



**HAL**  
open science

# ANOVEX: ANalysis Of Variability for heavy-tailed EXtremes

Stéphane Girard, Thomas Opitz, Antoine Usseglio-Carleve

► **To cite this version:**

Stéphane Girard, Thomas Opitz, Antoine Usseglio-Carleve. ANOVEX: ANalysis Of Variability for heavy-tailed EXtremes. 2023. hal-04200300v1

**HAL Id: hal-04200300**

**<https://hal.science/hal-04200300v1>**

Preprint submitted on 8 Sep 2023 (v1), last revised 31 May 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ANOVEX: ANalysis Of Variability for heavy-tailed EXtremes

S. Girard<sup>a</sup>, T. Opitz<sup>b</sup> & A. Usseglio-Carleve<sup>c</sup>

<sup>a</sup> Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

<sup>b</sup> Biostatistics and Spatial Processes, INRAE, 84914 Avignon, France

<sup>c</sup> Avignon Université, LMA UPR 2151, 84000 Avignon, France

## Abstract

Analysis of variance (ANOVA) is commonly employed to assess differences in the means of independent samples. However, it is unsuitable for evaluating differences in tail behaviour, especially when means do not exist or empirical estimation of moments is inconsistent due to heavy-tailed distributions. Here, we propose an ANOVA-like decomposition to analyse tail variability, allowing for flexible representation of heavy tails through a set of user-defined extreme quantiles, possibly located outside the range of observations. Building on the assumption of regular variation, we introduce a test for significant tail differences among multiple independent samples and derive its asymptotic distribution. We investigate the theoretical behaviour of the test statistics for the case of two samples, each following a Pareto distribution, and explore strategies for setting hyperparameters in the test procedure. To demonstrate the finite-sample performance, we conduct simulations that highlight generally reliable test behaviour for a wide range of situations. The test is applied to identify clusters of financial stock indices with similar extreme log-returns and to detect temporal changes in daily precipitation extremes at rain gauges in Germany.

**Keywords:** Extreme-value analysis, Analysis of variance, Heavy tails, Extreme quantiles, Hypothesis testing.

**2020 MSC:** 62G32, 62H15, 62E20.

## 1 Introduction

Let us consider data composed of  $J$  samples of independent and identically distributed (i.i.d.) realisations of random variables with cumulative distribution functions  $F_1, \dots, F_J$ , respectively, where  $J > 1$ . Given that  $F_j$  has mean  $\mu_j$  for  $j = 1, \dots, J$ , the classical ANalysis Of VAriance (ANOVA) tests the equality of these  $J$  means through a statistic based on the decomposition of the total variance into intra-class and inter-class variances [Scheffe, 1999]. However, this approach to detect data heterogeneity suffers from some limitations. First, the asymptotic properties of the test statistic are obtained under Gaussian assumptions. Moreover, testing the equality of means always requires the existence of the first moment, and even of the second moment in the ANOVA setting, whereas such conditions are not fulfilled by many distributions. Certain extensions relax these conditions by proposing, for instance, to test equality of medians, as in [Hollander et al., 2013, Chapter 6] where no finite moments

are required, or to use rank-transform procedures [Marden and Muyot, 1995], or to generalize ANOVA by assuming only the existence of low-order moments, as in [Rizzo and Székely, 2010]. However, these extensions focus on statistics describing only the central behaviour of the distribution, and are therefore unable to detect heterogeneity in the distribution tails.

An accurate specification of the tail is crucial to anticipate the consequences of rare events and is often required for risk management [Embrechts et al., 1999]. Distributions can indeed share the same mean or median but can show strongly contrasted tail behaviours. Many biological, environmental or physical phenomena are known to be heavy-tailed and possess power-law behaviour [Clauset et al., 2009]. In finance, the centered log-returns of a financial asset, previously commonly modeled using the Gaussian distribution, are nowadays rather modeled by a Student’s  $t$  distribution to account for their heavy tails [Jondeau et al., 2007]. In this context, a common assumption for risk management is to assume that the survival functions (s.f.) associated with  $F_1, \dots, F_J$  are regularly varying, *i.e.*,

$$\bar{F}_j(x) = 1 - F_j(x) = x^{-1/\xi_j} \mathcal{L}_j^{(j)}(x), \quad \frac{\mathcal{L}_j(tx)}{\mathcal{L}_j(x)} \rightarrow 1, \quad x \rightarrow \infty, \quad t > 0, \quad j = 1, \dots, J, \quad (1)$$

where  $\xi_j > 0$  is called the tail-index and  $\mathcal{L}_j$  is a slowly-varying function. In the extreme-value theory, condition (1) characterises the set of distributions with positive tail-index, also known as the Fréchet maximum domain of attraction, see de Haan and Ferreira [2006, Theorem 1.2.1]. Such distributions only admit moments of order less than  $1/\xi_j$ , making ANOVA infeasible in many situations.

Several alternative test procedures focusing on the tail indices  $\xi_1, \dots, \xi_J$  have been proposed in the literature, and we summarize them without claiming to be exhaustive. Mougeot and Tribouley [2010] developed an equality test of tail indices based on a nonparametric minimax point of view; Worms and Worms [2015] introduced a test for comparing tail indices for heavy-tailed distributions via an empirical likelihood ratio; Einmahl et al., [2016, Section 3] provided a test for different tail indices in the setting of heteroscedastic extremes. However, such tests based on tail indices do not take into account the slowly-varying functions  $\mathcal{L}_j$ , which could be modified by any positive factor without affecting nor the tail-index  $\xi_j$  neither the behaviour of classical tail-index estimators. Therefore, in practice, the tail of a heavy-tailed distribution is usually described by both the tail-index (also called shape parameter) and a scale parameter. In many applications [Davison and Smith, 1990], the Peaks-Over-Threshold (POT) method is used where the excesses above a high threshold are approximated by a Generalized Pareto distribution with both shape and scale parameters. We define extreme quantiles as having a probability level  $\alpha = \alpha_n \rightarrow 1$  as the sample size  $n$  tends to infinity. Two heavy-tailed distributions have asymptotically equivalent extreme quantiles if and only if they share the same shape and scale parameters. Extreme quantiles are therefore interesting tools for testing tail heterogeneity. For instance, Hoga [2017] used extreme quantiles to propose a test for a change point in a series of random variables, and Deuber et al. [2023] developed quantile-based tests for treatment effects with heavy-tailed distributions in a causal framework. However, such quantile-based tests usually consider only a single quantile, which again may be problematic since distributions with very different tail behaviours can have coinciding quantiles at a given probability level.

Inspired by the ANOVA procedure, we here propose to test for the equality of  $L$  extreme quantiles  $q_1(\alpha_{\ell,n}), \dots, q_J(\alpha_{\ell,n})$ ,  $\ell = 1, \dots, L$ , with  $L > 1$ , by decomposing their variability similarly to classical ANOVA. Using extreme quantiles, we focus on the most extreme possible events, for which classical statistical techniques that are directly based on the empirical distribution are not applicable. By jointly studying multiple quantiles, we obtain a precise characterization of tail behaviour and statistical

uncertainty at extreme levels. This approach is simple, does not require the existence of any moment, and is based on the general heavy-tail assumption (1) that encompasses a wide class of distributions commonly used in the literature on risk management, such as the Pareto, Generalized Pareto, Fréchet or Student’s  $t$  distributions.

The remainder of the paper is organized as follows. The new test statistic, called ANOVEX (*ANalysis of Variability in EXtremes*), is introduced in Section 2, based on some extreme quantile estimators. Section 3 provides the asymptotic properties of this test statistic and precisely defines the ANOVEX test procedure. The statistical power of ANOVEX is discussed in Section 4 thanks to approximations of type I and type II errors for three examples involving Pareto distributions, where  $J = 2$ . The simulation study in Section 5 highlights the test performance on various simulated settings. Two real data examples (financial and environmental data) are considered in Section 6. The proofs of the results are postponed to Appendix A, while some additional figures are provided in Appendix B.

## 2 Setting of ANOVEX

We consider  $J > 1$  independent samples  $E_j = \{X_i^{(j)}, i = 1, \dots, n_j\}$  with  $j = 1, \dots, J$  and possibly different sample sizes  $n_j > 1$ . We assume that the random variables in each  $E_j$  are mutually independent and identically distributed according to the cumulative distribution function  $F_j$ . Inspired by ANOVA, we propose to test the equality of the tail of  $X^{(1)}, \dots, X^{(J)}$ . We focus on the heavy-tailed case by assuming that condition (1) holds, such that each cumulative distribution function  $F_j$  is heavy-tailed with positive tail-index  $\xi_j > 0$ . More specifically, this property is denoted by  $F_j \in \mathcal{C}_1(\xi_j)$  with the following definition.

**Definition 1** (Class  $\mathcal{C}_1(\xi)$ ). *The cumulative distribution function  $F$  is said to belong to the class  $\mathcal{C}_1(\xi)$ ,  $\xi > 0$  if it satisfies condition (1), i.e., if  $\bar{F} := 1 - F$  is regularly-varying with index  $-1/\xi$  such that*

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(ty)}{\bar{F}(t)} = y^{-1/\xi} \text{ for all } y > 0.$$

This assumption is widely used in many risk management areas, including financial log-returns [Rachev, 2003], actuarial claim amounts [Mikosch, 1997] or environmental records, such as daily precipitations in Katz et al. [2002] or other natural phenomena in Pisarenko and Rodkin [2010]. Recall that a heavy-tailed distribution with tail-index  $\xi$  does not have moments of order greater than  $1/\xi$ , making the use of the classical ANOVA limited to the “reasonably” heavy-tailed distributions fulfilling  $\xi < 1/2$ . As a counter-example, the claim amounts caused by tornadoes have an estimated tail-index close to 1 in Daouia et al. [2023], and burnt areas of wildfires often have tail-index with values beyond  $1/2$  [Pereira and Turkman, 2019, Koh et al., 2023].

We here propose to compare extreme quantiles, which do not require the existence of any moment of the distribution. For asymptotic results, we consider sample sizes  $n_j$  increasing at the same rate, i.e.,  $n_j/n \rightarrow \lambda_j > 0$  as  $n \rightarrow \infty$ , for  $j = 1, \dots, J$ . We denote by  $q_j(\alpha) := \inf \{x \in \mathbb{R} : \alpha \leq F_j(x)\}$  the quantile of distribution  $F_j$  at probability level  $\alpha \in (0, 1)$ . The null hypothesis to be tested is

$$(H_0) \quad \text{For all } j \neq j' \in \{1, \dots, J\}^2, \quad q_{j'}(\alpha)/q_j(\alpha) \rightarrow 1 \text{ as } \alpha \rightarrow 1.$$

The tail behaviour of the  $J$  samples is summarized through  $J \times L$  estimators  $\tilde{q}_j(\alpha_{\ell,n})$  of extreme quantiles  $q_j(\alpha_{\ell,n})$ , where  $\alpha_{\ell,n} \rightarrow 1$  such that  $n(1 - \alpha_{\ell,n}) \rightarrow \tau_\ell \geq 0$  as  $n \rightarrow \infty$ ,  $\ell = 1, \dots, L$  and

$j = 1, \dots, J$ . The above conditions imply

$$\mathbb{P} \left( \max_{i=1, \dots, n_j} X_i^{(j)} \leq q_j(\alpha_{\ell, n}) \right) = \alpha_{\ell, n}^{n_j} \rightarrow \exp(-\lambda_j \tau_\ell),$$

as  $n \rightarrow \infty$ , and therefore such quantiles can lie beyond the observation range when  $\tau_\ell$  is small. Direct empirical estimation of such quantiles leads to a high estimation variance and/or bias, so that semi-parametric estimation is preferable. A classical approach is to use the Weissman estimator [Weissman, 1978], defined as follows:

$$\widehat{q}_j^{\text{W}}(\alpha_{\ell, n} | \beta_{j, n}) = \widehat{q}_j(\beta_{j, n}) \left( \frac{1 - \beta_{j, n}}{1 - \alpha_{\ell, n}} \right)^{\widehat{\xi}_j(\beta_{j, n})}, \quad (2)$$

where

- $\beta_{j, n}$  are intermediate probability levels, *i.e.*,  $\beta_{j, n} \rightarrow 1$  and  $n(1 - \beta_{j, n}) \rightarrow \infty$  as  $n \rightarrow \infty$ , for  $j = 1, \dots, J$ ;
- $\widehat{q}_j(\beta_{j, n}) = X_{[\beta_{j, n} n_j]_{n_j}}^{(j)}$  is the empirical quantile of level  $\beta_{j, n}$  in  $E_j$ ,  $j = 1, \dots, J$ , and
- $\widehat{\xi}_j(\beta_{j, n})$  is an estimator of  $\xi_j$ , such as the Hill estimator [Hill, 1975] based on the  $k_{j, n} := \lfloor (1 - \beta_{j, n}) n_j \rfloor$  largest observations from sample  $E_j$ ,  $j = 1, \dots, J$ :

$$\widehat{\xi}_j^{\text{H}}(\beta_{j, n}) = \frac{1}{k_{j, n}} \sum_{i=0}^{k_{j, n}-1} \log X_{n_j-i, n_j} - \log X_{n_j-k_{j, n}, n_j}. \quad (3)$$

Inspired by classical ANOVA, we aim at additively decomposing the variance of quantiles into contributions from the factor (*i.e.*, the sample in our setting) and from the different quantile levels. Moreover, we work in a log-scale to achieve statistically stable behaviour of differences between log-quantiles and their averages in the heavy-tailed setting. We write the overall sum-of-squares as

$$\Delta_n = \frac{1}{JL} \sum_{j=1}^J \sum_{\ell=1}^L (\log \widetilde{q}_j(\alpha_{\ell, n}) - \mu_{\alpha, n})^2, \quad \text{with} \quad \mu_{\alpha, n} = \frac{1}{JL} \sum_{j=1}^J \sum_{\ell=1}^L \log \widetilde{q}_j(\alpha_{\ell, n}).$$

With ANOVEX, we propose to decompose this extreme log-quantile variance into two parts: a variance  $\Delta_{1, n}$  due to the  $J$  different samples, and a variance  $\Delta_{2, n}$  due to the  $L$  different quantile levels. This decomposition is formally proven in the following proposition and is further illustrated in Figure 1. We construct the ANOVEX test statistic to reject ( $H_0$ ) if  $\Delta_{1, n}/\Delta_{2, n}$  exceeds an appropriately fixed threshold, *i.e.*, if the contribution of the inter-sample variance to the total variance is very strong. Therefore, the test statistic is straightforward to interpret and to compute. The definition of the threshold is investigated in the next section.

**Proposition 2.1.** *The decomposition  $\Delta_n = \Delta_{1, n} + \Delta_{2, n}$  holds where*

$$\begin{aligned} \Delta_{1, n} &= \frac{1}{L} \sum_{\ell=1}^L \Delta_{1, \ell, n} \quad (\text{variance due to the different samples}), \\ \Delta_{2, n} &= \frac{1}{L} \sum_{\ell=1}^L \left( \mu_{\alpha, n}^{(\ell)} - \mu_{\alpha, n} \right)^2 \quad (\text{variance due to the different quantile levels}), \end{aligned}$$

with the components indexed by  $\ell = 1, \dots, L$  given as

$$\mu_{\alpha,n}^{(\ell)} = \frac{1}{J} \sum_{j=1}^J \log \tilde{q}_j(\alpha_{\ell,n}) \quad (\text{mean estimation at level } \ell),$$

$$\Delta_{1,\ell,n} = \frac{1}{J} \sum_{j=1}^J \left( \log \tilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right)^2 \quad (\text{variance due to the different samples at level } \ell).$$

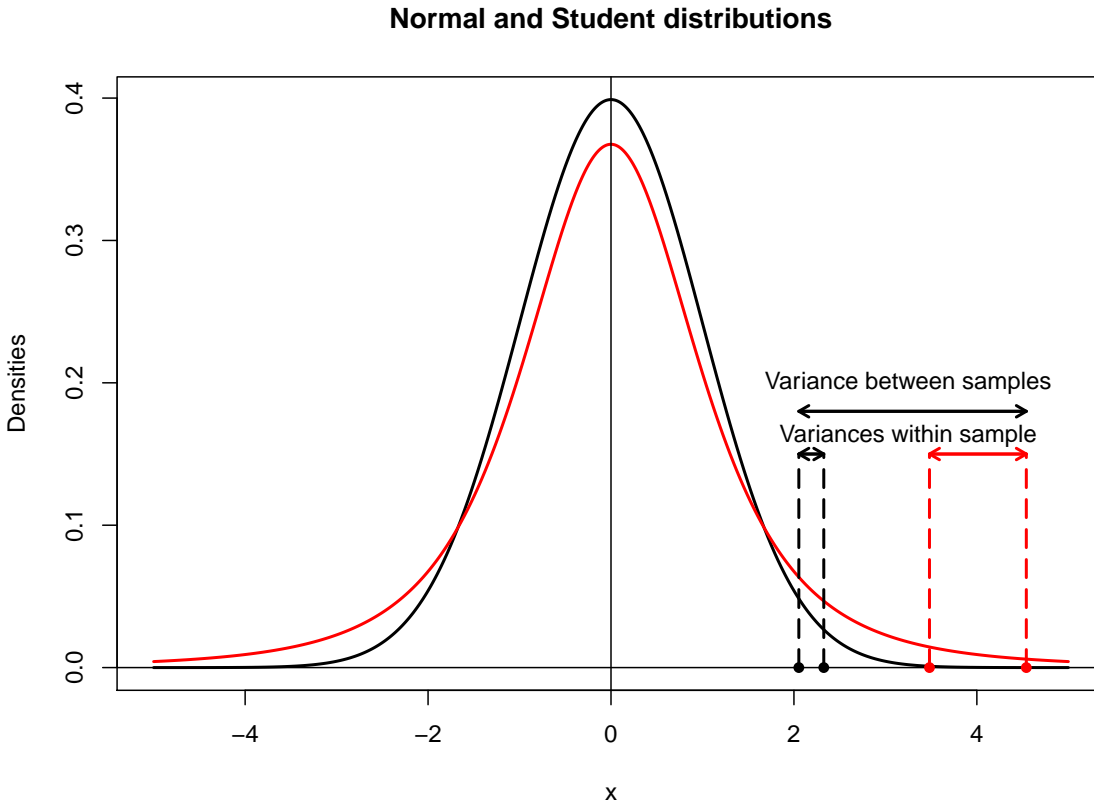


Figure 1: Illustration of the ANOVEX-decomposition of extreme log-quantile variance with  $J = 2$  for a standard normal (black curve) and a Student's  $t$  distribution with 3 degrees of freedom. The number of extreme quantiles is  $L = 2$  with probability levels chosen at 0.98 and 0.99.

### 3 Asymptotic distribution of the test statistic

To investigate the asymptotic properties of the ratio  $\Delta_{1,n}/\Delta_{2,n}$ , we need the following notations defined for any positive  $u_1, \dots, u_L$ :

$$\text{smlog}(u_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log(u_\ell)^2, \quad \text{and} \quad \text{varlog}(u_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log(u_\ell)^2 - \left( \frac{1}{L} \sum_{\ell=1}^L \log(u_\ell) \right)^2, \quad (4)$$

which can be viewed respectively as the empirical second moment and variance of  $(\log(u_1), \dots, \log(u_L))$ . Our first main result establishes the distribution of the test statistic under  $(H_0)$ , which requires an appropriate deterministic rescaling factor.

**Theorem 3.1.** *Assume  $F_j \in \mathcal{C}_1(\xi_j)$  for  $j = 1, \dots, J$  and the following conditions:*

- *Extreme probability levels: For all  $n \geq 1$  and  $\ell = 1, \dots, L$ , the probabilities  $\alpha_{\ell,n} \in (0, 1)$  satisfy  $\alpha_{\ell,n} \rightarrow 1$  as  $n \rightarrow \infty$ .*
- *Extreme quantile estimator:  $\tilde{q}_j(\alpha_{\ell,n})$  is estimator of the extreme quantile  $q_j(\alpha_{\ell,n})$ , computed on the  $n_j$ -sample  $E_j$  where  $n_j/n \rightarrow \lambda_j > 0$  as  $n \rightarrow \infty$ , and satisfies*

$$\bar{\sigma}_{\ell,n}^{-1} (\log(\tilde{q}_j(\alpha_{\ell,n})) - \log(q_j(\alpha_{\ell,n}))) \xrightarrow{d} \mathcal{N}(0, \xi_j^2), \quad (5)$$

for deterministic normalizing sequences  $\bar{\sigma}_{\ell,n} \rightarrow 0$  as  $n \rightarrow \infty$ , and for all  $j = 1, \dots, J$  and  $\ell = 1, \dots, L$ .

Then, under  $(H_0)$ ,

$$\frac{J \operatorname{varlog}((1 - \alpha_n)_{1:L})}{\bar{\sigma}_n^2} \frac{\Delta_{1,n}}{\Delta_{2,n}} \xrightarrow{d} \chi_{J-1}^2, \quad (6)$$

with  $\chi_{J-1}^2$  a chi-square random variable with  $J - 1$  degrees of freedom and notation

$$\bar{\sigma}_n^2 = \frac{1}{L} \sum_{\ell=1}^L \bar{\sigma}_{\ell,n}^2.$$

If, moreover,  $\alpha_{\ell,n} = 1 - \tau_\ell/n$  for  $\ell = 1, \dots, L$ , then (6) can be simplified as

$$\frac{J \operatorname{varlog}(\tau_{1:L})}{\bar{\sigma}_n^2} \frac{\Delta_{1,n}}{\Delta_{2,n}} \xrightarrow{d} \chi_{J-1}^2.$$

**Remark 1** (Use of other risk measures). *The result of Theorem 3.1 remains valid if the quantile is replaced by any other risk measure whose extreme estimator fulfills condition (5). Indeed, the Conditional Tail Expectation (also called Expected Shortfall, introduced in Artzner et al. [1999]), or the expectile [Newey and Powell, 1987], could be used, since they are asymptotically proportional to the extreme quantile [El-Methni et al., 2014, Daouia et al., 2018]. Here, we keep the focus on quantiles since they do not require a finite first moment.*

We now study an extreme quantile estimator satisfying the assumptions of Theorem 3.1. Condition (5) is fulfilled by combining the Weissman estimator (2) with the Hill estimator (3). However, a stronger second-order assumption, widely used in the extreme-value literature, is necessary. We denote it by  $\mathcal{C}_2(\xi, \rho, A)$ .

**Definition 2** (Class  $\mathcal{C}_2(\xi, \rho, A)$ ). *The distribution  $F$  belongs to the class  $\mathcal{C}_2(\xi, \rho, A)$  with tail-index  $\xi > 0$  and second-order parameter  $\rho < 0$ , if there exists a measurable auxiliary function  $A$  with constant sign, satisfying  $A(t) \rightarrow 0$  as  $t \rightarrow \infty$ , such that*

$$\lim_{t \rightarrow \infty} \frac{1}{A(1/\bar{F}(t))} \left( \frac{\bar{F}(ty)}{\bar{F}(t)} - y^{-1/\xi} \right) = y^{-1/\xi} \frac{y^{\rho/\xi} - 1}{\xi \rho}, \quad \text{for all } y > 0.$$

Beirlant et al. [2004] lists numerous distributions satisfying this assumption. The function  $A(\cdot)$  is regularly-varying with index  $\rho$ . Therefore, when  $\rho$  is small, say  $\rho \leq -1$  (including the Pareto, Burr, Fréchet distributions), the distribution tail is very close to a Pareto tail, and the Weissman estimator is very accurate. On the contrary, if  $\rho$  is close to 0, as for the Generalized Pareto distribution ( $\rho = -\xi$ ) with small positive tail-index, the Weissman approximation remains valid, but becomes very accurate only relatively far in the tail. Since the Weissman estimator is asymptotically Gaussian under the second-order condition  $\mathcal{C}_2(\xi, \rho, A)$ , condition (5) of Theorem 3.1 holds (see the proof in Appendix), and we get the following result.

**Corollary 3.1.** *Assume  $F_j \in \mathcal{C}_2(\xi_j, \rho_j, A_j)$  for  $j = 1, \dots, J$  according to Definition 2. Moreover, we assume the following conditions hold:*

- *Extreme probability levels: For all  $n \geq 1$  and  $\ell = 1, \dots, L$ ,  $\alpha_{\ell, n} \in (0, 1)$  satisfies  $\alpha_{\ell, n} \rightarrow 1$  as  $n \rightarrow \infty$ .*
- *Extreme quantile estimator:  $\widehat{q}_j^W(\alpha_{\ell, n} | \beta_{j, n})$  is the Weissman estimator (2) of the extreme quantile  $q_j(\alpha_{\ell, n})$  computed on the  $n_j$ -sample  $E_j$  such that, as  $n \rightarrow \infty$  and for all  $j = 1, \dots, J$  and  $\ell = 1, \dots, L$ ,*

$$\begin{aligned} n_j/n &\rightarrow \lambda_j > 0, \\ (1 - \beta_{j, n})/(1 - \beta_n) &\rightarrow 1/\lambda_j, \\ (1 - \alpha_{\ell, n})/(1 - \beta_n) &\rightarrow 0, \\ \sqrt{n(1 - \beta_n)}/\log((1 - \beta_n)/(1 - \alpha_{\ell, n})) &\rightarrow \infty. \end{aligned}$$

- *Second-order behaviour: the auxiliary function  $A$  satisfies*

$$\sqrt{n(1 - \beta_n)}A_j((1 - \beta_n)^{-1}) \rightarrow 0, \quad n \rightarrow \infty, \quad \text{for all } j = 1, \dots, J. \quad (7)$$

Then, under  $(H_0)$ ,

$$\frac{J \operatorname{varlog}((1 - \alpha_n)_{1:L}) n(1 - \beta_n)}{\frac{1}{L} \sum_{\ell=1}^L \log\left(\frac{1 - \beta_n}{1 - \alpha_{n, \ell}}\right)^2} \frac{\Delta_{1, n}}{\Delta_{2, n}} \xrightarrow{d} \chi_{J-1}^2, \quad n \rightarrow \infty. \quad (8)$$

If, moreover,  $\alpha_{\ell, n} = 1 - \tau_\ell/n$  for  $\ell = 1, \dots, L$ , then (8) can be simplified as

$$\frac{J \operatorname{varlog}(\tau_{1:L}) n(1 - \beta_n)}{S_n(\beta_n, \tau_{1:L})} \frac{\Delta_{1, n}}{\Delta_{2, n}} \xrightarrow{d} \chi_{J-1}^2, \quad n \rightarrow \infty, \quad (9)$$

where

$$S_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log\left(\frac{n(1 - \beta_n)}{\tau_\ell}\right)^2. \quad (10)$$

We shortly discuss the technical conditions involved in this result. First, the sample sizes  $n_1, \dots, n_J$  have to be asymptotically proportional. In practice, they should therefore not be too unbalanced, and the reference value  $n$  could be chosen as the average of  $n_j$  if they are not equal. Moreover, the intermediate levels  $\beta_{j, n}$  have to be chosen accordingly, such that the number of order statistics  $k_{j, n} = n_j(1 - \beta_{j, n})$  remains the same in all samples. Second, the condition (7) on the second-order behaviour



ensures that the Hill estimator is unbiased across all samples. Results for the more general case  $\sqrt{n(1-\beta_n)}A_j((1-\beta_n)^{-1}) \rightarrow c_j \in \mathbb{R}$  could also be derived but involve more complicated formulas. Moreover, bias-reduced versions of the Hill estimator with the same asymptotic variance were proposed in the literature [Caeiro et al., 2005], and this condition may thus be dropped or replaced by a rather weak condition on the auxiliary functions  $A_j$ .

The proposed ANOVEX test rejects  $(H_0)$  with asymptotic level  $\gamma \in (0, 1)$  when

$$T_n := \frac{J \operatorname{varlog}(\tau_{1:L}) n(1-\beta_n)}{S_n(\beta_n, \tau_{1:L})} \frac{\Delta_{1,n}}{\Delta_{2,n}} > \chi_{J-1, 1-\gamma}^2, \quad (11)$$

where  $\chi_{J-1, 1-\gamma}^2$  denotes the quantile of level  $1-\gamma$  of the chi-square distribution with  $J-1$  degrees of freedom. The test is asymptotically equivalent to

$$\frac{J \operatorname{varlog}(\tau_{1:L}) n(1-\beta_n)}{\log(n(1-\beta_n))^2} \frac{\Delta_{1,n}}{\Delta_{2,n}} > \chi_{J-1, 1-\gamma}^2.$$

For accuracy reasons, we adopt the first version (11) of  $T_n$  in the sequel.

To sum up, we have constructed a statistical test called ANOVEX to detect whether a set of tail quantiles is different across different distributions. Unlike classical ANOVA, ANOVEX does not require a finite variance or even mean and can be applied to a much wider class of heavy-tailed distributions.

Next, we analytically investigate the finite sample behaviour of the test on some simple examples to study the power of this test. A numerical investigation for more general settings is carried out in the simulation study in Section 5.

## 4 Examples of type-I and type-II error approximations

An attractive feature of ANOVEX is its simplicity: it only requires an extreme quantile estimator, such as the Weissman estimator. For the Pareto distribution, the prototype of regularly-varying distributions, this simple structure allows us to leverage Rényi's representation [Rényi, 1953] to provide very accurate approximations of the distribution of the test statistic  $T_n$ , as well as type I and type II errors, in the case of  $J=2$  independent samples.

We write  $\mathcal{P}(1/\xi)$  for the basic Pareto distribution having scale parameter 1, and therefore lower endpoint 1 of its support, and shape parameter  $1/\xi > 0$ , where  $\xi$  is the tail-index; its cumulative distribution function is  $1-x^{-1/\xi}$ ,  $x > 1$ .

Three examples are considered with  $J=2$  independent samples coming from either two identical Pareto distributions (Section 4.1), two Pareto distributions with different scale parameters (Section 4.2), or two Pareto distributions with different shape parameters (Section 4.3). We treat the case where  $\alpha_{\ell,n} = 1 - \tau_{\ell}/n$  for  $\ell = 1, \dots, L$ .

In addition to the normalizing constant  $S_n$  in (10), we need the following constants:

$$s_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \sqrt{1 + \log\left(\frac{n(1-\beta_n)}{\tau_{\ell}}\right)^2},$$

$$\text{and } \mathfrak{s}_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log\left(\frac{n}{\tau_{\ell}}\right) \sqrt{1 + \log\left(\frac{n(1-\beta_n)}{\tau_{\ell}}\right)^2}.$$

## 4.1 Identically distributed Pareto samples

Assume that both samples follow the same Pareto distribution. The ANOVEX test is supposed to wrongly reject the null hypothesis with asymptotic probability  $\gamma$  (type I error). We investigate the impact of the test parameters on this probability through an approximation of the rejection probability provided in the following proposition.

**Proposition 4.1.** *Consider two independent samples  $E_j = \{X_1^{(j)}, \dots, X_n^{(j)}\}$ ,  $j = 1, 2$ , of i.i.d. variables following the same Pareto distribution  $\mathcal{P}(1/\xi)$ ,  $\xi > 0$ . Assume the following:*

- (a) *Probability levels:  $(\beta_n)$  is an intermediate probability level such that  $(1 - \beta_n) \log(n) \rightarrow 0$  as  $n \rightarrow \infty$ .*
- (b) *Extreme quantile estimator:  $\hat{q}_j^W(\alpha_{\ell,n} | \beta_n)$  is the Weissman estimator (2) of the extreme quantile  $q_j(\alpha_{\ell,n})$ , based on the Hill estimator  $\hat{\xi}_j^H(\beta_n)$  defined in (3), where  $\alpha_{\ell,n} = 1 - \tau_\ell/n$ , for  $\ell = 1, \dots, L$  and  $j = 1, 2$ .*

Then, the probability  $p_n(\gamma) := \mathbb{P}_{H_0}(T_n > \chi_{1,1-\gamma}^2)$  to wrongly reject  $(H_0)$  with asymptotic level  $\gamma \in (0, 1)$  is given by

$$p_n(\gamma) = 2\bar{\Phi} \left( \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right) (1 + o(1)), \quad (12)$$

where  $\bar{\Phi}(\cdot)$  is the standard Gaussian survival function.

According to (12), the behaviour of  $p_n(\gamma)$  is driven by  $S_n(\beta_n, \tau_{1:L})$  but does not depend on the tail-index  $\xi$ . Clearly,  $p_n(\gamma) \xrightarrow{\geq} \gamma$  as  $n \rightarrow \infty$  in view of  $S_n(\beta_n, \tau_{1:L}) \sim (\log(n(1 - \beta_n)))^2$  as  $n \rightarrow \infty$ . Indeed, a first-order expansion yields

$$p_n(\gamma) = \gamma + \frac{\bar{\Phi}^{-1}(\gamma/2)\varphi(\bar{\Phi}^{-1}(\gamma/2))}{(\log(n(1 - \beta_n)))^2} (1 + o(1)),$$

where  $\varphi(\cdot)$  is the density of the standard Gaussian distribution. Therefore, we obtain a logarithmic rate of convergence with respect to  $n(1 - \beta_n)$ , the expected number of observations above  $q_1(\beta_n) = q_2(\beta_n)$ . An illustration of the accuracy of approximation (12) for Pareto distributions  $\mathcal{P}(1/\xi = 4)$  and  $\mathcal{P}(1/\xi = 1)$  is provided in the top panel of Figure 2, for  $n = 1,000$  and  $L = 2, \dots, 30$ . The first-order approximation (12) (solid blue curve) is compared to the empirical estimation of  $p_n(\gamma = 0.05)$  based on  $N = 10,000$  replications (dashed blue curve); see Section 5 for details and further examples. The approximation is fairly precise, especially for moderate values of  $L$ , and there arises only a small bias of observed type-I errors with respect to the nominal level  $\gamma$ . As expected,  $p_n(\gamma) \geq \gamma$  (horizontal black line for  $\gamma$ ). The true  $p_n(\gamma)$  does not seem to depend on  $\xi$  in these numerical experiments, in line with the analytical approximation (12). Finally, under the weak condition  $\log(n(1 - \beta_n)) \geq 3$ , the quantity  $S_n(\beta_n, \tau_{1:L})$  is a decreasing function of  $L$ , which entails that the approximation (12) is increasing with  $L$  in the case where  $\tau_\ell = \ell$ . The practical choice of  $L$  is discussed in Section 5.

## 4.2 Pareto samples with different scale parameters

We consider the case of  $J = 2$  Pareto samples of size  $n > 0$ , with same shape parameter  $1/\xi$ ,  $\xi > 0$ , and respective scale parameters 1 and  $\lambda_n > 0$ , where  $\lambda_n = 1$  yields the setting of Section 4.1. Here, we focus on the case where  $\lambda_n \xrightarrow{\neq} 1$  as  $n \rightarrow \infty$ , and we investigate the power of the ANOVEX test to reject  $(H_0)$  in this situation. The following result provides an approximation of the distribution of the test statistic.

**Proposition 4.2.** *Suppose that the assumptions (a, b) of Proposition 4.1 hold. Consider two independent samples  $E_1 = \{X_1^{(1)}, \dots, X_n^{(1)}\}$  and  $E_2 = \{X_1^{(2)}, \dots, X_n^{(2)}\}$ , where, for  $i = 1, \dots, n$ , the  $X_i^{(1)}$  s are i.i.d. from a Pareto distribution  $\mathcal{P}(1/\xi)$ ,  $\xi > 0$  and the  $X_i^{(2)}$  s are i.i.d. with  $X_i^{(2)} \stackrel{d}{=} \lambda_n X_i^{(1)}$ ,  $\lambda_n > 0$ ,  $\lambda_n \rightarrow 1$  as  $n \rightarrow \infty$  such that*

- (i)  $\frac{\log(n(1-\beta_n))}{(n(1-\beta_n))^{3/4}} \vee \sqrt{\frac{\log(n(1-\beta_n))}{n}} = o(\log(\lambda_n))$ ,
- (ii)  $\log(\lambda_n) = o(1/\sqrt{n(1-\beta_n)})$ .

Then, as  $n \rightarrow \infty$ ,

$$T_n \stackrel{d}{=} \left( \frac{\log(\lambda_n)^2 n(1-\beta_n)}{2\xi^2 S_n(\beta_n, \tau_{1:L})} - \frac{\sqrt{2n(1-\beta_n)} \log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\xi S_n(\beta_n, \tau_{1:L})} \Gamma + \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma^2 \right) \times \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(n(1-\beta_n))} \right) \right),$$

where  $\Gamma$  is a standard normal random variable.

This result yields an approximation of type II error associated with the ANOVEX test when the alternative hypothesis is formulated as

$$(H_{1,n}) \quad X_i^{(1)} \sim \mathcal{P}(1/\xi) \text{ and } X_i^{(2)} \stackrel{d}{=} \lambda_n X_i^{(1)}, \quad i = 1, \dots, n \text{ and } \lambda_n \rightarrow 1 \text{ as } n \rightarrow \infty,$$

as illustrated in the next remark.

**Remark 2.** *If assumption (i) of Proposition 4.2 is replaced by the slightly stronger condition*

$$(i') \quad \log(n(1-\beta_n))^2 / (n(1-\beta_n))^{3/4} \vee \sqrt{\frac{\log(n(1-\beta_n))^3}{n}} = o(\log(\lambda_n)),$$

then, the probability  $\mathbb{P}_{H_{1,n}}(T_n \leq \chi_{1,1-\gamma}^2)$  to accept  $(H_0)$  with asymptotic level  $\gamma \in (0, 1)$  may be approximated, for  $n$  large enough, by

$$\bar{\Phi}(\Omega_{1,n} - \sqrt{\Omega_{2,n}}) - \bar{\Phi}(\Omega_{1,n} + \sqrt{\Omega_{2,n}}) \quad (13)$$

where

$$\Omega_{1,n} = \frac{\log(\lambda_n) \sqrt{n(1-\beta_n)} s_n(\beta_n, \tau_{1:L})}{\sqrt{2\xi} (1 + S_n(\beta_n, \tau_{1:L}))},$$

$$\Omega_{2,n} = \frac{\log(\lambda_n)^2 n(1-\beta_n)}{2\xi^2} \frac{s_n(\beta_n, \tau_{1:L})^2 - 1 - S_n(\beta_n, \tau_{1:L})}{(1 + S_n(\beta_n, \tau_{1:L}))^2} + \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} \chi_{1,1-\gamma}^2 > 0.$$

First, remark that  $\Omega_{1,n} \rightarrow 0$  and  $\Omega_{2,n} \rightarrow \chi_{1,1-\gamma}^2$  as  $n \rightarrow \infty$ ; see the proof of Remark 2 in the Appendix for details. Thus,  $\mathbb{P}_{H_{1,n}}(T_n \leq \chi_{1,1-\gamma}^2) \rightarrow 1 - \gamma$  as  $n \rightarrow \infty$ , meaning that the asymptotic type II error is  $1 - \gamma$  when  $(H_{1,n})$  approaches  $(H_0)$ . Second, considering  $\lambda_n = 1$  yields

$$\begin{aligned} \bar{\Phi}(\Omega_{1,n} - \sqrt{\Omega_{2,n}}) - \bar{\Phi}(\Omega_{1,n} + \sqrt{\Omega_{2,n}}) &= 1 - 2\bar{\Phi} \left( \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right) \\ &= 1 - p_n(\gamma)(1 + o(1)), \end{aligned}$$

see (12), which is in accordance with the type I error provided by Proposition 4.1. Third, in the converse case where  $\lambda_n$  is large (*i.e.* when the two Pareto distributions are very different), then the approximation (13) tends to zero, and the ANOVEX test is likely to reject ( $H_0$ ). The same reasoning may be applied to  $\lambda_n^{1/\xi}$  which is the key quantity in  $\Omega_{1,n}$  and  $\Omega_{2,n}$ . For a fixed value of  $\lambda_n$ , the approximated probability is a decreasing function of  $\xi$ : Heavy tails are thus more easily discriminated. Adopting the classical choice  $\beta_n = 1 - c/\sqrt{n}$ ,  $c > 0$ , conditions of Proposition 4.2 and Remark 2 imply that  $\lambda_n$  converges to 1 not faster than  $n^{-3/8}$  (up to a logarithmic factor). This may be interpreted as the minimum gap between ( $H_0$ ) and ( $H_{1,n}$ ) that the ANOVEX test is able to discriminate.

Finally, an illustration of the accuracy of the approximation given in Remark 2 is provided on the second row of Figure 2 in Section 5 for  $\lambda_n = 1.2$  and different values of  $\xi \in \{0.15, 0.25, 0.35, 0.50\}$  (left panel) and for  $\xi = 0.25$  and  $\lambda_n \in \{1.1, 1.2, 1.3, 1.4\}$  (right panel). It appears that the approximation (13) is, in all considered cases, remarkably accurate. The numerical results also confirm the roles of  $\lambda_n$  and  $\xi$  in the type II error.

### 4.3 Pareto samples with different shape parameters

Finally, let us deal with the case of  $J = 2$  Pareto samples of size  $n > 0$ , with respective shape parameters  $1/\xi$  and  $1/(\xi\theta_n)$  where  $\xi, \theta_n > 0$ . Remark that  $\theta_n = 1$  yields back the setting of Paragraph 4.1. Similarly to Paragraph 4.2, we investigate the ability of the ANOVEX test to reject ( $H_0$ ) when  $\theta_n \xrightarrow{\neq} 1$  as  $n \rightarrow \infty$ . The next result provides an approximation of the test statistic distribution in this context.

**Proposition 4.3.** *Suppose that the assumptions (a,b) of Proposition 4.1 hold with the additional condition that  $\log(n) = O(\log(n(1 - \beta_n)))$ . Consider two independent samples  $E_1 = \{X_1^{(1)}, \dots, X_n^{(1)}\}$  and  $E_2 = \{X_1^{(2)}, \dots, X_n^{(2)}\}$ , where, for  $i = 1, \dots, n$ , the  $X_i^{(1)}$ s are i.i.d. from a Pareto distribution  $\mathcal{P}(1/\xi)$ ,  $\xi > 0$  and the  $X_i^{(2)}$ s are i.i.d. with  $X_i^{(2)} \stackrel{d}{=} (X_i^{(1)})^{\theta_n}$ ,  $\theta_n > 0$  and  $\theta_n \rightarrow 1$  as  $n \rightarrow \infty$  such that*

$$(i) \frac{\log(n(1-\beta_n))}{(n(1-\beta_n))^{3/4}} \vee \sqrt{\frac{\log(n(1-\beta_n))}{n}} = o(\log(\theta_n)),$$

$$(ii) \log(\theta_n) = o(1/\sqrt{n(1-\beta_n)}).$$

Then, as  $n \rightarrow \infty$ ,

$$T_n \stackrel{d}{=} 2 \left( \frac{n(1-\beta_n)(1-\theta_n)^2}{(1+\theta_n)^2} \frac{\text{smlog}(n/\tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} + 2 \frac{\sqrt{n(1-\beta_n)}\sqrt{1+\theta_n^2}(1-\theta_n)}{(1+\theta_n)^2} \frac{\mathfrak{s}_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma + \frac{(1+\theta_n^2)}{(1+\theta_n)^2} \frac{1+S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma^2 \right) \times \left( 1 + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(n(1-\beta_n))} \right) + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) \right),$$

where  $\Gamma$  is a standard normal random variable, and  $\text{smlog}(\cdot)$  is defined in Equation (4).

Note that the conditions on the sequence  $(\theta_n)$  are the same ones as those assumed for the sequence  $(\lambda_n)$  in the previous example. The same choices can thus be made, with similar interpretations. This result can be used to approximate the type II error associated with the ANOVEX test when the alternative hypothesis is

$$(H'_{1,n}) \quad X_i^{(1)} \sim \mathcal{P}(1/\xi) \text{ and } X_i^{(2)} \stackrel{d}{=} (X_i^{(1)})^{\theta_n}, \quad i = 1, \dots, n \text{ and } \theta_n \rightarrow 1 \text{ as } n \rightarrow \infty,$$

as illustrated in the next remark.

**Remark 3.** If assumption (i) of Proposition 4.3 is replaced by the slightly stronger condition

$$(i') \log(n(1 - \beta_n))^2 / (n(1 - \beta_n))^{3/4} \vee \sqrt{\frac{\log(n(1 - \beta_n))^3}{n}} = o(\log(\theta_n)),$$

then, the probability  $\mathbb{P}_{H'_{1,n}}(T_n \leq \chi_{1,1-\gamma}^2)$  to accept  $(H_0)$  with asymptotic level  $\gamma \in (0, 1)$  may be approximated by for  $n$  large enough by

$$\bar{\Phi}\left(\Psi_{1,n} - \sqrt{\Psi_{2,n}}\right) - \bar{\Phi}\left(\Psi_{1,n} + \sqrt{\Psi_{2,n}}\right) \quad (14)$$

where

$$\begin{aligned} \Psi_{1,n} &= \frac{\sqrt{n(1 - \beta_n)}(\theta_n - 1)\mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{1 + \theta_n^2}(1 + S_n(\beta_n, \tau_{1:L}))}, \\ \Psi_{2,n} &= \frac{(\theta_n - 1)^2 n(1 - \beta_n)\mathfrak{s}_n(\beta_n, \tau_{1:L})^2 - (1 + S_n(\beta_n, \tau_{1:L}))\text{smlog}(n/\tau_{1:L})}{(1 + \theta_n^2)(1 + S_n(\beta_n, \tau_{1:L}))^2} \\ &\quad + \frac{(1 + \theta_n)^2}{(1 + \theta_n^2)} \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} \frac{\chi_{1,1-\gamma}^2}{2} > 0. \end{aligned}$$

Here also, one can remark that  $\mathbb{P}_{H'_{1,n}}(T_n \leq \chi_{1,1-\gamma}^2)$  tends to  $1 - \gamma$  as  $n \rightarrow \infty$ . Moreover, taking  $\theta_n = 1$  leads to the formula of Proposition 4.1. However, unlike the approximation of Proposition 4.2, this probability is not related to the tail-index  $\xi$ . The accuracy of this approximation for several values of  $\theta_n$  is illustrated on the bottom left panel of Figure 2 in Section 5 for  $\xi = 0.25$  and  $\theta_n \in \{1.1, 1.2, 1.3, 1.4\}$ . Again, the curves associated with the approximated probability (14) are nearly superimposed to the curves associated with the empirical type II errors.

## 5 Simulation study

In this section, our methodology is applied on simulated examples. For each example, we simulate  $N = 10,000$  times  $J$  samples of size  $n = 1,000$ , and compute the test statistic (11) to reject, or not reject  $(H_0)$  with the confidence level  $\gamma = 0.05$ . For convenience, we take  $\alpha_{\ell,n} = 1 - \ell/n$  ( $\tau_\ell = \ell$ ) for  $\ell = 1, \dots, L$ , and consider several values of  $L$  (ranging from 2 to 30). We finally report the empirical rejection (or equivalently non-rejection) probabilities obtained through the  $N$  replications. Note that throughout the section, several heavy-tailed distributions are considered (all with tail-index  $\xi > 0$ ):

- The Pareto distribution with s.f.  $\bar{F}(x) = x^{-1/\xi}$  for  $x > 1$ ;
- The Generalized Pareto distribution (GPD) with s.f.  $\bar{F}(x) = (1 + \xi x)^{-1/\xi}$  for  $x > 0$ . Note that this distribution fulfills Assumption  $\mathcal{C}_2(\xi, \rho, A)$  with  $\rho = -\xi$ ;
- The Fréchet distribution with s.f.  $\bar{F}(x) = 1 - \exp(-x^{-1/\xi})$  for  $x > 0$ . Assumption  $\mathcal{C}_2(\xi, \rho, A)$  is also satisfied with  $\rho = -1$ ;
- The Burr( $\rho$ ) distribution with s.f.  $\bar{F}(x) = (1 + x^{-\rho/\xi})^{1/\rho}$  for  $x > 0$ . In this case,  $\mathcal{C}_2(\xi, \rho, A)$  is satisfied for any  $\rho < 0$ .

Firstly, we focus on the most simple case with  $J = 2$  samples of random variables  $X_1$  and  $X_2$ . In line with Sections 4.1, 4.2 and 4.3, we propose the following toy examples:

- ( $\mathcal{P}$ ) As in Proposition 4.1,  $X_1 \stackrel{d}{=} X_2 = \mathcal{P}(1/\xi)$ , *i.e.* both samples are i.i.d. replications of a Pareto distribution. In this setting, we are supposed to reject ( $H_0$ ) with probability  $\gamma = 0.05$ . Figure 2 gives an overview of the rejection probabilities as functions of  $L$ . We also consider the case of two identical Fréchet, Burr and GP distributions, and take  $\xi = 0.25$  (top left panel) and  $\xi = 1$  (top middle panel) in all cases. Note that, when  $\xi = 1$ , Burr and GPD distributions coincide.
- ( $\mathcal{P}_\lambda$ ) As in Proposition 4.2,  $X_1$  follows a Pareto distribution, and  $X_2 \stackrel{d}{=} \lambda_n X_1$ , where  $(\lambda_n)$  fulfills the conditions of Proposition 4.2. This case is interesting (and pretty complicated) since the extreme quantiles of both distributions asymptotically coincide (as  $n \rightarrow \infty$ ). However, in our finite sample size setting,  $\lambda_n \neq 1$  and we expect the test to reject ( $H_0$ ). The empirical non-rejection probabilities  $p_{\lambda_n}(0.05)$  as well as their approximations calculated in Section 4.2 are reported in Figure 2 for  $\xi \in \{0.15; 0.25; 0.35; 0.5\}$ ,  $\beta_n = 0.9$  ( $\approx 1 - 3/\sqrt{n}$ ),  $\lambda_n = 1.2$  ( $= 1 + 2/n^{1/3}$ ) (top right panel) and  $\xi = 0.25$ ,  $\lambda_n \in \{1.1; 1.2; 1.3; 1.4\}$  (bottom left panel).
- ( $\mathcal{P}_\theta$ ) As in Proposition 4.3,  $X_1$  follows a Pareto distribution, and  $X_2 \stackrel{d}{=} X_1^{\theta_n}$ , where  $(\theta_n)$  fulfills the conditions of Proposition 4.3. Here also, the two distributions (and quantiles) are asymptotically the same, but are slightly different with  $n = 1,000 < \infty$ . The empirical non-rejection probabilities  $p_{\theta_n}(0.05)$  as well as their approximations calculated in Section 4.3 are reported in Figure 2 (bottom middle panel) for  $\xi = 0.25$  and  $\theta_n = 1 + k/n^{1/3}$ ,  $1 \leq k \leq 4$ .
- ( $\mathcal{PB}$ )  $X_1$  follows a Pareto distribution and  $X_2$  follows a Burr distribution with second-order parameter  $\rho < 0$ . Such a distribution has a quantile function given by  $q(\alpha) = ((1-\alpha)^\rho - 1)^{-\xi/\rho}$ . Hence, ( $H_0$ ) is fulfilled, and the test is supposed to reject it with probability  $\gamma = 0.05$ . More precisely, when  $\rho$  is small, the Burr distribution is very close to a Pareto one, and one may expect an empirical rejection probability close to 0.05. On the contrary, when  $\rho$  is close to 0, the convergence towards 1 of the ratio of the two extreme quantiles is very slow, and a high rejection rate may be feared. Some results with  $\xi = 0.75$  and  $\rho \in \{-5; -1; -0.5\}$  are reported in Figure 2 (bottom right panel).
- ( $\mathcal{FB}$ )  $X_1$  follows a Fréchet distribution and  $X_2$  follows a Burr distribution with second-order parameter  $\rho < 0$ . Here also, ( $H_0$ ) is fulfilled, and a rejection rate around 0.05 is expected. The results for  $\rho = -1$  and  $\xi = 0.25$  (note that the results are not sensitive to  $\xi$ ) are reported in Figure 2 (bottom right panel).

Note that we limited ourselves to  $\beta_n = 0.9$  ( $\approx 1 - 3/\sqrt{n}$ ) everywhere for convenience, in order to fulfill the assumptions of Propositions 4.1, 4.2 and 4.3. It clearly appears that the rejection rate is increasing with  $L$ . Interestingly, note also that the type I error seems to be decreasing with  $\rho$ . Indeed, for small values of  $\rho$  (let us say  $\rho \leq -1$ ), the choice of  $L = 2$  seems to be the best calibrated one in terms of type I error. However, when  $\rho$  is close to 0, the rejection rate is too low if  $L$  is small, and a choice of a large  $L$  is thus more suited. Through additional simulations (considering  $\tau_\ell = \ell$  and several Burr distributions with different values of  $\rho$  or Generalized Pareto distributions), we observed that for  $n = 1,000$ , a choice of  $L \approx 20$  seems to be optimal when  $\rho = -0.75$  (*i.e.* the rejection rate is around 5%). Similarly,  $L \approx 50$  is tailored when  $\rho = -0.5$ , and  $L \approx 80$  in the challenging case  $\rho = -0.25$ . The results are similar for Burr and Generalized Pareto distributions, and are apparently only sensitive to  $\rho$ , and not to the distribution itself. A prior estimation of  $\rho$  (using for instance the estimators of Gomes and Martins [2002]) may thus be useful to select the parameter  $L$  yielding the best calibration of the

type I error. Finally, the type I and type II errors approximations calculated in Propositions 4.1, 4.2 and 4.3 are remarkably accurate, see the top panels and the bottom left one in Figure 2.

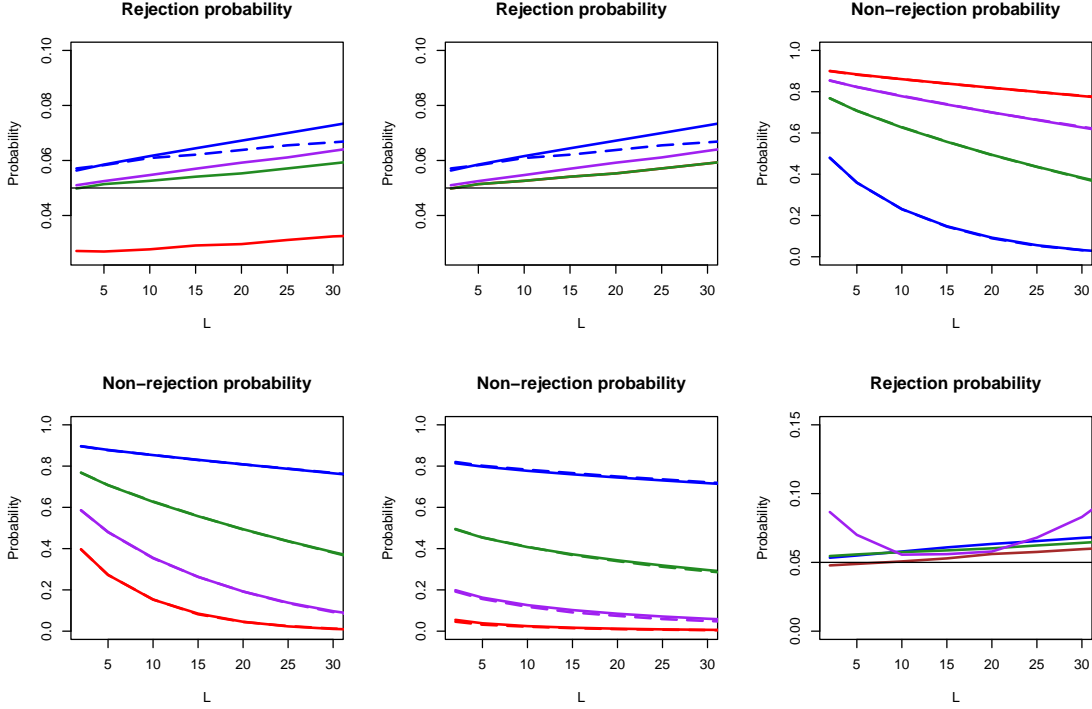


Figure 2: Top left and middle: Example ( $\mathcal{P}$ ) with  $\xi = 0.25$  (left) and  $\xi = 1$  (middle). Empirical (dashed curves) and approximated (solid curve) values of  $p_n(\gamma) = \mathbb{P}_{H_0}(T_n > \chi_{1,1-\gamma}^2)$  as functions of  $L$  for: Pareto vs Pareto (blue), Burr vs Burr (green), Fréchet vs Fréchet (purple) and GPD vs GPD (red). Top right: Example ( $\mathcal{P}_\lambda$ ). Empirical (dotted curves) and approximated (solid curves) values of  $p_{\lambda_n}(\gamma)$  as functions of  $L$ .  $\lambda_n = 1.2$  and  $\xi = 0.15$  (blue curves),  $\xi = 0.25$  (green curves),  $\xi = 0.35$  (purple curves) and  $\xi = 0.5$  (red curves). Bottom left:  $\xi = 0.25$  and  $\lambda_n = 1.1$  (blue curves),  $\lambda_n = 1.2$  (green curves),  $\lambda_n = 1.3$  (purple curves) and  $\lambda_n = 1.4$  (red curves). Bottom middle: Example ( $\mathcal{P}_\theta$ ) with  $\xi = 0.25$ . Empirical (dotted curves) and approximated (solid curves) values of  $p_{\theta_n}(\gamma)$  as functions of  $L$  with  $\theta_n = 1.1$  (green curves),  $\theta_n = 1.2$  (brown curves),  $\theta_n = 1.3$  (blue curves) and  $\theta_n = 1.4$  (red curves). Bottom right: Examples ( $\mathcal{PB}$ ) and ( $\mathcal{FB}$ ). Pareto vs Burr with  $\rho = -5$  (blue),  $\rho = -1$  (green),  $\rho = -0.5$  (purple) and  $\xi = 0.75$ . Fréchet vs Burr with  $\rho = -1$  and  $\xi = 0.25$  (brown). In all examples,  $n = 1,000$ ,  $\beta_n = 0.9$ ,  $\gamma = 0.05$  and  $\alpha_{\ell,n} = 1 - \ell/n$ ,  $\ell = 1, \dots, L$ .

Let us now consider some examples where  $J > 2$ , more specifically  $J \in \{5, 10, 15\}$ :

( $\mathcal{MF}$ )  $X_1, \dots, X_J$  all follow an unit ( $\xi = 1$ ) Fréchet distribution. We are thus supposed to reject ( $H_0$ ) with probability 0.05. Note that such a distribution has no mean, and the classical ANOVA is thus not applicable.

( $\mathcal{MM}$ ) For all  $j = 1, \dots, J$ ,  $X_j$  follows an unit Pareto distribution if  $j = 4k + 1$ , an unit Fréchet distribution if  $j = 4k + 2$ , an unit Burr(-1) distribution if  $j = 4k + 3$  and an unit GPD if

$j = 4(k + 1)$  ( $k \in \mathbb{N}$ ). The extreme quantiles of all these distributions are asymptotically equivalent and  $(H_0)$  is thus satisfied.

( $\mathcal{MP}$ )  $X_1, \dots, X_{J-1}$  all follow an unit Pareto distribution, and  $X_J$  follows a Pareto distribution with tail-index  $\theta$ . If  $\theta \neq 1$ , then  $(H_0)$  obviously has to be rejected. We propose to consider the cases  $\theta = 0.8$  and  $\theta = 1.2$ .

( $\mathcal{CP}$ )  $X_1, \dots, X_{J-1}$  all follow an unit Burr distribution (with  $\rho = -1$ ), and  $X_J \stackrel{d}{=} \pi X_1 + (1-\pi)\sqrt{X_1}$  is a contaminated Burr distribution ( $\pi \in [0, 1]$ ). When  $\pi < 1$ , the quantiles of  $X_J$  are asymptotically equivalent to  $\pi(1-\tau)^{-1}$  and differ from those of  $X_1, \dots, X_{J-1}$  which are asymptotically equivalent to  $(1-\tau)^{-1}$ , and we thus hope to reject  $(H_0)$ . We consider the cases  $\pi \in \{0.2, 0.5, 0.8\}$ .

The results are reported in Figure B.1. Unsurprisingly, the error rates are increasing with the number  $J$  of samples. Comparing the first two examples, it seems that the type I error is not significantly sensitive to the underlying distribution (for two distributions with the same second-order parameter  $\rho$ , confirming the observation made in the previous paragraph). Example ( $\mathcal{MP}$ ) shows that, in more than 65% of the replications, the ANOVEX procedure is able detect whether a sample over 5 has a slightly lower tail-index (0.8 vs 1 for the 4 other samples). The result drops at 50% when a sample has a slightly greater tail-index (1.2 vs 1). The ANOVEX test is thus efficient to discriminate samples with different tail indices. However, the latter is less efficient when all the samples share the same tail-index, but have different scale parameters. Indeed, example ( $\mathcal{CP}$ ) shows that, when the mixture parameter  $\pi$  is slightly lower than 1 (0.8), the test is rejected with a rate of only 10% (when  $J = 5, 10$  or 15). This rate is obviously much better when  $\pi = 0.5$  (more than 25 % when  $J = 5$ ) and  $\pi = 0.2$  (almost 100%).

## 6 Applications on real data

### 6.1 Analysis of stock market indices

In this first real data example, we study log-returns for  $J = 10$  stock market indices, namely AEX (Netherlands), CAC40 (France), Cboe UK 100 (United Kingdom), DAX (Germany), S&P Merval (Argentina), MOEX (Russia), NASDAQ (United States), Nikkei 225 (Japan), SSE Composite (China) and TA-125 (Israel). Since financial markets have strong connections, we expect to observe similar performances across different stock indices. Similarly to [Jondeau and Rockinger \[2003\]](#) and [Mougeot and Tribouley \[2010\]](#), we focus on the tail behaviour of log-returns for different stock indices and apply the ANOVEX procedure to test whether their tails are equal. Data were collected on Yahoo Finance by taking the last 1,001 adjusted closing prices before June 16, 2023 (included), leading to 10 samples of  $n = 1,000$  log-returns. An overview of the data is provided in Figure B.2 and Table 1. Figure B.3 also shows the exponential QQ-plot of the weighted log-spacings  $i \log(X_{n-i+1,n}^{(j)} / X_{n-i,n}^{(j)})$ , for  $i = 1, \dots, 100$  and  $j = 1, \dots, 10$ . The heavy-tail assumption seems valid here since all QQ-plots are close to a straight line with slope  $\xi_j$ , the worst linear fit being observed for NASDAQ and MOEX indices. We consider  $\beta_n = 0.9$ ,  $\alpha_{\ell,n} = 1 - \ell/n$  for all  $\ell = 1, \dots, L$ , and a confidence level of 95%, such that  $\gamma = 0.05$ .

A first naive ANOVEX test with  $L = 2$  and  $J = 10$  shows that  $(H_0)$  is clearly rejected, the  $p$ -value being  $1.51 \times 10^{-4}$ . Since the rejection rate is increasing with  $L$  (see Section 5), one can reasonably conclude that all tails are not equal. In Table 1, the S&P Merval index appears to be significantly



	Mean $10^{-4}$	Var. $10^{-4}$	$\widehat{\xi}_j^H(\beta_n)$	$\widehat{q}_j^H(\beta_n)$ $10^{-2}$		Mean $10^{-4}$	Var. $10^{-4}$	$\widehat{\xi}_j^H(\beta_n)$	$\widehat{q}_j^H(\beta_n)$ $10^{-2}$
AEX	2.81	1.67	0.343	1.40	MOEX	0.70	4.30	0.442	1.48
CAC 40	2.63	1.99	0.417	1.40	NASDAQ	5.55	2.92	0.376	1.85
Cboe UK 100	0.23	1.42	0.439	1.17	Nikkei 225	4.70	1.58	0.356	1.51
DAX	2.78	2.06	0.454	1.37	SSE Composite	1.06	1.10	0.401	1.16
S&P MERVAL	24.7	188	0.382	3.28	TA-125	2.14	1.44	0.334	1.32

Table 1: Application to stock market indices. Columns report empirical mean, variance, tail-index and intermediate  $\beta_n^{\text{th}}$  quantile of the  $n = 1,000$  log-returns for each financial index.

different from the others. This is confirmed by separately testing the S&P MERVAL against all the other indices: all bivariate ANOVEX tests are rejected with  $L = 2$ .

In the following, we exclude S&P MERVAL from the study and test the equality of the remaining  $J = 9$  tails. With  $L = 2$ , the ANOVEX test is no longer rejected, the  $p$ -value being 0.132, but we find that it is strongly rejected when  $L$  is increased to larger values. As suggested in Section 5, we propose to select  $L$  so as to calibrate the type I error. First, note that Table 1 reports the tail indices estimated around 0.3-0.4 for all of the 9 samples. Moreover, using the R function `mop` in the `Expectrem` package, the second-order parameter  $\rho$  is estimated at  $\hat{\rho} \simeq -0.7$  in all 9 samples. To choose an appropriate value of  $L$ , we perform a simulation study with 9 samples of size 1,000 following a Burr distribution with  $\gamma = 0.35$  and  $\rho = -0.7$ . We obtain that  $L = 30$  provides a type I error around  $\gamma = 0.05$ , and we therefore set  $L = 30$ . With this choice,  $(H_0)$  is rejected for the stock indices, with a  $p$ -value of  $2.3 \times 10^{-3}$ , and we conclude that the log-return tails are significantly different.

	AEX	CAC 40	Cboe UK 100	DAX	MOEX	NASDAQ	Nikkei 225	SSE Composite
AEX	×							
CAC 40	0.161	×						
Cboe UK 100	0.511	0.364	×					
DAX	0.064	0.638	0.200	×				
MOEX	0.030	0.453	0.111	0.764	×			
NASDAQ	0.004	0.165	0.021	0.334	0.503	×		
Nikkei 225	0.359	0.544	0.478	0.281	0.175	0.047	×	
SSE Composite	0.567	0.102	0.481	0.044	0.019	0.002	0.192	×
TA-125	0.496	0.043	0.268	0.015	0.005	0.0003	0.110	0.641

Table 2: Application to stock market indices. Pairwise ANOVEX-based  $p$ -values for 9 indices, here obtained with  $L = 25$ . Results smaller than 0.05 are reported in red.

To study more precisely which groups of stock indices show different tail behaviour, we study the  $p$ -values of the ANOVEX tests for pairs of samples reported in Table 2. The log-returns of the European stock market indices (AEX, CAC 40, Cboe UK 100 and DAX) could be gathered in a first group with homogeneous tail behaviour by further including Nikkei 225. Indeed, none of the pairwise tests is rejected for this group of 5 samples, and the  $p$ -value for the test of the whole group with  $J = 5$  is 0.335). These results are in line with results in Jondeau and Rockinger [2003], where one of the groups is composed of France, Germany, Japan, UK and US. The only difference here is that the ANOVEX test excludes the US index NASDAQ from this group. In the study of Mougeot and Tribouley [2010], the CAC 40, FTSE 100 (another UK stock market index) and Nikkei 225 are gathered together, *i.e.*,

they share similar tail indices. Based on the pairwise  $p$ -values, we can also pool the NASDAQ (which is significantly different from the AEX and Cboe UK 100) and the MOEX (which is also very different from the AEX) in a second group, and a last group could be composed of the SSE Composite, with tail behaviour different from DAX, MOEX and the NASDAQ, present in the other two groups) and TA-125, for which ( $H_0$ ) is most of the time rejected.

To sum up, the ANOVEX test allows us to split the stock market indices into four clusters based on the tail behaviour of their log-returns:

- Cluster 1: AEX, CAC 40, Cboe UK 100, DAX and Nikkei 225.
- Cluster 2: NASDAQ and MOEX.
- Cluster 3: SSE Composite and TA-125.
- Cluster 4: S&P Merval.

In view of the exponential QQ-plots of Figure B.3, Cluster 2 is composed of the two indices NASDAQ and MOEX, for which heavy-tail behaviour of log-returns may be questionable. Finally, we have compared the right-tail of the log-return distributions. For risk management issues, it may also be interesting to compare the left-tails (i.e. the losses). By doing the same approach, we find similar results (and almost the same clusters), with one difference: the Nikkei 225 belongs to Cluster 3 (with SSE Composite and TA-125) instead of Cluster 1.

## 6.2 Analysis of daily precipitation in Germany

To illustrate the behaviour of ANOVEX for  $J > 2$  groups, we use the test to detect nonstationary behaviour in extremes of daily accumulated precipitation across  $J = 6$  decades for observations collected at weather stations in Germany. Heavy-tailedness is generally accepted for daily precipitation measurements. We consider two periods: 1961–2020 with only few missing observations at 4342 weather stations, and 1901–1960 with sparser spatial coverage through 920 stations with few missing observations. For each combination of an observation series and period with available data, the different samples consist of the observations for the different decades, either  $(1901 + (j - 1) \times 10) - (1901 + (j \times 10))$  or  $(1961 + (j - 1) \times 10) - (1961 + (j \times 10))$ , with  $j \in \{1, 2, \dots, 6\}$  and  $J = 6$ . We test ( $H_0$ ) using the asymptotic  $\chi^2$ -distribution under ( $H_0$ ) with  $\gamma = 0.05$  and  $L = 10$ , and  $\beta_{j,n} = \beta_n$  is chosen to retain  $k_{j,n} = k = 100$  extreme order statistics for each sample  $j = 1, \dots, 6$ . Figure 3 reports results where we highlight gauges with significant nonstationarity across decades, and we also report relatively low  $p$ -values larger than  $\gamma = 0.05$ . For the 1961–2020 period, the proportion of gauges with ( $H_0$ ) rejected is around 5% and therefore of the order of  $\gamma$ , which corresponds to the expected number of type I errors under  $H_0$ . For 1901–1960, a higher proportion of 16% of stations has ( $H_0$ ) rejected. Unreported results for other choices of  $k$  and  $L$  show relatively stable behaviour of  $p$ -values over a range of  $k$  between 50 and 250 for fixed  $L$ , with slightly higher values for larger values of  $k$  for which estimation uncertainty is lower, thus indicating no strong sensitivity to the choice of  $k$ .  $p$ -values tend to be relatively lower when increasing  $L$  for fixed  $k$ , which is natural since the test then considers a larger number  $L$  of extreme quantiles, some of them estimated at lower probability levels, such that statistical uncertainty decreases and the power of the test increases; however, we would then test for differences in the distribution of events that may be less extreme than those considered with  $L = 10$ .

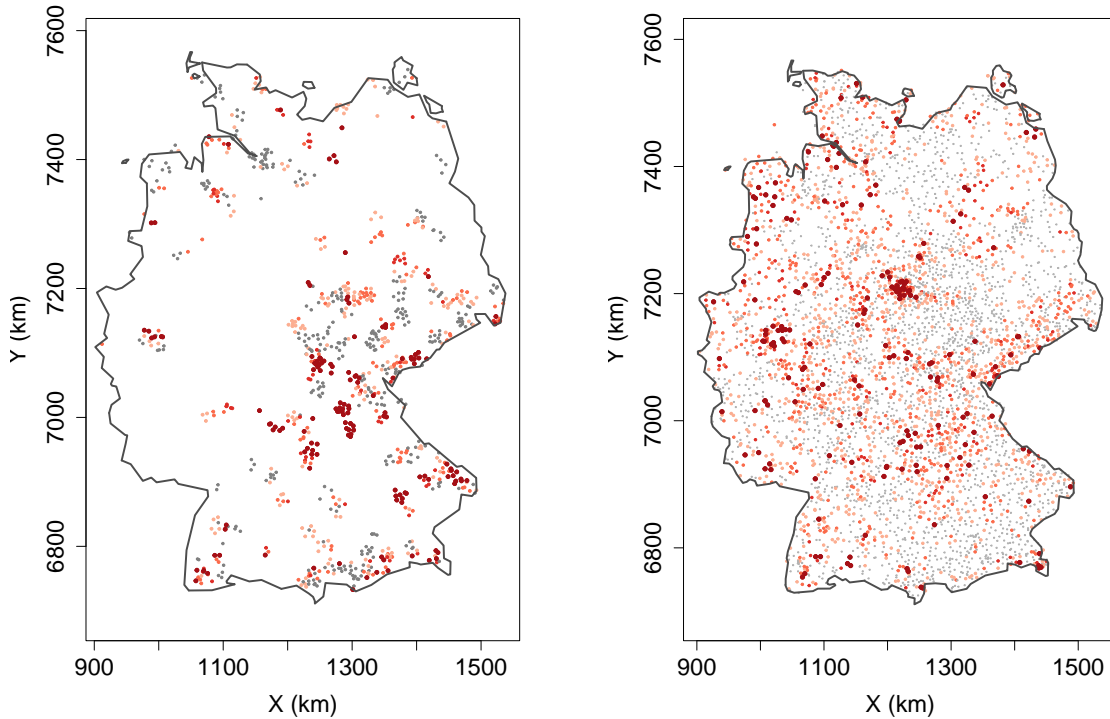


Figure 3: Application to German precipitation series. Left: 1901–1960 period; right: 1961–2020 period. Obtained  $p$ -values are shown differently according to intervals  $[0.0, 0.05]$  (dark red, big points),  $(0.05, 0.5]$  (lighter red, smaller points), and  $(0.5, 1]$  (grey, small points).

Follow-up work will include expert interpretations by climatologists in terms of observation biases and existing knowledge about local climate systems and potential climate-change effects.

## 7 Conclusion

We have introduced a new simple test for heterogeneity in the upper distribution tail, based on a decomposition of variability of extreme log-quantiles. This approach is called ANOVEX and draws inspiration from classical ANOVA. It allows putting focus on differences arising for the distribution of the most extreme possible events, for which uncertainties are usually high, especially with naive approaches not making use of extreme-value statistics for extrapolation. The simulation study highlighted that our test is an easy-to-use, interpretable and efficient tool for detecting different tail behaviours. It is a valuable complement to existing tests that focus on more specific alternative hypotheses, such as differences in the tail-index or, provided that tail indices coincide, in the tail scale. Our test is particularly powerful for detecting different tail scales in the case of small tail indices. We have thoroughly investigated its asymptotic properties, including the limit distribution of the test statistic, and we provide highly accurate second-order approximations of type I and type II errors for the class of Pareto distributions. Overall, ANOVEX shows promising capability in identifying tail heterogeneity and can be used for applications in various statistical analyses involving heavy-tailed data.

In future work, we aim to investigate how an automatic choice of  $L$  could be achieved, for example by establishing an explicit link between  $L$  and the second-order parameter  $\rho$ . We also plan to relax certain assumptions, notably by considering dependent data (dependence across groups, and/or serial dependence within groups). In the current study, we have considered samples with equal sizes in the error approximations and the simulation study. Although our theoretical result (Theorem 3.1) allows samples to have different sizes, biases may arise in practice, particularly when dealing with relatively small samples and strongly unbalanced sample sizes, resulting in type I error probabilities deviating significantly from the test level.

Furthermore, we plan to leverage the ANOVEX test statistic for change point detection. More generally, our test statistic provides a valuable tail dissimilarity measure for machine-learning tools, such as for making splits in regression trees and random forests. In recent related work, Maillard and Robert [2023] proposed estimation of a tree but they focus solely on the tail-index. Farkas et al. [2021] adopted a model-based approach for inferring trees by using the Generalized Pareto Distribution (GPD) for exceedances above a high threshold. Gnecco et al. [2022] developed GPD-based random forests; however, the tree structure and likelihood weights are computed in a preliminary step using traditional quantile regression forests. To extend these existing approaches, an ANOVEX-based procedure could offer a rapid and robust likelihood-free method for tail prediction, applicable to both trees and random forests.

## 8 Acknowledgements

This work is part of the ANOVEX project funded by the French national research institutes INRAE and Inria within the scope of their collaborations on the topic of environmental risks. This work is further supported by the French National Research Agency (ANR) in the framework of the Investissements d’Avenir Program (ANR-15-IDEX-02). S. Girard acknowledges the support of the Chair “Stress Test, Risk Management and Financial Steering”, led by the French Ecole Polytechnique and its Foundation and sponsored by BNP Paribas.

## References

- P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels. *Statistics of extremes: theory and applications*, volume 558. John Wiley & Sons, 2004.
- F. Caeiro, M.I. Gomes, and D. Pestana. Direct reduction of bias of the classical Hill estimator. *REVSTAT-Statistical Journal*, 3(2):113–136, 2005.
- A. Clauset, C.R. Shalizi, and M.E.J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.
- A. Daouia, S. Girard, and G. Stupfler. Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80:263–292, 2018.

- A. Daouia, G. Stupfler, and A. Usseglio-Carleve. Inference for extremal regression with dependent heavy-tailed data. *The Annals of Statistics*, 2023. To appear.
- A.C. Davison and R.L. Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 52(3):393–425, 1990.
- L. de Haan and A. Ferreira. *Extreme value theory: an introduction*, volume 21. Springer, 2006.
- D. Deuber, J. Li, S. Engelke, and M.H. Maathuis. Estimation and inference of extremal quantile treatment effects for heavy-tailed distributions. *Journal of the American Statistical Association*, 2023. To appear.
- J.H.J. Einmahl, L. Haan, and C. Zhou. Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(1):31–51, 2016.
- J. El-Methni, L. Gardes, and S. Girard. Nonparametric estimation of extreme risks from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, 41:988–1012, 2014.
- P. Embrechts, S.I. Resnick, and G. Samorodnitsky. Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41, 1999.
- S. Farkas, A. Heranval, O. Lopez, and M. Thomas. Generalized Pareto Regression Trees for extreme events analysis. *arXiv preprint arXiv:2112.10409*, 2021.
- N. Gnecco, E.M. Terefe, and S. Engelke. Extremal random forests. *arXiv preprint arXiv:2201.12865*, 2022.
- M.I. Gomes and M.J. Martins. “Asymptotically unbiased” estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5:5–31, 2002.
- B.M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975.
- Y. Hoga. Testing for changes in (extreme) VaR. *The Econometrics Journal*, 20(1):23–51, 2017.
- M. Hollander, D.A. Wolfe, and E. Chicken. *Nonparametric statistical methods*. John Wiley & Sons, 2013.
- E. Jondeau and M. Rockinger. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, 10(5):559–581, 2003.
- E. Jondeau, S.-H. Poon, and M. Rockinger. *Financial modeling under non-Gaussian distributions*. Springer Science & Business Media, 2007.
- R. Katz, M. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advances in Water Resources*, 25(8-12):1287–1304, 2002.
- B. Klartag. A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probability Theory and Related Fields*, 145(1):1–33, 2009.
- J. Koh, F. Pimont, J.-L. Dupuy, and T. Opitz. Spatiotemporal wildfire modeling through point processes with moderate and extreme marks. *The Annals of Applied Statistics*, 17(1):560–582, 2023.

- A. Maillard and C.Y. Robert. Tail index partition-based rules extraction with application to tornado damage insurance. *ASTIN Bulletin: The Journal of the IAA*, 53(2):258–284, 2023.
- J. Marden and M. Muyot. Rank tests for main and interaction effects in analysis of variance. *Journal of the American Statistical Association*, 90(432):1388–1398, 1995.
- T. Mikosch. Heavy-tailed modelling in insurance. *Communications in Statistics. Stochastic models*, 13(4):799–815, 1997.
- M. Mougeot and K. Tribouley. Procedure of test to compare the tail indices. *Annals of the Institute of Statistical Mathematics*, 62:383–412, 2010.
- W. Newey and J. Powell. Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847, 1987.
- J.M.C. Pereira and K.F. Turkman. Statistical models of vegetation fires: Spatial and temporal patterns. In *Handbook of Environmental and Ecological Statistics*, pages 401–420. Chapman and Hall/CRC, 2019.
- V. Pisarenko and M. Rodkin. *Heavy-tailed distributions in disaster analysis*, volume 30. Springer Science & Business Media, 2010.
- S. Rachev. *Handbook of heavy tailed distributions in finance: Handbooks in finance, Book 1*. Elsevier, 2003.
- A. Rényi. On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica*, 4(2):191–231, 1953.
- M. Rizzo and G. Székely. DISCO analysis: A nonparametric extension of analysis of variance. *The Annals of Applied Statistics*, 4(2):1034 – 1055, 2010.
- H. Scheffe. *The analysis of variance*, volume 72. John Wiley & Sons, 1999.
- I. Weissman. Estimation of parameters and large quantiles based on the  $k$  largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.
- J. Worms and R. Worms. A test for comparing tail indices for heavy-tailed distributions via empirical likelihood. *Communications in Statistics-Theory and Methods*, 44(15):3289–3302, 2015.

# Appendix

## A Proofs of the theoretical results

Some technical lemmas are collected in Paragraph A.1, they will reveal useful to prove the main results in Paragraph A.2.

### A.1 Preliminary results

We first show that, under  $(H_0)$ , if all cumulative distribution functions  $F_1, \dots, F_J$  are heavy-tailed, then, necessarily, their tail-indices are the same.

**Lemma A.1.** *Assume  $(H_0)$  holds. If  $F_j \in \mathcal{C}_1(\xi_j)$  for  $j = 1, \dots, J$  then*

$$(i) \quad \xi_1 = \dots = \xi_J =: \xi.$$

$$(ii) \quad \log q_j(\alpha) = -\xi \log(1 - \alpha)(1 + o(1)) \text{ for all } j = 1, \dots, J \text{ as } \alpha \rightarrow 1.$$

*Proof of Lemma A.1.* (i) Let us recall that  $F_j \in \mathcal{C}_1(\xi_j)$  implies that there exists a slowly-varying function  $L_j$  such that

$$q_j(\alpha) = (1 - \alpha)^{-\xi_j} L_j(1/(1 - \alpha)). \quad (15)$$

It straightforwardly follows that  $q_{j_0}(\alpha)/q_j(\alpha) \rightarrow 0$  as  $\alpha \rightarrow 1$  if  $\xi_j > \xi_{j_0}$  and  $q_{j_0}(\alpha)/q_j(\alpha) \rightarrow +\infty$  as  $\alpha \rightarrow 1$  if  $\xi_j < \xi_{j_0}$ . The result (i) is thus proved.

(ii) From Lemma A.1(i), one has  $\xi_1 = \dots = \xi_J =: \xi$  and therefore (15) can be rewritten as

$$\log(q_j(\alpha)) = -\xi \log(1 - \alpha) + \log(L_j(1/(1 - \alpha))) = -\xi \log(1 - \alpha)(1 + o(1)),$$

which proves the result.  $\square$

The second lemma provides an asymptotic equivalent of  $\Delta_{1,n}$  (defined in Proposition 2.1) as  $n \rightarrow \infty$ .

**Lemma A.2.** *Assume  $F_j \in \mathcal{C}_1(\xi_j)$  for  $j = 1, \dots, J$ .*

- For all  $n \geq 1$  and  $\ell = 1, \dots, L$ , let  $\alpha_{\ell,n} \in (0, 1)$  such that  $\alpha_{\ell,n} \rightarrow 1$  as  $n \rightarrow \infty$ .
- Let  $\tilde{q}_j(\alpha_{\ell,n})$  be an estimator of the extreme quantile  $q_j(\alpha_{\ell,n})$  computed on the  $n_j$ -sample  $E_j$  such that  $n_j/n \rightarrow \lambda_j > 0$  as  $n \rightarrow \infty$  and

$$\sigma_{j,\ell,n}^{-1} (\log \tilde{q}_j(\alpha_{\ell,n}) - \log q_j(\alpha_{\ell,n})) \xrightarrow{d} \mathcal{N}(0, \xi_j^2),$$

for some  $\sigma_{j,\ell,n} \rightarrow \infty$  as  $n \rightarrow \infty$ , and for all  $j = 1, \dots, J$  and  $\ell = 1, \dots, L$ .

Then, under  $(H_0)$ ,

$$\Delta_{1,n} = \frac{\xi^2}{L} \sum_{\ell=1}^L \left[ \frac{1}{J} \sum_{j=1}^J \sigma_{j,\ell,n}^2 Z_j^2 - \left( \frac{1}{J} \sum_{k=1}^J \sigma_{k,\ell,n} Z_k \right)^2 \right] (1 + o_{\mathbb{P}}(1)),$$

where  $Z_1, \dots, Z_J$  follow independent standard Gaussian distributions.

*Proof of Lemma A.2.* Let us first recall that for all  $\ell = 1, \dots, L$ ,

$$\Delta_{1,\ell,n} = \frac{1}{J} \sum_{j=1}^J \left( \log \tilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right)^2 \quad \text{with } \mu_{\alpha,n}^{(\ell)} = \frac{1}{J} \sum_{j=1}^J \log \tilde{q}_j(\alpha_{\ell,n}).$$

Under  $(H_0)$ , we have in view of Lemma A.1(i):

$$\log \tilde{q}_j(\alpha_{\ell,n}) = \log q_j(\alpha_{\ell,n}) + \sigma_{j,\ell,n} \xi Z_j (1 + o_{\mathbb{P}}(1)),$$

where  $Z_1, \dots, Z_J$  follow independent standard Gaussian distributions. Lemma A.1(ii) entails

$$\log \tilde{q}_j(\alpha_{\ell,n}) = -\xi \log(1 - \alpha_{\ell,n})(1 + o(1)) + \sigma_{j,\ell,n} \xi Z_j (1 + o_{\mathbb{P}}(1)),$$

and therefore,

$$\left( \log \tilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right)^2 = \xi^2 \left( \sum_{k=1}^J \kappa_{k,j} \sigma_{k,\ell,n} Z_k \right)^2 (1 + o_{\mathbb{P}}(1)), \quad \text{with } \kappa_{j,j} = 1 - \frac{1}{J} \text{ and } \kappa_{k,j} = -\frac{1}{J} \text{ if } k \neq j,$$

so that

$$\begin{aligned} \Delta_{1,\ell,n} &= \frac{J-1}{J^2} \xi^2 \sum_{j=1}^J \sigma_{j,\ell,n}^2 Z_j^2 (1 + o_{\mathbb{P}}(1)) - \frac{2}{J^2} \xi^2 \sum_{1 \leq j < k \leq J} \sigma_{j,\ell,n} \sigma_{k,\ell,n} Z_j Z_k (1 + o_{\mathbb{P}}(1)) \\ &= \xi^2 \left[ \frac{1}{J} \sum_{j=1}^J \sigma_{j,\ell,n}^2 Z_j^2 - \left( \frac{1}{J} \sum_{k=1}^J \sigma_{k,\ell,n} Z_k \right)^2 \right] (1 + o_{\mathbb{P}}(1)). \end{aligned}$$

Finally, without any further assumption, the term  $\Delta_{1,n}$  may be written as follows:

$$\Delta_{1,n} = \frac{\xi^2}{L} \sum_{\ell=1}^L \left[ \frac{1}{J} \sum_{j=1}^J \sigma_{j,\ell,n}^2 Z_j^2 - \left( \frac{1}{J} \sum_{k=1}^J \sigma_{k,\ell,n} Z_k \right)^2 \right] (1 + o_{\mathbb{P}}(1)),$$

which is the desired result.  $\square$

The third lemma provides an asymptotic equivalent of  $\Delta_{2,n}$  (defined in Proposition 2.1) as  $n \rightarrow \infty$ .

**Lemma A.3.** *Assume  $F_j \in \mathcal{C}_1(\xi_j)$  for  $j = 1, \dots, J$ .*

- For all  $n \geq 1$  and  $\ell = 1, \dots, L$ , let  $\alpha_{\ell,n} \in (0, 1)$  such that  $\alpha_{\ell,n} \rightarrow 1$  as  $n \rightarrow \infty$ .
- Let  $\tilde{q}_j(\alpha_{\ell,n})$  be an estimator of the extreme quantile  $q_j(\alpha_{\ell,n})$  computed on the  $n_j$ -sample  $E_j$  such that  $n_j/n \rightarrow \lambda_j > 0$  as  $n \rightarrow \infty$  and

$$\log \tilde{q}_j(\alpha_{\ell,n}) / \log q_j(\alpha_{\ell,n}) \xrightarrow{\mathbb{P}} 1, \quad (16)$$

for all  $j = 1, \dots, J$  and  $\ell = 1, \dots, L$ .

Then, under  $(H_0)$ ,

$$\Delta_{2,n} = \xi^2 \text{varlog}((1 - \alpha_n)_{1:L})(1 + o_{\mathbb{P}}(1)).$$



*Proof of Lemma A.3.* Combining (16) with Lemma A.1(ii) yields

$$\log \tilde{q}_j(\alpha_{\ell,n}) = (\log q_j(\alpha_{\ell,n})) (1 + o_{\mathbb{P}}(1)) = -\xi \log(1 - \alpha_{\ell,n})(1 + o_{\mathbb{P}}(1))$$

so that

$$\mu_{\alpha,n}^{(\ell)} = \frac{1}{J} \sum_{j=1}^J \log \tilde{q}_j(\alpha_{\ell,n}) = -\xi \log(1 - \alpha_{\ell,n})(1 + o_{\mathbb{P}}(1)),$$

for all  $\ell = 1, \dots, L$ , and consequently,

$$\mu_{\alpha,n} - \mu_{\alpha,n}^{(\ell)} = -\xi \frac{1}{L} \sum_{k=1}^L \tilde{\kappa}_{k,\ell} \log(1 - \alpha_{k,n})(1 + o_{\mathbb{P}}(1)),$$

where  $\tilde{\kappa}_{\ell,\ell} = 1 - L$  and  $\tilde{\kappa}_{k,\ell} = 1$  if  $k \neq \ell$ . Some straightforward calculations lead to

$$\begin{aligned} \Delta_{2,n} &= \frac{1}{L} \sum_{\ell=1}^L \left( \mu_{\alpha,n} - \mu_{\alpha,n}^{(\ell)} \right)^2 \\ &= \xi^2 \left( \frac{L-1}{L^2} \sum_{\ell=1}^L \log(1 - \alpha_{\ell,n})^2 - \frac{2}{L^2} \sum_{1 \leq \ell < \ell' \leq L} \log(1 - \alpha_{\ell,n}) \log(1 - \alpha_{\ell',n}) \right) (1 + o_{\mathbb{P}}(1)) \\ &= \xi^2 \left( \frac{1}{L} \sum_{\ell=1}^L \log(1 - \alpha_{\ell,n})^2 - \frac{1}{L^2} \sum_{\ell=1}^L \sum_{\ell'=1}^L \log(1 - \alpha_{\ell,n}) \log(1 - \alpha_{\ell',n}) \right) (1 + o_{\mathbb{P}}(1)) \\ &= \xi^2 \left[ \frac{1}{L} \sum_{\ell=1}^L \log(1 - \alpha_{\ell,n})^2 - \left( \frac{1}{L} \sum_{\ell=1}^L \log(1 - \alpha_{\ell,n}) \right)^2 \right] (1 + o_{\mathbb{P}}(1)), \end{aligned}$$

which is the expected result.  $\square$

The next lemma provides precise asymptotic representations associated with Hill estimators and intermediate quantiles computed on Pareto samples.

**Lemma A.4.** *Consider two independent samples  $E_1 = \{X_1^{(1)}, \dots, X_n^{(1)}\}$  and  $E_2 = \{X_1^{(2)}, \dots, X_n^{(2)}\}$ , both distributed from a Pareto distribution  $\mathcal{P}(1/\xi)$ ,  $\xi > 0$ . Let  $(\beta_n)$  be an intermediate probability level,  $\hat{q}_j(\beta_n) := X_{[n(1-\beta_n)]}^{(j)}$  and let  $\hat{\xi}_j^{\text{H}}(\beta_n)$  be the associated Hill estimators (3),  $j = 1, 2$ . Then, the following asymptotic representations hold:*

$$\hat{\xi}_j^{\text{H}}(\beta_n) \stackrel{d}{=} \xi + \frac{\xi}{\sqrt{n(1-\beta_n)}} \Gamma_{\xi,j} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) \right), \quad (17)$$

$$\log \left( \frac{\hat{q}_j(\beta_n)}{q_j(\beta_n)} \right) \stackrel{d}{=} \frac{\xi}{\sqrt{n(1-\beta_n)}} \Gamma_{q,j} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}}(1-\beta_n) \right), \quad (18)$$

$$\hat{\xi}_1^{\text{H}}(\beta_n) - \hat{\xi}_2^{\text{H}}(\beta_n) \stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_{\xi} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1-\beta_n)} \right) \right), \quad (19)$$

$$\log \left( \frac{\hat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\hat{q}_2(\beta_n)}{q_2(\beta_n)} \right) \stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1-\beta_n)} \right) + O_{\mathbb{P}}(1-\beta_n) \right), \quad (20)$$

where  $\Gamma_{q,j}$ ,  $\Gamma_{\xi,j}$ ,  $\Gamma_q$  and  $\Gamma_{\xi}$  are standard Gaussian random variables,  $j = 1, 2$ . Moreover, Equations (19) and (20) also hold for two Pareto samples with different scale parameters:  $X_i^{(2)} \stackrel{d}{=} \lambda X_i^{(1)}$ ,  $\lambda > 0$ ,  $i = 1, \dots, n$ .

*Proof of Lemma A.4.* Let us introduce  $k_n = \lfloor n(1-\beta_n) \rfloor$  to simplify the notations. Let  $j \in \{1, 2\}$ . First, Rényi's representation entails that the log-spacings  $\left(\log(X_{n-i,n}^{(j)}) - \log(X_{n-k_n,n}^{(j)})\right)$ ,  $i = 0, \dots, k_n - 1$  are independent and exponentially distributed. Hill estimators are thus Gamma distributed:

$$\widehat{\xi}_j^H(\beta_n) \stackrel{d}{=} \frac{1}{k_n} \sum_{i=1}^{k_n} \mathcal{E}_i^{(j)},$$

where  $\{\mathcal{E}_1^{(j)}, \dots, \mathcal{E}_{k_n}^{(j)}\}$  are i.i.d. realisations of an exponential distribution with mean  $\xi$ . Berry-Esseen Theorem thus yields

$$\widehat{\xi}_j^H(\beta_n) - \xi \stackrel{d}{=} \frac{\xi}{\sqrt{k_n}} \Gamma_{\xi,j} \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right)\right),$$

and (17) is proved. Second, one has, in view of Rényi's representation:

$$\log\left(\frac{\widehat{q}_j(\beta_n)}{q_j(\beta_n)}\right) \stackrel{d}{=} \mathcal{E}_{n-k_n,n}^{(j)} - \xi \log(n/k_n) \stackrel{d}{=} \sum_{i=k_n+1}^n \frac{\mathcal{E}_i^{(j)}}{i} - \xi \log(n/k_n),$$

and thus, introducing  $Y_i^{(j)} = (\mathcal{E}_i^{(j)} - \xi)/(\xi i)$ , the following expansion holds

$$\sqrt{k_n} \log\left(\frac{\widehat{q}_j(\beta_n)}{q_j(\beta_n)}\right) \stackrel{d}{=} \xi \sqrt{k_n} \sum_{i=k_n+1}^n Y_i^{(j)} + \xi \sqrt{k_n} \left(\sum_{i=k_n+1}^n \frac{1}{i} - \log(n/k_n)\right) =: \xi(A_n^{(j)} + B_n).$$

The well-known formula

$$\sum_{i=1}^n \frac{1}{i} = \log(n) + \gamma - \frac{1}{2n}(1 + o(1)),$$

(where  $\gamma$  is Euler's constant) entails that the non-random term can be controlled as  $B_n = O(1/\sqrt{k_n})$ . Letting  $\sigma_i^2 = \mathbb{E}((Y_i^{(j)})^2)$  and  $\rho_i = \mathbb{E}(|Y_i^{(j)}|^3)$ , Berry-Esseen Theorem for non identically distributed random variables shows that

$$\frac{\sum_{i=k_n+1}^n Y_i^{(j)}}{\sqrt{\sum_{i=k_n+1}^n \sigma_i^2}} = \Gamma_{q,j} + O_{\mathbb{P}}\left(\frac{\max_{i=k_n+1, \dots, n} \rho_i / \sigma_i^2}{\sqrt{\sum_{i=k_n+1}^n \sigma_i^2}}\right),$$

or equivalently,

$$A_n^{(j)} = \sqrt{k_n \sum_{i=k_n+1}^n \sigma_i^2} \Gamma_{q,j} + O_{\mathbb{P}}\left(\sqrt{k_n} \max_{i=k_n+1, \dots, n} \rho_i / \sigma_i^2\right).$$

Moreover,  $\sigma_i^2 = 1/i^2$ ,  $\rho_i = c/i^3$  with  $c > 0$  so that

$$\begin{aligned} \sum_{i=k_n+1}^n \sigma_i^2 &= \frac{1}{k_n} \left(1 + O\left(\frac{1}{k_n}\right)\right) - \frac{1}{n} \left(1 + O\left(\frac{1}{n}\right)\right) = \frac{1}{k_n} \left(1 + O\left(\frac{1}{k_n}\right) + O\left(\frac{k_n}{n}\right)\right), \\ \max_{i=k_n+1, \dots, n} \rho_i / \sigma_i^2 &= O\left(\frac{1}{k_n}\right), \end{aligned}$$

and therefore  $A_n^{(j)} = \Gamma_{q,j} + O_{\mathbb{P}}(1/\sqrt{k_n}) + O_{\mathbb{P}}(k_n/n)$ . All in all,

$$\sqrt{k_n} \log\left(\frac{\widehat{q}_j(\beta_n)}{q_j(\beta_n)}\right) \stackrel{d}{=} \xi \Gamma_{q,j} + O(1/\sqrt{k_n}) + O_{\mathbb{P}}(k_n/n),$$

and (18) is proved. Moreover, one has

$$\widehat{\xi}_1^H(\beta_n) - \widehat{\xi}_2^H(\beta_n) \stackrel{d}{=} \frac{(\mathcal{E}_1^{(1)} - \mathcal{E}_1^{(2)}) + \dots + (\mathcal{E}_{k_n}^{(1)} - \mathcal{E}_{k_n}^{(2)})}{k_n} \stackrel{d}{=} \frac{\mathcal{L}_1 + \dots + \mathcal{L}_{k_n}}{k_n},$$

where  $\{\mathcal{L}_1, \dots, \mathcal{L}_{k_n}\}$  are i.i.d. realisations of a centered Laplace distribution with variance  $2\xi^2$ . Since the Laplace distribution is log-concave, centered and symmetric, [Klartag, 2009, Theorem 1] may be applied to refine the Berry-Esseen bound with  $k_n$  (or equivalently  $n(1 - \beta_n)$ ) instead of  $\sqrt{k_n}$  (or equivalently  $\sqrt{n(1 - \beta_n)}$ ), hence the third result (19). Similarly,

$$\log\left(\frac{\widehat{q}_1(\beta_n)}{\widehat{q}_2(\beta_n)}\right) = \log\left(\frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)}\right) - \log\left(\frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)}\right) \stackrel{d}{=} \sum_{i=k_n+1}^n \frac{(\mathcal{E}_i^{(1)} - \mathcal{E}_i^{(2)})}{i} \stackrel{d}{=} \sum_{i=k_n+1}^n \frac{\mathcal{L}_i}{i},$$

and the second result of [Klartag, 2009, Theorem 1] can be used to establish the Berry-Esseen bound. Rewriting

$$\sqrt{k_n} \log\left(\frac{\widehat{q}_1(\beta_n)}{\widehat{q}_2(\beta_n)}\right) \stackrel{d}{=} \sqrt{k_n} \sqrt{\sum_{i=k_n+1}^n \frac{1}{i^2}} \sum_{i=k_n+1}^n \theta_{i,n} \mathcal{L}_i,$$

with  $\theta_{i,n} = \frac{1}{i\sqrt{\sum_{j=k_n+1}^n 1/j^2}}$  and  $\sum_{i=k_n+1}^n \theta_{i,n}^2 = 1$ , it thus follows:

$$\sum_{i=k_n+1}^n \theta_{i,n} \mathcal{L}_i = \sqrt{2\xi} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \sum_{i=k_n+1}^n \theta_{i,n}^4 \right) \right).$$

Straightforward calculations on Riemann series lead to

$$\begin{aligned} \sum_{i=k_n+1}^n \theta_{i,n}^4 &= \frac{\sum_{i=k_n+1}^n \frac{1}{i^4}}{\left(\sum_{i=k_n+1}^n \frac{1}{i^2}\right)^2} = \frac{\sum_{i=k_n+1}^{\infty} \frac{1}{i^4} - \sum_{i=n+1}^{\infty} \frac{1}{i^4}}{\left(\sum_{i=k_n+1}^{\infty} \frac{1}{i^2} - \sum_{i=n+1}^{\infty} \frac{1}{i^2}\right)^2} = \frac{O\left(\frac{1}{k_n^3}\right)}{\left(\frac{1}{k_n} + O\left(\frac{1}{k_n^2}\right) + O\left(\frac{1}{n}\right)\right)^2} = O\left(\frac{1}{k_n}\right), \\ \sqrt{k_n} \sum_{i=k_n+1}^n \frac{1}{i^2} &= \sqrt{k_n} \sqrt{\frac{1}{k_n} + O\left(\frac{1}{k_n^2}\right) + O\left(\frac{1}{n}\right)} = 1 + O\left(\frac{1}{k_n}\right) + O\left(\frac{k_n}{n}\right). \end{aligned}$$

Combining the previous expansions yields the expected result (20). To conclude, since the Hill estimator is scale invariant, Equation (19) also holds if  $X_i^{(2)} \sim \lambda \mathcal{P}(1/\xi)$ , for all  $\lambda > 0$  and  $i = 1, \dots, n$ . By noticing that, in this case,  $\widehat{q}_2(\beta_n)/(\lambda q_1(\beta_n)) \stackrel{d}{=} \widehat{q}_1(\beta_n)/q_1(\beta_n)$ , Equation (20) holds true as well.  $\square$

## A.2 Proofs of main results

*Proof of Proposition 2.1.* The representation follows from the usual way of decomposing sums of squares in ANOVA-like analyses:

$$\begin{aligned} \Delta_n &=: \Delta_{1,n} + \Delta_{2,n} + \Delta_{3,n} \\ &= \frac{1}{JL} \sum_{j=1}^J \sum_{\ell=1}^L \left( \log \widetilde{q}_j(\alpha_{\ell,n}) + \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n} \right)^2 \\ &= \frac{1}{JL} \sum_{j=1}^J \sum_{\ell=1}^L \left( \log \widetilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right)^2 + \frac{1}{JL} \sum_{j=1}^J \sum_{\ell=1}^L \left( \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n} \right)^2 \\ &+ \frac{2}{JL} \sum_{j=1}^J \sum_{\ell=1}^L \left( \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n} \right) \left( \log \widetilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right). \end{aligned} \tag{21}$$

The interaction term in (21) satisfies  $\Delta_{3,n} = 0$ , which can be shown by switching the summations with respect to  $j$  and  $\ell$ :

$$\begin{aligned}\Delta_{3,n} &= \frac{2}{JL} \sum_{\ell=1}^L \sum_{j=1}^J \left( \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n} \right) \left( \log \tilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right) \\ &= \frac{2}{L} \sum_{\ell=1}^L \left( \mu_{\alpha,n}^{(\ell)} - \mu_{\alpha,n} \right) \frac{1}{J} \sum_{j=1}^J \left( \log \tilde{q}_j(\alpha_{\ell,n}) - \mu_{\alpha,n}^{(\ell)} \right) = 0,\end{aligned}$$

following the definition of  $\mu_{\alpha,n}^{(\ell)}$ .  $\square$

*Proof of Theorem 3.1.* First, (5) and Lemma A.2 entail

$$\begin{aligned}\Delta_{1,n} &= \frac{\xi^2}{L} \sum_{\ell=1}^L \bar{\sigma}_{\ell,n}^2 \left[ \frac{1}{J} \sum_{j=1}^J Z_j^2 - \left( \frac{1}{J} \sum_{k=1}^J Z_k \right)^2 \right] (1 + o_{\mathbb{P}}(1)) \\ &\stackrel{d}{=} \frac{\xi^2}{J} \left( \frac{1}{L} \sum_{\ell=1}^L \bar{\sigma}_{\ell,n}^2 \right) \chi_{J-1}^2 (1 + o_{\mathbb{P}}(1)) \\ &= \frac{\xi^2}{J} \bar{\sigma}_n^2 \chi_{J-1}^2 (1 + o_{\mathbb{P}}(1)),\end{aligned}$$

where  $Z_1, \dots, Z_J$  follow independent standard Gaussian distributions. Second, remark that (5) and  $\bar{\sigma}_{\ell,n} \rightarrow 0$  imply (16) so that Lemma A.3 yields  $\Delta_{2,n} = \xi^2 \text{varlog}((1 - \alpha_n)_{1:L}) (1 + o_{\mathbb{P}}(1))$ , and consequently,

$$\frac{J \text{varlog}((1 - \alpha_n)_{1:L})}{\bar{\sigma}_n^2} \frac{\Delta_{1,n}}{\Delta_{2,n}} \stackrel{d}{=} \chi_{J-1}^2 (1 + o_{\mathbb{P}}(1)),$$

which proves the result.  $\square$

*Proof of Corollary 3.1.* Under  $(H_0)$ , [de Haan and Ferreira, 2006, Theorem 4.3.8] entails the following representation for all  $j = 1, \dots, J$  and  $\ell = 1, \dots, L$ :

$$\log(\widehat{q}_j^{\text{W}}(\alpha_{\ell,n} | \beta_{j,n})) = \log(q_j(\alpha_{\ell,n})) + \frac{\log\left(\frac{1 - \beta_{j,n}}{1 - \alpha_{\ell,n}}\right)}{\sqrt{n_j(1 - \beta_{j,n})}} \xi_j Z_j (1 + o_{\mathbb{P}}(1)), \quad (22)$$

where  $Z_1, \dots, Z_J$  are independent standard Gaussian random variables. Moreover, taking account of  $\xi_j = \xi$  from Lemma A.1(i),  $n_j/n \rightarrow \lambda_j$  and  $(1 - \beta_{j,n})/(1 - \beta_n) \rightarrow 1/\lambda_j$  as  $n \rightarrow \infty$  for all  $j = 1, \dots, J$  yields

$$\log(\widehat{q}_j^{\text{W}}(\alpha_{\ell,n} | \beta_{j,n})) = \log(q_j(\alpha_{\ell,n})) + \frac{\log\left(\frac{1 - \beta_n}{1 - \alpha_{\ell,n}}\right)}{\sqrt{n(1 - \beta_n)}} \xi Z_j (1 + o_{\mathbb{P}}(1)),$$

since  $\log\left(\frac{1 - \beta_{j,n}}{1 - \alpha_{\ell,n}}\right) \sim \log\left(\frac{1 - \beta_n}{1 - \alpha_{\ell,n}}\right)$  as  $n \rightarrow \infty$ . As a conclusion, (5) holds with  $\bar{\sigma}_{\ell,n} = \frac{\log\left(\frac{1 - \beta_n}{1 - \alpha_{\ell,n}}\right)}{\sqrt{n(1 - \beta_n)}}$  and the result follows from Theorem 3.1.  $\square$

*Proof of Proposition 4.1.* In the case where  $J = 2$ ,  $\Delta_{1,n}$  can be simplified as

$$\Delta_{1,n} = \frac{1}{4L} \sum_{\ell=1}^L (\log \tilde{q}_1(\alpha_{\ell,n}) - \log \tilde{q}_2(\alpha_{\ell,n}))^2.$$

Using the Weissman estimator (2) and the Hill estimator (3), we have, for  $j = 1, 2$ :

$$\log \widehat{q}_j^{\text{W}}(\alpha_{\ell,n} | \beta_n) = \log q_j(\alpha_{\ell,n}) + \log \left( \frac{1 - \beta_n}{1 - \alpha_{\ell,n}} \right) \left( \widehat{\xi}_j^{\text{H}}(\beta_n) - \xi \right) + \log \left( \frac{\widehat{q}_j(\beta_n)}{q_j(\beta_n)} \right), \quad (23)$$

and therefore

$$\begin{aligned} \log \widehat{q}_1^{\text{W}}(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^{\text{W}}(\alpha_{\ell,n} | \beta_n) &= \log \left( \frac{1 - \beta_n}{1 - \alpha_{\ell,n}} \right) \left( \widehat{\xi}_1^{\text{H}}(\beta_n) - \widehat{\xi}_2^{\text{H}}(\beta_n) \right) \\ &+ \log \left( \frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right). \end{aligned} \quad (24)$$

Since  $X_i^{(1)}$  and  $X_i^{(2)}$  are Pareto distributed for  $i = 1, \dots, n$ , Equations (19) and (20) of Lemma A.4 yield

$$\widehat{\xi}_1^{\text{H}}(\beta_n) - \widehat{\xi}_2^{\text{H}}(\beta_n) \stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1 - \beta_n)}} \Gamma_\xi \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) \right), \quad (25)$$

$$\log \left( \frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right) \stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1 - \beta_n)}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) + O_{\mathbb{P}}(1 - \beta_n) \right), \quad (26)$$

where  $\Gamma_q$  and  $\Gamma_\xi$  are two independent standard Gaussian random variables. Hence, plugging (25) and (26) into (24), it follows that

$$\begin{aligned} \log \widehat{q}_1^{\text{W}}(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^{\text{W}}(\alpha_{\ell,n} | \beta_n) &\stackrel{d}{=} \frac{\log \left( \frac{1 - \beta_n}{1 - \alpha_{\ell,n}} \right)}{\sqrt{n(1 - \beta_n)}} \sqrt{2}\xi \Gamma_\xi \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) \right) \\ &+ \frac{\sqrt{2}\xi}{\sqrt{n(1 - \beta_n)}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) + O_{\mathbb{P}}(1 - \beta_n) \right). \end{aligned}$$

Taking account of  $\alpha_{\ell,n} = 1 - \tau_\ell/n$ , the above equality can be rewritten as

$$\begin{aligned} &\log \widehat{q}_1^{\text{W}}(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^{\text{W}}(\alpha_{\ell,n} | \beta_n) \\ &\stackrel{d}{=} \frac{1}{\sqrt{n(1 - \beta_n)}} \left\{ \log \left( \frac{n(1 - \beta_n)}{\tau_\ell} \right) \sqrt{2}\xi \Gamma_\xi \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) \right) + \sqrt{2}\xi \Gamma_q \right. \\ &\quad \left. + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) + O_{\mathbb{P}}(1 - \beta_n) \right\} \\ &\stackrel{d}{=} \frac{1}{\sqrt{n(1 - \beta_n)}} \left\{ \sqrt{1 + \log \left( \frac{n(1 - \beta_n)}{\tau_\ell} \right)^2} \sqrt{2}\xi \Gamma + O_{\mathbb{P}} \left( \frac{\log(n(1 - \beta_n))}{n(1 - \beta_n)} \right) + O_{\mathbb{P}}(1 - \beta_n) \right\}, \end{aligned}$$

where  $\Gamma$  is a standard Gaussian random variable, and consequently,

$$\begin{aligned} \Delta_{1,n} &\stackrel{d}{=} \frac{1}{4L} \sum_{\ell=1}^L \left[ \frac{\sqrt{2}\xi}{\sqrt{n(1 - \beta_n)}} \sqrt{1 + \log \left( \frac{n(1 - \beta_n)}{\tau_\ell} \right)^2} \Gamma \right]^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) + O_{\mathbb{P}} \left( \frac{1 - \beta_n}{\log(n(1 - \beta_n))} \right) \right) \\ &\stackrel{d}{=} \frac{\xi^2 \Gamma^2}{2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{n(1 - \beta_n)} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1 - \beta_n)} \right) + O_{\mathbb{P}} \left( \frac{1 - \beta_n}{\log(n(1 - \beta_n))} \right) \right). \end{aligned}$$

Similarly, one can easily prove, thanks to the choice  $\alpha_{\ell,n} = 1 - \tau_\ell/n$ , that

$$\begin{aligned}\Delta_{2,n} &= \frac{1}{L^2} \left[ L \sum_{\ell=1}^L \left( \frac{\log \widehat{q}_1^W(\alpha_{\ell,n} | \beta_n) + \log \widehat{q}_2^W(\alpha_{\ell,n} | \beta_n)}{2} \right)^2 - \left( \sum_{\ell=1}^L \frac{\log \widehat{q}_1^W(\alpha_{\ell,n} | \beta_n) + \log \widehat{q}_2^W(\alpha_{\ell,n} | \beta_n)}{2} \right)^2 \right] \\ &= \text{varlog}(\tau_{1:L}) \left( \frac{\widehat{\xi}_1^H(\beta_n) + \widehat{\xi}_2^H(\beta_n)}{2} \right)^2 \\ &= \xi^2 \text{varlog}(\tau_{1:L}) \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) \right).\end{aligned}\quad (27)$$

Combining the two previous results, we get the following first-order approximation of  $\Delta_{1,n}/\Delta_{2,n}$ :

$$\frac{\Delta_{1,n}}{\Delta_{2,n}} \stackrel{d}{=} \frac{\Gamma^2}{2 \text{varlog}(\tau_{1:L})} \frac{1 + S_n(\beta_n, \tau_{1:L})}{n(1-\beta_n)} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(n(1-\beta_n))} \right) \right),$$

and, in view of (9) the test statistic  $T_n$  becomes, under the condition  $(1-\beta_n) \log(n) \rightarrow 0$  as  $n \rightarrow \infty$ :

$$\begin{aligned}T_n &= \Gamma^2 \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right) \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(n(1-\beta_n))} \right) \right) \\ &=: \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right) \Gamma_n^2,\end{aligned}$$

where  $\Gamma_n \xrightarrow{d} \Gamma$ . Then,

$$\begin{aligned}p_n(\gamma) &:= \mathbb{P}_{H_0} (T_n > \chi_{1,1-\gamma}^2) \\ &= 2\mathbb{P}_{H_0} \left( \Gamma_n > \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right) \\ &= 2\bar{\Phi} \left( \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right)\end{aligned}\quad (28)$$

$$+ 2(\Phi - \Phi_n) \left( \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right),\quad (29)$$

where  $\Phi_n$  is the cumulative distribution function of  $\Gamma_n$ . Clearly,  $S_n(\beta_n, \tau_{1:L}) \rightarrow \infty$  as  $n \rightarrow \infty$  and thus (28) converges to  $\gamma$ . Moreover, Dini's Theorem states that  $\Phi_n$  converges uniformly to  $\Phi$  on compact sets, and thus (29) converges to 0. As a conclusion, one has

$$p_n(\gamma) = 2\bar{\Phi} \left( \bar{\Phi}^{-1}(\gamma/2) \left( 1 + \frac{1}{S_n(\beta_n, \tau_{1:L})} \right)^{-1/2} \right) (1 + o(1)),$$

and the result is proved.  $\square$

*Proof of Proposition 4.2.* As a consequence of (23) in the proof of Proposition 4.1, one has

$$\begin{aligned}\log \widehat{q}_1^W(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^W(\alpha_{\ell,n} | \beta_n) &= \log \left( \frac{q_1(\alpha_{\ell,n})}{q_2(\alpha_{\ell,n})} \right) + \log \left( \frac{1-\beta_n}{1-\alpha_{\ell,n}} \right) \left( \widehat{\xi}_1^H(\beta_n) - \widehat{\xi}_2^H(\beta_n) \right) \\ &\quad + \log \left( \frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right).\end{aligned}$$

Besides,  $q_2(\alpha_{\ell,n}) = \lambda_n q_1(\alpha_{\ell,n})$  and Equations (19), (20) in Lemma A.4 entail:

$$\begin{aligned}\widehat{\xi}_1^H(\beta_n) - \widehat{\xi}_2^H(\beta_n) &\stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_\xi \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1-\beta_n)} \right) \right), \\ \log \left( \frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right) &\stackrel{d}{=} \frac{\sqrt{2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{n(1-\beta_n)} \right) + O_{\mathbb{P}}(1-\beta_n) \right),\end{aligned}$$

where  $\Gamma_\xi$  and  $\Gamma_q$  are standard Gaussian random variables. Taking account of  $\alpha_{\ell,n} = 1 - \tau_\ell/n$  and introducing  $k_n = n(1 - \beta_n)$  to simplify the notations yields

$$\begin{aligned}&\log \widehat{q}_1^W(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^W(\alpha_{\ell,n} | \beta_n) \\ &= -\log(\lambda_n) + \frac{\sqrt{2}\xi \log(k_n/\tau_\ell)}{\sqrt{k_n}} \Gamma_\xi \left( 1 + O_{\mathbb{P}} \left( \frac{1}{k_n} \right) \right) + \frac{\sqrt{2}\xi}{\sqrt{k_n}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{k_n} \right) + O_{\mathbb{P}}(1-\beta_n) \right) \\ &= -\log(\lambda_n) + \frac{\sqrt{2}\xi \sqrt{1 + \log(k_n/\tau_\ell)^2}}{k_n} \Gamma \left( 1 + O_{\mathbb{P}} \left( \frac{1}{k_n} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(k_n)} \right) \right),\end{aligned}$$

where  $\Gamma$  is a standard Gaussian random variable. Then, recalling that

$$S_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log \left( \frac{k_n}{\tau_\ell} \right)^2 \quad \text{and} \quad s_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \sqrt{1 + \log \left( \frac{k_n}{\tau_\ell} \right)^2},$$

it follows:

$$\begin{aligned}\Delta_{1,n} &= \frac{1}{4L} \sum_{\ell=1}^L \left( \log \widehat{q}_1^W(\alpha_{\ell,n} | \beta_n) - \log \widehat{q}_2^W(\alpha_{\ell,n} | \beta_n) \right)^2 \\ &= \frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma \left( 1 + O_{\mathbb{P}} \left( \frac{1}{k_n} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(k_n)} \right) \right) \\ &\quad + \frac{\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))}{2 k_n} \Gamma^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1}{k_n} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(k_n)} \right) \right).\end{aligned}$$

Assumption (i) entails  $\log(k_n)^2/k_n^2 \vee \log(k_n)/n = o(\log(\lambda_n)^2)$  while  $\log(n)(1 - \beta_n) \rightarrow 0$  implies  $\log(k_n)/n = o(1/k_n)$ . Remarking that  $S_n(\beta_n, \tau_{1:L}) \sim \log(k_n)^2$  and  $s_n(\beta_n, \tau_{1:L}) \sim \log(k_n)$  as  $n \rightarrow \infty$  yields

$$\begin{aligned}\Delta_{1,n} &= \frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))}{2 k_n} \Gamma^2 \\ &\quad + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^2} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(\lambda_n) \log(k_n)}{k_n^{3/2}} \right) + O_{\mathbb{P}} \left( \log(\lambda_n) \sqrt{k_n} \right).\end{aligned}$$

Assumption (i) implies in particular  $\log(\lambda_n) = O(\log(k_n)/\sqrt{k_n})$  which, in turn, entails that the third and fourth  $O_{\mathbb{P}}(\cdot)$  are respectively bounded above by the first and second ones. The above expansion can thus be simplified as

$$\Delta_{1,n} = \frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))}{2 k_n} \Gamma^2 + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^2} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right).$$

Similarly to (27) in the proof of Proposition 4.1, one has

$$\Delta_{2,n} = \xi^2 \text{varlog}(\tau_{1:L}) \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right).$$

Combining the previous two results, and since  $(\log k_n)^2/k_n^{3/2} = o(\log(\lambda_n)^2)$ , in view of (i), we get the following asymptotic expansion of  $\Delta_{1,n}/\Delta_{2,n}$ :

$$\begin{aligned}\frac{\Delta_{1,n}}{\Delta_{2,n}} &= \frac{\log(\lambda_n)^2}{4\xi^2 \text{varlog}(\tau_{1:L})} \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right)\right) - \frac{1}{\sqrt{2}\xi} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n} \text{varlog}(\tau_{1:L})} \Gamma \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right)\right) \\ &+ \frac{1}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n \text{varlog}(\tau_{1:L})} \Gamma^2 \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right)\right) + O_{\mathbb{P}}\left(\frac{\log(k_n)^2}{k_n^2}\right) + O_{\mathbb{P}}\left(\frac{\log(k_n)}{n}\right) \\ &= \frac{\log(\lambda_n)^2}{4\xi^2 \text{varlog}(\tau_{1:L})} - \frac{1}{\sqrt{2}\xi} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n} \text{varlog}(\tau_{1:L})} \Gamma + \frac{1}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n \text{varlog}(\tau_{1:L})} \Gamma^2 \\ &+ O_{\mathbb{P}}\left(\frac{\log(k_n)^2}{k_n^{3/2}}\right) + O_{\mathbb{P}}\left(\frac{\log(k_n)}{n}\right) + O_{\mathbb{P}}\left(\frac{\log(\lambda_n)^2}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{\log(\lambda_n) \log(k_n)}{k_n}\right).\end{aligned}$$

The third and fourth  $O_{\mathbb{P}}(\cdot)$  are bounded above by the first one since  $\log(\lambda_n) = O(\log(k_n)/\sqrt{k_n})$  in view of condition (ii), and consequently,

$$\begin{aligned}\frac{\Delta_{1,n}}{\Delta_{2,n}} &= \frac{\log(\lambda_n)^2}{4\xi^2 \text{varlog}(\tau_{1:L})} - \frac{1}{\sqrt{2}\xi} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n} \text{varlog}(\tau_{1:L})} \Gamma + \frac{1}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n \text{varlog}(\tau_{1:L})} \Gamma^2 \\ &+ O_{\mathbb{P}}\left(\frac{\log(k_n)^2}{k_n^{3/2}}\right) + O_{\mathbb{P}}\left(\frac{\log(k_n)}{n}\right) \\ &= \frac{1}{\xi^2 \text{varlog}(\tau_{1:L})} \left(\frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n} \Gamma^2\right) \\ &\times \left(1 + \frac{O_{\mathbb{P}}\left(\frac{\log(k_n)^2}{k_n^{3/2}}\right) + O_{\mathbb{P}}\left(\frac{\log(k_n)}{n}\right)}{\frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n} \Gamma^2}\right).\end{aligned}$$

Focusing on the denominator of the above term, condition (ii) shows that  $\sqrt{k_n} \log(\lambda_n) \rightarrow 0$  as  $n \rightarrow \infty$  leading to

$$\frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n} \Gamma^2 = \frac{\xi^2}{2} \frac{\log(k_n)^2}{k_n} \Gamma^2 (1 + o_{\mathbb{P}}(1)),$$

and thus

$$\begin{aligned}\frac{\Delta_{1,n}}{\Delta_{2,n}} &= \frac{1}{\xi^2 \text{varlog}(\tau_{1:L})} \left(\frac{\log(\lambda_n)^2}{4} - \frac{\xi}{\sqrt{2}} \frac{\log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}} \Gamma + \frac{\xi^2}{2} \frac{(1 + S_n(\beta_n, \tau_{1:L}))}{k_n} \Gamma^2\right) \\ &\times \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right)\right),\end{aligned}$$

or equivalently,

$$\begin{aligned}T_n &= \frac{2 \text{varlog}(\tau_{1:L}) k_n}{S_n(\beta_n, \tau_{1:L})} \frac{\Delta_{1,n}}{\Delta_{2,n}} \\ &= \left(\frac{\log(\lambda_n)^2 k_n}{2\xi^2 S_n(\beta_n, \tau_{1:L})} - \frac{\sqrt{2k_n} \log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\xi S_n(\beta_n, \tau_{1:L})} \Gamma + \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma^2\right) \\ &\times \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right)\right).\end{aligned}$$

The result follows.  $\square$



*Proof of Remark 2.* From Proposition 4.2, one has  $\mathbb{P}(T_n \leq \chi_{1,1-\gamma}^2) = \mathbb{P}(a_n \Gamma^2 + b_n \Gamma + C_n \leq 0)$  where

$$a_n = \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})},$$

$$b_n = -\frac{\sqrt{2k_n} \log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\xi S_n(\beta_n, \tau_{1:L})}$$

are two non random sequences and

$$C_n = \frac{\log(\lambda_n)^2 k_n}{2\xi^2 S_n(\beta_n, \tau_{1:L})} - \chi_{1,1-\gamma}^2 (1 + \varepsilon_n),$$

$$\varepsilon_n = O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right)$$

are two random variables. It straightforwardly follows that

$$\Omega_{1,n} := -\frac{b_n}{2a_n} = \frac{\sqrt{k_n} \log(\lambda_n) s_n(\beta_n, \tau_{1:L})}{\sqrt{2}\xi (1 + S_n(\beta_n, \tau_{1:L}))} \quad \text{and} \quad \frac{b_n^2}{4a_n^2} = \frac{k_n \log(\lambda_n)^2 s_n(\beta_n, \tau_{1:L})^2}{2\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))^2}.$$

Moreover, in view of (i) and  $\log(n)(1 - \beta_n) \rightarrow 0$ , one has:

$$-\frac{C_n}{a_n} = \chi_{1,1-\gamma}^2 \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} - \frac{\log(\lambda_n)^2 k_n}{2\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))} + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right).$$

Besides, (i') is equivalent to  $\log(k_n)^2/k_n^{3/4} \vee \sqrt{\log(k_n)^3/n} = o(\log(\lambda_n))$  and thus

$$\begin{aligned} \frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n} &= \frac{k_n \log(\lambda_n)^2}{2\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))^2} (s_n(\beta_n, \tau_{1:L})^2 - 1 - S_n(\beta_n, \tau_{1:L})) \\ &\quad + \chi_{1,1-\gamma}^2 \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right) \\ &= \Omega_{2,n} \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right)\right), \end{aligned}$$

where

$$\Omega_{2,n} = \frac{\log(\lambda_n)^2 k_n s_n(\beta_n, \tau_{1:L})^2 - 1 - S_n(\beta_n, \tau_{1:L})}{2\xi^2 (1 + S_n(\beta_n, \tau_{1:L}))^2} + \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} \chi_{1,1-\gamma}^2.$$

A second-order Taylor expansion shows that  $s_n(\beta_n, \tau_{1:L})^2 - 1 - S_n(\beta_n, \tau_{1:L}) \rightarrow 0$  as  $n \rightarrow \infty$ . Besides, since  $S_n(\beta_n, \tau_{1:L}) \sim (\log k_n)^2$  as  $n \rightarrow \infty$ , it follows from (ii) that  $\Omega_{2,n} \rightarrow \chi_{1,1-\gamma}^2$  as  $n \rightarrow \infty$ . As a consequence,  $\Omega_n$  is positive for  $n$  large enough. Hence,

$$\begin{aligned} \mathbb{P}(T_n \leq \chi_{1,1-\gamma}^2) &= \mathbb{P}\left(\Gamma \in \left[-\frac{b_n}{2a_n} - \sqrt{\frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n}}, -\frac{b_n}{2a_n} + \sqrt{\frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n}}\right]\right) \\ &= \mathbb{P}\left(\Gamma \in \left[\Omega_{1,n} \pm \sqrt{\Omega_{2,n}} \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right)\right)\right]\right), \end{aligned}$$

and the proposed approximation follows.  $\square$

*Proof of Proposition 4.3.* Following the steps of the proof of Proposition 4.2, one has:

$$\begin{aligned} \log\left(\frac{\widehat{q}_1^W(\alpha_{\ell,n} | \beta_n)}{\widehat{q}_2^W(\alpha_{\ell,n} | \beta_n)}\right) &= \log\left(\frac{q_1(\alpha_{\ell,n})}{q_2(\alpha_{\ell,n})}\right) + \log\left(\frac{1 - \beta_n}{1 - \alpha_{\ell,n}}\right) \left(\widehat{\xi}_1^H(\beta_n) - \widehat{\xi}_2^H(\beta_n)\right) + \log\left(\frac{\widehat{q}_1(\beta_n) q_2(\beta_n)}{q_1(\beta_n) \widehat{q}_2(\beta_n)}\right) \\ &\quad + \xi \log\left(\frac{1 - \beta_n}{1 - \alpha_{\ell,n}}\right) (\theta_n - 1). \end{aligned} \tag{30}$$

First, we establish a result similar to those of Lemma A.4 adapted to our setting. Equations (17) and (18) yield:

$$\widehat{\xi}_2^H(\beta_n) \stackrel{d}{=} \frac{\theta_n}{k_n} \sum_{i=1}^{k_n} \mathcal{E}_i^{(2)} \stackrel{d}{=} \theta_n \xi + \frac{\theta_n \xi}{\sqrt{n(1-\beta_n)}} \Gamma_{\xi,2} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) \right),$$

since  $\{\mathcal{E}_1^{(2)}, \dots, \mathcal{E}_{k_n}^{(2)}\}$  are i.i.d. realisations of an exponential distribution with mean  $\xi$ . It thus comes

$$\widehat{\xi}_1^H(\beta_n) - \widehat{\xi}_2^H(\beta_n) \stackrel{d}{=} (1-\theta_n)\xi + \frac{\sqrt{1+\theta_n^2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_{\xi} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) \right).$$

Similarly,

$$\log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right) \stackrel{d}{=} \theta_n \left\{ \mathcal{E}_{n-k_n, n}^{(2)} - \xi \log(n/k_n) \right\} \stackrel{d}{=} \frac{\theta_n \xi}{\sqrt{n(1-\beta_n)}} \Gamma_{q,j} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}}(1-\beta_n) \right),$$

and one thus has

$$\log \left( \frac{\widehat{q}_1(\beta_n)}{q_1(\beta_n)} \right) - \log \left( \frac{\widehat{q}_2(\beta_n)}{q_2(\beta_n)} \right) \stackrel{d}{=} \frac{\sqrt{1+\theta_n^2}\xi}{\sqrt{n(1-\beta_n)}} \Gamma_q \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{n(1-\beta_n)}} \right) + O_{\mathbb{P}}(1-\beta_n) \right).$$

Moreover,  $q_2(\alpha_{\ell, n}) = q_1(\alpha_{\ell, n})^{\theta_n}$  and  $q_1(\alpha_{\ell, n}) = (\tau_{\ell}/n)^{-\xi}$ , hence

$$\log(q_1(\alpha_{\ell, n})/q_2(\alpha_{\ell, n})) = \xi(\theta_n - 1) \log(\tau_{\ell}/n)$$

and replacing in (30) yields

$$\log \left( \frac{\widehat{q}_1^W(\alpha_{\ell, n} | \beta_n)}{\widehat{q}_2^W(\alpha_{\ell, n} | \beta_n)} \right) = \xi(1-\theta_n) \log \left( \frac{n}{\tau_{\ell}} \right) + \frac{\sqrt{1 + \log \left( \frac{k_n}{\tau_{\ell}} \right)^2}}{\sqrt{k_n}} \xi \sqrt{1 + \theta_n^2} \Gamma \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) + O_{\mathbb{P}} \left( \frac{1-\beta_n}{\log(k_n)} \right) \right),$$

where we have introduced  $k_n = n(1-\beta_n)$ . Using (i), one has  $\log(k_n)/k_n^{3/4} \vee \sqrt{\log(k_n)/n} = o(\log(\theta_n))$  and taking account of  $\log(n)(1-\beta_n) \rightarrow 0$  as  $n \rightarrow \infty$ , it follows

$$\begin{aligned} \Delta_{1,n} &= \frac{\xi^2(1-\theta_n)^2 \text{smlog}(n/\tau_{1:L})}{4} + \frac{\sqrt{1+\theta_n^2}\xi^2(1-\theta_n)\mathfrak{s}_n(\beta_n, \tau_{1:L})\Gamma}{2\sqrt{k_n}} + \frac{(1+S_n(\beta_n, \tau_{1:L}))\xi^2(1+\theta_n^2)\Gamma^2}{4k_n} \\ &+ O_{\mathbb{P}} \left( \frac{\log(\theta_n) \log(n)^2}{k_n} \right) + O_{\mathbb{P}} \left( \log(\theta_n) \log(n) \sqrt{\frac{1-\beta_n}{n}} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^{3/2}} \right), \end{aligned}$$

where it is recalled that

$$\mathfrak{s}_n(\beta_n, \tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log \left( \frac{n}{\tau_{\ell}} \right) \sqrt{1 + \log \left( \frac{k_n}{\tau_{\ell}} \right)^2} \quad \text{and} \quad \text{smlog}(n/\tau_{1:L}) = \frac{1}{L} \sum_{\ell=1}^L \log \left( \frac{n}{\tau_{\ell}} \right)^2.$$

Note that the condition  $\log(\theta_n) = o(1/\sqrt{k_n})$  in (ii) ensures that the first and second  $O_{\mathbb{P}}$  are bounded from above by the two others, hence

$$\begin{aligned} \Delta_{1,n} &= \frac{\xi^2(1-\theta_n)^2 \text{smlog}(n/\tau_{1:L})}{4} + \frac{\sqrt{1+\theta_n^2}\xi^2(1-\theta_n)\mathfrak{s}_n(\beta_n, \tau_{1:L})\Gamma}{2\sqrt{k_n}} + \frac{(1+S_n(\beta_n, \tau_{1:L}))\xi^2(1+\theta_n^2)\Gamma^2}{4k_n} \\ &+ O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^{3/2}} \right), \end{aligned}$$

In addition, straightforward calculations lead to

$$\Delta_{2,n} = \text{varlog}(\tau_{1:L}) \left( \frac{\widehat{\xi}_1^H(\beta_n) + \widehat{\xi}_2^H(\beta_n)}{2} \right)^2 = \xi^2 \text{varlog}(\tau_{1:L}) \left( \frac{1 + \theta_n}{2} \right)^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right).$$

Combining the previous two results, the following asymptotic expansion of  $\Delta_{1,n}/\Delta_{2,n}$  follows:

$$\begin{aligned} \frac{\Delta_{1,n}}{\Delta_{2,n}} &= \frac{\text{smlog}(n/\tau_{1:L})(1 - \theta_n)^2}{\text{varlog}(\tau_{1:L})(1 + \theta_n)^2} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right) + \frac{2\sqrt{1 + \theta_n^2}(1 - \theta_n) \mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}(1 + \theta_n)^2 \text{varlog}(\tau_{1:L})} \Gamma \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right) \\ &+ \frac{1 + \theta_n^2}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{\text{varlog}(\tau_{1:L})k_n} \Gamma^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^{3/2}} \right). \end{aligned}$$

Condition (ii) implies  $\log(\theta_n)^2 \log(n)^2 / \sqrt{k_n} = o(\log(k_n)^2 / k_n^{3/2})$  leading to

$$\begin{aligned} \frac{\Delta_{1,n}}{\Delta_{2,n}} &= \frac{\text{smlog}(n/\tau_{1:L})(1 - \theta_n)^2}{\text{varlog}(\tau_{1:L})(1 + \theta_n)^2} + \frac{2\sqrt{1 + \theta_n^2}(1 - \theta_n) \mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}(1 + \theta_n)^2 \text{varlog}(\tau_{1:L})} \Gamma + \frac{1 + \theta_n^2}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{\text{varlog}(\tau_{1:L})k_n} \Gamma^2 \\ &+ O_{\mathbb{P}} \left( \frac{\log(\theta_n)^2 \log(n)^2}{\sqrt{k_n}} \right) + O_{\mathbb{P}} \left( \frac{\log(\theta_n) \log(n) \log(k_n)}{k_n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^{3/2}} \right) \\ &= \frac{\text{smlog}(n/\tau_{1:L})(1 - \theta_n)^2}{\text{varlog}(\tau_{1:L})(1 + \theta_n)^2} + \frac{2\sqrt{1 + \theta_n^2}(1 - \theta_n) \mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}(1 + \theta_n)^2 \text{varlog}(\tau_{1:L})} \Gamma + \frac{1 + \theta_n^2}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{\text{varlog}(\tau_{1:L})k_n} \Gamma^2 \\ &+ O_{\mathbb{P}} \left( \frac{\log(k_n)}{n} \right) + O_{\mathbb{P}} \left( \frac{\log(k_n)^2}{k_n^{3/2}} \right) \\ &= \frac{\text{smlog}(n/\tau_{1:L})(1 - \theta_n)^2}{\text{varlog}(\tau_{1:L})(1 + \theta_n)^2} + \frac{2\sqrt{1 + \theta_n^2}(1 - \theta_n) \mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{k_n}(1 + \theta_n)^2 \text{varlog}(\tau_{1:L})} \Gamma \\ &+ \frac{1 + \theta_n^2}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{\text{varlog}(\tau_{1:L})k_n} \Gamma^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1 - \beta_n}{\log(k_n)} \right) + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right). \end{aligned}$$

Therefore, the test statistics can be written as

$$\begin{aligned} T_n &= \frac{2 \text{varlog}(\tau_{1:L}) k_n}{S_n(\beta_n, \tau_{1:L})} \frac{\Delta_{1,n}}{\Delta_{2,n}} \\ &= \frac{2 \text{smlog}(n/\tau_{1:L}) k_n (1 - \theta_n)^2}{S_n(\beta_n, \tau_{1:L}) (1 + \theta_n)^2} + \frac{4 \mathfrak{s}_n(\beta_n, \tau_{1:L}) \sqrt{k_n} \sqrt{1 + \theta_n^2} (1 - \theta_n) \Gamma}{S_n(\beta_n, \tau_{1:L}) (1 + \theta_n)^2} \\ &+ \frac{2(1 + \theta_n^2)}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma^2 \left( 1 + O_{\mathbb{P}} \left( \frac{1 - \beta_n}{\log(k_n)} \right) + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right), \end{aligned}$$

or, under the condition  $\log(n)(1 - \beta_n) \rightarrow 0$ ,

$$\begin{aligned} T_n &= 2 \left( \frac{\text{smlog}(n/\tau_{1:L}) k_n (1 - \theta_n)^2}{S_n(\beta_n, \tau_{1:L}) (1 + \theta_n)^2} + 2 \frac{\mathfrak{s}_n(\beta_n, \tau_{1:L}) \sqrt{k_n} \sqrt{1 + \theta_n^2} (1 - \theta_n) \Gamma}{S_n(\beta_n, \tau_{1:L}) (1 + \theta_n)^2} + \frac{(1 + \theta_n^2)}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \Gamma^2 \right) \\ &\times \left( 1 + O_{\mathbb{P}} \left( \frac{1 - \beta_n}{\log(k_n)} \right) + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k_n}} \right) \right), \end{aligned}$$

and the result is proved.  $\square$

*Proof of Remark 3.* Let us introduce the two non-random sequences

$$\begin{aligned} a_n &= 2 \frac{(1 + \theta_n^2)}{(1 + \theta_n)^2} \frac{1 + S_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})}, \\ b_n &= 4 \sqrt{k_n} \frac{\mathfrak{s}_n(\beta_n, \tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \frac{\sqrt{1 + \theta_n^2} (1 - \theta_n)}{(1 + \theta_n)^2}, \end{aligned}$$

as well as the two random variables

$$C_n = 2k_n \frac{\text{smlog}(n/\tau_{1:L})}{S_n(\beta_n, \tau_{1:L})} \frac{(1 - \theta_n)^2}{(1 + \theta_n)^2} - \chi_{1,1-\gamma}^2 (1 + \varepsilon_n),$$

$$\varepsilon_n = O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right),$$

so that  $\mathbb{P}(T_n \leq \chi_{1,1-\gamma}^2) = \mathbb{P}(a_n \Gamma^2 + b_n \Gamma + C_n \leq 0)$ . First,

$$\Psi_{1,n} := -\frac{b_n}{2a_n} = \frac{\sqrt{k_n}(\theta_n - 1)\mathfrak{s}_n(\beta_n, \tau_{1:L})}{\sqrt{1 + \theta_n^2}(1 + S_n(\beta_n, \tau_{1:L}))} \quad \text{and} \quad \frac{b_n^2}{4a_n^2} = \frac{k_n(\theta_n - 1)^2 \mathfrak{s}_n^2(\beta_n, \tau_{1:L})}{(1 + \theta_n^2)(1 + S_n(\beta_n, \tau_{1:L}))^2},$$

while

$$-\frac{C_n}{a_n} = \frac{\chi_{1,1-\gamma}^2}{2} \frac{(1 + \theta_n)^2 S_n(\beta_n, \tau_{1:L})}{(1 + \theta_n^2)(1 + S_n(\beta_n, \tau_{1:L}))} - \frac{(\theta_n - 1)^2 k_n \text{smlog}(n/\tau_{1:L})}{(1 + \theta_n^2)(1 + S_n(\beta_n, \tau_{1:L}))} + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right),$$

hence, under (i'), one has  $\log(k_n)^2/k_n^{3/4} \vee \sqrt{\log(k_n)^3/n} = o(\log(\theta_n))$ , and therefore

$$\frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n} = \Psi_{2,n} \left( 1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(k_n)}\right) \right),$$

where

$$\Psi_{2,n} = \frac{(\theta_n - 1)^2 k_n \mathfrak{s}_n(\beta_n, \tau_{1:L})^2 - (1 + S_n(\beta_n, \tau_{1:L})) \text{smlog}(n/\tau_{1:L})}{(1 + \theta_n^2)(1 + S_n(\beta_n, \tau_{1:L}))^2} + \frac{(1 + \theta_n)^2}{(1 + \theta_n^2)} \frac{S_n(\beta_n, \tau_{1:L})}{1 + S_n(\beta_n, \tau_{1:L})} \frac{\chi_{1,1-\gamma}^2}{2}.$$

It appears that  $\Psi_{2,n} \rightarrow \chi_{1,1-\gamma}^2$  as  $n \rightarrow \infty$  under the condition  $\log(n) = O(\log(k_n))$  and thus  $\Psi_{2,n} > 0$  for  $n$  large enough. As a consequence, one has

$$\begin{aligned} \mathbb{P}(T_n \leq \chi_{1,1-\gamma}^2) &= \mathbb{P}\left(\Gamma \in \left[-\frac{b_n}{2a_n} - \sqrt{\frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n}}, -\frac{b_n}{2a_n} + \sqrt{\frac{b_n^2}{4a_n^2} - \frac{C_n}{a_n}}\right]\right) \\ &= \mathbb{P}\left(\Gamma \in \left[\Psi_{1,n} \pm \sqrt{\Psi_{2,n}} \left(1 + O_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1 - \beta_n)}}\right) + O_{\mathbb{P}}\left(\frac{1 - \beta_n}{\log(n(1 - \beta_n))}\right)\right)\right]\right), \end{aligned}$$

leading to the proposed approximation.  $\square$

## B Additional figures

This section gathers some additional figures related to the numerical experiments conducted in Sections 5 and 6.

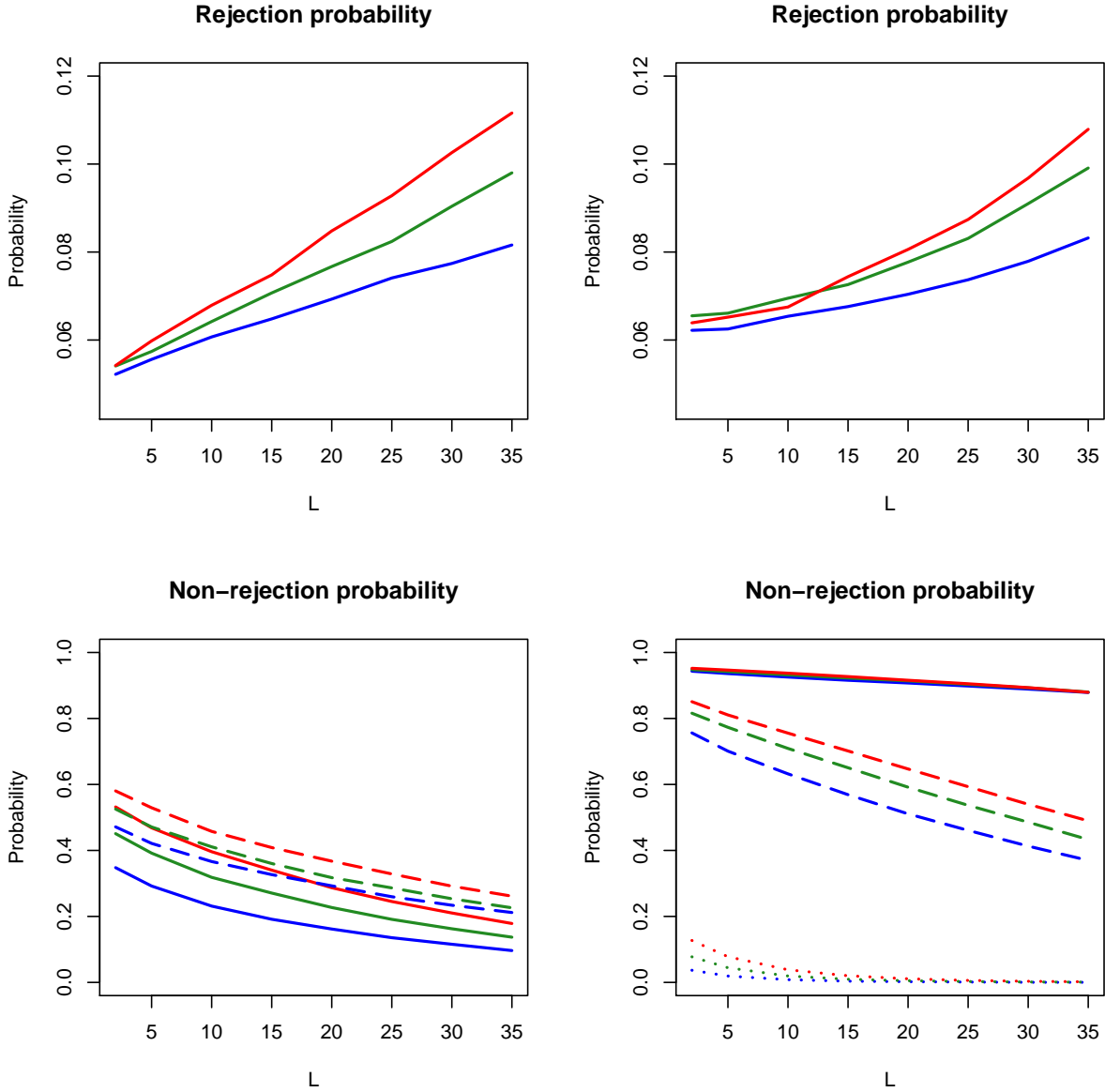


Figure B.1: From left to right, top to bottom: examples  $(\mathcal{MF})$ ,  $(\mathcal{MM})$ ,  $(\mathcal{MP})$  with  $\theta = 0.8$  (solid curves) and  $\theta = 1.2$  (dashed curves), and  $(\mathcal{CP})$  with  $\pi = 0.2$  (dotted curves),  $\pi = 0.5$  (dashed curves) and  $\pi = 0.8$  (solid curves). In all examples,  $n = 1,000$ ,  $\beta_n = 0.9$  and  $J = 5$  (blue curves),  $J = 10$  (green curves) and  $J = 15$  (red curves).

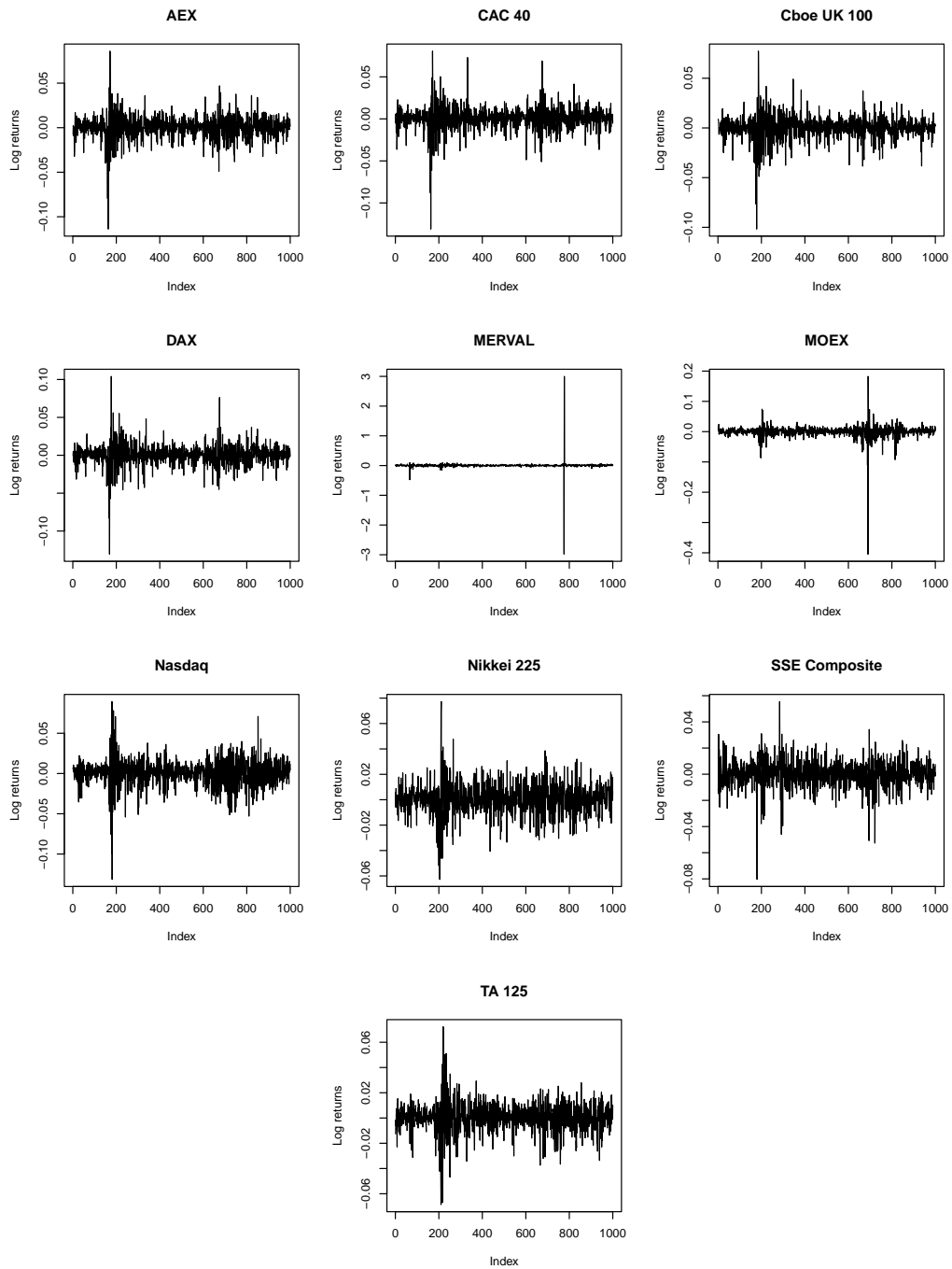


Figure B.2:  $n = 1,000$  log-returns associated with  $J = 10$  stock market indices.

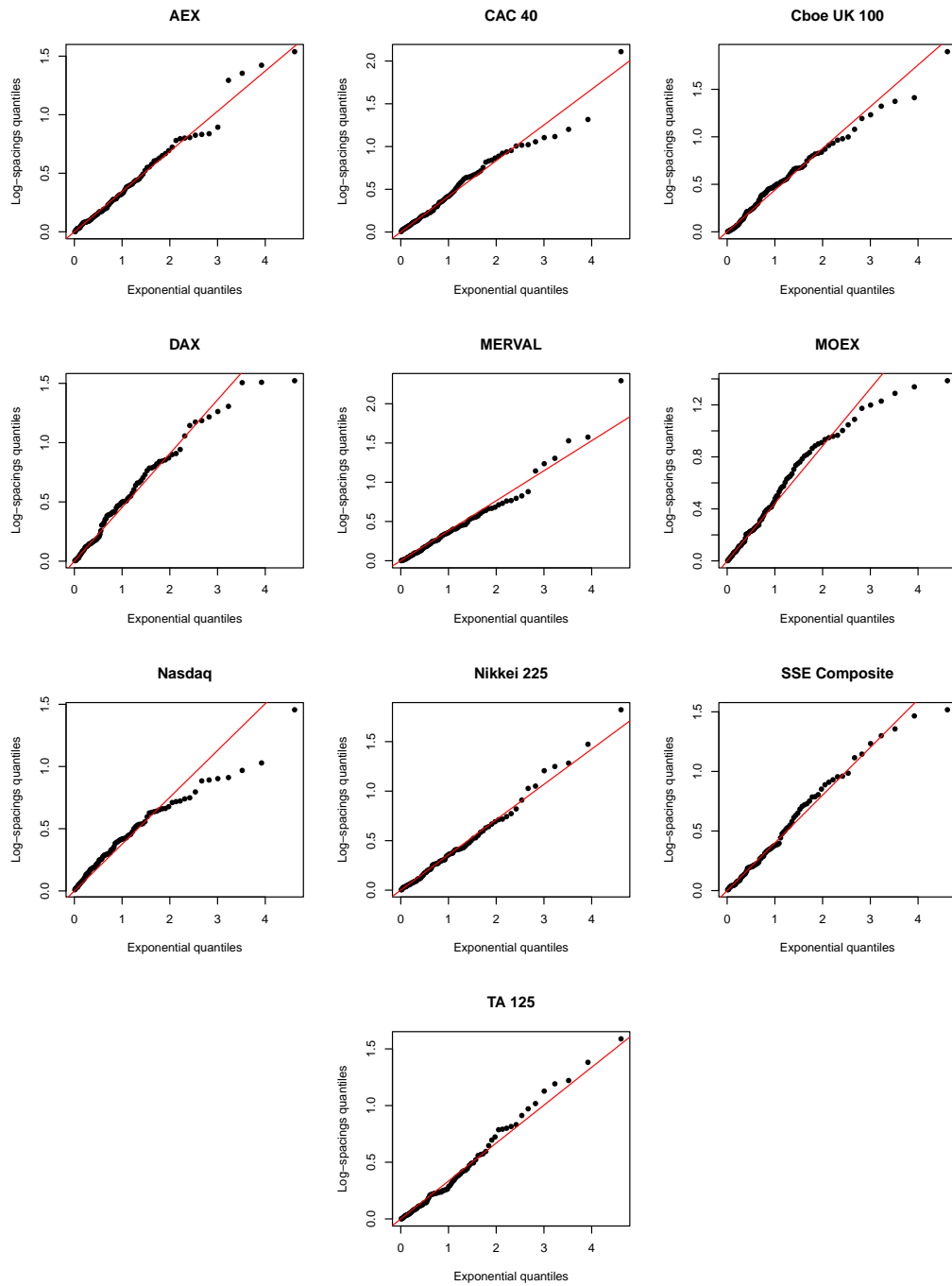


Figure B.3: Exponential QQ-plots of the weighted log-spacings associated with the daily log-returns computed from 10 financial indices.