Extremal covariance matrices

by

Youssouph Cissokho

Department of Mathematics and Statistics

Faculty of Science

University of Ottawa.

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfilment of the requirements

For the MSc degree in

Mathmatics and Statistics[1]

[1]The MSc program is a joint program with Carleton University, administered by the Ottawa- Carleton Institute of Mathematics and Statistics

Abstract

The tail dependence coefficient (TDC) is a natural tool to describe extremal dependence. Estimation of the tail dependence coefficient can be performed via empirical process theory. In case of extremal independence, the limit degenerates and hence one cannot construct a test for extremal independence. In order to deal with this issue, we consider an analog of the covariance matrix, namely the extremogram matrix, whose entries depend only on extremal observations. We show that under the null hypothesis of extremal independence and for finite dimension $d \geq 2$, the largest eigenvalue of the sample extremogram matrix converges to the maximum of d independent normal random variables. This allows us to conduct an hypothesis testing for extremal independence by means of the asymptotic distribution of the largest eigenvalue. Simulation studies are performed to further illustrate this approach.

Declaration

I, the undersigned, hereby declare that the work contained in this essay is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Acknowledgements

First and formost, I would like to thank God for his grace and mercy he has given me to be able to finish this work. I would like to express my deep gratitude to **Professor Rafal Kulik**, my research supervisor, for his valuable and constructive suggestions during the planning and development of this research work. His willingness to give his time so generously has been really appreciated. My grateful thanks are also extended to all my family members.

Contents

List of Figures vi

List of Tables viii

1 Introduction 1

2 Preliminaries 4

2.1 Regular variation . . . 4

2.1.1 Regularly varying functions . . . 4

2.1.2 Bounds and limits . . . 6

2.1.3 Regularly varying random variables . . . 7

2.1.4 Vague convergence . . . 9

2.1.5 Regularly varying random vectors . . . 10

2.1.6 Tail dependence coefficient (TDC) . . . 11

2.2 Weak convergence in metric spaces . . . 17

2.2.1 Lindeberg's condition . . . 18

2.2.2 Tightness via entropy . . . 19

2.2.3 Auxiliary results . . . 21

3 Weak convergence of the tail empirical process 22

3.1 Tail empirical process . . . 22

3.1.1 Proof of Theorem 3.1.3 . . . 24

4 Estimation and testing for extremal (In)dependence 28

4.1 Estimation of Tail Dependence Coefficient (TDC) . . . 29

4.2 Testing for Extremal (In)dependence using TDC . . . 33

4.3 Extremogram matrix: bivariate case . . . 35
4.3.1 Asymptotics for eigenvalues under extremal independence . . . 37
4.3.2 Testing for Extremal Independence using random matrices . . . 40
4.4 Extremogram matrix: d-dimensional case . . . 42
4.4.1 Testing for Extremal Independence using random matrices . . . 43
4.5 Tail empirical process with random levels . . . 44
4.6 Implementation: Simulation studies . . . 50
4.7 Real data analysis . . . 52

Conclusion and further work **59**

References **61**

List of Figures

2.1 Simulated samples of size 10000 of two i.i.d. Pareto random variables 13

2.2 Simulated samples of size 10000 of two dependent random variables 15

4.1 Asymtotic confidence intervals for the tail dependence coefficient for Example 4.1.3. Blue line: the estimator $\widetilde{\beta}_n^{(X,Y)}(1)$ plotted against different choices of the threshold u_n; red dashed lines - confidence interval using (4.19); yellow line - confidence interval using the true value of $\beta^{(X,Y)}(1)$. 34

4.2 The density function of $M(s = 1.2)$ (blue dashed line) and the density of the normal distribution with $\mu = \mathbb{E}(M(s))$ and $\sigma^2 = Var(M(s))$ (red line). 41

4.3 The scatter plot (left panel) and the estimator of the TDC (right panel) for X and Y drown independently from Pareto distribution with $\alpha = 4$. 51

4.4 The scatter plot (left panel) and the estimator of the TDC (right panel) for the model $Y = Z_2 + Z_3$, $X = Z_1 + Z_2$, where Z_i, $i = 1, 2, 3$ are independent Pareto with $\alpha = 4$. 52

4.5 The bivariate t distribution: the scatter plot (left panel) and the estimator of the TDC (right panel) plotted against the order statistics. 52

4.6 Time series plot of NASDAQ (left panel) and S&P500 (right panel). 53

4.7 QQ-plot of NASDAQ (left panel) and S&P500 (right panel). 54

4.8 The absolute log returns of daily stock prices for S&P500 vs. NASDAQ data set: the scatter plot (top left panel), Hill plot (top right panel) of NASDAQ (red dashed line) and S&P 500 (blue dashed line) with CI, the densities (bottom left panel) and the estimator of the TDC (bottom right panel). 55

4.9 Time series plot of German Mark (left panel) and French Franc (right panel). 56

4.10 QQ-plot of German Mark (left panel) and French Franc (right panel). 57

4.11 The absolute log returns of daily exchange-rates for German Mark and French Franc data set: the scatter plot (top left panel), Hill plot (top right panel) of German Mark (red line) and French Franc (blue line), the densities (bottom left panel) and the estimator of the TDC (bottom right panel). 58

List of Tables

4.1 Testing for extremal independence. The 95% asymptotic confidence intervals are CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.$ $\Big(-0.25, 0.56\Big)$, CI(1.3)=$\Big(M_{.025}(1.3), M_{.975}(1.3)\Big) = \Big(-0.23, 0.51\Big)$ and CI(1.5)=$\Big(M_{.025}(1.5), M_{.975}(1.5)\Big) = \Big(-0.16, 0.35\Big)$ for $s = 1.2, 1.3, 1.5$, $u_n = n^{\delta}$ and $\delta = .10$. The "+" sign indicates that T_1 belongs to the CI. 42

4.2 Testing for extremal independence. The 95% asymptotic confidence interval for $s = 1.2$ is CI=$\Big(M_{.025}, M_{.975}\Big) = \Big(-0.25; 0.56\Big)$. The "+" sign indicates that T_2 belongs to the CI and the numbers in parentheses correspond to the true TDC. 50

4.3 Testing for extremal independence. The 95% asymptotic confidence interval for $s = 1.3$ is CI=$\Big(M_{.025}, M_{.975}\Big) = \Big(-0.23, 0.51\Big)$. The "+" sign indicates that T_2 belongs to the CI and the numbers in parentheses correspond to the true TDC. 50

4.4 Testing for extremal independence for the stock prices data set. The 95% asymptotic confidence interval is CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.2)\Big) = \Big(-0.24; 0.53\Big)$. The "-" sign indicates that T_2 doesn't belong to the CI. 54

4.5 Testing for extremal independence for the exchange-rates data set. The 95% asymptotic confidence interval is CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.2)\Big) = \Big(-0.25; 0.55\Big)$. The "+" sign indicates that T_2 belongs to the CI. 55

Chapter 1

Introduction

The study of the behavior of extreme events becomes one of the priorities in recent years, because the occurence of these events may cause damage in many sectors (e.g. finance, economic, insurance, loss of human lives, etc.). For univariate random variables a good theoretical toll for the statistical modeling of such events is embedded in the framework of Extreme Value Theory (EVT) with the famous Fisher-Tippett-Gnedenko theorem that describes the limiting behavior for the properly normalized maxima. In a multivariate case, one of the many tools used to study the extreme events is the so-called **tail dependence coefficient**. If extremes in a multivariate vector do not occur together, then we have **extremal independence**. Otherwise, we have **extremal dependence**.

The main goal of this thesis is to test a null hypothesis for extremal independence by means of the tail dependence coefficient. To estimate the tail dependence coefficient, we use a nonparametric approach. It is more advantageous in a sense that it avoids any misidentification about the underlying distribution.

Related studies on the tail dependence coefficient.

Asymptotic properties of the estimators of the tail dependence coefficient are investigated and reviewed in [10]. The authors prove limit theorems for the proposed estimators under known and unknown marginal distributions. In our case we work under the assumption of regular variation and consider the estimator which is obtained from the definition of the TDC (2.4) by replacing the cdfs with their empirical versions. As such, we provide limit theorems under a weaker technical assumptions, but restricting a class of bivariate vectors. More specifically, our Theorem 4.1.1 corresponds to Theorem 5 in [10]. Similarly, our Theorem 4.5.3 corresponds Theorem 6 in [10]. The estimator used in [10] is based on the rank order statistic in contrast to our estimator used in Theorem 4.5.3 which is based on the intermediate order statistic. Our contribution,

as compared to [10], is a much simpler proof of the asymptotic normality. However, we do not claim any novelty here.

Testing for extremal independence.

A test for extremal independence is considered, where the null hypothesis is formulated as follows:

$$H_0\text{: } \textit{There is an extremal independence} \text{ vs.}$$

$$H_1\text{: } \textit{There is an extremal dependence.}$$

However, under the null hypothesis of extremal independence, the limit for the estimator of the tail dependence coefficient is degenerated and as such the estimator cannot be used directly to construct a test for extremal independence. This is the problem shared with virtually all estimators of the extremal independence - they degenerate under extremal independence. To overcome this problem, we introduce an analog of the covariance matrix, namely the **extremogram matrix**, whose entries depend only on the extremes. Random matrices, especially for high dimensional problems, became very popular in the last several years; see [1]. **We use random matrices, for the first time, in a novel context of extremal independence**. We study the asymptotic properties of the largest eigenvalue of the sample extremogram matrix. We show that under the null hypothesis of extremal independence, the largest sample eigenvalue converges (properly normalized) to the maximum of a finite number of normal random variables. As such, **for the first time, we can construct a proper test for extremal independence.**

Organisation of the thesis.

The purpose of Chapter 2 is to review some known results on regular variation, vague and weak convergence. This chapter is organized as follows, we start in Section 2.1 with a survey on regular variation. Important results and bounds, such as Karamata Theorem, Potter's bound, Breiman Lemma and the uniform convergence theorem, are presented. Thereafter, we introduce the concept of vague convergence and connect it to the notion of regularly varying random vectors. With the help of regular variation, we define the notion of the tail dependence coefficient.

In Section 2.2 we discuss weak convergence in metric spaces. Some important results like Lindeberg condition are presented. We also introduce the notion of tightness and uniform equicontinuity using a sophisticated entropy method.

The results in this chapter are not new and can be found in literature as mentioned in each section.

In Chapter 3, we discuss weak convergence of tail empirical processes of an i.i.d. regularly random vectors using the results developed in the previous chapter. The main result is Theorem 3.1.3. We do not claim originality of this result.

In Chapter 4, the nonparametric estimation of the tail dependence coefficient (TDC) and an hypothesis testing for extremal independence are developed. The advantage of the nonparametric approach is that it avoids any misidentification about the underlying distribution in contrast to the semiparametric or parametric approach. First, in Section 4.1, we discuss the estimation of the TDC by the standard approach, i.e. the estimator is obtained from the definition of the TDC where the cdfs are replaced with their empirical versions; see (4.1). We prove limit theorems for the estimator with deterministic levels. We do not claim the originality of this result. The limiting result can be used to construct confidence intervals for the tail dependence coefficient. However, under the null hypothesis of extremal independence, the limit is degenerated (see Corollary 4.1.2 and Section 4.2) and as such the estimator cannot be used to construct a test for extremal independence. This is the problem shared with virtually all estimators of the extremal (in)dependence - they degenerate under extremal independence. See also [10], Theorems 5 and 6.

To avoid this drawback, we consider an analog of the covariance matrix, namely the extremogram matrix, whose entries depend only on the extremes. Its sample counterpart is obtained by plugging-in the estimators of the tail dependence coefficient. We use random matrices, for the first time, in a novel context of extremal independence.

We work under the finite dimensional case, say $d = 2$ in Section 4.3, and an extension to arbitrary but finite dimension $d \geq 2$ in Section 4.4. In both cases, we prove that the largest eigenvalue of the sample extremogram matrix converges (after a proper normalization) in distribution to the maximum of a finite number of independent Gaussian random variables with explicit mean and variance. Having that in hand, we are now ready to conduct an hypothesis testing by means of the distribution of the largest eigenvalue of the sample extremogram matrix. Section 4.5 deals with the estimation of the tail dependence coefficient with data based, random levels. We obtain the same limit theorems as in the deterministic levels case. The transition to random levels follows the path as in [7]. However, the transition to random matrices is based on the original author's contribution.

Simulation studies are conducted in Section 4.6, while the real data analysis in Section 4.7.

Chapter 2

Preliminaries

A purpose of this chapter is to review some known results on regular variation, vague and weak convergence. This chapter is organized as follows, we start in Section 2.1 with a survey of regular variation. Important results and bounds, such as Karamata Theorem, Potter's bound, Breiman Lemma and the uniform convergence theorem, are presented. Thereafter, we introduce the concept of vague convergence and connect it to the notion of regularly varying random vectors. With the help of regular variation, we define the notion of the tail dependence coefficient.

In Section 2.2 we discuss weak convergence in a metric space. Some important results like Lindeberg condition are presented. We also introduce the notion of tightness and uniform equicontinuity using a sophisticated entropy method.

All the results in this chapter are not new and can be found in literature as indicated in each section.

2.1 Regular variation

In this section we recall the concept of regular variation of functions and random vectors and connect it to vague convergence. With help of regular variation, we will define the tail dependence coefficient. Several examples are given. The results presented in this section are known and can be found in e.g. [11] and [9]. We provide some proofs for completeness.

2.1.1 Regularly varying functions

We begin this section by introducing a basic definition of slowly varying functions followed by regularly varying functions and finally we end up with regularly varying random variables.

Let ℓ be a positive measurable function defined on $[0, \infty)$. Then ℓ is called *slowly varying* at infinity, denoted

by SV_∞, if

$$\lim_{x\to\infty} \frac{\ell(tx)}{\ell(x)} = 1,\ \forall\, t > 0\ .$$

A positive measurable function f defined on $[0,\infty)$ is said to be *regularly varying* at infinity with index $\alpha \in \mathbb{R}$ if $f(x) \neq 0$ for large x and

$$\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^\alpha,\ \forall\, t > 0\ .$$

We denote by $RV_\infty(\alpha)$ the set of regularly varying functions at infinity with index α. In fact, every regularly varying function can be expressed as

$$f(x) = x^\alpha \ell(x),$$

where ℓ is slowly varying.

Lemma 2.1.1 *The convergence in* (2.1.1) *is uniform. More specifically:*

(i) *if* $\alpha > 0$, *the convergence is uniform* $[0, b]$, $0 \leq b < \infty$;

(ii) *if* $\alpha < 0$, *the convergence is uniform on* $[b, \infty], 0 < b < \infty$;

(iii) *if* $\alpha = 0$, *the convergence is uniform on* $[a, b]$, $0 < a \leq b < \infty$.

Example 2.1.2 The function $f(x) = x^\delta$, for $\delta \in \mathbb{R}\setminus\{0\}$ is regularly varying at infinity with index δ. Indeed for all $x, t > 0$,

$$\frac{f(tx)}{f(x)} = \frac{(tx)^\delta}{x^\delta} = t^\delta.$$

Similarly, functions like $f(x) = x^\delta \log(x)$, $f_1(x) = x^\delta(\log(x))^\beta$ and $f_2(x) = x^\delta \log(\log(x))$, $\delta \neq 0, \beta \in \mathbb{R}$ are also regularly varying functions.

The powers of nonnegative regularly varying functions are also regularly varying, the product (resp. the composition) of two nonnegative regularly varying functions with indices α_1 and α_2 is regularly varying with index $\alpha_1 + \alpha_2$ (resp. $\alpha_1\alpha_2$). Also, the sum two nonnegative regularly varying functions is regularly varying with index $\max\{\alpha_1, \alpha_2\}$. In other words if f, f_1 and f_2 are positive measurable functions, then

- if $f_1 \in RV_\infty(\alpha_1)$, $f_2 \in RV_\infty(\alpha_2) \Rightarrow$
$$\begin{cases} f_1 f_2 \in RV_\infty(\alpha_1 + \alpha_2). \\ f_1 \circ f_2 \in RV_\infty(\alpha_1\alpha_2) \quad (if \quad \lim_{x\to\infty} f_2(x) = \infty). \\ f_1 + f_2 \in RV_\infty(\max\{\alpha_1, \alpha_2\}). \end{cases}$$

- if $f \in RV_\infty(\alpha) \Rightarrow \ \forall\, \delta \in \mathbb{R},\ f^\delta \in RV_\infty(\alpha\delta)$.
- if $f \in RV_\infty(\alpha)$, then for $\alpha \neq 0$, *as* $x \to \infty \Rightarrow$ $\begin{cases} f(x) \to \infty & \text{if } \alpha > 0, \\ f(x) \to 0 & \text{if } \alpha < 0. \end{cases}$

2.1.2 Bounds and limits

We briefly provide some useful and well known results such as Potter's bounds, or Karamata's theorem.

Lemma 2.1.3 *Suppose $g \in RV_\infty(\alpha)$ and $\ell \in SV_\infty$, locally bounded away from zero and ∞, $\alpha \in \mathbb{R} \setminus \{0\}$. Take $\delta \in (0,1)$. Then there exists t_0 such that for all $x > 0$ and $t \geq t_0$,*

$$(1-\delta)x^{\alpha-\delta} \leq \frac{g(tx)}{g(x)} \leq (1+\delta)x^{\alpha+\delta}$$

and

$$(1-\delta)x^{-\delta} \leq \frac{\ell(tx)}{\ell(x)} \leq (1+\delta)x^{\delta}.$$

In what follows we write $f(x) \sim g(x)$ whenever $\lim_{x\to\infty} f(x)/g(x) = 1$.

Lemma 2.1.4 (Karamata's theorem) *Let f be locally bounded on $[0,+\infty)$, positive and regularly varying function with index $\alpha \in \mathbb{R}$ and let $\gamma \in \mathbb{R}$.*

(1) $\alpha + \gamma > -1$, then

$$\int_1^\infty t^\gamma f(t)\,dt = \infty \ \text{ and } \ \int_1^x t^\gamma f(t)\,dt \sim (\gamma+\alpha+1)^{-1}x^{\gamma+1}f(x).$$

(2) If $\alpha + \gamma < -1$, then

$$\int_1^\infty t^\gamma f(t)\,dt < \infty \ \text{ and } \ \int_x^\infty t^\gamma f(t)\,dt \sim -(\gamma+\alpha+1)^{-1}x^{\gamma+1}f(x).$$

(3) If $\alpha + \gamma = -1$, then $\ell(x) = \int_1^x t^\gamma |f(t)|\,dt$ is slowly varying and $x^\gamma f(x) = o(\ell(x))$.

Proof: Note that the proof is not original, but we provide it in order to present the techniques.

Since $f \in RV_\infty(\alpha)$, then there exists a slowly varying function ℓ such that $f(x) = x^\alpha \ell(x)$. Assume that $\alpha + \gamma > -1$. Then

$$\frac{1}{x^{\gamma+1}f(x)}\int_1^x t^\gamma f(t)\,dt = \int_{1/x}^1 s^\gamma \frac{f(sx)}{f(x)}\,ds = \int_{1/x}^1 \frac{g(sx)}{g(x)}\,ds,$$

where $g(x) = x^\gamma f(x)$. The function g is regularly varying with index $\gamma + \alpha$. If $\alpha + \gamma > 0$, then by Lemma 2.1.1(i),

$$\frac{g(sx)}{g(x)} \to s^{\gamma+\alpha}, \text{ uniformly on } [0,1].$$

Hence, by the uniform convergence theorem, the above integral converges to the desired result. If $\alpha + \gamma \in (-1, 0]$, we cannot conclude the result directly, since the convergence $g(sx)/g(x)$ is uniform on $[a, 1]$, for $a > 0$. There exists $\epsilon > 0$ small enough so that $\alpha + \gamma + 1 - \epsilon > 0$, we have

$$\int_{1/x}^1 s^\gamma \frac{f(sx)}{f(x)}\, ds = \int_{1/x}^1 s^{\epsilon-1} s^{\gamma+1-\epsilon} \frac{f(sx)}{f(x)}\, ds = \int_{1/x}^1 s^{\epsilon-1} \frac{g(sx)}{g(x)} ds,$$

where $g(x) = x^{\gamma+1-\epsilon} f(x) \in RV_\infty(\alpha + \gamma + 1 - \epsilon)$. Now we have

$$g(sx)/g(x) \to s^{\alpha+\gamma+1-\epsilon}, \text{ uniformly on } [0,1].$$

Therefore, (1) is proven. The proof of (2) is similar.

□

2.1.3 Regularly varying random variables

In this section we apply the concept of regular variation to random variables and their distributions.

Definition 2.1.5 *A cdf (cumulative distribution function) (resp. right tail distribution) of a random variable X denoted by F_X (resp. $\overline{F}_X$) is defined as*

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \le x), \quad \forall\, x \in \mathbb{R}. \\ \overline{F}_X(x) &= 1 - F_X(x) = \mathbb{P}(X > x), \quad \forall\, x \in \mathbb{R}. \end{aligned}$$

Note that $\overline{F}_X$ is also called the survival function. In what follows, we may use F or F_X accordingly.

Definition 2.1.6 *A nonnegative random variable X is said to be regularly varying at infinity with index α if and only if its tail distribution function is regularly varying with index $-\alpha$ ($\alpha > 0$), that is*

$$\lim_{x\to\infty} \frac{\overline{F}_X(tx)}{\overline{F}_X(x)} = \frac{\mathbb{P}(X > tx)}{\mathbb{P}(X > x)} = t^{-\alpha}.$$

The parameter α is called the tail index. It measures the heaviness of the tail X. We will write $X \in RV_\infty(\alpha)$

or $\overline{F}_X \in RV_\infty(-\alpha)$.

We state important properties of regularly varying random variables. Some of these properties parallel the results in Section 2.1.2.

Lemma 2.1.7 *Let X be a nonnegative random variable with distribution function F. If $\overline{F} \in RV_\infty(-\alpha)$, $\alpha > 0$. Then $E(X^\beta) < \infty$ if $\beta < \alpha$ and $E(X^\beta) = \infty$ if $\beta > \alpha$. Moreover,*

$$\begin{cases} \lim_{x\to\infty} \frac{\int_0^x u^\beta \, dF(u)}{x^\beta \overline{F}(x)} = \frac{\alpha}{\alpha-\beta} \quad if \quad \beta < \alpha \\ \\ \lim_{x\to\infty} \frac{\int_0^x u^\beta \, dF(u)}{x^\beta \overline{F}(x)} = \frac{\alpha}{\beta-\alpha} \quad if \quad \beta > \alpha. \end{cases}$$

Lemma 2.1.8 (Potter's bound) *Let X be a nonnegative random variable with distribution function F and $RV_\infty(\alpha)$. Then $\forall\, \epsilon > 0\ \exists\, C > 1$: $\forall\, x \geq 0\,, y \geq 0$,*

$$\frac{\overline{F}(y^{-1}x)}{\overline{F}(x)} = \frac{\mathbb{P}(yX > x)}{\mathbb{P}(X > x)} \leq C \max(1, y)^{\epsilon+\alpha}\ .$$

Lemma 2.1.9 (Breiman's Lemma) *Assume that X and Y are two independent nonnegative random variables. If $X \in RV_\infty(\alpha)$, $\alpha > 0$ and $E(Y^{\alpha+\epsilon}) < \infty$ for some $\epsilon > 0$, then $XY \in RV_\infty(\alpha)$ and*

$$\lim_{x\to\infty} \frac{\mathbb{P}(YX > x)}{\mathbb{P}(X > x)} = E(Y^\alpha).$$

In what follows, we define a quantile function. It will play an important role in the thesis. We first provide its definition, followed by the transfer of regular variation.

Definition 2.1.10 *We define the quantile function of a random variable X by*

$$F^{\leftarrow}(q) = \inf\{x \in \mathbb{R} :\ F(x) \geq q\}, \qquad q \in (0, 1),$$

with the convention that the infimum of a empty set is ∞.

Let recall some of properties of the quantile functions:

- $F^{\leftarrow}$ is non-decreasing on (0,1) and left continuous;
- $F^{\leftarrow}(F(x)) \leq x$ for any $x \in \mathbb{R}$;
- $F(F^{\leftarrow}(q)) \geq q$ for any $q \in (0, 1)$.

The following function will be of interest:

$$Q(u) = F^{\leftarrow}\left(1 - \frac{1}{u}\right), \ \ \forall u \geq 1.$$

For this function we have:

$$\overline{F} \in RV_{\infty}(-\alpha) \Rightarrow Q \in RV_{\infty}(\alpha^{-1}).$$

We will work with random vectors with regularly varying marginals. Then, several situations can happen according to different behaviour of the marginals:

1. $X \overset{d}{=} Y$, i.e. $\mathbb{P}(X > x) = \mathbb{P}(Y > x) = x^{-\alpha}\ell(x)$, for $\ell(x) \in SV_{\infty}$.
2. $\mathbb{P}(X > x) \sim \mathbb{P}(Y > x)$, i.e. $\mathbb{P}(X > x) = x^{-\alpha}\ell(x)$ and $\mathbb{P}(Y > x) = x^{-\alpha}\tilde{\ell}(x)$ with $\ell(x), \tilde{\ell}(x) \in SV_{\infty}$ and $\lim_{x\to\infty} \ell(x)/\tilde{\ell}(x) = 1$.
3. $\mathbb{P}(X > x) = x^{-\alpha}\ell(x)$ and $\mathbb{P}(Y > x) = x^{-\alpha}\tilde{\ell}(x)$ with $\ell(x), \tilde{\ell}(x) \in SV_{\infty}$ and $\lim_{x\to\infty} \ell(x)/\tilde{\ell}(x) = \infty$.
4. $\mathbb{P}(X > x) = x^{-\alpha}\ell(x)$ and $\mathbb{P}(Y > x) = x^{-\beta}\tilde{\ell}(x)$ with $\alpha < \beta$ and $\ell(x), \tilde{\ell}(x) \in SV_{\infty}$.

We note that in the last two situations X has a heavier tail than Y, that is $\mathbb{P}(X > x)/\mathbb{P}(Y > x) \to \infty$ as $x \to \infty$. We will not deal with such situations in the thesis, we will focus on the first two cases. We will write $X \overset{d}{\approx} Y$ to indicate that the tails of X and Y are asymptotically the same.

2.1.4 Vague convergence

The notion of vague convergence appears as a useful tool in studying regular variation, while weak convergence is a tool to study the limiting behavior of random variables. We summarize some useful definitions, results and properties of vague and weak convergence. Let E be a locally compact topological space with countable base; often it is safe to think of E as a finite dimensional Euclidean space or $\mathbb{R}$. Let ν be a measure on E. If $\nu(K) < \infty$ for all relatively compact sets $K \subseteq E$, then ν is called a Radon measure.

Definition 2.1.11 (Vague convergence) *Let $\boldsymbol{M}_+(E)$ be the set of all nonnegative Radon measures on E and $\boldsymbol{C}_K^+(E)$ be the set of all continuous functions $f : E \to \mathbb{R}_+$ with compact support. Let ν and ν_n be Radon measures on E. We say that the sequence ν_n converges vaguely to ν, written $\nu_n \xrightarrow{v} \nu$, if*

$$\int f \, d\nu_n \to \int f \, d\nu,$$

for all $f \in \boldsymbol{C}_K^+(E)$.

The following result gives a link between regular variation and vague convergence.

Proposition 2.1.12 *Assume that a nonnegative random variable X, with distribution function F, is regularly varying with index α. Then there exists a sequence of constants a_n such that, as $n \to \infty$, $n\overline{F}(a_n x) \to x^{-\alpha}$ and*

$$\nu_n(\cdot) := n\mathbb{P}(a_n^{-1} X \in \cdot) \xrightarrow{v} \nu(\cdot),$$

where $\nu(x, \infty] = x^{-\alpha}$ and the convergence holds in $\boldsymbol{M}_+((0, \infty])$, the set of all nonnegative Radon measures on $(0, \infty]$.

2.1.5 Regularly varying random vectors

We begin this section by introducing the notion of multivariate regularly varying random vectors and some of their properties.

Definition 2.1.13 *An $\mathbb{R}^d-$valued random vector $\mathbb{X}$ and its distribution are said to be regularly varying with index $\alpha > 0$, written as $\mathbb{X} \in RV_\infty(\alpha)$, if there exists a Radon measure ν on the Borel σ-field $\mathcal{B}(\overline{\mathbb{R}}_0^d)$ of $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \backslash \{0\}$ such that*

$$\frac{\mathbb{P}(x^{-1}\mathbb{X} \in \cdot)}{\mathbb{P}(\|\mathbb{X}\| > x)} \xrightarrow{v} \nu(\cdot), \text{ as } x \to \infty,$$

where $\|\cdot\|$ is a vector norm in $\mathbb{R}^d$. The limiting measure is homogeneous, $\nu(xC) = x^{-\alpha}\nu(C)$ for any $x > 0$ and all Borel sets $C \in \overline{R}_0^d$ bounded away from zero. This definition is equivalent to the sequential definition of regular variation: there exists a scaling sequence $c_n \to \infty$ such that

$$n\mathbb{P}(c_n^{-1}\mathbb{X} \in \cdot) \xrightarrow{v} \nu$$

on $\overline{\mathbb{R}}^d \backslash \{0\}$.

Example 2.1.14 If X and Y are i.i.d. nonnegative regularly varying random variables such that $X \overset{d}{\approx} Y$, then $\mathbb{X} = (X, Y)$ is a regularly varying vector. Indeed, if A is of the form $A = ([0, s] \times [0, t])^c \subset \mathbb{R}_+^2 \setminus \{(0, 0)\}$, then the limiting measure is given (up to a constant) by

$$\nu(A) = s^{-\alpha} + t^{-\alpha}.$$

Indeed,

$$\begin{aligned}\frac{\mathbb{P}((X,Y)\in xA)}{\mathbb{P}(X>x)} &= \frac{\mathbb{P}(X>sx \text{ or } Y>tx)}{\mathbb{P}(X>x)}\\ &= \frac{\mathbb{P}(X>sx)}{\mathbb{P}(X>x)}+\frac{\mathbb{P}(Y>tx)}{\mathbb{P}(X>x)}-\frac{\mathbb{P}(X>sx)}{\mathbb{P}(X>x)}\mathbb{P}(Y>tx).\end{aligned}$$

Since X and Y are $RV_\infty(\alpha)$, we have as $x\to\infty$,

$$\frac{\mathbb{P}(X>sx)}{\mathbb{P}(X>x)}\to s^{-\alpha}\,,\quad \frac{\mathbb{P}(Y>tx)}{\mathbb{P}(X>x)}\to t^{-\alpha}\,,\quad \frac{\mathbb{P}(X>sx)}{\mathbb{P}(X>x)}\mathbb{P}(Y>tx)\to 0.$$

Therefore

$$\frac{\mathbb{P}((X,Y)\in xA)}{\mathbb{P}(X>x)}\to \nu\big(A\big)=s^{-\alpha}+t^{-\alpha},\ \text{as } x\to\infty. \tag{2.1}$$

Hence, the result follows. The latter result means that the limiting measure is concentrated on the axis. In other words, X and Y cannot be big at the same time. Furthermore, for any norm $\|\cdot\|$, we have

$$\mathbb{P}(\|\mathbb{X}\|>x)\sim\mathbb{P}(X>x)\times\aleph,\ \text{ as } x\to\infty,$$

where $\aleph$ is any constant. Indeed, if e.g. $\|\mathbb{X}\|=\|\mathbb{X}\|_\infty=\max(X,Y)$, then

$$\mathbb{P}(\|\mathbb{X}\|>x)=\mathbb{P}(\max(X,Y)>x)=\mathbb{P}(X>x \text{ or } Y>x).$$

So dividing both sides of the above equation by $\mathbb{P}(X>x)$ gives us

$$\frac{\mathbb{P}(\|\mathbb{X}\|>x)}{\mathbb{P}(X>x)}=\frac{\mathbb{P}(X>x \text{ or } Y>x)}{\mathbb{P}(X>x)}\to 2,\ \text{ as } x\to\infty,$$

using (2.1) with $(s,t)=(1,1)$.

2.1.6 Tail dependence coefficient (TDC)

When dealing with joint extreme events, one of the most important tools used for this purpose is the notion of tail dependence.

Definition 2.1.15 *Let (X,Y) be a bivariate random vector with marginal distributions F_X and F_Y. We define the tail dependence coefficient (*TDC*) between X and Y by*

$$\lambda_{\text{TDC}}^{(X,Y)}=\lim_{p\to 0}\frac{\mathbb{P}(X>F_X^{\leftarrow}(1-p),Y>F_Y^{\leftarrow}(1-p))}{p}\in[0,1], \tag{2.2}$$

where $F_X^{\leftarrow}$ and $F_Y^{\leftarrow}$ denote the quantile functions of X and Y, respectively.

If there is no risk for misinterpretation, we will write λ_{TDC} instead of $\lambda_{\text{TDC}}^{(X,Y)}$. Note that λ_{TDC} plays an important role since it quantifies the extremal relation between the variables. In other words, the bigger λ_{TDC} is, the stronger is the extremal dependence. We say that:

- if $\lambda_{\text{TDC}} = 0$, then there is no extremal dependence between X and Y. We say that X and Y are **asymptotically (extremally) independent**;
- if $\lambda_{\text{TDC}} \in (0, 1]$, then there is extremal dependence between X and Y. We say that X and Y are **asymptotically (extremally) dependent**.

We note that (2.2) is equivalent to

$$\lambda_{\text{TDC}} = \lim_{p \to 0} \frac{1}{\mathbb{P}(U > 1-p)} \mathbb{P}(U > 1-p, V > 1-p), \tag{2.3}$$

where $U = F_X(X)$ and $V = F_Y(Y)$ are standard uniform random variables. If we assume further that X and Y are s.t. $X \overset{d}{\approx} Y$, then (2.3) can be rewritten as follows:

$$\lambda_{\text{TDC}} = \lim_{x \to \infty} \frac{\mathbb{P}(X > x, Y > x)}{\mathbb{P}(X > x)}. \tag{2.4}$$

Indeed, if we assume for simplicity that F_X and F_Y are continuous and strictly increasing, then we have

$$\begin{aligned}
\lambda_{\text{TDC}} &= \lim_{p \to 0} \frac{1}{\mathbb{P}(U > 1-p)} \mathbb{P}(U > 1-p, V > (1-p)) \\
&= \lim_{p \to 0} \frac{1}{\mathbb{P}(F_X^{\leftarrow}(U) > F_X^{\leftarrow}(1-p))} \mathbb{P}(F_X^{\leftarrow}(U) > F_X^{\leftarrow}(1-p), F_X^{\leftarrow}(V) > F_X^{\leftarrow}(1-p)) \\
&= \lim_{p \to 0} \frac{1}{\mathbb{P}(F_X^{\leftarrow}(U) > F_X^{\leftarrow}(1-p))} \mathbb{P}(F_X^{\leftarrow}(U) > F_X^{\leftarrow}(1-p), F_Y^{\leftarrow}(V) > F_Y^{\leftarrow}(1-p)) \qquad (X \overset{d}{\approx} Y) \\
&= \lim_{p \to 0} \frac{1}{\mathbb{P}(F_X^{\leftarrow}(U) > x_p)} \mathbb{P}(F_X^{\leftarrow}(U) > x_p, F_Y^{\leftarrow}(V) > x_p) \qquad (x_p = F_X^{\leftarrow}(1-p)) \\
&= \lim_{x \to \infty} \frac{1}{\mathbb{P}(X > x)} \mathbb{P}(X > x, Y > x).
\end{aligned}$$

The formula (2.4) also makes sense when X and Y are asymptotically the same. Therefore, in this thesis, we will work with (2.4).

Example 2.1.16 Assume that X and Y are i.i.d. nonnegative regularly varying random variables and $X \overset{d}{\approx} Y$. Then $\lambda_{\text{TDC}} = 0$. Indeed,

$$\lambda_{\mathrm{TDC}} = \lim_{x\to\infty} \frac{\mathbb{P}(X > x, Y > x)}{\mathbb{P}(X > x)} = \lim_{x\to\infty} \mathbb{P}(Y > x) = 0.$$

See the Figure 2.1 for an example where X and Y are independently drawn from Pareto distribution. We can see that the large observations do not occur jointly.

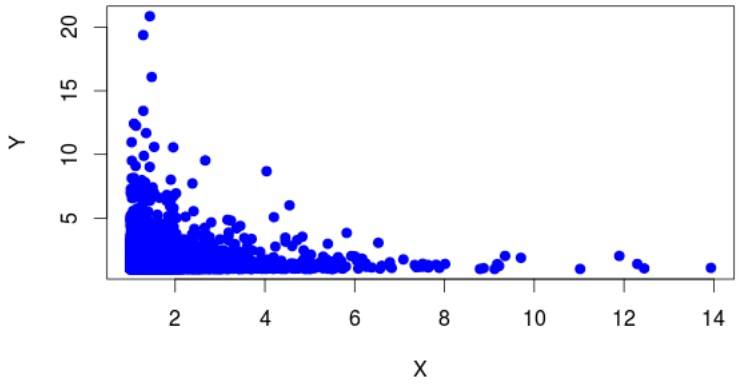

Figure 2.1: Simulated samples of size 10000 of two i.i.d. Pareto random variables

Example 2.1.17 Assume that $Z_i, i = 1, 2, 3$, are i.i.d. nonnegative regularly varying random variables with the same index α, i.e.

$$\mathbb{P}(Z_i > x) = x^{-\alpha}\ell(x),\ i = 1, 2, 3,$$

where $\ell(x)$ is a slowly varying function. Define $X = Z_1 + Z_2$ and $Y = Z_2 + Z_3$. Then $\lambda_{\mathrm{TDC}} = \frac{1}{2}$.
Note first that

$$\lim_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(Z_i > x)} = 2, \quad i = 1, 2, 3. \tag{2.5}$$

In fact, observe first that

$$\{Z_1 > x\} \cup \{Z_2 > x\} \subseteq \{Z_1 + Z_2 > x\}$$

thus

$$\frac{\mathbb{P}(Z_1 + Z_2 > x)}{\mathbb{P}(Z_1 > x)} \geq \frac{\mathbb{P}(Z_1 > x \text{ or } Z_2 > x)}{\mathbb{P}(Z_1 > x)} \xrightarrow{(2.1)} 2, \quad as \quad x \to \infty.$$

On the other hand, for $\epsilon \in (0, 1, 2)$, we have

$$\{Z_1 + Z_2 > x\} \subseteq \{Z_1 > x(1-\epsilon)\} \cup \{Z_2 > x(1-\epsilon)\} \cup \{Z_1 > x\epsilon, Z_2 > x\epsilon\}$$

therefore,

$$\frac{\mathbb{P}(Z_1 + Z_2 > x)}{\mathbb{P}(Z_1 > x)} \leq \frac{\mathbb{P}(Z_1 > x \text{ or } Z_2 > x \text{ or } \{Z_1 > x\epsilon, Z_2 > x\epsilon\})}{\mathbb{P}(Z_1 > x)} \xrightarrow{(2.1)} 2, \quad as \quad x \to \infty.$$

Leading to the desired result.

Coming back to our example, let $\delta \in (0,1)$, we have

$$\begin{aligned}
\frac{\mathbb{P}(X > x, Y > x)}{\mathbb{P}(X > x)} &= \frac{\mathbb{P}(Z_1 + Z_2 > x, Z_2 + Z_3 > x)}{\mathbb{P}(X > x)} \\
&= \frac{\mathbb{P}(Z_1 + Z_2 > x, Z_2 + Z_3 > x, Z_2 > \delta x)}{\mathbb{P}(X > x)} + \frac{\mathbb{P}(Z_1 + Z_2 > x, Z_2 + Z_3 > x, Z_2 \leq \delta x)}{\mathbb{P}(X > x)} \\
&\leq \frac{\mathbb{P}(Z_2 > \delta x)}{\mathbb{P}(X > x)} + \frac{\mathbb{P}(Z_1 > x(1-\delta))}{\mathbb{P}(X > x)} \mathbb{P}(Z_3 > x(1-\delta)) \\
&\leq \frac{\mathbb{P}(Z_2 > \delta x)}{\mathbb{P}(X > \delta x)} \frac{\mathbb{P}(X > \delta x)}{\mathbb{P}(X > x)} + \frac{\mathbb{P}(Z_1 > x(1-\delta))}{\mathbb{P}(X > x(1-\delta))} \frac{\mathbb{P}(X > x(1-\delta))}{\mathbb{P}(X > x)} \mathbb{P}(Z_3 > x(1-\delta)) \\
&\to \frac{1}{2}\delta^{-\alpha} + \frac{1}{2}(1-\delta)^{-\alpha} \times 0, \; as \; x \to \infty.
\end{aligned}$$

Hence

$$\lim_{\delta \to 1} \lim_{x \to \infty} \frac{\mathbb{P}(X > x, Y > x)}{\mathbb{P}(X > x)} \leq \frac{1}{2}\,. \tag{2.6}$$

On the other hand,

$$\frac{\mathbb{P}(X > x, Y > x)}{\mathbb{P}(X > x)} = \frac{\mathbb{P}(Z_1 + Z_2 > x, Z_2 + Z_3 > x)}{\mathbb{P}(X > x)} \geq \frac{\mathbb{P}(Z_2 > x)}{\mathbb{P}(X > x)} \xrightarrow{(2.5)} \frac{1}{2}, \quad as \quad x \to \infty. \tag{2.7}$$

Combining (2.6) and (2.7), we get $\lambda_{\text{TDC}} = \frac{1}{2}$. □

Figure 2.2 shows the graph of X and Y defined in this example, where Z_i are drown independently from Pareto distribution. We can see that some of the large values X and Y coincide.

Example 2.1.18 Let Z_1, Z_2 and V be independent nonnegative random variables such that $Z_1, Z_2 \in$

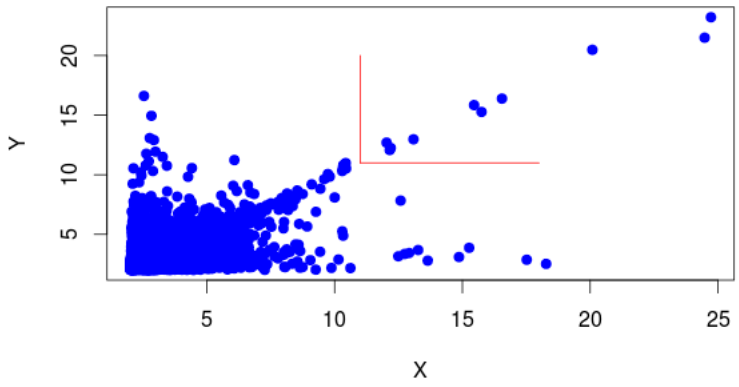

Figure 2.2: Simulated samples of size 10000 of two dependent random variables

$RV_\infty(\alpha)$, $Z_1 \stackrel{d}{=} Z_2$, $E(V^{\alpha+\epsilon}) < \infty$, for some $\epsilon > 0$. Define

$$X = VZ_1 \text{ and } Y = VZ_2.$$

Then the tail dependence coefficient vanishes, $\lambda_{\text{TDC}} = 0$, which just means that X and Y are dependent but asymptotically independent. In fact, note first that by Breiman's Lemma (Theorem 2.1.9), X and Y are $RV_\infty(\alpha)$:

$$\mathbb{P}(X > x) \sim E(V^\alpha)\mathbb{P}(Z_1 > x), \quad \mathbb{P}(Y > x) \sim E(V^\alpha)\mathbb{P}(Z_2 > x).$$

We have

$$\begin{aligned}
\lambda_{\text{TDC}} &= \lim_{x\to\infty} \frac{\mathbb{P}(X>x, Y>x)}{\mathbb{P}(X>x)} = \lim_{x\to\infty} \frac{\mathbb{P}(VZ_1>x, VZ_2>x)}{\mathbb{P}(X>x)} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(VZ_1>x, VZ_2>x)}{E(V^\alpha)\mathbb{P}(Z_1>x)} \\
&= \lim_{x\to\infty} \frac{E\big(\mathbb{P}(VZ_1>x, VZ_2>x)|V\big)}{E(V^\alpha)\mathbb{P}(Z_1>x)} \\
&= \lim_{x\to\infty} \frac{1}{E(V^\alpha)} E\left(\frac{\overline{F}_{Z_1}(\frac{x}{V})}{\overline{F}_{Z_1}(x)} \overline{F}_{Z_1}\left(\frac{x}{V}\right)\right) \\
&= \frac{1}{E(V^\alpha)} E\left(\lim_{x\to\infty} \frac{\overline{F}_{Z_1}(\frac{x}{V})}{\overline{F}_{Z_1}(x)} \lim_{x\to\infty} \overline{F}_{Z_1}\left(\frac{x}{V}\right)\right) \\
&= \frac{1}{E(V^\alpha)} E(V^\alpha \times 0) = 0.
\end{aligned}$$

Hence $\lambda_{\text{TDC}} = 0$.

Note that the argument used in order to exchange the limit and the expectation comes from the fact that Z_1, Z_2 are regular varying, and by application of Potter's bound (Theorem 2.1.8) and the dominated convergence theorem, using the assumption $E(V^{\alpha+\epsilon}) < \infty$. □

Example 2.1.19 Let Z_1, Z_2 and V be independent nonnegative random variables such that $Z_1, Z_2 \in RV_\infty(\alpha)$, $Z_1 \overset{d}{=} Z_2$, $V \in RV_\infty(\beta)$ with $\beta < \alpha$. This means that V has a heavier tail than Z_1 and Z_2, that is

$$\frac{\mathbb{P}(V>x)}{\mathbb{P}(Z_1>x)} \to \infty$$

and $E(V^{\alpha+\epsilon}) = \infty$, for all $\epsilon > 0$. We also note that $E(Z_1^{\beta+\epsilon}) < \infty$ for some $\epsilon > 0$. Define

$$X = VZ_1 \text{ and } Y = VZ_2.$$

Then the tail dependence coefficient $\lambda_{\text{TDC}} \neq 0$. By Breiman's Lemma

$$\mathbb{P}(X>x) \sim E(Z_1^\beta)\mathbb{P}(V>x), \quad \mathbb{P}(Y>x) \sim E(Z_2^\beta)\mathbb{P}(V>x).$$

Thus, we have

$$\begin{aligned}
\lambda_{\text{TDC}} &= \lim_{x\to\infty}\frac{\mathbb{P}(X>x,Y>x)}{\mathbb{P}(X>x)} = \lim_{x\to\infty}\frac{\mathbb{P}(VZ_1>x,VZ_2>x)}{\mathbb{P}(X>x)} \\
&= \lim_{x\to\infty}\frac{E\left(\mathbb{P}(VZ_1>x,VZ_2>x)|(Z_1,Z_2)\right)}{E(Z^\alpha)\mathbb{P}(V>x)} = \lim_{x\to\infty}\frac{E\big(\mathbb{P}(V>\frac{x}{Z_1},V>\frac{x}{Z_2})|(Z_1,Z_2)\big)}{E(Z^\alpha)\mathbb{P}(V>x)} \\
&= \lim_{x\to\infty}\frac{E\left(\mathbb{P}\big(V>x(\frac{1}{Z_1}\vee\frac{1}{Z_2})\big)|(Z_1,Z_2)\right)}{E(Z^\alpha)\mathbb{P}(V>x)} \\
&= \lim_{x\to\infty}\frac{1}{E(Z^\alpha)}E\left(\frac{\overline{F}_V(x\big(\frac{1}{Z_1}\vee\frac{1}{Z_2}\big))}{\overline{F}_V(x)}\right) = \frac{1}{E(Z^\alpha)}E\left(\lim_{x\to\infty}\frac{\overline{F}_V(x\big(\frac{1}{Z_1}\vee\frac{1}{Z_2}\big))}{\overline{F}_V(x)}\right) \qquad (C1) \\
&= \frac{1}{E(Z^\alpha)}E\left(\big(\frac{1}{Z_1}\vee\frac{1}{Z_2}\big)^{-\beta}\right) = \frac{1}{E(Z^\alpha)}E\big(\big(Z_1\wedge Z_2\big)^\beta\big)\neq 0.
\end{aligned}$$

Hence $\lambda_{\text{TDC}}\neq 0$, which implies asymptotic dependence between X and Y. Note that (C1) comes from the fact that

$$\frac{\overline{F}_V(x\big(\frac{1}{Z_1}\vee\frac{1}{Z_2}\big))}{\overline{F}_V(x)}\to\big(\frac{1}{Z_1}\vee\frac{1}{Z_2}\big)^{-\beta},\ x\to\infty,$$

and by application of Potter's bound and the dominated convergence theorem we can exchange the expectation with the limit, on account of $E(Z_1^{\beta+\epsilon})<\infty$.

Remark 2.1.20 From these couple of examples, we have learned that independence or dependence *does not* always imply extremal independence or dependence.

2.2 Weak convergence in metric spaces

In this section we discuss weak convergence in a metric space. Some important results like Lindeberg condition are presented. We also introduce the notion of tightness and uniform equicontinuity using a sophisticated entropy method. These results are known from the literature and can be found for example in [2], [12] and [9, p.57-100].

Let $(\mathbb{S},d)$ be a complete, separable metric space, equipped with the Borel σ-field $\mathfrak{F}$, generated by open sets.

Definition 2.2.1 (Weak convergence) *Let $\mathbb{P}_n$, $n\geq 0$, be a sequence of probability measures on $\mathbb{S}$. We*

say that $\mathbb{P}_n$ *converges weakly to* $\mathbb{P}$*, written* $\mathbb{P}_n \Rightarrow \mathbb{P}$*, if* $\mathbb{P}_n(f) \to \mathbb{P}(f)$*, for every* $f \in \boldsymbol{C}(\mathbb{S})$*, the set of bounded, continuous function* f *mapping from* $\mathbb{S}$ *to* $\mathbb{R}$*.*

ξ is called a random element of $\mathbb{S}$ if it is a measurable function from a probability space $(\Omega, \mathfrak{B}, \mathbb{P})$ into a metric space $(\mathbb{S}, \mathfrak{F}, d)$. Furthermore, if $\mathbb{S}$ equals $\mathbb{R}$, $\mathbf{C}$, $\mathbb{R}^n$, $n > 0$, then ξ is called a random variable, a random function and a random vector respectively.

Definition 2.2.2 (Convergence in distribution) *Let* ξ_n *be a sequence of random elements with values in a metric space* $(\mathbb{S}, \mathfrak{F})$*, defined on possibly different probability spaces* $(\Omega_n, \mathfrak{B}_n, \mathbb{P}_n)$*. We say that* ξ_n *converges in distribution to* ξ *on* $(\Omega, \mathfrak{B}, \mathbb{P})$*, if*

$$\mathbb{E}(f(\xi_n)) \to \mathbb{E}(f(\xi)),$$

for every $f \in \boldsymbol{C}(\mathbb{S})$*. We write* $\xi_n \xrightarrow{d} \xi$*.*

We note that the definitions above are virtually the same. The second one describes weak convergence of random elements ξ_n with laws $\mathbb{P}_n$.
The following theorem gives a characterization of weak convergence.

Theorem 2.2.3 (Skorokhod representation theorem) *Let* $\zeta_1, \zeta_2, ...$ *be a random elements in* $(\mathbb{S}, d)$ *such that* $\zeta_n \xrightarrow{d} \zeta$*. Then there exists a random elements* $\xi \stackrel{d}{=} \zeta$*,* $\xi_n \stackrel{d}{=} \zeta_n$*,* $n \in \mathbb{N}$ *such that* $\xi_n \xrightarrow{a.s} \xi$*, on a suitable probability space.*

Note that this theorem is interesting since it allows one to replace the convergence in distribution with the convergence almost surely.

2.2.1 Lindeberg's condition

Assume that for each n, the random elements $\xi_{n,1}, \ldots, \xi_{n,m_n}$, is independent. Suppose that

$$\mathbb{E}(\xi_{n,j}) = 0, \quad \mathbb{E}(\xi_{n,j}^2) = \sigma_{n,j}^2 < +\infty, \text{ and } s_n^2 = \sum_{j=1}^{m_n} \sigma_{n,j}^2.$$

If for any $\epsilon > 0$

$$\lim_{n \to +\infty} \sum_{j=1}^{m_n} \frac{1}{s_n^2} \mathbb{E}(|\xi_{n,j}|^2 \mathbb{I}_{\{|\xi_{n,j}| > \epsilon s_n\}}) = 0, \tag{2.8}$$

then

$$S_n / s_n = \sum_{j=1}^{m_n} \xi_{n,j} / s_n \Rightarrow N(0,1).$$

There is a stronger condition than Lindeberg's, called the Lyapunov condition, which says that if the moments of order $2+\delta$, for some $\delta > 0$ exists, then

$$\lim_{n\to+\infty} \sum_{j=1}^{m_n} E(|\xi_{n,j}|^{2+\delta}) = 0.$$

Remark 2.2.4 *If Lyapunov's condition is true, then so is Lindeberg's condition.*

In fact for any $\epsilon, \delta > 0$ such that $|Y| > \epsilon$, we have

$$Y^2 \mathbb{I}_{\{|Y|>\epsilon\}} \leq \frac{|Y|^{2+\delta}}{\epsilon^\delta} \Longrightarrow \mathbb{E}(Y^2 \mathbb{I}_{\{|Y|>\epsilon\}}) < \epsilon^{-\delta} \mathbb{E}(|Y|^{2+\delta}).$$

If $\mathbb{E}(|\xi_{n,j}|^{2+\delta}) < +\infty$, then the sum in (2.8) is at most

$$\frac{1}{\epsilon^\delta s_n^{2+\delta}} \sum_{j=1}^{m_n} \mathbb{E}(|\xi_{n,j}|^{2+\delta}).$$

2.2.2 Tightness via entropy

Entropy is a useful tool to prove tightness of a sequence of independent random processes. The method is "dimension free".

Let $\mathcal{G}$ be a class of functions. Let $\xi_n,\ n \geq 1$ be a sequence of random elements mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\ell^\infty(\mathcal{G})$, the set of bounded functions indexed by $\mathcal{G}$, endowed with the norm $||F||_{\mathcal{G}} = \sup_{g\in\mathcal{G}} |F(g)|$, where F is a map defined in $\ell^\infty(\mathcal{G})$.

Example 2.2.5 *Let $X_j,\ j \geq 1$ be a standard uniform i.i.d. random variables. Define*

$$\xi_n(t) = \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{n} \mathbb{I}_{\{X_j \leq t\}} - t \right).$$

Then ξ_n are random elements in $\ell^\infty(\mathcal{G})$, where $\mathcal{G}$ is the class of indicators of $[0,t]$, $t \in [0,1]$.

Definition 2.2.6 *Let ξ_n be a sequence of random maps indexed by a class of function $\mathcal{G}$, with values in $\ell^\infty(\mathcal{G})$ endowed with a semi-metric ρ induced with a supremum norm:*

- *The sequence is asymptotically tight if for each $\epsilon > 0$, there exists a compact set $K \subset \ell^\infty(\mathcal{G})$ such that*

$$\limsup_{n\to\infty} \mathbb{P}(\xi_n \notin K^\delta) < \epsilon, \quad \textit{for every } \ \delta > 0$$

where $K^\delta = \{g \in \ell^\infty(\mathcal{G}) : \rho(g, K) < \delta\}$.

- *The sequence is asymptotically uniformly* ρ*-equicontinuous (in probability) if for every* $\epsilon, \eta > 0$*, there exists* $\delta > 0$*, such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\sup_{g,f\in\mathcal{G}:\, \rho(g,f)<\delta} \left|\xi_n(f) - \xi_n(g)\right| > \epsilon \right) < \eta.$$

For the purpose of this thesis, these two concepts are in fact equivalent.

Proposition 2.2.7 *A sequence* ξ_n *is asymptotically tight if and only if* $\xi_n(g)$ *is tight in* $\mathbb{R}$ *for every* $g \in \mathcal{G}$ *and there exists a semi-metric* ρ *on* $\mathcal{G}$ *such that* $(\mathcal{G}, \rho)$ *is totally bounded and* ξ_n *is asymptotically uniformly* ρ*-equicontinuous (in probability).*

Let $(\xi_{n,j}(f),\ f \in \mathcal{G})$ be independent stochastic processes indexed by a common semi-metric space $(\mathcal{G}, \rho)$ Define the random semi-metric by

$$d_n(f, g) = \sum_{j=1}^{n} \Big(\xi_{n,j}(f) - \xi_{n,j}(g)\Big)^2.$$

The following theorem is a simplified version of Theorem 2.11.1 in [12].

Theorem 2.2.8 *For each* n*, let* $\xi_{n,1}, \ldots, \xi_{n,m_n}$ *be independent stochastic processes with finite second moments indexed by a totally bounded semi-metric space* $(\mathcal{G}, \rho)$*. Assume that*

(i) The maps

$$(x_1, \ldots, x_{m_n}) \to \sup_{\rho(f,g)<\delta} \Big| \sum_{j=1}^{m_n} e_j \big(\xi_{n,j}(f) - \xi_{n,j}(g)\big)^r \Big|,\ r = 1, 2,$$

are measurable for every $\delta > 0$*, every vector* $(e_1, \ldots, e_{m_n}) \in \{-1, 0, 1\}^{m_n}$ *and every* $n \in \mathbb{N}$.

(ii) For every $\epsilon > 0$,

$$\lim_{n\to+\infty} \sum_{j=1}^{m_n} \mathbb{E}\Big(\|\xi_{n,j}\|_{\mathcal{G}}^2 \mathbb{I}_{\{\|\xi_{n,j}\|_{\mathcal{G}} > \epsilon\}} \Big) = 0. \tag{2.9}$$

(iii) It holds

$$\lim_{\epsilon\to 0} \limsup_{n\to+\infty} \sup_{f,g:\rho(f,g)<\epsilon} \sum_{j=1}^{m_n} \mathbb{E}\Big(\xi_{n,j}(f) - \xi_{n,j}(g)\Big)^2 = 0. \tag{2.10}$$

(iv) For every $\epsilon > 0$,

$$\lim_{\delta \to 0} \lim_{n \to +\infty} \mathbb{P}\left(\int_0^{\delta} \sqrt{\log N(\epsilon', \mathcal{G}, d_n)}\, d\epsilon' > \epsilon \right) = 0, \tag{2.11}$$

where $N(\epsilon', \mathcal{G}, d_n)$ is the minimal number of balls $\{g : d_n(g, f) < \epsilon'\}$ of radius ϵ' needed to cover the set $\mathcal{G}$.

Then the sequence

$$\sum_{j=1}^{m_n} \Big(\xi_{n,j} - \mathbb{E}(\xi_{n,j}) \Big)$$

is asymptotically ρ-equicontinuous and hence tight.

The latter condition (condition (iv)) is not easy to check due to the random metric d_n. However, the spaces $\mathcal{G}$ we are going to consider are linearly ordered and then this condition is satisfied for free.

2.2.3 Auxiliary results

The following result describes convergence of inverses.

Lemma 2.2.9 (Vervaat lemma - convergence of inverses) *(i) Suppose that $\xi_0(\cdot)$ is a continuous function and $\xi_j(\cdot)$, $j \geq 1$ are a decreasing functions on $[a, b]$. Moreover, let g be defined on $[a, b]$ with a positive derivative g'. Let γ_n be a sequence of positive values such that $\gamma_n \to 0$ as n goes to infinity and*

$$\frac{\xi_n(s) - g(s)}{\gamma_n} \to \xi_0(s), \;\; n \to +\infty,$$

uniformly on $[a, b]$. Then

$$\frac{\xi_n^{\leftarrow}(s) - g^{\leftarrow}(s)}{\gamma_n} \to -(g^{\leftarrow})'(s)\xi_0(g^{\leftarrow}(s)), \;\; n \to +\infty,$$

uniformly on $[g(a), g(b)]$, where $g^{\leftarrow}$, $\xi_n^{\leftarrow}$ are inverse functions (right- or left-continuous or defined in any way consistent with the monotonicity).

(ii) If H_n, $n \geq 0$, are nondecreasing functions on $\mathbb{R}$ with range on $[a, d]$ such that $H_n \to H$, then $H_n^{\leftarrow} \to H^{\leftarrow}$.

Chapter 3

Weak convergence of the tail empirical process

In this section we discuss weak convergence of tail empirical processes of an i.i.d. regularly random vectors. The main result is Theorem 3.1.3. Although the result has not been stated in the literature in its form, we do not claim any originality, since it deals with i.i.d. random vectors. The proof is provided for completeness.

3.1 Tail empirical process

Let (X_j, Y_j), $j = 1 \dots, n$, be an i.i.d. sequence of random vectors sampled from (X, Y). We assume for simplicity that $X \overset{d}{\approx} Y$. Let F be the marginal distribution of X. Let u_n be a non decreasing sequence such that $u_n \to +\infty$ and $n\overline{F}(u_n) \to +\infty$. For $s_0 \in (0, 1)$, define

$$\widetilde{\beta}_n^{(X,Y)}(s) = \frac{1}{n\overline{F}(u_n)} \sum_{j=1}^{n} \mathbb{I}_{\{X_j > su_n, Y_j > su_n\}}, \ s \geq s_0,$$

which will be called **the tail empirical function**. Furthermore, define

$$\beta_n^{(X,Y)}(s) = \mathbb{E}\Big(\widetilde{\beta}_n^{(X,Y)}(s)\Big) = \frac{\mathbb{P}(X > su_n, Y > su_n)}{\mathbb{P}(X > u_n)} \tag{3.1}$$

and

$$\beta^{(X,Y)}(s) = \lim_{n \to +\infty} \beta_n^{(X,Y)}(s). \tag{3.2}$$

We assume that the limit in (3.2) exists which is certainly true under bivariate regular variation.

Remark 3.1.1 Note that if we let $s = 1$ in (3.2), then $\beta^{(X,Y)}(1) = \lambda_{\mathrm{TDC}}$, is the tail dependence coefficient between X and Y, defined in Section 2.1.6; cf. (2.4).

Definition 3.1.2 *We define the tail empirical processes*

$$G_n^{(X,Y)}(s) = \sqrt{n\overline{F}(u_n)} \left\{ \widetilde{\beta}_n^{(X,Y)}(s) - \beta_n^{(X,Y)}(s) \right\} \tag{3.3}$$

$$H_n^X(s) = G_n^{(X,X)}(s) = \sqrt{n\overline{F}(u_n)}(\widetilde{\delta}_n^X(s) - \delta_n^X(s)), \ s \geq s_0, \tag{3.4}$$

$$H_n^Y(s) = G_n^{(Y,Y)}(s) = \sqrt{n\overline{F}(u_n)}(\widetilde{\delta}_n^Y(s) - \delta_n^Y(s)), \ s \geq s_0,$$

where

$$\widetilde{\delta}_n^X(s) = \frac{1}{n\overline{F}(u_n)} \sum_{j=1}^n \mathbb{I}_{\{X_j > su_n\}}\ , \quad \widetilde{\delta}_n^Y(s) = \frac{1}{n\overline{F}(u_n)} \sum_{j=1}^n \mathbb{I}_{\{Y_j > su_n\}}, \tag{3.5}$$

and

$$\delta^Y(s) = \delta^X(s) = \lim_{n\to+\infty} \delta_n^X(s) = \lim_{n\to+\infty} \mathbb{E}[\widetilde{\delta}_n^X(s)] = s^{-\alpha}\ . \tag{3.6}$$

The main result of this section is the following weak convergence result. The result is not new, but we provide a proof for completeness.

Theorem 3.1.3 *Let (X_j, Y_j), $j = 1, ..., n$ be i.i.d. regularly varying random vectors such that $X_j \overset{d}{\approx} Y_j$. Let F be a distribution function of X_1. Let u_n be a non decreasing sequence such that $u_n \to +\infty$ and $n\overline{F}(u_n) \to +\infty$. Then*

$$G_n^{(X,Y)}(s) \Rightarrow G^{(X,Y)}(s), \ \ n \to +\infty, \tag{3.7}$$

$$H_n^X(s) \Rightarrow H^X(s), \quad H_n^Y(s) \Rightarrow H^Y(s), \qquad n \to +\infty, \tag{3.8}$$

in $\ell^\infty[s_0, +\infty)$ endowed with the sup-norm. The convergences hold jointly and $G^{(X,Y)}(\cdot)$, $H^Y(\cdot)$ and $H^X(\cdot)$ are Gaussian processes with covariance functions given respectively by

$$\boldsymbol{Cov}(G^{(X,Y)}(s_1), G^{(X,Y)}(s_2)) = \beta^{(X,Y)}(s_1 \vee s_2) \tag{3.9}$$

And

$$\boldsymbol{Cov}(H^Y(s), H^Y(t)) = \boldsymbol{Cov}(H^X(s), H^X(t)) = (s \vee t)^{-\alpha}. \tag{3.10}$$

3.1.1 Proof of Theorem 3.1.3

Fix $s_0 \in (0,1)$. Let $s \in [s_0, +\infty)$, define

$$K_{n,j}(s) = \frac{\sqrt{n\overline{F}(u_n)}}{n\overline{F}(u_n)} \left(\mathbb{I}_{\{X_j > su_n, Y_j > su_n\}} - \mathbb{P}(X_j > su_n, Y_j > su_n) \right), \ j = 1, \ldots, n.$$

Then we have

$$G_n^{(X,Y)}(s) = \sum_{j=1}^{n} K_{n,j}(s) = \sqrt{n\overline{F}(u_n)} \Big(\widetilde{\widetilde{\beta}}_n^{(X,Y)}(s) - \beta_n^{(X,Y)}(s) \Big).$$

The proof requires establishing the finite dimensional convergence and the tightness. For the former, we use the Lindeberg conditions. We first calculate the limiting covariance.

Lemma 3.1.4 *Under the conditions of Theorem 3.1.3,*

$$\lim_{n \to +\infty} \boldsymbol{Cov}(G_n^{(X,Y)}(s_1), G_n^{(X,Y)}(s_2)) = \beta^{(X,Y)}(s_1 \vee s_2), \tag{3.11}$$

$$\lim_{n \to +\infty} \boldsymbol{Var}(G_n^{(X,Y)}(s)) = \beta^{(X,Y)}(s). \tag{3.12}$$

Proof: We have

$$\begin{aligned}
&\mathbf{Cov}(G_n^{(X,Y)}(s_1), G_n^{(X,Y)}(s_2)) = \mathbf{Cov}\left(\sum_{j=1}^{n} K_{n,j}(s_1), \sum_{j=1}^{n} K_{n,j}(s_2) \right) \\
&= \mathbf{Cov}\left(\frac{1}{\sqrt{n\overline{F}(u_n)}} \sum_{j=1}^{n} \left(\mathbb{I}_{\{X_j > s_1u_n, Y_j > s_1u_n\}} - \mathbb{P}(X_j > s_1u_n, Y_j > s_1u_n) \right), \right. \\
&\quad \left. \frac{1}{\sqrt{n\overline{F}(u_n)}} \sum_{j=1}^{n} \left(\mathbb{I}_{\{X_j > s_2u_n, Y_j > s_2u_n\}} - \mathbb{P}(X_j > s_2u_n, Y_j > s_2u_n) \right) \right) \\
&= \frac{1}{\overline{F}(u_n)} \mathbb{E}\left(\mathbb{I}_{\{X_1 > s_1u_n, Y_1 > s_1u_n\}} \mathbb{I}_{\{X_j > s_2u_n, Y_j > s_2u_n\}} \right) - o(1) \\
&\sim \beta^{(X,Y)}(s_1 \vee s_2), \text{ as } n \to \infty,
\end{aligned}$$

where $o(1)$ stands for

$$\frac{\mathbb{P}(X_1 > s_1u_n, Y_1 > s_1u_n)}{\overline{F}(u_n)} \mathbb{P}(X_1 > s_2u_n, Y_1 > s_2u_n).$$

Hence (3.11) and (3.12) hold for all $s, s_1, s_2 \in [s_0, +\infty)$. □

The finite dimensional convergence is stated in Proposition 3.1.5.

Proposition 3.1.5 *Under the conditions of Theorem 3.1.3 we have for any finite set* $\{s_1, \cdots, s_k\} \in \mathbb{R}$,

$$\Big(G_n^{(X,Y)}(s_1), \ldots, G_n^{(X,Y)}(s_k)\Big) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

where the limiting covariance matrix is given by $\Sigma_{(k\times k)} = \big[\beta^{(X,Y)}(s_i \vee s_j)\big]_{i,j=1}^{k}$.
This result also implies the convergence of the margins, in distribution, to a centered normal random variable with limiting cavariances defined in Lemma 3.1.4.

Proof: Note first that $\mathbb{E}(K_{n,j}(s)) = 0$. Next, we show that, as $n \to +\infty$,

$$\sum_{j=1}^{n} \mathbb{E}\Big(K_{n,j}^2(s)\mathbb{I}_{\{|K_{n,j}(s)|>\epsilon\}}\Big) \to 0, \quad \forall \epsilon > 0.$$

By Hölder's inequality for each $j \geq 1$,

$$\begin{aligned}
\mathbb{E}\Big(K_{n,j}^2(s)\mathbb{I}_{\{|K_{n,j}(s)|>\epsilon\}}\Big) &\leq \mathbb{E}^{1/2}(K_{n,j}^4(s))\mathbb{E}^{1/2}\Big(\mathbb{I}_{\{|K_{n,j}(s)|>\epsilon\}}\Big) \\
&\leq \mathbb{E}^{1/2}(K_{n,j}^4(s))\mathbb{P}^{1/2}(|K_{n,j}(s)| > \epsilon).
\end{aligned}$$

Since

$$\begin{aligned}
\mathbb{E}(K_{n,j}^4(s)) &= \frac{1}{n^2\overline{F}^2(u_n)}\mathbb{E}\Big(\mathbb{I}_{\{X_j>su_n, Y_j>su_n\}} - \mathbb{P}(X_j > su_n, Y_j > su_n)\Big)^4 \\
&\leq C\frac{\mathbb{P}(X_j > su_n, Y_j > su_n)}{n^2\overline{F}^2(u_n)}, \quad \text{for some} \quad C > 0.
\end{aligned}$$

We have

$$\mathbb{E}^{1/2}(K_{n,j}^4(s)) \leq C^{1/2}\frac{\mathbb{P}^{1/2}(X_j > su_n, Y_j > su_n)}{n\overline{F}(u_n)}.$$

Also, by Markov inequality, we have,

$$\mathbb{P}^{1/2}(|K_{n,j}(s)| > \epsilon) \leq \frac{\mathbb{E}^{1/2}(K_{n,j}^4(s))}{\epsilon^2} \leq \frac{1}{\epsilon^2}C^{1/2}\frac{\mathbb{P}^{1/2}(X_j > su_n, Y_j > su_n)}{n\overline{F}(u_n)}.$$

This implies that

$$\sum_{j=1}^{n} \mathbb{E}\Big(K_{n,j}^2(s)\mathbb{I}_{\{X_j>su_n, Y_j>su_n\}}\Big) \leq \frac{C}{\epsilon^2 n\overline{F}(u_n)}\frac{\mathbb{P}(X_1 > su_n, Y_1 > su_n)}{\overline{F}(u_n)}.$$

Therefore,

$$\sum_{j=1}^{n}\mathbb{E}\Big(K_{n,j}^2(s)\mathbb{I}_{\{X_j>su_n,Y_j>su_n\}}\Big)\to 0,\quad as\ \ n\to+\infty.$$

Consequently, the Lindeberg condition holds. This, together with the Cramer-Wold device, finishes the proof.

□

Next, we prove tightness.

Proposition 3.1.6 *Under the conditions of Theorem 3.1.3, the sequence* $\{G_n^{(X,Y)}(\cdot), n\geq 1\}$ *is tight.*

Proof: By Theorem 2.2.8, it is enough to check (2.9) and (2.10). Consider the case of the class of the indicator functions

$$\mathcal{G}=\{f_s=\mathbb{I}_{\{(s,+\infty)\}},\ s\geq s_0\}.$$

Define for every $f_s\in\mathcal{G}$,

$$Z_{n,j}(f_s)=Z_{n,j}(s)=\frac{1}{\sqrt{n\overline{F}(u_n)}}\mathbb{I}_{\{X_j>su_n,Y_j>su_n\}},\quad K_{n,j}(s)=Z_{n,j}(s)-\mathbb{E}(Z_{n,j}(s))$$

and the semi-metric ρ by

$$\rho(f_s,f_t)=\rho(s,t)=3\Big|\beta^{(X,X)}(s)-\beta^{(X,X)}(t)\Big|+3\Big|\beta^{(Y,Y)}(s)-\beta^{(Y,Y)}(t)\Big|=6\Big|\beta^{(X,X)}(s)-\beta^{(X,X)}(t)\Big|,$$

where $\beta^{(X,X)}(s)$ and $\beta^{(Y,Y)}(s)$ are defined in (3.2). The convergence in (2.9) follows straightforwardly from the application of Lindeberg's condition as in the proof of Proposition 3.1.5, since $\|Z_{n,j}\|_{\mathcal{G}}^2=\frac{\mathbb{I}_{\{X_j>s_0u_n,Y_j>s_0u_n\}}}{n\overline{F}(u_n)}$.

For (2.10), note first that for $t>s\geq s_0$ we have

$$\Big|\mathbb{I}_{\{X_j>su_n,Y_j>su_n\}}-\mathbb{I}_{\{X_j>tu_n,Y_j>tu_n\}}\Big|\leq\mathbb{I}_{\{su_n<X_j<tu_n\}}+\mathbb{I}_{\{su_n<Y_j<tu_n\}}.$$

By the above inequality,

$$\begin{aligned}\sum_{j=1}^{n}\mathbb{E}\Big([Z_{n,j}(s)-Z_{n,j}(t)]^2\Big)&\leq\frac{1}{\overline{F}(u_n)}\mathbb{E}\Big([\mathbb{I}_{\{su_n<X_j<tu_n\}}+\mathbb{I}_{\{su_n<Y_j<tu_n\}}]^2\Big)\\&\leq 3\Big\{\beta_n^{(X,X)}(s)-\beta_n^{(X,X)}(t)+\beta_n^{(Y,Y)}(s)-\beta_n^{(Y,Y)}(t)\Big\}.\end{aligned}$$

Fix $\epsilon > 0$, then since as n goes infinity, $\beta_n^{(X,Y)}(s) \to \beta^{(X,Y)}(s)$,

$$\lim_{\delta \to 0} \limsup_{n \to +\infty} \sup_{\rho(s,t)<\delta} \sum_{j=1}^{n} \mathbb{E}[Z_{n,j}(s) - Z_{n,j}(t)]^2) \leq \epsilon + \lim_{\delta \to 0} \sup_{\rho(s,t)<\delta} \rho(s,t) < \epsilon$$

which leads to the desired result since ϵ is arbitrary. □

Chapter 4

Estimation and testing for extremal (In)dependence

This chapter deals with the nonparametric estimation of the tail dependence coefficient (TDC) and an hypothesis testing for extremal independence. The advantage of the nonparametric approach is that it avoids any misidentification about the underlying distribution in contrast to the semiparametric or parametric approach.

In Section 4.1, we discuss the estimation of the TDC by the standard approach, i.e. the estimator is obtained from the definition of the TDC where the cdfs are replaced with their empirical versions; see (4.1). We prove limit theorems for the estimator with deterministic levels. The limiting result can be used to construct confidence intervals for the tail dependence coefficient. However, under the null hypothesis of extremal independence, the limit is degenerated (see Corollary 4.1.2 and Section 4.2) and as such the estimator cannot be used to construct a test for extremal independence. This is the problem shared with virtually all estimators of the extremal (in)dependence - they degenerate under extremal independence. See [10], Theorems 5 and 6.

To avoid this drawback, we consider an analog of the covariance matrix, namely the extremogram matrix, whose entries depend only on the extremes, i.e. the tail dependence coefficient. Random matrices, especially for high dimensional problems, became very popular in the last several years; see [1]. We use them, for the first time, in a novel context of extremal independence.

We work under the finite dimensional case, say $d = 2$ in Section 4.3, and an extension to arbitrary but finite dimension $d \geq 2$ in Section 4.4. In both cases, we prove that the largest eigenvalue of its sample counterpart converges in distribution to the maximum of two independent Gaussian random variables with

explicit mean and variance. Having that in hand, we are now ready to conduct an hypothesis testing by means of the distribution of the largest eigenvalue of the sample extremogram matrix.

Estimators of the tail dependence coefficient with data based, random levels, are presented in Section 4.5. We obtain the same limit theorems as in the deterministic levels case. Simulation studies are conducted in Section 4.6. Note that the transition from deterministic levels to random levels follows the path as in [7]. However the extension to random matrices is a new idea and is the original author's contribution.

4.1 Estimation of Tail Dependence Coefficient (TDC)

Let (X_j, Y_j), $j = 1 \ldots, n$, be an i.i.d. sequence of regularly varying nonnegative random vectors sampled from (X, Y). Assume that $X \overset{d}{\approx} Y$. Recall that (cf. (3.1))

$$\beta_n^{(X,Y)}(s) = \frac{\mathbb{P}(X > su_n, Y > su_n)}{\mathbb{P}(X > u_n)} \quad \text{and} \quad \beta^{(X,Y)}(s) = \lim_{n\to+\infty} \beta_n^{(X,Y)}(s),$$

where $u_n \to \infty$ as $n \to \infty$. Our goal is to estimate $\beta^{(X,Y)}(s)$, using the empirical estimate defined as:

$$\widetilde{\beta}_n^{(X,Y)}(s) = \frac{\sum_{j=1}^n \mathbb{I}_{\{X_j > su_n, Y_j > su_n\}}}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}}. \tag{4.1}$$

From the theoretical point of view, the estimation result will be stated for $s \in [s_0, +\infty)$, $s_0 \in (0, 1)$ however, the main interest is the estimation when s is in the neighbourhood of 1. We first need to find the asymptotic behavior of

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)\Big). \tag{4.2}$$

To do so, the approach is to use the tail empirical process $G_n^{(X,Y)}(\cdot)$ defined in (3.3) and express the process in (4.2) as a function of $G_n^{(X,Y)}$. The following theorem is an immediate consequence of Theorem 3.1.3.

Theorem 4.1.1 *Assume that (X_j, Y_j), $j = 1, \ldots, n$ are i.i.d. regularly varying vectors of nonnegative random variables such that $X_j \overset{d}{\approx} Y_j$. Assume moreover that*

$$\lim_{n\to+\infty} \sqrt{n\overline{F}(u_n)} \sup_{s\geq s_0} \Big|\beta_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)\Big| = 0. \tag{4.3}$$

Then

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)\Big) \Rightarrow G_*^{(X,Y)}(s) := G^{(X,Y)}(s) - \beta^{(X,Y)}(s)H^X(1) \tag{4.4}$$

in $\ell^\infty[s_0, \infty)$.

Note that by Proposition 3.1.5, we have for any fixed $s \geq s_0,\ s_0 \in (0,1)$, the convergence in (4.4) holds in distribution and the limiting distribution is a centrered normal with variance defined in (4.13). This can be used to construct asymptotic confidence interval for the corresponding estimator.

Proof: We have,

$$\begin{aligned}
&\sqrt{n\overline{F}(u_n)}(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)) = \sqrt{n\overline{F}(u_n)}\left(\frac{\sum_{j=1}^n \mathbb{I}_{\{X_j > su_n, Y_j > su_n\}}}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} - \beta^{(X,Y)}(s)\right) \\
&= \sqrt{n\overline{F}(u_n)}\left(\frac{\sum_{j=1}^n \left(\mathbb{I}_{\{X_j > su_n, Y_j > su_n\}} - \mathbb{P}(X_1 > su_n, Y_1 > su_n)\right)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} + \frac{n\mathbb{P}(X_1 > su_n, Y_1 > su_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} - \beta^{(X,Y)}(s)\right) \\
&= \frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}}\left(\frac{1}{\sqrt{n\overline{F}(u_n)}}\sum_{j=1}^n \left(\mathbb{I}_{\{X_j > su_n, Y_j > su_n\}} - \mathbb{P}(X_1 > su_n, Y_1 > su_n)\right)\right) \\
&\quad + \sqrt{n\overline{F}(u_n)}\left(\frac{n\mathbb{P}(X_1 > su_n, Y_1 > su_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} - \beta^{(X,Y)}(s)\right) \\
&= \frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} G_n^{(X,Y)}(s) + \sqrt{n\overline{F}(u_n)}\left(\frac{n\mathbb{P}(X_1 > su_n, Y_1 > su_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} - \beta^{(X,Y)}(s)\right) \\
&= A_n G_n^{(X,Y)}(s) + B_n,
\end{aligned}$$

where

$$A_n = \frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} \quad \text{and} \quad B_n = \sqrt{n\overline{F}(u_n)}\left(\frac{n\mathbb{P}(X_1 > su_n, Y_1 > su_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}} - \beta^{(X,Y)}(s)\right).$$

Since $G_n^{(X,Y)}(s) \Rightarrow G^{(X,Y)}(s)$, we have

$$\frac{1}{n\overline{F}(u_n)}\left(\sum_{j=1}^n \left(\mathbb{I}_{\{X_j > su_n, Y_j > su_n\}} - \mathbb{P}(X_1 > su_n, Y_1 > su_n)\right)\right) \xrightarrow{\mathbb{P}} 0,$$

uniformly on compact subsets of $[s_0, +\infty)$, $s_0 \in (0,1)$ where $\xrightarrow{\mathbb{P}}$ means convergence in probability. This also implies

$$\frac{\sum_{j=1}^n \mathbb{I}_{\{X_j > su_n, Y_j > su_n\}}}{n\overline{F}(u_n)} \to \beta^{(X,Y)}(s), \tag{4.5}$$

in probability for each s. If we let $s = 1$ and replace Y by X, then (4.5) gives

$$\frac{\sum_{j=1}^n \mathbb{I}_{\{X_j > u_n\}}}{n\overline{F}(u_n)} \xrightarrow{\mathbb{P}} 1.$$

By the continuous mapping theorem, the reciprocal

$$A_n = \frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j>u_n\}}} \xrightarrow{\mathbb{P}} 1. \tag{4.6}$$

Also by Slutsky's theorem, convergence of $G_n^{(X,Y)}$ and (4.6) give

$$A_n G_n^{(X,Y)}(s) \Rightarrow G^{(X,Y)}(s). \tag{4.7}$$

Now let us have a look at B_n. We have

$$\begin{aligned} B_n &= \sqrt{n\overline{F}(u_n)}\left(\frac{n\mathbb{P}(X_1>su_n, Y_1>su_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j>u_n\}}} - \beta^{(X,Y)}(s)\right) \\ &= \sqrt{n\overline{F}(u_n)}\left(\frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j>u_n\}}} - 1\right)\frac{\mathbb{P}(X_1>su_n, Y_1>su_n)}{\overline{F}(u_n)} \\ &\quad + \sqrt{n\overline{F}(u_n)}\left(\frac{\mathbb{P}(X_1>su_n, Y_1>su_n)}{\overline{F}(u_n)} - \beta^{(X,Y)}(s)\right). \end{aligned}$$

Since $G_n^{(X,Y)}(s) \Rightarrow G^{(X,Y)}(s)$, for all $s \geq s_0$, then in particular, for $s = 1$,

$$H_n^X(1) := G_n^{(X,X)}(1) = \sqrt{n\overline{F}(u_n)}\left(\frac{\sum_{j=1}^n \mathbb{I}_{\{X_j>u_n\}}}{n\overline{F}(u_n)} - 1\right) \xrightarrow{d} G^{(X,X)}(1) =: H^X(1),$$

$$H_n^Y(1) := G_n^{(Y,Y)}(1) = \sqrt{n\overline{F}(u_n)}\left(\frac{\sum_{j=1}^n \mathbb{I}_{\{Y_j>u_n\}}}{n\overline{F}(u_n)} - 1\right) \xrightarrow{d} G^{(Y,Y)}(1) =: H^Y(1),$$

jointly with the convergence of $G_n^{(X,Y)}$ to $G^{(X,Y)}$. Application of the delta method gives

$$\sqrt{n\overline{F}(u_n)}\left(\frac{n\overline{F}(u_n)}{\sum_{j=1}^n \mathbb{I}_{\{X_j>u_n\}}} - 1\right) \xrightarrow{d} -H^X(1).$$

Therefore,

$$B_n \Rightarrow -\beta^{(X,Y)}(s)H^X(1). \tag{4.8}$$

Hence, combining (4.8) and (4.7), bearing in mind the joint convergence, we have

$$\sqrt{n\overline{F}(u_n)}\left(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)\right) \Rightarrow G^{(X,Y)}(s) - \beta^{(X,Y)}(s)H^X(1). \tag{4.9}$$

□

Properties of the limiting process $G_*^{(X,Y)}(s)$

First, we recall from Lemma 3.1.4 that

$$\mathbf{Cov}\Big(G^{(X,Y)}(s_1), G^{(X,Y)}(s_2)\Big) = \beta^{(X,Y)}(s_1 \vee s_2).$$

Using the same method as in the proof of Lemma 3.1.4 and letting $\Theta_n(x,y) = \frac{\mathbb{P}(X>xu_n, Y>yu_n)}{\overline{F}(u_n)}$ and $\Theta(x,y) = \lim_{n\to\infty}\Theta_n(x,y)$, then we can further conclude that

$$\begin{aligned}
\mathbf{Cov}\Big(G^{(X,Y)}(s_1), G^{(X,X)}(s_2)\Big) &= \mathbf{Cov}\Big(G^{(X,Y)}(s_1), H^X(s_2)\Big) = \Theta(s_1 \vee s_2, s_1), &(4.10)\\
\mathbf{Cov}\Big(G^{(X,Y)}(s_1), G^{(Y,Y)}(s_2)\Big) &= \mathbf{Cov}\Big(G^{(X,Y)}(s_1), H^Y(s_2)\Big) = \Theta(s_1, s_1 \vee s_2). &(4.11)
\end{aligned}$$

This together with the fact that $X \overset{d}{\approx} Y$, implies that for $s \in [s_0, +\infty)$,

$$\begin{aligned}
\mathbf{Cov}\Big(G_*^{(X,Y)}(s), G_*^{(X,Y)}(s)\Big) &= \beta^{(X,Y)}(s)\Big(1 - 2\Theta(s\vee 1, s) + \beta^{(X,Y)}(s)\Big). &(4.12)\\
\mathbf{Cov}\Big(G_*^{(X,Y)}(1), G_*^{(X,Y)}(1)\Big) &= \beta^{(X,Y)}(1)\Big(1 - \beta^{(X,Y)}(1)\Big). &(4.13)
\end{aligned}$$

Thus, we have a very important corollary which shows that we cannot construct a test for extremal independence using the proposed estimator of the tail dependence coefficient (see also Section 4.2).

Corollary 4.1.2 *Under the extremal independence, $G_*^{(X,Y)}(s)$ is degenerated, that is*

$$\boldsymbol{Var}(G_*^{(X,Y)}(s)) = 0. \tag{4.14}$$

Moreover,

$$\begin{aligned}
\mathbf{Cov}\Big(G_*^{(X,Y)}(s), G_*^{(X,X)}(s)\Big) &= \beta^{(X,Y)}(s)\Big(1 - (s\vee 1)^{-\alpha}\Big) + s^{-\alpha}\Big(\beta^{(X,Y)}(s) - \Theta(s\vee 1, s)\Big). &(4.15)\\
\mathbf{Cov}\Big(G_*^{(X,Y)}(s), G_*^{(Y,Y)}(s)\Big) &= \beta^{(X,Y)}(s)\Big(1 - \Theta(1, s)\Big) + s^{-\alpha}\Big(\beta^{(X,Y)}(s)\beta^{(X,Y)}(1) - \Theta(s, 1)\Big). &(4.16)\\
\mathbf{Cov}\Big(G_*^{(Y,Y)}(s), G_*^{(Y,Y)}(s)\Big) &= \mathbf{Cov}\Big(G_*^{(X,X)}(s), G_*^{(X,X)}(s)\Big) = s^{-\alpha}\Big(1 - 2(s\vee 1)^{-\alpha} + s^{-\alpha}\Big). &(4.17)\\
\mathbf{Cov}\Big(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\Big) &= \beta^{(X,Y)}(s)\Big(1 - \Theta(1, s)\Big) + s^{-\alpha}\Big(\beta^{(X,Y)}(s) - \Theta(s\vee 1, s)\Big). &(4.18)
\end{aligned}$$

Therefore regardless whether we have the extremal dependence or independence,

$$\mathbf{Cov}\Big(G_*^{(X,Y)}(1), G_*^{(X,X)}(1)\Big) = \mathbf{Cov}\Big(G_*^{(X,Y)}(1), G_*^{(Y,Y)}(1)\Big) = 0.$$

Confidence intervals

From Theorem 4.1.1, we know that $\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\beta}_n^{(X,Y)}(1)-\beta^{(X,Y)}(1)\Big)$ converges in distribution to $G_*^{(X,Y)}(1) \sim \mathcal{N}\Big\{0, \beta^{(X,Y)}(1)\Big(1-\beta^{(X,Y)}(1)\Big)\Big\}$, on account of Proposition 3.1.5. Hence, we can construct the $100(1-\theta)\%$ an asymptotic confidence interval for $\beta^{(X,Y)}(1)$ as

$$\left[\widetilde{\beta}_n^{(X,Y)}(1) - Z_{\theta/2}\frac{1}{\sqrt{n\overline{F}(u_n)}}\ ;\ \widetilde{\beta}_n^{(X,Y)}(1) + Z_{\theta/2}\frac{1}{\sqrt{n\overline{F}(u_n)}}\right],$$

where $Z_{\theta/2}$ is the quantile of

$$G_*^{(X,Y)}(1) = G^{(X,Y)}(1) - \beta^{(X,Y)}(1)H^X(1)\ ,$$

i.e. $\mathbb{P}(G_*^{(X,Y)}(1) > Z_{\theta/2}) = \theta/2$. Recalling that $G_*^{(X,Y)}(1)$ is a centered normal random variable with the variance given in (4.13) and replacing therein $\beta^{(X,Y)}(s)$ with $\widetilde{\beta}_n^{(X,Y)}(s)$, the confidence interval can be constructed as

$$\left[\widetilde{\beta}_n^{(X,Y)}(1) - z_{\theta/2}\sqrt{\frac{\widetilde{\beta}_n^{(X,Y)}(1)(1-\widetilde{\beta}_n^{(X,Y)}(1))}{n\overline{F}(u_n)}}\ ;\ \widetilde{\beta}_n^{(X,Y)}(1) + z_{\theta/2}\sqrt{\frac{\widetilde{\beta}_n^{(X,Y)}(1)(1-\widetilde{\beta}_n^{(X,Y)}(1))}{n\overline{F}(u_n)}}\right], \tag{4.19}$$

where $z_{\theta/2}$ is the standard normal percentile.

Example 4.1.3 *[**Example** 2.1.17 continued]* We simulate 1000 observations from the model given in Example 2.1.17 where $Z_j's$ are independent Pareto with $\alpha = 4$ and $X = Z_1 + Z_2$, $Y = Z_2 + Z_3$. In this case, the tail dependent coefficient is $\beta^{(X,Y)}(1) = 0.5$. Figure 4.1 shows the plot of the confidence intervals against the threshold $u_n = n^\delta$ which is such that $n\bar{F}(u_n) = n^{1-\alpha\delta} \to +\infty$, that is $\delta \in (0, 1/\alpha)$.

Note that the true confidence interval (with $\beta^{(X,Y)}(1) = 0.5$) and the estimated confidence interval are almost not distinguishable.

4.2 Testing for Extremal (In)dependence using TDC

In this section we indicate problems related to testing for extremal (in)dependence between two random variables using their tail dependence coefficient.

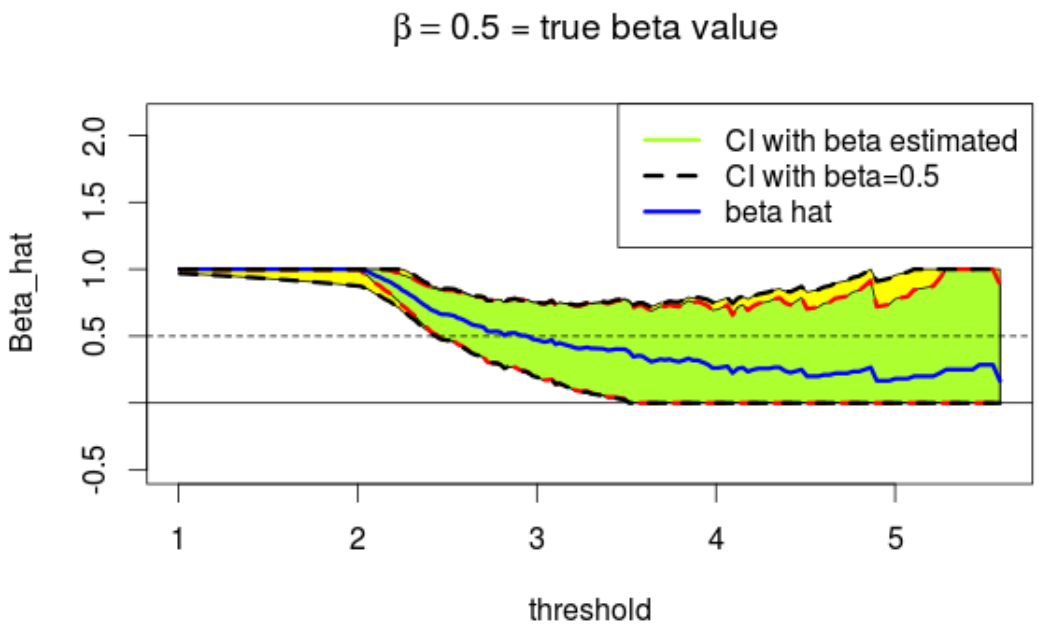

Figure 4.1: Asymtotic confidence intervals for the tail dependence coefficient for Example 4.1.3. Blue line: the estimator $\widetilde{\beta}_n^{(X,Y)}(1)$ plotted against different choices of the threshold u_n; red dashed lines - confidence interval using (4.19); yellow line - confidence interval using the true value of $\beta^{(X,Y)}(1)$.

The test for independence would have the form:

$$H_0: \quad \beta^{(X,Y)}(\cdot) = 0 \quad vs \quad H_1: \quad \beta^{(X,Y)}(\cdot) \neq 0,$$

where

$$\beta^{(X,Y)}(s) = \lim_{n \to +\infty} \frac{\mathbb{P}(X > su_n, Y > su_n)}{\mathbb{P}(X > u_n)}.$$

To conduct the test, we could use $\widetilde{\beta}_n^{(X,Y)}(1)$ defined in (4.1). **However, under H_0 the limit of the estimator $\widetilde{\beta}_n^{(X,Y)}(1)$ of $\beta^{(X,Y)}(1)$ is degenerated, hence we cannot construct a formal test**; see Corollary 4.1.2. On the other hand, we can test for specific values of the tail dependence coefficient:

$$H_0: \quad \beta^{(X,Y)}(\cdot) = \beta_0 \quad vs \quad H_1: \quad \beta^{(X,Y)}(\cdot) \neq \beta_0.$$

If $\beta_0 \neq 0$, then to conduct the test we can use the confidence interval (4.19). But we have some issues with this approach since:

- the distribution F is unknown and u_n is arbitrary;

- we cannot extend this procedure to higher dimension, the problem we will encounter here is the multiple testing which involves the "Type I errors" or "False positive" which is the probability of rejecting H_0 while it is true. For example, assume that we want to perform 50 tests at the level of significance $\theta = 5\%$ that is for each test we have 5% chance of making a "Type I errors". If all the null hypothesis are true, then the expected number of "Type I error" is 2.5. Moreover, if all tests are assumed to be independent, then we have a binomial distribution with $n = 50$ and $p = 5\%$, therefore

$$\mathbb{P}(\text{at least one "Type I error"}) = 1 - \mathbb{P}(\text{ no "Type I error"}) = 1 - (1 - 0.05)^{50} = 92.3\%.$$

So, as we can see, with 50 tests, we have 92.3% chance of having at least one "Type I error."
The first issue is addressed using radom levels, while the second issue is tackled using random matrices approach. As we will see, the latter approach will allow us to construct a formal test for extremal independence.

4.3 Extremogram matrix: bivariate case

The covariance matrix of a random vector is a standard object summarizing all the dependence between several its components. Since the tail dependence coefficient describes extremal *dependence*, we can think of introducing an extremal counterpart to the covariance matrix, namely the extremogram matrix (we borrow the *extremogram* terminology from [3]).

The section is organized as follows. We first define the extremogram matrix and its sample counterpart. Then we prove, under the null hypothesis of the extremal independence, that the distribution of the largest eigenvalue of the sample extremogram matrix converges to the maximum of two independent Gaussian random variables; see Theorem 4.3.4. Then, we apply the limiting result to construct a test for independence. Theory is illustrated by simulated data.

The material presented in this section is the author original contribution.

Let (X, Y) be a random vector.

Definition 4.3.1 (Extremogram matrix) *We define the extremogram matrix by*

$$\Delta(s) = \begin{pmatrix} \beta^{(X,X)}(s) & \beta^{(X,Y)}(s) \\ \beta^{(Y,X)}(s) & \beta^{(Y,Y)}(s) \end{pmatrix}. \tag{4.20}$$

It is an analog of the covariance matrix but it depends only on extreme values. If $s = 1$, then we have

$$\Delta(1) := \begin{pmatrix} 1 & \beta^{(X,Y)}(1) \\ \beta^{(Y,X)}(1) & 1 \end{pmatrix}.$$

A sample counterpart to the covariance matrix is the sample covariance matrix. We extend this idea to the extremal dependence.

Definition 4.3.2 (Sample extremogram matrix) *We call the sample extremogram matrix, denoted by $\widetilde{\Delta}_n(s)$, the matrix whose entries are estimators of the tail dependence coefficients, i.e.*

$$\widetilde{\Delta}_n(s) = \begin{pmatrix} \widetilde{\beta}_n^{(X,X)}(s) & \widetilde{\beta}_n^{(X,Y)}(s) \\ \widetilde{\beta}_n^{(Y,X)}(s) & \widetilde{\beta}_n^{(Y,Y)}(s) \end{pmatrix}, \tag{4.21}$$

where

$$\widetilde{\beta}_n^{(U,V)}(s) = \frac{\sum_{j=1}^n \mathbb{I}_{\{U_j > su_n, V_j > su_n\}}}{\sum_{j=1}^n \mathbb{I}_{\{U_j > u_n\}}}$$

with U, V equal X or Y.

We note that the extremogram matrix and the sample extremogram matrix are symmetric.
The next goal is to establish the asymptotic distribution of the eigenvalues of the sample extremogram matrix.

Why random matrices are important?

Before we continue, let us give a brief motivation. Attention are given to spectral properties of large dimensional random matrices known as Random Matrix Theory (RMT) because of their interesting properties and statistical applications. The use of the limiting properties of eigenvalues originates from quantum mechanics where they are utilized to describe energy levels of particles in a large system and also serve as finite dimensional approximation of infinite dimensional operator. From statistical point of view, they may be used to correct traditional tests or estimators which fail in large dimension. Furthermore, in the principal component analysis (PCA), the first k principal components correspond to the k largest eigenvalues of the sample covariance matrix (e.g. [6], [1], [4])

Asymptotics for sample extremogram matrices

By (4.4), we have

$$\sqrt{n\overline{F}(u_n)}(\mathrm{vec}(\widetilde{\Delta}_n(s)) - \mathrm{vec}(\Delta(s))) \Rightarrow \mathrm{vec}(\mathcal{H}(s)). \tag{4.22}$$

where

$$\mathcal{H}(s) = \begin{pmatrix} G_*^{(X,X)}(s) & G_*^{(X,Y)}(s) \\ G_*^{(Y,X)}(s) & G_*^{(Y,Y)}(s) \end{pmatrix}, \tag{4.23}$$

,

$$G_*^{(U,V)}(s) = G^{(U,V)}(s) - \beta^{(U,V)}(s) H^U(1).$$

and vec(A) is a vector obtained by stacking the columns of the matrix A on top of one another. Before we prove the result on limiting distribution for the sample eigenvalues, we briefly introduce the operator norm of matrices.

Operator norms. Let $\Sigma = [\sigma_{ij}]_{i,j=1}^d \in \mathbb{R}^{d\times d}$ and $|\cdot|$ be a norm in $\mathbb{R}^d$. Matrix norms are defined as follows:

$$\| \Sigma \|_2 = \max\{|\Sigma x| : \ |x| = 1\} \quad \textit{and} \quad \| \Sigma \|_\infty = \max_{1\le i\le d} \sum_{j=1}^d |\sigma_{ij}|. \tag{4.24}$$

It holds that

$$\| \Sigma \|_2^2 = \max_{1\le i\le d} |\delta_i| \quad \textit{and} \quad \| \Sigma \|_2 \le \| \Sigma \|_\infty, \tag{4.25}$$

where δ_i are eigenvalues of Σ. Recall the Weyl inequality for two matrices Σ, Γ: if δ_i and γ_i, $i = 1, \cdots, d$ are ordered eigenvalues of Σ and Λ, respectively. Then

$$\max_{1\le i\le d} |\delta_i - \gamma_i| \le \| \Sigma - \Gamma \|_2 .$$

4.3.1 Asymptotics for eigenvalues under extremal independence

Again, the goal is to test for extremal independence between X and Y. Therefore, we test

$$H_0 : \ \beta^{(X,Y)}(\cdot) = 0 \quad vs \quad H_1 : \ \beta^{(X,Y)}(\cdot) \neq 0.$$

Limiting distribution of the sample eigenvalues under the null hypothesis. Under the null hypothesis we have

$$\widetilde{\Delta}_n(s) = \begin{pmatrix} \widetilde{\beta}_n^{(X,X)}(s) & \widetilde{\beta}_n^{(X,Y)}(s) \\ \widetilde{\beta}_n^{(Y,X)}(s) & \widetilde{\beta}_n^{(Y,Y)}(s) \end{pmatrix}, \ \widetilde{\Delta}_n(1) = \begin{pmatrix} 1 & \widetilde{\beta}_n^{(X,Y)}(1) \\ \widetilde{\beta}_n^{(Y,X)}(1) & 1 \end{pmatrix},$$

$$\Delta(s) = \begin{pmatrix} s^{-\alpha} & 0 \\ 0 & s^{-\alpha} \end{pmatrix}, \quad \Delta(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We recall that under H_0 the process $G_*^{(X,Y)}(\cdot)$ is degenerated. Therefore under H_0 the matrix $\mathcal{H}$ defined in (4.23) becomes

$$\mathcal{H}(s) = \begin{pmatrix} G_*^{(X,X)}(s) & 0 \\ 0 & G_*^{(Y,Y)}(s) \end{pmatrix}.$$

We also recall that for $s = 1$ we have $G_*^{(X,X)}(1) = G_*^{(Y,Y)}(1) = 0$, hence it is important to keep $s \neq 1$ above. In summary,

Corollary 4.3.3 *Assume that the conditions of Theorem 4.1.1 are satisfied. Then, under H_0*

$$\sqrt{n\overline{F}(u_n)} \begin{pmatrix} \widetilde{\beta}_n^{(X,X)}(s) - s^{-\alpha} \\ \widetilde{\beta}_n^{(X,Y)}(s) \\ \widetilde{\beta}_n^{(Y,X)}(s) \\ \widetilde{\beta}_n^{(Y,Y)}(s) - s^{-\alpha} \end{pmatrix} \Longrightarrow \begin{pmatrix} G_*^{(X,X)}(s) \\ 0 \\ 0 \\ G_*^{(Y,Y)}(s) \end{pmatrix}.$$

We recall that $s = 1$ leads to the tail dependence coefficient. However, if $s = 1$ then

$$\sqrt{n\overline{F}(u_n)} \begin{pmatrix} 0 \\ \widetilde{\beta}_n^{(X,Y)}(1) \\ \widetilde{\beta}_n^{(Y,X)}(1) \\ 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, under H_0, the limiting distribution does not provide any valuable information for hypothesis testing. To avoid this drawback, we have to consider $s \neq 1$ or the supremum over a given set. Let I be a compact set, $I \subseteq \mathbb{R}_+$ such that $1 \in I$. Then we have the following result.

Theorem 4.3.4 *Assume that $s \neq 1$. Let $\widetilde{\lambda}_{n(1)}(s) \geq \widetilde{\lambda}_{n(2)}(s)$ be the ordered eigenvalues of $\widetilde{\Delta}_n(s)$ and $\lambda_{(1)}(s) = \lambda_{(2)}(s) = s^{-\alpha}$ be the eigenvalues of $\Delta(s)$. Then, under H_0 and if the conditions of Theorem 4.1.1 hold, we have*

$$\sqrt{n\overline{F}(u_n)}\left(\widetilde{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\right) \Rightarrow \max\left(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\right).$$

Proof:

Define a diagonal matrix $\widetilde{\Delta}_n^*(s)$ whose diagonal elements are those of $\widetilde{\Delta}_n(s)$, that is

$$\widetilde{\Delta}_n^*(s) = \begin{pmatrix} \widetilde{\beta}_n^{(X,X)}(s) & 0 \\ 0 & \widetilde{\beta}_n^{(Y,Y)}(s) \end{pmatrix}.$$

Then, it holds by Corollary 4.3.3 that

$$\sqrt{n\overline{F}(u_n)}\mathrm{vec}\Big(\widetilde{\Delta}_n(s) - \widetilde{\Delta}_n^*(s)\Big) = \sqrt{n\overline{F}(u_n)} \begin{pmatrix} 0 \\ \widetilde{\beta}_n^{(X,Y)}(s) \\ \widetilde{\beta}_n^{(Y,X)}(s) \\ 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, we should expect that the $\|\cdot\|_\infty$ norm for $\widetilde{\Delta}_n(s) - \widetilde{\Delta}_n^*(s)$ goes to zero. In fact,

Lemma 4.3.5 *Under the conditions of Theorem 4.3.4:*

$$\sup_{s\in I} \sqrt{n\overline{F}(u_n)} \parallel \widetilde{\Delta}_n(s) - \widetilde{\Delta}_n^*(s) \parallel_\infty = o_p(1).$$

Proof: We have by (4.24) that

$$\begin{aligned} &\sup_{s\in I} \sqrt{n\overline{F}(u_n)} \parallel \widetilde{\Delta}_n(s) - \widetilde{\Delta}_n^*(s) \parallel_\infty = \sup_{s\in I} \sqrt{n\overline{F}(u_n)} \max\Big(\widetilde{\beta}_n^{(X,Y)}(s), \widetilde{\beta}_n^{(Y,X)}(s)\Big) \\ &\leq \sup_{s\in I} \sqrt{n\overline{F}(u_n)}\widetilde{\beta}_n^{(X,Y)}(s) + \sup_{s\in I} \sqrt{n\overline{F}(u_n)}\widetilde{\beta}_n^{(Y,X)}(s) = o_p(1), \end{aligned}$$

on account of Corollary 4.3.3. □

By (4.25) and Lemma 4.3.5 we have

$$\sqrt{n\overline{F}(u_n)} \parallel \widetilde{\Delta}_n(s) - \widetilde{\Delta}_n^*(s) \parallel_2 = o_p(1), \tag{4.26}$$

that is

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\lambda}_{n(i)}(s) - \widetilde{\lambda}_{n(i)}^*(s)\Big) = o_p(1)\ ,$$

where $\widetilde{\lambda}_{n(1)}^*(s) \geq \widetilde{\lambda}_{n(2)}^*(s)$ are the ordered eigenvalues of $\widetilde{\Delta}_n^*(s)$. That is

$$\widetilde{\lambda}_{n(1)}^*(s) = \max(\widetilde{\beta}_n^{(X,X)}(s), \widetilde{\beta}_n^{(Y,Y)}(s)) \qquad \text{and} \qquad \lambda_{(1)}(s) = \max(s^{-\alpha}, s^{-\alpha}) = s^{-\alpha}.$$

$$\widetilde{\lambda}_{n(2)}^*(s) = \min(\widetilde{\beta}_n^{(X,X)}(s), \widetilde{\beta}_n^{(Y,Y)}(s)) \qquad \text{and} \qquad \lambda_{(2)}(s) = s^{-\alpha}.$$

We continue with the proof of Theorem 4.3.4. By application of Corollary 4.3.3 to $\widetilde{\Delta}_n^*(s)$ and $\Delta_0(s)$, we get

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\beta}_n^{(X,X)}(s) - s^{-\alpha}\Big) \Rightarrow G_*^{(X,X)}(s) \tag{4.27}$$

and

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\beta}_n^{(Y,Y)}(s) - s^{-\alpha}\Big) \Rightarrow G_*^{(Y,Y)}(s). \tag{4.28}$$

We now want to obtain the limiting distribution of $\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\lambda}_{n(1)}^*(s) - \lambda_{(1)}(s)\Big)$. From (4.27) and (4.28), we deduce that

$$\begin{aligned}\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\lambda}_{n(1)}^*(s) - \lambda_{(1)}(s)\Big) &= \sqrt{n\overline{F}(u_n)}\max\Big(\widetilde{\beta}_n^{(X,X)}(s) - s^{-\alpha}, \widetilde{\beta}_n^{(Y,Y)}(s) - s^{-\alpha}\Big) \\ &\Rightarrow \max\Big(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\Big). \end{aligned} \tag{4.29}$$

Hence, combining (4.26) and (4.29), it holds

$$\sqrt{n\overline{F}(u_n)}\Big(\widetilde{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\Big) \Rightarrow \max\Big(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\Big),$$

which leads to the desired result. □

4.3.2 Testing for Extremal Independence using random matrices

Define the test statistic for $s \in [s_0, +\infty)$, $s_0 \in (0,1)$ as

$$T_1(s) = \sqrt{n\overline{F}(u_n)}\Big(\widetilde{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\Big). \tag{4.30}$$

Then under H_0 and for any fixed s, $T_1(s)$ has asymptotically the same distribution as $M(s) = \max\Big(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\Big)$. We know that $G_*^{(X,X)}(s)$ and $G_*^{(Y,Y)}(s)$ are jointly normal random variables. By (4.18), under H_0, the random variables are uncorrelated (for $s \geq s_0$). Hence, they are independent. By [8], The distribution function of $M(s)$ can be written as

$$F_{M(s)}(x) = \Phi^2\Big(\frac{x}{\sigma(s)}\Big), \quad x \in \mathbb{R}.$$

Hence,

$$f_{M(s)}(x) = \frac{2}{\sigma(s)}\phi\Big(\frac{x}{\sigma(s)}\Big)\Phi\Big(\frac{x}{\sigma(s)}\Big), \quad x \in \mathbb{R},$$

where $\sigma^2(s) = \mathrm{Var}\Big(\mathrm{G}_*^{(\mathrm{X},\mathrm{X})}(\mathrm{s})\Big) \overset{(4.17)}{=} \mathrm{s}^{-\alpha}\Big(1 - 2(\mathrm{s} \vee 1)^{-\alpha} + \mathrm{s}^{-\alpha}\Big)$, $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the pdf and the cdf of the standard normal distribution. The mean and the variance of $M(s)$ are

$$\mathbb{E}(M(s)) = \sigma(s)\pi^{-1/2}, \qquad \mathrm{Var(M(s))} = \sigma^2(\mathrm{s})\pi^{-1}(\pi - 1).$$

Therefore, the .025 and .975 quantiles of $M(s)$ can be calculated as

$$M_{.025}(s) = \sigma(s)\Phi^{-1}(\sqrt{.025}), \qquad M_{.975}(s) = \sigma(s)\Phi^{-1}(\sqrt{.975})\ .$$

Note that the distribution of $M(s)$ is not symmetric, see Figure 4.2. The hypothesis testing can be performed as follows: If $\widetilde{\theta}_n(s) \in \Big(M_{.025}(s), M_{.975}(s)\Big)$, then we fail to reject H_0, othewise, we reject H_0.

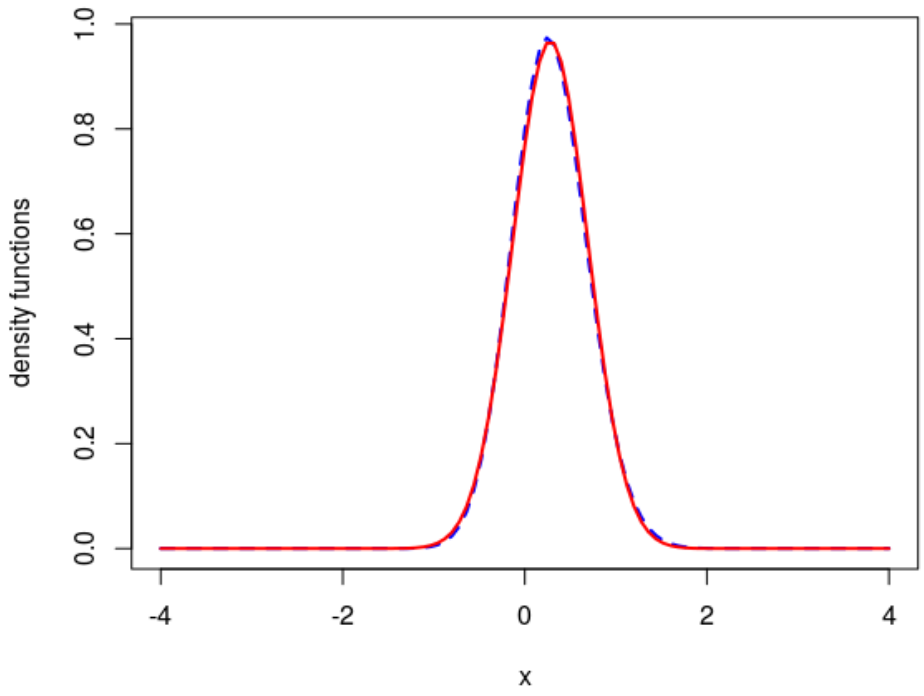

Figure 4.2: The density function of $M(s = 1.2)$ (blue dashed line) and the density of the normal distribution with $\mu = \mathbb{E}(M(s))$ and $\sigma^2 = Var(M(s))$ (red line).

We apply these results to the models defined in Examples 2.1.16, 2.1.17 and the model $Y = \phi X + \sigma|Z|$ where $\phi \in (0, 1)$, X is Pareto and Z is standard normal independent of X.

We simulate 1000 independent observations from (X, Y) from each of these models. Random variables

Model //True TDC	Ind. Pareto	Dep. Pareto	Third model
	$\lambda_{TDC} = 0$	$\lambda_{TDC} = 0.5$	$\lambda_{TDC} = 0.8^4$
$T_1(1.2)$	0.12	3.65	1.34
CI(1.2)	+	-	-
$T_1(1.3)$	0	2.18	1.10
CI(1.3)	+	-	-
$T_1(1.5)$	0	0.83	0.61
CI(1.5)	+	-	-

Table 4.1: Testing for extremal independence. The 95% asymptotic confidence intervals are CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.2)\Big) = \Big(-0.25, 0.56\Big)$, CI(1.3)=$\Big(M_{.025}(1.3), M_{.975}(1.3)\Big) = \Big(-0.23, 0.51\Big)$ and CI(1.5)=$\Big(M_{.025}(1.5), M_{.975}(1.5)\Big) = \Big(-0.16, 0.35\Big)$ for $s = 1.2, 1.3, 1.5$, $u_n = n^\delta$ and $\delta = .10$. The "+" sign indicates that T_1 belongs to the CI.

X and Y are Pareto with the parameter $\alpha = 4$. For the third model, we chose $\phi = 0.8$ and $\sigma = 0.1$. We calculate the test statistics $T_1(s)$ with $u_n = n^\delta$, $\delta = .10$ and $s = 1.2, 1.3, 1.5$. The simulation results are summarized in Table 4.1. As we can see, the simulated results confirmed the theoretical results obtained in the previous section. Indeed, for independent Pareto, we fail to reject H_0 for any $s \in \{1.2, 1.3, 1.5\}$, while for the dependent Pareto (Example 2.1.17) and the third model mentioned above, we reject the null hypothesis.

A comprehensive simulation study will be done using the *practical estimators*, based on the order statistics, not on the deterministic threshold u_n.

A good parameter choice For a good asymptotic confidence set, the parameter s has to be chosen roughly in $[0.7; 1) \cup (1; 1.5]$. However, large values of s produce a narrower asymptotic confidence sets and small values of s generate a wider asymptotic confidence sets.

4.4 Extremogram matrix: d-dimensional case

The following section contains an extension of the two dimensional case to higher but finite dimension, say $d \geq 2$. To this end, we will conveniently introduce the following notation.

Definition 4.4.1 *Let $X \in \mathbb{R}^d$. Let $X_j = (X_{1j}, \ldots, X_{dj})$, $j = 0 \ldots, n$, be an i.i.d. sequence of regularly varying nonnegative random vectors. Assume that $X_{1j} \stackrel{d}{=} ... \stackrel{d}{=} X_{dj}$. Define the tail dependence coefficient*

between the k^{th} and ℓ^{th} component of X_0 as

$$\beta^{(k,\ell)}(s) = \lim_{n\to+\infty} \beta_n^{(k,\ell)}(s) = \lim_{n\to+\infty} \frac{\mathbb{P}(X_{k0} > su_n, X_{\ell 0} > tu_n)}{\mathbb{P}(X_{10} > u_n)}, \quad k,\ell = 1,\dots,d,$$

and its empirical estimate by

$$\widetilde{\beta}_n^{(k,\ell)}(s) = \frac{\sum_{j=1}^n \mathbb{I}_{\{X_{kj} > su_n, X_{\ell j} > tu_n\}}}{\sum_{j=1}^n \mathbb{I}_{\{X_{kj} > su_n\}}}.$$

We also define the extremogram matrix and its sample estimate as follows

$$\Delta(s) = \left(\beta^{(k,\ell)}(s)\right)_{k,\ell=1}^d, \qquad \widetilde{\Delta}_n(s) = \left(\widetilde{\beta}_n^{(k,\ell)}(s)\right)_{k,\ell=1}^d.$$

4.4.1 Testing for Extremal Independence using random matrices

Again, the goal is to test for extremal independence between the components of the vector X_0. Therefore, we test

$$H_0: \ \beta^{(k,\ell)}(\cdot) = 0 \quad vs \quad H_1: \ \beta^{(k,\ell)}(\cdot) \neq 0 \ \text{ for all } k \neq \ell\ .$$

Limiting distribution of sample eigenvalues in higher but finite dimension $d \geq 2$. Note first that under the null hypothesis, $\beta^{(k,\ell)}(s) = 0$, for $k \neq \ell$. We have the following theorem which is an extension of the two dimensional case (see Theorem 4.3.4).

Theorem 4.4.2 *Let $X_j, j \geq 1$ be a random sequence as in Definition 4.4.1. Let $\widetilde{\lambda}_{n(1)}(s) \geq \ldots \geq \widetilde{\lambda}_{n(d)}(s)$ be the ordered eigenvalues of $\widetilde{\Delta}_n(s)$ and $\lambda_{(1)}(s) = \ldots = \lambda_{(d)}(s) = s^{-\alpha}$ be the eigenvalues of $\Delta(s)$. Assume that for all $k,\ell = 1,\dots,d$ we have*

$$\lim_{n\to+\infty} \sqrt{n\overline{F}(u_n)} \sup_{s>\epsilon>0} |\beta_n^{(k,\ell)}(s) - \beta^{(k,\ell)}(s)| = 0. \tag{4.31}$$

Then, under H_0 we have

$$\sqrt{n\overline{F}(u_n)}\left(\widetilde{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\right) \Rightarrow \max_{1\leq k\leq d}\left(G_*^{(k,k)}(s)\right), \tag{4.32}$$

where $G_*^{(k,k)}(s)$, $k = 1,\dots,d$, are independent Gaussian processes with the covariance $\mathrm{Cov}(\mathrm{G}_*^{(\mathrm{k,k})}(\mathrm{s}), \mathrm{G}_*^{(\mathrm{k,k})}(\mathrm{s})) = \mathrm{s}^{-\alpha}\left(1 - 2(\mathrm{s}\vee 1)^{-\alpha} + \mathrm{s}^{-\alpha}\right)$.

Proof: The same as in the two dimensional case.

4.5 Tail empirical process with random levels

In this section, we replace the deterministic levels u_n with random, data based, levels and obtain the same limit theorems as in the deterministic levels case. The transition from deterministic to random levels follows the same path as in [7]. Then, we apply again these results to hypothesis testing using random matrices. This approach is again the original author's contribution.

The process considered in (3.3) depends on the deterministic levels u_n which are chosen such that

$$u_n \to +\infty \text{ and } n\overline{F}(u_n) \to +\infty.$$

In other words the choice of u_n depends on the index of regular variation α, which is unknown. To tackle this drawback, empirical processes with random levels are considered. Assume for a moment that F is known, is continuous and strictly increasing. We consider $X_{n:1} \leq \ldots \leq X_{n:n}$, the order statistics from the sample X_j, $j = 1, \ldots, n$. For the given u_n, choose a sequence of integers $k = n\overline{F}(u_n)$ depending on n (the dependence in n is omitted for the notation point of view). Consequently, $k = k_n \to \infty$ and $k/n \to 0$. Then $F^{\leftarrow}(1 - \frac{k}{n}) = u_n$, where $F^{\leftarrow}$ is the inverse of F. Furthermore, let

$$\widehat{F}_n^{\leftarrow}(u) = \inf\{y : \widehat{F}_n(y) > u\}.$$

The empirical estimator of $u_n = F^{\leftarrow}(1-\frac{k}{n})$ is given by $\widehat{F}_n^{\leftarrow}(1-\frac{k}{n}) = X_{n:n-k}$. We call $X_{n:n-k}$ the *intermediate order statistics*. In conclusion, u_n can be approximated by $X_{n:n-k}$. This motivates the following data-driven estimator of the tail dependence coefficient between X and Y:

$$\widehat{\beta}_n^{(X,Y)}(s) = \frac{1}{k}\sum_{j=1}^{n} \mathbb{I}_{\{X_j > sX_{n:n-k}, Y_j > sX_{n:n-k}\}}. \tag{4.33}$$

We note that $\widehat{\beta}_n^{(X,Y)}(s)$ is just $\widetilde{\beta}_n^{(X,Y)}(s)$ defined in (4.1), where u_n is replaced with $X_{n:n-k}$.

The goal is to obtain the limiting theory for $\widehat{\beta}_n^{(X,Y)}(s)$. In order to do so, we need to obtain the joint convergence of the intermediate order statistics and the tail empirical process $G_n^{(X,Y)}(s)$.

Recall from (3.5) and (3.6) that

$$\widetilde{\delta}_n^X(s) = \frac{1}{n\overline{F}(u_n)}\sum_{j=1}^{n} \mathbb{I}_{\{X_j > su_n\}}\ , \quad \widetilde{\delta}_n^Y(s) = \frac{1}{n\overline{F}(u_n)}\sum_{j=1}^{n} \mathbb{I}_{\{Y_j > su_n\}},$$

and

$$\delta(s) := \delta^Y(s) = \delta^X(s) = \lim_{n\to+\infty} \delta_n^X(s) = \lim_{n\to+\infty} \mathbb{E}[\widetilde{\delta}_n^X(s)] = \lim_{n\to\infty} \overline{F}(su_n)/\overline{F}(u_n) = s^{-\alpha} .$$

We make the following assumption:

$$\frac{d\delta_n}{ds}(s) \to \frac{d\delta}{ds}(s) = -\alpha s^{-\alpha-1}, \tag{4.34}$$

uniformly in a neighborhood of 1.

Proposition 4.5.1 *Assume that the conditions of Theorem 4.1.1 hold. If moreover condition (4.34) is satisfied, then*

$$\sqrt{k}\left(\left\{\frac{X_{n:n-k}}{u_n} - 1\right\}, G_n^{(X,Y)}(\cdot)\right) \Rightarrow \left(\alpha^{-1}H^X(1), G^{(X,Y)}(\cdot)\right),$$

in $\mathbb{R} \times \ell^\infty[s_0, +\infty)$.

Proof: Recall from (3.3) and (3.4)

$$G_n^{(X,Y)}(s) = \sqrt{n\overline{F}(u_n)}\left\{\widetilde{\beta}_n^{(X,Y)}(s) - \beta_n^{(X,Y)}(s)\right\}, \ s \geq s_0,$$
$$H_n^X(s) = \sqrt{n\overline{F}(u_n)}(\widetilde{\delta}_n(s) - \delta_n(s)), \ s \geq s_0.$$

By Theorem 3.1.3 and Skorokhod representation theorem (Theorem 2.2.3), there exist sequences of processes $\zeta_n^{(X,Y)} \stackrel{d}{=} G_n^{(X,Y)}$, $\zeta^{(X,Y)} \stackrel{d}{=} G^{(X,Y)}$ and $\zeta_n \stackrel{d}{=} H_n^X$, $\zeta \stackrel{d}{=} H$ on some probability space such that

$$\zeta_n^{(X,Y)}(\cdot) \to \zeta^{(X,Y)}(\cdot), \quad \zeta_n(\cdot) \to \zeta(\cdot), \tag{4.35}$$

almost surely, uniformly on a compact subsets of $[s_0, +\infty)$, $s_0 \in (0,1)$. Then (4.35) and the composition mapping theorem yield

$$\zeta_n(\delta_n^{\leftarrow}(s)) = \sqrt{k}\left(\widetilde{\delta}_n \circ \delta_n^{\leftarrow}(s) - s\right) \to \zeta(\delta^{\leftarrow}(s)), \ \ a.s, \tag{4.36}$$

uniformly on compact subsets of $[s_0, +\infty)$. By application of Vervaat's Lemma 2.2.9, with $1/\gamma_n = k^{1/2}$, $\xi_n(s) = \widetilde{\delta}_n \circ \delta_n^{\leftarrow}(s)$, $g(s) = s$ and $\xi_0(s) = \zeta \circ \delta^{\leftarrow}(s)$, we have

$$\zeta_n(\delta_n^{\leftarrow}(s)) = \sqrt{k}\left(\left(\widetilde{\delta}_n \circ \delta_n^{\leftarrow}\right)^{\leftarrow}(s) - s\right) \to -\zeta(\delta^{\leftarrow}(s)), \ \ a.s,$$

uniformly on compact subsets of $[s_0, +\infty)$. That is

$$\sqrt{k}\Big(\delta_n \circ \widetilde{\delta}_n^{\leftarrow}(s) - s\Big) \to -\zeta(\delta^{\leftarrow}(s)), \quad a.s. \tag{4.37}$$

Observe now that $\widetilde{\delta}_n^{\leftarrow}(1) = X_{n:n-k}/u_n$ and $\delta_n^{\leftarrow}(1) = 1$, then by Taylor's expansion around 1 of $\delta_n \circ \widetilde{\delta}_n^{\leftarrow}$, we have

$$\begin{aligned}
\delta_n(\widetilde{\delta}_n^{\leftarrow}(1)) - 1 &= \delta_n(\widetilde{\delta}_n^{\leftarrow}(1)) - \delta_n(\delta_n^{\leftarrow}(1)) \\
&= \frac{d\delta_n}{ds}(\delta_n^{\leftarrow}(1))\Big(\widetilde{\delta}_n^{\leftarrow}(1) - \delta_n^{\leftarrow}(1)\Big)(1 + o(1)) \\
&= \frac{d\delta_n}{ds}(\delta_n^{\leftarrow}(1))\Big(X_{n:n-k}/u_n - 1\Big)(1 + o(1)).
\end{aligned} \tag{4.38}$$

Hence, by (4.37), (4.38) and (4.34), we have

$$\sqrt{k}\Big(\delta_n \circ \widetilde{\delta}_n^{\leftarrow}(1) - 1\Big) = -\alpha\sqrt{k}\big(X_{n:n-k}/u_n - 1\big)(1 + o(1)) \;\; i.e. \;\; \sqrt{k}\Big\{X_{n:n-k}/u_n - 1\Big\} \xrightarrow{a.s.} \alpha^{-1}\zeta(1). \tag{4.39}$$

Since (4.39) and (4.35) hold almost surely, then they hold jointly. Therefore, the convergence holds also in the original probability space. □

Remark 4.5.2 *If* (4.39) *holds, then* $X_{n:n-k}/u_n - 1 \to 0$, *in probability.*

Recall from (4.33) that

$$\widehat{\beta}_n^{(X,Y)}(s) = \frac{1}{k}\sum_{j=1}^{n} \mathbb{I}_{\{X_j > sX_{n:n-k}, Y_j > sX_{n:n-k}\}}, \quad \widehat{\delta}_n(s) = \frac{1}{k}\sum_{j=1}^{n} \mathbb{I}_{\{X_j > X_{n:n-k}\}}. \tag{4.40}$$

Consider the following empirical processes

$$\widehat{G}_n^{(X,Y)}(s) = \sqrt{k}\left(\widehat{\beta}_n^{(X,Y)}(s)) - \beta^{(X,Y)}(s)\right), \quad \widehat{H}_n^{X}(s) = \sqrt{k}\Big(\widehat{\delta}_n(s) - \delta(s)\Big), \;\; s \geq s_0. \tag{4.41}$$

($\widehat{H}_n^{Y}(s)$ is defined similarly). We have the weak convergence of the tail empirical processes defined in (4.41).

Theorem 4.5.3 *Suppose that the conditions of Theorem 4.1.1 hold, where* $n\overline{F}(u_n)$ *is replaced by* k. *Furthermore, assume that* (4.34) *holds and* $(d/ds)\beta^{(X,Y)}(s)$ *exists and is finite in a neighborhood of* 1. *Then,*

$$\widehat{G}_n^{(X,Y)}(s) \Rightarrow G^{(X,Y)}(s) - \beta^{(X,Y)}(s)H^{X}(1), \tag{4.42}$$

as $n \to +\infty$ *in* $\ell^\infty[s_0, +\infty)$. *Moreover,*

$$\widehat{H}_n^X(s) \Rightarrow H^X(s) - s^{-\alpha} H^X(1), \quad \widehat{H}_n^Y(s) \Rightarrow H^Y(s) - s^{-\alpha} H^Y(1) \tag{4.43}$$

and the convergences hold jointly.

Proof: Recall that

$$G_n^{(X,Y)}(s) = \sqrt{n\overline{F}(u_n)} \left(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s) \right), \ s \geq s_0,$$

where

$$\widetilde{\beta}_n^{(X,Y)}(s) = \frac{1}{n\overline{F}(u_n)} \sum_{j=1}^{n} \mathbb{I}_{\{X_j > su_n, Y_j > su_n\}}.$$

We deduce that

$$\widehat{G}_n^{(X,Y)} = G_n^{(X,Y)}(sX_{n:n-k}/u_n).$$

We have

$$\begin{aligned}
\widehat{G}_n^{(X,Y)}(s) &= \sqrt{k} \left(\widehat{\beta}_n^{(X,Y)}(s) - \beta_n^{(X,Y)}(sX_{n:n-k}/u_n) \right) \\
&\quad + \sqrt{k} \left(\beta_n^{(X,Y)}(sX_{n:n-k}/u_n) - \beta^{(X,Y)}(sX_{n:n-k}/u_n) \right) + \sqrt{k} \left(\beta^{(X,Y)}(sX_{n:n-k}/u_n) - \beta^{(X,Y)}(s) \right) \\
&= G_n^{(X,Y)}(sX_{n:n-k}/u_n) + J_1(s) + J_2(s).
\end{aligned}$$

By Theorem 3.1.3, Remark 4.5.2 and bearing in mind the joint convergence, the first term converges weakly to $G^{(X,Y)}(\cdot)$. The second term $J_1(s)$ vanishes by assumption (4.3). For $J_2(s)$, we apply the delta method to (4.39) to have

$$\sqrt{k} \left\{ \beta^{(X,Y)}(sX_{n:n-k}/u_n) - \beta^{(X,Y)}(s) \right\} \to -\beta^{(X,Y)}(s) H^X(1).$$

This leads to the desired result. □

Remark 4.5.4 (Comparison between Theorem 4.1.1 and Theorem 4.5.3) *For Theorem 4.1.1:*

$$\sqrt{n\overline{F}(u_n)} \left(\widetilde{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s) \right) \Rightarrow G_*^{(X,Y)}(s) := G^{(X,Y)}(s) - \beta^{(X,Y)}(s) H^X(1).$$

For Theorem 4.5.3:

$$\sqrt{k} \left(\widehat{\beta}_n^{(X,Y)}(s) - \beta^{(X,Y)}(s) \right) \Rightarrow G_*^{(X,Y)}(s) := G^{(X,Y)}(s) - \beta^{(X,Y)}(s) H^X(1).$$

The limits are the same. As a consequence, the estimator of the tail dependence coefficient with deterministic levels and random normalization ($\widetilde{\beta}_n^{(X,Y)}$) has the same limit as the estimator with random levels and deterministic normalization ($\widehat{\beta}_n^{(X,Y)}$).

Now we want to use the above result to derive the asymptotic distribution of the corresponding sample extremogram matrix.
Recall the matrix $\widetilde{\Delta}_n(s)$ from (4.21). Consider the following matrices:

$$\widehat{\Delta}_n(s) = \begin{pmatrix} \widehat{\beta}_n^{(X,X)}(s) & \widehat{\beta}_n^{(X,Y)}(s) \\ \widehat{\beta}_n^{(X,Y)}(s) & \widehat{\beta}_n^{(Y,Y)}(s) \end{pmatrix}, \quad \widehat{\Delta}_n(1) = \begin{pmatrix} 1 & \widehat{\beta}_n^{(Y,X)}(1) \\ \widehat{\beta}_n^{(X,Y)}(1) & 1 \end{pmatrix}.$$

Note that $\widehat{\Delta}_n(s)$ is just $\widetilde{\Delta}_n(s)$, where $\widetilde{\beta}_n^{(X,Y)}$ is replaced with $\widehat{\beta}_n^{(X,Y)}$. By Theorem 4.5.3, we have

$$\sqrt{k}\left\{\text{vec}(\widehat{\Delta}_n(s)) - \text{vec}(\Delta(s))\right\} \Rightarrow \text{vec}(\mathcal{H}(s)),$$

where

$$\mathcal{H}(s) = \begin{pmatrix} G_*^{(X,X)}(s) & G_*^{(X,Y)}(s) \\ G_*^{(Y,X)}(s) & G_*^{(Y,Y)}(s) \end{pmatrix}$$

and

$$G_*^{(U,V)}(s) = G^{(U,V)}(s) - \beta^{(U,V)}(s)H^U(1),$$

defined in (4.23). Recall that

$$H_0: \ \beta^{(X,Y)}(\cdot) = 0 \quad vs \quad H_1: \ \beta^{(X,Y)}(\cdot) \neq 0.$$

Under the null hypothesis, keeping in mind that under H_0, the limiting process is degenerated, we have

$$\Delta_0(s) = \begin{pmatrix} s^{-\alpha} & 0 \\ 0 & s^{-\alpha} \end{pmatrix}, \quad \mathcal{H}_0(s) = \begin{pmatrix} G_*^{(X,X)}(s) & 0 \\ 0 & G_*^{(Y,Y)}(s) \end{pmatrix}.$$

Now, we can extend Theorems 4.3.4 and 4.4.2 to random levels.

Theorem 4.5.5 *Assume that the conditions of Theorem 4.5.3 are satisfied. Let $\widehat{\lambda}_{n(1)}(s) \geq \widehat{\lambda}_{n(2)}(s)$ be the ordered eigenvalues of $\widehat{\Delta}_n(s)$ and $\lambda_{(1)}(s) = \lambda_{(2)}(s) = s^{-\alpha}$ be the eigenvalues of $\Delta(s)$. Assume that*

$$\lim_{n\to+\infty} \sqrt{k} \sup_{s>\epsilon>0} \left|\beta_n^{(X,Y)}(s) - \beta^{(X,Y)}(s)\right| = 0.$$

Then under H_0 we have

$$\sqrt{k}\left(\widehat{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\right) \Rightarrow \max\left(G_*^{(X,X)}(s), G_*^{(Y,Y)}(s)\right),$$

where $G_*^{(X,X)}(s) = H^X(s) - s^{-\alpha}H^X(1)$ and $G_*^{(Y,Y)}(s) = H^Y(s) - s^{-\alpha}H^Y(1)$.

Theorem 4.5.6 *Let $X_j, j \geq 1$ be a random sequence as in Definition 4.4.1. Let $\widehat{\lambda}_{n(1)}(s) \geq \ldots \geq \widehat{\lambda}_{n(d)}(s)$ be the ordered eigenvalues of $\widehat{\Delta}_n(s)$ and $\lambda_{(1)}(s) = \ldots = \lambda_{(d)}(s) = s^{-\alpha}$ be the eigenvalues of $\Delta(s)$. Assume that for all $m, \ell = 1, \ldots, d$ we have*

$$\lim_{n\to+\infty} \sqrt{k} \sup_{s>\epsilon>0} |\beta_n^{(m,\ell)}(s) - \beta^{(m,\ell)}(s)| = 0.$$

Then, under H_0 we have

$$\sqrt{k}\left(\widehat{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\right) \Rightarrow \max_{1\leq\ell\leq d}\left(G_*^{(\ell,\ell)}(s)\right),$$

where $G_*^{(\ell,\ell)}(s)$, $\ell = 1, \cdots, d$, are as in Theorem 4.4.2.

Hypothesis testing with random levels We run the same simulations as in Table 4.1 using the test statistic obtained wih random levels for the models used therein. Indeed, we calculate the test statistics for $s \in [s_0, +\infty)$, $s_0 \in (0, 1)$

$$T_2(s) = \sqrt{k}\left(\widehat{\lambda}_{n(1)}(s) - \lambda_{(1)}(s)\right).$$

The value of k ranges from $0.01n$ to $0.1n$, and $s = 1.2$ (see Table (4.2)) and $s = 1.3$ (see Table (4.3)). As we can see, the simulated results confirm the theoretical results obtained in this section. Indeed, for independent Pareto, we fail to reject H_0, while for the dependent Pareto (Example 2.1.17) and the third model mentioned above, we reject the null hypothesis.

Model(True TDC)	k	0.01n	0.02n	0.03n	0.04n	0.05n	0.06n	0.07n	0.08n	0.09n	0.1n
Ind. Pareto (0)	T_2	0	0	0	0	0	0.13	0.24	0.34	0.32	0.30
	CI	+	+	+	+	+	+	+	+	+	+
Dep. Pareto (0.5)	T_2	0.63	0.89	0.73	0.79	0.85	0.90	1.08	1.45	1.37	1.41
	CI	-	-	-	-	-	-	-	-	-	-
Third model (0.8^4)	T_2	0.32	0.67	0.91	1.11	1.27	1.29	1.44	1.68	1.79	1.81
	CI	+	-	-	-	-	-	-	-	-	-

Table 4.2: Testing for extremal independence. The 95% asymptotic confidence interval for $s = 1.2$ is CI=$\Big(M_{.025}, M_{.975}\Big) = \Big(-0.25; 0.56\Big)$. The "+" sign indicates that T_2 belongs to the CI and the numbers in parentheses correspond to the true TDC.

Model(True TDC)	k	0.01n	0.02n	0.03n	0.04n	0.05n	0.06n	0.07n	0.08n	0.09n	0.1n
Ind. Pareto (0)	T_2	0	0	0	0	0	0	0	0.11	0.11	0.30
	CI	+	+	+	+	+	+	+	+	+	+
Dep. Pareto (0.5)	T_2	0.63	0.44	0.73	0.63	0.57	0.65	0.60	0.67	0.63	0.70
	CI	-	+	-	-	-	-	-	-	-	-
Third model (0.8^4)	T_2	0	0.22	0.73	0.79	0.99	1.03	1.08	1.12	1.16	1.30
	CI	+	+	-	-	-	-	-	-	-	-

Table 4.3: Testing for extremal independence. The 95% asymptotic confidence interval for $s = 1.3$ is CI=$\Big(M_{.025}, M_{.975}\Big) = \Big(-0.23, 0.51\Big)$. The "+" sign indicates that T_2 belongs to the CI and the numbers in parentheses correspond to the true TDC.

4.6 Implementation: Simulation studies

In this section, we perform some simulations studies to support our theoretical results. We deal with the estimation of the tail dependence coefficient defined in (2.4) as

$$\lambda_{\text{TDC}} = \lim_{x\to\infty} \frac{\mathbb{P}(X > x, Y > x)}{\overline{F}(x)},$$

where X and Y have the same distribution. We make use of the estimator $\widehat{\beta}^{(X,Y)}(1)$ defined in (4.33). The estimator is computed for different values k, where k is the number of order statistics and plotted against the order statistics $X_{n:1}, \ldots, X_{n:n}$, being arranged in increasing order. The choice of the threshold k is not addressed here but can be found in the literature (e.g. [10]). The estimator captures the dependence

structure.

Example 4.6.1 (Independent Pareto) *We simulate 1000 independent observations from* (X, Y)*, where* X *and* Y *are independent Pareto-distributed with* $\alpha = 4$*. In this case the tail dependence coefficient is just* 0*. Figure 2.2 shows the estimated values of the tail dependence coefficient computed using the estimator* (4.33) *for different values of* k*, where* k *is the number of order statistics used.*

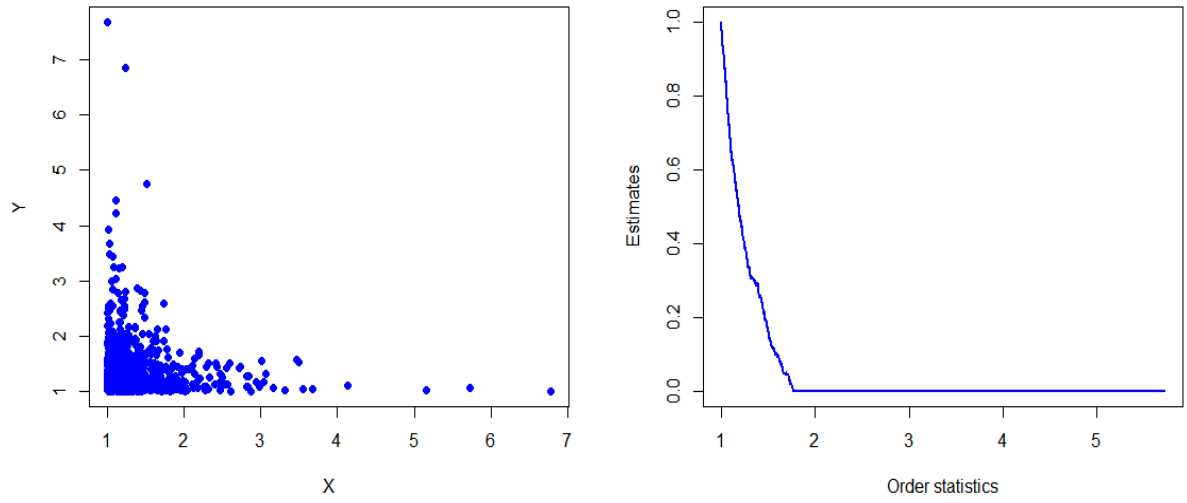

Figure 4.3: The scatter plot (left panel) and the estimator of the TDC (right panel) for X and Y drown independently from Pareto distribution with $\alpha = 4$.

Example 4.6.2 (Dependent Pareto) *We simulate 1000 independent observations from the model* (X, Y)*, where* $Y = Z_2 + Z_3$*,* $X = Z_1 + Z_2$*, and* Z_i*,* $i = 1, 2, 3$ *are independent Pareto with* $\alpha = 4$*. In this case the tail dependence coefficient is* 0.5*. The results are displayed on Figure 4.4. This dependence is captured by the estimator.*

Example 4.6.3 (Bivariate t**)** *Figure 4.5 shows the estimate of the tail dependence coefficient for the bivariate* t *distribution i.e* $(X, Y) = \sqrt{Z}(|U_1|, |U_2|)$*, where* α/Z *is chi-square distributed with* $\alpha = 4$ *degrees of freedom and* U_1*,* U_2 *are standard normal with correlation* $\rho = 0.9$*. The tail dependence coefficient in this case is* 0.63*, see [5]. The scatter plot indicates strong dependence in the upper and lower tail which is confirmed by the estimation of the tail dependence coefficient.*

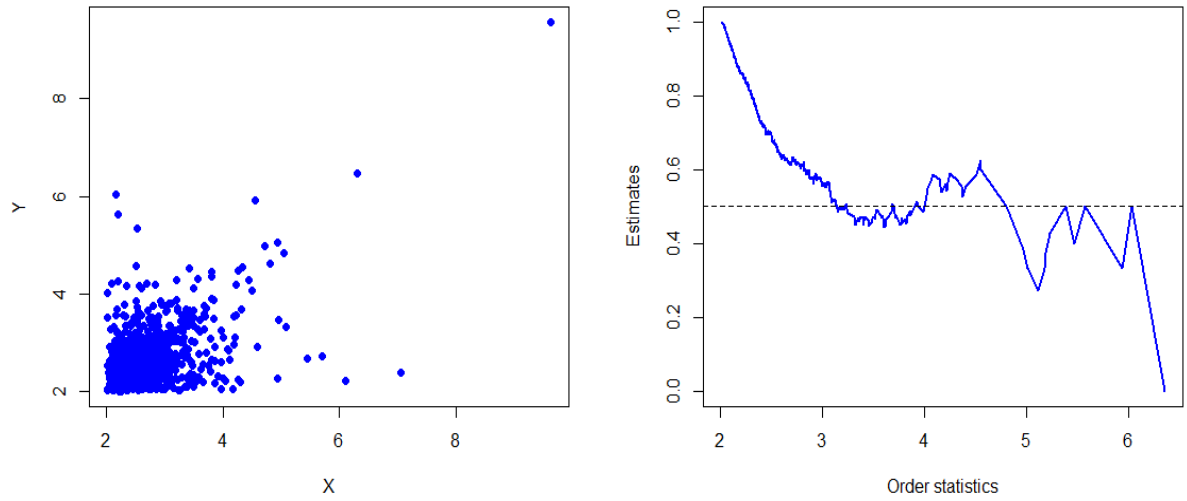

Figure 4.4: The scatter plot (left panel) and the estimator of the TDC (right panel) for the model $Y = Z_2 + Z_3$, $X = Z_1 + Z_2$, where Z_i, $i = 1, 2, 3$ are independent Pareto with $\alpha = 4$.

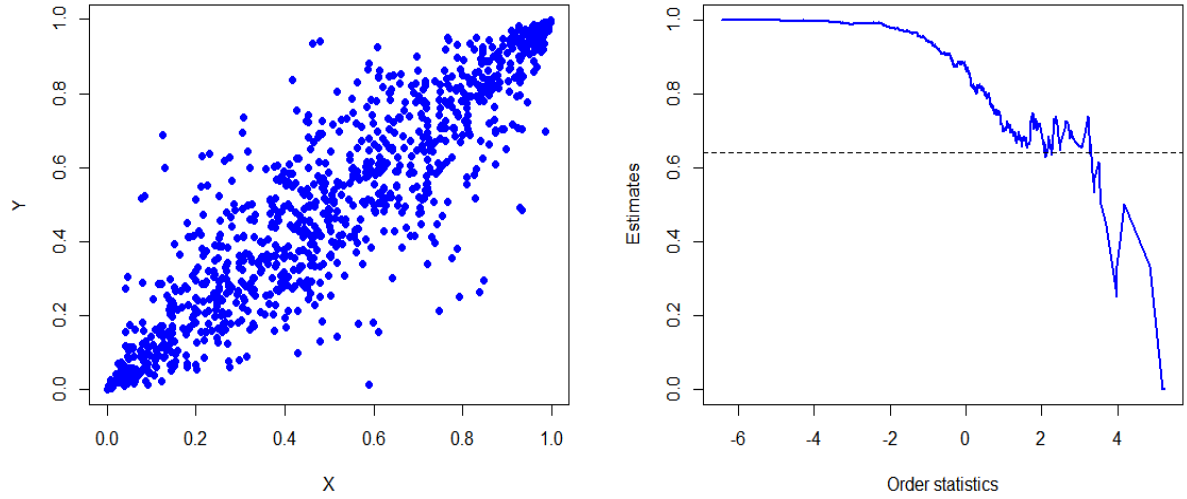

Figure 4.5: The bivariate t distribution: the scatter plot (left panel) and the estimator of the TDC (right panel) plotted against the order statistics.

4.7 Real data analysis

In this section we apply our method to two financial data sets.

Example 4.7.1 (Stock prices data set) *The first data set contains the absolute log returns of daily stock*

prices for S&P500 and NASDAQ from January 20, 2015 to September 19, 2017. The data set consist of 673 records (Source: Yahoo Finance). The QQ-plots of these data sets, see Figure 4.7, show a tilted S-shape which stretches out in the extremes indicating the heaviness of the tail of the data. In other words, they have fatter tails than normal distribution. Hence, we cannot assume normality. Moreover, the scatter plot, see Figure 4.8 (top left panel), indicates a strong dependence in the upper tail, which is confirmed by the estimation of the tail dependence coefficient, Figure 4.8 (bottom right panel). Furthermore, on Figure 4.8 we estimate the tail index for S&P500 and NASDAQ using the classical Hill estimator. We can assume that they have the same tail index and its estimate was chosen as 2.7. *The assumption that the two datasets come approximatively from the same populations (in the tails) are verified in Figure 4.8 (bottom left panel). Hypothesis testing for extremal independence for this data set is performed and the results are summarized in Table 4.4.*

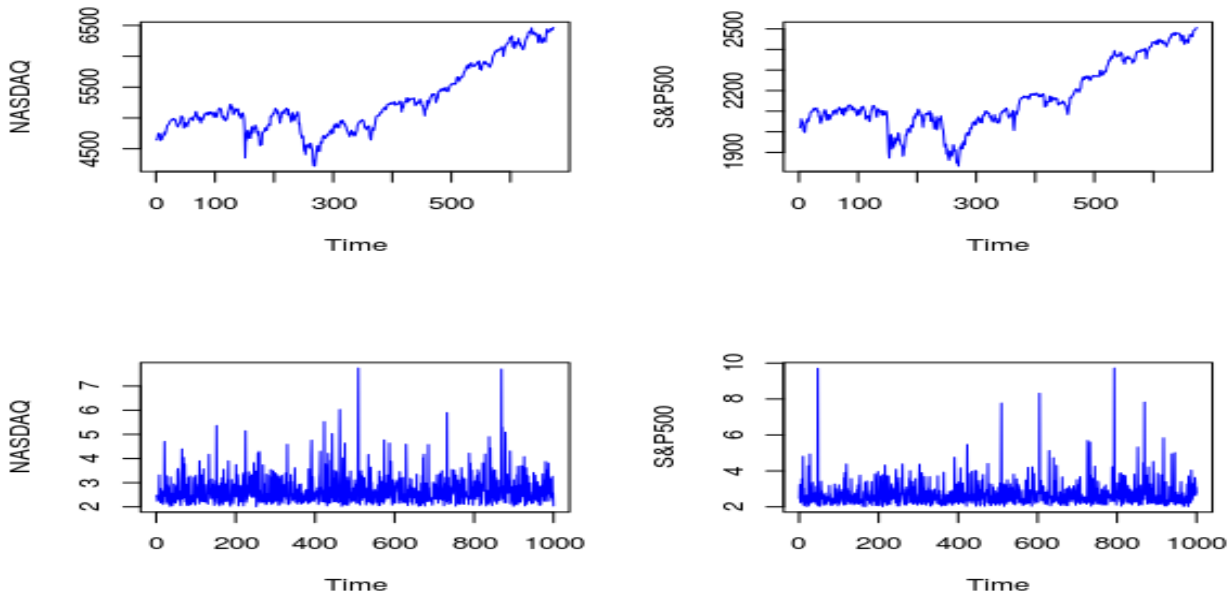

Figure 4.6: Time series plot of NASDAQ (left panel) and S&P500 (right panel).

Example 4.7.2 (Exchange-rates data set) *The second data set contains the daily exchange-rates of the German and French currencies against the US dollar for the time period 1979-1998. The data set consist of 4774 records (see datamarket.com). Figure 4.10 shows again that we have heavy-tailed data sets. Moreover, the scatter plot, see Figure 4.11 (top left panel), indicates strong dependence in the upper tail, which is*

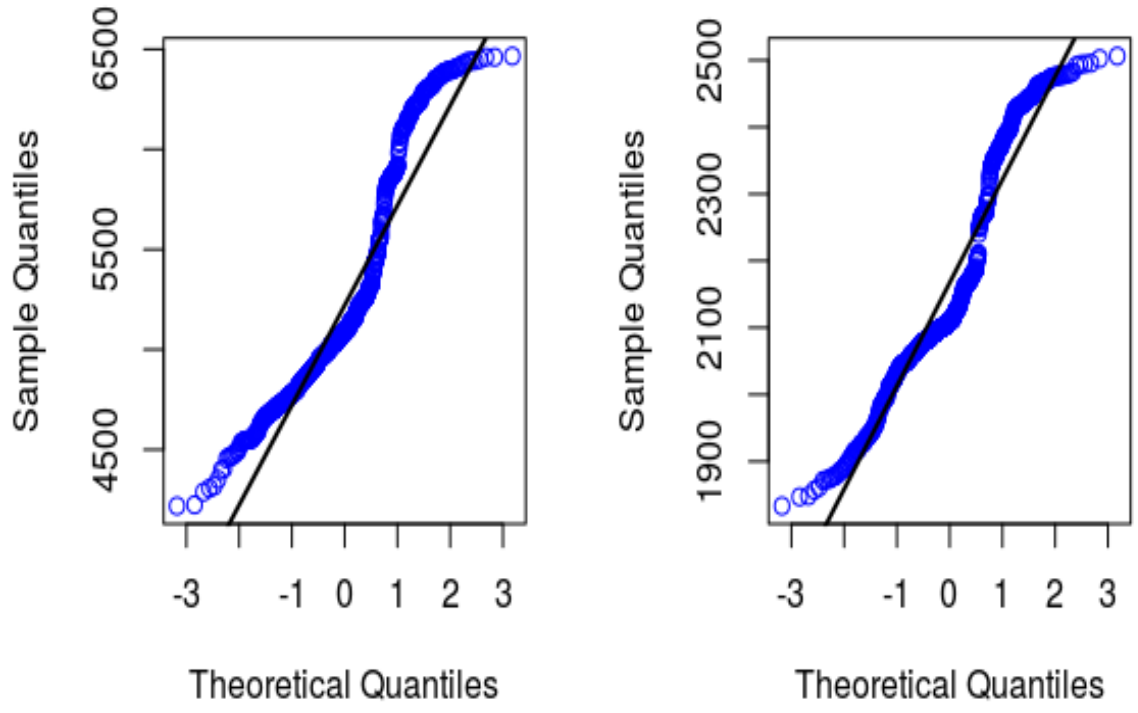

Figure 4.7: QQ-plot of NASDAQ (left panel) and S&P500 (right panel).

s value	k	0.01n	0.02n	0.03n	0.04n	0.05n	0.06n	0.07n	0.08n	0.09n	0.1n
s=1.2	$T_2(1.2)$	3.16	4.47	5.48	6.32	7.07	7.75	8.31	8.94	9.49	9.95
	CI(1.2)	-	-	-	-	-	-	-	-	-	-

Table 4.4: Testing for extremal independence for the stock prices data set. The 95% asymptotic confidence interval is CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.2)\Big) = \Big(-0.24; 0.53\Big)$. The "-" sign indicates that T_2 doesn't belong to the CI.

confirmed by the estimation of the TDC, Figure 4.11 (bottom right panel). Furthermore, on Figure 4.11 we estimate the tail index for the German and French currencies using the classical Hill estimator. We can assume that they have the same tail index and its estimate was chosen as 3. *The assumption that the two datasets come approximatively from the same populations (in the tails) are verified in Figure 4.8 (bottom left panel).*

Hypothesis testing for extremal independence for this data set is performed and the results are summarized in Table 4.5.

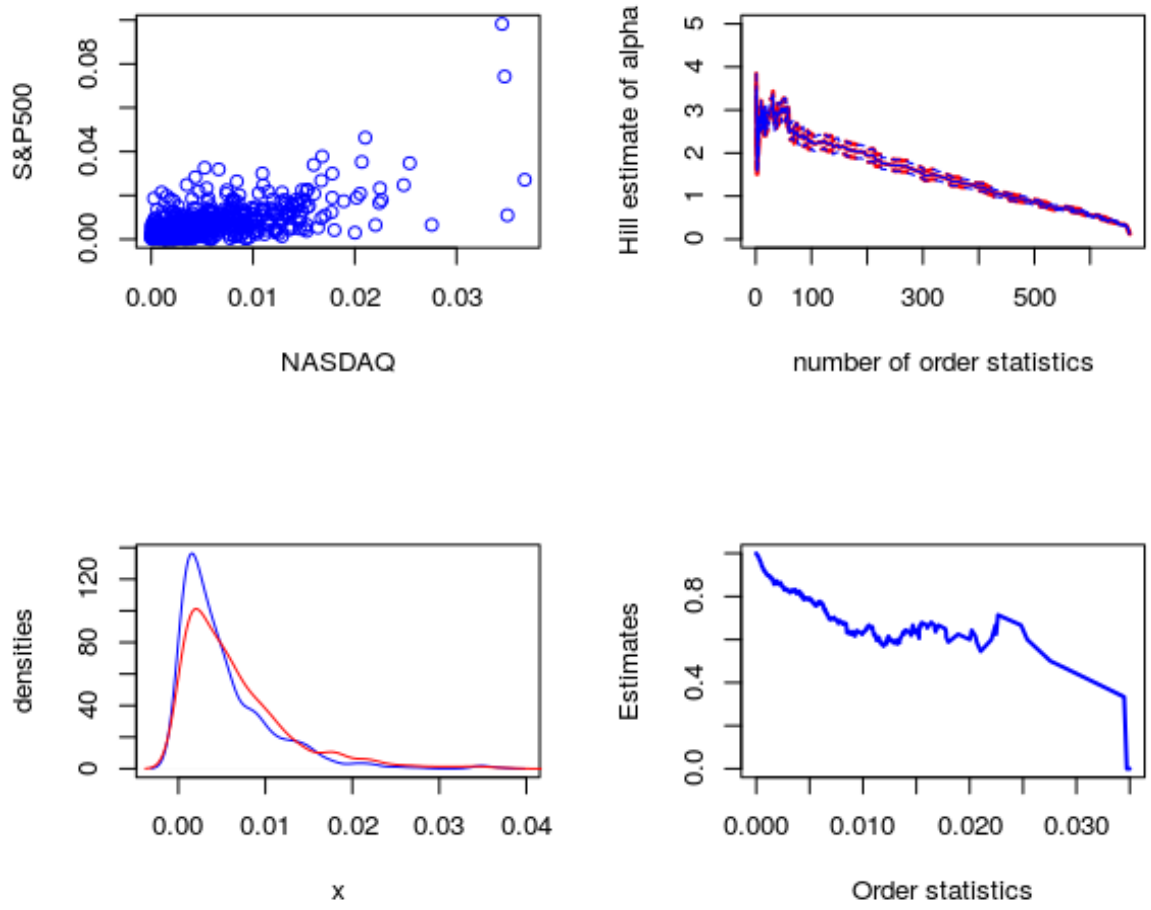

Figure 4.8: The absolute log returns of daily stock prices for S&P500 vs. NASDAQ data set: the scatter plot (top left panel), Hill plot (top right panel) of NASDAQ (red dashed line) and S&P 500 (blue dashed line) with CI, the densities (bottom left panel) and the estimator of the TDC (bottom right panel).

s value	k	0.01n	0.02n	0.03n	0.04n	0.05n	0.06n	0.07n	0.08n	0.09n	0.1n
s=1.2	$T_2(1.2)$	1.17	3.18	4.93	5.44	6.81	7.04	8.03	8.91	9.95	10.89
	CI(1.2)	-	-	-	-	-	-	-	-	-	-

Table 4.5: Testing for extremal independence for the exchange-rates data set. The 95% asymptotic confidence interval is CI(1.2)=$\Big(M_{.025}(1.2), M_{.975}(1.2)\Big) = \Big(-0.25; 0.55\Big)$. The "+" sign indicates that T_2 belongs to the CI.

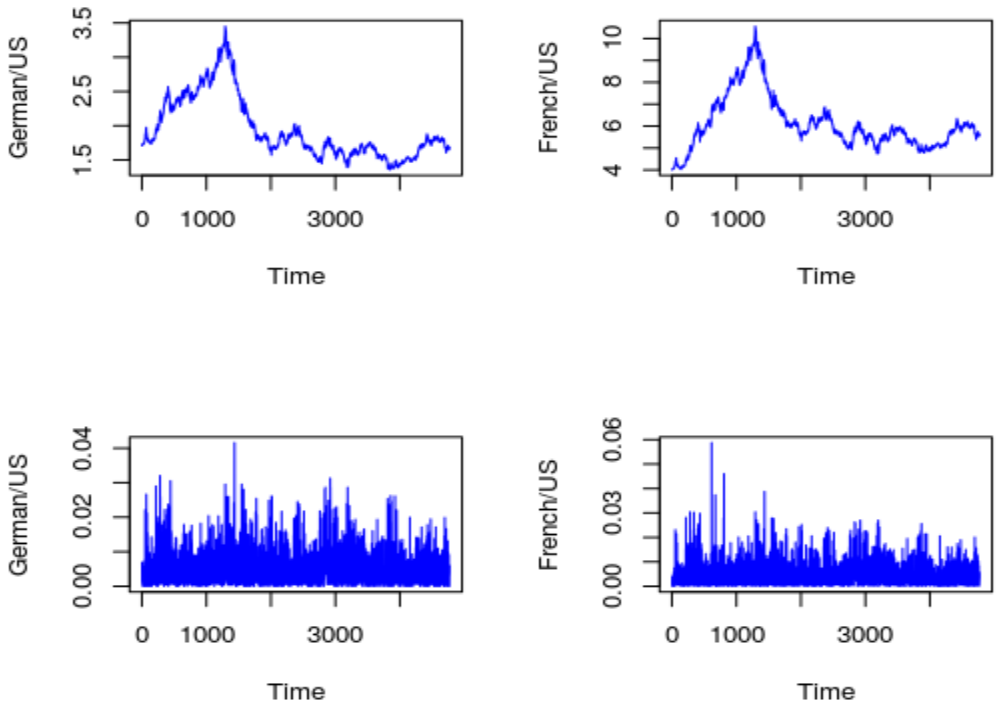

Figure 4.9: Time series plot of German Mark (left panel) and French Franc (right panel).

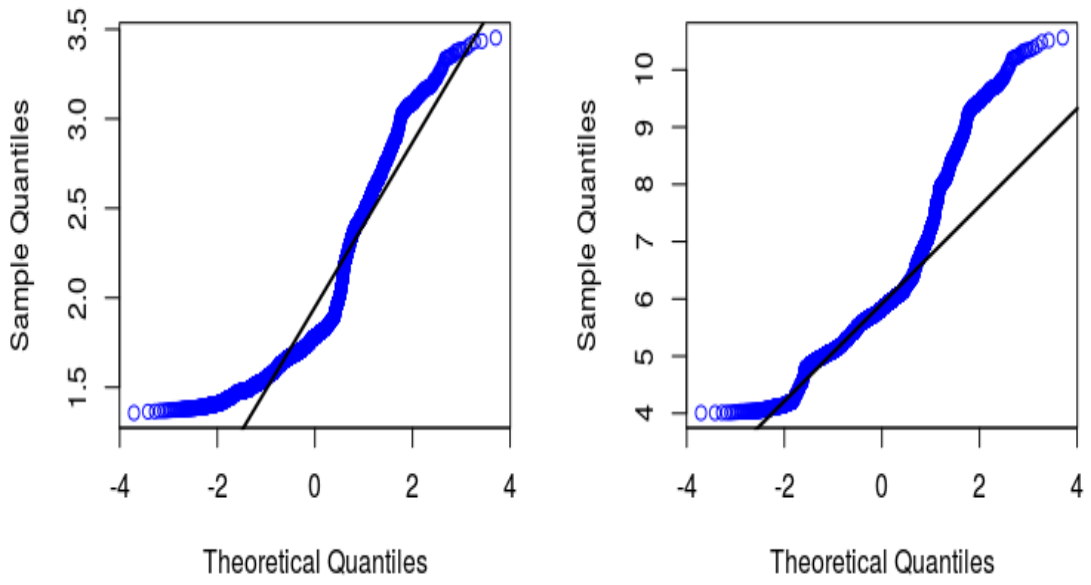

Figure 4.10: QQ-plot of German Mark (left panel) and French Franc (right panel).

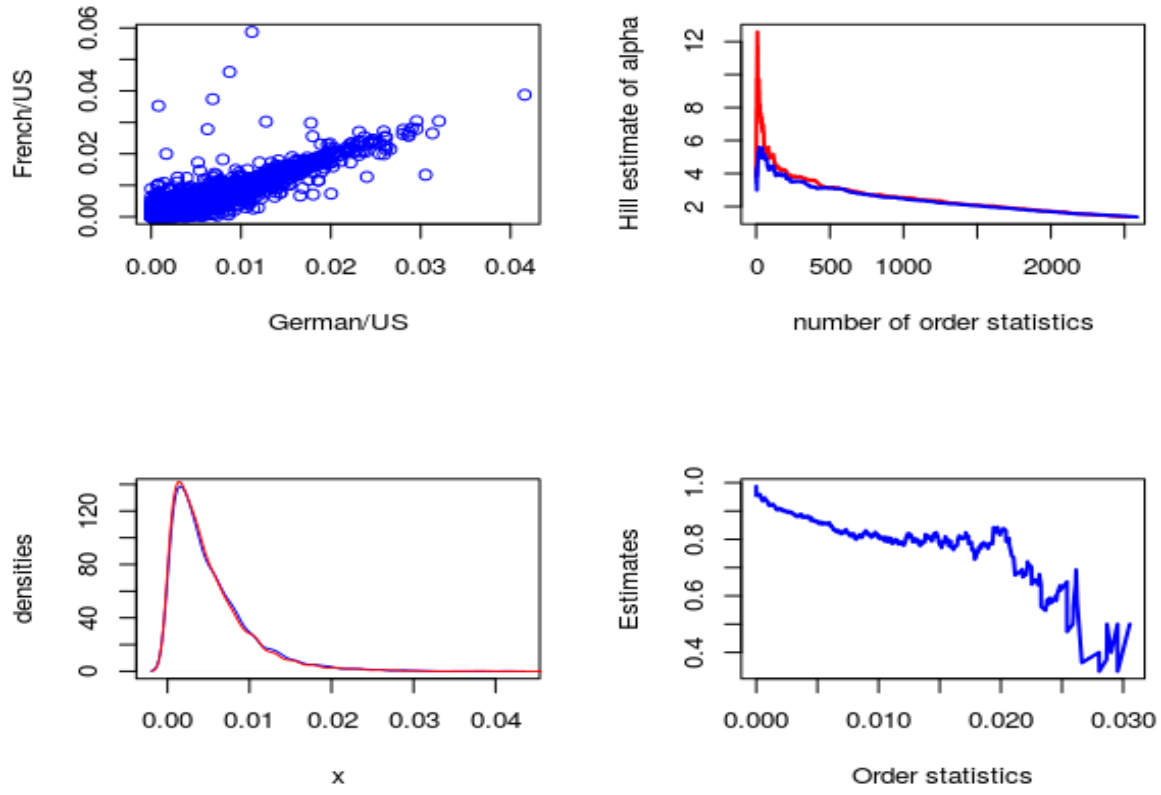

Figure 4.11: The absolute log returns of daily exchange-rates for German Mark and French Franc data set: the scatter plot (top left panel), Hill plot (top right panel) of German Mark (red line) and French Franc (blue line), the densities (bottom left panel) and the estimator of the TDC (bottom right panel).

Conclusion and further work

The tail dependence coefficient is a good tool to describe the amount of dependence in the extremes as it is shown in this paper. We propose a nomparametric estimator of the TDC and prove consistency and asymptotic normality. Since the limit of the estimator is degenerated then we cannot construct a formal test for extremal independence. To avoid this drawback, we consider an analog of the sample covariance matrix, namely the sample extremogram matrix of dimension d, whose entries depend only on the extremes i.e. the tail dependence coefficient. We work under the finite dimensional case, say $d \geq 2$, and we prove that the largest eigenvalue converges is distribution to the maximum of d independent Gaussian random variables. Having that in hand, we are now ready to conduct an hypothesis testing by means of the distribution of the largest eigenvalue of the sample extremogram matrix and the results are quite satifactory.

The extension of the extremogram matrix to higher dimension d that grows with n, that is $d = d_n \to +\infty$ as $n \to +\infty$, and $\lambda_{n(1)}(s) = \max\left(\widetilde{\beta}_n^{(1,1)}(s), \cdots, \widetilde{\beta}_n^{(d_n,d_n)}(s)\right)$ is under study.

Bibliography

[1] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.

[2] Patrick Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.

[3] Richard A. Davis and Thomas Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 11 2009.

[4] Richard A. Davis, Thomas Mikosch, and Oliver Pfaffel. Asymptotic theory for the sample covariance matrix of a heavy-tailed multivariate time series. *Stochastic Process. Appl.*, 126(3):767–799, 2016.

[5] Stefano Demarta and Alexander McNeil. The t copula nad related copulas. *Unpublished manuscript*, 2004.

[6] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Annals of statistics*, pages 295–327, 2001.

[7] Rafal Kulik and Philippe Soulier. The tail empirical process for long memory stochastic volatility sequences. *Stochastic Process. Appl.*, 121(1):109–134, 2011.

[8] Saralees Nadarajah and Samuel Kotz. Exact distribution of the max/min of two gaussian random variables. *IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS*, VOL. 16(NO. 2), FEBRUARY 2008.

[9] S. Resnick. *Heavy-Tail Phenomena:Probabilistic and Statistical Modelling*. Springer, 2007.

[10] Rafael Schmidt and Ulrich Stadtmüller. Non-parametric estimation of tail dependence. *Scand. J. Statist.*, 33(2):307–335, 2006.

[11] E. Seneta. *Functions of Regular Variation. Lecture Notes in Mathematics*. Springer, New York, 506 edition, 1976.

[12] Aad W. van der Vaart and Jon A. Wellner. *Weak convergence and empirical processes*. Springer, New York, 1996.