Abstract We consider extreme value analysis for independent but non-identically distributed observations. In particular, the observations do not share the same extreme value index. This situation is related to, but differs from, heteroscedastic extremes in Einmahl et al. (2016). Compared to the heteroscedastic extremes, our model allows for a broader class in which tails of the probability distributions of different observations are of different order. In other words, we are dealing with distributions that differ much more than the heteroscedastic extremes. Assuming continuously changing extreme value indices, we provide a non-parametric estimate for the functional extreme value index. Besides estimating the extreme value index locally, we also provide a global estimator for the trend and its joint asymptotic property. The global asymptotic property can be used for testing a pre-specified parametric trend in the extreme value indices. In particular, it can be applied to test whether the extreme value index remains at a constant level across all observations.

Keywords: Hill estimator; peak over threshold; local and global estimation; stochastic integral

MSC 2010 subject classifications: 62G32; 62G10
1 Introduction

Extreme value analysis makes statistical inference on the tail region of a distribution function. Balkema and de Haan (1974) shows that extreme observations above a high threshold follows approximately a scaled generalized Pareto distribution (GPD). Consequently, one main parameter governs the tail behavior: the shape parameter in the GPD, also known as the extreme value index. Estimation of this parameter is therefore of prime importance for tail inference. Classical extreme value analysis assumes that the observations are independent and identically distributed (IID). Recent studies aim at dealing with the case when observations are drawn from different distributions. In this paper, we aim at dealing with non IID observations: we consider a continuously changing extreme value index and try to estimate the functional extreme value index accurately.

Consider a set of distribution functions $F_s(x)$ for $s \in [0,1]$ and independent random variables $X_i \sim F_{\frac{i}{n}}(x)$ for $i = 1, \cdots, n$. Here $F_s(x)$ is assumed to be continuous with respect to $s$ and $x$. To perform extreme value analysis, assume that $F_s \in D_{\gamma(s)}$, where $D_{\cdot}$ denotes the max-domain of attraction with corresponding extreme value index. In this paper, we consider the case that $\gamma(s)$ is a positive continuous function on $[0,1]$. The goal is to estimate the function $\gamma(s)$ and test the hypothesis that $\gamma(s) = \gamma_0(s)$ for some given function $\gamma_0$, based on the observations $X_1, \cdots, X_n$.

The idea for estimating $\gamma(s)$ locally is similar to the kernel density estimation. More specifically, we will use only observations $X_i$ in the $h$–neighborhood of $s$, i.e. $i \in I_n(s) = \{i : \frac{i}{n} - s \leq h\}$, where $h := h(n)$ is the bandwidth such that as $n \to \infty$, $h \to 0$ and $nh \to \infty$. Correspondingly there will be $[2nh]$ observations for $s \in [h, 1-h]$. To apply any known estimators for the extreme value index, such as the Hill estimator, we choose an intermediate sequence $k := k(n)$ such that as $n \to \infty$, $k \to \infty$ and $k/n \to 0$. Then we use the top $[2kh]$ order statistics among the $[2nh]$ local observations in the $h$–neighborhood of $s$ to estimate $\gamma(s)$. The local Hill estimator for $\gamma(s)$ is defined as follows. Rank the
observations $X_i$ with $i \in I_n(s)$ as $X_{1, [2nh]}^s \leq \cdots \leq X_{[2nh], [2nh]}^s$. Then,

$$
\hat{\gamma}_H(s) := \frac{1}{2kh} \sum_{i \in I_n(s)} (\log X_i - \log X_{[2nh] - [2kh], [2nh]})^+.
$$

(1.1)

We start with considering the local asymptotic normality. Under some suitable conditions for $k$ and $h$, we can show that, as $n \to \infty$, for each fixed $s \in (0, 1)$,

$$
\sqrt{2kh} (\hat{\gamma}_H(s) - \gamma(s)) \overset{d}{\to} N(0, (\gamma(s))^2).
$$

This result is comparable with the asymptotic normality of the Hill estimator, but now the estimation is based on observations with different extreme value indices. The speed of convergence is $\sqrt{2kh}$ because only the top $[2kh]$ order statistics are used in the estimation.

Next, we consider testing the hypothesis that $\gamma(s) = \gamma_0(s)$ for all $s \in [0, 1]$. Although we are able to estimate the function $\gamma$ locally, since the local estimators use only local observations, their asymptotic limits are obviously independent. That prevents us from constructing a testing procedure. In addition, the local estimators converges with a slow speed of convergence $1/\sqrt{2kh}$. In order to achieve the stated goal, we consider the estimation of $\Gamma(s) = \int_0^s \gamma(u)du$ and test the equivalent hypothesis that $\Gamma(s) = \Gamma_0(s) = \int_0^s \gamma_0(u)du$.

The function $\Gamma(s)$ is estimated by aggregating the local estimators of $\gamma(s)$ to a “global estimator” as follows. Consider the series $s_t = (2t - 1)h$ for $t = 1, 2, \cdots, [1/2h]$. Then the intervals $I(s_t)$ for $t = 1, 2, \cdots, [1/2h]$ form a partition of $[0, 1]$. The estimator of $\Gamma(s)$ is defined as

$$
\hat{\Gamma}_H(s) = 2h \sum_{s_t \leq s} \hat{\gamma}_H(s_t).
$$

(1.2)

Note that $\hat{\Gamma}_H(s)$ is a stochastic process defined on $[0, 1]$. We shall show that, under the same conditions on $k$ and $h$ used in the local estimation, there exists a series of Brownian motions $\{W_n(t)\}$ such that, as $n \to \infty$,

$$
\sup_{0 \leq s \leq 1} \left| \sqrt{k} \left( \hat{\Gamma}_H(s) - \Gamma(s) \right) - \int_0^s \gamma(u)dW_n(u) \right| \overset{P}{\to} 0.
$$
There are two notable features in this asymptotic relation. Firstly, the convergence is uniformly for all \( s \in [0, 1] \). Secondly, the speed of convergence for the estimators \( \hat{\Gamma}_H(s) \) is \( 1/\sqrt{k} \). From these two features, it is possible to construct efficient testing methods to test the null hypothesis that \( \Gamma(s) = \Gamma_0(s) \) for all \( s \in [0, 1] \) with some given \( \Gamma_0(s) \).

Our approach can be regarded as a combination of kernel density estimation and extreme value statistics. To prove the local and global asymptotic normality, we need to combine two limiting procedures as the number of observations tending to infinity. Firstly, the observations used are from a \( h \)-neighborhood that is shrinking. Secondly, within each \( h \)-neighborhood, we need to apply a threshold to all observations that is increasing towards infinity. If the \( h \)-neighborhood shrinks too fast, there will be no sufficient observations in each neighborhood for statistical inference. If it shrinks too slowly, we would have involved too many observations with vary different extreme value indices such that the estimation is distorted. Therefore, the two limiting procedures have to be balanced such that the resulting estimators possess proper asymptotic properties.

For that purpose, we assume some conditions regarding the choice of \( k \) and \( h \) that are related to the speed of variation of the distribution function \( F_s \) and the continuity of the extreme value index \( \gamma(s) \); see conditions (2.2)–(2.5). The first two conditions (2.2) and (2.3) are typical assumptions in kernel density estimation and extreme value statistics respectively. The third condition (2.4) ensures that the extreme value index function \( \gamma(s) \) is sufficiently smooth. In other words, observations in the \( h \)-neighborhood have extreme value indices that are not too far off. Notice that \( \gamma(s) \) governs the parametric structure of the limit only. In order that a unified “threshold” can be applied to all observations in the \( h \)-neighborhood, we further assume the smoothness of the intermediate quantiles as in condition (2.5). These conditions are not too restrictive, see Example 2.1 below.

The most close studies to our approach are Gardes and Girard (2010) and Goegebeur et al. (2014). The setups of these two studies are similar to our analysis albeit formulated in a conditional setup. The former focuses on non–stochastic covariates, whereas the latter focuses on random covariates. Both approaches proposed estimators using observations locally and established the local asymptotic normality only. To obtain the
local asymptotic normality, the conditions assumed in these two studies are quite different from our conditions and the results obtained therein also differ from our results. Besides, we attempt to establish a global result for \( \hat{\Gamma}(s) \). The global asymptotic result is necessary for conducting hypothesis testing.

Our paper is also related to, but differs from, heteroscedastic extremes. Einmahl et al. (2016) models the tail region of distributions of non IID observations by considering the quotient between tails of different distributions and a common tail. By assuming that such quotients stay positive and finite as one goes further into the tail, the asymptotic constant is called “scedasis”. Within such a framework, the extreme value index remains unchanged across the non IID observations. Compared to the heteroscedastic extremes, we allow for continuously changing extreme value index and try to estimate the functional extreme value index accurately. In our case, the tails of the probability distributions are of different order, i.e. the quotient between the tails of the distributions at two locations with different extreme value indices tends to either zero or infinity. In other words, we are dealing with distributions that differ much more than in the heteroscedastic extremes. Therefore, our situation cannot be handled in the same way as in heteroscedastic extremes.

This paper is also related to the literature dealing with the variation or trend in extreme value index when considering a purely parametric model such as the generalized extreme value distribution (GEV) or the GPD. Firstly, one may model the trend in the parameters of such models as a specific functional of the covariates; see e.g. Smith (1989) for the GEV model and Davison and Smith (1990) for the GPD model. Secondly, the trend can also be non–parametrically estimated using various local estimation techniques; see, e.g. Davison and Ramesh (2000) using the local likelihood method, Hall and Tajvidi (2000) using the local linearization method, among others. Compared to all these studies, we do not impose a fully parameterized model and therefore maintain a semi–parametric approach.

Lastly, this paper contributes to the literature on testing the null hypothesis of constant extreme value index. Quintos et al. (2001) and more recently Hoga (2017) considered
testing a change point in the tail index. Einmahl et al. (2016) proposed two tests for the same purpose. Nevertheless, in all these studies, the main asymptotic result for the constructed tests is under a relatively more restrictive null than having constant extreme value index only. In contrast, we consider a wider null hypothesis potentially including models with constant extreme value index that are excluded from the null of the two existing studies. In addition, our study allows for testing the null hypothesis of having a general pre-specified trend in the extreme value index beyond the constant function, such as \( \gamma(s) = \gamma_0(s) \) for all \( s \).

We demonstrate the performance of our testing procedure by extensive simulation studies. In addition, our estimation procedure for the \( \gamma(s) \) function is also valid when the function differs from a constant function. We apply our developed method to the losses of the S&P 500 index. The testing results show that we do not reject the constant extreme value index in the period from 1988 to 2012 but do reject this null in a longer period from 1963 to 2012.

The paper is organized as follows. Our main theorem regarding the local and global estimators are presented in Section 2. The testing procedure is established in Section 3 with simulations. Section 4 is devoted to the application. Proofs are postponed to the Appendix.

2 Main theorem

We need the following conditions for obtaining the asymptotic properties of the local and global estimators.

Firstly, we assume the usual second order condition, but uniformly for all \( s \in [0, 1] \) as follows. Denote \( U_s = (1/(1 - F_s))^- \) as the quantile function corresponding to the distribution \( F_s \). Suppose there exists a continuous negative function \( \rho(s) \) on \( [0, 1] \) and a set of auxiliary functions \( A_s(t) \) that are continuous with respect to \( s \), such that

\[
\lim_{t \to \infty} \frac{U_s(tx) - x^{\gamma(s)}}{A_s(t)} = x^{\gamma(s)} \frac{x^{\rho(s)} - 1}{\rho(s)},
\]

(2.1)
holds uniformly for all $s \in [0,1]$ and $x > 1/2$. A similar uniform second order condition has been adopted in Einmahl and Lin (2006).

Next, we require that the intermediate sequence $k$ and the bandwidth $h$ are properly chosen as follows: there exists some positive constant $\varepsilon > 0$ such that as $n \to \infty$,

$$h = h_n \to 0, \quad k = k_n \to \infty, \quad k_n/n \to 0, \quad k_n h_n^{1+\varepsilon} \to \infty, \quad k_n h_n^2 \to 0 \quad (2.2)$$

$$\Delta_{1,n} := \sqrt{k} \sup_{0 \leq s \leq 1} \left| A_s \left( \frac{n}{k} \right) \right| \to 0, \quad (2.3)$$

$$\Delta_{2,n} := \sqrt{k} \log k \sup_{|s_1 - s_2| \leq 2h} |\gamma(s_1) - \gamma(s_2)| \to 0. \quad (2.4)$$

$$\Delta_{3,n} := \sup_{|s_1 - s_2| \leq h} \left| \frac{U_{s_1}(\frac{x}{N}) - 1}{A_{s_2}(\frac{x}{N})} \right| \to 0, \quad (2.5)$$

The first condition ensures that the number of high order statistics used in each local interval tends to infinity. The second condition is the one usually required for extreme value analysis in order to guarantee to have no asymptotic bias in the estimator. The third condition states that $(1 - \frac{k}{N})$-quantiles of distributions are sufficiently smooth in short $h$-intervals. The last condition assumes that $k$ is compatible with the $h$-variation in the $\gamma$ function.

We first show through an example that the assumptions are consistent and not too restrictive.

**Example 2.1** Suppose $U_s(t) = t^{\gamma(s)}(1 + t^{\rho(s)})$. Then, we get that

$$\frac{U_s(tx) - x^{\gamma(s)}}{U_s(t) - x^{\rho(s)}} = x^{\gamma(s)} \left( x^{\rho(s)} - 1 \right) \left( 1 + t^{\rho(s)} \right)^{-1}.$$  

Further, we assume that $\gamma$ and $\rho$ are continuous functions such that $\gamma(s) > 0$, $\rho(s) < 0$ and $\gamma(s) + \rho(s) \neq 0$ for all $s \in [0,1]$. Denote $\tilde{\rho} = \sup_{0 \leq s \leq 1} \rho(s) < 0$. In addition, we assume Lipschitz continuity for the $\gamma$ and $\rho$ functions: for any $s_1, s_2 \in [0,1]$, $|\gamma(s_1) - \gamma(s_2)| < c_1 |s_1 - s_2|$ and $|\rho(s_1) - \rho(s_2)| < c_2 |s_1 - s_2|$ for some $c_1, c_2 > 0$.

Firstly, condition (2.1) holds with $A_s(t) = t^{\rho(s)}$. 

7
Secondly, we find necessary conditions to ensure condition (2.5). Note that

$$\frac{U_{s_1}(\frac{n}{k}) - 1}{U_{s_2}(\frac{n}{k}) A_{s_2} \left( \frac{n}{k} \right)} = \left[ (\frac{n}{k})^{\gamma(s_1) - \gamma(s_2)} - 1 + (\frac{n}{k})^{\gamma(s_1) - \gamma(s_2) + \rho(s_1) - \rho(s_2)} - 1 \right] \left( 1 + (\frac{n}{k})^{\rho(s_2)} \right)^{-1}.$$ 

Therefore, (2.5) holds if as \( n \to \infty \),

$$h \left( \log \left( \frac{n}{k} \right) \right) \left( \frac{n}{k} \right)^{-\bar{\beta}} \to 0,$$  

(2.6)

We will find proper choices for \( k = k_n \) and \( h = h_n \) such that (2.6) holds together with other constraints on these two series later.

Thirdly, condition (2.3) holds if \( \sqrt{k} \left( \frac{n}{k} \right)^{\bar{\beta}} \to 0 \) as \( n \to \infty \) for some \( \varepsilon > 0 \).

Fourthly, condition (2.4) holds if \( h \sqrt{k} \log k \to 0 \) as \( n \to \infty \).

Combining all aforementioned sufficient conditions, we verify that conditions (2.2)–(2.5) holds for the following choice of \( k \) and \( h \). First choose, \( k = k_n = n^\eta \) for some \( \eta \in \left( \frac{-\bar{\beta}}{\rho + 1}, \frac{-2\bar{\beta}}{2\rho + 1} \right) \). The fact that \( \eta < \frac{-2\bar{\beta}}{2\rho + 1} \) implies that \( \sqrt{k} \left( \frac{n}{k} \right)^{\bar{\beta} + \varepsilon} \to 0 \) as \( n \to \infty \) for some \( \varepsilon > 0 \). Next choose \( h = h_n = \left( \frac{e}{\log n} \right)^{\frac{\bar{\beta}}{\rho + 1}} \), which verifies that as \( n \to \infty \), \( h \sqrt{k} \log k \to 0 \). The latter also implies that as \( n \to \infty \), \( kh^{1+\varepsilon} \to 0 \) as required in (2.2). In addition, \( \eta > \frac{-\bar{\beta}}{\rho + 1} \) implies that as \( n \to \infty \), \( kh^{1+\varepsilon} \to 0 \) for some \( \varepsilon > 0 \), which is required in (2.2). Lastly the limit relation (2.6) holds obviously.

To conclude, we have shown that our required conditions are consistent. Notice that this example can be easily generalized to \( U_s(t) = C(s) t^{\gamma(s)} (1 + D(s) t^{\rho(s)}) \), with some proper continuous function \( C(s) > 0 \) and \( D(s) \) on \([0,1]\). Further the power in the Lipschitz continuity condition of \( \gamma \) and \( \rho \) is not necessarily one, but any positive number. Therefore, our required conditions are not too restrictive.

The following theorem gives the local asymptotic normality for the estimator \( \hat{\gamma}_H(s) \) defined in (1.1).

**Theorem 2.2** Let \( X_1, X_2, \cdots, X_n \) be independent random variables with distributions \( X_i \sim F_x(x) \), where \( F_x(x) \) is a set of distribution functions defined on \( s \in [0,1] \). Assume that \( F_x(x) \) is continuous with respect to \( s \) and \( x \) and \( F_x \in D_{\gamma(s)} \) where \( \gamma(s) \) is a positive
continuous function on $[0, 1]$. Assume the conditions (2.1)–(2.5). Then as $n \to \infty$, we have that for all $s \in [h, 1-h], \sqrt{2kh} (\hat{\gamma}_H(s) - \gamma(s)) \xrightarrow{d} N(0, (\gamma(s))^2)$. The same conditions guarantee the asymptotic normality of the global estimator $\hat{\Gamma}_H(s)$ defined in (1.2).

**Theorem 2.3** Assume the same conditions as in Theorem 2.2. Then there exists a series of Brownian motions $W_n(s)$ such that as $n \to \infty$, 

$$
\sup_{s \in [0, 1]} \left| \sqrt{k} \left( \hat{\Gamma}_H(s) - \Gamma(s) \right) - \int_0^s \gamma(u) dW_n(u) \right| \xrightarrow{P} 0.
$$

**3 Testing trends in extreme value indices**

Theorem 2.3 provides the possibility to test if the extreme value indices follow a specific trend, i.e. $H_0 : \gamma(s) = \gamma_0(s)$ for all $s \in [0, 1]$, with some given function $\gamma_0$. Similar to testing the specific trend in the “skedasis” function in Einmahl et al. (2016), we apply an equivalent test to test $H_0 : \Gamma(s) = \Gamma_0(s) \in [0, 1]$ for all $s$, where $\Gamma_0(s) = \int_0^s \gamma_0(u) du$. Clearly, one may construct a Kolmogorov-Smirnov (KS) type test with the testing statistic defined as 

$$
T := \sup_{s \in [0, 1]} \left| \hat{\Gamma}_H(s) - \Gamma_0(s) \right|
$$

Then, Theorem 2.3 implies that under the null hypothesis $H_0 : \Gamma(s) = \Gamma_0(s)$ for all $s \in [0, 1]$, as $n \to \infty$, 

$$
\sqrt{kT} \xrightarrow{d} \sup_{s \in [0, 1]} \left| \int_0^s \gamma(u) dW(u) \right|,
$$

where $W(u)$ is a standard Brownian motion defined on $[0, 1]$.

It is often of interest to test whether the extreme value index remains constant over time, without prior knowledge on the constant extreme value index, i.e. $H_0 : \gamma(s) = \gamma$ for all $s \in [0, 1]$ without specifying $\gamma$. In this case, one may use $\hat{\Gamma}_H(1)$ as an estimator of
the constant extreme value index $\gamma$ and define the testing statistic as

$$\tilde{T} := \sup_{s \in [0,1]} \left| \frac{\hat{\gamma}_{H}(s)}{\hat{\gamma}_{H}(1)} - s \right|$$

It is straightforward to show that under the null hypothesis $H_0 : \gamma(s)$ is a constant for all $s \in [0,1]$, as $n \to \infty$,

$$\sqrt{k}\tilde{T} \overset{d}{\to} \sup_{s \in [0,1]} |B(s)|,$$

where $B(s)$ is a standard Brownian bridge defined on $[0,1]$. Note that the limit distribution is identical to that in the KS test.

We run a simulation study to demonstrate the finite sample performance of the testing procedure using $\tilde{T}$. In all our simulations we generate $m$ samples with $n$ observations in each sample, where $m = 2000$ and $n = 5000$. For the two parameters $k$ and $h$, we choose several combinations between $k = 100, 200$ and $h = 0.025, 0.04$.

For each sample, we simulate the observations from the following data generating process

$$X_i = Z_i^{1/\gamma(i/n)}, \quad i = 1, 2, \cdots, n,$$

where $\{Z_i\}_{i=1}^n$ are i.i.d. observations from the standard Fréchet distribution with distribution function $F(x) = \exp(-1/x)$ for $x > 0$. For the function $\gamma(s)$ we consider a linear trend as $\gamma(s) = 1 + bs$. If $b = 0$, it resembles the i.i.d. case, i.e. the null hypothesis that the extreme value indices remain constant holds. We consider two alternative cases for which $b = 1$ and $b = 2$. In total, we have 12 sets of simulations due to the various choices of $k$, $h$ and $b$.

For each simulated sample $j$ we apply the test $\tilde{T}$ in Section 3 to test whether the extreme value indices remain constant and obtain the corresponding p-value, $p_j$, for $j = 1, 2, \cdots, m$. For the simulations based on $b = 0$, i.e. when the null hypothesis holds, we plot the histograms of the simulated p-values across all $m$ samples. Figure 1 presents four plots corresponding to four choices of $(k, h)$. We observe that in general the simulated p-values follow a standard uniform distribution on $[0,1]$. Although some deviations are observed for large p-values, particularly for $k = 100$, since p-values are
usually compared to low significance levels such as 1% or 5%, the deviations will not distort the Type I error under the null hypothesis much. This confirms the validity of our test under the null hypothesis.

Figure 1: Histograms of the p-values under the null hypothesis

\[ h = 0.025 \quad \text{and} \quad h = 0.04 \]

\[ k = 200 \]

\[ k = 100 \]

NOTE: The simulated observations are generated from the standard Fréchet distribution with sample size \( n = 5000 \). For each sample, the test \( \tilde{T} \) in Section 3 is applied with the corresponding \( k \) and \( h \) to obtain a p-value. Each graph presents the histogram of the p-values obtained from \( m = 2000 \) samples. The dash line indicates the standard uniform distribution.

Next, for all sets of simulations, we calculate the rejection rate based on each significance level \( \alpha \) as \( \# \{ j : p_j < \alpha \} / m \) for \( \alpha = 0.01, 0.05 \) and 0.1. The rejection rates are reported in Table 3.

In the first panel, we observe that under the null hypothesis, the rejection rates, i.e. the Type I error, are close to the significance levels. Between the two choices of \( h \), the differences are marginal. Between the two choices of \( k \), some over rejections are observed with \( k = 200 \). This might be due to a potential bias in the local estimator \( \hat{\gamma}(s) \) when using a high \( k \).

In the next two panels, the rejections rates can be read as the power of the test.
Table 1: Rejection rates in simulations

<table>
<thead>
<tr>
<th></th>
<th>$k = 200$</th>
<th>$k = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h = 0.025$</td>
<td>$h = 0.04$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.129</td>
</tr>
<tr>
<td>$b = 0$</td>
<td>0.5</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>0.016</td>
</tr>
<tr>
<td>$b = 1$</td>
<td>0.5</td>
<td>0.371</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>0.258</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>0.5</td>
<td>0.760</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>0.500</td>
</tr>
</tbody>
</table>

NOTE: The simulated observations are generated from Fréchet distribution with varying extreme value index defined by $\gamma(s) = 1 + bs$. For each sample with sample size $n = 5000$, the test $\tilde{T}$ in Section 3 is applied with the corresponding $k$ and $h$ to obtain a p-value. The rejection rates are calculated as the number of samples with a p-value below the significance level $\alpha$, divided by the total number of samples $m = 2000$.

Between the two choices of $h$, $h = 0.04$ leads to a slightly higher power, though the improvement is marginal. Between the two choices of $k$, $k = 200$ leads to a much higher power. Therefore choosing a higher $k$ faces the tradeoff between a potential overrejection under the null and the strong power under the alternative.

Lastly, when comparing across the two panels, the power is higher for $b = 2$. This is in line with the intuition that the test is more powerful to detect larger deviation from the null hypothesis of having constant extreme value index.

Finally, for the two alternative: $b = 1$ and $b = 2$, we plot the average of the estimated $\gamma(s)$ across the $m$ samples and its corresponding 95% confidence interval. There are two ways to construct the confidence interval. Firstly, we use the asymptotic theory in Theorem 2.2 to construct the confidence interval based on the averaged estimate of $\gamma(s)$. Secondly, we can obtain an empirical confidence interval from the $m$ estimates. The comparison between the two provides a validation of our asymptotic theory. In this exercise, we fix $k = 200$ and $h = 0.025$.

Figure 2 shows the estimation result for the alternative models. Firstly, the average estimates across $m$ sample resembles the true value of the $\gamma(s)$ function. Secondly, the confidence intervals derived from our asymptotic theory coincide with that obtained from the simulation indicating the validity of our asymptotic theory. In both cases the empirical
confidence interval is shifted slightly upwards compared to the theoretical confidence interval. This is potentially due to the presence of a minor positive bias term neglected in the asymptotic theory.

Figure 2: Estimated varying extreme value indices

(i) $b = 1$

![Graph](image)

(ii) $b = 2$

![Graph](image)

NOTE: The simulated observations are generated from Fréchet distribution with varying extreme value index defined by $\gamma(s) = 1 + bs$. The upper and lower panel show the result for $b = 1$ and $b = 2$ respectively. For each sample with sample size $n = 5000$, the local estimator $\gamma(s)$ in (1.1) is applied with $k = 200$ and $h = 0.025$. The solid line presents the average of the local estimators across $m = 2000$ samples. The dash line in the middle indicates the true value of the $\gamma(s)$ function. The dotted lines indicates the upper and lower bounds of the theoretical 95% confidence intervals, derived from Theorem 2.2, using the averaged estimate at each location $s$. The dash-dotted lines indicate the upper and lower bounds of the empirical 95% confidence intervals obtained from the 2.5% and 97.5% quantiles among the $m$ estimates.
4 Application

We apply our developed method to the losses of stock returns. We intend to test whether the extreme value indices of the losses remain unchanged over time. If rejecting the null, we estimate the time variation in the extreme value indices.

We employ the same dataset as in Einmahl et al. (2016), i.e. the S&P 500 index, from 1988 till 2012. We construct daily loss returns defined as $X_t = \log(P_t/P_{t+1})$, where $P_t$ is the index on day $t$. This results in a sample with 6302 observations. Since the sample size is similar to the simulation study, we choose $h = 0.04$.

Similar to Einmahl et al. (2016), we also test the constancy of the extreme value indices over a subperiod from 1988 to 2012 (6302 observations). The result is shown in the upper panel of Figure 3. We do not reject the null hypothesis for $k$ up to 600. This result differs from the conclusion in Einmahl et al. (2016) where the constancy of the extreme value index in the period from 1988 to 2012 was rejected. We interpret the difference by the fact that we are testing a broader null. Notice that in the heteroskedastic extremes model in Einmahl et al. (2016), the skedasis function is assumed to be bounded away from 0 and $+\infty$, whereas our model potentially allows for the unbounded case in the null hypothesis.

Since the null hypothesis is not rejected for the sample from 1988 to 2012, we consider an extended sample from 1963 to 2012 (12586 observations). With such a large sample, we choose a lower $h = 0.02$ in line with the theoretical requirement on $h$. The result is shown in the lower panel of Figure 3. We reject having a constant extreme value index during this long period for $k$ ranging from 250 to 500. Hence, we conclude that there is a change of extreme value index during the period from 1963 to 2012.

Finally, we plot the estimated $\Gamma(s)$ function and the corresponding 95% confidence band uniformly for all $s \in [0,1]$ in Figure 4 for the period from 1963 to 2012. In this analysis, we use $k = 400$, for which the null hypothesis of having a constant extreme value index was rejected. We obtain the confidence band in two ways.

Without having prior information on the shape of $\Gamma(s)$, we obtain from Theorem 2.3
Figure 3: Testing the constancy of the extreme value indices: S&P 500 index

(i) 1988–2012

![Graph showing p-values for k values from 200 to 1000, for observations in 1988–2012 (6302 observations, h = 0.04).]

(ii) 1963–2012

![Graph showing p-values for k values from 200 to 1000, for observations in 1963–2012 (12586 observations, h = 0.02).]

NOTE: The two figures show the results for testing the constancy of extreme value indices based on the loss returns of the S&P 500 index. The test $\tilde{T}$ in Section 3 is applied with various levels of $k$. The figures show the p-values of the test against the corresponding $k$. The upper and lower panels show the result based on observations in 1988–2012 (6302 observations, $h = 0.04$) and 1963-2012 (12586 observations, $h = 0.02$) respectively. The dash line indicates the level of p-value at 0.05.

that as $n \to \infty$, 

$$\sup_{s \in [0,1]} \left| \sqrt{k} \left( \hat{\Gamma}_H (s) - \Gamma(s) \right) \right| \overset{d}{\to} \sup_{s \in [0,1]} \left| \int_0^s \gamma(u) dW(u) \right|,$$

where $W(u)$ is a Brownian Motion. We simulate the quantile of the limit and use that for constructing the uniform confidence band. Since the limit distribution involves the func-
tion \( \gamma(s) \), we plug the estimate of the \( \gamma(s) \) function into the stochastic integral and simulate the statistic \( \sup_{s \in [0,1]} \left| \int_0^s \hat{\gamma}(u) dW(u) \right| \) one million times. Then we take the numerical 95\% quantile among the one million simulations, denoted as \( q(0.95) \). Eventually, the uniform confidence band can be constructed as \( [\hat{\Gamma}(s) - q(0.95) / \sqrt{k}, \hat{\Gamma}(s) + q(0.95) / \sqrt{k}] \). The upper and lower bound for the confidence band are shown by the dotted lines in the upper panel of Figure 4. The dash line indicates the function \( s \hat{\Gamma}(1) \), which corresponds to the case \( \gamma(s) = \gamma \) for all \( s \in [0,1] \). The dash line lies always within the confidence band which is seemingly contradictory to our testing result. Notice that the construction of the uniform confidence band is not based on the null hypothesis in the testing analysis. The width of the band is therefore relatively wider due to the stochastic integral.

To be consistent with the testing procedure, we construct uniform confidence band under the null hypothesis that \( \gamma(s) = \gamma \) for all \( s \in [0,1] \) in the lower panel of Figure 4. More specifically, we use the quantile of the limit distribution \( \hat{\Gamma}(1) \sup_{s \in [0,1]} |B(s)| \) where \( B(s) \) is a Browanian bridge, instead of the simulated \( q(0.95) \) in determining the uniform width of the confidence band. The result shows that the straight line \( s \hat{\Gamma}(1) \) falls out of the uniform confidence band in the region \( s \in [0.4, 0.5] \). This is in line with our testing result.
Figure 4: Estimating the $\Gamma(s)$ function: S&P 500 index

(i) Confidence band without prior knowledge on $\gamma(s)$

(ii) Confidence band with assuming constant extreme value index

NOTE: The two figures show the estimated $\Gamma(s)$ function based on the loss returns of the S&P 500 index in 1963–2012 (12586 observations). The solid line shows the point estimates with $h = 0.02$ and $k = 300$. The dash line is a straight line $s\Gamma(1)$ for $s \in [0, 1]$. The dotted lines indicate the upper and lower bounds of the 95% confidence interval uniformly for all $s \in [0, 1]$. In the upper panel, the uniform confidence band is constructed without using any prior knowledge on the shape of $\Gamma(s)$. In the lower panel, the uniform confidence band is constructed under the assumption that $\gamma(s)$ is constant for all $s \in [0, 1]$. 
A Appendix: Proofs

We start with presenting auxiliary results that are necessary for the proof of our main theorems. Then we establish asymptotic properties for the “local tail empirical process”. Finally, we provide the proof for the two main theorems.

A.1 Auxiliary results

The following lemma shows that under the conditions of Theorem 2.2, the quantile function in the regularly variation property is exchangeable in a $h$–neighborhood. Denote $q_n = k^{1+\varepsilon}$ for any $\varepsilon > 0$.

**Lemma A.1** Under the conditions of Theorem 2.2, as $n \to \infty$,

$$\sqrt{k} \sup_{|s_1-s_2| \leq h, 1/2 \leq x \leq q_n} \left| \frac{U_{s_1}(\frac{n}{k}x)}{U_{s_2}(\frac{n}{k})} x^{\gamma(s_2)} - 1 \right| \to 0. \quad (A.1)$$

**Proof of Lemma A.1.** Write

$$\frac{U_{s_1}(\frac{n}{k}x)}{U_{s_2}(\frac{n}{k})} x^{\gamma(s_2)} = \frac{U_{s_1}(\frac{n}{k}x)}{U_{s_1}(\frac{n}{k})} \cdot \frac{U_{s_1}(\frac{n}{k})}{U_{s_2}(\frac{n}{k})} \cdot x^{\gamma(s_1)-\gamma(s_2)} =: I_1 \cdot I_2 \cdot I_3$$

Firstly, the second order condition (2.1) ensures that (cf. Theorem 2.3.8 in de Haan and Ferreira (2006)) uniformly for all $s_1 \in [0,1], x \geq 1/2$, as $n \to \infty$,

$$\left| \frac{I_1 - 1}{A_{s_1}(\frac{n}{k})} \right| = O(1).$$

Together with the condition (2.3), we get that as $n \to \infty$,

$$\sqrt{k} \sup_{s_1 \in [0,1], x \geq 1/2} |I_1 - 1| = o(1).$$

Secondly, the conditions (2.3) and (2.5) ensure that $n \to \infty$,

$$\sqrt{k} \sup_{|s_1-s_2| \leq h} |I_2 - 1| = o(1).$$
Lastly, from the condition (2.4), we get that as $n \to \infty$

$$\sqrt{k} \sup_{|s_1 - s_2| \leq h, 1/2 \leq x \leq q_n} |\log x| |\gamma(s_1) - \gamma(s_2)| < 2\sqrt{k} |\log k| \sup_{|s_1 - s_2| \leq h} |\gamma(s_1) - \gamma(s_2)| \to 0.$$  

It implies that $\sqrt{k} \sup_{|s_1 - s_2| \leq h, 1/2 \leq x \leq q_n} |I_3 - 1| = o(1)$.

The lemma is then proved by combining the three components. $\blacksquare$

For any given $s \in [0, 1]$, denote $\mathcal{U}_{s,n} = \max\{i: |s_i - s| \leq h\} U_{i/n}$ and $\mathcal{U}_{s,n} = \min\{i: |s_i - s| \leq h\} \mathcal{U}_{i/n}$. Obviously $\mathcal{U}_{s,n} \leq U_s \leq \mathcal{U}_{s,n}$. The following is a direct corollary from Lemma A.1.

**Corollary A.2** As $n \to \infty$,

$$\Delta_{4,n} := \sqrt{k} \sup_{0 \leq s \leq 1} \sup_{1/2 \leq x \leq q_n} \left| \frac{\mathcal{U}_{s,n} \left( \frac{n}{k} x \right)}{U_s \left( \frac{n}{k} x \right)} x^{\gamma(s)} - 1 \right| \to 0,$$

$$\Delta_{5,n} := \sqrt{k} \sup_{0 \leq s \leq 1} \sup_{1/2 \leq x \leq q_n} \left| \frac{\mathcal{U}_{s,n} \left( \frac{n}{k} x \right)}{U_s \left( \frac{n}{k} x \right)} x^{\gamma(s)} - 1 \right| \to 0.$$  

### A.2 Proof of Theorem 2.2

We prove the theorem by constructing upper and lower bounds for the local estimator $\hat{\gamma}_H(s)$ at a fixed $s$.

The local estimator is based on observations $X_i$ with $i \in I_n(s) = \{i: |s_i - s| \leq h\}$. Write $X_i = U_{i/n}(Z_i)$ where $\{Z_i: i \in I_n(s)\}$ are $[2nh]$ i.i.d. standard Pareto distributed random variables. To construct the local Hill estimator, we rank the observations $\{X_i: i \in I_n(s)\}$ into order statistics as $X_{1, [2nh]}^{(s)} \leq \cdots \leq X_{[2nh], [2nh]}^{(s)}$. We also rank $\{Z_i: i \in I_n(s)\}$ into order statistics as $Z_{1, [2nh]}^{(s)} \leq \cdots \leq Z_{[2nh], [2nh]}^{(s)}$. Notice that since the $U_{i/n}$ functions are different, the order statistic $X_{j, [2nh]}^{(s)}$ may not correspond to the order statistic $Z_{j, [2nh]}^{(s)}$ for $j = 1, 2, \cdots, [2nh]$.

Nevertheless, recall the notations $\mathcal{U}_{s,n} = \max\{i: |s_i - s| \leq h\} U_{i/n}$ and $\mathcal{U}_{s,n} = \min\{i: |s_i - s| \leq h\} \mathcal{U}_{i/n}$. The inequalities

$$\mathcal{U}_{s,n}(Z_{j, [2nh]}^{(s)}) \leq X_{j, [2nh]}^{(s)} \leq \mathcal{U}_{s,n}(Z_{j, [2nh]}^{(s)}).$$
for all $j = 1, 2, \cdots, [2nh]$, are obtained as follows. Take the upper bound as an example. Since there are $j$ random variables among $\{Z_i : i \in I_n(s)\}$ that are lower or equal to $Z_{j,[2nh]}^{(s)}$, and the $U$-functions corresponding to these $Z_i$ are all bounded below $U$, we get that there are at least $j$ random variables among $\{X_i : i \in I_n(s)\}$ that are bounded below by $U_{s,n}(Z_{j,[2nh]}^{(s)})$. This proves the inequality for the upper bound. A similar argument can be made for the lower bound.

Therefore, we get an upper bound for $\hat{\gamma}_H(s)$ as follows,

$$\hat{\gamma}_H(s) = \frac{1}{2kh} \sum_{j=1}^{[2kh]} \left( \log X_{[2nh]-j+1,[2nh]}^{(s)} - \log X_{[2kh],[2nh]}^{(s)} \right) \leq \frac{1}{2kh} \sum_{j=1}^{[2kh]} \left( \log U_{s,n}(Z_{[2nh]-j+1,[2nh]}^{(s)}) - \log U_{s,n}(Z_{[2kh],[2nh]}^{(s)}) \right) = \frac{1}{2kh} \sum_{j=1}^{[2kh]} \left( \log \frac{U_{s,n}(Z_{[2nh]-j+1,[2nh]}^{(s)})}{U_{s,n}(Z_{[2kh],[2nh]}^{(s)})} \right). \quad (A.2)$$

We would now apply Corollary A.2 to bound the two terms in (A.2). For that purpose it is necessary to check that $\frac{k}{n} Z_{[2nh]-j+1,[2nh]}^{(s)} \in [1/2, q_n]$ for all $j = 1, 2, \cdots, [2kh] + 1$.

For the lower bound of $1/2$, it follows from the fact that $\frac{k}{n} Z_{[2nh]-j+1,[2nh]}^{(s)} \overset{P}{\to} 1$, as $n \to \infty$. For the upper bound, notice that, as $n \to \infty$, with probability 1,

$$\frac{k}{n} Z_{[2nh],[2nh]}^{(s)} \leq k \frac{\max_{1 \leq i \leq n} Z_i}{n} = kO_P(1) \leq k^{1+\epsilon} = q_n.$$

Here, we consider the maxima over all $Z_i$ across all observations, and use the fact that $\frac{\max_{1 \leq i \leq n} Z_i}{n} = O_P(1)$ as $n \to \infty$. In this way, we have verified the upper bound $q_n$ uniformly for all $s \in [h, 1 - h]$.

Now, we are ready to apply Corollary A.2 to the two terms in (A.2) and continue the
inequality as
\[
\hat{\gamma}_H(s) \leq \frac{1}{2kh} \sum_{j=1}^{[2kh]} \left( \gamma(s) \log \frac{Z_{[2nh]-j+1,[2nh]}^{(s)}}{n/k} + \log \left( 1 + \frac{\Delta_{4,n}}{\sqrt{k}} \right) 
- \gamma(s) \log \frac{Z_{[2nh],[2nh]}^{(s)}}{n/k} - \log \left( 1 - \frac{\Delta_{5,n}}{\sqrt{k}} \right) \right)
\]
\[
= \gamma(s) \frac{1}{2kh} \left( \sum_{j=1}^{[2kh]} \log \frac{Z_{[2nh]-j+1,[2nh]}^{(s)}}{Z_{[2kh],[2nh]}^{(s)}} \right) + \frac{[2kh]}{2kh} \left( \log \left( 1 + \frac{\Delta_{4,n}}{\sqrt{k}} \right) - \log \left( 1 - \frac{\Delta_{5,n}}{\sqrt{k}} \right) \right)
\]
\[
= \gamma(s) J_n(s) + \frac{1}{\sqrt{k}} o(1),
\] (A.3)

where \( J_n(s) = \frac{1}{2kh} \sum_{j=1}^{[2kh]} \log \frac{Z_{[2nh]-j+1,[2nh]}^{(s)}}{Z_{[2kh],[2nh]}^{(s)}} \). Note that the remainder part \( \frac{1}{\sqrt{k}} o(1) \) is non-stochastic, and the \( o(1) \) term depends only on \( \Delta_{4,n} \) and \( \Delta_{5,n} \) which is uniformly negligible for all \( s \in [h, 1-h] \).

Similarly, one can establish a lower bound for \( \hat{\gamma}_H(s) \) as
\[
\hat{\gamma}_H(s) \geq \gamma(s) J_n(s) + \frac{1}{\sqrt{k}} o(1).
\]

Hence, we get that
\[
\sqrt{2kh} (\hat{\gamma}_H(s) - \gamma(s)) = \sqrt{2kh} \gamma(s) (J_n(s) - 1) + \sqrt{2h} o(1).
\]

To obtain the theorem, we only need to show that \( \sqrt{2kh}(J_n(s) - 1) \) converges to a standard normal distribution as \( n \to \infty \).

If we disregard the order, then the set \( \left\{ \log \frac{Z_{[2nh]-j+1,[2nh]}^{(s)}}{Z_{[2kh],[2nh]}^{(s)}} : 1 \leq j \leq [2kh] \right\} \) is a set of i.i.d. standard exponentially distributed random variables. We denote them, without loss
of generality, as \( \{ E_j(s) \}_{j=1}^{[2kh]} \). From the central limit theorem, we get that

\[
\sqrt{2kh} (J_n(s) - 1) = \sqrt{2kh} \left( \frac{1}{2kh} \sum_{j=1}^{[2kh]} \log E_j(s) - 1 \right)
\]

\[
= \sqrt{\frac{2kh}{2kh}} \cdot \sqrt{2kh} \left( \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \log E_j(s) - 1 \right) + \frac{2kh - 2kh}{\sqrt{2kh}}
\]

\[
\xrightarrow{d} N(0, 1),
\]

as \( n \to \infty \). Notice that here we use the fact that \( kh \to \infty \), as \( n \to \infty \).

### A.3 Proof of Theorem 2.3

Next, we prove the global asymptotic normality in Theorem 2.3, which is a uniform result over all \( s \in [0, 1] \). We deal with \( s \in [0, h] \) and \( s \in [h, 1] \) differently. We first handle the more difficult case \( s \in [h, 1] \), while dealing with \( s \in [0, h] \) at the end.

Recall the definition of \( \hat{\Gamma}_H(s) \) in (1.2) as

\[
\hat{\Gamma}_H(s) = 2h \sum_{s_t \leq s} \hat{\gamma}_H(s),
\]

where \( s_t = (2t - 1)h \) for \( t = 1, 2, \cdots, \lfloor 1/2h \rfloor \).

We first revisit the proof of Theorem 2.2 in order to get the inequality (A.3) uniformly for all \( s_t, t = 1, 2, \cdots, \lfloor 1/2h \rfloor \). One important step to achieve this goal is to prove the statement of Corollary A.2 uniformly for all \( s_t \). The following lemma guarantees that one can do this.

**Lemma A.3** Under the same conditions as in Theorem 2.3, as \( n \to \infty \)

\[
P \left( \frac{k}{n} Z_{[2nh]-j+1, [2nh]}(s_t) \in [1/2, q_n] \text{ for all } j = 1, 2, \cdots, [2kh] + 1 \text{ and all } t = 1, 2, \cdots, \lfloor 1/2h \rfloor \right) \to 1.
\]

**Proof of Lemma A.3.** The validity of the upper bound \( q_n \) in the lemma has already been proved in the proof of Theorem 2.2 as follows. As \( n \to \infty \), with probability tending
to 1, $\frac{k}{n} Z^{(s)}_{[2nh],[2nh]} \leq q_n$ holds uniformly for all $s$. Consequently, it holds for all $s_t$.

For the lower bound, it is difficult to have a similar result for all $s$. We therefore prove the following weaker version: as $n \to \infty$, with probability tending to 1, $\frac{k}{n} Z^{(s)}_{[2nh],[2kh],[2nh]} \geq 1/2$ holds for all $t = 1, 2, \cdots, [1/2h]$. Notice that the order statistics $Z^{(s)}_{[2nh],[2kh],[2nh]}$ are independent and identically distributed across $t = 1, 2, \cdots, [1/2h]$. This implies that

$$P\left(\frac{k}{n} Z^{(s)}_{[2nh],[2kh],[2nh]} \geq 1/2 \text{ for all } t = 1, 2, \cdots, [1/2h]\right) = \left(P\left(\frac{k}{n} Z^{(s)}_{[2nh],[2kh],[2nh]} \geq 1/2\right)\right)^{[1/2h]} .$$

We continue to establish the lower bound for the probability regarding $s_1$ using a generalized Chebyshev’s inequality as follows. Without loss of generality, we omit the superscript $(s_1)$. Since $Z_{[2nh],[2kh],[2nh]}$ is the $([2nh] - [2kh])$-th order statistic among $[2nh]$ i.i.d. standard Pareto distributed random variables, $1/Z_{[2nh],[2kh],[2nh]}$ is the $([2kh] + 1)$-th order statistic among $[2nh]$ i.i.d. uniformly distributed random variables, denoted as $U_{[2kh]+1,[2nh]}$, and $P\left(\frac{k}{n} Z_{[2nh],[2kh],[2nh]} \geq 1/2\right) = P\left(\frac{n}{k} U_{[2kh]+1,[2nh]} \leq 2\right)$. Hence,

$$P\left(\frac{n}{k} U_{[2kh]+1,[2nh]} > 2\right) \leq P\left(\left|\frac{n}{k} U_{[2kh]+1,[2nh]} - 1\right| > 1\right) \leq M_l,$$

where $M_l := \mathbb{E}\left|\frac{n}{k} U_{[2kh]+1,[2nh]} - 1\right|^l$, for any $l \in \mathbb{N}$.

Lastly, we calculate the quantity $M_l$. Recall the Inequality (1.1) in Theorem 1 in Csörgő et al. (1986). There exists a proper probability space on which one can define a series of i.i.d. $U(0,1)$ random variables $U_1, U_2, \cdots$ and a sequence of Brownian bridges $B_i(s), i = 1, 2, \cdots$ such that

$$P\left(\sup_{0 \leq u \leq d_n/n} \left|n^{1/2}(u - U_{[un],[n]} - B_n(u))\right| \geq n^{-1/2}(a \log d_n + x)\right) \leq be^{-cx},$$

for all $n_0 \leq d_n \leq n$ and $x > 0$, where $n_0$, $a$, $b$ and $c$ are suitably chosen constants.

We apply this inequality to the $[2nh]$ uniform random variables. Corresponding, we replace $n$ by $[2nh]$. Further we take a specific $u = d_n/[2nh]$. We get that

$$P\left(\left|[2nh]^{1/2}\left(\frac{d_n}{[2nh]} - U_{d_n,[2nh]}\right) - B_n\left(\frac{d_n}{[2nh]}\right)\right| \geq ([2nh])^{-1/2}(a \log d_n + x)\right) \leq be^{-cx},$$
which can be organized as
\[
P \left( \left| d_n^{1/2} \left( 1 - \frac{2nh}{d_n} U_{d_n,2nh} \right) - \sqrt{\frac{2nh}{d_n}} B_n \left( \frac{d_n}{2nh} \right) \right| \geq d_n^{-1/2} \left( a \log d_n + x \right) \right) \leq b e^{-cx}.
\] (A.4)

We will later apply this inequality with substituting \( d_n \) by \( [2kh] + 1 \). Such a substitution is allowed due to the fact that \( kh \to \infty \) as \( n \to \infty \).

Since as \( n \to \infty \), \( \frac{n}{k} \frac{[2kh]+1}{[2nh]} \to 1 \), we get that
\[
\left| \frac{n}{k} U_{[2kh]+1,2nh} - 1 \right| \leq 2 \left( \frac{[2nh]}{[2kh]+1} U_{[2kh]+1,2nh} - 1 \right) + 1
\]
\[
\leq 2 \frac{[2nh]}{2kh} U_{[2kh]+1,2nh} - 1 + 1
\]
\[
\leq 2 \frac{[2nh]}{d_n^{1/2}} \left( \sqrt{\frac{[2nh]}{d_n}} B_n \left( \frac{d_n}{2nh} \right) + \Theta_n + d_n^{-1/2} \right),
\]
where \( d_n = [2kh] + 1 \) and \( \Theta_n = \left| d_n^{1/2} \left( 1 - \frac{2nh}{d_n} U_{d_n,2nh} \right) - \sqrt{\frac{2nh}{d_n}} B_n \left( \frac{d_n}{2nh} \right) \right| \). From applying inequality (A.4) with \( d_n = [2kh] + 1 \), we get that
\[
P \left( \Theta_n \geq d_n^{-1/2} \left( a \log d_n + x \right) \right) \leq b e^{-cx}.
\] (A.5)

Finally, we can calculate the moment \( M_l \) as
\[
M_l = E \left| \frac{n}{k} U_{[2kh]+1,2nh} - 1 \right|^l \leq 2^l d_n^{1/2} E \left( \left| \sqrt{\frac{[2nh]}{d_n}} B_n \left( \frac{d_n}{2nh} \right) + \Theta_n + d_n^{-1/2} \right|^l \right)
\]
\[
\leq 2^l d_n^{1/2} \left( E \left| \sqrt{\frac{[2nh]}{d_n}} B_n \left( \frac{d_n}{2nh} \right) \right|^l + E \Theta_n + d_n^{-l/2} \right)
\]

The term \( \sqrt{\frac{[2nh]}{d_n}} B_n \left( \frac{d_n}{2nh} \right) \) follows a normal distribution with mean zero and variance
\[
\frac{[2nh]}{d_n} \left( 1 - \frac{d_n}{[2nh]} \right) = 1 - \frac{d_n}{[2nh]} < 1.
\]
It implies that \( E \left| \sqrt{\frac{[2nh]}{d_n}} B_n \left( \frac{d_n}{2nh} \right) \right|^l < C_l \) for a
constant $C_l$. For the other term $\Theta_n$, we calculate its moment by using (A.5) as follows:

$$E\Theta_n^l = \int_0^\infty P(\Theta_n^l > x)dx = \int_0^\infty P(\Theta_n > x)lx^{l-1}dx$$

$$\leq \int_0^{d_n^{-1/2}a \log d_n} P(\Theta_n > x)lx^{l-1}dx + \int_0^{\infty} be^{-cx}l (d_n^{-1/2}(a \log d_n + x))^{l-1} d_n^{-1/2}dx$$

$$\leq (d_n^{-1/2}a \log d_n)^l + 2^{l-1}l (d_n^{-1/2} \int_0^{\infty} be^{-cx}(a \log d_n)^{l-1} dx + d_n^{-1/2} \int_0^{\infty} bx^{l-1}e^{-cx}dx)$$

$$= \left(\frac{a \log d_n}{d_n^{l/2}}\right)^l + 2^{l-1}l \left(\frac{b}{c} \frac{(a \log d_n)^{l-1}}{d_n^{l/2}} + \frac{b}{c} (l-1)! \right)$$

Since $d_n \to \infty$ as $n \to \infty$, we get that $E\Theta_n^l \to 0$. By combining the three terms, we get that for any $l$,

$$M_l \leq d_n^{-l/2}D_l = \frac{D_l}{(2kh^l + 1)^{l/2}} \leq \frac{D_l}{(2kh)^{l/2}}$$

where $D_l$ is a constant that only depends on $l$.

With this upper bound on $M_l$, we can continue with the generalized Chebyshev’s inequality to obtain that

$$P\left(\frac{k}{n}Z_{[2nh]-[2kh],[2nh]} \geq 1/2\right) = 1 - P\left(\frac{n}{k}U_{[2kh]-1,[2nh]} > 2\right) \geq 1 - M_l \geq 1 - \frac{D_l}{(2kh)^{l/2}}.$$

Hence,

$$P\left(\frac{k}{n}Z_{[2nh]-[2kh],[2nh]} \geq 1/2 \text{ for all } t = 1, 2, \cdots, [1/2h]\right) \geq \left(1 - \frac{D_l}{(2kh)^{l/2}}\right)^{[1/2h]}.$$

As $n \to \infty$, we have that $\frac{D_l}{(2kh)^{l/2}}^{1/[2h]} \leq D_l^{1/(2h^{l/2})}$, for some positive constant $D'_l$. As $n \to \infty$, since $h \to 0$ and $kh^{1+\varepsilon} \to \infty$, by taking an integer $l > 2/\varepsilon$, we get that $k^{l/2}h^{l/2+1} \to \infty$, which implies that $\left(1 - \frac{D_l}{(2kh)^{l/2}}\right)^{[1/2h]} \to 1$. We thus get that

$$\lim_{n \to \infty} P\left(\frac{k}{n}Z_{[2nh]-[2kh],[2nh]} \geq 1/2 \text{ for all } t = 1, 2, \cdots, [1/2h]\right) = 1,$$

which gives the uniform lower bound.

We continue with the proof of Theorem 2.3. Recall that in the proof of Theorem 2.2, inequality (A.3) is proved for each given $s$. From the uniform bounds for all $s_t$
given in Lemma A.3, we can now obtain the inequality (A.3) uniformly for all $s_t$, $t = 1, 2, \cdots, [1/2h]$ as

$$
\hat{\gamma}_H(s_t) \leq \gamma(s_t)J_n(s_t) + \frac{1}{\sqrt{k}}o(1),
$$

where $J_n(s_t) = \frac{1}{2kh} \sum_{j=1}^{[2kh]} \log \frac{Z^{(s_t)}_{[2nh]-j+1, [2nh]}}{Z^{(s_t)}_{[2kh], [2nh]}}$ and the $o(1)$ term depends only on $\Delta_{4,n}$ and $\Delta_{5,n}$ which tends to zero uniformly for all $s_t$.

Again, we regard the set \( \left\{ \log \frac{Z^{(s_t)}_{[2nh]-j+1, [2nh]}}{Z^{(s_t)}_{[2kh], [2nh]}} : 1 \leq j \leq [2kh] \right\} \) as a set of i.i.d. standard exponentially distributed random variables denoted as \( \{ E^{(s_t)}_j : 1 \leq j \leq [2kh] \} \). Since the sets $I(s_t)$ are disjoint, we can collect all $E^{(s_t)}_j$ for all $1 \leq j \leq [2kh]$ and $t = 1, 2, \cdots, [1/2h]$ and regard them as $[2kh][1/2h]$ i.i.d. standard exponentially distributed random variables.

We continue with the upper bound for all $\hat{\gamma}_H(s_t)$ to establish an upper bound for $\hat{\Gamma}_H(s)$ as

$$
\hat{\Gamma}_H(s) = 2h \sum_{s_t \leq s} \hat{\gamma}_H(s_t) \leq 2h \sum_{s_t \leq s} \gamma(s_t) \frac{1}{2kh} \sum_{j=1}^{[2kh]} E^{(s_t)}_j + s \frac{1}{\sqrt{k}}o(1),
$$

which implies that

$$
\sqrt{k} \left( \hat{\Gamma}_H(s) - 2h \sum_{s_t \leq s} \gamma(s_t) \right) \leq \frac{1}{\sqrt{k}} \sum_{s_t \leq s} \gamma(s_t) \sum_{j=1}^{[2kh]} (E^{(s_t)}_j - 1) + s \cdot o(1).
$$

With establishing a similar lower bound, we get that as $n \to \infty$,

$$
\sup_{s \in [h, 1]} \left| \sqrt{k} \left( \hat{\Gamma}_H(s) - 2h \sum_{s_t \leq s} \gamma(s_t) \right) - \frac{1}{\sqrt{k}} \sum_{s_t \leq s} \gamma(s_t) \sum_{j=1}^{[2kh]} (E^{(s_t)}_j - 1) \right| \overset{P}{\to} 0 \quad (A.6)
$$

The theorem is proved for $s \in [h, 1]$ by combining (A.6) with the following two limit relations: as $n \to \infty$,

$$
\sup_{s \in [h, 1]} \sqrt{k} \left| 2h \sum_{s_t \leq s} \gamma(s_t) - \int_0^s \gamma(u)du \right| \to 0; \quad (A.7)
$$

26
and there exists a series of Brownian motion $W_n(s)$ such that, as $n \to \infty$,

$$
\sup_{s \in [h, 1]} \left| \frac{1}{\sqrt{k}} \sum_{s_t \leq s} \gamma(s_t) \sum_{j=1}^{[2kh]} (E_j^{(s_t)} - 1) - \int_0^s \gamma(u) dW_n(u) \right| \overset{P}{\to} 0. \quad (A.8)
$$

Firstly, we handle the deterministic relation (A.7). Denote $t_s = \max \{t : s_t \leq s\}$.

The following inequality gives a uniform upper bound for the term $|2h \sum_{s_t \leq s} \gamma(s_t) - \int_0^s \gamma(u) du|$.

$$
\left| 2h \sum_{s_t \leq s} \gamma(s_t) - \int_0^s \gamma(u) du \right| = \left| 2h \sum_{t=1}^{t_s} \gamma(s_t) - \sum_{t=1}^{t_s} \int_{s_t-h}^{s_t+h} \gamma(u) du - \int_{s_t-h}^{s_t} \gamma(u) du \right|
\leq \sum_{t=1}^{t_s} \int_{s_t-h}^{s_t+h} |\gamma(s_t) - \gamma(u)| du + \bar{\gamma} 2h \leq \frac{\Delta_{2,n}}{\sqrt{k \log k}} + \bar{\gamma} h,
$$

where $\bar{\gamma} = \sup_{0 \leq s \leq 1} \gamma(s) > 0$. Here, we used the definition of $\Delta_{2,n}$ in (2.4). Since $kh^2 \to 0$ as $n \to \infty$, we get the limit relation (A.7).

Next, we handle (A.8) by constructing a partial sum process as follows. The random variables $E_j^{(s_t)}$ for $j = 1, 2, \cdots, [2kh]$ and $t = 1, 2, \cdots, [1/2h]$ is a sequence of $m = [2kh][1/2h]$ independent standard exponential random variables. For convenience, we relabel $E_j^{(s_t)} - 1$ as $Y_1, Y_2, \cdots, Y_m$, i.e. $Y_{(t-1)[2kh]+j} = E_j^{(s_t)} - 1$. In addition, it is clear that $\mathbb{E}(e^{Y_t}) = e^{-t} \frac{1}{1-t}$ for all $t < 1$. Consider the partial sum process

$$
S_n(u) = \frac{1}{\sqrt{m}} \sum_{i=1}^{[mu]} Y_i,
$$

for $u \in [0, 1]$. By verifying the condition in Theorem 2.2 (ii) in Csörgő and Horváth (1993), we obtain from that theorem that there exists a series of Brownian motions $W_n(u)$ and a constant $C$ such that

$$
S_n(u) = W_n(u) + \theta_n(u),
$$
where as \( n \to \infty \), with probability tending to one,

\[
\sup_{u \in [0,1]} |\theta_n(u)| \leq C \frac{\log m}{\sqrt{m}}
\]

holds. Notice that \( m/k \to 1 \) as \( n \to \infty \). With a constant \( C' > C \), we have that as \( n \to \infty \), with probability tending to one,

\[
\sup_{u \in [0,1]} |\theta_n(u)| \leq C' \frac{\log k}{\sqrt{k}}
\]

(A.9)

Recall the notation \( t_s = \max\{t : s_t \leq s\} \). Write

\[
\frac{1}{\sqrt{k}} \sum_{s_t \leq s} \gamma(s_t) \sum_{j=1}^{[2kh]} (E_j - 1)
\]

\[
= \frac{\sqrt{m}}{\sqrt{k}} \sum_{t=1}^{t_s} \gamma(s_t)(S_n(s_t + h) - S_n(s_t - h))
\]

\[
= \sqrt{\frac{m}{k}} \left( \sum_{t=1}^{t_s} \gamma(s_t)(W_n(s_t + h) - W_n(s_t - h)) \right) + \sqrt{\frac{m}{k}} \left( \sum_{t=1}^{t_s} \gamma(s_t)(\theta_n(s_t + h) - \theta_n(s_t - h)) \right)
\]

\[
=: I_1(s) + I_2(s)
\]

We first deal with \( I_2(s) \). Note that \( s_t + h = 2th \) and \( s_t - h = 2(t - 1)h \), by using (A.9), we get that

\[
|I_2(s)| = \sqrt{\frac{m}{k}} \left( \sum_{t=1}^{t_s-1} \sup_{u \in [0,1]} |\theta_n(u)| |\gamma(s_t) - \gamma(s_{t+1})| + \gamma(s_t) |\theta_n(2th)| \right)
\]

\[
\leq 2 \sup_{u \in [0,1]} |\theta_n(u)| \left[ [1/2h] \frac{\Delta_{2,n}}{\sqrt{k \log k}} + \bar{\gamma} \right]
\]

\[
\leq 2 \sup_{u \in [0,1]} |\theta_n(u)| \left[ [1/2h] \frac{\Delta_{2,n}}{\sqrt{k \log k}} + \bar{\gamma} \right]
\]

\[
\leq 2C' \frac{\log k}{\sqrt{k}} \left[ [1/2h] \frac{\Delta_{2,n}}{\sqrt{k \log k}} + \bar{\gamma} \right]
\]

\[
\leq 2C' \frac{\log k}{\sqrt{k}} \left( \frac{\Delta_{2,n}}{2kh} + \bar{\gamma} \frac{\log k}{\sqrt{k}} \right).
\]

Since the right hand side does not depend on \( s \) and converges to zero as \( n \to \infty \), we get
that \( \sup_{s \in [h, 1]} |I_2(s)| \xrightarrow{P} 0 \) as \( n \to \infty \).

Next, we deal with \( I_1(s) \). We show that it is close to the stochastic integral in the theorem. Firstly, we ignore the term \( \sqrt{\frac{m}{k}} \) and consider the difference between the random part and the stochastic integral. Using the Cauchy’s inequality, we get that

\[
\left| \sum_{t=1}^{s} \gamma(s_t)(W_n(s_t + h) - W_n(s_t - h)) - \int_0^s \gamma(u) dW_n(u) \right| \\
\leq \left| \sum_{t=1}^{s} \int_{s_t-h}^{s_t+h} (\gamma(s_t) - \gamma(u)) dW_n(u) \right| + \left| \int_0^s \gamma(u) dW_n(u) \right| \\
\leq \sqrt{\sum_{t=1}^{s} \int_{s_t-h}^{s_t+h} (\gamma(s_t) - \gamma(u))^2 du} \sqrt{\sum_{t=1}^{s} (dW_n(u))^2} + \int_0^s (\gamma(u))^2 du \int_0^s (dW_n(u))^2 \\
= \frac{\Delta_{2,n}^2}{\sqrt{k} \log k} + \sqrt{h \gamma} \\
= \frac{\Delta_{2,n}^2}{\sqrt{k} \log k} + \sqrt{h \gamma}
\]

Here we use the fact that \( \int_0^1 (dW_n(u))^2 = 1 \). Clearly, the upper bound converges to zero as \( n \to \infty \) uniformly for all \( s \in [h, 1] \). Hence we get that

\[
\sup_{s \in [h, 1]} \left| I_1(s) - \int_0^s \gamma(u) dW_n(u) \right| \\
\leq \sqrt{\frac{m}{k}} \sup_{s \in [h, 1]} \left| \sum_{t=1}^{s} \gamma(s_t)(W_n(s_t + h) - W_n(s_t - h)) - \int_0^s \gamma(u) dW_n(u) \right| \\
+ \left( \sqrt{\frac{m}{k}} - 1 \right) \sup_{s \in [h, 1]} \left| \int_0^s \gamma(u) dW_n(u) \right|
\]

Since \( \frac{m}{k} \to 1 \) as \( n \to \infty \), we have that the first term in the right hand side converges to zero in probability as \( n \to \infty \). In addition, the second term in the right hand side also converges to zero in probability because \( \sup_{s \in [h, 1]} \left| \int_0^s \gamma(u) dW_n(u) \right| = O_p(1) \) as \( n \to \infty \). Hence we get that, as \( n \to \infty \).

\[
\sup_{s \in [h, 1]} \left| I_1(s) - \int_0^s \gamma(u) dW_n(u) \right| \xrightarrow{P} 0.
\]

The limit relation (A.8) is proved by combining \( I_1(s) \) and \( I_2(s) \). Consequently, by com-
bining (A.6), (A.7) and (A.8), we get that, as \( n \to \infty \),

\[
\sup_{s \in [h,1]} \left| \sqrt{k} \left( \hat{\Gamma}_H(s) - \Gamma(s) \right) - \int_0^s \gamma(s) dW_n(s) \right| \xrightarrow{P} 0.
\]

To complete the proof, we now handle the part for \( s \in [0,h) \). Notice that for \( s < h \) \( \hat{\Gamma}_H(s) = 0 \), \( \Gamma(s) \leq \bar{\gamma}h \), we get that as \( n \to \infty \),

\[
\sup_{s \in [0,h]} \left| \sqrt{k} \left( \hat{\Gamma}_H(s) - \Gamma(s) \right) \right| \leq \sqrt{k} \bar{\gamma}h \to 0.
\]

We only need to check that as \( n \to \infty \), \( \sup_{s \in [0,h]} \left| \int_0^s \gamma(u) dW_n(u) \right| \xrightarrow{P} 0 \). This is achieved by using the Cauchy’s inequality to obtain that for all \( s \in [0, h] \),

\[
\left| \int_0^s \gamma(u) dW_n(u) \right| \leq \sqrt{\int_0^s (\gamma(u))^2 du} \cdot \sqrt{\int_0^s (dW_n(u))^2} \leq \bar{\gamma} \sqrt{h} \cdot \sqrt{h} = \bar{\gamma}h,
\]

where the upper bound tends to zero as \( n \to \infty \).
References


