$\lim_{x\to\infty} \frac{1-F_{n,i}(x)}{1-F(x)} = \lim_{x\to\infty} \frac{1-\exp\big(-\sigma(\frac{i}{n})/\{x-\mu(\frac{i}{n})\}\big)}{1-\exp(-1/x)} = \sigma\big(\tfrac{i}{n}\big),$$

yielding $c = \sigma$. We show that Conditions (B0)-(B6) are met. Condition (B0) and (B2) are obviously fulfilled, provided the scedasis function c is sufficiently regular. Condition (B1) can be seen to hold as follows. Since $(W_t)_{t\in\mathbb{Z}}$ is a moving maximum process, its tail process exists by Theorem 13.5.5 in Kulik and Soulier (2020), which implies that it is regularly varying by Theorem 2.1 in Basrak and Segers (2009). Then, $Z_t = 1/(1-U_t) = 1/\{1-F(W_t)\}$ is regularly varying with index $\alpha = 1$ according to Lemma 2.1 in Drees et al. (2015). By Berghaus and Bücher (2018), page 2322, $(W_t)_{t\in\mathbb{Z}}$, and hence also $(U_t)_{t\in\mathbb{Z}}$, is geometrically β-mixing, whence Condition (B3) holds. In that reference it is further shown that their Condition 2.1(ii) holds for $\delta = 1$, implying that our Condition (B4) also holds for $\delta = 1$ in view of the fact that, $rk = o(n)$ by Condition (B3). This also yields that $\mathrm{E}[|\sum_{s=1}^r \mathbf{1}(1-\frac{k}{n}x \ge U_s > 1-\frac{k}{n}y)|^3] \lesssim r\frac{k}{n}(y-x)$ for all $0 \le x \le y \le c_\infty(L)$, for n large enough (such that $rk/n \le 1$), which implies (B5). Finally, Condition (B6) can be seen to hold for $A(x) = x^{-1}$.

3. Estimation of the (integrated) scedasis function

In this section, we provide weak convergence results for estimators for the scedasis function c and its integrated version C; see also Einmahl et al. (2016) for related results in the serial independent case. Throughout, let $X_{n,1} \le \ldots \le X_{n,n}$ denote the order statistic of $X_1^{(n)}, \ldots, X_n^{(n)}$.

First, for the estimation of the scedasis function, we apply a kernel density estimator. Let K be a continuous and symmetric function on $[-1,1]$ with $K(x) = 0$ for $|x| > 1$ and $\int_{-1}^1 K(x)\,\mathrm{d}x = 1$. Let $h = h_n > 0$ denote a bandwidth paramater. Since we are also concerned with the estimation of c near the boundaries of the interval $[0,1]$, we make use of the boundary-corrected kernel K_b of K (see Jones, 1993): for $s \in [0,1]$, set

$$\tilde{c}_n(s) = \frac{1}{kh}\sum_{i=1}^n \mathbf{1}(X_i^{(n)} > X_{n,n-k})\; K_b\Big(\frac{s-i/n}{h}, s\Big),$$

where $k = k_n$ is from Condition (B3) and where K_b is defined as follows. First, for $p \in [0,1]$, let

$$a_j(p) = \int_{-1}^p x^j K(x)\,\mathrm{d}x, \quad b_j(p) = \int_{-p}^1 x^j K(x)\,\mathrm{d}x.$$

For $s \le h$, write $s = ph$ and let

$$K_b(x,s) = \frac{a_2(p) - a_1(p)x}{a_0(p)a_2(p) - a_1^2(p)} K(x), \quad x \in [-1,1],$$

and for $s \ge 1-h$, write $s = 1-ph$ and let

$$K_b(x,s) = \frac{b_2(p) - b_1(p)x}{b_0(p)b_2(p) - b_1^2(p)} K(x), \quad x \in [-1,1],$$

and for $s \in (h, 1-h)$, let $K_b(x,s) = K(x)$ for $x \in [-1,1]$. Note that K_b is depending on n, which we have suppressed from the notation.

To obtain asymptotic normality of the introduced estimator we additionally impose the following condition.

(B7) **Bandwidth.** The bandwidth sequence $h = h_n > 0$ satisfies $h \to 0$ and $kh \to \infty$. Further, $k^{1/5}h \to \lambda \geq 0$ and $r = o(\sqrt{kh})$ and $h \geq k^{-1/3}$.

The first three conditions in (B7) are standard bandwidth conditions that have also been imposed in Proposition 2 in Einmahl et al. (2016) to establish asymptotic normality of the scedasis estimator at point $s = 1$. The condition $r = o(\sqrt{kh})$ is slightly stronger than $r = o(\sqrt{k})$ from Condition (B3), which is used in Theorem 3.2 below to derive asymptotic normality of the estimator for the integrated scedasis function, where the rate of convergence is $\sqrt{k}$. Finally, the condition $h \geq k^{-1/3}$ is required for technical reasons in the proof (together with $r = o(n/k)$ from Condition (B3), it implies $kr^2 = o(n^2h^3)$, which will be used throughout the proofs); note that it is satisfied for the standard MSE optimal bandwidth choice of the order $k^{-1/5}$ (Tsybakov, 2009).

Theorem 3.1. *Suppose that Conditions (B0)-(B7) hold for $L = 2$ and let $c \in C^2([0,1])$. Let the function K be Lipschitz-continuous and symmetric on $[-1,1]$ with $K(x) = 0$ for $|x| > 1$ and $\int_{-1}^{1} K(x)\,\mathrm{d}x = 1$. If k satisfies $\sqrt{k}A(\frac{n}{2k}) \to 0$, then, for any $s \in [0,1]$ and as $n \to \infty$,*

$$\sqrt{kh}\{\tilde{c}_n(s) - c(s)\} \rightsquigarrow \mathcal{N}(\mu_s, \sigma_s^2),$$

where

$$\mu_s = \lambda^{5/2}\frac{c''(s)}{2}a(s), \quad \sigma_s^2 = c(s)\eta(s)\left\{d_0(1,1) + 2\sum_{h=1}^{\infty} d_h(1,1)\right\}$$

and where, recalling the tail process $(Y_t)_{t\in\mathbb{N}_0}$ associated with $(Z_t)_{t\in\mathbb{Z}}$ from (2.1),

$$d_h(x,x') = \mathrm{P}\Big(Y_0 > \frac{1}{x}, Y_h > \frac{1}{x'}\Big) \tag{3.1}$$

and $a(0) = \int_{-1}^{0} K_b(x,0)x^2\,\mathrm{d}x, \eta(0) = \int_{-1}^{0} K_b^2(x,0)\,\mathrm{d}x, a(1) = \int_0^1 K_b(x,1)x^2\,\mathrm{d}x, \eta(1) = \int_0^1 K_b^2(x,1)\,\mathrm{d}x$ and, for $s \in (0,1)$,

$$a(s) = \int_{-1}^{1} K(x)x^2\,\mathrm{d}x, \qquad \eta(s) = \int_{-1}^{1} K^2(x)\,\mathrm{d}x.$$

It is part of the assertion that the series defining σ_s^2 is convergent. The result may further be extended to cover the cases $s = s_n = ph$ and $s = s_n = 1 - ph$ for some $p \in (0,1]$; details are omitted for the sake of brevity.

Next, we analyze an estimator for the integrated scedasis function $C(s) = \int_0^s c(x)\,\mathrm{d}x$, that was also investigated in Einmahl et al. (2016). Define the estimator for C as

$$\hat{C}_n(s) = \frac{1}{k}\sum_{i=1}^{\lfloor ns\rfloor} \mathbf{1}\big(X_i^{(n)} > X_{n,n-k}\big), \quad s \in [0,1].$$

Theorem 3.2. *Suppose that Conditions (B0)-(B6) hold for $L = 1$ and that k satisfies $\sqrt{k}A(\frac{n}{k}) \to 0$. Then, as $n \to \infty$,*

$$\big\{\sqrt{k}\big(\hat{C}_n(s) - C(s)\big)\big\}_{s\in[0,1]} \rightsquigarrow \big\{\mathbb{S}(s,1) - C(s)\mathbb{S}(1,1)\big\}_{s\in[0,1]}$$

in $(\ell^\infty([0,1]), \|\cdot\|_\infty)$, where $\mathbb{S}$ denotes a tight, centered Gaussian process on $[0,1]^2$ with covariance given by

$$\mathfrak{c}((s,x),(s',x')) = C(s\wedge s')\Big\{d_0(x,x') + \sum_{h=1}^{\infty}\big(d_h(x,x') + d_h(x',x)\big)\Big\}, \tag{3.2}$$

where d_h is defined in (3.1). It is part of the assertion that the above series is convergent.

4. Testing for heteroscedastic extremes

In the following we construct tests that allow to detect whether the time series exhibits heteroscedasticity of extremes. Here, the extremes are homoscedastic (i.e., not heteroscedastic) if the scedasis function satisfies $c \equiv 1$ or if, equivalently, the integrated scedasis function satisfies $C(s) = s$ for all $s \in [0, 1]$. Thus, we test

$$H_0 : C(s) = s \quad \text{for all} \quad s \in [0, 1], \qquad H_1 : C(s) \neq s \quad \text{for some} \quad s \in [0, 1].$$

To this purpose, we pursue two approaches, where one is based on a bootstrap-procedure and the other uses a self-normalization technique. Let $\mathbb{C}_n(s) = \sqrt{k}\{\hat{C}_n(s) - s\}$, $s \in [0, 1]$, such that, by Theorem 3.2, under H_0 and as $n \to \infty$,

$$\big\{\mathbb{C}_n(s)\big\}_{s\in[0,1]} \rightsquigarrow \big\{\mathbb{S}(s, 1) - s\mathbb{S}(1, 1)\big\}_{s\in[0,1]}$$

in $(\ell^\infty([0, 1]), \|\cdot\|_\infty)$. Note that $\mathbb{S}(\cdot, 1)$ is a tight, centered Gaussian process on $[0, 1]$ satisfying $\operatorname{Cov}(\mathbb{S}(s, 1), \mathbb{S}(t, 1)) = (s \wedge t)\sigma^2$, $s, t \in [0, 1]$, where $\sigma^2 = d_0(1, 1) + 2\sum_{h=1}^{\infty} d_h(1, 1)$ and d_h is defined in Theorem 3.2, which implies that under H_0, as $n \to \infty$,

$$\mathbb{C}_n \rightsquigarrow \sigma\mathbb{B} \quad \text{in } (\ell^\infty([0, 1]), \|\cdot\|_\infty), \tag{4.1}$$

where $\mathbb{B}$ denotes a Brownian Bridge on $[0, 1]$.

For both approaches take the block length parameter r from Condition (B3) (which now becomes a hyperparameter of the statistical method; see Drees, 2015 and Kulik and Soulier, 2020 for a similar approach), set $m = \lfloor n/r \rfloor$ and let

$$I_j = \{(j - 1)r + 1, \dots, jr\}, \quad j = 1, \dots, m,$$

be the j-th block of size r.

We start with the bootstrap, more precisley, we use a multiplier block bootstrap. Let $B \in \mathbb{N}$ denote the number of boostrap repetitions and let $(\xi_1^{(b)}, \dots, \xi_m^{(b)})_{b=1,\dots,B}$ be i.i.d. and independent from $(X_i^{(n)})_i$, with $\mathrm{E}[\xi_j^{(b)}] = 0, \mathrm{E}[(\xi_j^{(b)})^2] = 1$ and $|\xi_j^{(b)}| \le M$ for some constant $M > 0$ for all $j = 1, \dots, m$ and $b = 1, \dots, B$ (for instance, $\xi_j^{(b)}$ is Rademacher distributed). Set

$$\mathbb{C}_{n,\xi}^{(b)}(s) = \mathbb{D}_{n,\xi}^{(b)}(s) - \hat{C}_n(s)\mathbb{D}_{n,\xi}^{(b)}(1),$$

where

$$\mathbb{D}_{n,\xi}^{(b)}(s) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} (\xi_j^{(b)} - \bar{\xi}^{(b)}) \sum_{i\in I_j} \mathbf{1}\Big(X_i^{(n)} > X_{n,n-k}\Big)\mathbf{1}(\tfrac{i}{n} \le s), \quad s \in [0, 1],$$

and $\bar{\xi}^{(b)} = m^{-1} \sum_{j=1}^{m} \xi_j^{(b)}$. Note that we may write

$$\mathbb{D}_{n,\xi}^{(b)}(s) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \xi_j^{(b)} \Big\{Y_{n,j}(s) - \frac{1}{m} \sum_{\ell=1}^{m} Y_{n,\ell}(s)\Big\}$$

with $Y_{n,j}(s) = \sum_{i\in I_j} \mathbf{1}\big(X_i^{(n)} > X_{n,n-k}\big)\mathbf{1}(\frac{i}{n} \le s)$, which is akin to the process considered in Formula (2.3) in Drees (2015).

Theorem 4.1. *Suppose that Conditions (B0)-(B6) hold for $L = 1$ and that k satisfies $\sqrt{k}A(\frac{n}{k}) \to 0$. Then, as $n \to \infty$,*

$$\left(\mathbb{C}_n, \mathbb{C}_{n,\xi}^{(1)}, \ldots, \mathbb{C}_{n,\xi}^{(B)}\right) \rightsquigarrow \left(\mathbb{C}, \mathbb{C}^{(1)}, \ldots, \mathbb{C}^{(B)}\right) \quad \text{in } (\ell^\infty([0,1]), \|\cdot\|_\infty)^{B+1},$$

where $\mathbb{C}(s) = \mathbb{S}(s,1) - C(s)\mathbb{S}(1,1)$ and $\mathbb{C}^{(1)}, \ldots, \mathbb{C}^{(B)}$ are independent copies of $\mathbb{C}$.

The previous theorem may alternatively be formulated as a conditional limit theorem, see Section 3.6 in van der Vaart and Wellner (1996) or Section 10 in Kosorok (2008) for details on that mode of convergence when applied to non-measurable stochastic processes. More precisely, by Lemma 3.11 in Bücher and Kojadinovic (2019), the weak convergence relation in the previous theorem is equivalent to the fact that $\sup_{h \in \mathrm{BL}_1(\ell^\infty([0,1]))} \left| \mathrm{E}\left[h(\mathbb{C}_{n,\xi}^{(1)}) \mid X_{n,1}, \ldots, X_{n,n}\right] - \mathrm{E}[h(\mathbb{C})]\right| = o_{\mathrm{P}}(1)$ and that $\mathbb{C}_{n,\xi}^{(1)}$ is asymptotically measurable, where $\mathrm{BL}_1(\ell^\infty([0,1])$ denotes the set of real valued Lipschitz functions on $\ell^\infty([0,1])$ with Lipschitz constant 1 that are bounded by 1. We prefer to work with the unconditional statement from Theorem 4.1, as it is more intuitive.

We propose to test for $H_0 : c \equiv 1$ based on the test statistics

$$S_{n,1} = \|\mathbb{C}_n\|_\infty, \quad T_{n,1} = \int_0^1 \mathbb{C}_n(s)^2 \,\mathrm{d}s.$$

In view of Theorem 4.1, the corresponding bootstrap quantities are given by

$$S_{n,1}^{(b)} = \|\mathbb{C}_{n,\xi}^{(b)}\|_\infty, \quad T_{n,1}^{(b)} = \int_0^1 \mathbb{C}_{n,\xi}^{(b)}(s)^2 \,\mathrm{d}s, \quad b = 1, \ldots, B.$$

For $\alpha \in (0,1)$, let $\hat{q}_{n,B,S}(1-\alpha)$ and $\hat{q}_{n,B,T}(1-\alpha)$ denote the empirical $(1-\alpha)$-quantile of $S_{n,1}^{(1)}, \ldots, S_{n,1}^{(B)}$ and $T_{n,1}^{(1)}, \ldots, T_{n,1}^{(B)}$, respectively. The test procedures are then defined as

$$\varphi_{n,B,S}(\alpha) = \mathbf{1}\big(S_{n,1} > \hat{q}_{n,B,S}(1-\alpha)\big), \quad \varphi_{n,B,T}(\alpha) = \mathbf{1}\big(T_{n,1} > \hat{q}_{n,B,T}(1-\alpha)\big).$$

Corollary 4.2. *Suppose that Conditions (B0)-(B6) hold for $L = 1$ and that k satisfies $\sqrt{k}A(\frac{n}{k}) \to 0$. Let $\alpha \in (0,1)$. Then, if $H_0 : c \equiv 1$ is met,*

$$\lim_{n,B\to\infty} \mathrm{P}(\varphi_{n,B,S}(\alpha) = 1) = \alpha, \quad \lim_{n,B\to\infty} \mathrm{P}(\varphi_{n,B,T}(\alpha) = 1) = \alpha.$$

Further, if $H_1 : c \not\equiv 1$ is met, then, for any $B \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathrm{P}(\varphi_{n,B,S}(\alpha) = 1) = 1, \quad \lim_{n\to\infty} \mathrm{P}(\varphi_{n,B,T}(\alpha) = 1) = 1.$$

Next, we introduce tests based on the concept of self-normalization. The basic idea is to consider the quotient of two statistics, such that the unknown variance factor σ in (4.1) cancels out. To do this, we take two of the bootstrap-quantities from Theorem 4.1, and define

$$S_{n,2} = \frac{\|\mathbb{C}_n\|_\infty}{\|\mathbb{C}_{n,\xi}^{(1)} - \mathbb{C}_{n,\xi}^{(2)}\|_\infty}, \quad T_{n,2} = \frac{\int_0^1 \mathbb{C}_n^2(s) \,\mathrm{d}s}{\int_0^1 \big(\mathbb{C}_{n,\xi}^{(1)}(s) - \mathbb{C}_{n,\xi}^{(2)}(s)\big)^2 \,\mathrm{d}s}.$$

By Theorem 4.1 we know that under H_0, as $n \to \infty$,

$$S_{n,2} \rightsquigarrow S_2 := \frac{\|\mathbb{B}\|_\infty}{\|\mathbb{B}^{(1)} - \mathbb{B}^{(2)}\|_\infty}, \quad T_{n,2} \rightsquigarrow T_2 := \frac{\int_0^1 \mathbb{B}(s)^2 \,\mathrm{d}s}{\int_0^1 \big(\mathbb{B}^{(1)}(s) - \mathbb{B}^{(2)}(s)\big)^2 \,\mathrm{d}s},$$

where $\mathbb{B}$, $\mathbb{B}^{(1)}$ and $\mathbb{B}^{(2)}$ are independent Brownian Bridges on $[0,1]$. For $\alpha \in (0,1)$, let $q_S(1-\alpha)$ and $q_T(1-\alpha)$ be the $(1-\alpha)$-quantile of S_2 and T_2, respectively. The corresponding test procedures are given by

$$\varphi_{n,S}(\alpha) = \mathbf{1}\big(S_{n,2} > q_S(1-\alpha)\big), \quad \varphi_{n,T}(\alpha) = \mathbf{1}\big(T_{n,2} > q_T(1-\alpha)\big).$$

Corollary 4.3. *Suppose that Conditions (B0)-(B6) hold for $L = 1$ and that k satisfies $\sqrt{k}A(\frac{n}{k}) \to 0$. Let $\alpha \in (0,1)$. Then, if $H_0 : c \equiv 1$ is met,*

$$\lim_{n\to\infty} \mathrm{P}(\varphi_{n,S}(\alpha) = 1) = \alpha, \quad \lim_{n\to\infty} \mathrm{P}(\varphi_{n,T}(\alpha) = 1) = \alpha$$

Further, if $H_1 : c \not\equiv 1$ is met, then

$$\lim_{n\to\infty} \mathrm{P}_{H_1}(\varphi_{n,S}(\alpha) = 1) = 1, \quad \lim_{n\to\infty} \mathrm{P}(\varphi_{n,T}(\alpha) = 1) = 1.$$

5. Assessing the serial dependence

Within our basic model described in the introduction, the dynamics of the time series extremes are governed by the stationary time series $(Z_t)_{t\in\mathbb{Z}}$ from Condition (B1). There are many interesting statistical problems related to those dynamics which are worth to be investigated like, e.g., estimating the distribution of the tail process (see Davis et al., 2018 for stationary observations) or estimation of general cluster functionals (see Section 10 in Kulik and Soulier, 2020 for stationary observations). Throughout, we restrict attention to estimating the extremal index θ, which may be regarded as the most traditional parameter associated with the serial dependence.

Recall that the extremal index $\theta \in (0,1]$ of $(Z_t)_t$ exists iff the same is true for $(U_t)_t$ (in that case, the indices are equal), and that the latter requires that, for any $\tau > 0$, there exists a sequence $(u_n(\tau))_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} n\{1-u_n(\tau)\} = \tau$ and

$$\lim_{n\to\infty} \mathrm{P}\big(\max_{1\le i\le n} U_i \le u_n(\tau)\big) = e^{-\theta\tau}. \tag{5.1}$$

One can further show that, if the extremal index exists, then (5.1) holds for any sequence $(u_n(\tau))_n$ with $\lim_{n\to\infty} n\{1-u_n(\tau)\} = \tau$. Subsequently, we choose $u_n(\tau) = 1-\tau/n$.

For estimating θ, we divide the finite stretch of observations $X_1^{(n)},\dots,X_n^{(n)}$ into non-overlapping successive blocks of size $q = q_n$, i.e., into blocks

$$I_j' = \{(j-1)q+1,\dots,jq\}, \quad j = 1,\dots,k',$$

where $k' = \lfloor n/q \rfloor$. For $j = 1,\dots,k'$, set

$$Z_{n,j} = q\big\{1 - \max_{i\in I_j'} F(X_i^{(n)})\big\}, \quad \hat{Z}_{n,j} = q\big\{1 - \max_{i\in I_j'} \hat{F}_n(X_i^{(n)})\big\}, \tag{5.2}$$

where $\hat{F}_n(x) = n^{-1}\sum_{i=1}^n \mathbf{1}(X_i^{(n)} \le x)$ denotes the empirical c.d.f. of $X_1^{(n)},\dots,X_n^{(n)}$. Note that, in view of (1.1), for sufficiently large $x \in \mathbb{R}$,

$$\mathrm{E}[1-\hat{F}_n(x)] = \frac{1}{n}\sum_{i=1}^n \{1-F_{n,i}(x)\} = \{1-F(x)\}\frac{1}{n}\sum_{i=1}^n \{c(i/n)+o(1)\} \approx 1-F(x)$$

(ignoring the possible non-uniformity in (1.1) for the moment), whence $\hat{Z}_{n,j}$ can be regarded as an observable counterpart of $Z_{n,j}$.

In the following, we will show that $Z_{n,1+\lfloor\xi k'\rfloor}$, $\xi \in [0,1)$, asymptotically follows an exponential distribution with parameter depending on θ, this result being the basis for our estimation procedure for θ, see Lemma 5.2. To prove this, we impose the subsequent conditions.

(B8) **Extremal Index.** The stationary time series $(U_t)_{t\in\mathbb{Z}}$ from Condition (B1) is assumed to have an extremal index $\theta \in (0,1]$.

(B9) **Blocking sequences and mixing.** The blocksize q is chosen in such a way that it satisfies $q = o(\sqrt{n})$ and $n\beta(q) = o(q)$ as $n \to \infty$.

(B10) **Uniform integrability.** For some $\delta_1 > 0$,

$$\limsup_{n\to\infty} \sup_{\xi\in(0,1)} \mathrm{E}\left[\left|Z_{n,1+\lfloor \xi k'\rfloor}\right|^{2+\delta_1}\right] < \infty.$$

Condition (B10) is imposed to deduce uniform integrability of the $Z^2_{n,1+\lfloor \xi k'\rfloor}$; it will imply convergence of the corresponding first and second moments.

Remark 5.1. We exemplarily show that the above conditions hold for the location-scale model from Example 2.1 with $(W_t)_{t\in\mathbb{Z}}$ chosen as the max-autoregressive process defined in (2.2). First, the process $(W_t)_{t\in\mathbb{Z}}$ has an extremal index θ given by $\theta = 1-\lambda$ (Beirlant et al., 2004, Chapter 10), such that $(U_t)_{t\in\mathbb{Z}}$ also has extremal index θ and Condition (B8) holds. Further, by Berghaus and Bücher (2018), page 2322, $(W_t)_{t\in\mathbb{Z}}$, and hence also $(U_t)_{t\in\mathbb{Z}}$, is geometrically β-mixing, whence Condition (B9) is fulfilled for appropriate choice of q. Regarding Condition (B10), we have, for $j \in \mathbb{N}$,

$$\begin{aligned} Z_{n,j} &= \bar{Z}_{n,j} \frac{1 - \max_{i\in I'_j} F(X_i^{(n)})}{1 - \max_{i\in I'_j} F_{n,i}(X_i^{(n)})} \\ &\le \bar{Z}_{n,j} \frac{1 - \exp\big(-(c_{\min}\max_{i\in I'_j} W_i + \inf_{s\in[0,1]}\mu(s))^{-1}\big)}{1-\exp\big(-(\max_{i\in I'_j} W_i)^{-1}\big)}, \end{aligned} \tag{5.3}$$

where $\bar{Z}_{n,j} = q\{1 - \max_{i\in I'_j} U_i\}$. Note that the distribution of the right-hand side in the last display is independent of $j \in \mathbb{N}$. By induction, $\mathrm{P}\big(\max_{i=1,\dots,b} W_i \le x\big) = F(x)^{1+\theta(b-1)} = \exp(-\{1+\theta(b-1)\}/x)$ for $x > 0, b \in \mathbb{N}$, such that $\max_{i=1,\dots,q} W_i$ converges to ∞ in probability. Therefore, any absolute moment of the second factor of the right-hand side in (5.3) converges. Further, it is shown in Example 6.1 in Berghaus and Bücher (2017), see the proof of their Condition 2.1(vi) holds, that $\limsup_{n\to\infty} \mathrm{E}\left[\bar{Z}_{n,1}^{\delta'}\right] < \infty$ for any $\delta' > 0$. Along with inequality (5.3), Hölder's inequality implies that Condition (B10) holds.

Lemma 5.2. *Fix $\xi \in [0,1)$. Suppose that Conditions (B0)-(B2), (B6) and (B8)-(B9) hold. Then, $Z_{n,1+\lfloor \xi k'\rfloor} \rightsquigarrow \mathrm{Exp}(\theta c(\xi))$ as $n \to \infty$.*

This result motivates estimators based on the method of moments, see Northrop (2015); Berghaus and Bücher (2018) for the stationary case. Consider the (unobservable) random variable

$$T_n = \frac{1}{k'}\sum_{j=1}^{k'} Z_{n,j}.$$

Then, for $\varphi_n : [0,1] \to \mathbb{R}, \varphi_n(\xi) = \sum_{j=1}^{k'} \mathrm{E}\left[Z_{n,j}\right]\mathbf{1}\big(\xi \in \big[\frac{j-1}{k'}, \frac{j}{k'}\big)\big)$, we obtain

$$\mathrm{E}[T_n] = \frac{1}{k'}\sum_{j=1}^{k'} \mathrm{E}\left[Z_{n,j}\right] = \int_0^1 \varphi_n(\xi)\,\mathrm{d}\xi.$$

By Condition (B10) and Lemma 5.2, for any fixed $\xi \in [0,1)$, we have $\varphi_n(\xi) = \mathrm{E}\left[Z_{n,1+\lfloor \xi k'\rfloor}\right] \to \mathrm{E}[V_\xi], n \to \infty$, where $V_\xi \sim \mathsf{Exp}(\theta c(\xi))$. Since $\sup_{n\in\mathbb{N}} \|\varphi_n\|_\infty < \infty$

by Condition (B10), the dominated convergence theorem implies

$$\mathrm{E}\left[T_n\right] = \int_0^1 \varphi_n(\xi)\,\mathrm{d}\xi \to \int_0^1 \mathrm{E}\left[V_\xi\right]\,\mathrm{d}\xi = \frac{1}{\theta}\int_0^1 \frac{1}{c(\xi)}\,\mathrm{d}\xi.$$

Recall that the function c is positive and continuous on $[0,1]$; thus there is a positive number $c_{\min}$ such that $c(s) > c_{\min}$ for all $s \in [0,1]$. Therefore, it is advisable to also truncate $\tilde{c}_n$ from below, say by considering $\hat{c}_n = \max(\tilde{c}_n, \kappa)$ with some small, positive constant $\kappa > 0$. Subsequently, we assume that $0 < \kappa < c_{\min}$. Now, let us estimate $\tau = \int_0^1 c(\xi)^{-1}\,\mathrm{d}\xi$ by $\hat{\tau}_n = \int_0^1 \hat{c}_n(\xi)^{-1}\,\mathrm{d}\xi$. Since $\mathrm{E}[T_n] \to \theta^{-1}\tau$, a sensible, observable method of moments estimator for θ is given by

$$\hat{\theta}_n = \hat{T}_n^{-1}\hat{\tau}_n, \quad \text{where} \quad \hat{T}_n = \frac{1}{k'}\sum_{j=1}^{k'} \hat{Z}_{n,j}.$$

The subsequent theorem yields concistency of this estimator; its finite-sample properties are studied in Section 7.

Theorem 5.3. *Suppose that Conditions (B0)-(B2) hold. Assume $c \in C^2([0,1])$ and let the function K in the definition of $\hat{c}_n$ be Lipschitz-continuous.*

- *(a) If additionally Conditions (B3)-(B7) hold for $L = 2$ and if k satisfies $\sqrt{k}A(\frac{n}{2k}) \to 0$, then $\hat{\tau}_n = \tau + o_{\mathrm{P}}(1)$ as $n \to \infty$.*
- *(b) If additionally Conditions (B3)-(B6) hold for $k = k'$ and for all $L \in \mathbb{N}$, and if k' satisfies $\sqrt{k'}A(\frac{n}{Lk'}) \to 0$ for all $L \in \mathbb{N}$, and if Conditions (B8)-(B10) hold, then $\hat{T}_n = \theta^{-1}\tau + o_{\mathrm{P}}(1)$ as $n \to \infty$.*

In particular, if all of the above conditions are met, then $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ as $n \to \infty$.

6. Weak convergence of the (simple) STEP

Functional weak convergence of the subsequent processes will be essential for proving the asymptotic results in the previous sections. Precisely, we are interested in the simple sequential tail empirical process (simple STEP) $\mathbb{S}_n$ and the sequential tail empirical process (STEP) $\mathbb{F}_n$ defined as

$$\mathbb{S}_n(s,x) = \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]} \mathbf{1}\{U_i^{(n)} > 1 - \tfrac{k}{n}c(\tfrac{i}{n})x\} - xC(s)\right\}, \tag{6.1}$$

$$\mathbb{F}_n(s,x) = \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]} \mathbf{1}\left\{X_i^{(n)} > V\left(\frac{n}{kx}\right)\right\} - xC(s)\right\}, \tag{6.2}$$

where $(s,x) \in [0,1] \times [0,\infty)$ and where $V = (\frac{1}{1-F})^{-1}$.

Proposition 6.1. *Suppose that Conditions (B0)–(B3) hold. Fix some constant $L \in \mathbb{N}$ and suppose that Conditions (B4) and (B5) hold for L. Then, as $n \to \infty$,*

$$\mathbb{S}_n \rightsquigarrow \mathbb{S} \quad in \quad (\ell^\infty([0,1]\times[0,L]), \|\cdot\|_\infty),$$

where $\mathbb{S}$ denotes a tight, centered Gaussian process on $[0,1]\times[0,L]$ with covariance $\mathfrak{c}((s,x),(s',x'))$ as defined in (3.2).

Proposition 6.2. *Suppose that Conditions (B0)–(B3) and (B6) hold. Fix some constant $L \in \mathbb{N}$ and suppose that Conditions (B4) and (B5) hold for L. If k satisfies*

$\sqrt{k}A(\frac{n}{Lk}) \to 0$ *as* $n \to \infty$, *then*

$$\sup_{(s,x)\in[0,1]\times[0,L]} |\mathbb{F}_n(s,x) - \mathbb{S}_n(s,x)| = o_{\mathrm{P}}(1).$$

As a consequence, $\mathbb{F}_n \rightsquigarrow \mathbb{S}$ *in* $(\ell^\infty([0,1]\times[0,L]), \|\cdot\|_\infty)$.

7. Finite-Sample Results

A simulation study is carried out to analyze the finite-sample performance of the introduced methods. Results are presented for scaled versions of two common time series models. Define the following functions, later resulting in different scedasis functions.

(i) $c_{1,\beta}(s) = \beta + 2(1-\beta)s$,

(ii) $c_{2,\beta}(s) = (\beta + 4(1-\beta)s)\mathbf{1}(s \in [0, 0.5]) + (4 - 3\beta - 4(1-\beta)s)\mathbf{1}(s \in (0.5, 1])$.

Note that $c_{1,\beta}$ is a straight line connecting the points $(0,\beta)$ and $(1, 2-\beta)$, while $c_{2,\beta}$ is a polygonal chain with vertices $(0,\beta)$ $(1/2, 2-\beta)$ and $(1,\beta)$.

We consider the following scale models.

- The ARMAX-model: Let $(W_t)_t$ be an ARMAX-process as specified in (2.2). We consider $\lambda \in \{0, 0.25\}$; note that $\lambda = 0$ corresponds to the i.i.d. case. Denote the c.d.f. of W_t by F, which is the c.d.f. of the standard Fréchet-distribution. For $j \in \{1,2\}$ and $i \in \{1,\dots,n\}$, let
$$X_i^{(n)} = c_{j,\beta}(\tfrac{i}{n})W_i.$$
By Example 2.1, the scedasis function c is equal to $c_{j,\beta}$. Further, for $j \in \{1,2\}$, consider
$$X_i^{(n)} = \tilde{c}_{j,\beta}(\tfrac{i}{n}, W_i)W_i := \big\{\mathbf{1}(W_i < p) + c_{j,\beta}(\tfrac{i}{n})\ \mathbf{1}(W_i \ge p)\big\}W_i,$$
where p is the 80%-quantile of F. In this model, the scale transformation introduced by $c_{j,\beta}$ only effects the observations exceeding the large threshold p. One can easily see that the scedasis function c is equal to $c_{j,\beta}$.
- The ARCH-model: Let $(W_t)_t$ be an ARCH-process, i.e.,
$$W_t = (2\times 10^{-5} + \lambda W_{t-1}^2)^{1/2} V_t, \quad t \in \mathbb{Z},$$
where $\lambda \in (0,1)$ and $(V_t)_{t\in\mathbb{Z}}$ is an i.i.d. sequence of $\mathcal{N}(0,1)$-distributed random variables. We consider $\lambda = 0.7$. By Theorem 1.1 in de Haan et al. (1989) the c.d.f. F of W_t satisfies $1 - F(x) \sim dx^{-\kappa'}$ as $x \to \infty$ for some constant $d > 0$, with κ' (approximately) given by $\kappa' = \kappa'(\lambda) = 1.586$; see Table 3.2 in that reference. For $j \in \{1,2\}$ and $i \in \{1,\dots,n\}$, let
$$X_i^{(n)} = c_{j,\beta}(\tfrac{i}{n})^{1/\kappa'}\ W_i.$$
The scedasis function c is equal to $c_{j,\beta}$. Further, similar as for the ARMAX-model, consider
$$X_i^{(n)} = \tilde{c}_{j,\beta}(\tfrac{i}{n}, W_i)W_i := \big\{\mathbf{1}(W_i < p) + c_{j,\beta}(\tfrac{i}{n})^{1/\kappa'}\ \mathbf{1}(W_i \ge p)\big\}W_i$$
for $j \in \{1,2\}$, where p is the 80%-quantile of F. A straightforward calculation shows that the scedasis function c is equal to $c_{j,\beta}$ as well.

Note that the ARMAX model with $\lambda = 0$ corresponds to the case that the observations are independent. We call this case simply the independent model.

In the subsequent simulation study, the parameter β of the scedasis functions, is set to $\beta = 1, 0.75, 0.5, 0.25$. In each case, the sample size is fixed to $n = 2000$ and

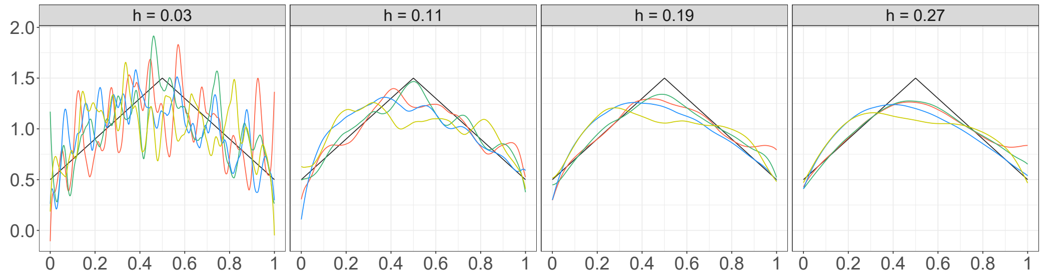

FIGURE 1. The scedasis function $c_{2,\beta}$ with $\beta = 0.5$ (black line) and the estimator $\tilde{c}_n$ evaluated at four exemplary time series generated from the ARCH-model.

the performance of the statistical methods is assessed based on $N = 1000$ simulation runs each if not mentioned otherwise.

7.1. **Estimation of the scedasis function.** We start by briefly considering the behavior of the kernel estimator for the scedasis function. For the sake of brevity, we restrict the presentation to the ARCH-model with scedasis function $c_{2,\beta}$ with $\beta = 0.5$; the behavior within the other models was found to be very similar. In Figure 1, we depict the estimator $\tilde{c}_n$ for four exemplary time series, where we use the biweight kernel K

$$K(x) = \frac{15}{16}(1 - x^2)^2, \quad x \in [-1, 1], \tag{7.1}$$

$k = 400$ and consider bandwidths $h \in \{0.03, 0.11, 0.19, 0.27\}$. We observe typical over-fitting (under-smoothing) for small values of h and under-fitting (over-smoothing) for large values of h. Note in particular that the estimator no longer captures the peak of $c_{2,\beta}(s)$ at $s = 0.5$ for $h = 0.27$. Visual inspection suggests that reasonably good choices for the bandwidth lie in the interval $[0.1, 0.2]$; an observation that was confirmed in simulations regarding the other models described in the previous section.

7.2. **Testing for heteroscedastic extremes.** We next study the performance of the introduced test procedures. Recall that both the tests based on the multiplier block bootstrap and the ones relying on the method of self-normalization depend on a multiplier sequence $(\xi_i)_i$, for which we choose an i.i.d. Rademacher sequence. The following results are based on $B = 300$ bootstrap replicates. We consider block sizes $q \in \{4, 8\}$ and number of exceedances $k \in \{100, 200\}$, which corresponds to 5% or 10% of the total observations, respectively. The test level is set to $\alpha = 0.05$.

Since the Cramér-von-Mises-type test statistics (i.e., $\varphi_{n,B,T}$ and $\varphi_{n,T}$) were found to be superior to the Kolmogorov-Smirnov-type test statistics (i.e., $\varphi_{n,B,S}$ and $\varphi_{n,S}$), we only present results for the former. Here, we refer to $\varphi_{n,B,T}$ simply as the bootstrap, and to $\varphi_{n,T}$ as the self-normalization. All rejection percentages are presented in Table 1.

We start by discussing the behavior of the tests under $H_0 : C(s) = s$ for all $s \in [0, 1]$; note that $\beta = 1$ represents being under H_0 for all data generating processes under consideration. We also present results for the Cramér-von-Mises-type test from Einmahl et al. (2016), which was designed for the case of independent data and is here denoted by EdHZ. One can see that our tests hold their level and, as

Model	β	Bootstrap with $(k,r)=$ (100, 4)	(100, 8)	(200, 4)	(200, 8)	Self-Normalization with $(k,r)=$ (100, 4)	(100, 8)	(200, 4)	(200, 8)	EdHZ with $k=$ 100	200
Panel (A): Scale model $c_{1,\beta}$											
Indep.	1.0	3.9	2.2	1.8	0.6	4.2	2.3	2.3	1.5	4.7	4.1
	0.75	23.3	17.9	34.6	22.5	16.5	12.7	21.8	15.6	26.9	46.2
	0.5	78.2	71.5	95.8	91.2	51.2	44.0	70.0	60.7	81.9	98.2
	0.25	98.9	98.4	100.0	99.9	84.5	76.5	94.7	89.4	99.0	100.0
ARMAX	1.0	6.6	4.5	4.4	2.7	5.5	4.0	3.4	3.1	14.8	12.1
	0.75	21.4	17.3	30.7	20.8	14.6	12.2	18.7	14.3	35.4	50.7
	0.5	65.6	59.7	88.3	79.2	42.5	36.7	59.3	48.3	77.4	96.0
	0.25	94.7	91.8	99.6	99.4	71.5	66.5	90.6	82.4	97.6	100.0
ARCH	1.0	7.9	5.0	4.4	1.7	5.7	4.5	3.6	3.0	16.1	11.1
	0.75	48.0	38.4	51.0	37.1	30.7	24.4	28.8	25.3	61.6	66.2
	0.5	94.7	92.0	98.5	96.5	73.6	67.6	81.2	69.9	98.0	99.6
	0.25	99.9	99.8	100.0	100.0	94.3	91.2	96.8	93.2	100.0	100.0
Panel (B): Scale model $c_{2,\beta}$											
Indep.	1.0	3.9	2.2	1.8	0.6	4.2	2.3	2.3	1.5	4.7	4.1
	0.75	7.3	4.3	5.6	2.2	6.7	4.6	5.6	3.2	5.8	7.4
	0.5	29.4	19.6	52.0	32.3	17.0	11.2	25.6	17.3	20.1	55.3
	0.25	78.6	68.6	98.2	92.6	42.1	33.6	57.9	48.4	68.0	98.8
ARMAX	1.0	6.6	4.5	4.4	2.7	5.5	4.0	3.4	3.1	14.8	12.1
	0.75	9.6	6.8	9.4	5.7	7.7	6.2	7.4	5.2	17.6	21.9
	0.5	27.5	18.3	41.2	26.4	16.3	12.0	20.2	16.2	36.9	64.8
	0.25	62.8	53.7	90.6	79.6	32.9	26.3	50.7	40.5	74.3	97.0
ARCH	1.0	7.9	5.0	4.4	1.7	5.7	4.5	3.6	3.0	16.1	11.1
	0.75	19.8	13.0	14.2	6.4	13.7	8.1	10.7	5.6	27.6	26.6
	0.5	66.5	53.5	73.7	52.7	35.1	25.6	37.9	24.7	73.1	86.6
	0.25	96.2	92.7	99.5	97.5	65.1	56.4	69.3	60.0	98.5	99.9
Panel (C): Scale model $\tilde{c}_{1,\beta}$											
Indep.	1.0	3.9	2.2	1.8	0.6	4.2	2.3	2.3	1.5	4.7	4.1
	0.75	24.3	18.4	34.8	22.8	15.8	12.5	24.2	14.4	26.9	46.2
	0.5	78.2	71.0	96.0	91.4	49.2	46.2	71.0	58.4	81.9	98.2
	0.25	99.3	98.6	100.0	100.0	80.1	75.9	94.0	89.1	99.0	100.0
ARMAX	1.0	6.6	4.5	4.4	2.7	5.5	4.0	3.4	3.1	14.8	12.1
	0.75	21.9	17.7	30.3	20.6	16.0	12.7	17.9	12.2	35.4	50.7
	0.5	65.8	57.9	88.4	80.1	41.6	36.0	61.2	50.1	77.4	96.0
	0.25	94.8	91.7	99.6	99.5	71.5	61.2	88.4	83.3	97.6	100.0
ARCH	1.0	7.9	5.0	4.4	1.7	5.7	4.5	3.6	3.0	16.1	11.1
	0.75	46.5	37.9	52.0	36.2	31.4	24.4	29.4	21.7	61.6	66.2
	0.5	94.6	92.0	98.7	96.4	76.9	66.1	79.7	68.8	98.0	99.6
	0.25	99.9	100.0	100.0	100.0	93.9	89.8	97.6	94.2	100.0	100.0
Panel (D): Scale model $\tilde{c}_{2,\beta}$											
Indep.	1.0	3.9	2.2	1.8	0.6	4.2	2.3	2.3	1.5	4.7	4.1
	0.75	6.9	4.0	6.4	2.0	5.8	5.2	5.9	3.3	5.8	7.4
	0.5	28.8	20.1	51.6	33.0	17.6	10.6	24.9	17.9	20.1	55.3
	0.25	78.4	69.0	98.5	92.2	39.3	32.5	58.6	46.2	68.0	98.8
ARMAX	1.0	6.6	4.5	4.4	2.7	5.5	4.0	3.4	3.1	14.8	12.1
	0.75	10.5	7.0	10.1	5.3	7.1	6.4	7.1	4.4	17.6	21.9
	0.5	25.8	19.1	42.0	25.9	17.3	10.6	23.7	14.1	36.9	64.8
	0.25	63.6	53.3	90.8	79.5	33.9	26.2	48.7	40.8	74.3	97.0
ARCH	1.0	7.9	5.0	4.4	1.7	5.7	4.5	3.6	3.0	16.1	11.1
	0.75	19.8	12.3	14.8	6.8	13.9	7.2	11.3	6.9	27.6	26.6
	0.5	65.7	53.4	73.3	53.8	36.4	26.5	36.8	27.4	73.1	86.6
	0.25	96.3	93.1	99.6	97.8	64.5	54.9	70.8	59.7	98.5	99.9

TABLE 1. Empirical rejection percentage of the test procedures.

expected, that the test from Einmahl et al. (2016) holds its level in the independent model, but fails to do so in the other dependent models.

Next, we consider the performance under the alternatives. One can see that the power of the tests increases with decreasing β, which is to be expected since a

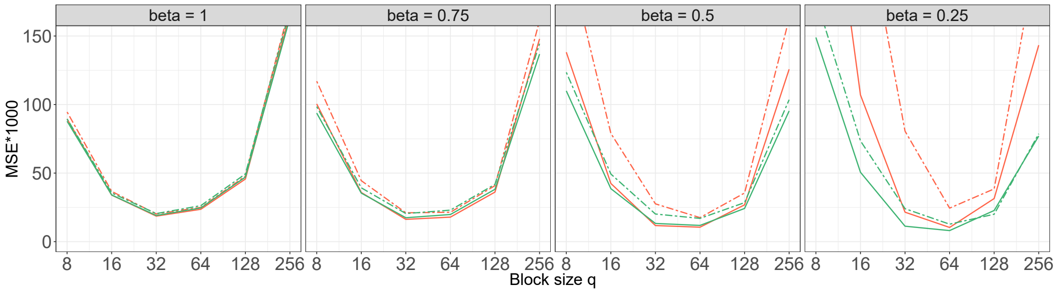

FIGURE 2. Mean squared error, multiplied by 10^3, for the estimation of θ in the ARCH-model with scedasis function $c_{2,\beta}$ for $\hat\theta_{n1}$ (orange lines) and $\hat\theta_{n2}$ (green lines) with $k = 400$ (solid lines) and $k = 300$ (dotted lines).

decrease in β results in a stronger deviation of $c_{j,\beta}$ from the null hypothesis that the scedasis function equals one. In general, the power of the bootstrap-test is uniformly higher than the power of the test based on self-normalization, but both exhibit high power for $\beta = 0.25$. Recall again that the self-normalization test only requires evaluation of $\mathbb{C}_{n,\xi}^{(b)}$ for $b \in \{1, 2\}$, while the expression must be evaluated a large number of times for the bootstrap test (we choose $B = 300$). With regard to the choice of k and r the highest power is usually attained for $k = 200$ and $r = 4$.

7.3. **Estimation of the extremal index.** We finally briefly evaluate the performance of the estimator for the extremal index. For comparison, we also introduce a second estimator for θ based on the method of moments, which may also be motivated by Lemma 5.2: under the notation of Section 5, consider the (unobservable) random variable

$$T_{n2} = \frac{1}{k'} \sum_{j=1}^{k'} Z_{n,j} \; c\big(\tfrac{j}{k'}\big).$$

Note that $\mathrm{E}\left[Z_{n,1+\lfloor \xi k' \rfloor}\right] c\big(\frac{1+\lfloor \xi k' \rfloor}{k'}\big) \to \mathrm{E}[V_\xi] c(\xi) = \frac{1}{\theta}$, where $V_\xi \sim \mathrm{Exp}(\theta c(\xi))$, by continuity of c, Condition (B10) and Lemma 5.2. Then, as in Section 5, it follows that

$$\mathrm{E}[T_{n2}] = \frac{1}{k'} \sum_{j=1}^{k'} \mathrm{E}\left[Z_{n,j} c\big(\tfrac{j}{k'}\big)\right] \to \int_0^1 \frac{1}{\theta}\, \mathrm{d}\xi = \frac{1}{\theta}.$$

Therefore, another sensible method of moments estimators for θ is given by

$$\hat\theta_{n2} = \left\{ \frac{1}{k'} \sum_{s=1}^{k'} \hat Z_{n,s} \; \hat c_n\Big(\frac{s}{k'}\Big) \right\}^{-1}.$$

We only present results for the ARCH-model; the ARMAX- and independent model were found to yield very similar results. Note that for $\lambda = 0.7$ in the ARCH-model we have $\theta = 0.721$, see Table 3.2 in de Haan et al. (1989).

In what follows, the block size q is chosen from the set $\{8, 16, 32, 64, 128, 256\}$ (recall that $k' = \lfloor n/q \rfloor$) and the number of exceedances $k \in \{300, 400\}$ are considered. (Here, slightly larger values of k turned out to work better than in the context of testing for heteroscedastic extremes above.) Regarding the kernel density estimator, we set $\kappa = 0.1$, set the bandwidth to $h = 0.2$ and use the biweight kernel from (7.1).

In Figure 2, the mean squared error (MSE) of $\hat\theta_{n1}$ and $\hat\theta_{n2}$ is plotted as a function of the block size q, where the true scedasis function is given by $c_{2,\beta}$ for different values of β. One can see that the MSE-curves are mostly U-shaped, and that a minimum value is reached at an intermediate blocksize of $q \in \{32, 64\}$. Further, in most scenarios the alternative estimator $\hat\theta_{n2}$ outperforms the estimator $\hat\theta_{n1}$, and the larger number of exceedances $k = 400$ seems to work better than $k = 300$ in terms of minimal MSE-values. The same observations were found for the other scedasis functions $c_{1,\beta}$, $\tilde c_{1,\beta}$ and $\tilde c_{2,\beta}$.

8. Proofs

For space considerations, we only present the proofs for the theoretical results from Section 6, which are central to all other proofs. The remaining proofs for Sections 3-5 are collected in a supplementary material.

Proof of Proposition 6.1. Recall that $c_\infty(L) = 1 + L\|c\|_\infty$. For $i \in \{1, \dots, n\}$ and $n \in \mathbb{N}$, define

$$X'_{n,i} = \left(\frac{U_i^{(n)} - (1 - \frac{k}{n}c_\infty(L))}{\frac{k}{n}}\right)_+ = \max\left(\frac{U_i^{(n)} - (1 - \frac{k}{n}c_\infty(L))}{\frac{k}{n}}, 0\right) \tag{8.1}$$

and let $v_n = \Pr(X'_{n,i} \neq 0) = \frac{k}{n}c_\infty(L)$. We may then write

$$\mathbb{S}_n(s,x) = \frac{1}{\sqrt{k}}\sum_{i=1}^{\lfloor sn\rfloor}\Big\{\mathbf{1}(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x) - \Pr(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x)\Big\} + \sqrt{k}\Big\{\frac{1}{n}\sum_{i=1}^{\lfloor ns\rfloor} c(\tfrac{i}{n}) - C(s)\Big\}x \equiv \mathbb{S}_{n,1}(s,x) + \mathbb{S}_{n,2}(s,x).$$

As a consequence of (B2), the term $\mathbb{S}_{n,2}$ converges to zero, uniformly in s and x, and we are left with investigating $\mathbb{S}_{n,1}$. We are going to identify that process with an empirical cluster process, see Drees and Rootzén (2010). In the following we set $L = 1$; the proof for arbitrary $L \in \mathbb{N}$ follows analogously. We also write $c_\infty = c_\infty(1)$.

Recall that $1 < r < n$ denotes an integer sequence converging to infinity such that $r = o(n)$ as $n \to \infty$. Let $Y_{n,j}$ denote the jth block of consecutive values of $X'_{n,1}, \dots, X'_{n,n}$, i.e.,

$$Y_{n,j} = (X'_{n,i})_{i\in I_j}, \quad I_j = \{(j-1)r + 1, \dots, jr\}, \quad j = 1, \dots, m = \lfloor n/r\rfloor.$$

We may then write

$$\begin{aligned}\mathbb{S}_{n,1}(s,x) &= c_\infty^{1/2}\Big\{\frac{1}{\sqrt{nv_n}}\sum_{i=1}^{rm}\mathbf{1}\big\{X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, \tfrac{i}{n} \le s\big\} \\ &\qquad - \mathrm{E}\,\mathbf{1}\big\{X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, \tfrac{i}{n} \le s\big\}\Big\} + o_{\mathrm{P}}(1) \\ &= c_\infty^{1/2}\Big\{\frac{1}{\sqrt{nv_n}}\sum_{j=1}^{m}\big\{\tilde f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}[\tilde f_{j,n,s,x}(Y_{n,j})]\big\}\Big\} + o_{\mathrm{P}}(1)\end{aligned}$$

where the $o_\mathrm{P}(1)$ is due to the fact that $mr \neq n$ in general, and where $\tilde f_{j,n,s,x}$ denotes the cluster functional (see Drees and Rootzén, 2010, Definition 2.1)

$$\tilde f_{j,n,s,x}(y_1,\dots,y_\ell) = \sum_{i=1}^{\ell} \mathbf{1}(y_i > c_\infty - c(\tfrac{(j-1)r+i}{n})x, \tfrac{(j-1)r+i}{n} \le s), \qquad \ell \in \mathbb{N}.$$

Hence, we need to show functional weak convergence of $\{\tilde{\mathbb{Z}}_n(s,x)\}_{(s,x)}$, where

$$\tilde{\mathbb{Z}}_n(s,x) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m} \big\{\tilde f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}[\tilde f_{j,n,s,x}((Y_{n,j})]\big\}. \tag{8.2}$$

Unfortunately, results from Drees and Rootzén (2010) are not directly applicable, as functions f depending on n (and, even more complicated, on j) are not allowed in their theory. Before proceeding, note that we may slightly redefine $\tilde f_{j,n,s,x}$. Indeed, let $\mathbb{Z}_n$ be defined analogously to $\tilde{\mathbb{Z}}_n$, but in terms of

$$f_{j,n,s,x}(y_1,\dots,y_\ell) = \mathbf{1}(j \le \lfloor sm \rfloor) g_{j,n,x}(y_1,\dots,y_\ell) \tag{8.3}$$

where

$$g_{j,n,x}(y_1,\dots,y_\ell) = \sum_{i=1}^{\ell} \mathbf{1}(y_i > c_\infty - c(\tfrac{(j-1)r+i}{n})x), \qquad \ell \in \mathbb{N}.$$

Now, for all $s,x \in [0,1]$,

$$\begin{aligned} |\mathbb{Z}_n(s,x) - \tilde{\mathbb{Z}}_n(s,x)| &\le 2\frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m}\sum_{i=1}^{r} |\mathbf{1}(\tfrac{j}{m} \le s) - \mathbf{1}(\tfrac{(j-1)r+i}{n} \le s)| \\ &\le 2\frac{r}{\sqrt{nv_n}} \sum_{j=1}^{m} \mathbf{1}(\tfrac{j-1}{m} < s \le \tfrac{j+1}{m}) \le 4\frac{r}{\sqrt{nv_n}}. \end{aligned} \tag{8.4}$$

Recalling $v_n = \frac{k}{n}c_\infty$, we have $r = o(\sqrt{nv_n})$ by (B3). As a consequence, we have shown that

$$\mathbb{S}_n = c_\infty^{1/2}\mathbb{Z}_n + o_\mathrm{P}(1) \quad \text{in} \quad \ell^\infty([0,1]\times[0,L]), \tag{8.5}$$

such that it is sufficient to show that the process $\mathbb{Z}_n$ converges to $c_\infty^{-1/2}\mathbb{S}$.

Consider weak convergence of the fidis of $\mathbb{Z}_n$ first, and for that purpose let us first assume that the blocks $Y_{n,1},\dots,Y_{n,m}$ are independent. The general case will be reduced to the independent case by the Bernstein blocking technique below. Under the assumption of independent blocks, we may apply the Cramér-Wold device and the classical Lindeberg CLT (Billingsley, 1995, Theorem 27.2). We need to show that

$$\lim_{n\to\infty} \mathfrak{c}_n((s,x),(s',x')), = c_\infty^{-1}\mathfrak{c}((s,x),(s',x')), \tag{8.6}$$

where $\mathfrak{c}$ is defined in (3.2) and where

$$\mathfrak{c}_n((s,x),(s',x')) = \frac{1}{nv_n}\sum_{j=1}^{m} \mathrm{Cov}\big(f_{j,n,s,x}(Y_{n,j}), f_{j,n,s',x'}(Y_{n,j})\big), \tag{8.7}$$

and that the Lindeberg condition is satisfied, that is, for any $(s,x) \in [0,1]^2$ and any $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{nv_n}\sum_{j=1}^{m} \mathrm{E}\Big[\{f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}\, f_{j,n,s,x}(Y_{n,j})\}^2$$

$$\mathbf{1}(|f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}\, f_{j,n,s,x}(Y_{n,j})| > \varepsilon\sqrt{nv_n})\Big] = 0.$$

Observing that $|f_{j,n,s,x}| \le r$, the Lindeberg condition is actually a simple consequence of the assumption $r = o(\sqrt{nv_n})$ in (B3), see also Corollary 3.6 in Drees and Rootzén (2010) for a similar argumentation.

It remains to prove (8.6), and for that purpose, we follow arguments from the proof of Remark 3.7 and Corollary 4.2 in Drees and Rootzén (2010). First of all, since $|\,\mathrm{E}[f_{j,n,s,x}(Y_{n,j})]| \le r\mathrm{P}(X_{n,1} \ne 0) = rv_n$ for all $j = 1,\dots,m$ and $s, x \in [0,1]$, we have that

$$\begin{aligned}\mathfrak{c}_n((s,x),(s',x')) &= \frac{1}{nv_n}\sum_{j=1}^{m}\mathrm{E}[f_{j,n,s,x}(Y_{n,j})f_{j,n,s',x'}(Y_{n,j})] + O(rv_n)\\ &= \frac{r}{n}\sum_{j=1}^{m}\mathbf{1}(j \le \lfloor (s\wedge s')m\rfloor)A_n(j) + O(rv_n) \end{aligned} \tag{8.8}$$

where

$$A_n(j) = \frac{1}{rv_n}\mathrm{E}\left[g_{j,n,x}(Y_{n,j})g_{j,n,x'}(Y_{n,j})\right]$$

and where the remainder is $o(1)$ by (B3).

Let us next calculate $A_n(j)$. For that purpose, recall the notion of the length of the core of a cluster y, denoted by $L(y)$, see Definition 2.1 in Drees and Rootzén (2010). Let $K > 0$ be a constant and decompose

$$\begin{aligned}A_n(j) = \frac{1}{rv_n}&\mathrm{E}\left[g_{j,n,x}(Y_{n,j})g_{j,n,x'}(Y_{n,j})\mathbf{1}(L(Y_{n,j}) \le K)\right]\\ &+ \frac{1}{rv_n}\mathrm{E}\left[g_{j,n,x}(Y_{n,j})g_{j,n,x'}(Y_{n,j})\mathbf{1}(L(Y_{n,j}) > K)\right]\\ = S_{n,K}&(j) + R_{n,K}(j).\end{aligned}$$

By stationarity, we have

$$\begin{aligned}R_{n,K}(j) &\le \frac{1}{rv_n}\sum_{i,i'=1}^{r}\mathrm{P}\Big(X'_{n,i} > 0, X'_{n,i'} > 0, L(Y_{n,1}) > K\Big)\\ &\le \frac{1}{rv_n}\mathrm{E}\Big[\Big(\sum_{i=1}^{r}\mathbf{1}(X'_{n,i} > 0)\Big)^2\mathbf{1}(L(Y_{n,1}) > K)\Big]\\ &\le \Big\{\frac{1}{rv_n}\mathrm{E}\Big[\Big(\sum_{i=1}^{r}\mathbf{1}(X'_{n,i} > 0)\Big)^{2+\delta}\Big]\Big\}^{2/(2+\delta)}\Big\{\frac{1}{rv_n}\mathrm{P}(L(Y_{n,1}) > K)\Big\}^{\delta/(2+\delta)}.\end{aligned}$$

Thus, as a consequence of (B4) and Lemma 5.2(vii) in Drees and Rootzén (2010), which is applicable by (B3), we obtain that

$$\lim_{K\to\infty}\limsup_{n\to\infty}\sup\{R_{n,K}(j) : j = 1,\dots,m\} = 0. \tag{8.9}$$

Further, for any $j \in \{1,\dots,m\}$,

$$\begin{aligned}S_{n,K}(j) &= \frac{1}{rv_n}\sum_{i,i'\in J_{n,j}}\mathrm{P}\Big(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, X'_{n,i'} > c_\infty - c(\tfrac{i'}{n})x', L(Y_{n,j}) \le K\Big)\\ &= \frac{1}{rv_n}\sum_{\substack{i,i'\in J_{n,j},\\ |i-i'|\le K}}\mathrm{P}\Big(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, X'_{n,i'} > c_\infty - c(\tfrac{i'}{n})x', L(Y_{n,j}) \le K\Big)\end{aligned}$$

$$= S'_{n,K}(j) + R'_{n,K}(j),$$

where

$$S'_{n,K}(j) = \frac{1}{rv_n} \sum_{\substack{i,i' \in J_{n,j}, \\ |i-i'| \le K}} \mathrm{P}\Big(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, X'_{n,i'} > c_\infty - c(\tfrac{i'}{n})x'\Big),$$

$$R'_{n,K}(j) = \frac{1}{rv_n} \sum_{\substack{i,i' \in J_{n,j}, \\ |i-i'| \le K}} \mathrm{P}\Big(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x, X'_{n,i'} > c_\infty - c(\tfrac{i'}{n})x', L(Y_{n,j}) > K\Big).$$

By similar calculations as in (8.9), we have that

$$\lim_{K\to\infty} \limsup_{n\to\infty} \sup\{R'_{n,K}(j) : j = 1, \dots, m\} = 0. \tag{8.10}$$

Further, by Lemma 9.1 and uniform continuity of c,

$$S'_{n,K}(j) = \frac{1}{c_\infty r} \sum_{i \in J_{n,j}} c(\tfrac{i}{n}) \Big\{ d_0(x,x') + \sum_{h=1}^{K\wedge(r-i)} \{d_h(x,x') + d_h(x',x)\} \Big\} + o(1)$$

$$= \frac{c(\frac{j-1}{m})}{c_\infty} D_K(x,x') + o(1),$$

where the $o(1)$ is uniform in $x, x' \in [0,1]$ and $j = 1, \dots, m$ and where

$$D_K(x,x') = d_0(x,x') + \sum_{h=1}^{K} \{d_h(x,x') + d_h(x',x)\}.$$

Assembling terms, we have

$$A_n(j) = \frac{c(\frac{j-1}{m})}{c_\infty} D_K(x,x') + R_{n,K}(j) + R'_{n,K}(j) + o(1)$$

where the $o(1)$ is uniform in $j = 1, \dots, m$ and $x, x' \in [0,1]$.

As a consequence of the latter display and (8.8), we obtain that

$$\mathfrak{c}_n((s,x),(s',x')) = \mathfrak{c}_{n,K}((s,x),(s',x')) + \mathfrak{r}_{n,K}((s,x),(s',x')) + o(1),$$

where

$$\mathfrak{c}_{n,K}((s,x),(s',x')) = c_\infty^{-1} \frac{r}{n} \sum_{j=1}^{\lfloor (s\wedge s')m \rfloor} c(\tfrac{j-1}{m}) D_K(x',x)$$

$$\mathfrak{r}_{n,K}((s,x),(s',x')) = c_\infty^{-1} \frac{r}{n} \sum_{j=1}^{\lfloor (s\wedge s')m \rfloor} \{R_{n,K}(j) + R'_{n,K}(j)\}$$

By (8.9) and (8.10), we have

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathfrak{r}_{n,K}((s,x),(s',x')) = 0.$$

Further,

$$\lim_{n\to\infty} \mathfrak{c}_{n,K}((s,x),(s',x')) = \frac{C(s\wedge s')}{c_\infty} D_K(x,x').$$

We may finally apply Lemma 9.2 to conclude that (8.6) is met.

The next step consists of getting rid of the assumption of independence of blocks. Recall that $1 < \ell_n < r$ denotes an integer sequence converging to infinity such that $\ell_n = o(r)$. We may then write $\mathbb{Z}_n(s,x) = \mathbb{Z}_n^+(s,x) + \mathbb{Z}_n^-(s,x)$, where

$$\mathbb{Z}_n^+(s,x) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lfloor sm \rfloor} \sum_{i=(j-1)r+1}^{jr-\ell_n} \mathbf{1}(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x) - \Pr(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x)$$

$$\mathbb{Z}_n^-(s,x) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lfloor sm \rfloor} \sum_{i=jr-\ell_n+1}^{jr} \mathbf{1}(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x) - \Pr(X'_{n,i} > c_\infty - c(\tfrac{i}{n})x).$$

Further, for $n \in \mathbb{N}$, let $Y^*_{n,1}, \dots, Y^*_{n,m}$ denote an i.i.d. sequence, where $Y^*_{n,1}$ is equal in distribution to $Y_{n,1}$. Let $\mathbb{Z}_n^*, \mathbb{Z}_n^{+,*}$ and $\mathbb{Z}_n^{-,*}$ be defined analogously to $\mathbb{Z}_n, \mathbb{Z}_n^+$ and $\mathbb{Z}_n^-$, but in terms of $Y^*_{n,1}, \dots, Y^*_{n,m}$. We will show that:

(i) For any $s, x \in [0,1]$, we have $\mathbb{Z}_n^{-,*}(s,x) = o_{\mathrm{P}}(1)$ and $\mathbb{Z}_n^-(s,x) = o_{\mathrm{P}}(1)$.

(ii) The fidis of $\mathbb{Z}_n^{+,*}$ converge weakly if and only if the fidis of $\mathbb{Z}_n^+$ converge weakly. In that case, the weak limits coincide.

As a consequence, the asymptotic distribution of the fidis of $\mathbb{Z}_n$ coincides with the asymptotic distribution of the fidis of $\mathbb{Z}_n^*$, and the latter has already been derived above.

Proof of (i). Let us first show that $\mathbb{Z}_n^{-,*}(s,x) = o_{\mathrm{P}}(1)$, which follows if we show that $\mathrm{Var}(\mathbb{Z}_n^{-,*}(s,x)) = o(1)$. Now, by stationarity,

$$\mathrm{E}\Big\{\sum_{i=1}^{r} \mathbf{1}(X'_{n,i} \neq 0)\Big\}^2 \geq \mathrm{E} \sum_{j=1}^{\lfloor r/\ell_n \rfloor} \Big\{ \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \mathbf{1}(X'_{n,i} \neq 0)\Big\}^2$$
$$= \Big\lfloor \frac{r}{\ell_n} \Big\rfloor \mathrm{E}\Big\{\sum_{i=1}^{\ell_n} \mathbf{1}(X'_{n,i} \neq 0)\Big\}^2. \tag{8.11}$$

As a consequence, by independence of blocks, stationarity and (B4),

$$\mathrm{Var}(\mathbb{Z}_n^{-,*}(s,x)) \leq \frac{m}{nv_n} \mathrm{Var}\Big(\sum_{i=1}^{\ell_n} \mathbf{1}(X'_{n,i} \neq 0)\Big) = O(\ell_n/r),$$

which converges to 0 by the assumption on ℓ_n.

Now, consider $\mathbb{Z}_n^-(s,x)$. Split the sum into two sums $\mathbb{Z}_n^{-,\mathrm{even}}(s,x)$ and $\mathbb{Z}_n^{-,\mathrm{odd}}(s,x)$, according to whether j is even or odd. It suffices to show that each of these sums is $o_{\mathrm{P}}(1)$. We only consider the sum over the even blocks; the argumentation for the odd blocks is similar. Now, since the observations making up the even numbered blocks are separated by r observations, we may follow the argumentation in Eberlein (1984) to obtain that

$$d_{\mathrm{TV}}(P^{(Y_{n,2j})_{1\leq j\leq \lfloor m/2\rfloor}}, P^{(Y^*_{n,2j})_{1\leq j\leq \lfloor m/2\rfloor}}) \leq \lfloor m/2 \rfloor \beta(r), \tag{8.12}$$

where d_{TV} denotes the total variation distance between two probability laws. Since $m\beta(\ell_n) = o(1)$ by (B3), the latter display is $o(1)$. As a consequence, $\mathbb{Z}_n^{-,\mathrm{even}}(s,x) = \mathbb{Z}_n^{-,\mathrm{even},*}(s,x) + o_{\mathrm{P}}(1)$. Finally, $\mathbb{Z}_n^{-,\mathrm{even},*}(s,x) = o_{\mathrm{P}}(1)$ by the same reasoning as for $\mathbb{Z}_n^{-,*}$.

Proof of (ii). Note that $\mathbb{Z}_n^+$ only depends on $(Y_{n,j}^{(r-\ell_n)})_{1\leq j\leq m}$, where $Y_{n,j}^{(r-\ell_n)}$ consists of the first $r-\ell_n$ coordinates of $Y_{n,j}$. A similar assertion holds for $\mathbb{Z}_n^{+,*}$, which is defined in terms of $((Y_{n,j}^*)^{(r-\ell_n)})_{1\le j\le m}$. The assertion in (ii) follows from the fact that

$$d_{\mathrm{TV}}(P^{(Y_{n,j}^{(r-\ell_n)})_{1\le j\le m}}, P^{((Y_{n,j}^*)^{(r-\ell_n)})_{1\le j\le m}}) \le m\beta(\ell_n) \to 0$$

by assumption and since the respective shortened blocks are separated by ℓ_n observations.

It remains to show asymptotic tightness. For that purpose, decompose $\mathbb{Z}_n = \mathbb{Z}_n^{\mathrm{even}} + \mathbb{Z}_n^{\mathrm{odd}}$ and likewise $\mathbb{Z}_n^* = \mathbb{Z}_n^{\mathrm{even},*} + \mathbb{Z}_n^{\mathrm{odd},*}$ into sums over even and odd numbered blocks. Clearly, asymptotic tightness of $\{\mathbb{Z}_n(s,x)\}_{(s,x)\in[0,1]^2}$ follows from asymptotic tightness of $\{\mathbb{Z}_n^{\mathrm{even}}(s,x))\}_{(s,x)\in[0,1]^2}$ and $\{\mathbb{Z}_n^{\mathrm{odd}}(s,x)\}_{(s,x)\in[0,1]^2}$. We only consider the even numbered blocks. In view of (8.12), it is further sufficient to show asymptotic tightness of $\{\mathbb{Z}_n^{\mathrm{even},*}(s,x)\}_{(s,x)\in[0,1]^2}$. To reduce the notational complexity, we instead prove asymptotic tightness of $\{\mathbb{Z}_n^*(s,x)\}_{(s,x)\in[0,1]^2}$. For that purpose, we apply Theorem 11.16 in Kosorok (2008), with t in that theorem replaced by (s,x), and with

$$f_{n,j}(\omega;(s,x)) = \mathbf{1}(\tfrac{j}{m} \le s) \times \frac{1}{\sqrt{nv_n}} \sum_{i\in J_{n,j}} \mathbf{1}\big(X'^*_{n,i}(\omega) > c_\infty - c(\tfrac{i}{n})x\big),$$

where ω is an element of the underlying probability space on which the $X'^*_{n,i}$ are defined. We need to show that

(1) $\{f_{n,j} : j = 1, \dots, m\}$ is almost measurable Suslin (AMS);
(2) the $\{f_{n,j}\}$ are manageable with envelopes $\{F_{n,j}\}$ given through

$$F_{n,j}(\omega) := \frac{1}{\sqrt{nv_n}} \sum_{i\in J_{n,j}} \mathbf{1}(X'^*_{n,i}(\omega) \ne 0);$$

(3) $\lim_{n\to\infty} \mathrm{E}\{\mathbb{Z}_n^*(s,x)\mathbb{Z}_n^*(s',x')\}$ exists for all $(s,x),(s',x') \in [0,1]^2$;
(4) $\limsup_{n\to\infty} \sum_{j=1}^m \mathrm{E}\, F_{n,j}^2 < \infty$;
(5) $\lim_{n\to\infty} \sum_{j=1}^m \mathrm{E}\, F_{n,j}^2 \mathbf{1}(F_{n,j} > \varepsilon) = 0$ for all $\varepsilon > 0$;
(6) $\rho(s,x;s',x') = \lim_{n\to\infty} \rho_n(s,x;s',x')$ exists for every $(s,x),(s',x') \in [0,1]^2$, where

$$\rho_n(s,x;s',x') := \left\{ \sum_{j=1}^m \mathrm{E} \left| f_{n,j}(\cdot;s,x) - f_{nj}(\cdot;s',x') \right|^2 \right\}^{1/2}. \tag{8.13}$$

[In that case, ρ defines a semimetric on $[0,1]^2$.] Moreover, $\rho_n(s_n,x_n;s_n',x_n') \to 0$ for all sequences $(s_n,x_n)_{n\in\mathbb{N}}, (s_n',x_n')_{n\in\mathbb{N}} \subset [0,1]^2$ such that $\rho(s_n,x_n;s_n',x_n') \to 0$.

Proof of (1)**.** By Lemma 11.15 in Kosorok (2008), the triangular array $\{f_{n,j}\}$ is AMS provided it is separable, that is, provided that, for every $n \in \mathbb{N}$, there exists a countable subset $S_n \subset [0,1]^2$ such that

$$\mathrm{P}^*\Bigg(\sup_{(s,x)\in[0,1]} \inf_{(s',x')\in S_n} \sum_{j=1}^m \{f_{n,j}(\omega;s,x) - f_{n,j}(\omega;s',x')\}^2 > 0 \Bigg) = 0.$$

Define $S_n := (\mathbb{Q} \cap [0,1])^2$ for all $n \in \mathbb{N}$. Then, for every element ω of the underlying probability space and for every $(s,x) \in [0,1]^2$, there exists $(s',x') \in S_n$ such that

$$\sum_{j=1}^m \{f_{n,j}(\omega;s,x) - f_{n,j}(\omega;s',x')\}^2 = 0.$$

Proof of (2)**.** By Theorem 11.17(iv) in Kosorok (2008), it suffices to prove that the triangular arrays $\{\tilde f_{n,j}(\omega;x) = \frac{1}{\sqrt{nv_n}} \sum_{i\in J_{n,j}} \mathbf{1}\big(X'^{*}_{n,i}(\omega) > c_\infty - c(\frac{i}{n})x\big) : x \in [0,1]\}_{j=1,\dots,m}$ and $\{\tilde g_{nj}(\omega;s) = \mathbf{1}(j/m \le s) : s \in [0,1]\}_{j=1,\dots,m}$ are manageable with respective envelopes $\{F_{n,j}(\omega)\}_{j=1,\dots,m}$ and $\{\tilde G_{nj}(\omega) \equiv 1\}_{j=1,\dots,m}$. Following the discussion on Page 221 in Kosorok (2008), these two assertions are consequences of the fact that both $\tilde f_{n,j}$ and $\tilde g_{n,j}$ are increasing in x and s, respectively.

Proof of (3), (4) **and** (5)**.** Condition (3) is simply the calculation of $\mathfrak{c}((s,x),(s',x'))$ above. Condition (4) is a consequence of (B4). Moreover, the assumption $r = o(\sqrt{nv_n})$ in (B3) implies (5).

Proof of (6)**.** Let

$$\sigma^2(x,x') = d_0(x,x') + \sum_{h=1}^{\infty} \big(d_h(x,x') + d_h(x',x)\big).$$

For $(s,x),(s',x') \in [0,1]^2$, let $\bar x = x$ if $s \ge s'$ and $\bar x = x'$ else. Then, by similar arguments that lead to (8.6), we have

$$\begin{aligned}
&\rho_n^2(s,x;s',x')\\
&= \frac{1}{nv_n}\Bigg\{ \sum_{j=1}^{\lfloor (s\wedge s')m\rfloor} \mathrm{E}\left\{g_{j,n,x}(Y_{n,j}) - g_{j,n,x'}(Y_{n,j})\right\}^2 + \sum_{j=\lfloor (s\wedge s')m\rfloor+1}^{\lfloor (s\vee s')m\rfloor} \mathrm{E}\left\{g_{j,n,\bar x}(Y_{n,j})\right\}^2 \Bigg\}\\
&= c_\infty^{-1}\Big\{ C(s\wedge s')\{\sigma^2(x,x) - 2\sigma^2(x,x') + \sigma^2(x',x')\}\\
&\qquad\qquad + \{C(s\vee s') - C(s\wedge s')\}\sigma^2(\bar x,\bar x)\Big\} + o(1)\\
&= \rho^2((s,x),(s',x')) + o(1),
\end{aligned}$$

for any fixed $(s,x),(s',x') \in [0,1]^2$. In order to show the convergence along sequences as claimed in (6), it is sufficient to show that the convergence in the last display is in fact uniform. Note that the argumentation used for pointwise convergence does not imply uniform convergence, due to the pointwise nature of the main argument, Lemma 9.2.

Let $t_j = (s_j, x_j, s_j', x_j') \in [0,1]^4, j = 1,2$. Suppose we have shown that

$$|\rho_n^2(t_1) - \rho_n^2(t_2)| \lesssim H_n(t_1,t_2) \tag{8.14}$$

with

$$H_n(t_1,t_2) = h_0(|x_1 - x_2| + q_n) + h_0(|x_1' - x_2'| + q_n) + |s_1 - s_2| + |s_1' - s_2'| + q_n,$$

where q_n denotes a sequence converging to zero (independent of t_1,t_2), where h_0 denotes a continuous, non-negative, increasing function on $[0,1]$ with $h_0(0) = 0$ and where the symbol '$\lesssim$' means that the left-hand side is bounded by a constant multiple of the right-hand side, the constant being independent of n, t_1, t_2. By pointwise convergence of ρ_n^2, we then also have

$$|\rho^2(t_1) - \rho^2(t_2)| \lesssim H(t_1,t_2), \tag{8.15}$$

where

$$H(t_1,t_2) = h_0(|x_1 - x_2|) + h_0(|x_1' - x_2'|) + |s_1 - s_2| + |s_1' - s_2'|.$$

Now, let $\varepsilon > 0$ be given. Then, by uniform continuity of h_0, there exists $\delta > 0$ such that $H_n(t_1,t_2) < \varepsilon$ and $H(t_1,t_2) < \varepsilon$ for all $\|t_1 - t_2\|_2 < \delta$ and for all n sufficiently large. Choose a finite grid of points $t^{(1)}, \dots, t^{(p)}$ such that each point $t \in [0,1]^4$ lies in the open ball of radius δ with center $t^{(j)}$, for some $j = 1, \dots, p$. Then,

$$|\rho_n^2(t) - \rho^2(t)| \le |\rho_n^2(t) - \rho_n^2(t^{(j)})| + |\rho_n^2(t^{(j)}) - \rho^2(t^{(j)})| + |\rho^2(t^{(j)}) - \rho^2(t)|$$
$$\lesssim 2\varepsilon + \max_{j=1}^{p} |\rho_n^2(t^{(j)}) - \rho^2(t^{(j)})|.$$

The upper bound does not depend on t, and converges to 2ε for $n \to \infty$ by pointwise convergence. Since $\varepsilon > 0$ was arbitrary, we obtain that $\rho_n^2 \to \rho^2$ uniformly.

It remains to show (8.14). Let $s_j^\vee = s_j \vee s_j'$ and $s_j^\wedge = s_j \wedge s_j'$. Up to symmetry, we need to distinguish three cases:

$$s_2^\vee \le s_1^\wedge, \quad s_2^\wedge \le s_1^\wedge \le s_2^\vee \le s_1^\vee, \quad s_2^\wedge \le s_1^\wedge \le s_1^\vee \le s_2^\vee.$$

For brevity, we only consider the first case, and make the further assumption that $s_2 < s_2' < s_1 < s_1'$. Introduce the notation $G_j(x) = g_{n,j,x}(Y_{n,j})$. We may then write

$$\rho_n^2(t_1) - \rho_n^2(t_2) = a_{n1} + a_{n2} + a_{n3} + a_{n4},$$

where

$$a_{n1} = \tfrac{1}{nv_n} \textstyle\sum_{j=1}^{\lfloor s_2 m \rfloor} \mathrm{E}\Big[\big\{G_j(x_1) - G_j(x_1')\big\}^2 - \big\{G_j(x_2) - G_j(x_2')\big\}^2\Big],$$
$$a_{n2} = \tfrac{1}{nv_n} \textstyle\sum_{\lfloor s_2 m \rfloor + 1}^{\lfloor s_2' m \rfloor} \mathrm{E}\Big[\big\{G_j(x_1) - G_j(x_1')\big\}^2 - \big\{G_j(\bar x_2)\big\}^2\Big],$$
$$a_{n3} = \tfrac{1}{nv_n} \textstyle\sum_{\lfloor s_2' m \rfloor + 1}^{\lfloor s_1 m \rfloor} \mathrm{E}\Big[\big\{G_j(x_1) - G_j(x_1')\big\}^2\Big],$$
$$a_{n4} = \tfrac{1}{nv_n} \textstyle\sum_{\lfloor s_1 m \rfloor + 1}^{\lfloor s_1' m \rfloor} \mathrm{E}\Big[\big\{G_j(\bar x_1)\big\}^2\Big].$$

Note that $\mathrm{E}\{G_j(x)\}^2 \le \mathrm{E}\{\sum_{i \in J_{n,1}} \mathbf{1}(X_{n,i}' > 0\}^2 = O(rv_n)$, uniformly in x and $j = 1, \dots, m$, by Condition (B4). Hence,

$$|a_{n2}| \lesssim \frac{\lfloor s_2' m \rfloor - \lfloor s_2 m \rfloor}{m} \le s_2' - s_2 + m^{-1} \le |s_1 - s_2| + m^{-1}.$$

Similarly, $|a_{n3}|$ and $|a_{n4}|$ are bounded by a constant multiple of $|s_1' - s_2'| + m^{-1}$. It remains to treat $|a_{n1}|$. The triangular inequality and the Cauchy-Schwarz-inequality imply that each summand of $|a_{n1}|$ can be bounded by

$$\mathrm{E}\Big[\big|G_j(x_1) - G_j(x_1') + G_j(x_2) - G_j(x_2')\big| \cdot \big|G_j(x_1) - G_j(x_1') - G_j(x_2) + G_j(x_2')\big|\Big]$$
$$\le \Big\{\mathrm{E}\,\big|G_j(x_1) - G_j(x_1') + G_j(x_2) - G_j(x_2')\big|^2\Big\}^{1/2}$$
$$\times \Big[\Big\{\mathrm{E}\,\big|G_j(x_1) - G_j(x_2)\big|^2\Big\}^{1/2} + \Big\{\mathrm{E}\,\big|G_j(x_1') - G_j(x_2')\big|^2\Big\}^{1/2}\Big]$$

The first factor is of the order $O((rv_n)^{1/2})$ by Condition (B4), uniformly in $j = 1, \dots, m$ and the x-arguments. Regarding the second factor note that, by Hölder-continuity of c as assumed in Condition (B2), we have

$$0 < c(\tfrac{jr}{n}) - K_c(\tfrac{r}{n})^{1/2} \le c(\tfrac{i}{n}) \le c(\tfrac{jr}{n}) + K_c(\tfrac{r}{n})^{1/2} \qquad \forall\, i \in J_{n,j},$$

for sufficiently large n. Without loss of generality, let $x_1 \le x_2$. Then, by monotonicity and Condition (B5),

$$\mathrm{E}\,\big|G_j(x_1) - G_j(x_2)\big|^2$$
$$= \mathrm{E}\Big\{\sum_{i \in J_{n,j}} \mathbf{1}(c_\infty - c(\tfrac{i}{n})x_1 \ge X_{n,i}' > c_\infty - c(\tfrac{i}{n})x_2)\Big\}^2$$

$$\le \mathrm{E}\Big\{ \sum_{i\in J_{n,j}} \mathbf{1}(c_\infty - \{c(\tfrac{jr}{n}) - K_c(\tfrac{r}{n})^{1/2}\}x_1 \ge X'_{n,i} > c_\infty - \{c(\tfrac{jr}{n}) + K_c(\tfrac{r}{n})^{1/2}\}x_2\Big\}^2$$
$$\le h\big(c(\tfrac{jr}{n})(x_2-x_1) + K_c(\tfrac{r}{n})^{1/2}(x_1+x_2)\big)r\tfrac{k}{n}$$
$$\le h\big(c_\infty(x_2-x_1) + 2K_c m^{-1/2}\big)rv_n \tag{8.16}$$

As a consequence,

$$|a_{n1}| \lesssim h^{1/2}(c_\infty|x_1-x_2| + 2K_c m^{-1/2}) + h^{1/2}(c_\infty|x'_1-x'_2| + 2K_c m^{-1/2})$$

which finally proves (8.14) with $h_0(x) = h^{1/2}(c_\infty x)$ and $q_n = 2K_c m^{-1/2}/c_\infty$. □

Proof of Proposition 6.2. Let $(s,x) \in [0,1]\times[0,L]$. Set $\varepsilon_n(x) = V(n/(kx)) = F^{-1}(1-kx/n)$ such that, almost surely,

$$\mathbb{F}_n(s,x) = \sqrt{k}\Big\{\frac{1}{k}\sum_{i=1}^n \mathbf{1}\Big\{U_i^{(n)} > F_{n,i}\Big(V\Big(\frac{n}{kx}\Big)\Big)\Big\}\mathbf{1}(i/n\le s) - xC(s)\Big\}$$
$$= \sqrt{k}\Big\{\frac{1}{k}\sum_{i=1}^n \mathbf{1}\Big\{U_i^{(n)} > 1 - \frac{kx}{n}\frac{1-F_{n,i}(\varepsilon_n(x))}{1-F(\varepsilon_n(x))}\Big\}\mathbf{1}(i/n\le s) - xC(s)\Big\}.$$

According to Condition (B6), there exist $y_0 < x^*$ and $\tau > 0$ such that, for all $y > y_0, n\in\mathbb{N}, 1\le i\le n$,

$$c(i/n)\Big\{1 - \frac{\tau}{c_{\min}}A\Big(\frac{1}{1-F(y)}\Big)\Big\} \le \frac{1-F_{n,i}(y)}{1-F(y)} \le c(i/n)\Big\{1 + \frac{\tau}{c_{\min}}A\Big(\frac{1}{1-F(y)}\Big)\Big\}.$$

Since $\varepsilon_n(x)\to x^*$, this implies, for n large enough,

$$\Big\{U_i^{(n)} > 1 - c(i/n)(1-\delta_n)\frac{kx}{n}\Big\} \subseteq \Big\{U_i^{(n)} > 1 - \frac{1-F_{n,i}(\varepsilon_n(x))}{1-F(\varepsilon_n(x))}\frac{kx}{n}\Big\}$$
$$\subseteq \Big\{U_i^{(n)} > 1 - c(i/n)(1+\delta_n)\frac{kx}{n}\Big\}, \tag{8.17}$$

where $\delta_n = \sup_{x\in(0,L]}\frac{\tau}{c_{\min}}A\big(\frac{n}{kx}\big) = \frac{\tau}{c_{\min}}A\big(\frac{n}{kL}\big)$. As a consequence, by the definition of $\mathbb{S}_n$ in (6.1), almost surely

$$\mathbb{S}_n(s, x(1-\delta_n)) - \sqrt{k}\delta_n xC(s) \le \mathbb{F}_n(s,x) \le \mathbb{S}_n(s, x(1+\delta_n)) + \sqrt{k}\delta_n xC(s).$$

Therefore,

$$\sup_{(s,x)\in[0,1]\times[0,L]} |\mathbb{F}_n(s,x) - \mathbb{S}_n(s,x)| \le 2w_{\delta_n}(\mathbb{S}_n) + 2\sqrt{k}\delta_n,$$

where, for $\delta > 0$,

$$w_\delta(\mathbb{S}_n) = \sup_{(s,y),(s,z)\in[0,1]^2:|y-z|<\delta} |\mathbb{S}_n(s,y) - \mathbb{S}_n(s,z)|. \tag{8.18}$$

Now, since $\sqrt{k}\delta_n \le \sqrt{k}\frac{\tau}{c_{\min}}A\big(\frac{n}{kL}\big) = o(1)$ by Condition (B6), it suffices to show that, for any $\varepsilon > 0$,

$$\limsup_{n\to\infty} \mathrm{P}(w_{\delta_n}(\mathbb{S}_n) > \varepsilon) = 0. \tag{8.19}$$

For arbitrary $\delta > 0$, we have

$$\mathrm{P}(w_{\delta_n}(\mathbb{S}_n) > \varepsilon) = \mathrm{P}(w_{\delta_n}(\mathbb{S}_n) > \varepsilon, \delta_n < \delta) + \mathrm{P}(w_{\delta_n}(\mathbb{S}_n) > \varepsilon, \delta_n \ge \delta)$$
$$\le \mathrm{P}(w_\delta(\mathbb{S}_n) > \varepsilon) + o(1).$$

In the following we set $L = 1$ in order to be able to refer to the proof of Proposition 6.1 in an easier manner; the general case $L\in\mathbb{N}$ can again be shown analogously.

Consider the semimetric ρ on $[0,1]^2$ defined in the proof of Proposition 6.1, see (8.13). By Theorem 11.16 in Kosorok (2008), we know that $[0,1]^2$ is totally bounded under ρ. Further, by (8.15), $\rho((s,y),(s,z)) \lesssim h_0^{1/2}(|y-z|) \le h_0^{1/2}(\delta)$ for all $s,y,z \in [0,1]$ with $|y-z|<\delta$, where h_0 is non-decreasing and continuous with $h_0(0)=0$. Consequently,

$$\begin{aligned}&\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \mathrm{P}(w_\delta(\mathbb{S}_n)>\varepsilon)\\ &\le \lim_{\delta\downarrow 0}\limsup_{n\to\infty} \mathrm{P}\Big(\sup_{(s,y),(s,z)\in[0,1]^2:\rho((s,y),(s,z))<\delta} |\mathbb{S}_n(s,y)-\mathbb{S}_n(s,z)|>\varepsilon\Big),\end{aligned}$$

which equals 0 by Theorem 7.19 and Theorem 11.16 in Kosorok (2008), the latter being applicable because of the proof of Proposition 6.1. □

9. Auxiliary Results

Lemma 9.1. *Under the assumptions of Proposition 6.1, for any fixed $h \ge 0$, we have that*

$$\sup_{x,x'\in[0,L]}\sup_{i=1,\dots n}\left|\frac{1}{v_n}\mathrm{P}\Big(X'_{n,i}>c_\infty-c(\tfrac{i}{n})x, X'_{n,i+h}>c_\infty-c(\tfrac{i+h}{n})x'\Big)-\frac{c(\frac{i}{n})}{c_\infty}d_h(x,x')\right|$$

converges to 0 as $n\to\infty$, where $v_n=v_n(L)=\frac{k}{n}c_\infty$ with $c_\infty=c_\infty(L)$.

Proof. First note that, as a consequence of (2.1) and the continuous mapping theorem, for any $\ell\in\mathbb{N}$ and with $c_n=c_\infty\frac{k}{n}$,

$$\begin{aligned}&\mathrm{P}\big((X'_{n,1}\dots,X'_{n,\ell})\in dx \mid X'_{n,1}>0\big)\\ &\quad=\mathrm{P}\Big(\big(c_\infty\{1-(c_nZ_1)^{-1}\}_+,\dots,c_\infty\{1-(c_nZ_\ell)^{-1}\}_+\big)\in dx \,\Big|\, Z_1>c_n^{-1}\Big)\\ &\quad\rightsquigarrow \mathrm{P}((W_1,\dots,W_\ell)\in dx),\end{aligned}$$

where $W_j=c_\infty(1-1/Y_{j-1})_+$. Note that W_1 is standard uniform on $(0,c_\infty)$ and that $W_j\ge 0$ may have an atom at zero and is absolutely continuous on $(0,c_\infty)$, for $j\ge 2$. A simple extension of Lemma 2.11 in van der Vaart (1998) implies that

$$\sup_{x_1,\dots,x_\ell>0}\big|\mathrm{P}(X'_{n,1}>x_1\dots,X'_{n,\ell}>x_\ell \mid X'_{n,1}>0)-\mathrm{P}(W_1>x_1,\dots,W_\ell>x_\ell)\big|=o(1).$$

Thus, for $h\ge 0$ fixed, by uniform continuity of c and $r=o(n)$,

$$\begin{aligned}&\frac{1}{v_n}\mathrm{P}\Big(X'_{n,i}>c_\infty-c(\tfrac{i}{n})x, X'_{n,i+h}>c_\infty-c(\tfrac{i+h}{n})x'\Big)\\ &\qquad=\mathrm{P}\Big(X'_{n,i}>c_\infty-c(\tfrac{i}{n})x, X'_{n,i+h}>c_\infty-c(\tfrac{i+h}{n})x' \mid X'_{n,i}>0\Big)\\ &\qquad=\mathrm{P}\Big(W_1>c_\infty-c(\tfrac{i}{n})x, W_{h+1}>c_\infty-c(\tfrac{i+h}{n})x'\Big)+o(1)\\ &\qquad=\mathrm{P}\Big(W_1>c_\infty-c(\tfrac{i}{n})x, W_{h+1}>c_\infty-c(\tfrac{i}{n})x'\Big)+o(1),\end{aligned}$$

where the $o(1)$ is uniform in $i=1,\dots,n$ and $x,x'\in[0,L]$. Further, by the spectral decomposition of $(Y_t)_{t\in\mathbb{N}_0}$ (Theorem 3.1 in Basrak and Segers, 2009), that is $(Y_t)_{t\in\mathbb{N}_0}=(Y_0\Theta_t)_{t\in\mathbb{N}_0}$ for some process $(\Theta_t)_{t\in\mathbb{N}_0}$ independent of Y_0 and with $\Theta_0=1$, we obtain, by a change of variable,

$$\begin{aligned}&\mathrm{P}\Big(W_1>c_\infty-c(\tfrac{i}{n})x, W_{h+1}>c_\infty-c(\tfrac{i}{n})x'\Big)\\ &\qquad=\mathrm{P}\Big(Y_0>\frac{c_\infty}{c(\frac{i}{n})x}, Y_h>\frac{c_\infty}{c(\frac{i}{n})x'}\Big)\end{aligned}$$

$$= \int_1^\infty \mathrm{P}\Big(Y_0 > \frac{c_\infty}{c(\frac{i}{n})x}, Y_0\Theta_h > \frac{c_\infty}{c(\frac{i}{n})x'} \mid Y_0 = y\Big) y^{-2}\, dy$$
$$= \int_{\frac{c_\infty}{c(i/n)x}}^\infty \mathrm{P}\Big(\Theta_h > \frac{c_\infty}{yc(\frac{i}{n})x'}\Big) y^{-2}\, dy$$
$$= \frac{c(\frac{i}{n})}{c_\infty} \int_{1/x}^\infty \mathrm{P}\Big(\Theta_h > \frac{1}{zx'}\Big) z^{-2}\, dz$$
$$= \frac{c(\frac{i}{n})}{c_\infty} \mathrm{P}\Big(Y_0 > \frac{1}{x}, Y_h > \frac{1}{x'}\Big) = \frac{c(\frac{i}{n})}{c_\infty} d_h(x, x'),$$

which implies the assertion. □

Lemma 9.2. *Let $(a_n)_{n\in\mathbb{N}}$, $(c_k)_{k\in\mathbb{N}}$ and $(r_{n,k})_{(n,k)\in\mathbb{N}^2}$ be sequences satisfying*

$$a_n = c_k + r_{n,k} \quad \text{and} \quad \lim_{k\to\infty} \limsup_{n\to\infty} |r_{n,k}| = 0.$$

Then $(c_k)_{k\in\mathbb{N}}$ and $(a_n)_{n\in\mathbb{N}}$ are converging, and the respective limits are equal.

Proof. Let $R_k = \limsup_{n\to\infty} r_{n,k}$. Along a subsequence, we have $\lim_{\ell\to\infty} r_{n_\ell,k} = R_k$. Hence, $a = \lim_{\ell\to\infty} a_{n_\ell} = c_k + R_k$ exists, and therefore $\lim_{k\to\infty} c_k = a$. Finally, $\limsup_{n\to\infty} a_n \le c_k + R_k \to a$ and $\liminf_{n\to\infty} a_n \ge c_k - R_k \to a$ as $k \to \infty$. □

Acknowledgements

This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation, which is gratefully acknowledged. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. The authors are grateful to Chen Zhou for helpful discussions and suggestions.

References

Aue, A. and L. Horváth (2013). Structural breaks in time series. *J. Time Series Anal. 34*(1), 1–16.

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

Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004). *Statistics of extremes: Theory and Applications.* Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd.

Berbee, H. C. P. (1979). *Random walks with stationary increments and renewal theory*, Volume 112 of *Mathematical Centre Tracts.* Amsterdam: Mathematisch Centrum.

Berghaus, B. and A. Bücher (2017). Goodness-of-fit tests for multivariate copula-based time series models. *Econometric Theory 33*(2), 292–330.

Berghaus, B. and A. Bücher (2018). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Ann. Statist. 46*(5), 2307–2335.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Billingsley, P. (1995). *Probability and measure* (Third ed.). Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York. A Wiley-Interscience Publication.

Brockwell, P. J. and R. A. Davis (1991). *Time series: theory and methods* (Second ed.). Springer Series in Statistics. Springer-Verlag, New York.

Bücher, A., P. Kinsvater, and I. Kojadinovic (2017). Detecting breaks in the dependence of multivariate extreme-value distributions. *Extremes 20*(1), 53–89.

Bücher, A. and I. Kojadinovic (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *Journal of Theoretical Probability 32*(3), 1145–1165.

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer Series in Statistics. Springer-Verlag London, Ltd., London.

Dahlhaus, R. (2012). Locally stationary processes. In T. Subba Rao, S. Subba Rao, and C. Rao (Eds.), *Time Series Analysis: Methods and Applications*, Volume 30 of *Handbook of Statistics*, pp. 351–413. Elsevier.

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

Davison, A. C. and R. L. Smith (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological) 52*(3), 393–425.

de Haan, L., S. I. Resnick, H. Rootzén, and C. G. de Vries (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl. 32*(2), 213–224.

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

de Haan, L. and C. Zhou (2021). Trends in extreme value indices. *Journal of the American Statistical Association 116*(535), 1265–1279.

Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl. Probab. 10*(4), 1274–1301.

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

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

Drees, H., J. Segers, and M. Warchoł (2015). Statistics for tail processes of markov chains. *Extremes 18*(3), 369–402.

Eberlein, E. (1984). Weak convergence of partial sums of absolutely regular sequences. *Statist. Probab. Lett. 2*(5), 291–293.

Einmahl, J. H. J., L. de Haan, and C. Zhou (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78*(1), 31–51.

Einmahl, J. H. J., A. Ferreira, L. de Haan, C. Neves, and C. Zhou (2022). Spatial dependence and space–time trend in extreme events. *Ann. Statist. 50*(1), 30–52.

Hall, P. and N. Tajvidi (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci. 15*(2), 153–167.

Hoga, Y. (2017). Change point tests for the tail index of β-mixing random variables. *Econometric Theory 33*(4), 915–954.

Hoga, Y. (2018). A structural break test for extremal dependence in β-mixing random vectors. *Biometrika 105*(3), 627–643.

Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Statistics and computing 3*(3), 135–146.

Kley, T., S. Volgushev, H. Dette, and M. Hallin (2016). Quantile spectral processes: Asymptotic analysis and inference. *Bernoulli 22*(3), 1770–1807.

Kojadinovic, I. and P. Naveau (2017). Detecting distributional changes in samples of independent block maxima using probability weighted moments. *Extremes 20*(2), 417–450.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference*. Springer Series in Statistics. New York: Springer.

Kulik, R. and P. Soulier (2020). *Heavy-Tailed Time Series*. Springer Series in Operations Research and Financial Engineering. Springer New York.

Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete 65*(2), 291–306.

Leadbetter, M. R., G. Lindgren, and H. Rootzén (1983). *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York-Berlin.

Northrop, P. J. (2015). An efficient semiparametric maxima estimator of the extremal index. *Extremes 18*(4), 585–603.

Segers, J. (2003). Functionals of clusters of extremes. *Adv. in Appl. Probab. 35*(4), 1028–1045.

Tsybakov, A. B. (2009). *Introduction to nonparametric estimation*. Springer Series in Statistics. Springer, New York. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

van der Vaart, A. W. (1998). *Asymptotic Statistics*, Volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes - Springer Series in Statistics*. New York: Springer.

SUPPLEMENTARY MATERIAL ON "STATISTICS FOR HETEROSCEDASTIC TIME SERIES EXTREMES"

AXEL BÜCHER AND TOBIAS JENNESSEN

Abstract. This supplementary material contains the remaining proofs for the main paper. Proofs for Sections 3-5 are presented in Sections A-C, respectively. Some auxiliary results are collected in Section D.

Appendix A. Proofs for Section 3

Proof of Theorem 3.1. Fix $s \in [0,1]$. By definition, $K_b(\cdot,0)$ and $K_b(\cdot,1)$ do not depend on n, and the same is true for $K_b(\cdot,s)$ with $s \in (0,1)$ and sufficiently large n; we then have $K_b(\cdot,s) = K$. Let

$$\Psi_n(x) = k^{-1} \sum_{i=1}^n \mathbf{1}\Big(X_i^{(n)} > V\Big(\frac{n}{kx}\Big)\Big)$$

such that $\mathbb{F}_n(1,x) = \sqrt{k}\{\Psi_n(x) - x\}$. By Proposition 6.2, $\{\mathbb{F}_n(1,x)\}_{x\in[0,1]} \rightsquigarrow \{\mathbb{S}(1,x)\}_{x\in[0,1]}$ in $(\ell^\infty([0,1]), \|\cdot\|_\infty)$. Note that $\Psi_n^{-1}(x) = nk^{-1}\big\{1 - F(X_{n,n-\lfloor kx\rfloor})\big\}$, such that

$$\Big\{\sqrt{k}\Big(nk^{-1}(1 - F(X_{n,n-\lfloor kx\rfloor})) - x\Big) + \mathbb{F}_n(1,x)\Big\}_{x\in[0,1]} = o_{\mathrm{P}}(1)$$

by the functional delta-method applied to the inverse map (see Theorem 3.9.4 in van der Vaart and Wellner, 1996). In particular, for $y_n = nk^{-1}\{1 - F(X_{n,n-k})\}$ we obtain

$$\sqrt{k}(y_n - 1) = -\mathbb{F}_n(1,1) + o_{\mathrm{P}}(1) \rightsquigarrow -\mathbb{S}(1,1), \tag{A.1}$$

yielding

$$\mathrm{P}\big(h^{1/4}k^{1/2}|y_n - 1| \le 1\big) \to 1. \tag{A.2}$$

Let $K_b^+(\cdot,s)$ and $K_b^-(\cdot,s)$ denote the positive and negative part of $K_b(\cdot,s)$, respectively, and define, for $y > 0$,

$$c_n^{\pm}(y,s) = \frac{1}{kh}\sum_{i=1}^n \mathbf{1}\Big(X_i^{(n)} > V\Big(\frac{n}{ky}\Big)\Big)K_b^{\pm}\Big(\frac{s - i/n}{h}, s\Big),$$

such that $\tilde{c}_n(s) = c_n^+(y_n,s) - c_n^-(y_n,s)$. Note that $c_n^{\pm}(\cdot,s)$ is monotonically increasing; therefore on the event $\{h^{1/4}k^{1/2}|y_n - 1| \le 1\}$ in (A.2) we have

$$c_n^+(y^-,s) - c_n^-(y^+,s) \le c_n^+(y_n,s) - c_n^-(y_n,s) \le c_n^+(y^+,s) - c_n^-(y^-,s).$$

where $y^{\pm} = 1 \pm (k^{1/2}h^{1/4})^{-1}$.

The proof of the theorem is finished once we have shown

$$\sqrt{kh}\big\{c_n^+(y^+,s) - c_n^-(y^-,s) - c(s)\big\} \rightsquigarrow \mathcal{N}(\mu_s, \sigma_s^2), \tag{A.3}$$

$$\sqrt{kh}\big\{c_n^+(y^-,s) - c_n^-(y^+,s) - c(s)\big\} \rightsquigarrow \mathcal{N}(\mu_s, \sigma_s^2). \tag{A.4}$$

April 21, 2022

We restrict ourselves to proving (A.3), the assertion in (A.4) can be treated analogously. Set

$$d_n^{\pm}(y,s) = \frac{1}{kh}\sum_{i=1}^n \mathbf{1}\Big(U_i^{(n)} > 1 - c(i/n)\frac{ky}{n}\Big)K_b^{\pm}\Big(\frac{s-i/n}{h},s\Big)$$

and let us first show that

$$\sqrt{kh}\big\{c_n^+(y^+,s) - c_n^-(y^-,s) - d_n^+(y^+,s) + d_n^-(y^-,s)\big\} = o_{\mathrm{P}}(1). \tag{A.5}$$

which is a consequence of

$$\sqrt{kh}\big\{c_n^+(y^+,s) - d_n^+(y^+,s)\big\} = o_{\mathrm{P}}(1), \quad \sqrt{kh}\big\{c_n^-(y^-,s) - d_n^-(y^-,s)\big\} = o_{\mathrm{P}}(1). \tag{A.6}$$

We only prove the first assertion in (A.6), the second one follows by similar arguments. By the same arguments that lead to (8.17), defining $\varepsilon_n = F^{-1}(1-ky^+/n)$, we have

$$\Big\{U_i^{(n)} > 1 - c(i/n)(1-w_n)\frac{ky^+}{n}\Big\} \subseteq \Big\{U_i^{(n)} > 1 - \frac{1-F_{n,i}(\varepsilon_n)}{1-F(\varepsilon_n)}\frac{ky^+}{n}\Big\}$$
$$\subseteq \Big\{U_i^{(n)} > 1 - c(i/n)(1+w_n)\frac{ky^+}{n}\Big\},$$

where $w_n = \frac{\tau}{c_{\min}}A\big(\frac{1}{1-F(\varepsilon_n)}\big) = \frac{\tau}{c_{\min}}A\big(\frac{n}{ky^+}\big)$. Consequently, rewriting

$$c_n^+(y,s) = \frac{1}{kh}\sum_{i=1}^n \mathbf{1}\Big(U_i^{(n)} > F(\varepsilon_n)\Big) = \frac{1}{kh}\sum_{i=1}^n \mathbf{1}\Big\{U_i^{(n)} > 1 - \frac{1-F_{n,i}(\varepsilon_n)}{1-F(\varepsilon_n)}\frac{ky^+}{n}\Big\},$$

(which is true a.s.), we have

$$d_{n,-}^+(y^+,s) \le c_n^+(y^+,s) \le d_{n,+}^+(y^+,s), \tag{A.7}$$

where

$$d_{n,\pm}^+(x,s) = \frac{1}{kh}\sum_{i=1}^n K_b^+\Big(\frac{s-i/n}{h},s\Big)\mathbf{1}\Big(U_i^{(n)} > 1 - c(i/n)(1\pm w_n)\frac{kx}{n}\Big).$$

As a consequence of (A.7), the proof of the first assertion in (A.6) is finished once we show that

$$\sqrt{kh}\big\{d_{n,\pm}^+(y^+,s) - d_n^+(y^+,s)\big\} = o_{\mathrm{P}}(1). \tag{A.8}$$

For that purpose, note that

$$\begin{aligned}
&\mathrm{E}\left[\sqrt{kh}|d_{n,\pm}^+(y^+,s) - d_n^+(y^+,s)|\right] \\
&\le \frac{1}{\sqrt{kh}}\sum_{i=1}^n K_b^+\Big(\frac{s-i/n}{h},s\Big)\mathrm{E}\left[\Big|\mathbf{1}\Big(U_i^{(n)} > 1 - c(i/n)(1\pm w_n)\frac{ky^+}{n}\Big)\right. \\
&\qquad\qquad \left. - \mathbf{1}\Big(U_i^{(n)} > 1 - c(i/n)\frac{ky^+}{n}\Big)\Big|\right] \\
&\le 2w_n y^+\sqrt{kh}\frac{1}{nh}\sum_{i=1}^n K_b^+\Big(\frac{s-i/n}{h},s\Big)c(i/n) \\
&= \frac{2\tau}{c_{\min}}y^+\sqrt{kh}A\Big(\frac{n}{ky^+}\Big)\frac{1}{nh}\sum_{i=1}^n K_b^+\Big(\frac{s-i/n}{h},s\Big)c(i/n) \\
&= \frac{2\tau}{c_{\min}}y^+\sqrt{kh}A\Big(\frac{n}{ky^+}\Big)\Big\{c(s)\eta_1(s) - hc'(s)\eta_2(s) + \frac{h^2}{2}c''(s)\eta_3(s) + o(h^2) + O\Big(\frac{1}{nh}\Big)\Big\}
\end{aligned}$$

by Lemma A.1, where η_i is defined as in this lemma with K replaced by $K_b^+(\cdot, s)$. The term in the last line of the above display converges to zero since $\sqrt{kh}A(n/(ky^+)) \le \sqrt{k}A(n/(2k)) \to 0$ by assumption. This proves (A.8) and hence (A.5) as argued above.

In the next step, we enforce a block structure, later allowing us to apply mixing conditions and show asymptotic independence of blocks. Let r from Condition (B3) denote the length of a block, and for simplicity we assume $m = n/r \in \mathbb{N}$ (otherwise, a potential remainder block of less than r observations can be shown to be asymptotically negligible). Set

$$e_n^{\pm}(y,s) = \frac{1}{kh}\sum_{j=1}^m K_b^{\pm}\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j}\mathbf{1}\Big(U_t^{(n)} > 1 - c(\tfrac{t}{n})\frac{ky}{n}\Big).$$

Subsequently, we show

$$\sqrt{kh}\big\{d_n^+(y^+,s) - d_n^-(y^-,s) - e_n^+(y^+,s) + e_n^-(y^-,s)\big\} = o_{\mathrm{P}}(1). \tag{A.9}$$

Write

$$\begin{aligned}
&\mathrm{E}\left[\sqrt{kh}\big|d_n^+(y^+,s) - e_n^+(y^+,s)\big|\right] \\
&\le \frac{1}{\sqrt{kh}}\sum_{j=1}^m\sum_{t\in I_j}\mathrm{P}\Big(U_t^{(n)} > 1 - c(\tfrac{t}{n})\frac{ky^+}{n}\Big)\Big|K_b^+\Big(\frac{s-t/n}{h}, s\Big) - K_b^+\Big(\frac{s-j/m}{h}, s\Big)\Big| \\
&= \sqrt{\frac{k}{h}}\frac{y^+}{n}\sum_{j=1}^m\sum_{l=0}^{r-1} c\Big(\frac{jr-l}{n}\Big)\Big|K_b^+\Big(\frac{s-\frac{jr-l}{n}}{h}, s\Big) - K_b^+\Big(\frac{s-\frac{jr}{n}}{h}, s\Big)\Big|.
\end{aligned}$$

Since $K_b^{\pm}(\cdot, s)$ does not depend on n for sufficiently large n and is Lipschitz-continuous, say with constant L', the above can be bounded by

$$L'y^+\sqrt{\frac{k}{h}}\frac{r}{n^2h}\sum_{j=1}^m\sum_{l=0}^{r-1} c\Big(\frac{jr-l}{n}\Big) = L'y^+\frac{k^{1/2}r}{n^2h^{3/2}}\sum_{j=1}^n c(\tfrac{j}{n})$$

which converges to zero by Condition (B7). Analogously, $\mathrm{E}\left[\sqrt{kh}\big|d_n^-(y^-,s) - e_n^-(y^-,s)\big|\right] = o(1)$, implying that (A.9) holds. Together with (A.5), we have shown that

$$\sqrt{kh}\big\{c_n^+(y^+,s) - c_n^-(y^-,s) - c(s)\big\} = \sqrt{kh}\big\{e_n^+(y^+,s) - e_n^-(y^-,s) - c(s)\big\} + o_{\mathrm{P}}(1),$$

whence the assertion in (A.3) is shown once we prove that

$$H_n = \sqrt{kh}\big\{e_n^+(y^+,s) - e_n^-(y^-,s) - c(s)\big\} \rightsquigarrow \mathcal{N}(\mu_s, \sigma_s^2). \tag{A.10}$$

The assertion in (A.10) in turn is a consequence of

$$\lim_{n\to\infty}\mathrm{E}[H_n] = \mu_s, \qquad H_n - \mathrm{E}[H_n] \rightsquigarrow \mathcal{N}(0, \sigma_s^2). \tag{A.11}$$

We start by proving the assertion regarding $\mathrm{E}[H_n]$ in (A.11). For that purpose, write

$$\begin{aligned}
&\mathrm{E}\Big[e_n^+(y^+,s) - e_n^-(y^-,s)\Big] \\
&= \frac{1}{kh}\sum_{j=1}^m K_b^+\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j}\mathrm{P}\Big(U_t^{(n)} > 1 - c(\tfrac{t}{n})\frac{ky^+}{n}\Big) \\
&\qquad - K_b^-\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j}\mathrm{P}\Big(U_t^{(n)} > 1 - c(\tfrac{t}{n})\frac{ky^-}{n}\Big)
\end{aligned}$$

$$= \frac{1}{nh} \sum_{j=1}^{m} \sum_{t \in I_j} c(\tfrac{t}{n}) \Big\{ K_b^+\Big(\frac{s-j/m}{h}, s\Big) y^+ - K_b^-\Big(\frac{s-j/m}{h}, s\Big) y^- \Big\}$$
$$= \frac{1}{mh} \sum_{j=1}^{m} c(\tfrac{j}{m}) \Big\{ K_b^+\Big(\frac{s-j/m}{h}, s\Big) y^+ - K_b^-\Big(\frac{s-j/m}{h}, s\Big) y^- \Big\} + o\big((kh)^{-1/2}\big),$$

where the $o\big((kh)^{-1/2}\big)$-term is due to c being Lipschitz-continuous and $kr^2 = o(n^2h)$, which holds by Condition (B7) and $r = o(n/k)$ from (B3).

Hence, the above calculation and Lemma A.1 imply that

$$\frac{1}{\sqrt{kh}} \mathrm{E}[H_n] = y^+ \Big(c(s)\eta_1^+(s) - hc'(s)\eta_2^+(s) + \frac{h^2}{2} c''(s)\eta_3^+(s) + o(h^2) + O(\tfrac{1}{mh}) \Big)$$
$$-y^- \Big(c(s)\eta_1^-(s) - hc'(s)\eta_2^-(s) + \frac{h^2}{2} c''(s)\eta_3^-(s) + o(h^2) + O(\tfrac{1}{mh}) \Big)$$
$$-c(s) + o\big((kh)^{-1/2}\big),$$

where η_i^+ and η_i^- are defined as η_i in Lemma A.1 but with K replaced by $K_b^+(\cdot, s)$ and $K_b^-(\cdot, s)$, respectively (note that the latter two functions do not depend on s or n as argued at the beginning of this proof). Next, note that $|y^\pm - 1| = (k^{1/2}h^{1/4})^{-1} = o\big(1/\sqrt{kh}\big)$ and $o(h^2) + O\big(1/(mh)\big) = o\big(1/\sqrt{kh}\big)$ due to $k^{1/5}h \to \lambda$ and $kr^2 = o(n^2h)$, which follows from Conditions (B7) and (B3). As a consequence,

$$\mathrm{E}[H_n] = \sqrt{kh} \Big(c(s)\big(\eta_1^+(s) - \eta_1^-(s) - 1\big) - hc'(s)\big(\eta_2^+(s) - \eta_2^-(s)\big) + \frac{h^2}{2} c''(s)\big(\eta_3^+(s) - \eta_3^-(s)\big) \Big) + o(1). \tag{A.12}$$

Note that $K_b(\cdot, s) = K_b^+(\cdot, s) - K_b^-(\cdot, s)$. First, let $s \in (0,1)$. For n large enough such that $h < s < 1-h$, we have $K_b(x, s) = K(x)$, $x \in [-1, 1]$, and

$$\eta_1^+(s) - \eta_1^-(s) = \int_{-1}^{1} K(x)\,\mathrm{d}x = 1, \quad \eta_2^+(s) - \eta_2^-(s) = \int_{-1}^{1} K(x)x\,\mathrm{d}x = 0,$$
$$\eta_3^+(s) - \eta_3^-(s) = \int_{-1}^{1} K(x)x^2\,\mathrm{d}x = a(s).$$

Second, for $s = 1$, the construction of the boundary kernel implies (Jones, 1993)

$$\eta_1^+(s) - \eta_1^-(s) = \int_{0}^{1} K_b(x,1)\,\mathrm{d}x = 1, \quad \eta_2^+(s) - \eta_2^-(s) = \int_{0}^{1} K_b(x,1)x\,\mathrm{d}x = 0,$$
$$\eta_3^+(s) - \eta_3^-(s) = \int_{0}^{1} K_b(x,1)x^2\,\mathrm{d}x = a(1).$$

And for $s = 0$, we have

$$\eta_1^+(s) - \eta_1^-(s) = \int_{-1}^{0} K_b(x,0)\,\mathrm{d}x = 1, \quad \eta_2^+(s) - \eta_2^-(s) = \int_{-1}^{0} K_b(x,0)x\,\mathrm{d}x = 0,$$
$$\eta_3^+(s) - \eta_3^-(s) = \int_{-1}^{0} K_b(x,0)x^2\,\mathrm{d}x = a(0).$$

Altogether, these equalities and equation (A.12) yield $\lim_{n\to\infty} \mathrm{E}[H_n] = \frac{\lambda^{5/2}}{2} c''(s)a(s) = \mu_s$ for any $s \in [0,1]$, as asserted in (A.11), where we again used $k^{1/5}h \to \lambda$ from (B7).

Next, consider the assertion on the right-hand side of (A.11). For that purpose, recall $c_\infty = c_\infty(2) = 1 + 2\|c\|_\infty$ and $X'_{n,i}$ from (8.1) with $L = 2$. We may then rewrite $e_n^\pm$ as

$$e_n^\pm(y^\pm, s) = \frac{1}{kh}\sum_{j=1}^m K_b^\pm\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^\pm\big).$$

We are going to apply a big-block-small-block technique. For that purpose, let

$$I_j^B = \{(j-1)r+1, \dots, jr-\ell_n\}, \quad I_j^S = \{jr-\ell_n+1, \dots, jr\},$$

where the sequence $(\ell_n)_n$ is from Condition (B3). Set

$$e_{n,B}^\pm(y,s) = \frac{1}{kh}\sum_{j=1}^m K_b^\pm\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j^B}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y\big) - \mathrm{P}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y\big),$$

$$e_{n,S}^\pm(y,s) = \frac{1}{kh}\sum_{j=1}^m K_b^\pm\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j^S}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y\big) - \mathrm{P}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y\big).$$

As a consequence, we may write

$$H_n - \mathrm{E}[H_n] = \sqrt{kh}\big\{e_{n,B}^+(y^+,s) - e_{n,B}^-(y^-,s)\big\} + \sqrt{kh}\big\{e_{n,S}^+(y^+,s) - e_{n,S}^-(y^-,s)\big\},$$

whence the assertion on the right-hand side of (A.11) follows if we prove that

$$H_{n1} := \sqrt{kh}\big\{e_{n,S}^+(y^+,s) - e_{n,S}^-(y^-,s)\big\} = o_{\mathrm{P}}(1), \tag{A.13}$$

$$H_{n2} := \sqrt{kh}\big\{e_{n,B}^+(y^+,s) - e_{n,B}^-(y^-,s)\big\} \rightsquigarrow \mathcal{N}(0,\sigma_s^2) \tag{A.14}$$

We start by proving (A.13), for which it suffices to show that $\operatorname{Var}\big(\sqrt{kh}\{e_{n,S}^+(y^+,s) - e_{n,S}^-(y^-,s)\}\big) = o(1)$. For $n \in \mathbb{N}$ and $j \in \{1,\dots,m\}$, let $V_{n,j} = (X'_{n,t})_{t\in I_j^S}$, and note that $e_{n,S}^\pm$ is a function of $(V_{n,j})_{j=1,\dots,m}$. Further, let $(V_{n,j}^*)_{j=1,\dots,m}$ denote an i.i.d. sequence, where $V_{n,j}^*$ is equal in distribution to $V_{n,j}$. Finally, let $e_{n,S}^{\pm,*}$ be defined as $e_{n,S}^\pm$, but in terms of $(V_{n,j}^*)_{j=1,\dots,m}$ instead of $(V_{n,j})_{j=1,\dots,m}$. First, we show the assertion in (A.13) with $e_{n,S}^\pm$ replaced by $e_{n,S}^{\pm,*}$. By independence of blocks, we may write

$$\begin{aligned}
&\operatorname{Var}\big(\sqrt{kh}\{e_{n,S}^{+,*}(y^+,s) - e_{n,S}^{-,*}(y^-,s)\}\big) \\
&\le \frac{1}{kh}\sum_{j=1}^m \mathrm{E}\Bigg[\Bigg\{K_b^+\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j^S}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^+\big) \\
&\qquad\qquad - K_b^-\Big(\frac{s-j/m}{h}, s\Big)\sum_{t\in I_j^S}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^-\big)\Bigg\}^2\Bigg] \\
&\le \frac{1}{kh}\sum_{j=1}^m \Bigg\{K_b^+\Big(\frac{s-j/m}{h}, s\Big)^2 \mathrm{E}\Bigg[\Bigg\{\sum_{t\in I_j^S}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^+\big)\Bigg\}^2\Bigg] \\
&\qquad\qquad + K_b^-\Big(\frac{s-j/m}{h}, s\Big)^2 \mathrm{E}\Bigg[\Bigg\{\sum_{t\in I_j^S}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^-\big)\Bigg\}^2\Bigg]\Bigg\}
\end{aligned}$$

$$\leq \frac{1}{kh} \operatorname{E}\Big[\Big\{\sum_{t\in I_1^S} \mathbf{1}(X'_{n,t}\neq 0)\Big\}^2\Big] \sum_{j=1}^m K_b^+\Big(\frac{s-j/m}{h},s\Big)^2 + K_b^-\Big(\frac{s-j/m}{h},s\Big)^2, \tag{A.15}$$

where the last step is due to stationarity. As in (8.11), we have

$$\operatorname{E}\Big[\Big\{\sum_{t\in I_1^S} \mathbf{1}(X'_{n,t}\neq 0)\Big\}^2\Big] \leq \frac{\ell_n}{r} \operatorname{E}\Big[\Big\{\sum_{t=1}^r \mathbf{1}(X'_{n,t}\neq 0)\Big\}^2\Big] = O\Big(\frac{\ell_n k}{n}\Big),$$

where the last bound follows from Condition (B4). As a consequence, the expression in (A.15) can be bounded by

$$O\Big(\frac{\ell_n m}{n}\Big)\frac{1}{mh}\sum_{j=1}^m K_b^+\Big(\frac{s-j/m}{h},s\Big)^2 + K_b^-\Big(\frac{s-j/m}{h},s\Big)^2$$
$$= O\Big(\frac{\ell_n}{r}\Big)\Big\{\int_{\frac{s-1}{h}}^{s/h} K_b^+(x,s)^2 + K_b^-(x,s)^2 \,\mathrm{d}x + O\Big(\frac{1}{mh}\Big)\Big\}, \tag{A.16}$$

which converges to zero due to $\ell_n = o(r)$ and $mh\to\infty$, since $kh\to\infty$ and $m\gg k$ by $r=o(n/k)$ in Condition (B7) and (B3), respectively. Hence, $\sqrt{kh}\{e^{+,*}_{n,S}(y^+,s) - e^{-,*}_{n,S}(y^-,s)\} = o_{\mathrm{P}}(1)$. The same argumentation that was used in the proof of Proposition 6.1 can be used to deduce (A.13).

It remains to prove (A.14). For that purpose, write

$$H_{n2} = \sqrt{kh}\{e^+_{n,B}(y^+,s) - e^-_{n,B}(y^-,s)\} = \frac{1}{\sqrt{kh}}\sum_{j=1}^m f_{j,n}(s),$$

where, for $j\in\{1,\dots,m\}$,

$$f_{j,n}(s) = K_b^+\Big(\frac{s-j/m}{h},s\Big)\sum_{t\in I_j^B}\Big\{\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^+\big) - \mathrm{P}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^+\big)\Big\}$$
$$-K_b^-\Big(\frac{s-j/m}{h},s\Big)\sum_{t\in I_j^B}\Big\{\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^-\big) - \mathrm{P}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^-\big)\Big\}.$$

Note that $f_{j,n}(s)$ is centered and depends on the block I_j^B only, such that the observations making up $f_{j,n}(s)$ and $f_{i,n}(s)$ are separated by at least ℓ_n observations for $j\neq i$. By the same arguments given in the proof of Proposition 6.1 we can assume that $f_{1,n}(s),\dots,f_{m,n}(s)$ are independent. As a consequence, we may apply the classical Lindeberg Central Limit Theorem. The Lindeberg condition is satisfied, if for any $\varepsilon>0$,

$$\lim_{n\to\infty}\frac{1}{kh}\sum_{j=1}^m \operatorname{E}\Big[f_{j,n}(s)^2\mathbf{1}\big(|f_{j,n}(s)|>\varepsilon\sqrt{kh}\big)\Big] = 0.$$

Since $|f_{j,n}(s)| \lesssim r-\ell_n \leq r$, the Lindeberg condition already follows from the assumption $r=o(\sqrt{kh})$ in (B7), see Corollary 3.6 in Drees and Rootzén (2010) for a similar argumentation.

It remains to prove that $\lim_{n\to\infty}\operatorname{Var}(H_{n2}) = \sigma_s^2$. Let

$$d_{n,j}(y) = \sum_{t\in I_j^B}\mathbf{1}\big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y\big),$$

such that

$$\begin{aligned}\operatorname{Var}(H_{n2}) &= \frac{1}{kh}\sum_{j=1}^{m}\operatorname{Var}\Big(K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big)\\ &= \frac{1}{kh}\sum_{j=1}^{m}\operatorname{E}\Big[\Big\{K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big\}^2\Big]\\ &\qquad - \operatorname{E}\Big[K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big]^2\\ &=: A_n - B_n.\end{aligned}$$

By stationarity,

$$|\operatorname{E}[d_{n,j}(y^\pm)]| \le \sum_{t\in I_j^B}\operatorname{P}(X'_{n,t}\neq 0) = (r-\ell_n)\frac{k}{n}c_\infty(L),$$

implying

$$\begin{aligned}|B_n| &\le \frac{1}{kh}\sum_{j=1}^{m}\Big\{K_b^+\Big(\frac{s-j/m}{h},s\Big)^2\operatorname{E}[d_{n,j}(y^+)]^2 + K_b^-\Big(\frac{s-j/m}{h},s\Big)^2\operatorname{E}[d_{n,j}(y^-)]^2\Big\}\\ &\le c_\infty(L)^2\frac{k}{m}\frac{1}{mh}\sum_{j=1}^{m}K_b^+\Big(\frac{s-j/m}{h},s\Big)^2 + K_b^-\Big(\frac{s-j/m}{h},s\Big)^2,\end{aligned}$$

which converges to zero by the previous calculation in (A.16) and $k/m = o(1)$ by Condition (B3). As a consequence,

$$\operatorname{Var}(H_{n2}) = A_n + o(1). \tag{A.17}$$

Next, write $A_n = m^{-1}\sum_{j=1}^m A_n(j)$, where

$$A_n(j) = \frac{m}{kh}\operatorname{E}\Big[\Big\{K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big\}^2\Big].$$

Recall the definition of the length of the core of a cluster y, denoted by $L(y)$, see Definition 2.1 in Drees and Rootzén (2010). For some constant $K>0$, writing $Y_{n,j} = (X'_{n,t})_{t\in I_j^B}$ for $j\in\{1,\dots,m\}$, we may then decompose $A_n(j) = S_{n,K}(j) + R_{n,K}(j)$, where

$$\begin{aligned}&S_{n,K}(j)\\ &= \frac{m}{kh}\operatorname{E}\Big[\Big\{K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big\}^2\mathbf{1}\big(L(Y_{n,j})\le K\big)\Big],\\ &R_{n,K}(j)\\ &= \frac{m}{kh}\operatorname{E}\Big[\Big\{K_b^+\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^+) - K_b^-\Big(\frac{s-j/m}{h},s\Big)d_{n,j}(y^-)\Big\}^2\mathbf{1}\big(L(Y_{n,j})> K\big)\Big].\end{aligned}$$

We have

$$\begin{aligned}R_{n,K}(j) &\le \frac{m}{kh}K_b^+\Big(\frac{s-j/m}{h},s\Big)^2\operatorname{E}\Big[d_{n,j}^2(y^+)\mathbf{1}\big(L(Y_{n,j})>K\big)\Big]\\ &\quad + \frac{m}{kh}K_b^-\Big(\frac{s-j/m}{h},s\Big)^2\operatorname{E}\Big[d_{n,j}^2(y^-)\mathbf{1}\big(L(Y_{n,j})>K\big)\Big].\end{aligned}$$

The two summands summands on the right-hand side can be written as

$$\frac{m}{kh}K_b^{\pm}\Big(\frac{s-j/m}{h},s\Big)^2\sum_{q,t\in I_j^B}\mathrm{P}\Big(X'_{n,t}>c_\infty-c(\tfrac{t}{n})y^{\pm},$$
$$X'_{n,q}>c_\infty-c(q/n)y^{\pm},L(Y_{n,j})>K\Big)$$
$$\leq\frac{m}{kh}K_b^{\pm}\Big(\frac{s-j/m}{h},s\Big)^2\sum_{q,t\in I_j^B}\mathrm{P}\Big(X'_{n,t}>0,X'_{n,q}>0,L(Y_{n,j})>K\Big)$$
$$=\frac{m}{kh}K_b^{\pm}\Big(\frac{s-j/m}{h},s\Big)^2\,\mathrm{E}\Big[\Big\{\sum_{t=1}^{r}\mathbf{1}(X'_{n,t}>0)\Big\}^2\mathbf{1}\big(L(Y_{n,1})>K\big)\Big]$$
$$\leq\frac{1}{h}K_b^{\pm}\Big(\frac{s-j/m}{h},s\Big)^2\Big\{\frac{m}{k}\,\mathrm{E}\Big[\Big\{\sum_{t=1}^{r}\mathbf{1}(X'_{n,t}>0)\Big\}^{2+\delta}\Big]\Big\}^{\frac{2}{2+\delta}}\Big\{\frac{m}{k}\mathrm{P}\big(L(Y_{n,1})>K\big)\Big\}^{\frac{\delta}{2+\delta}}$$

by Hölder's inequality. Consequently, we obtain

$$\frac{1}{m}\sum_{j=1}^{m}R_{n,K}(j)\leq\Big\{\frac{m}{k}\,\mathrm{E}\Big[\Big\{\sum_{t=1}^{r}\mathbf{1}(X'_{n,t}>0)\Big\}^{2+\delta}\Big]\Big\}^{\frac{2}{2+\delta}}\Big\{\frac{m}{k}\mathrm{P}\big(L(Y_{n,1})>K\big)\Big\}^{\frac{\delta}{2+\delta}}$$
$$\times\frac{1}{mh}\sum_{j=1}^{m}K_b^{+}\Big(\frac{s-j/m}{h},s\Big)^2+K_b^{-}\Big(\frac{s-j/m}{h},s\Big)^2.$$

By Condition (B4) and Lemma 5.2 (vii) in Drees and Rootzén (2010), which is applicable by Condition (B3) (note that their v_n is $v_n=\Pr(X'_{n,i}\neq 0)=\frac{k}{n}c_\infty(2)$ in our notation), we have

$$\lim_{K\to\infty}\limsup_{n\to\infty}\frac{1}{m}\sum_{j=1}^{m}R_{n,K}(j)=0.\tag{A.18}$$

Next, consider the term $S_{n,K}(j)$ with $j\in\{1,\dots,m\}$, which may be written as

$$S_{n,K}(j)$$
$$=\frac{m}{kh}\Big\{K_b^{+}\Big(\frac{s-j/m}{h},s\Big)^2\,\mathrm{E}\Big[d_{n,j}^2(y^+)\mathbf{1}\big(L(Y_{n,j})\leq K\big)\Big]$$
$$+K_b^{-}\Big(\frac{s-j/m}{h},s\Big)^2\,\mathrm{E}\Big[d_{n,j}^2(y^-)d_{n,j}(y^-)\mathbf{1}\big(L(Y_{n,j})\leq K\big)\Big]$$
$$-2K_b^{+}\Big(\frac{s-j/m}{h},s\Big)K_b^{-}\Big(\frac{s-j/m}{h},s\Big)\,\mathrm{E}\Big[d_{n,j}(y^+)d_{n,j}(y^-)\mathbf{1}\big(L(Y_{n,j})\leq K\big)\Big]\Big\}$$
$$=\frac{m}{kh}\Big\{K_b^{+}\Big(\frac{s-j/m}{h},s\Big)^2\sum_{q,t\in I_j^B,|q-t|\leq K}\mathrm{P}\Big(X'_{n,t}>c_\infty-c(\tfrac{t}{n})y^{+},$$
$$X'_{n,q}>c_\infty-c(\tfrac{q}{n})y^{+},L(Y_{n,j})\leq K\Big)$$
$$+K_b^{-}\Big(\frac{s-j/m}{h},s\Big)^2\sum_{q,t\in I_j^B,|q-t|\leq K}\mathrm{P}\Big(X'_{n,t}>c_\infty-c(\tfrac{t}{n})y^{-},$$
$$X'_{n,q}>c_\infty-c(\tfrac{q}{n})y^{-},L(Y_{n,j})\leq K\Big)$$

$$- 2K_b^+\Big(\frac{s-j/m}{h},s\Big)K_b^-\Big(\frac{s-j/m}{h},s\Big) \sum_{q,t\in I_j^B,|q-t|\le K} \mathrm{P}\Big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y^+, X'_{n,q} > c_\infty - c(\tfrac{q}{n})y^-, L(Y_{n,j}) \le K\Big)\Big\}.$$

Let $S'_{n,K}(j)$ be defined exactly as the right-hand side of the previous display, but with the probability terms replaced by

$$\mathrm{P}\Big(X'_{n,t} > c_\infty - c(\tfrac{t}{n})y, X'_{n,q} > c_\infty - c(\tfrac{q}{n})y'\Big).$$

for $y, y' \in \{y^+, y^-\}$ (i.e., we omit the additional condition $L(Y_{n,j}) \le K$ everywhere). Further, let $R'_{n,K}(j) = S'_{n,K}(j) - S_{n,K}(j)$. By the same arguments that were used for $R_{n,K}$ above, one can show that

$$\lim_{K\to\infty} \limsup_{n\to\infty} \frac{1}{m}\sum_{j=1}^m R'_{n,K}(j) = 0. \tag{A.19}$$

Regarding the remaining terms $S'_{n,K}(j)$ we obtain, by uniform continuity of c and Lemma 9.1,

$$S'_{n,K}(j) = S''_{n,K}(j) + R''_n(j)o(1),$$

where the $o(1)$ is uniform in $j = 1, \dots, m$, where

$$R''_n(j) = \frac{1}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 + \frac{1}{h}K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 - \frac{2}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)$$

and where

$$\begin{aligned} S''_{n,K}(j) = {}& \frac{1}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 \frac{1}{r}\sum_{t\in I_j^B} c(\tfrac{t}{n})\Big\{d_0(y^+,y^+) + 2\sum_{q=1}^{K\wedge(jr-\ell_n-t)} d_q(y^+,y^+)\Big\} \\ &+ \frac{1}{h}K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 \frac{1}{r}\sum_{t\in I_j^B} c(\tfrac{t}{n})\Big\{d_0(y^-,y^-) + 2\sum_{q=1}^{K\wedge(jr-\ell_n-t)} d_q(y^-,y^-)\Big\} \\ &- \frac{2}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big) \\ &\quad \cdot \frac{1}{r}\sum_{t\in I_j^B} c(\tfrac{t}{n})\Big\{d_0(y^+,y^-) + \sum_{q=1}^{K\wedge(jr-\ell_n-t)} d_q(y^+,y^-) + d_q(y^-,y^+)\Big\}. \end{aligned}$$

By Lemma A.1 (and a straightforward extension of this lemma to the case of a product of kernels) we have $m^{-1}\sum_{j=1}^m R''_n(j) = O(1)$. Moreover,

$$\begin{aligned} S''_{n,K}(j) = {}& \frac{1}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 c(\tfrac{j}{m})D_K(y^+,y^+) + \frac{1}{h}K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 c(\tfrac{j}{m})D_K(y^-,y^-) \\ &- \frac{2}{h}K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)c(\tfrac{j}{m})D_K(y^+,y^-) + o(1), \end{aligned}$$

where the $o(1)$ is uniform in $j = 1, \dots, m$ (and in $y^+, y^- \in [0,2]$ as arbitrary inputs), and

$$D_K(x,x') = d_0(x,x') + \sum_{q=1}^K d_q(x,x') + d_q(x',x).$$

Therefore,

$$\frac{1}{m}\sum_{j=1}^{m} S'_{n,K}(j)$$
$$= \frac{1}{mh}\sum_{j=1}^{m} c(\tfrac{j}{m})\Big\{K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 D_K(y^+,y^+) + K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)^2 D_K(y^-,y^-)$$
$$- 2K_b^+\Big(\frac{s-\frac{j}{m}}{h},s\Big)K_b^-\Big(\frac{s-\frac{j}{m}}{h},s\Big)D_K(y^+,y^-)\Big\} + o(1),$$

which converges to $c(s)\eta(s)D_K(1,1)$ by a straightforward extension of Lemma A.1 to the case of a product of kernels. Further, note that $D_K(y,y')$ converges to $D_K(1,1)$ for $y,y' \in \{y^+,y^-\}$, since d_q is continuous in $(1,1)$ by Theorem 2 and its subsequent discussion in Segers (2003). Finally, since

$$A_n = \frac{1}{m}\sum_{j=1}^{m} A_n(j) = \frac{1}{m}\sum_{j=1}^{m} S'_{n,K}(j) - \frac{1}{m}\sum_{j=1}^{m} R'_{n,K}(j) + \frac{1}{m}\sum_{j=1}^{m} R_{n,K}(j)$$

Lemma 9.2 and (A.18) and (A.19) imply $\lim_{n\to\infty} \operatorname{Var}(H_{n2}) = \lim_{n\to\infty} A_n = \sigma_s^2$, where we used (A.17). □

Lemma A.1. *Assume $c \in C^2([0,1])$. Let K be a Lipschitz-continuous function on $[-1,1]$ with $K(x) = 0$ for $|x| > 1$. Further, let $h = h_n > 0$ satisfy $h \to 0$ and $nh \to \infty$ for $n \to \infty$. Then, for any $s \in [0,1]$, as $n \to \infty$,*

$$\frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{s-i/n}{h}\Big)c(i/n) = c(s)\eta_1(s) - hc'(s)\eta_2(s) + \frac{h^2}{2}c''(s)\eta_3(s) + o(h^2) + O\Big(\frac{1}{nh}\Big),$$
$$\frac{1}{nh}\sum_{i=1}^{n} K^2\Big(\frac{s-i/n}{h}\Big)c(i/n) = c(s)\eta_4(s) + O(h) + O\Big(\frac{1}{nh}\Big).$$

where

$$\eta_1(s) = \mathbf{1}(s\le h)\int_{-1}^{s/h} K(x)\mathrm{d}x + \mathbf{1}(h<s<1-h)\int_{-1}^{1} K(x)\mathrm{d}x + \mathbf{1}(s\ge 1-h)\int_{\frac{s-1}{h}}^{1} K(x)\mathrm{d}x,$$
$$\eta_2(s) = \mathbf{1}(s\le h)\int_{-1}^{s/h} K(x)x\mathrm{d}x + \mathbf{1}(h<s<1-h)\int_{-1}^{1} K(x)x\mathrm{d}x + \mathbf{1}(s\ge 1-h)\int_{\frac{s-1}{h}}^{1} K(x)x\mathrm{d}x,$$
$$\eta_3(s) = \mathbf{1}(s\le h)\int_{-1}^{s/h} K(x)x^2\mathrm{d}x + \mathbf{1}(h<s<1-h)\int_{-1}^{1} K(x)x^2\mathrm{d}x + \mathbf{1}(s\ge 1-h)\int_{\frac{s-1}{h}}^{1} K(x)x^2\mathrm{d}x,$$
$$\eta_4(s) = \mathbf{1}(s\le h)\int_{-1}^{s/h} K^2(x)\mathrm{d}x + \mathbf{1}(h<s<1-h)\int_{-1}^{1} K^2(x)\mathrm{d}x + \mathbf{1}(s\ge 1-h)\int_{\frac{s-1}{h}}^{1} K^2(x)\mathrm{d}x.$$

Proof. A Riemann sum approximation implies

$$\frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{s-i/n}{h}\Big)c(i/n) = \int_{\frac{s-1}{h}}^{s/h} K(x)c(s-hx)\,\mathrm{d}x + O\Big(\frac{1}{nh}\Big).$$

Next, by Taylor's theorem, there exists some $\tau_x \in [0,1]$ such that

$$\int_{\frac{s-1}{h}}^{s/h} K(x)c(s-hx)\,\mathrm{d}x$$

$$= c(s)\int_{\frac{s-1}{h}}^{s/h} K(x)\,\mathrm{d}x - hc'(s)\int_{\frac{s-1}{h}}^{s/h} K(x)x\,\mathrm{d}x + \frac{h^2}{2}c''(s)\int_{\frac{s-1}{h}}^{s/h} K(x)x^2\,\mathrm{d}x + \frac{h^2}{2}\int_{\frac{s-1}{h}}^{s/h} K(x)x^2\big\{c''(s - xh\tau_x) - c''(s)\big\}\,\mathrm{d}x.$$

Since K has compact support and c'' is continuous, the dominated convergence theorem implies that the last integral is of the order $o(h^2)$. If $s < 1-h$, we have $\frac{s-1}{h} < -1$, and for $s > h$, we have $s/h > 1$. This allows to rewrite the boundaries of the integral accordingly in view of the fact that K has support $[-1,1]$.

For the second assertion, write

$$\begin{aligned}
&\frac{1}{nh}\sum_{i=1}^{n} K^2\Big(\frac{s-i/n}{h}\Big)c(i/n) \\
&= \int_{\frac{s-1}{h}}^{s/h} K^2(x)c(s-hx)\,\mathrm{d}x + O\Big(\frac{1}{nh}\Big) \\
&= c(s)\int_{\frac{s-1}{h}}^{s/h} K^2(x)\,\mathrm{d}x + \int_{\frac{s-1}{h}}^{s/h} K^2(x)\big\{c(s-hx) - c(s)\big\}\,\mathrm{d}x + O\Big(\frac{1}{nh}\Big) \\
&= c(s)\int_{\frac{s-1}{h}}^{s/h} K^2(x)\,\mathrm{d}x + O(h) + O\Big(\frac{1}{nh}\Big),
\end{aligned}$$

where the last step is again due to the dominated convergence theorem. □

Proof of Theorem 3.2. As at the beginning of the proof of Theorem 3.1, let $y_n = nk^{-1}\big\{1 - F(X_{n,n-k})\big\}$. The definition of the STEP $\mathbb{F}_n$ in (6.2) allows to write

$$\sqrt{k}\{\hat{C}_n(s) - C(s)\} = \mathbb{F}_n(s, y_n) + C(s)\sqrt{k}(y_n - 1)$$

By the proof of Theorem 3.1, see (A.1), we know that

$$\sqrt{k}(y_n - 1) = -\mathbb{F}_n(1,1) + o_{\mathrm{P}}(1).$$

Suppose we have shown that

$$\sup_{s\in[0,1]} |\mathbb{F}_n(s, y_n) - \mathbb{F}_n(s,1)| = o_{\mathrm{P}}(1). \tag{A.20}$$

Then, by the previous three displays, uniformly in s,

$$\sqrt{k}\{\hat{C}_n(s) - C(s)\} = \mathbb{F}_n(s,1) - C(s)\mathbb{F}_n(1,1) + o_{\mathrm{P}}(1), \tag{A.21}$$

which implies the assertion since $\{\mathbb{F}_n(s,1)\}_{s\in[0,1]} \rightsquigarrow \{\mathbb{S}(s,1)\}_{s\in[0,1]}$ in $(\ell^\infty([0,1]), \|\cdot\|_\infty)$ by Proposition 6.2. It remains to prove (A.20). Note that

$$\sup_{s\in[0,1]} |\mathbb{F}_n(s,y_n) - \mathbb{F}_n(s,1)| \le \sup_{(s,y),(s,z)\in[0,1]^2: |y-z|<|y_n-1|} |\mathbb{F}_n(s,y) - \mathbb{F}_n(s,z)|,$$

For any $\varepsilon > 0$ and $\mu \in (0, 1/2)$, we have $\mathrm{P}(k^\mu|y_n - 1| < \varepsilon) \to 1$. Thus, on this event the above supremum can be bounded by

$$\sup_{(s,y),(s,z)\in[0,1]^2: |y-z|<\delta_n} |\mathbb{F}_n(s,y) - \mathbb{F}_n(s,z)|.$$

where $\delta_n := \varepsilon k^{-\mu} \downarrow 0$. Analogously to showing (8.19) in the proof of Proposition 6.2, we obtain that the last expression is asymptotically negligible (note that the same semimetric used in the proof of Proposition 6.2 can be applied here by Theorem 7.19 in Kosorok (2008) and Proposition 6.2 and the proof of tightness in the proof of Proposition 6.1, which made Theorem 11.16 in Kosorok (2008) applicable for $\mathbb{S}_n$). □

Appendix B. Proofs for Section 4

For $b \in \mathbb{N}$ and $(s, x) \in [0, 1]^2$ let

$$\mathbb{S}_{n,\xi}^{(b)}(s, x) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \xi_j^{(b)} \sum_{i \in I_j} \mathbf{1}\Big(U_i^{(n)} > 1 - \frac{kx}{n} c(\tfrac{i}{n})\Big)\mathbf{1}(\tfrac{i}{n} \le s), \tag{B.1}$$

$$\mathbb{F}_{n,\xi}^{(b)}(s, x) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \xi_j^{(b)} \sum_{i \in I_j} \mathbf{1}\Big(X_i^{(n)} > V\Big(\frac{n}{kx}\Big)\Big)\mathbf{1}(\tfrac{i}{n} \le s)$$

denote *bootstrap-versions* of the (simple) STEP defined in (6.1) and (6.2).

Proposition B.1. *Suppose that Conditions (B0)-(B6) hold for $L = 1$. Then, for any $B \in \mathbb{N}$ and as $n \to \infty$,*

$$\big(\mathbb{S}_n, \mathbb{S}_{n,\xi}^{(1)}, \dots, \mathbb{S}_{n,\xi}^{(B)}\big) \rightsquigarrow \big(\mathbb{S}, \mathbb{S}^{(1)}, \dots, \mathbb{S}^{(B)}\big) \quad \text{in} \quad \big(\ell^\infty([0, 1]^2), \|\cdot\|_\infty\big)^{B+1},$$

where $\mathbb{S}^{(1)}, \dots, \mathbb{S}^{(B)}$ are independent copies of $\mathbb{S}$ from Proposition 6.1.

Proposition B.2. *Suppose that Conditions (B0)-(B6) hold for $L = 1$. Then, for any $b \in \mathbb{N}$ and as $n \to \infty$,*

$$\sup_{(s,x)\in[0,1]^2} |\mathbb{F}_{n,\xi}^{(b)}(s, x) - \mathbb{S}_{n,\xi}^{(b)}(s, x)| = o_{\mathrm{P}}(1).$$

As a consequence, by Proposition 6.2 and B.1, for any $B \in \mathbb{N}$ and as $n \to \infty$,

$$\big(\mathbb{F}_n, \mathbb{F}_{n,\xi}^{(1)}, \dots, \mathbb{F}_{n,\xi}^{(B)}\big) \rightsquigarrow \big(\mathbb{S}, \mathbb{S}^{(1)}, \dots, \mathbb{S}^{(B)}\big) \quad \text{in} \quad \big(\ell^\infty([0, 1]^2), \|\cdot\|_\infty\big)^{B+1},$$

where $\mathbb{S}^{(1)}, \dots, \mathbb{S}^{(B)}$ are independent copies of $\mathbb{S}$ from Proposition 6.1.

Proof of Theorem 4.1. Define

$$\tilde{\mathbb{C}}_{n,\xi}^{(b)}(s) = \tilde{\mathbb{D}}_{n,\xi}^{(b)}(s) - \hat{C}_n(s)\tilde{\mathbb{D}}_{n,\xi}^{(b)}(1),$$

where

$$\tilde{\mathbb{D}}_{n,\xi}^{(b)}(s) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \xi_j^{(b)} \sum_{i \in I_j} \mathbf{1}\Big(X_i^{(n)} > X_{n,n-k}\Big)\mathbf{1}(\tfrac{i}{n} \le s), \quad s \in [0, 1].$$

Recall that $y_n = nk^{-1}\big(1 - F(X_{n,n-k})\big)$ converges to 1 in probability by (A.1). For $b \in \mathbb{N}$ and $s \in [0, 1]$, we have

$$\begin{aligned}
\tilde{\mathbb{C}}_{n,\xi}^{(b)}(s) &= \mathbb{F}_{n,\xi}^{(b)}(s, y_n) - \hat{C}_n(s)\mathbb{F}_{n,\xi}^{(b)}(1, y_n) \\
&= \mathbb{F}_{n,\xi}^{(b)}(s, 1) - C(s)\mathbb{F}_{n,\xi}^{(b)}(1, 1) \\
&= \mathbb{S}_{n,\xi}^{(b)}(s, 1) - C(s)\mathbb{S}_{n,\xi}^{(b)}(1, 1)
\end{aligned}$$

where the third equality is a consequence of Proposition B.2 and where the second equality is a consequence of $\sup_{s\in[0,1]} |\hat{C}_n(s) - C(s)| = o_{\mathrm{P}}(1)$ by Theorem 3.2 and

$$\sup_{s\in[0,1]} |\mathbb{F}_{n,\xi}^{(b)}(s, y_n) - \mathbb{F}_{n,\xi}^{(b)}(s, 1)| = o_{\mathrm{P}}(1),$$

which can be seen to hold by the same argumentation as in the proof of Theorem 3.2 for showing (A.20), since $\mathbb{F}_{n,\xi}^{(b)} \rightsquigarrow \mathbb{S}$ in $\big(\ell^\infty([0,1]^2), \|\cdot\|_\infty\big)$ by Proposition B.2. Hence, (A.21) and Proposition 6.2 imply the representation

$$\Big(\sqrt{k}(\hat{C}_n - C), \tilde{\mathbb{C}}_{n,\xi}^{(1)}, \dots, \tilde{\mathbb{C}}_{n,\xi}^{(B)}\Big) = \Big(\mathbb{S}_n(s,1) - C(s)\mathbb{S}_n(1,1), \\ \mathbb{S}_{n,\xi}^{(1)}(s,1) - C(s)\mathbb{S}_{n,\xi}^{(1)}(1,1), \dots, \mathbb{S}_{n,\xi}^{(B)}(s,1) - C(s)\mathbb{S}_{n,\xi}^{(B)}(1,1)\Big) + o_{\mathrm{P}}(1).$$

By Proposition B.1 and the continuous mapping theorem, the previous expression weakly converges to

$$\Big(\mathbb{S}(s,1) - C(s)\mathbb{S}(1,1), \mathbb{S}^{(1)}(s,1) - C(s)\mathbb{S}^{(1)}(1,1), \dots, \mathbb{S}^{(B)}(s,1) - C(s)\mathbb{S}^{(B)}(1,1)\Big) \\ = \Big(\mathbb{C}, \mathbb{C}^{(1)}, \dots, \mathbb{C}^{(B)}\Big)$$

in $(\ell^\infty([0,1]), \|\cdot\|_\infty)^{B+1}$. Finally, since

$$\mathbb{D}_{n,\xi}^{(b)}(s) = \tilde{\mathbb{D}}_{n,\xi}^{(b)}(s) - \sqrt{k}\bar{\xi}^{(b)}\hat{C}_n(s)$$

and $\hat{C}_n(1) = 1$, we have $\mathbb{C}_{n,\xi}^{(b)} = \tilde{\mathbb{C}}_{n,\xi}^{(b)}$, which proves the theorem. □

Proof of Corollary 4.2. By Theorem 4.1 and the Continuous Mapping Theorem, we have that, under H_0, as $n \to \infty$,

$$\big(S_{n,1}, S_{n,1}^{(1)}, \dots, S_{n,1}^{(B)}\big) \rightsquigarrow \big(\|\mathbb{C}\|_\infty, \|\mathbb{C}^{(1)}\|_\infty, \dots, \|\mathbb{C}^{(B)}\|_\infty\big),$$
$$\big(T_{n,1}, T_{n,1}^{(1)}, \dots, T_{n,1}^{(B)}\big) \rightsquigarrow \Big(\int_0^1 \mathbb{C}(s)^2 \,\mathrm{d}s, \int_0^1 \mathbb{C}^{(1)}(s)^2 \,\mathrm{d}s, \dots, \int_0^1 \mathbb{C}^{(B)}(s)^2 \,\mathrm{d}s\Big).$$

Note that $\mathbb{C} = \sigma\mathbb{B}$ in distribution by (4.1), where $\mathbb{B}$ is a Brownian Bridge on $[0,1]$, which implies that $\|\mathbb{C}\|_\infty$ and $\int_0^1 \mathbb{C}(s)^2 \,\mathrm{d}s$ are continuous random variables. Further note that $(\xi_1^{(b)}, \dots, \xi_m^{(b)})_{b=1,\dots,B}$ are i.i.d. Overall, Lemma 4.2 in Bücher and Kojadinovic (2019) is applicable, which proves the assertion under H_0. For the assertion under H_1 let us consider $\varphi_{n,B,S}$; $\varphi_{n,B,T}$ can be treated analogously. Note that under H_1,

$$k^{-1/2} S_{n,1} = \sup_{s\in[0,1]} |\hat{C}_n(s) - s| \xrightarrow{\mathbb{P}} \sup_{s\in[0,1]} |C(s) - s| > 0$$

and $r^{-1/2} S_{n,1}^{(b)} = r^{-1/2}\|\mathbb{C}_{n,\xi}^{(b)}\|_\infty = o_{\mathrm{P}}(1)$, such that $S_{n,1}^{(b)} = O_{\mathrm{P}}(r^{1/2})$ for any $b \in \{1, \dots, B\}$. The claim follows since $r = o(k)$ by Condition (B3). □

Proof of Corollary 4.3. The proof of the statement regarding the null hypothesis is an immediate consequence of Theorem 4.1. Under H_1, one can easily show that $S_{n,2}, T_{n,2}$ converge to ∞ in probability, which implies the respective assertion under H_1. □

Proof of Proposition B.1. Fix $b \in \{1, \dots, B\}$. We only show weak convergence of $(\mathbb{S}_n, \mathbb{S}_{n,\xi}^{(b)})$; the joint weak convergence of all $B+1$ components can be shown analogously. In the following, we omit the upper index (b) at all instances. Recall $c_\infty(L) = 1 + L\|c\|_\infty$ and $X'_{n,i}$ from (8.1) and $v_n = \Pr(X'_{n,i} \neq 0) = \frac{k}{n}c_\infty(L)$. For $(s,x) \in [0,1]^2$, write

$$\mathbb{S}_{n,\xi}(s,x) = c_\infty(L)^{1/2}\tilde{\mathbb{Z}}_{n,\xi}(s,x) + R_{n,\xi}(s,x), \tag{B.2}$$

where

$$\tilde{\mathbb{Z}}_{n,\xi}(s,x) = \frac{1}{\sqrt{nv_n}}\sum_{j=1}^m \xi_j \sum_{i\in I_j} \mathbf{1}(\tfrac{i}{n}\le s)\Big\{\mathbf{1}(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x) - \Pr(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x)\Big\},$$

$$R_{n,\xi}(s,x) = \frac{x\sqrt{k}}{n}\sum_{j=1}^m \xi_j \sum_{i\in I_j} \mathbf{1}(\tfrac{i}{n}\le s)c(\tfrac{i}{n}).$$

First, we show that $R_{n,\xi} = o_{\mathrm{P}}(1)$. Due to Condition (B3) one can easily show that it suffices to prove $\tilde{R}_{n,\xi} = o_{\mathrm{P}}(1)$, where

$$\tilde{R}_{n,\xi}(s) = \sum_{j=1}^m f_{n,j}(s), \quad f_{n,j}(s) = \frac{\sqrt{k}}{n}\xi_j\mathbf{1}(\tfrac{j}{m}\le s)\sum_{i\in I_j} c(\tfrac{i}{n}), \quad s\in[0,1].$$

First, for $s\in[0,1]$, we have $\mathrm{E}[\tilde{R}_{n,\xi}(s)] = 0$ and

$$\mathrm{Var}\left(\tilde{R}_{n,\xi}(s)\right) = \frac{k}{n^2}\sum_{j=1}^m \Big\{\mathbf{1}(\tfrac{j}{m}\le s)\sum_{i\in I_j} c(\tfrac{i}{n})\Big\}^2 \le \|c\|_\infty^2 \frac{kr^2m}{n^2} = \|c\|_\infty^2\frac{kr}{n} = o(1)$$

by Condition (B3), such that $\tilde{R}_{n,\xi}(s) = o_{\mathrm{P}}(1)$ for any fixed $s\in[0,1]$. It remains to show tightness of $\tilde{R}_{n,\xi}$. To this, we will apply Lemma A.1 from Kley et al. (2016) with $\psi(x) = x^2$, $\bar{\eta} = 2/m$, $T=[0,1]$ and $d(s,t) = |s-t|$. Note that the Orlicz-norm with $\psi(x) = x^2$ coincides with the L_2-norm $\|\cdot\|_2$. First, for all $|s-t|\ge\bar{\eta}/2 = 1/m$, we have

$$\|\tilde{R}_{n,\xi}(s) - \tilde{R}_{n,\xi}(t)\|_2 \le 2\|c\|_\infty\frac{kr}{n}|s-t| \le |s-t|$$

for sufficiently large n by Condition (B3). By Lemma A.1 in Kley et al. (2016), for any $\delta>0, \eta\ge\bar{\eta}$, there exists a random variable S' and a constant $K'>\infty$, such that

$$\mathrm{P}\Big(\sup_{d(s,t)<\delta}|\tilde{R}_{n,\xi}(s) - \tilde{R}_{n,\xi}(t)| > 2\varepsilon\Big) \le \mathrm{P}(|S'|>\varepsilon) + \mathrm{P}\Big(\sup_{d(s,t)\le\bar{\eta}}|\tilde{R}_{n,\xi}(s) - \tilde{R}_{n,\xi}(t)| > \varepsilon/2\Big), \tag{B.3}$$

for all $\varepsilon>0$, where

$$\mathrm{P}(|S'|>\varepsilon) \le \Big(\frac{8K'}{\varepsilon}\Big)^2\Big(\int_{1/m}^\eta D(x,d)^{1/2}\mathrm{d}x + (\delta + 4/m)D(\eta,d)\Big)^2.$$

Here, $D(\cdot,d)$ denotes the *packing number* on $([0,1],d)$ and satisfies $D(x,d)\le 4x^{-1}+1$, $x>0$, see van der Vaart and Wellner (1996), page 98. Thus,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathrm{P}(|S'|>\varepsilon) \le \Big(\frac{8K'}{\varepsilon}\Big)^2\Big(\int_0^\eta (4x^{-1}+1)^{1/2}\mathrm{d}x\Big)^2. \tag{B.4}$$

Further, we have

$$|\tilde{R}_{n,\xi}(s) - \tilde{R}_{n,\xi}(t)| \le M\|c\|_\infty\frac{\sqrt{k}}{m}\sum_{j=1}^m |\mathbf{1}(j/m\le s) - \mathbf{1}(j/m\le t)|,$$

where $|\mathbf{1}(j/m\le s) - \mathbf{1}(j/m\le t)| = \mathbf{1}(s\wedge t < j/m \le s\vee t)$ does not equal zero for at most two different $j\in\{1,\dots,m\}$, if $d(s,t) = |s-t|\le\bar{\eta} = 2/m$. Consequently,

$$\mathrm{P}\Big(\sup_{d(s,t)\le\bar{\eta}}|\tilde{R}_{n,\xi}(s) - \tilde{R}_{n,\xi}(t)| > \varepsilon/2\Big) \le \mathbf{1}\Big(2M\|c\|_\infty\tfrac{\sqrt{k}}{m} > \varepsilon/2\Big),$$

which converges to zero as $n \to \infty$ since $rk = o(n)$ by Condition (B3). Altogether, by (B.3) and (B.4), we have shown, for all $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \mathrm{P}\Big(\sup_{d(s,t)<\delta} |\tilde R_{n,\xi}(s) - \tilde R_{n,\xi}(t)| > 2\varepsilon \Big) \le \Big(\frac{8K'}{\varepsilon}\Big)^2 \Big(\int_0^\eta (4x^{-1}+1)^{1/2} \mathrm{d}x \Big)^2,$$

which can be made arbitrarily small by choosing η accordingly. This concludes the proof of $R_{n,\xi} = o_{\mathrm{P}}(1)$.

By equation (B.2) we obtain $\mathbb{S}_{n,\xi} = c_\infty(L)^{1/2} \tilde{\mathbb{Z}}_{n,\xi} + o_{\mathrm{P}}(1)$. Since $|\xi_j| \le M$ the same calculation as in (8.4) in the proof of Proposition 6.1 (for treating $\mathbb{Z}_n$ and $\tilde{\mathbb{Z}}_n$) yields $\sup_{(s,x)\in[0,1]^2} |\tilde{\mathbb{Z}}_{n,\xi}(s,x) - \mathbb{Z}_{n,\xi}(s,x)| = o_{\mathrm{P}}(1)$ with

$$\begin{aligned} \mathbb{Z}_{n,\xi}(s,x) &= \frac{1}{\sqrt{nv_n}} \sum_{j=1}^m \xi_j \mathbf{1}(\tfrac{j}{m} \le s) \sum_{i\in I_j} \Big\{ \mathbf{1}(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x) \\ &\qquad\qquad\qquad - \Pr(X'_{n,i} > c_\infty(L) - c(\tfrac{i}{n})x) \Big\} \\ &= \frac{1}{\sqrt{nv_n}} \sum_{j=1}^m \xi_j \big\{ f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}[f_{j,n,s,x}(Y_{n,j})] \big\}, \end{aligned}$$

where $Y_{n,j} = (X'_{n,i})_{i\in I_j}$, $j = 1, \dots, m$, and $f_{j,n,s,x}$ is defined as in (8.3). Further, by (8.5) we know that $\mathbb{S}_n = c_\infty(L)^{1/2} \mathbb{Z}_n + o_{\mathrm{P}}(1)$, where

$$\mathbb{Z}_n(s,x) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^m \big\{ f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}[f_{j,n,s,x}(Y_{n,j})] \big\},$$

as defined after (8.2), the only difference to $\mathbb{Z}_{n,\xi}$ being the multipliers ξ_j. It remains to show weak convergence of $(\mathbb{Z}_n, \mathbb{Z}_{n,\xi})$. We start with the corresponding weak convergence of the fidis. Since $\xi_1, \dots, \xi_m$ are independent with $|\xi_j| \le M$ and independent of $Y_{n,1}, \dots, Y_{n,m}$ the proof is analogous to the one of Proposition 6.1. Let us just calculate the covariance function for independent blocks $Y_{n,1}, \dots, Y_{n,m}$. Note that $\mathrm{E}[\xi_j] = 0$ and $\mathrm{E}[\xi_j^2] = 1$. For $(s,x), (s',x') \in [0,1]^2$, we obtain

$$\frac{1}{nv_n} \sum_{j=1}^m \mathrm{Cov}\,\big(f_{j,n,s,x}(Y_{n,j}), \xi_j f_{j,n,s',x'}(Y_{n,j}) \big) = 0$$

and

$$\frac{1}{nv_n} \sum_{j=1}^m \mathrm{Cov}\,\big(\xi_j f_{j,n,s,x}(Y_{n,j}), \xi_j f_{j,n,s',x'}(Y_{n,j}) \big) = \frac{1}{nv_n} \sum_{j=1}^m \mathrm{E}\,\big[f_{j,n,s,x}(Y_{n,j}) f_{j,n,s',x'}(Y_{n,j}) \big],$$

which equals $\mathfrak{c}_n((s,x),(s',x'))$ defined in (8.7) and converges to $\mathfrak{c}((s,x),(s',x'))$ from Proposition 6.1 by the corresponding proof.

With regard to the asymptotic tightness, note that by Lemma 1.4.3 in van der Vaart and Wellner (1996) it suffices to show asymptotic tightness of $\mathbb{Z}_n$ and $\mathbb{Z}_{n,\xi}$ separately. Asymptotic tightness of $\mathbb{Z}_n$ has been shown in the proof of Proposition 6.1. Concerning the asymptotic tightness of $\mathbb{Z}_{n\xi}$, the proof follows analogously. Here, the conditions (1)-(5) in the proof of Proposition 6.1 can immediately be seen to hold since $|\xi_j| \le M$, and condition (6) follows since the function ρ_n is the same as before due to $\mathrm{E}[\xi_j^2] = 1$. □

Proof of Proposition B.2. Let $s, x \in [0,1]$ and $b \in \{1, \dots, B\}$. Set $\varepsilon_n(x) = V(n/(kx)) = F^{-1}(1 - kx/n)$ such that, almost surely,

$$\mathbb{F}_{n,\xi}^{(b)}(s,x) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \xi_j^{(b)} \sum_{i \in I_j} \mathbf{1}\Big\{U_i^{(n)} > 1 - \frac{kx}{n} \frac{1 - F_{n,i}(\varepsilon_n(x))}{1 - F(\varepsilon_n(x))}\Big\} \mathbf{1}(\tfrac{i}{n} \le s).$$

Note that $|\xi_j^{(b)}| \le M$. Then, by relation (8.17) and the definition of $\mathbb{S}_{n,\xi}^{(b)}$ and $\mathbb{S}_n$ in (B.1) and (6.1), respectively, we obtain

$$\begin{aligned}
&|\mathbb{F}_{n,\xi}^{(b)}(s,x) - \mathbb{S}_{n,\xi}^{(b)}(s,x)| \\
&\le \frac{M}{\sqrt{k}} \sum_{j=1}^{m} \sum_{i \in I_j} \mathbf{1}(\tfrac{i}{n} \le s) \Big| \mathbf{1}\Big(U_i^{(n)} > 1 - c(\tfrac{i}{n})(1+\delta_n)\frac{kx}{n}\Big) - \mathbf{1}\Big(U_i^{(n)} > 1 - c(\tfrac{i}{n})\frac{kx}{n}\Big) \Big| \\
&\qquad + \Big| \mathbf{1}\Big(U_i^{(n)} > 1 - c(\tfrac{i}{n})(1-\delta_n)\frac{kx}{n}\Big) - \mathbf{1}\Big(U_i^{(n)} > 1 - c(\tfrac{i}{n})\frac{kx}{n}\Big) \Big| \\
&= M\big\{\mathbb{S}_n(s, x(1+\delta_n)) + \sqrt{k}C(s)x(1+\delta_n) - (\mathbb{S}_n(s,x) + \sqrt{k}C(s)x) \\
&\qquad - (\mathbb{S}_n(s, x(1-\delta_n)) + \sqrt{k}C(s)x(1-\delta_n)) + \mathbb{S}_n(s,x) + \sqrt{k}C(s)x\big\} \\
&= M\big\{\mathbb{S}_n(s, x(1+\delta_n)) - \mathbb{S}_n(s, x(1-\delta_n)) + 2C(s)x\sqrt{k}\delta_n\big\},
\end{aligned}$$

where δ_n is defined after (8.17). Consequently,

$$\sup_{(s,x) \in [0,1]^2} |\mathbb{F}_{n,\xi}^{(b)}(s,x) - \mathbb{S}_{n,\xi}^{(b)}(s,x)| \le M w_{2\delta_n}(\mathbb{S}_n) + 2M\sqrt{k}\delta_n,$$

where $w_\delta(\mathbb{S}_n)$ is defined in (8.18) in the proof of Proposition 6.2. There, it is further shown that $w_{2\delta_n}(\mathbb{S}_n) = o_{\mathrm{P}}(1)$ and $\sqrt{k}\delta_n = o(1)$ by Condition (B6), which implies the assertion. □

Appendix C. Proofs for Section 5

Proof of Lemma 5.2. For $x > 0$ write

$$\begin{aligned}
\mathrm{P}(Z_{n,1+\lfloor \xi k' \rfloor} \ge x) &= \mathrm{P}\Big(\max_{i \in I'_{1+\lfloor \xi k' \rfloor}} F(X_i^{(n)}) \le 1 - x/q\Big) \\
&= \mathrm{P}\Big(X_i^{(n)} \le F^{-1}(1 - x/q) \ \text{ for all } \ i \in I'_{1+\lfloor \xi k' \rfloor}\Big) \\
&= \mathrm{P}\Big(Z_i \le \frac{1}{1 - F_{n,i}}\big(F^{-1}(1 - x/q)\big) \ \text{ for all } \ i \in I'_{1+\lfloor \xi k' \rfloor}\Big).
\end{aligned}$$

By Corollary D.2 the last term equals

$$\mathrm{P}\Big(Z_i \le \frac{q}{c(\xi)x} \ \text{ for all } \ i \in I'_{1+\lfloor \xi k' \rfloor}\Big) + o(1) = \mathrm{P}\Big(\max_{i \in I'_{1+\lfloor \xi k' \rfloor}} U_i \le 1 - \frac{c(\xi)x}{q}\Big) + o(1),$$

which converges to $\exp(-\theta c(\xi)x)$ by (5.1). □

Proof of Theorem 5.3. We start with part (a). By Theorem 3.1 we know that $\tilde{c}_n(x) = c(x) + o_{\mathrm{P}}(1)$ for any $x \in [0,1]$, and the continuous mapping theorem implies that $\hat{c}_n(x)^{-1} = \max(\tilde{c}_n(x), \kappa)^{-1} = c(x)^{-1} + o_{\mathrm{P}}(1)$ for any $x \in [0,1]$. Since $\hat{c}_n \ge \kappa$, we obtain $\mathrm{E}\big[|\hat{c}_n(x)^{-1}|^p\big] \le \kappa^{-p} < \infty$ for any $p > 0$ and $x \in [0,1]$. By Example 2.21 in van der Vaart (1998), this implies $\mathrm{E}\big[\hat{c}_n(x)^{-1}\big] \to c(x)^{-1}$ for any $x \in [0,1]$, such that

$$\mathrm{E}[\hat{\tau}_n] = \int_0^1 \mathrm{E}\big[\hat{c}_n(x)^{-1}\big] \,\mathrm{d}x \to \int_0^1 c(x)^{-1} \,\mathrm{d}x = \tau$$

by the dominated convergence theorem. Next, we show that $\mathrm{Var}(\hat\tau_n) = o(1)$. Note that $\hat c_n(x)^{-1}\hat c_n(y)^{-1} = c(x)^{-1}c(y)^{-1} + o_{\mathrm{P}}(1)$ for any $x, y \in [0,1]$ and $\mathrm{E}\left[|\hat c_n(x)^{-1}\hat c_n(y)^{-1}|^p\right] \le \kappa^{-2p}$ for any $p > 0$ and $x, y \in [0,1]$, such that as above $\mathrm{E}\left[\hat c_n(x)^{-1}\hat c_n(y)^{-1}\right] = c(x)^{-1}c(y)^{-1} + \mathrm{o}(1)$. Thus, by Fubini's theorem

$$\begin{aligned}\mathrm{Var}(\hat\tau_n) &= \int_0^1\int_0^1 \mathrm{E}\left[\hat c_n(x)^{-1}\hat c_n(y)^{-1}\right]\,\mathrm{d}x\mathrm{d}y - \Big(\int_0^1 \mathrm{E}\left[\hat c_n(x)^{-1}\right]\,\mathrm{d}x\Big)^2 \\ &\to \int_0^1\int_0^1 c(x)^{-1}c(y)^{-1}\,\mathrm{d}x\mathrm{d}y - \Big(\int_0^1 c(x)^{-1}\,\mathrm{d}x\Big)^2 = 0.\end{aligned}$$

The assertion in (a) follows from Markov's inequality.

We continue with part (b). Write $\hat T_n = S_{n1} + S_{n2} + S_{n3}$, where

$$S_{n1} = \frac{1}{k'}\sum_{j=1}^{k'} \hat Z_{n,j} - Z_{n,j}, \quad S_{n2} = \frac{1}{k'}\sum_{j=1}^{k'} Z_{n,j} - \mathrm{E}[Z_{n,j}], \quad S_{n3} = \frac{1}{k'}\sum_{j=1}^{k'} \mathrm{E}[Z_{n,j}].$$

First, we show $S_{n3} \to \tau/\theta$. Write $S_{n3} = \int_0^1 \varphi_n(\xi)\,\mathrm{d}\xi$, where

$$\varphi_n(\xi) = \mathrm{E}[Z_{n,1+\lfloor \xi k'\rfloor}] \to (\theta c(\xi))^{-1}$$

by Lemma 5.2 and uniform integrability, which follows from (B10). Hence, the dominated convergence theorem implies that $S_{n3} \to \int_0^1 (\theta c(\xi))^{-1}\,\mathrm{d}\xi = \tau/\theta$; note $\sup_{n\in\mathbb{N}} \|\varphi_n\|_\infty < \infty$ by Condition (B10).

In the following, we prove $S_{n1} = o_{\mathrm{P}}(1)$ and $S_{n2} = o_{\mathrm{P}}(1)$, and start with S_{n2}. Split S_{n2} into S_{n2}^{even} and S_{n2}^{odd}, which are defined as S_{n2} but with j only ranging over the even or odd numbers in $\{1, \dots, k'\}$, respectively. It suffices to show that S_{n2}^{even} and S_{n2}^{odd} are asymptotically negligible. We only treat S_{n2}^{even}; the proof for S_{n2}^{odd} is similar.

For $n \in \mathbb{N}$, let $(Z^*_{n,j})_{j=1,\dots,k'}$ denote an independent sequence with $Z^*_{n,j}$ being equal in distribution to $Z_{n,j}$ for $j = 1, \dots, k'$. Since the observations making up the even numbered blocks are separated by at least q observations, we may follow the argumentation in Eberlein (1984) to obtain

$$d_{\mathrm{TV}}\Big(P^{(Z_{n,2j})_{1\le j\le \lfloor k'/2\rfloor}},\ P^{(Z^*_{n,2j})_{1\le j\le \lfloor k'/2\rfloor}}\Big) \le \lfloor k'/2\rfloor\beta(q),$$

where d_{TV} denotes the total variation distance between two probability laws. Since $k'\beta(q) = o(1)$ by (B9), the above expression converges to zero as well, and $S_{n2}^{\mathrm{even}} = S_{n2}^{\mathrm{even},*} + o_{\mathrm{P}}(1)$, where $S_{n2}^{\mathrm{even},*}$ is defined as S_{n2}^{even} but in terms of $(Z^*_{n,j})_j$. Finally, $\mathrm{E}[S_{n2}^{\mathrm{even},*}] = 0$ and

$$\mathrm{Var}(S_{n2}^{\mathrm{even},*}) = \frac{1}{(k')^2}\sum_{j=1, j\ \mathrm{even}}^{k'} \mathrm{Var}(Z^*_{n,j}) \le \frac{1}{(k')^2}\sum_{j=1}^{k'} \mathrm{Var}(Z^*_{n,j}) = \frac{1}{k'}\int_0^1 g_n(\xi)\,\mathrm{d}\xi,$$

where $g_n(\xi) = \mathrm{Var}(Z_{n,1+\lfloor \xi k'\rfloor}) \to (\theta c(\xi))^{-2}$ by Lemma 5.2 and uniform integrability from (B10), which implies $\mathrm{Var}(S_{n2}^{\mathrm{even},*}) = o(1)$ and $S_{n2}^{\mathrm{even},*} = o_{\mathrm{P}}(1)$.

It remains to show $S_{n1} = o_{\mathrm{P}}(1)$. Note that the STEP $\mathbb{F}_n$ from (6.1) with $k = k'$ satisfies

$$\mathbb{F}_n(1, q(1 - F(x))) = q\sqrt{k'}\big\{F(x) - \hat F_n(x)\big\},$$

which yields $\hat{Z}_{n,j} - Z_{n,j} = \frac{1}{\sqrt{k'}}\mathbb{F}_n(1, Z_{n,j})$ for $j = 1, \dots, k'$ by the definition of $Z_{n,j}$ and $\hat{Z}_{n,j}$ in (5.2). Therefore,

$$S_{n1} = \frac{1}{k'}\sum_{j=1}^{k'} \frac{1}{\sqrt{k'}}\mathbb{F}_n(1, Z_{n,j}) = \int_0^\infty \mathbb{F}_n(1, x)\,\mathrm{d}H_n(x),$$

where $H_n(x) = (k')^{-3/2}\sum_{j=1}^{k'}\mathbf{1}(Z_{n,j} \le x)$. Note that $\sup_{x\in[0,T]}|H_n(x)| \le (k')^{-1/2} = o(1)$ for any $T \in \mathbb{N}$. Under the imposed conditions, Proposition 6.2 is applicable for $k = k'$, yielding $\{\mathbb{F}_n(1,x)\}_{x\in[0,T]} \rightsquigarrow \{\mathbb{S}(1,x)\}_{x\in[0,T]}$ in $(\ell^\infty([0,T]), \|\cdot\|_\infty)$, such that

$$S_{n1}(T) := \int_0^T \mathbb{F}_n(1,x)\,\mathrm{d}H_n(x) = o_{\mathrm{P}}(1), \quad T \in \mathbb{N},$$

by Lemma C.8 in Berghaus and Bücher (2017). By Theorem 4.2 in Billingsley (1968), the proof of $S_{n1} = o_{\mathrm{P}}(1)$ is finished once we show that, for any $\delta > 0$,

$$\lim_{T\to\infty}\limsup_{n\to\infty} P\big(|S_{n1} - S_{n1}(T)| > \delta\big) = 0.$$

Set $f_n(x,z) = \mathbf{1}(x > V(q/z)) - z/q$, such that $S_{n1} = \frac{1}{(k')^2}\sum_{i=1}^{k'}\sum_{j=1}^{n} f_n(X_j^{(n)}, Z_{n,i})$. Write $S_{n1} - S_{n1}(T) = A_{n,T} + B_{n,T} + C_{n,T}$, where

$$A_{n,T} = \frac{1}{(k')^2}\sum_{i=1}^{k'}\sum_{j\in\{i-1,i,i+1\}}\sum_{s\in I'_j} f_n(X_s^{(n)}, Z_{n,i})\mathbf{1}(Z_{n,i} \ge T),$$

$$B_{n,T} = \frac{1}{(k')^2}\sum_{i=1}^{k'-2}\sum_{j=i+2}^{k'}\sum_{s\in I'_j} f_n(X_s^{(n)}, Z_{n,i})\mathbf{1}(Z_{n,i} \ge T),$$

$$C_{n,T} = \frac{1}{(k')^2}\sum_{i=3}^{k'}\sum_{j=1}^{i-2}\sum_{s\in I'_j} f_n(X_s^{(n)}, Z_{n,i})\mathbf{1}(Z_{n,i} \ge T).$$

First, $|A_{n,T}| \le 3q/k' = o(1)$ by Condition (B9). It remains to show, for any $\delta > 0$,

$$\lim_{T\to\infty}\limsup_{n\to\infty} P(|B_{n,T}| > \delta) = 0, \quad \lim_{T\to\infty}\limsup_{n\to\infty} P(|C_{n,T}| > \delta) = 0.$$

We only consider $C_{n,T}$; $B_{n,T}$ can be treated similarly. Write

$$C_{n,T} = \frac{1}{k'}\sum_{i=3}^{k'} \varphi_{n,i-2}(Z_{n,i})\mathbf{1}(Z_{n,i} \ge T)$$

with

$$\varphi_{n,i-2}(z) = \frac{1}{k'}\sum_{j=1}^{i-2}\sum_{s\in I'_j} f_n(X_s^{(n)}, z).$$

For fixed $i \in \{3,\dots,k'\}$, consider the expectation $\mathrm{E}\left[|\varphi_{n,i-2}(Z_{n,i})|\mathbf{1}(Z_{n,i} \ge T)\right]$. By Berbee's coupling lemma (Berbee, 1979), we may construct a random variable $Z^*_{n,i}$ independent of $((X_s^{(n)})_{s\in I'_j})_{j=1,\dots,i-2}$ and equal in distribution to $Z_{n,i}$ with

$$\mathrm{P}(Z^*_{n,i} \ne Z_{n,i}) = \beta\big(\sigma(Z_{n,i}), \sigma\big(((X_s^{(n)})_{s\in I'_j})_{j=1,\dots,i-2}\big)\big) \le \beta(q).$$

Hence,

$$\mathrm{E}\left[|\varphi_{n,i-2}(Z_{n,i})|\mathbf{1}(Z_{n,i} \ge T)\right] = \mathrm{E}\left[|\varphi_{n,i-2}(Z^*_{n,i})|\mathbf{1}(Z^*_{n,i} \ge T)\right]$$

$$+\,\mathrm{E}\left[\left\{|\varphi_{n,i-2}(Z_{n,i})|\mathbf{1}(Z_{n,i}\geq T)-|\varphi_{n,i-2}(Z^*_{n,i})|\mathbf{1}(Z^*_{n,i}\geq T)\right\}\mathbf{1}(Z_{n,i}\neq Z^*_{n,i})\right].$$

Since $|\varphi_{n,i-2}|\leq q$ the second summand can be bounded by $2q\mathrm{P}(Z_{n,i}\neq Z^*_{n,i})\leq 2q\beta(q)\leq 2k'\beta(q)=o(1)$ by Condition (B9), uniformly in i. Now, consider the first summand in the above display, for which we first treat $\mathrm{E}[|\varphi_{n,i-2}(z)|]$ for $T\leq z\leq q$ (note that $Z^*_{n,i}\leq q$ a.s.). We have

$$\mathrm{E}[|\varphi_{n,i-2}(z)|]\leq\frac{1}{k'}\sum_{j=1}^{i-2}\sum_{s\in I'_j}\mathrm{E}\left[|f_n(X_s^{(n)},z)|\right]\leq z+\frac{1}{k'}\sum_{j=1}^{i-2}\sum_{s\in I'_j}\mathrm{P}\big(X_s^{(n)}>V(q/z)\big).$$

Since $\mathrm{P}\big(X_s^{(n)}>V(q/z)\big)=1-F_{n,s}\big((1-F)^{-1}(z/q)\big)$ and by Condition (B6), there exists some $\tau>0$ such that, for all $s\leq n$ and n large enough,

$$\mathrm{P}\big(X_s^{(n)}>V(q/z)\big)<\frac{z}{q}c\Big(\frac{s}{n}\Big)\Big\{1+\tfrac{\tau}{c_{\min}}A\Big(\frac{q}{z}\Big)\Big\}.$$

As a consequence, uniformly in i,

$$\mathrm{E}[|\varphi_{n,i-2}(z)|]\leq z+z\|c\|_\infty\Big\{1+\tfrac{\tau}{c_{\min}}A\Big(\frac{q}{z}\Big)\Big\}.$$

Since A is eventually decreasing, the last expression can be bounded by

$$z\big[1+\|c\|_\infty\big\{1+\tfrac{\tau}{c_{min}}A(1)\big\}\big]$$

for $T\leq z\leq q$. After conditioning on $Z^*_{n,i}$ we thus obtain with the Cauchy-Schwarz-inequality

$$\mathrm{E}\left[|\varphi_{n,i-2}(Z^*_{n,i})|\mathbf{1}(Z^*_{n,i}\geq T)\right]\leq\Big[1+\|c\|_\infty\Big\{1+\tfrac{\tau}{c_{min}}A(1)\Big\}\Big]\,\mathrm{E}\left[Z^*_{n,i}\mathbf{1}(Z^*_{n,i}\geq T)\right].$$

Since $Z^*_{n,i}$ has the same distribution as $Z_{n,i}$ and by the Cauchy Schwartz inequality, we have thus found the bound

$$\begin{aligned}\mathrm{E}[|C_{n,T}|]&\leq o(1)+\frac{1}{k'}\sum_{i=1}^{k'}\Big[1+\|c\|_\infty\Big\{1+\tfrac{\tau}{c_{min}}A(1)\Big\}\Big]\,\mathrm{E}\left[Z_{n,i}\mathbf{1}(Z_{n,i}\geq T)\right]\\&\lesssim o(1)+\int_0^1 g_n(\xi)\,\mathrm{d}\xi,\end{aligned}$$

where $g_n(\xi)=\mathrm{E}[Z^2_{n,1+\lfloor\xi k'\rfloor}]^{1/2}\mathrm{P}(Z_{n,1+\lfloor\xi k'\rfloor}\geq T)^{1/2}$ converges to $\mathrm{E}[V_\xi^2]^{1/2}\mathrm{P}(V_\xi\geq T)^{1/2}$ as $n\to\infty$ for $V_\xi\sim\mathrm{Exp}(\theta c(\xi))$ by Lemma 5.2 and Condition (B10). Altogether,

$$\lim_{T\to\infty}\limsup_{n\to\infty}\leq\lim_{T\to\infty}\int_0^1\mathrm{E}[V_\xi^2]^{1/2}P(V_\xi\geq T)^{1/2}\,\mathrm{d}\xi=0,$$

which implies (b). □

Appendix D. Auxiliary Results

Lemma D.1. *Fix $\xi\in[0,1)$ and $x>0$. Under Conditions (B0)-(B2), (B6) and (B9), $A_n=B_n+o_{\mathrm{P}}(1)$ and $B_n=C_n+o_{\mathrm{P}}(1)$ as $n\to\infty$, where*

$$A_n=\sum_{i\in I'_{1+\lfloor\xi k'\rfloor}}\mathbf{1}\Big(Z_i>\frac{1}{1-F_{n,i}}\big(F^{-1}(1-x/q)\big)\Big),$$

$$B_n=\sum_{i\in I'_{1+\lfloor\xi k'\rfloor}}\mathbf{1}\Big(Z_i>\frac{q}{c(i/n)x}\Big),\qquad C_n=\sum_{i\in I'_{1+\lfloor\xi k'\rfloor}}\mathbf{1}\Big(Z_i>\frac{q}{c(\xi)x}\Big).$$

Proof. For the first part of the lemma, first, note that since c is a positive and continuous function on $[0,1]$, there exist $v, w > 0$ such that $v < c(s) < w$ for all $s \in [0,1]$. By Condition (B6) there are real numbers $y_0 < x^*$ and $\tau > 0$ such that for all $y > y_0$, $n \in \mathbb{N}$ and $1 \le i \le n$,

$$c(i/n)\Big\{1 - \frac{\tau}{v}A\Big(\frac{1}{1-F(y)}\Big)\Big\} < \frac{1-F_{ni}(y)}{1-F(y)} < c(i/n)\Big\{1 + \frac{\tau}{v}A\Big(\frac{1}{1-F(y)}\Big)\Big\}.$$

Set $y_n = F^{-1}(1-x/q)$ and $w_n = \frac{\tau}{v}A\big(\frac{1}{1-F(y_n)}\big) = \frac{\tau}{v}A\big(\frac{q}{x}\big)$. Thus, for n large enough (such that $y_n > y_0$) we have for all $1 \le i \le n$,

$$\Big\{Z_i \ge \frac{q}{c(i/n)x}(1-w_n)^{-1}\Big\} \subseteq \Big\{Z_i \ge \frac{q}{x}\frac{1-F(y_n)}{1-F_{n,i}(y_n)}\Big\} \subseteq \Big\{Z_i \ge \frac{q}{c(i/n)x}(1+w_n)^{-1}\Big\}.$$

Since

$$A_n = \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \mathbf{1}\Big(Z_i > \frac{q}{x}\frac{1-F(y_n)}{1-F_{n,i}(y_n)}\Big),$$

this implies $B_n^- \le A_n \le B_n^+$, where

$$B_n^{\pm} = \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \mathbf{1}\Big(Z_i > \frac{q}{c(i/n)x}(1 \pm w_n)^{-1}\Big).$$

Next, we have

$$\begin{aligned}
\mathrm{E}[|B_n^{\pm} - B_n|] &\le \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \mathrm{E}\Big[\Big|\mathbf{1}\Big(Z_i > \frac{q}{c(i/n)x}(1 \pm w_n)^{-1}\Big) - \mathbf{1}\Big(Z_i > \frac{q}{c(i/n)x}\Big)\Big|\Big] \\
&\le \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \Big\{\mathrm{P}\Big(\frac{q}{c(i/n)x}(1 \pm w_n)^{-1} < Z_i \le \frac{q}{c(i/n)x}\Big) \\
&\qquad\qquad + \mathrm{P}\Big(\frac{q}{c(i/n)x} < Z_i \le \frac{q}{c(i/n)x}(1 \pm w_n)^{-1}\Big)\Big\}.
\end{aligned}$$

Let us consider the case with the plus-sign. Note that $w_n > 0$. Recalling that Z_i is Pareto-distributed the above expression reduces to

$$w_n \frac{x}{q} \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} c(i/n) \le w_n x \|c\|_\infty,$$

which converges to 0 since $w_n \to 0$ by Condition (B6). The case with the minus-sign can be treated analogously. Hence, we have shown $B_n^{\pm} - B_n \to 0$ in $L_1(\mathrm{P})$ as $n \to \infty$. The assertion follows from $B_n^- \le A_n \le B_n^+$.

For the second part of the lemma write

$$\begin{aligned}
\mathrm{E}[|B_n - C_n|] &\le \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \mathrm{E}\Big[\Big|\mathbf{1}\Big(Z_i > \frac{q}{c(i/n)x}\Big) - \mathbf{1}\Big(Z_i > \frac{q}{c(\xi)x}\Big)\Big|\Big] \\
&\le \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} \Big\{\mathrm{P}\Big(\frac{q}{c(i/n)x} < Z_i \le \frac{q}{c(\xi)x}\Big) + \mathrm{P}\Big(\frac{q}{c(\xi)x} < Z_i \le \frac{q}{c(i/n)x}\Big)\Big\} \\
&= \frac{x}{q} \sum_{i \in I'_{1+\lfloor \xi k' \rfloor}} |c(i/n) - c(\xi)|,
\end{aligned}$$

where the last equation is due to the fact that Z_i is Pareto-distributed. Further, by Condition (B2), we have

$$\begin{aligned}\frac{1}{q}\sum_{i\in I'_{1+\lfloor \xi k'\rfloor}} |c(i/n)-c(\xi)| &\le \frac{K_c}{q}\sum_{i\in I'_{1+\lfloor \xi k'\rfloor}} |i/n-\xi|^{1/2}\\ &= \frac{K_c}{q}\sum_{j=1}^{q}\Big|\frac{\lfloor \xi k'\rfloor r+j}{n}-\xi\Big|^{1/2}\\ &= \frac{K_c}{qn^{1/2}}\sum_{j=1}^{q}\Big|(\lfloor \xi k'\rfloor-\xi k')q+j\Big|^{1/2} \le \frac{\sqrt{2}K_c(q)^{1/2}}{n^{1/2}} = o(1)\end{aligned}$$

by Condition (B9). Therefore, $B_n - C_n \to 0$ in $L_1(\mathrm{P})$ as $n\to\infty$, which implies the second assertion. □

Corollary D.2. *Fix $\xi\in[0,1)$ and $x>0$. Under Condition (B0)-(B2), (B6) and (B9),*

$$\mathrm{P}\Big(Z_i \le \frac{q}{c(i/n)x} \ \ \textit{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) - \mathrm{P}\Big(Z_i \le \frac{q}{c(\xi)x} \ \ \textit{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) = o(1),$$

and

$$\begin{aligned}\mathrm{P}\Big(Z_i \le \frac{1}{1-F_{n,i}}\big(F^{-1}(1-x/q)\big) \ \ \textit{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big)&\\ - \mathrm{P}\Big(Z_i \le \frac{q}{c(i/n)x} \ \ \textit{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) &= o(1), \quad n\to\infty.\end{aligned}$$

Proof. Note that

$$\begin{aligned}\mathrm{P}\Big(Z_i \le \frac{1}{1-F_{n,i}}\big(F^{-1}(1-x/q)\big) \ \ \text{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) &= \mathrm{P}(A_n=0),\\ \mathrm{P}\Big(Z_i \le \frac{q}{c(i/n)x} \ \ \text{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) &= \mathrm{P}(B_n=0),\\ \mathrm{P}\Big(Z_i \le \frac{q}{c(\xi)x} \ \ \text{for all}\ \ i\in I'_{1+\lfloor \xi k'\rfloor}\Big) &= \mathrm{P}(C_n=0).\end{aligned}$$

Therefore,

$$\begin{aligned}|P(A_n=0)-P(B_n=0)| &= P(A_n=0,B_n>0)+P(A_n>0,B_n=0)\\ &\le 2P(|A_n-B_n|>1/2) = o(1)\end{aligned}$$

by Lemma D.1. And $|P(B_n=0)-P(C_n=0)| = o(1)$ can be shown analogously. □

3 Outlook

In this chapter, a potential continuation of this work and some open research questions are briefly presented.

First, the disjoint and sliding blocks estimators for the extremal index θ and the limiting cluster size distribution π in Chapter 2.1 and 2.2 were shown to asymptotically follow a centered normal distribution. Here, the limiting variance formulas were too complicated to allow for some simple estimation based on the plug-in principle. However, for further statistical inference on θ and π, e.g., for the construction of confidence intervals, estimators for the asymptotic variance formulas are needed. In the case of the extremal index, one approach would be to estimate the disjoint blocks estimator's variance based on an asymptotic expansion of the estimator derived from the proof of its asymptotic normality, as it was done in Berghaus and Bücher (2018). Since for the proposed estimators for θ the difference between the disjoint and sliding blocks variance only depends on θ, such an estimator could then be used for estimation of the sliding blocks variance as well. However, the proof of consistency of these variance estimators is quite elaborate in Berghaus and Bücher (2018)(Proposition 4.1), under even stronger mixing conditions, and, in particular, this approach does not work for the estimation of the asymptotic variance of the sliding blocks estimator for π since the difference between the corresponding disjoint and sliding blocks variance is more complicated. A more general approach, which could also be used to approximate the limiting distributions of the estimators for θ and π, consists of bootstrap methods such as the dependent/block multiplier bootstrap (Drees, 2015; Bücher and Kojadinovic, 2016). Studying bootstrap procedures to approximate the limiting distributions of statistics based on sliding blocks maxima seems especially appealing since, in many cases, deducting inference when sliding blocks are involved turns out to be difficult and there is no universal approach yet (Drees and Neblung, 2021).

Furthermore, the proposed estimators for θ and π depend on a block length parameter, which was seen to have a notable impact on the estimation accuracy in simulation studies. Here, it would be interesting to analyze estimators that aggregate over multiple block sizes, in order to achieve more robustness in this parameter and to possibly improve upon the single block length case; such an improvement has been observed in Zou et al. (2021) in a different context.

Besides, in Chapter 2.1, no estimator for the extremal index could be identified to be overall superior. It would be interesting to investigate which minimal asymptotic variance can be achieved by estimators relying on the considered rank-based samples, and whether there are estimators for the extremal index that are semiparametrically efficient.

In Chapter 2.3, a model was considered where the observations are serially dependent and their marginal tails are proportional to each other, described by the scedasis function. Estimators for the scedasis function and the integrated scedasis function and tests for detecting heteroscedasticity in the extremes were considered. Under domain of attraction conditions, one could further analyze estimators for the extreme value index or high quantiles at a specific time point, as in Einmahl et al. (2016) in the case of independent observations. Further, the introduced estimators for the extremal index θ of the stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ constitute a modification of the estimator from Berghaus and Bücher (2018). Here, one could also study method of moments estimators as in Chapter 2.1. In addition, it would be interesting to investigate estimators for θ that are not of block maxima type but rely on the runs or inter-exceedance times method instead. Besides, the proposed estimators for θ and the tests for heteroscedasticity in the extremes, which are based on a block bootstrap, both rely on the construction of disjoint blocks. In view of the findings in Chapter 2.1 and 2.2, it would be worth to examine whether the above methods can be improved upon by using sliding blocks.

This last point raises the more general question of the superiority of the sliding blocks method over the disjoint blocks method, and why it seems to be more advantageous for statistics of block maxima type than for ones of peak-over-threshold type (Cissokho and Kulik, 2020; Drees and Neblung, 2021). Another aspect of future research would be a weakening of the underlying assumptions needed for the proposed methods to work. This concerns relaxing the mixing conditions in Chapter 2, especially when beta-mixing is involved, and a further weakening of the regularity condition imposed on the time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ in Chapter 2.3. Since this time series was assumed to be stationary, the serial dependence of the overlying time series $(X_t^{(n)})_{t\in\mathbb{Z}}$ is unable to change over time. A step towards allowing for a more dynamic model that also permits a (smoothly) changing serial dependence over time would be to incorporate the concept of local stationarity (Vogt, 2012; Dahlhaus, 2012).

References

Aue, A. and Horváth, L. (2013). Structural breaks in time series. *Journal of Time Series Analysis*, 34(1):1–16.

Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5):792–804.

Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Processes and their Applications*, 119(4):1055–1080.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* John Wiley & Sons Ltd., Chichester.

Berbee, H. C. P. (1979). *Random walks with stationary increments and renewal theory.* Mathematisch Centrum, Amsterdam.

Berghaus, B. and Bücher, A. (2017). Goodness-of-fit tests for multivariate copula-based time series models. *Econometric Theory*, 33(2):292–330.

Berghaus, B. and Bücher, A. (2018). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *The Annals of Statistics*, 46(5):2307–2335.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Billingsley, P. (1979). *Probability and Measure.* Wiley, New York.

Billingsley, P. (1995). *Probability and Measure.* Wiley, New York, third edition.

Bradley, R. C. (1983). Approximation theorems for strongly mixing random variables. *Michigan Mathematical Journal*, 30(1):69–81.

Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, 2:107–144.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer, New York, second edition.

Bücher, A. and Jennessen, T. (2020). Method of moments estimators for the extremal index of a stationary time series. *Electronic Journal of Statistics*, 14(2):3103–3156.

Bücher, A. and Jennessen, T. (2022a). Statistical analysis for stationary time series at extreme levels: New estimators for the limiting cluster size distribution. *Stochastic Processes and their Applications*, 149:75–106.

Bücher, A. and Jennessen, T. (2022b). Statistics for Heteroscedastic Time Series Extremes. *arXiv:2204.09534*.

Bücher, A., Kinsvater, P., and Kojadinovic, I. (2017). Detecting breaks in the dependence of multivariate extreme-value distributions. *Extremes*, 20(1):53–89.

Bücher, A. and Kojadinovic, I. (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, 22(2):927–968.

Bücher, A. and Kojadinovic, I. (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *Journal of Theoretical Probability*, 32(3):1145–1165.

Bücher, A. and Segers, J. (2014). Extreme value copula estimation based on block maxima of a multivariate stationary time series. *Extremes*, 17(3):495–528.

Bücher, A. and Segers, J. (2018a). Inference for heavy tailed stationary time series based on sliding blocks. *Electronic Journal of Statistics*, 12(1):1098–1125.

Bücher, A. and Segers, J. (2018b). Maximum likelihood estimation for the Fréchet distribution based on block maxima extracted from a time series. *Bernoulli*, 24(2):1427–1462.

Bücher, A. and Zanger, L. (2021). On the disjoint and sliding block maxima method for piecewise stationary time series. *arXiv:2110.15576*.

Bücher, A. and Zhou, C. (2018). A horse racing between the block maxima method and the peak-over-threshold approach. *arXiv:1807.00282*.

Cai, J. J. (2019). A nonparametric estimator of the extremal index. *arXiv:1911.06674*.

Capéraà, P., Fougères, A.-L., and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84(3):567–577.

Cissokho, Y. and Kulik, R. (2020). Estimation of cluster functionals for regularly varying time series: sliding blocks estimators. *Electronic Journal of Statistics*, 15(1):2777–2831.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer, London.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, 25(1):1–37.

Dahlhaus, R. (2012). Locally stationary processes. In Subba Rao, T., Subba Rao, S., and Rao, C., editors, *Time Series Analysis: Methods and Applications*, pages 351–413. Elsevier, Amsterdam.

Dahlhaus, R. and Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis*, 19(6):629–655.

Darsow, W. F., Nguyen, B., and Olsen, E. T. (1992). Copulas and Markov processes. *Illinois Journal of Mathematics*, 36(4):600–642.

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

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3):393–425.

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

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

de Haan, L., Resnick, S. I., Rootzén, H., and de Vries, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Processes and their Applications*, 32(2):213–224.

de Haan, L. and Zhou, C. (2021). Trends in extreme value indices. *Journal of the American Statistical Association*, 116(535):1265–1279.

Dehling, H., Durieu, O., and Volny, D. (2009). New techniques for empirical processes of dependent data. *Stochastic Processes and their Applications*, 119(10):3699–3718.

Dehling, H. and Philipp, W. (2002). Empirical Process Techniques for Dependent Data. In Dehling, H., Mikosch, T., and Sørensen, M., editors, *Empirical Process Techniques for Dependent Data*, pages 3–113. Birkhäuser, Boston.

Dombry, C. (2015). Existence and consistency of the maximum likelihood estimators for the extreme value index within the block maxima framework. *Bernoulli*, 21(1):420–436.

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

Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *The Annals of Applied Probability*, 10(4):1274–1301.

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

Drees, H. and Knežević, M. (2020). Peak-over-threshold estimators for spectral tail processes: random vs deterministic thresholds. *Extremes*, 23(3):465–491.

Drees, H. and Neblung, S. (2021). Asymptotics for sliding blocks estimators of rare events. *Bernoulli*, 27(2):1239–1269.

Drees, H. and Rootzén, H. (2010). Limit theorems for empirical processes of cluster functionals. *The Annals of Statistics*, 38(4):2145–2186.

Drees, H., Segers, J., and Warchoł, M. (2015). Statistics for tail processes of Markov chains. *Extremes*, 18(3):369–402.

Eberlein, E. (1984). Weak convergence of partial sums of absolutely regular sequences. *Statistics & Probability Letters*, 2(5):291–293.

Einmahl, J. H. J., de Haan, L., and Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(1):31–51.

Einmahl, J. H. J., Ferreira, A., de Haan, L., Neves, C., and Zhou, C. (2022). Spatial dependence and space–time trend in extreme events. *The Annals of Statistics*, 50(1):30–52.

Ferreira, A. and de Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *The Annals of Statistics*, 43(1):276–298.

Ferreira, H. and Ferreira, M. (2018). Estimating the extremal index through local dependence. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(2):587–605.

Ferreira, M. (2018). Analysis of estimation methods for the extremal index. *Electronic Journal of Applied Statistical Analysis*, 11(1):296–306.

Ferro, C. A. T. (2003). Statistical methods for clusters of extreme values. *Ph. D. thesis, University of Lancaster.*

Ferro, C. A. T. and Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):545–556.

Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *The Annals of Statistics*, 37(5B):2990–3022.

Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press, New York.

Hall, P. and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 15(2):153–167.

Hoga, Y. (2017). Change point tests for the tail index of β-mixing random variables. *Econometric Theory*, 33(4):915–954.

Hoga, Y. (2018). A structural break test for extremal dependence in β-mixing random vectors. *Biometrika*, 105(3):627–643.

Hsing, T. (1984). Point processes associated with extreme value theory. *Ph. D. thesis, University of North Carolina at Chapel Hill.*

Hsing, T. (1991a). Estimating the parameters of rare events. *Stochastic Processes and their Applications*, 37(1):117–139.

Hsing, T. (1991b). On tail index estimation using dependent data. *The Annals of Statistics*, 19(3):1547–1569.

Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *The Annals of Statistics*, 21(4):2043–2071.

Hsing, T., Hüsler, J., and Leadbetter, M. R. (1988). On the exceedance point process for a stationary sequence. *Probability Theory and Related Fields*, 78(1):97–112.

Ibragimov, R. and Sharakhmetov, S. (2002). The exact constant in the Rosenthal inequality for random variables with mean zero. *Theory of Probability & Its Applications*, 46(1):127–132.

Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Statistics and Computing*, 3:135–146.

Klein Tank, A. and Können, G. (2003). Trends in indices of daily temperature and precipitation extremes in europe, 1946–99. *Journal of Climate*, 16(22):3665–3680.

Kley, T., Volgushev, S., Dette, H., and Hallin, M. (2016). Quantile spectral processes: Asymptotic analysis and inference. *Bernoulli*, 22(3):1770–1807.

Kojadinovic, I. and Naveau, P. (2017). Detecting distributional changes in samples of independent block maxima using probability weighted moments. *Extremes*, 20(2):417–450.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York.

Kulik, R. and Soulier, P. (2020). *Heavy-Tailed Time Series*. Springer, New York.

Kulik, R., Soulier, P., and Wintenberger, O. (2019). The tail empirical process of regularly varying functions of geometrically ergodic Markov chains. *Stochastic Processes and their Applications*, 129(11):4209–4238.

Laurini, F. and Tawn, J. A. (2003). New estimators for the extremal index and other cluster characteristics. *Extremes*, 6(3):189–211.

Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(2):291–306.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

Leadbetter, M. R. and Rootzén, H. (1988). Extremal theory for stochastic processes. *The Annals of Probability*, 16(2):431–478.

Mefleh, A. (2018). Contributions to extreme value theory: Trend detection for heteroscedastic extremes. *Ph. D. thesis, Université Bourgogne Franche-Comté*.

Moloney, N. R., Faranda, D., and Sato, Y. (2019). An overview of the extremal index. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(2):022101.

Móricz, F. (1982). A general moment inequality for the maximum of partial sums of single series. *Acta Scientiarum Mathematicarum*, 44(1-2):67–75.

Naveau, P., Guillou, A., Cooley, D., and Diebolt, J. (2009). Modelling pairwise dependence of maxima in space. *Biometrika*, 96(1):1–17.

Neblung, S. A. (2021). Estimators for temporal dependence of extremes. *Ph. D. thesis, Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky.*

Northrop, P. J. (2015). An efficient semiparametric maxima estimator of the extremal index. *Extremes*, 18(4):585–603.

O'Brien, G. L. (1987). Extreme values for stationary and Markov sequences. *The Annals of Probability*, 15(1):281–291.

Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications. *The Annals of Applied Probability*, 4(2):529–548.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131.

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute, 49(2), Proceedings of the 43rd session, Buenos Aires*, pages 859–878, 894–902.

Rémillard, B., Papageorgiou, N., and Soustra, F. (2012). Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis*, 110:30–42.

Resnick, S. and Stărică, C. (1998). Tail index estimation for dependent data. *The Annals of Applied Probability*, 8(4):1156–1183.

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

Robert, C. Y. (2009a). Asymptotic distributions for the intervals estimators of the extremal index and the cluster-size probabilities. *Journal of Statistical Planning and Inference*, 139(9):3288–3309.

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

Robert, C. Y., Segers, J., and Ferro, C. A. T. (2009). A sliding blocks estimator for the extremal index. *Electronic Journal of Statistics*, 3:993–1020.

Rootzén, H. (2009). Weak convergence of the tail empirical process for dependent sequences. *Stochastic Processes and their Applications*, 119(2):468–490.

Scotto, M. G., Barbosa, S. M., and Alonso, A. M. (2011). Extreme value and cluster analysis of european daily temperature series. *Journal of Applied Statistics*, 38(12):2793–2804.

Segers, J. (2003). Functionals of clusters of extremes. *Advances in Applied Probability*, 35(4):1028–1045.

Shao, Q.-M. and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *The Annals of Probability*, 24(4):2098–2127.

Smith, R. L. and Weissman, I. (1994). Estimating the extremal index. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(3):515–528.

Süveges, M. (2007). Likelihood estimation of the extremal index. *Extremes*, 10(1-2):41–55.

Süveges, M. and Davison, A. C. (2010). Model misspecification in peaks over threshold analysis. *The Annals of Applied Statistics*, 4(1):203–221.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer, New York.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *The Annals of Statistics*, 40(5):2601–2633.

Weissman, I. and Novak, S. Y. (1998). On blocks and runs estimators of the extremal index. *Journal of Statistical Planning and Inference*, 66(2):281–288.

Zolina, O., Simmer, C., Belyaev, K., Kapala, A., and Gulev, S. (2009). Improving estimates of heavy and extreme precipitation using daily records from european rain gauges. *Journal of Hydrometeorology*, 10(3):701–716.

Zou, N., Volgushev, S., and Bücher, A. (2021). Multiple block sizes and overlapping blocks for multivariate time series extremes. *The Annals of Statistics*, 49(1):295–320.

Author contribution statement

In the following, the individual contributions of the authors of the articles included in this thesis are outlined.

1) Bücher, A. and Jennessen, T. (2020). Method of moments estimators for the extremal index of a stationary time series. *Electronic Journal of Statistics*, 14(2):3103-3156.

   The initial idea in this article to investigate estimators for the extremal index based on the method of moments and the transformed block maxima was developed by the first author, who also proposed the CFG- and MAD-estimator and wrote the introduction. The second author suggested analyzing the ROOT-estimator and established the theoretical results. The simulation study was performed by the second author with the help of a student assistant. Apart from the introduction, the manuscript was mostly drafted by the second author. Throughout the project, the first author supported its development, and made several improvements and corrections leading to the final version of the article.

2) Bücher, A. and Jennessen, T. (2022). Statistical analysis for stationary time series at extreme levels: New estimators for the limiting cluster size distribution. *Stochastic Processes and their Applications*, 149:75-106.

   The first author proposed the estimator for the limiting cluster size distribution in this article and wrote the introduction. The concrete elaboration and the theoretical results were established by the second author. He was also responsible for the simulation study and drafted most of the manuscript. Throughout the project, the first author supported its development, and made several improvements and corrections leading to the final version of the article. Both authors contributed equally to the extensive revision of the article upon its first submission for publication in the above-mentioned journal.

3) Bücher, A. and Jennessen, T. (2022). Statistics for Heteroscedastic Time Series Extremes.

   Submitted for publication in *Bernoulli* and publicly available on *arXiv*:2204.09534.

The general setting in this article as an extension of the setting in Einmahl et al. (2016) to the case of dependent data was proposed by the first author and Chen Zhou. The first author developed and proved the weak convergence result for the simple STEP (Proposition 6.1). The remaining theoretical results were established by the second author, who also performed the simulation study. Throughout the project, the first author supported its development, and made several improvements and corrections leading to the final version of the article.

Eidesstattliche Versicherung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorgelegten oder ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 03. Mai 2022 Tobias Jennessen