Regularly Varying Time Series with Long Memory: Probabilistic Properties and Estimation

by

Clémonell Lord Baronat
BILAYI-BIAKANA

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Ph.D. degree in
Mathematics and Statistics[1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

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

Contents

1 Introduction **3**

1.1 Motivation and Goals . . . 3

1.2 Contribution and Structure . . . 5

1.2.1 Chapter 2: Mathematical Foundations . . . 6

1.2.2 Chapter 3: LMSV . . . 6

1.2.3 Chapter 4: Tail Empirical Processes . . . 7

1.2.4 Chapter 5: Risk measures . . . 7

1.2.5 Chapter 6: Simulations . . . 8

1.2.6 Chapter 7: Conclusion . . . 8

1.2.7 Summary of the contribution . . . 8

2 Mathematical Foundations **9**

2.1 Introductory Comments . . . 9

2.2 General Inverses of Monotone Functions . . . 9

2.3 Regular Variation . . . 11

2.3.1 Slowly Varying Functions . . . 11

2.3.2 Regularly Varying Functions . . . 14

2.3.3 Regularly Varying Random Variables . . . 18

2.4 Second-Order Regular Variation . . . 20

2.4.1 Second-Order Slowly Varying Functions . . . 21

2.4.2 Second-Order Regularly Varying Functions . . . 25

2.4.3 Second-Order Regularly Varying Random Variables . . . 28

2.5 Weak Convergence of Probability Measures . . . 31

2.5.1 Random Elements of Metric Spaces . . . 31

2.5.2 Skorokhod Topology . . . 33

2.5.3 Convergence of Functionals of Stochastic Processes . . . 36

2.6 Stationary Processes . . . 37
2.6.1 Stationarity and Ergodicity . . . 37
2.6.2 Linear Processes . . . 38
2.6.3 Martingales . . . 38
2.7 Long Memory Processes . . . 39
2.7.1 Different Types of Memory . . . 40
2.7.2 ARMA and FARIMA Models . . . 41
2.7.3 Hermite Polynomials . . . 42
2.7.4 Limit Theorems for Gaussian Long Memory Processes . . . 47
2.8 Leverage Effect . . . 49
2.9 Concluding Remarks . . . 51

3 Long Memory Stochastic Volatility Model with Leverage 52
3.1 Introductory Comments . . . 52
3.2 Model: Description and Assumptions . . . 53
3.3 Transfer Theorems and Technicalities . . . 55
3.4 No-bias Conditions . . . 63
3.5 Examples . . . 65
3.6 Concluding Remarks . . . 67

4 Limit Theorems for the Tail Empirical Processes 69
4.1 Introductory Comments . . . 69
4.2 Deterministic Levels . . . 71
4.2.1 Methodology . . . 71
4.2.2 Weak Convergence of the Martingale Part . . . 72
4.2.3 Weak Convergence of the Long Memory Part . . . 89
4.2.4 Weak convergence of the TEP with Deterministic Levels . . . 92
4.3 Random Levels . . . 94
4.3.1 Weak Convergence of the TEP with Random Levels . . . 94
4.3.2 Weak Convergence of Integral Functionals . . . 101
4.3.3 Tail Index Estimation - Harmonic Moment Estimators . . . 108
4.4 Concluding Remarks . . . 111

5 Estimation of Financial Risk Measures 112
5.1 Introductory Comments . . . 112
5.2 Coherent Financial Risk Measures . . . 113

5.3 Value-at-Risk (VaR) 115
5.3.1 Estimation of VaR 115
5.3.2 Limit Theorems for VaR 116
5.4 Expected Shortfall (ES) 122
5.4.1 Estimation of Expected Shortfall 123
5.4.2 Limit Theorems for Expected Shortfall 124
5.5 Concluding Remarks 129

6 Simulation Studies **130**
6.1 Introductory Comments 130
6.2 Simulations for i.i.d data 131
6.2.1 Experiment 1: Hill Plots for Hill Estimator 131
6.2.2 Experiment 2: Hill Plots for Value-at-Risk 132
6.2.3 Experiment 3: Hill Plots for Expected Shortfall 132
6.3 Simulations for the LMSV Model 133
6.3.1 Experiment 1: Hill Plots for Hill Estimator 133
6.3.2 Experiment 2: Hill Plots for Value-at-Risk 133
6.3.3 Experiment 3: Hill Plots for Expected Shortfall 134
6.4 Monte Carlo Simulations for i.i.d data 134
6.4.1 Experiment 1: Boxplots for Hill Estimator 134
6.4.2 Experiment 2: Boxplots for Value-at-Risk 134
6.4.3 Experiment 3: Boxplots for Expected Shortfall 135
6.5 Monte Carlo Simulations for LMSV data 135
6.5.1 Experiment 1: Boxplots for Hill Estimator 135
6.5.2 Experiment 2: Boxplots for Value-at-Risk 136
6.5.3 Experiment 3: Boxplots for Expected Shortfall 136
6.6 Simulations for the LMSV Model with Leverage 137
6.7 Figures 138

7 Conclusion **165**
7.1 Comments on Assumptions and Technicalities 165
7.2 Further Research Directions 167

Acknowledgements

I am heartily grateful to Professors **Rafał Kulik** and **Gail Ivanoff** for being truly a blessing for me. I have been fortunate to have you both as supervisors. I have enjoyed every milestone of your guidance and learned so many things from you! I cannot say thank you enough for your tremendous support! You are not only great professional academics but also wonderful persons.

Abstract

We consider tail empirical processes for long memory stochastic volatility models with heavy tails and leverage. We show a dichotomous behaviour for the tail empirical process with fixed levels, according to the interplay between the long memory parameter and the tail index; leverage does not play a role. On the other hand, the tail empirical process with random levels is not affected by either long memory or leverage. The tail empirical process with random levels is used to construct a family of estimators of the tail index, including the famous Hill estimator and harmonic mean estimators. The limiting behaviour of these estimators is not affected by either long memory or leverage. Furthermore, we consider estimators of risk measures such as Value-at-Risk and Expected Shortfall. In these cases, the limiting behaviour is affected by long memory, but it is not affected by leverage. The theoretical results are illustrated by simulation studies.

Chapter 1

Introduction

1.1 Motivation and Goals

For financial data, such as returns on investments, we want to estimate extremal quantities such as the probability of exceeding a very high level, high quantiles, or so-called risk measures. Such quantities are of importance in other fields as well: for example, environmental science (e.g. flood control) and engineering (e.g. risk assessment).

Financial data typically exhibit three widely accepted features not reflected in classical statistical models. Returns on investments are uncorrelated, but their squares, or absolute values, are (highly) correlated. Such behaviour is known as *long range dependence or long memory.* Log-returns are *heavy tailed*, that is - some moments of the log-returns are infinite. Finally, past returns and future volatility are negatively dependent. This phenomenon is referred to as the *leverage effect.* Typically, rising asset prices are accompanied by declining volatility, and vice versa. The leverage effect has been well documented in the economics literature. *Any mathematical model approximating the evolution of asset price should be able to generate long memory, heavy tails and the leverage effect.* This can be done through the use of *stochastic volatility models.*

In this class of stochastic processes, log-returns $\{X_j\}$ are modelled as follows:

$$X_j = \sigma_j Z_j,$$

where $\{Z_j\}$ is a sequence of independent, identically distributed (i.i.d.) random variables and $\{\sigma_j^2\}$ is the conditional variance or, more generally, a certain process which describes the volatility. In such a process, long memory is typically modelled through the sequence $\{\sigma_j\}$, while the tails can be modelled either through the sequence $\{Z_j\}$ or through $\{\sigma_j\}$, or both. The well known GARCH processes (Nobel Prize in Economics, 2003) belong to this class of models. In this case, the volatility sequence $\{\sigma_j\}$ is heavy tailed unless the distribution of Z_0 has finite support, and leverage can be present. However, long memory of squares cannot be modelled by the GARCH process. See [18] for more details on GARCH models.

Consequently, to capture long range dependence, the so-called long memory stochastic volatility (LMSV) model was introduced in [15]. An overview of stochastic volatility models with long range dependence and their basic properties is given in [24] and in [25]. In the original LMSV model, $\{Z_j\}$ is a sequence of i.i.d. standard normal random variables, independent of the volatility sequence $\{\sigma_j\}$, assumed to be of the form $\sigma_j = \exp(Y_j)$, where $\{Y_j\}$ is a long memory Gaussian sequence. However, the independence assumption excludes the possibility of modelling leverage effects.

Thus, motivated by the discussion above, in this thesis we consider the long memory stochastic volatility model with leverage by allowing the sequences $\{Z_j\}$ and $\{\sigma_j\}$ to be dependent. Heavy tails are modelled through the sequence $\{Z_j\}$ and long memory through the sequence $\{\sigma_j\}$ which is of a general form $\sigma_j = \phi(Y_j)$, where ϕ is a measurable function and $\{Y_j\}$ is a long memory Gaussian sequence. The model allows a general dependence structure between $\{Y_j\}$ and $\{Z_j\}$.

To address the problem of estimating extremal values, we consider the so-called *tail empirical process* (TEP), a variation of the classical empirical process that takes into account only large values. These limiting results are not only of theoretical interest, but are applicable to different statistical procedures based on intermediate extremes. It should be noted that the mathematical theory of the TEP in the case of dependent random variables is much more involved than that of the usual empirical process and has only been studied since the beginning of the 21st century. Indeed, in the case of independent, identically distributed random variables, the asymptotic theory is given in [34]. The corresponding theory for weakly dependent sequences is considered in [31], [30], [29], [48], [42], [9], often under ad-hoc conditions. The most advanced theory is presented in [32]. See also [23] for an extensive review in the i.i.d. case.

In this thesis, our goal is to study weak convergence of the tail empirical processes associated with heavy tailed long memory stochastic volatility sequences with leverage. In [39] the authors considered heavy tailed, long memory stochastic volatility models and obtained asymptotic results for the tail empirical processes. This was extended later to the multiparameter situation in [40]. However, in the latter two articles leverage was excluded, greatly simplifying theoretical considerations. As evidenced in [41], the presence of long memory, heavy tails and leverage may greatly affect the limiting behaviour of relevant statistics.

As such, this thesis can be viewed as an extension of [39]. However, it should be emphasized that the presence of leverage in the model creates additional theoretical challenges.

1.2 Contribution and Structure

In the setting described above, we show a dichotomous limiting behaviour for the tail empirical process that depends on the interplay between the strength of long memory and the heaviness of the tails. Surprisingly, the effect of leverage is negligible in the limit and hence the results are comparable to those in [39] where leverage is not present. **The extension of the asymptotic theory for tail empirical processes to the model with leverage is the first major contribution of the thesis.**

However, it should be pointed out clearly that the extension from models without leverage to those with leverage is highly nontrivial from a theoretical point of view. In [39] the authors were able to exploit the conditional independence of the sequence $\{X_j\}$ given $\{Y_j\}$. Here this approach is not applicable and instead we use the Doob decomposition of the tail empirical process into martingale and long memory parts. This makes the proof of tightness technically very involved.

The TEP is then used to produce estimators of various quantities related to extremal values. The first one, the so-called tail index, measures the heaviness of the tail of the distribution of X_0. The TEP is used to construct a family of estimators of the tail index, including the famous Hill estimator (see [23] for results in the i.i.d. case) and harmonic mean estimators (see [7] again for results in the i.i.d. case). Surprisingly, as already noted in [39], the effect of long memory vanishes in the limit. This is very important from a practical perspective.

The results on the tail empirical processes and tail index estimation are included in the paper [8].

Furthermore, we use the tail empirical process to construct estimators for two financial risk measures: Value-at-Risk and Expected Shortfall. It turns out that, in contrast to the estimators of the tail index, the limiting behaviour is very complicated and depends on an interplay between long memory and heaviness of the tails. These results are new and thus **the asymptotic behaviour of the estimators of the risk measures in the case of LMSV models with leverage is the second major contribution of this thesis.**

This thesis consists of seven chapters and is organized as follows.

1.2.1 Chapter 2: Mathematical Foundations

Throughout Chapter 2, we present the mathematical background that will be used in the remainder of the thesis. These concepts and tools include regular variation, second-order regular variation, weak convergence, second-order stationary processes, long memory processes (Sections 2.3 to 2.7), etc. This chapter ends with a discussion of the leverage effect (Section 2.8), which, as noted before, is an important feature exhibited by financial time series. It is worth mentioning that we have revisited the notion of second-order regular variation (Section 2.4). **This is the major contribution in this chapter** and it warrants a future publication.

The main references for this chapter are: [13, 23] (regular variation), [10, 11, 12, 51, 22] (weak convergence), [17, 5, 6] (long memory).

1.2.2 Chapter 3: LMSV

In Chapter 3, we introduce the long memory stochastic volatility model (LMSV) with leverage; see (3.1). We describe the model and state the relevant assumptions (Section 3.2). The main contributions of this chapter are the so-called transfer theorems (Section 3.3) and results related to no-bias conditions (Lemmas 3.4.1 and 3.4.2). Some existing results such as Lemmas 3.3.3 and 3.3.7 have been adapted to the second-order regular variation framework of this thesis. We illustrate these assumptions via two examples in Section 3.5.

1.2.3 Chapter 4: Tail Empirical Processes

In Chapter 4, we look into the limiting behaviour of the tail empirical processes associated with the LMSV model with leverage. Our contribution in this chapter is twofold. **From a theoretical point of view, the most important contribution is the proof of weak convergence of the tail empirical process (with fixed and random levels) in the presence of heavy tails, long memory and leverage**; see Theorems 4.2.18 and 4.3.4. Due to the complicated dependence structure of the process, the proof is not at all straightforward. From a practical point of view, the key result is that the asymptotic behaviour of the tail empirical process with random levels is unaffected by the presence of long memory and/or leverage in the model, and so in certain applications log-returns $\{X_j\}$ may be handled exactly as if they were i.i.d. heavy-tailed random variables. This greatly enhances the utility of the LMSV model with leverage considered here.

The limiting behaviour of integral functionals of the TEP is considered in Theorem 4.3.9. These functionals are used to construct a family of estimators of the tail index, including the famous Hill estimator and harmonic mean estimators (Section 4.3.3), whose asymptotic normality is studied in Theorem 4.3.16.

The results from this chapter are published in [8].

1.2.4 Chapter 5: Risk measures

In Chapter 5, we estimate financial risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES). This is done under the assumption that returns of a portfolio are heavy-tailed long memory sequences with leverage. We define the estimators of VaR in (5.4a)-(5.4b) and subsequently study their asymptotic behaviour in Proposition 5.3.2 and Theorems 5.3.3 and 5.3.5. Furthermore, we study the estimators of Expected Shortfall in Proposition 5.4.5 and Theorems 5.4.6 and 5.4.8. It turns out that the limiting behaviour of these estimators is very complicated and depends on a fine interplay between long memory and tails. This stems from the fact that the limiting behaviour of these estimators needs to take into account asymptotics of both intermediate order statistics and estimators of the tail index.

All the results presented in this chapter are new in the context of long memory models.

1.2.5 Chapter 6: Simulations

To illustrate our theory, we perform extensive numerical studies in Chapter 6.

1.2.6 Chapter 7: Conclusion

Chapter 7 concludes the thesis. In Section 7.1, we discuss in detail the technical assumptions imposed on the LMSV model. We propose future research directions in Section 7.2.

1.2.7 Summary of the contribution

- The results on second-order regular variation (Section 2.4) are partially new and warrant a short publication.
- Most of the results in Section 3.3 are new except Lemmas 3.3.3 and 3.3.7.
- The results in Chapter 4 are new and have been accepted for publication in the *Electronic Journal of Statistics* ([8]).
- The results in Chapter 5 are completely new and are derived from the tools developed in Chapters 3 and 4. We are preparing a publication based on these results.
- The theoretical results are illustrated by simulation studies in Chapter 6.

Chapter 2

Mathematical Foundations

2.1 Introductory Comments

This chapter provides the mathematical background and tools required in the upcoming chapters. The sort of analytical tools that we are mainly concerned with are regular variation, second-order regular variation, weak convergence of probability measures and Hermite polynomials.

This chapter consists of the following sections. In section 2.3, we discuss regular variation in terms of both real valued functions and random variables. In section 2.4, we introduce a stronger property - second-order regular variation. In section 2.5, we present some existing results about weak convergence of probability measures. In section 2.6, we briefly discuss second-order stationary processes. In section 2.7, we discuss long memory time series. Also, we review limit theorems for Hermite polynomials, the main tool used for the analysis of long memory Gaussian sequences. Finally in section 2.8, we discuss the leverage effect - a feature exhibited by financial time series.

2.2 General Inverses of Monotone Functions

Let $I = [a, b] \subset \mathbb{R}$.

Definition 2.2.1. *Let f be a real-valued, nondecreasing, right continuous function defined on I. The generalized inverse of f, denoted by $f^{\leftarrow}$, is defined as follows:*

$$f^{\leftarrow}(y) = \inf\{x \in I : f(x) \geq y\}, \tag{2.1}$$

for all $y \in \mathbb{R}$ *for which there exists* $x \in I$ *such that* $f(x) \geq y$*. Otherwise,* $f^{\leftarrow}(y) = b$*.*

It follows from Definition 2.2.1 that $f^{\leftarrow}$ is nondecreasing and left continuous. In addition, $\lim_{y \to -\infty} f^{\leftarrow}(y) = a$ and $\lim_{y \to \infty} f^{\leftarrow}(y) = b$.

Proposition 2.2.2. *Let* f *be a real-valued, nondecreasing, right continuous function defined on* I*. Then, for* $x \in I$ *and for every* $y \in \mathbb{R}$*,*

$$\begin{aligned} y \leq f(x) &\Leftrightarrow f^{\leftarrow}(y) \leq x, \\ y > f(x) &\Leftrightarrow f^{\leftarrow}(y) > x, \\ f\left(f^{\leftarrow}(y)\right) &\geq y. \end{aligned}$$

Definition 2.2.3. *Let* f *be a real valued, nondecreasing, left continuous function defined on* I*. The generalized inverse of* f*, denoted by* $f^{\rightarrow}$*, is defined as follows:*

$$f^{\rightarrow}(y) = \sup\{x \in I : f(x) \leq y\}, \tag{2.2}$$

for all $y \in \mathbb{R}$ *for which there exists* $x \in I$ *such that* $f(x) \leq y$*. Otherwise,* $f^{\rightarrow}(y) = a$*.*

It follows from Definition 2.2.3, $f^{\rightarrow}$ is nondecreasing and right continuous.

Proposition 2.2.4. *Let* f *be a real-valued, nondecreasing, left continuous function defined on* I*. Then, for* $x \in I$ *and for every* $y \in \mathbb{R}$*,*

$$y \leq f(x) \Leftrightarrow f^{\rightarrow}(y) \geq x.$$

It is worthwhile to mention that if f is continuous, then $f^{\leftarrow}$ and $f^{\rightarrow}$ coincide. In this case,

$$f\left(f^{\leftarrow}(y)\right) = f^{\leftarrow}\left(f(y)\right) = y.$$

The next result plays a major role when it comes to deriving the limit theorems for quantile processes (inverses of empirical processes).

Lemma 2.2.5. ***(Vervaat's lemma)****[23, p.357]*
Let $(f_n(t))_n$ *be a sequence of nondecreasing functions on an interval* $[a, b]$*. Let* g *be a function on the same interval with a nonnegative derivative,* g'*. Let* $(\delta_n)_n$ *be a sequence of nonnegative real numbers such that* $\delta_n \to 0$*, as* $n \to \infty$ *and there exists a continuous function* h *such that*

$$\frac{f_n(t) - g(t)}{\delta_n} \underset{n \to \infty}{\to} h(t),$$

uniformly on $[a, b]$. *Then,*

$$\frac{f_n^*(t) - g^*(t)}{\delta_n} \underset{n\to\infty}{\to} -(g^*)'(t) h\left(g^*(t)\right),$$

uniformly on $[g(a), g(b)]$, *where* $g^*, f_1^*, f_2^*, \ldots$ *are inverse functions (right or left continuous or defined in any way consistent with monotonicity).*

2.3 Regular Variation

The concept of regular variation was initially introduced in 1930 by Karamata. Since then, regular variation has been extensively studied and finds numerous applications in finance, economy, hydrology, applied probability, etc. This analytical property is required when dealing with heavy-tailed phenomena as well as domains of attraction; see [47].

This section is structured as follows. First, we present slowly varying functions, the cornerstone of this theory. Second, we tackle regularly varying real valued functions. Finally, we discuss regularly varying random variables.

2.3.1 Slowly Varying Functions

Definition 2.3.1. *[13, p.6-8]*
A measurable function $\ell : (0, +\infty) \to (0, +\infty)$ *is slowly varying at infinity if for all* $t > 0$,

$$\frac{\ell(xt)}{\ell(x)} \underset{x\to\infty}{\to} 1. \tag{2.3}$$

The convergence in (2.3) is uniform in t on each compact set in $(0, \infty)$. In the sequel, SV_∞ denotes the set of all slowly varying functions at infinity. Here are some noteworthy examples of slowly varying functions: nonnegative constants, logarithms, and iterated logarithms functions. For brevity, we will refer to functions satisfying (2.3) simply as *slowly varying functions.*

Theorem 2.3.2 (Karamata representation theorem). *[13, p.12]*
The function ℓ *is slowly varying if and only if it may be written in the form*

$$\ell(x) = c(x) \exp\left(\int_a^x \frac{\eta(y)}{y}\, dy\right),\ x \geq a, \tag{2.4}$$

for some $a > 0$, *where* $x \mapsto c(x)$ *is measurable and* $c(x) \to c > 0$, $\eta(x) \to 0$, *as* $x \to \infty$.

Example 2.3.3. The function $x \mapsto \ln x \in SV_\infty$. Indeed,

$$\ln x = \exp\left(\int_e^x \frac{dt}{t\ln t}\right) . \tag{2.5}$$

Remark 2.3.4. The value of a in this theorem is unimportant since ℓ, c, η may be altered at will on finite intervals. One may choose working with $a = 1$ or $a = 0$ on taking $\eta = 0$ on a neighbourhood of 0 to avoid divergence of integral at the origin, and one may assume that c is bounded. See [13, p.12].

Next, we gather some closure properties that slowly varying functions satisfy.

Proposition 2.3.5. *[13, p.16]*
Let ℓ, ℓ_1, $\ell_2, \dots, \ell_k$ be nonnegative measurable functions on $(0, \infty)$.

i)- $\ell_1, \ell_2 \in SV_\infty \Rightarrow \ell_1 + \ell_2 \in SV_\infty$.

ii)- $\ell_1, \ell_2 \in SV_\infty \Rightarrow \ell_1\ell_2 \in SV_\infty$.

iii)- $\ell_1, \ell_2 \in SV_\infty \Rightarrow \dfrac{\ell_1}{\ell_2} \in SV_\infty$.

iv)- $\ell \in SV_\infty \Rightarrow \ell^\gamma \in SV_\infty$, *for all* $\gamma \in \mathbb{R}$. *In particular,* $\ell \in SV_\infty \Rightarrow 1/\ell \in SV_\infty$.

v)- $\ell \in SV_\infty \Rightarrow \forall\gamma > 0,\ x^\gamma\ell(x) \to \infty$; $x^{-\gamma}\ell(x) \to 0$ *as* $x \to \infty$.
This property provides an insight on the asymptotic behaviour of a slowly varying function. It highlights that slowly varying functions are dominated by power functions.

vi)- If $\ell_1, \dots \ell_k \in SV_\infty$ *and* $r(x_1, \dots, x_k)$ *is a rational function with positive coefficients, then* $r(\ell_1(x), \dots, \ell_k(x)) \in SV_\infty$.

Proof. i)- Assume that $\ell_i \in SV_\infty$ for $i = 1, 2$. Equivalently, for all $\epsilon > 0$, there exists $M_i = M_i(\epsilon) > 0$ such that

$$x \geq M_i \Rightarrow |\ell_i(xt) - \ell_i(x)| < \epsilon\ell_i(x).$$

We consider $h(x,t) := (\ell_1(xt) + \ell_2(xt)) - (\ell_1(x) + \ell_2(x))$.
There exists $M_1 \vee M_2 = M(\epsilon) > 0$ such that $\forall x > 0$, $x \geq M(\epsilon)$ implies

$$\begin{aligned} |h(x,t)| &\leq |\ell_1(xt) - \ell_1(x)| + |\ell_2(xt) - \ell_2(x)| \\ &\leq \epsilon\left(\ell_1(x) + \ell_2(x)\right), \end{aligned}$$

which is equivalent to writing $\ell_1 + \ell_2 \in SV_\infty$.

ii)-iv)- These three proofs are straightforward applications of Definition 2.3.1.

v)- Assume that $\ell \in SV_\infty$. By Theorem 2.3.2, it holds that for all $x > 0$,

$$\ell(x) = c(x) \exp\left(\int_a^x \frac{\eta(y)}{y}\, dy\right),$$

for some $a > 0$ and $\eta(x) \to 0, c(x) \to c_0 > 0$ as $x \to \infty$. Since $\eta(x) \to 0$, as $x \to \infty$ then for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that $x \geq \delta \Rightarrow |\eta(x)| < \epsilon$. Therefore for all $y \geq a \vee \delta$ we have $-\epsilon \leq \eta(y) \leq \epsilon$. Consequently, we have

$$-\epsilon \ln\left(\frac{x}{a}\right) \leq \int_a^x \frac{\eta(y)}{y}\, dy \leq \epsilon \ln\left(\frac{x}{a}\right)$$
$$\left(\frac{x}{a}\right)^{-\epsilon} \leq \exp\left(\int_a^x \frac{\eta(y)}{y}\, dy\right) \leq \left(\frac{x}{a}\right)^{\epsilon}.$$

Hence for any $\gamma > 0$, by choosing $\epsilon < \gamma$, we conclude that $x^\gamma \ell(x) \to \infty$ and $x^{-\gamma}\ell(x) \to 0$, as $x \to \infty$.

vi)- Without loss of generality, let a_i, $b_i > 0$, $\sum_{i=1}^k b_i x_i \neq 0$ and

$$r(x_1, \ldots, x_k) = \sum_{i=1}^k a_i x_i / \sum_{i=1}^k b_i x_i.$$

Assuming that $\ell_1, \ldots, \ell_k \in SV_\infty$, we obtain for all $x > 0$,

$$r(\ell_1(x), \ldots, \ell_k(x)) = \frac{\sum_{i=1}^k a_i \ell_i(x)}{\sum_{i=1}^k b_i \ell_i(x)}.$$

This proves that $r(\ell_1(x), \ldots, \ell_k(x)) \in SV_\infty$ as a result of i)-iii).

□

Remark 2.3.6. The nonnegativity assumption in Proposition 2.3.5 plays a major role. In general, SV_∞ is not closed under subtraction. Although $x \mapsto \ln x$ and $x \mapsto \ln(1+x) \in SV_\infty$, but neither $x \mapsto \ln x - \ln(1+x)$ nor $x \mapsto \ln(1+x) - \ln x$ is slowly varying. In fact, setting $h(x) = \ln x - \ln(1+x)$ and applying l'Hospital's rule to the following ratio of functions yields that

$$\lim_{x\to\infty} \frac{h(xt)}{h(x)} = \lim_{x\to\infty} \frac{\ln\left(\frac{xt}{1+xt}\right)}{\ln\left(\frac{x}{1+x}\right)} = \lim_{x\to\infty} \frac{1+x}{1+xt} = \frac{1}{t} \neq 1.$$

Theorem 2.3.7 (Potter's bounds). *[13, p.25]*
Let ℓ be a slowly varying function. Then, for all $C > 1$, $\epsilon > 0$ there exists $\eta = \eta(C, \epsilon) \geq 0$ such that for $x \geq \eta$, $y \geq \eta$,

$$\frac{\ell(y)}{\ell(x)} \leq C \left[\left(\frac{y}{x}\right)^{\epsilon} \vee \left(\frac{y}{x}\right)^{-\epsilon} \right]. \tag{2.6}$$

Furthermore, if ℓ is bounded away from 0 and ∞ on every compact subset of $(0, \infty)$, then for all $\epsilon > 0$, there exists $C(\epsilon) > 1$ such that for $x > 0$, $y > 0$,

$$\frac{\ell(y)}{\ell(x)} \leq C(\epsilon) \left[\left(\frac{y}{x}\right)^{\epsilon} \vee \left(\frac{y}{x}\right)^{-\epsilon} \right]. \tag{2.7}$$

Theorem 2.3.8 (Karamata's Theorem). *[13, p.26; p.27], [44, p.9]*
Let ℓ be a slowly varying function and $a > 0$. Then

- *for $\gamma > -1$, as $x \to \infty$,*

$$\int_a^x u^{\gamma} \ell(u)\, du \sim \frac{x^{\gamma+1}}{\gamma+1} \ell(x). \tag{2.8}$$

- *for $\gamma < -1$, as $x \to \infty$,*

$$\int_x^{\infty} u^{\gamma} \ell(u)\, du \sim -\frac{x^{\gamma+1}}{\gamma+1} \ell(x)\ , \tag{2.9}$$

where $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$.

2.3.2 Regularly Varying Functions

Definition 2.3.9. *[13, p.19]*
A measurable function $f : (0, \infty) \to (0, \infty)$ is said to be regularly varying at infinity with index $\gamma \in \mathbb{R}$ if for all $t > 0$,

$$\frac{f(xt)}{f(x)} \underset{x\to\infty}{\to} t^{\gamma}. \tag{2.10}$$

We will denote by $RV_{\infty}(\gamma)$, the set of all regularly varying functions at infinity with index γ. The parameter γ is called the **index of regular variation**. The convergence in (2.10) is uniform in t, on compact subsets of $(0, \infty)$. From (2.10), a slowly varying function is regularly varying function at infinity with null index. Typical examples of regularly varying functions are: $x \mapsto x^{\gamma}$, $x \mapsto x^{\gamma} \ln(1 + x)$.

Theorem 2.3.10 (**Karamata's Representation**). *[13, p.21]*
A measurable function f is regularly varying at infinity with index γ if and only if there exists a slowly varying function ℓ as in Theorem 2.3.2 such that for all $x > 0$,

$$f(x) = x^{\gamma}\ell(x). \tag{2.11}$$

Theorem 2.3.11 (**Uniform Convergence Theorem**). *[13, p.22-23], [47, p.41]*
If f is a regularly varying function at infinity with index $\gamma \in \mathbb{R}$, then (2.10) holds locally uniformly in t on compact intervals. If $\gamma < 0$, then uniform convergence holds on intervals of the form $[a, \infty]$, $a > 0$. If $\gamma > 0$, then uniform convergence holds on intervals $(0, a]$, provided f is bounded on $(0, a]$, for all $a > 0$.

The next result is about some closure properties of regularly varying functions.

Proposition 2.3.12. *[13, p.22; p.26]*
Let f, f_1, $f_2, \ldots, f_k$ be measurable functions on $(0, \infty)$.

i)- $f_1 \in RV_\infty(\gamma_1), f_2 \in RV_\infty(\gamma_2) \Rightarrow f_1 + f_2 \in RV_\infty(\max(\gamma_1, \gamma_2))$.

ii)- $f_1 \in RV_\infty(\gamma_1), f_2 \in RV_\infty(\gamma_2) \Rightarrow f_1 f_2 \in RV_\infty(\gamma_1 + \gamma_2)$.

iii)- $f_1 \in RV_\infty(\gamma_1), f_2 \in RV_\infty(\gamma_2) \Rightarrow f_1 \circ f_2 \in RV_\infty(\gamma_1\gamma_2)$.

iv)- $f \in RV_\infty(\gamma) \Rightarrow \forall\lambda \in \mathbb{R},\ (f)^{\lambda} \in RV_\infty(\lambda\gamma)$. *In particular,*
$f \in RV_\infty(\gamma) \Rightarrow 1/f \in RV_\infty(-\gamma)$.

v)- $f \in RV_\infty(\gamma), \gamma \neq 0 \Rightarrow$ *as* $x \to \infty$, $f(x) \to \infty$ *if* $\gamma > 0$ *and* $f(x) \to 0$ *if* $\gamma < 0$.

Proof. i)- Assume that $f_i \in RV_\infty(\gamma_i)$ for $i = 1, 2$. Definition 2.3.9 and Theorem 2.3.10 yield that for all $x > 0$, $t > 0$, $f_i(x) = x^{\gamma_i}\ell_i(x)$, $\ell_i \in SV_\infty$ and

$$\frac{f_i(xt)}{f_i(x)} \underset{x\to\infty}{\to} t^{\gamma_i}.$$

If $\gamma_1 = \gamma_2$, then the proof is immediate. Without loss of generality, assume that $\gamma_1 < \gamma_2$. We have

$$\frac{(f_1 + f_2)(xt)}{(f_1 + f_2)(x)} = \frac{f_1(xt)/f_2(x) + f_2(xt)/f_2(x)}{1 + f_1(x)/f_2(x)}$$

On account of Proposition 2.3.5 [iii) and v)], it holds that

$$\frac{f_1(x)}{f_2(x)} = x^{-(\gamma_2-\gamma_1)}\frac{\ell_1(x)}{\ell_2(x)} \underset{x\to\infty}{\to} 0.$$

This implies that $\frac{f_1(xt)}{f_2(x)} \to 0$, as $x \to \infty$. Therefore, we get

$$\frac{(f_1+f_2)(xt)}{(f_1+f_2)(x)} \underset{x\to\infty}{\to} t^{\gamma_2}.$$

This means that $f_1 + f_2 \in RV_\infty(\max(\gamma_1, \gamma_2))$.

ii)- Assume that $f_i \in RV_\infty(\gamma_i)$ for $i = 1, 2$. Again by Theorem 2.3.10, we have

$$f_i(x) = x^{\gamma_i}\ell_i(x), \text{ where } \ell_i \in SV_\infty.$$

Therefore, $(f_1 f_2)(x) = x^{\gamma_1+\gamma_2}(\ell_1 \times \ell_2)(x)$. It holds by Proposition 2.3.5 [i] that $\ell_1 \times \ell_2 \in SV_\infty$. Thus, $f_1 f_2$ is a regularly varying function with index $\gamma_1 + \gamma_2$, and hence, the conclusion follows (cf. Theorem 2.3.10).

iii)- Assume that $f_i \in RV_\infty(\gamma_i)$ for $i = 1, 2$. Again by Theorem 2.3.10, we have for $x > 0$, $f_i(x) = x^{\gamma_i}\ell_i(x)$, where $\ell_i \in SV_\infty$. Therefore,

$$\begin{aligned}
&(f_1 \circ f_2)(x) = f_1(f_2(x)) = (x^{\gamma_2}\ell_2(x))^{\gamma_1}\ell_1(x^{\gamma_2}\ell_2(x)) \\
&= x^{\gamma_1\gamma_2}(\ell_2(x))^{\gamma_1}\ell_1(x^{\gamma_2}\ell_2(x)) = x^{\gamma_1\gamma_2}\ell(x),
\end{aligned}$$

where $x \mapsto \ell(x) := (\ell_2(x))^{\gamma_1}\ell_1(x^{\gamma_2}\ell_2(x))$. We claim that $\ell \in SV_\infty$. In fact, $\ell_2 \in SV_\infty$, so does $x \mapsto (\ell_2(x))^{\gamma_1}$ for $\gamma_1 > 0$. Set $\rho(x) = x^{\gamma_2}\ell_2(x)$. Moreover, if $x \mapsto \ell_1(\rho(x)) \in SV_\infty$, then $x \mapsto \ell(x) \in SV_\infty$, as the product of two slowly varying functions. It remains to show that $x \mapsto \ell_1(\rho(x)) \in SV_\infty$. Since $\ell_1 \in SV_\infty$, then by Theorem 2.3.2, we have for all $x > 0$,

$$\ell_1(x) = c(x)\exp\left(\int_a^x \frac{\eta(y)}{y}\,dy\right),$$

for some $a > 0$ and $\eta(x) \to 0$, $c(x) \to c_0 > 0$ as $x \to \infty$. Therefore

$$\frac{\ell_1(\rho(xt))}{\ell_1(\rho(x))} = \exp\left(\int_{\rho(x)}^{\rho(xt)} \frac{\eta(y)}{y}\,dy\right).$$

We recall that $\eta(x) \to 0$ as $x \to \infty$. Therefore for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that $x \geq \delta \Rightarrow -\epsilon < \eta(x) < \epsilon$. Consequently, we obtain

$$-\epsilon\ln\left(\frac{\rho(xt)}{\rho(x)}\right) < \int_{\rho(x)}^{\rho(xt)} \frac{\eta(y)}{y}\,dy < \epsilon\ln\left(\frac{\rho(xt)}{\rho(x)}\right)$$

and hence for x sufficiently large,

$$-\epsilon\gamma_2 \ln t < \int_{\rho(x)}^{\rho(xt)} \frac{\eta(y)}{y}\, dy < \epsilon\gamma_2 \ln t.$$

Since ϵ is arbitrary, then we conclude that $\displaystyle\int_{\rho(x)}^{\rho(xt)} \frac{\eta(y)}{y}\, dy \to 0$, as $x \to \infty$. Thus,

$$\frac{\ell_1(\rho(xt))}{\ell_1(\rho(x))} \underset{x\to\infty}{\to} 1.$$

iv)- Assume that $f \in RV_\infty(\gamma)$. So, for all $x > 0$, $f(x) = x^\gamma \ell(x)$, where $\ell \in SV_\infty$. Consequently, for all $\lambda \in \mathbb{R}$, $f^\lambda(x) = x^{\gamma\lambda}\ell^\lambda(x)$. Since $\ell \in SV_\infty$, so does ℓ^λ. Therefore, f^λ is a regularly varying function with index $\gamma\lambda$. Thus, the desired result holds (cf. Theorem 2.3.10). In particular, taking $\lambda = -1$ implies $1/\ell \in RV_\infty(-\gamma)$.

v)- Assume that $f \in RV_\infty(\gamma)$. So, for all $x > 0$, $f(x) = x^\gamma \ell(x)$, where $\ell \in SV_\infty$. The rest of the proof works identically as in the proof of Proposition 2.3.5 [v].

$\square$

Theorem 2.3.13 (**Potter's bounds**). *[13, p.25]*
If f is a regularly varying function with index γ, i.e. $f(x) = x^\gamma \ell(x)$, then for all $C > 1$, $\epsilon > 0$; there exists $\eta = \eta(C, \epsilon) \geq 0$ such that for $x, y \geq \eta$,

$$\frac{f(y)}{f(x)} \leq C\left(\left(\frac{y}{x}\right)^{\gamma+\epsilon} \vee \left(\frac{y}{x}\right)^{\gamma-\epsilon}\right). \tag{2.12}$$

Furthermore, if ℓ is bounded away from 0 and ∞ on every compact subset of $(0, \infty)$, then for all $\epsilon > 0$, there exists $C(\epsilon) > 1$ such that for $x > 0$, $y > 0$,

$$\frac{f(y)}{f(x)} \leq C(\epsilon)\left(\left(\frac{y}{x}\right)^{\gamma+\epsilon} \vee \left(\frac{y}{x}\right)^{\gamma-\epsilon}\right). \tag{2.13}$$

Theorem 2.3.14 (**Karamata's Theorem**). *[44, p.799], [13, p.26]*
Let $\gamma > -1$. If $f \in RV_\infty(\gamma)$ and integrable on (a, x) for any $x > 0$ and some $a > 00$, then $\displaystyle\int_a^x f(u)\, du \in RV_\infty(\gamma + 1)$ and

$$\int_a^x f(u)\, du \sim \frac{x f(x)}{\gamma + 1}, \ \ \text{as } x \to \infty. \tag{2.14}$$

The next result provides a suitable framework under which the generalized inverse function of a nondecreasing function is regularly varying.

Recall that for a real-valued, nondecreasing, right continuous function f defined on $I = [a, b]$, its generalized inverse function $f^{\leftarrow}$ is defined by

$$f^{\leftarrow}(y) = \inf\{x \in I : f(x) \geq y\}. \tag{2.15}$$

Proposition 2.3.15. *[23, p.367]*
If f is regularly varying at infinity with index $\gamma > 0$, then $f^{\leftarrow}$ is regularly varying at infinity with index $1/\gamma$.

2.3.3 Regularly Varying Random Variables

In what follows, random variables that we are going to be dealing with are assumed to be on a common probability space $(\Omega, \mathcal{F}, P)$. For any random variable W with distribution function F_W, we define its tail distribution function, $\bar{F}_W$, by

$$\bar{F}_W(x) := 1 - F_W(x).$$

Unless otherwise stated, we assume $\gamma > 0$ in what follows.

For simplicity, we consider nonnegative random variables only.

Definition 2.3.16. *[44]*
A random variable W is said to be regularly varying at infinity with index γ, if its tail distribution function $\bar{F}_W$ is regularly varying with index $-\gamma$, that is for all $t > 0$,

$$\frac{\bar{F}_W(xt)}{\bar{F}_W(x)} \underset{x\to\infty}{\rightarrow} t^{-\gamma}. \tag{2.16}$$

The parameter γ is called the **tail index** and measures the heaviness of the tail of X. The smaller γ is, the heavier is the right tail of the distribution of X. *The Pareto distribution, and distributions of the absolute values of t and Cauchy random variables all have regularly varying tails.* In what follows, we collect some properties satisfied by regularly varying random variables.

Remark 2.3.17. If W is a regularly varying random variable with tail distribution function $\bar{F}_W$, then Theorem 2.3.10 yields that

$$\bar{F}_W(x) = x^{-\gamma}\ell_W(x), \tag{2.17}$$

for some $\ell \in SV_\infty$. Furthermore since $\bar{F}_W$ is monotone non-increasing and bounded above by 1, it follows that ℓ_W is bounded away from 0 and ∞ on any compact interval of $(0, \infty)$. In later discussions, the following result will play a key role in justifying the interchange between a limit and an integral.

Lemma 2.3.18 (**Potter's bounds**). *If W is a random variable such that $\bar{F}_W \in RV_\infty(-\gamma)$, then for all $\epsilon > 0$ there exists $C(\epsilon) \geq 1$ such that for all $x, y > 0$,*

$$\frac{\bar{F}_W(x/y)}{\bar{F}_W(x)} \leq C(\epsilon)\left(y^{\gamma+\epsilon} \vee y^{\gamma-\epsilon}\right). \tag{2.18}$$

Proof. Since $\bar{F}_W \in RV_\infty(-\gamma)$, then (2.17) and Theorem 2.3.7 yield

$$\frac{\bar{F}_W(x/y)}{\bar{F}_W(x)} = y^\gamma \frac{\ell_W(x/y)}{\ell_W(x)} \leq C(\epsilon)\left(y^{\gamma+\epsilon} \vee y^{\gamma-\epsilon}\right).$$

□

Theorem 2.3.19 (**Potter's bounds**). *[50]*
Let W be a regularly varying random variable with index $\gamma > 0$. Then for all $\epsilon > 0$, there exists $C(\epsilon) > 1$ such that for all $x \geq 0$, $y > 0$,

$$\frac{\bar{F}_W\,(x/y)}{\bar{F}_W(x)} \leq C(\epsilon)\max(1, y^{\gamma+\epsilon}). \tag{2.19}$$

Proof. We assume that W is a regularly varying random variable with index $\gamma > 0$.
If $x = 0$, then the desired bound holds with $C = 1$. From here on, we assume that $x > 0$.
If $y \leq 1$, then $\bar{F}_W\,(x/y) \leq \bar{F}_W(x)$. Hence, the desired bound holds with $C = 1$.
If $y > 1$, then (2.17) ensures that

$$\frac{\bar{F}_W\,(x/y)}{\bar{F}_W(x)} = \frac{(x/y)^{-\gamma}\ell_W(x/y)}{x^{-\gamma}\ell_W(x)} = y^\gamma \frac{\ell_W(x/y)}{\ell_W(x)}.$$

So, by (2.7), we have $\forall \epsilon > 0$, there exists $C = C(\epsilon) > 1$ such that for $x > 0$, $y > 1$,

$$\frac{\bar{F}_W\,(x/y)}{\bar{F}_W(x)} \leq C y^\gamma \left(\frac{x/y}{x}\right)^{-\epsilon} = C y^{\gamma+\epsilon}.$$

□

The next result is known in the literature as *Breiman's Lemma*. It provides a framework under which a product of two random variables is regularly varying. This is critical for later proofs.

Lemma 2.3.20 (Breiman's Lemma). *[50, p.49]*
Let V and W be two independent nonnegative random variables such that $\bar{F}_V$ is regularly varying with index $-\gamma$. If there exists $\epsilon > 0$ such that $E(W^{\gamma+\epsilon}) < \infty$, then

$$\frac{P(VW > x)}{P(V > x)} \underset{x\to\infty}{\to} E(W^\gamma). \tag{2.20}$$

The meaning of (2.20) is that VW is regularly varying with index γ.

Proof. Assume that V is a regularly varying random variable with index γ. Define the family of functions $(G_x)_{x>0}$ by

$$G_x(y) := \frac{\bar{F}_V\,(x/y)}{\bar{F}_V(x)}\;, \quad y > 0\;.$$

Clearly, as x goes to ∞, G_x converges to y^γ. Moreover by (2.19), we have: $\forall\epsilon > 0$, $\exists$ $C = C(\epsilon) > 1$ such that for $x \geq 0$, $y > 0$,

$$G_x(y) \leq C(\max(1, y))^{\gamma+\epsilon}.$$

Since $E(W^{\gamma+\epsilon}) < \infty$, the dominated convergence theorem ensures $E\,(G_x(W)) \underset{x\to\infty}{\to} E\,(W^\gamma)$, which is equivalent to (2.20). ☐

Remark 2.3.21. We wrap up this section by discussing regular variation of other common functions related to either the distribution function, F_W, or the tail distribution function, $\bar{F}_W$, of a random variable W. These functions of interest are the following generalized inverses in the sense of (2.15):

$$U_W(t) = (1/\bar{F}_W)^{\leftarrow}(t)\,, \tag{2.21a}$$

$$Q_W(t) = F_W^{\leftarrow}\,(1 - 1/t)\,, \quad t > 1. \tag{2.21b}$$

Assume that $\bar{F}_W$ is strictly decreasing on the range of W. This implies that $1/\bar{F}_W$ is strictly increasing. Therefore, if $\bar{F}_W$ is regularly varying with index $-\gamma$, then U_W is regularly varying with index $1/\gamma$, by Proposition 2.3.15. In addition, if $\bar{F}_W$ is strictly decreasing and continuous on the range of W, U_W coincides with Q_W. **This assumption is in effect from now on.**

2.4 Second-Order Regular Variation

In this section we introduce a second order regular variation, a stronger form of regular variation.

Let g be a regularly varying function with index $\gamma \in \mathbb{R}$. The first use of second order regular variation is to control the difference

$$\left| \frac{g(xt)}{g(x)} - t^{\gamma} \right| . \tag{2.22}$$

The second use is related to an approximation of high quantiles. Second-order regular variation finds numerous applications in risk management and many other fields.

This section is structured as follows. First, we present second-order slowly varying functions. Second, we discuss second-order regularly varying real valued functions. Finally, we discuss second-order regularly varying random variables.

2.4.1 Second-Order Slowly Varying Functions

There are various definitions of second order slow and regular variation in the literature. The most common approach is via the asymptotic behaviour of the function, while another approach is via a representation of the function. In the case of regular variation the two approaches are equivalent; cf. the limiting behaviour in (2.10) and Karamata's Representation in Theorem 2.3.10.

For our purposes, the most suitable second order condition is the representation defined in [39]. See also [28]. As we will demonstrate below, this approach implies second order variation in terms of the limiting behaviour. However, equivalence appears to be an open question.

Definition 2.4.1 ([39], [28])**.** *A measurable function h is said to be second order slowly varying at infinity with index $\rho < 0$ and the rate function η if for all $x > 0$,*

$$h(x) = c^* \exp\left(\int_1^x \frac{\eta(u)}{u}\, du \right); \tag{2.23}$$

with $c^* > 0$, $\eta : (0, \infty) \to \mathbb{R}$ is a bounded regularly varying function at infinity with index ρ or $\eta(x) = 0$, for all x. Further, η is either nonnegative or nonpositive.

In the case that $\eta(x) = 0$ for all x sufficiently large, we set $\rho = -\infty$. We denote by $2SV_\infty(\rho, \eta)$, the set of all second-order slowly varying functions at infinity.

Example 2.4.2. Consider the following function

$$\begin{aligned} h(x) &= \frac{\alpha + x^{-\beta}}{\alpha+1}, \quad x, \alpha, \beta > 0 \\ &= \exp\left(\ln\left(\frac{\alpha + x^{-\beta}}{\alpha+1}\right)\right) \\ &= \exp\left(\int_1^x \frac{-\beta u^{-\beta-1}}{\alpha + u^{-\beta}}\, du\right). \end{aligned}$$

Therefore $h \in 2SV_\infty(-\beta, \eta)$, with $\eta(x) = -\beta/(1 + \alpha x^{\beta})$. Notice that η is nonpositive and bounded by β. In addition, $\eta \in RV_\infty(-\beta)$.

Second-order slowly varying functions satisfy the following equivalent representation stated, but not proven, in Eq. (66) of [39].

Lemma 2.4.3 ([39])**.** *If h is second-order slowly varying at infinity, then for all $x > 0$,*

$$h(x) = h(1) + \left(\int_1^x \frac{\eta(u)h(u)}{u}\, du\right). \tag{2.24}$$

Conversely, if (2.24) is satisfied, then (2.23) holds.

Proof. Notice that (2.23) implies $c^* = h(1)$. The derivative of (2.24) yields

$$h(1)\frac{\eta(x)}{x} \exp\left(\int_1^x \frac{\eta(u)}{u}\, du\right) = h'(x).$$

Using (2.23) and integrating this equation over the interval from 1 to t yields

$$\int_1^t \frac{\eta(x)h(x)}{x}\, dx = h(t) - h(1).$$

□

The following result on the asymptotic behaviour of functions in $2SV_\infty(\rho, \eta)$ appears to be new. In fact, (2.25) is frequently used as a definition of second order slow variation. See Section 2.3 in [23] and Remark 2.4.6 below.

Lemma 2.4.4. *Assume that (2.23) holds. Then for $t \geq 1$,*

$$\frac{h(xt)/h(x) - 1}{\eta(x)} \underset{x\to\infty}{\to} \frac{t^\rho - 1}{\rho}. \tag{2.25}$$

Proof. Assume without loss of generality that $x \geq 1$. It follows from (2.23) that

$$\begin{aligned} \frac{h(xt)}{h(x)} - 1 &= \exp\left(\int_x^{xt} \frac{\eta(u)}{u}\, du\right) - 1 \\ &= \exp\left(\eta(x) \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv\right) - 1, \end{aligned}$$

where the last equality holds by setting $v = u/x$. Since $\rho < 0$ and $\eta \in RV_\infty(\rho)$, then

$$\eta(x) \underset{x\to\infty}{\to} 0 \text{ and } \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv \underset{x\to\infty}{\to} \int_1^t v^{\rho-1}\, dv = \frac{t^\rho - 1}{\rho}\ ,$$

where the integral above converges thanks to Theorem 2.3.11. Therefore, as $x \to \infty$,

$$\begin{aligned} \frac{h(xt)}{h(x)} - 1 &= \eta(x) \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv + \frac{1}{2} \exp\left(\xi^*\right) \left(\eta(x) \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv\right)^2 \\ \frac{h(xt)/h(x) - 1}{\eta(x)} &= \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv \left(1 + \frac{1}{2} \exp\left(\xi^*\right) \left(\eta(x) \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv\right)\right). \end{aligned}$$

Note that the Lagrange remainder ensures that

$$0 \leq \xi^* \leq \int_1^t \left|\frac{\eta(xv)}{v\eta(x)}\right|\, dv.$$

It follows that for t fixed, $\exp(\xi^*)$ is bounded as a function of x and therefore

$$\frac{1}{2} \exp\left(\xi^*\right) \left(\eta(x) \int_1^t \frac{\eta(xv)}{v\eta(x)}\, dv\right) = o(1).$$

□

Next, we investigate some closure properties of second-order slowly varying functions that appear to be new.

Lemma 2.4.5. *Let h, h_1 and h_2 be measurable functions.*

i) If $h_1 \in 2SV_\infty(\rho_1, \eta_1)$, $h_2 \in 2SV_\infty(\rho_2, \eta_2)$ and η_1, η_2 have the same sign, then

$$h_1 + h_2 \in 2SV_\infty(\max(\rho_1, \rho_2), \eta_1 + \eta_2),$$

ii) If $h_1 \in 2SV_\infty(\rho_1, \eta_1)$, $h_2 \in 2SV_\infty(\rho_2, \eta_2)$ and η_1, η_2 are of the same sign, then

$$h_1 h_2 \in 2SV_\infty(\max(\rho_1, \rho_2), \eta_1 + \eta_2).$$

iii) If $h \in 2SV_\infty(\rho, \eta)$, then $h^p \in 2SV_\infty(\rho, p\eta)$, for all $p \in \mathbb{R}$.

Proof. In what follows, we assume that $h_i \in 2SV_\infty(\rho_i, \eta_i)$, for $i = 1, 2$. Therefore,

$$h_i(x) = h_i(1) + \int_1^x \frac{\eta_i(u)h_i(u)}{u}\, du,$$

where $\eta_i : (0, \infty) \to \mathbb{R}$ is either nonpositive or nonnegative, bounded and regularly varying function at infinity with index ρ_i.

i) Since $h_i \in 2SV_\infty(\rho_i, \eta_i)$, for $i = 1, 2$, then for all $x > 0$,

$$\begin{aligned}(h_1 + h_2)(x) &= h_1(1) + h_2(1) + \int_1^x \frac{\eta_1(u)h_1(u) + \eta_2(u)h_2(u)}{u}\, du \\ &= (h_1 + h_2)(1) + \int_1^x \frac{\eta_1(u)h_1(u) + \eta_2(u)h_2(u)}{(h_1(u) + h_2(u))u}(h_1(u) + h_2(u))\, du\ .\end{aligned}$$

Since for $i = 1, 2$, $\eta_i \in RV_\infty(\rho_i)$, then for all $x > 0$, $\eta_i(x) = x^{\rho_i} h_i^*(x)$, for some slowly varying function h_i^*. Therefore for all $x > 0$,

$$\begin{aligned}\eta(x) &:= \frac{\eta_1(x)h_1(x) + \eta_2(x)h_2(x)}{(h_1(x) + h_2(x))} \\ &= x^{\rho_1}\frac{h_1^*(x)h_1(x)}{h_1(x) + h_2(x)} + x^{\rho_2}\frac{h_2^*(x)h_2(x)}{h_1(x) + h_2(x)}.\end{aligned}$$

The latter decomposition shows that η is a sum of two regularly varying functions with indices ρ_1 and ρ_2, respectively. Thus, $\eta \in RV_\infty(\max(\rho_1, \rho_2))$, by Proposition 2.3.12. In addition, since for all $x > 0$,

$$|\eta(x)| \le \left|\frac{h_1(x)}{h_1(x) + h_2(x)}\right| |\eta_1(x)| + \left|\frac{h_2(x)}{h_1(x) + h_2(x)}\right| |\eta_2(x)| \le |\eta_1(x)| + |\eta_2(x)|,$$

then η is bounded. Notice that if η_1 and η_2 are of the same sign over $(0, \infty)$, then so is η as h_1 and h_2 are nonnegative by Theorem 2.3.10. Notice that $\eta(x) \underset{x\to\infty}{\sim} \eta_1(x) + \eta_2(x)$.

ii) Since $h_i \in 2RV_\infty(\gamma, \rho_i, \eta_i)$, for $i = 1, 2$, then for all $x > 0$,

$$(h_1h_2)(x) = c_1^* c_2^* \exp\left(\int_1^x \frac{(\eta_1 + \eta_2)(t)}{t}\, dt\right).$$

Further, η_1 and η_2 are bounded, nonnegative, nonincreasing, so is $\eta_1 + \eta_2$. Finally, $\eta_1 + \eta_2 \in RV_\infty(\max(\rho_1, \rho_2))$, thanks to Proposition 2.3.12.

iii) Since $h \in 2SV_\infty(\rho, \eta)$, then (2.23) yields that for all $x > 0$ and $p \in \mathbb{R}$,

$$h^p(x) = (c^*)^p \exp\left(\int_1^x \frac{p\eta(u)}{u}\, du\right).$$

This equivalent to writing $h^p \in 2SV_\infty(\rho, p\eta)$, since $p\eta$ is either nonpositive or nonnegative, bounded on $(0, \infty)$ and regularly varying with index ρ.

□

Remark 2.4.6. Functions satisfying (2.25) are referred to as *extended regularly varying with index ρ and the rate function η* in [23]. In other words, any second-order slowly varying function is extended regularly varying. As such, Theorem B.2.18 in [23] holds and yields the following result.

Lemma 2.4.7 ([23])**.** *If h is a second-order slowly varying function with index $\rho < 0$ and rate function η, then for all $\epsilon, \delta > 0$ there exists $x_0 = x_0(\epsilon, \delta) > 0$ such that for all $x > x_0$, $t > x_0/x$,*

$$\left|\frac{h(xt) - h(x)}{\eta_0(x)} - \frac{t^\rho - 1}{\rho}\right| \leq \epsilon t^\rho \max\left(t^\delta, t^{-\delta}\right) \ ,$$

where the rate function $\eta_0(x) = -\rho\{h(\infty) - h(x)\}$ with $h(\infty) = \lim_{x\to\infty} h(x)$.

2.4.2 Second-Order Regularly Varying Functions

In analogy to the relation between slow and regular variation, we extend second order slow variation to second order regular variation.

Definition 2.4.8 ([28], [39])**.** *A measurable function g is said to be second order regularly varying at infinity with indices $\gamma \in \mathbb{R}$, $\rho < 0$ and the rate function η if for all $x > 0$,*

$$g(x) = c^* x^\gamma \exp\left(\int_1^x \frac{\eta(u)}{u}\, du\right); \tag{2.26}$$

with $c^* > 0$, $\eta : (0, \infty) \to \mathbb{R}$ is a bounded regularly varying function at infinity with index ρ or $\eta(x) = 0$, for all x. Further, η is either nonnegative or nonpositive.

Henceforth $2RV_\infty(\gamma, \rho, \eta)$ denotes the set of all functions g such that (2.26) holds. In the case that $\eta(x) = 0$ for all x sufficiently large, we set $\rho = -\infty$. The rate function η is the driving force of the concept of second-order regular variation.

Example 2.4.9. Set $g(x) = 0$ for $0 < x < 1$. Let $\alpha > 0, \gamma > 1$. We have for all $x \geq 1$,

$$\begin{aligned} g(x) &= \frac{1}{2}\left(x^{-\alpha} + x^{-\alpha\gamma}\right) \\ &= x^{-\alpha} \exp\left(\ln\left(\frac{1 + x^{-\alpha(\gamma-1)}}{2}\right)\right) \\ &= x^{-\alpha} \exp\left(\int_1^x \frac{\alpha(\gamma-1)t^{-\alpha(\gamma-1)-1}dt}{1 + t^{-\alpha(\gamma-1)}}\right). \end{aligned}$$

Therefore, $g \in 2RV_\infty(-\alpha, -\alpha(\gamma-1), \eta)$, where the rate function is defined as follows

$$\eta(x) = x^{-\alpha(\gamma-1)} \frac{\alpha(\gamma-1)}{1 + x^{-\alpha(\gamma-1)}} = \frac{\alpha(\gamma-1)}{1 + x^{\alpha(\gamma-1)}}. \tag{2.27}$$

Note that η is nonnegative, bounded by $\alpha(\gamma-1)$. In addition, $\eta \in RV_\infty(-\alpha(\gamma-1))$.

Remark 2.4.10. Second-order regular variation implies regular variation. However, the converse is not true. In fact, $x \mapsto x^\alpha \ln x$, with $\alpha \in \mathbb{R}$, is regularly varying at infinity with index α but fails to be second-order regularly varying at infinity. Although $\eta(x) = 1/\ln x$ is nonnegative on $[e, \infty)$ and bounded above by 1, the Karamata representation (2.5) shows that $\eta \in SV_\infty$.

The next result linking second-order regular variation to second-order slow variation follows directly from Lemma 2.4.3.

Corollary 2.4.11. *If $g \in 2RV_\infty(\gamma, \rho, \eta)$, then there exists $h \in 2SV_\infty(\rho, \eta)$ such that*

$$g(x) = x^\gamma \left(h(1) + \int_1^x \frac{\eta(u)h(u)}{u}\, du\right), \tag{2.28}$$

where $\eta : (0, \infty) \to \mathbb{R}$ is a bounded regularly varying function at infinity with index ρ or $\eta(x) = 0$, for all x. Further, η is either nonnegative or nonpositive.

Lemma 2.4.12. *Assume that (2.26) holds. Then for $x \geq 1$, $t > 0$,*

$$\frac{g(xt)/g(x) - t^\gamma}{\eta(x)} \underset{x\to\infty}{\to} t^\gamma \frac{t^\rho - 1}{\rho}. \tag{2.29}$$

Remark 2.4.13. We note that sometimes (2.29) is considered as the definition of second-order regular variation. See Remark B.3.15 in [23].

Proof. Assume without loss of generality that $x \geq 1$. It follows from (2.26) that

$$\begin{aligned}\frac{g(xt)}{g(x)} - t^{\gamma} &= t^{\gamma}\left(\exp\left(\int_x^{xt}\frac{\eta(u)}{u}\,du\right) - 1\right)\\ &= t^{\gamma}\left(\exp\left(\eta(x)\int_1^{t}\frac{\eta(xv)}{v\eta(x)}\,dv\right) - 1\right),\end{aligned}$$

where the last equality holds by setting $v = u/x$. The remainder of the proof is analogous to that of Lemma 2.4.4.

□

Again, in analogy to second-order slow variation, we now investigate closure properties of second-order regularly varying functions.

Lemma 2.4.14. *Let g, g_1 and g_2 be measurable functions.*

i) If $g_i \in 2RV_\infty(\gamma_i, \rho_i, \eta_i)$, for $i = 1, 2$ and η_1, η_2 are of the same sign, then

$$g_1 g_2 \in 2RV_\infty(\gamma_1 + \gamma_2, \max(\rho_1, \rho_2), \eta_1 + \eta_2).$$

ii) If $g_1 \in 2RV_\infty(\gamma, \rho_1, \eta_1)$, $g_2 \in 2RV_\infty(\gamma, \rho_2, \eta_2)$ and η_1, η_2 have the same sign, then

$$g_1 + g_2 \in 2RV_\infty(\gamma, \max(\rho_1, \rho_2), \eta_1 + \eta_2),$$

iii) If $g \in 2RV_\infty(\gamma, \rho, \eta)$, then $g^p \in 2RV_\infty(p\gamma, \rho, p\eta)$, for all $p \in \mathbb{R}$.

The proof of this result is analogous to the proof of Lemma 2.4.5. We are finally ready to state a bound for second-order regularly varying functions. Notice from Corollary 2.4.11 that if $g \in 2RV_\infty(\gamma, \rho, \eta)$, then $h(x) = g(x)/x^\gamma \in 2SV_\infty(\rho, \eta)$. Therefore Lemma 2.4.7 applied to $h(x) = g(x)/x^\gamma$ and some simple manipulations yield:

Lemma 2.4.15 (Theorem 2.3.6 in [23])**.** *If g is second order regularly varying at infinity with indices $\gamma \in \mathbb{R}$, $\rho < 0$ and the rate function η, then for all $\epsilon, \delta > 0$ there exists $x_0 = x_0(\epsilon, \delta) > 0$ such that for all $x > x_0$, $t > x_0/x$,*

$$\left|\frac{g(xt)/g(x) - t^\gamma}{B_0(x)} - t^\gamma\frac{t^\rho - 1}{\rho}\right| \leq \epsilon t^{\gamma+\rho}\max\left(t^\delta, t^{-\delta}\right) \ , \tag{2.30}$$

where $B_0(x) = -\rho\left(\lim_{x\to\infty} g(x)/x^\gamma - g(x)/x^\gamma\right)$.

2.4.3 Second-Order Regularly Varying Random Variables

Unless otherwise stated, α, ρ and c^* are assumed to be strictly positive real numbers in this whole subsection. We continue to assume that all random variables are nonnegative.

Definition 2.4.16. *A random variable W with tail distribution function $\bar{F}_W$ is said to be second-order regularly varying with indices $-\alpha, -\rho$ and the rate function η if $\bar{F}_W \in 2RV_\infty(-\alpha, -\rho, \eta)$, that is for all $x > 0$,*

$$\bar{F}_W(x) = c^* x^{-\alpha} \exp\left(\int_1^x \frac{\eta(u)}{u}\, du\right). \tag{2.31}$$

Example 2.4.17. If W is Burr-distributed, then its tail distribution functions is

$$\begin{aligned}
\bar{F}_W(x) &= (1 + x^\beta)^{-\alpha}, \quad x, \alpha, \beta > 0 \\
&= x^{-\alpha\beta} \exp\left(-\alpha \ln\left(1 + x^{-\beta}\right)\right) \\
&= x^{-\alpha\beta} \exp(-\alpha \ln 2) \exp\left(-\alpha\{\ln\left(1 + x^{-\beta}\right) - \ln 2\}\right) \\
&= 2^{-\alpha} x^{-\alpha\beta} \exp\left(\alpha \int_1^x \frac{-\beta u^{-\beta-1}}{1 + u^{-\beta}}\, du\right).
\end{aligned}$$

Therefore, $\bar{F}_W \in 2RV_\infty(-\alpha\beta, -\beta, \eta)$, where for all $x > 0$, η is defined by

$$\eta(x) = \frac{\alpha\beta}{1 + x^\beta}.$$

Notice that η is nonnegative and bounded by $\alpha\beta$. In addition, $\eta \in RV_\infty(-\beta)$.

Remark 2.4.18. Clearly, if $\bar{F}_W \in 2RV_\infty(-\alpha, -\rho, \eta)$ with $\eta \not\equiv 0$ then by Lemma 2.4.4,

$$\frac{\bar{F}_W(xt)/\bar{F}_W(x) - t^{-\alpha}}{\eta(x)} \underset{x\to\infty}{\to} t^{-\alpha} \frac{1 - t^{-\rho}}{\rho}. \tag{2.32}$$

Thanks to second-order regular variation we can control the speed of convergence of the ratio $\bar{F}_W(xt)/\bar{F}_W(x)$, as $x \to \infty$. Pareto distributions are included in the definition of second regular variation by allowing $\eta(x) = 0$ for $x \geq 1$ and in this case

$$\bar{F}_W(xt)/\bar{F}_W(x) - t^{-\alpha} = 0, \ \forall t, x \geq 1.$$

Lemma 2.4.19 ([39], p. 130)**.** *Assume that $\bar{F}_W \in 2RV_\infty(-\alpha, -\rho, \eta)$. Then for any $\epsilon > 0$, there exists $C = C(\epsilon) > 0$ such that for all $t > 0, x \geq 1$,*

$$\left|\frac{\bar{F}_W(xt)}{\bar{F}_W(x)} - t^{-\alpha}\right| \leq C|\eta(x)| \left(t^{-\alpha-\rho+\epsilon} \vee t^{-\alpha-\rho-\epsilon}\right). \tag{2.33}$$

In what follows, we discuss second-order regular variation of U_W defined in (2.21a).

Remark 2.4.20. If $U_W \in 2RV_\infty(1/\alpha, -\rho^\dagger, \eta^\dagger)$ ($\alpha > 0$, $\rho^\dagger > 0$) then by Definition 2.4.8

$$U_W(x) = c^\dagger x^{1/\alpha} \exp\left(\int_1^x \frac{\eta^\dagger(u)}{u}\, du\right); \tag{2.34}$$

where $c^\dagger > 0$ and $\eta^\dagger \in RV_\infty(-\rho^\dagger)$ is either a nonpositive or nonnegative bounded function. Further, by Lemma 2.4.4, $U_W \in 2RV_\infty(1/\alpha, -\rho^\dagger, \eta^\dagger)$ implies

$$\frac{U_W(xt)/U_W(x) - t^{1/\alpha}}{\eta^\dagger(x)} \underset{x\to\infty}{\rightarrow} t^{1/\alpha}\frac{1 - t^{-\rho^\dagger}}{\rho^\dagger}. \tag{2.35}$$

The following result is a re-statement of Theorem 2.3.9 in [23] and holds under the weaker assumption (2.35).

Lemma 2.4.21. *Assume that $U_W \in 2RV_\infty(1/\alpha, -\rho^\dagger, \eta^\dagger)$. Then for all $\epsilon, \delta > 0$ there exists $x_0 = x_0(\epsilon, \delta) > 1$ such that for all $x > x_0$, $t > x_0/x$,*

$$\left|\frac{U_W(xt)/U_W(x) - t^{1/\alpha}}{D_0(x)} - t^{1/\alpha}\frac{1 - t^{-\rho^\dagger}}{\rho^\dagger}\right| \le \epsilon t^{1/\alpha - \rho^\dagger} \max\left(t^\delta, t^{-\delta}\right)\ , \tag{2.36}$$

where

$$D_0(x) = \rho^\dagger \left\{\frac{x^{1/\alpha}}{U_W(x)} \lim_{s\to\infty} \frac{U_W(s)}{s^{1/\alpha}} - 1\right\}\ .$$

Finally, how does second-order regular variation of $\bar{F}_W$ apply to U_W? We were not able to show equivalence between (2.31) and (2.34), however, we can justify equivalence between (2.32) and (2.35). This question is addressed in the next two lemmas (see Exercise 2.11 in [23]).

Lemma 2.4.22. *Assume that $\bar{F}_W$ is strictly decreasing and continuous. If (2.35) holds with $1/\alpha$, $-\rho^\dagger$, $\eta^\dagger$, then (2.32) holds with $-\alpha$, $-\rho = -\alpha\rho^\dagger$, η, where*

$$\eta(x) = \alpha^2 \eta^\dagger(U_W^{\leftarrow}(x)). \tag{2.37}$$

Proof. Assume that (2.35) holds. This is equivalent to writing

$$\frac{U_W(U_W^{\leftarrow}(x)t)/x - t^{1/\alpha}}{\eta^\dagger(U_W^{\leftarrow}(x))} \underset{x\to\infty}{\rightarrow} t^{1/\alpha}\frac{1 - t^{-\rho^\dagger}}{\rho^\dagger}.$$

For $x > 0$, let

$$G_x(t) = \frac{U_W(U_W^{\leftarrow}(x)t)}{x}, \quad t > 0 .$$

Since $\bar{F}_W$ is strictly decreasing and continuous, we have

$$G_x^{\leftarrow}(t) = \frac{U_W^{\leftarrow}(xt)}{U_W^{\leftarrow}(x)} = \frac{\bar{F}_W(x)}{\bar{F}_W(xt)},$$

where the last equality holds by exploiting (2.21a). So, Vervaat's Lemma 2.2.5 yields

$$\frac{\bar{F}_W(x)/\bar{F}_W(xt) - t^{\alpha}}{\eta^{\dagger}(U_W^{\leftarrow}(x))} \underset{x\to\infty}{\to} -t^{\alpha}\frac{1-t^{-\alpha\rho^{\dagger}}}{\rho^{\dagger}/\alpha}.$$

As a consequence, the desired result follows by the Taylor expansion:

$$\frac{\bar{F}_W(xt)/\bar{F}_W(x) - t^{-\alpha}}{\alpha^2\eta^{\dagger}(U_W^{\leftarrow}(x))} \underset{x\to\infty}{\to} t^{-2\alpha}\left(-t^{\alpha}\frac{1-t^{-\alpha\rho^{\dagger}}}{\rho^{\dagger}\alpha^2/\alpha}\right) = t^{-\alpha}\frac{1-t^{-\alpha\rho^{\dagger}}}{\alpha\rho^{\dagger}}.$$

□

Lemma 2.4.23. *Assume that $\bar{F}_W$ is strictly decreasing and continuous. If (2.32) holds with $-\alpha$, $-\rho$, η, then (2.35) holds with $1/\alpha$, $-\rho^{\dagger} = -\rho/\alpha$, $\eta^{\dagger}$, where*

$$\eta^{\dagger}(x) = \frac{\eta(\bar{F}_W^{\leftarrow}(x))}{\alpha^2}. \tag{2.38}$$

Proof. Assume that $\bar{F}_W \in 2RV_{\infty}(-\alpha, -\rho, \eta)$. This implies that

$$\frac{\bar{F}_W(xt)/\bar{F}_W(x) - t^{-\alpha}}{\eta(x)} \underset{x\to\infty}{\to} t^{-\alpha}\frac{1-t^{-\rho}}{\rho}.$$

This is equivalent to writing

$$\frac{\bar{F}_W(\bar{F}_W^{\leftarrow}(x)t)/x - t^{-\alpha}}{\eta(\bar{F}_W^{\leftarrow}(x))} \underset{x\to\infty}{\to} t^{-\alpha}\frac{1-t^{-\rho}}{\tilde{\rho}}.$$

Under the assumption $\bar{F}_W$ is strictly decreasing and continuous, we have

$$\left(\frac{\bar{F}_W(\bar{F}_W^{\leftarrow}(x)\cdot)}{x}\right)^{\leftarrow} = \frac{\bar{F}_W^{\leftarrow}(x\cdot)}{\bar{F}_W^{\leftarrow}(x)} = \frac{U_W(x)}{U_W(x\cdot)},$$

where the last equality holds by virtue of (2.21a). So, Vervaat's Lemma 2.2.5 yields

$$\frac{U_W(x)/U_W(xt) - t^{-1/\alpha}}{\eta(\bar{F}_W^{\leftarrow}(x))} \underset{x\to\infty}{\to} t^{-1/\alpha}\frac{1-t^{\rho/\alpha}}{\alpha\rho}.$$

Consequently, the desired result follows by the Taylor expansion

$$\frac{U_W(xt)/U_W(x) - t^{1/\alpha}}{\eta(\bar{F}_W^{\leftarrow}(x))/\alpha^2} \underset{x\to\infty}{\to} -t^{2/\alpha}\left(t^{-1/\alpha}\frac{1-t^{\rho/\alpha}}{\alpha\rho/\alpha^2}\right) = -t^{1/\alpha}\frac{t^{\rho/\alpha}-1}{\rho/\alpha}.$$

□

We summarize this section as follows. The main point of the corollary below is that in the following sections we will work with one single assumption on second order regular variation that will imply the desired rate of convergence conditions for $\bar{F}_W$ and U_W.

Corollary 2.4.24. *Assume that $\bar{F}_W$ is strictly decreasing and continuous. If $\bar{F}_W \in 2RV_\infty(-\alpha, -\rho, \eta)$ then (2.32) and (2.35) hold, the latter with $1/\alpha$, $-\rho^\dagger = -\rho/\alpha$ and $\eta^\dagger$ given in (2.38).*

2.5 Weak Convergence of Probability Measures

This section is organized as follows. We briefly present random elements of metric spaces. Then, we discuss the Skorokhod topology and convergence of functionals of stochastic processes. We wrap up with the following concepts: martingales, stationarity and ergodicity.

2.5.1 Random Elements of Metric Spaces

Let (S, d) be a metric space. As a regularity condition, we assume that (S, d) is *separable*, meaning that there is a countable dense subset K such that

$$\forall x \in S,\ \forall \epsilon > 0,\ \exists\ y \in K \text{ such that } d(x, y) < \epsilon. \tag{2.39}$$

Let $\mathcal{O}$ be the class of *open subsets* of S. Define the Borel σ-field $\mathcal{B}(S)$ to be the smallest σ-field generated by $\mathcal{O}$, that is, $\mathcal{B}(S) = \sigma(\mathcal{O})$. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space.

Definition 2.5.1. *[51, p.77]*
A random element X of $(S, \mathcal{B}(S))$ is a measurable mapping from $(\Omega, \mathcal{F}, P)$ to $(S, \mathcal{B}(S))$.

The probability distribution of X is the image probability measure, $P \circ X^{-1}$, on $(S, \mathcal{B}(S))$, induced by X , that is, for all $A \in \mathcal{B}(S)$,

$$P \circ X^{-1}(A) = P(X \in A). \tag{2.40}$$

where P is the probability measure on the underlying probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.5.2. *[51, p.77]*
Let (S, d) be a separable metric space endowed with the Borel σ-field $\mathcal{B}(S)$ on S. We say

that a sequence of probability measures $(P_n)_n$ on (S,d) converges weakly to a probability measure P on (S,d), and we write $P_n \Rightarrow P$, if

$$\int_S f\,dP_n \underset{n\to\infty}{\to} \int_S f\,dP, \tag{2.41}$$

for all f in $C(S)$, the space of all continuous bounded real valued functions on S.

The metric d enters in by determining which functions f are continuous on S.

Definition 2.5.3. *[51, p.77]*
We say that a sequence of random elements $(X_n)_n$ of a metric space (S,d) converges in distribution or converges weakly to a random element X of (S,d), and we write $X_n \Rightarrow X$, if their corresponding image probability measures converge weakly, that is, if

$$P\circ X_n^{-1} \Rightarrow P\circ X^{-1} \text{ on } (S,d). \tag{2.42}$$

Consequently, by virtue of Definition 2.5.3, it holds that

Corollary 2.5.4. *[51, p.78]*

$$X_n \Rightarrow X \Leftrightarrow E\left(f\left(X_n\right)\right) \underset{n\to\infty}{\to} E\left(f\left(X\right)\right),\ \text{for all } f\in C(S). \tag{2.43}$$

Theorem 2.5.5. ***(Skorokhod representation theorem)****[51, p.78]*
Suppose that X_n and X are random elements of $(S,\mathcal{B}(S))$. Moreover if,

$$X_n \Rightarrow X \text{ in } (S,d).$$

Then, there exist random elements Y_n and Y, defined on $(\Omega,\mathcal{F},P)$, such that

$$Y_n \overset{d}{=} X_n,\ Y \overset{d}{=} X, \tag{2.44}$$

$$Y_n \xrightarrow[n\to\infty]{a.s} Y, \tag{2.45}$$

where $\overset{d}{=}$ is understood to mean equality in distribution.

Theorem 2.5.6. ***(Cramer-Wold device)****[51, p.104]*
An arbitrary random vector $(X_{n,1},\dots,X_{n,k})$ in $\mathbb{R}^k$ converges in distribution to random vector $(X_1,\dots,X_k)$ in $\mathbb{R}^k$, and we write

$$(X_{n,1},\dots,X_{n,k}) \Rightarrow (X_1,\dots,X_k) \text{ in } \mathbb{R}^k, \tag{2.46}$$

if and only if, for all $a_1,\dots,a_k\in\mathbb{R}$,

$$\sum_{j=1}^k a_j X_{n,j} \Rightarrow \sum_{j=1}^k a_j X_j \text{ in } \mathbb{R}. \tag{2.47}$$

2.5.2 Skorokhod Topology

We present in this subsection spaces $D[0,1]$, $D[a,b]$ and $D(0,\infty)$ endowed with the Skorokhod J_1 topology. These spaces are suitable when it comes to dealing with weak convergence of stochastic processes. In this exposure, the standard Skorokhod space $D[0,1]$ is our focal point. Indeed, results on this functional space can be extended to more general aforementioned Skorokhod's spaces (cf. Remark 2.5.13). See [10], [12].

Definition 2.5.7. *[51]*
An element of $D = D[0,1]$ is a right continuous $\mathbb{R}$-valued function with left limits defined on $[0,1]$. We refer to such a function as càdlàg, a French acronym standing for continue à droite, limite à gauche.

To define one of the distances between functions in D, let:

- e be the identity map on $[0,1]$ and
- $\Lambda = \{\lambda \in [0,1]^{[0,1]} : \lambda \text{ is strictly increasing }, \lambda \text{ and } \lambda^{\leftarrow} \text{ are continuous}\}$.

Then, the distance between two functions in D is measured by the *standard J_1 metric*, d_{J_1}, defined as follows:

$$d_{J_1}(x,y) = \inf_{\lambda\in\Lambda}\{\|x\circ\lambda - y\| \vee \|\lambda - e\|\}, \tag{2.48}$$

where $a \vee b = \max(a,b)$ and $\|x\| = \sup\limits_{0\le t\le 1}|x(t)|$ is the uniform norm on $C[0,1]$, the space of continuous functions on $[0,1]$.

The following theorem provides concrete criteria for convergence in distribution of random elements X_n and X of the space D. It is a simplified version of Theorem 15.6 of [10], appropriate when the limit X is continuous.

Theorem 2.5.8. *[10, p. 28], [43, Lemma 2.5.6]*
Let $(X_n)_n$ be a sequence of random elements with values in $D[0,1]$. Suppose that

$$(X_n(t_1), \ldots X_n(t_k)) \xrightarrow{d} (X(t_1), \ldots, X(t_k)), \tag{2.49}$$

for all $t_1, \ldots, t_k \in [0,1]$, that $P(X(1) \neq X(1-)) = 0$, and that

$$P(|X_n(t) - X_n(t_1)| \geq \lambda, |X_n(t_2) - X_n(t)| \geq \lambda) \leq \frac{1}{\lambda^{2\gamma}}(F_n(t_2) - F_n(t_1))^{2\alpha} \tag{2.50}$$

for $t_1 \leq t \leq t_2$ and $n \geq 1$, where $\gamma \geq 0$, $\alpha > \frac{1}{2}$, and the functions F_n are all monotone increasing or all monotone decreasing, and the sequence $(F_n)_n$ converges (uniformly on $[0,1]$) to a monotone continuous function F. Then

$$X_n \Rightarrow X \text{ in } D[0,1]. \tag{2.51}$$

In what follows, we present a criterion under which the limit of distribution functions remains a distribution function. This property is known as tightness.

Definition 2.5.9. *[12]*
A sequence of distribution functions $(F_n)_n$ is said to be tight if

$$\forall \epsilon > 0,\ \exists\ a, b\ (a < b) : \forall n \geq 1, F_n(a) < \epsilon \text{ and } F_n(b) > 1 - \epsilon.$$

Definition 2.5.10. *[10]*
A sequence of probability measures $(\mu_n)_n$ on a separable metric space is said to be tight if for all $\epsilon > 0$, there exists a compact set $K = K(\epsilon)$ such that

$$\mu_n(K) > 1 - \epsilon.$$

The following theorem is due to Prokhorov and provides the necessary and sufficient conditions for tightness.

Theorem 2.5.11. *[12]*
A sequence of probability measures $(\mu_n)_n$ on a separable metric space is tight if and only if for any subsequence μ_{n_k}, there exists a further subsequence that converges weakly to a probability measure.

Corollary 2.5.12. *[10]*
Let $(\mu_n)_n$ be a tight sequence of probability measures on a separable metric space. Assume that any convergent subsequence μ_{n_k} converges weakly to the same probability measure μ. Then μ_n converges weakly to μ.

The following criterion, in terms of moments, is an alternative working version of (2.50) when it comes to checking tightness on $D[0,1]$:

$$E\left(|X_n(t) - X_n(t_1)|^{\gamma} |X_n(t_2) - X_n(t)|^{\gamma}\right) \leq \left(F_n(t_2) - F_n(t_1)\right)^{2\alpha}. \tag{2.52}$$

Remark 2.5.13. More generally, for A, any subinterval (open or closed) of $(0,\infty)$, $D(A)$ is the class of càdlàg real-valued functions on A. The previous definitions and results are easily extended from $D[0,1]$ to $D[a,b]$, for all $0 < a < b < \infty$. Let $(X_n)_n$ and X be elements of $D(0,\infty)$ and let $X_n 1\!\!1_{[a,b]}$ and $X 1\!\!1_{[a,b]}$ be their respective restrictions to $[a,b]$, $0 < a < b < \infty$. If for all $0 < a < b < \infty$,

$$X_n 1\!\!1_{[a,b]} \Rightarrow X 1\!\!1_{[a,b]} \text{ in } D[a,b], \tag{2.53}$$

then as in Section 16 of [12], it follows that

$$X_n \Rightarrow X \text{ in } D(0,\infty). \tag{2.54}$$

This criterion is appropriate when the limit, X, is continuous, as will always be the case in this thesis.

We end this discussion by stating an alternative tightness criterion proposed in [22]. This will be used later in this thesis.

Theorem 2.5.14. *[22, p.2]*
Let $(\xi_n)_{n\geq 1}$ be real valued stochastic processes defined on $[0,1]$ and whose paths are in the Skorokhod space $D[0,1]$ almost surely. Furthermore, let all the finite dimensional distributions of $(\xi_n)_n$ converge, as $n \to \infty$, to the corresponding ones of a process ξ. Assume that there are constants $1 < \delta \leq \gamma$, $c > 0$, and a nonnegative sequence $c_n \to 0$, as $n \to \infty$ such that, for all $n \geq 1$, we have

$$E\left(|\xi_n(0)|^\gamma\right) \leq c, \tag{2.55}$$

$$E\left(|\xi_n(t) - \xi_n(s)|^\gamma\right) \leq c|t-s|^\delta, \tag{2.56}$$

whenever $|t-s| \geq c_n$. Furthermore, assume that the processes $(\xi_n)_n$ can be written as the differences of nondecreasing processes $(\xi_n^\circ)_n$ and $(\xi_n^{\circ\circ})_n$, and let the processes $(\xi_n^{\circ\circ})_n$ be such that:

$$\max_{j=1,\dots,l_n} |\xi_n^{\circ\circ}(t_{j+1}) - \xi_n^{\circ\circ}(t_j)| \xrightarrow[n\to\infty]{p} 0, \tag{2.57}$$

where $t_j = jc_n$, for all $j = 0,1,\dots,l_n$ with $l_n := [1/c_n]$ and $t_{l_n+1} := 1$. Then the processes $(\xi_n)_n$ converge weakly to ξ in $D[0,1]$. Moreover, the limiting stochastic process ξ has continuous paths almost surely.

Remark 2.5.15. We make the following observations:

1. The statement of Theorem 2.5.14 is also valid when $\xi_n, n \geq 1$ can be written as the differences of nonincreasing processes ξ_n° and $\xi_n^{\circ\circ}$.

2. The statement of Theorem 2.5.14 is also valid for $D[a,b]$, $-\infty < a < b < \infty$, with

$$c = c_{a,b},\ t_j = a + jc_n,\ j = 0, 1, \dots, l_n$$

where l_n is the integer part of $b - a/c_n$ that is $l_n := [(b-a)/c_n]$ and $t_{l_n+1} := b$.

2.5.3 Convergence of Functionals of Stochastic Processes

The power of weak convergence theory comes from the fact that once a basic weak convergence result has been established, then many other weak limits can be derived from it, often using only continuity.

Theorem 2.5.16. *[51, p.85]*
If $X_n \Rightarrow X$ in (S,d) and $g : (S,d) \to (S', d')$ is continuous, then

$$g(X_n) \Rightarrow g(X) \text{ in } (S', d'). \tag{2.58}$$

Lemma 2.5.17. *[12]*
If $X_n \xrightarrow[n\to\infty]{d} X$ and $Y_n \xrightarrow[n\to\infty]{P} a$, where $a \in \mathbb{R}$. Then,

$$(X_n, Y_n) \xrightarrow[n\to\infty]{d} (X, a). \tag{2.59}$$

The next result is Lemma 2.5.10 in [43].

Lemma 2.5.18. *Let $D_1(0,\infty)$ be the set of non-increasing functions in $D(0,\infty)$ and $C_1(0,\infty)$ be the set of continuous non-increasing and positive functions in $C(0,\infty)$. Then, any map defined in $D(0,\infty) \times C_1(0,\infty)$ is continuous at functions in $C(0,\infty) \times C_1(0,\infty)$.*

The next result is Corollary 2.5.11 in [43].

Corollary 2.5.19. *For each n, let Γ_n be random elements of $D(0,\infty)$ and Φ_n random elements of $D_1(0,\infty)$. Suppose that*

$$(\Gamma_n, \Phi_n) \xrightarrow[n\to\infty]{d} (\Gamma, \Phi)$$

and $P(\Gamma \in C(0,\infty)) = P(\Phi \in C_1(0,\infty)) = 1$. Then,

$$\Gamma_n \circ \Phi_n \xrightarrow[n\to\infty]{d} \Gamma \circ \Phi \text{ in } D(0,\infty).$$

Theorem 2.5.20. *[11, p.332]*
Let $M \geq 1$ and $(X_n^{(M)}, Y_n)_n$ be a sequence of random elements of a metric space (S, d). If for each M, and for all $\epsilon \geq 0$, the following hold

$$X_n^{(M)} \Rightarrow X^{(M)} \text{ as } n \to \infty, \tag{2.60a}$$

$$X^{(M)} \Rightarrow X \text{ as } M \to \infty, \tag{2.60b}$$

$$\lim_{u\to\infty} \limsup_{n\to\infty} P\left(d\left(X_n^{(M)}, Y_n\right) \geq \epsilon\right) = 0. \tag{2.60c}$$

Then, as $n \to \infty$, Y_n converges weakly to X, that is, $Y_n \Rightarrow X$.

2.6 Stationary Processes

In this section, we discuss a number of concepts about stationary processes with finite second moment.

2.6.1 Stationarity and Ergodicity

In this subsection, we briefly present two properties of time series: stationarity and ergodicity.

Definition 2.6.1. *[33, p.328]*
A random sequence $(X_n)_n$ is strictly stationary if for every $k \geq 1$, the shifted sequence $(X_{n+k})_n$ has the same distribution, that is, for each m,

$$(X_k, \ldots, X_{k+m}) \stackrel{d}{=} (X_0, \ldots, X_m).$$

The sequence $(X_n)_n$ is said to be of second order if $Var(X_j) < \infty$, $\forall\ j$.

Definition 2.6.2. *[33, p.328]*
A stationary sequence $(X_n)_n$ is ergodic if every shift-invariant event has probability 0 or 1.

Theorem 2.6.3. *[33, p.328]*
Let $(X_n)_n$ be a strictly stationary and ergodic sequence and f be a measurable function. If $E(|f(X)|) < \infty$, then with probability 1,

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \underset{n\to\infty}{\to} E(f(X_0)). \tag{2.61}$$

Theorem 2.6.4. *[46, p.10]*
If $(X_n)_n$ is a strictly stationary and ergodic sequence and $g : \mathbb{R}^\infty \to \mathbb{R}$ is a measurable transformation such that $Y_k = g(X_k, X_{k-1}, \ldots)$, then $(Y_n)_n$ is a strictly stationary ergodic sequence.

2.6.2 Linear Processes

Definition 2.6.5. *[6, p. 43-44]*
Let $\{\varepsilon_j, j \in \mathbb{Z}\}$ be a sequence of independent and identically distributed random variables. A causal linear process or infinite order moving average process $\{X_j, j \in \mathbb{Z}\}$ is defined by

$$X_j = \mu + \sum_{k=0}^{\infty} a_k \varepsilon_{j-k},$$

where $\mu \in \mathbb{R}$ and $\{a_k, k \in \mathbb{Z}\}$ is a sequence of real numbers.

For simplicity, we assume hereafter that $\mu = 0$. Causality is an important property for predicting future values of the process. The next result is an immediate consequence of Theorem 2.6.4.

Theorem 2.6.6. *[46, p. 10]*
Let $\{\varepsilon_j, j \in \mathbb{Z}\}$ be a sequence of independent and identically distributed random variables with mean zero and finite variance. If $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$, then the linear transformation

$$X_k = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{k-j}$$

is a second-order strictly stationary ergodic process.

2.6.3 Martingales

For the purpose of this thesis, we only present two results: Rosenthal's inequality in Theorem 2.6.8 and the martingale Central Limit Theorem 2.6.9.

Definition 2.6.7. *[46, p. 11-12]*
Let $(X_n)_n$ be a sequence of integrable random variables on a probability space $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}_k\}$ be an increasing sequence of sub-σ-algebras of $\mathcal{F}$ such that X_k is $\mathcal{F}_k$-measurable.

- *A sequence* $\{X_k\}$ *is a martingale relative to* $\mathcal{F}_k$ *if for all* $k = 1, 2, \ldots$

$$E\left(X_k \| \mathcal{F}_{k-1}\right) = X_{k-1}, \ a.s.$$

- *A sequence* $\{X_k\}$ *is a martingale difference if*

$$E\left(X_k \| \mathcal{F}_{k-1}\right) = 0, \ a.s.$$

Theorem 2.6.8 (Rosenthal's inequality)**.** *[36, p.23-24]*
If $\{(X_n, \mathcal{F}_n)\}$ *is a martingale difference sequence,* $S_n = \sum_{i=1}^{n} X_i$*, and* $2 \leq p < \infty$*, then there exists a constant* C_p *depending only on* p *such that*

$$E\left(|S_n|^p\right) \leq C_p \left(E\left(\sum_{i=1}^{n} E\left(X_i^2 \| \mathcal{F}_{i-1}\right) \right)^{\frac{p}{2}} + \sum_{i=1}^{n} E\left(|X_i|^p\right) \right) .$$

Theorem 2.6.9. *[38]*
Let $\{(X_{nk}, \mathcal{F}_{n,k})\}$*,* $k = 1, 2, \ldots; n = 1.2, \ldots$ *be an array of random variables such that for each* n*,* $\{(X_{nk}, \mathcal{F}_{n,k})\}_k$ *is a martingale difference sequence. If for all* $\epsilon > 0$*,*

$$\begin{cases} \displaystyle\sum_{j=1}^{k_n} E\left(X_{nj}^2 \mathbb{1}_{\{|X_{nj}| > \epsilon\}} \| \mathcal{F}_{n,j-1}\right) \xrightarrow[n\to\infty]{p} 0, \\ \displaystyle\sum_{j=1}^{k_n} Var\left(X_{nj}^2 \| \mathcal{F}_{n,j-1}\right) \xrightarrow[n\to\infty]{p} 1. \end{cases} \tag{2.62}$$

Then,

$$\sum_{j=1}^{k_n} X_{nj} \xrightarrow[n\to\infty]{d} \mathcal{N},$$

where $\mathcal{N}$ *stands for a standard normal random variable.*

2.7 Long Memory Processes

Trend, seasonality, cycles, stationarity, are commonly accepted qualitative features of time series. In the early eighties, long memory was added to this list thanks to the contribution of Granger et al. [35]. Since then, it is well known that time series tend to have memories: short and long. Our focal point in this thesis is long memory. There are various mathematical ways of defining the existing types of memories of a time series. For instance in the time domain approach, measures of dependence such as autocovariance and autocorrelation functions are used. In the frequency approach, the notion of spectral density is used. In this thesis, we only consider the time domain approach.

2.7.1 Different Types of Memory

Definition 2.7.1. *[5, p.42]*
Let $\{X_j, j \in \mathbb{Z}\}$ be a second order stationary process with autocovariance function (ACF)

$$\gamma(k) = Cov(X_j, X_{j+k}) = Cov(X_0, X_k).$$

- *$\{X_j, j \in \mathbb{Z}\}$ is said to have long memory (long range dependence) if its ACF, is not absolutely summable, that is*

$$\sum_{k=-\infty}^{\infty} |\gamma(k)| = \infty.$$

- *$\{X_j, j \in \mathbb{Z}\}$ is said to have short memory if its ACF is absolutely summable, that is*

$$\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty. \tag{2.63}$$

For the purposes of this thesis, we confine our attention to a more restrictive definition of long range dependence.

Definition 2.7.2. *[5, p.42]*
Let $\{X_j, j \in \mathbb{Z}\}$ be a second order stationary process with autocovariance function γ. If there exists a long memory parameter $d \in (0, 1/2)$ and a slowly varying function ℓ_γ such that

$$\gamma(k) = Cov(X_j, X_{j+k}) \sim k^{-(1-2d)}\ell_\gamma(k), \tag{2.64}$$

then $\{X_j, j \in \mathbb{Z}\}$ is called a stationary process with long memory.

Next, we apply these definitions to second-order stationary causal linear processes.

Definition 2.7.3. *[6, p.45]*
Let $\{X_j, j \in \mathbb{Z}\}$ be a second-order stationary causal linear process.

- $\{X_j, j \in \mathbb{Z}\}$ *has long memory if*

$$a_j = j^{d-1}\ell_a(j), \quad (0 < d < 1/2), \tag{2.65}$$

where ℓ_a is a slowly varying function.

- $\{X_j, j \in \mathbb{Z}\}$ *has short memory if*

$$\sum_{j=0}^{\infty} |a_j| < \infty \text{ and } \sum_{j=0}^{\infty} a_j \neq 0. \tag{2.66}$$

Remark 2.7.4. Note that persistence, strong dependence and long-range dependence are alternative terminologies for long memory. Short memory means that correlations are quickly decaying. On the other hand, long memory means that correlations are slowly decaying.

2.7.2 ARMA and FARIMA Models

In this subsection, we present examples of short- and long-memory processes such as the ARMA and FARIMA models. We focus on second order processes only.

Definition 2.7.5. *[49]*
A time series $\{X_j, j \in \mathbb{Z}\}$ is said to be an AutoRegressive Moving Average process with orders p and q and we write ARMA(p,q) if it is stationary and

$$\phi_p(B)X_j = \theta_q(B)\varepsilon_j, \tag{2.67}$$

where $\{\varepsilon_j, j \in \mathbb{Z}\}$ is sequence of i.i.d. random variables with mean zero and finite variance, B is the backward shift operator, that is

$$B^j X_i = B^{j-1}(BX_i) = X_{i-j}, \quad j = 0, \pm1, \pm2, \ldots \tag{2.68}$$

and where the autoregressive and moving average operators are respectively defined by

$$\phi_p(B) = 1 - \sum_{j=1}^{p} \phi_j B^j \text{ and } \theta_q(B) = 1 + \sum_{j=1}^{p} \theta_j B^j \tag{2.69}$$

and are assumed to have no common roots.

Second order ARMA models have summable covariances and hence have short memory [49].

FARIMA, or Fractionally AutoRegressive Integrated Moving Average, models were introduced by Granger et al. [35], back in the early eighties. These models are alternatively called ARFIMA (AutoRegressive Fractionally Integrated Moving Average).

Definition 2.7.6. *[6]*
A time series $\{X_j, j \in \mathbb{Z}\}$ is said to be a Fractionally AutoRegressive Integrated Moving Average process and we write FARIMA(p, d, p) if

$$\phi_p(B)(1 - B)^d X_j = \theta_q(B)\varepsilon_j \tag{2.70}$$

where $\{\varepsilon_j, j \in \mathbb{Z}\}$ is sequence of i.i.d. random variables with mean zero and finite variance, $d \in (0, 1/2)$ is a rational number called the differencing order or long memory parameter.

The FARIMA$(0, d, 0)$ model is a basic example of a long memory time series [6], [46].

Remark 2.7.7. The rationale behind the name FARIMA is that $(1-B)^d X_j \sim ARMA(p, q)$.

Notice that if $\phi_p(z) = \theta_q(z) = 1$, then we obtain the simplest autoregressive fractionally integrated model known as the factional noise and defined as follows:

$$(1 - B)^d\, X_j = \varepsilon_j \; . \tag{2.71}$$

Thus

$$X_j = (1 - B)^{-d}\, \varepsilon_j = \sum_{i=1}^{\infty} \psi_i B^i \varepsilon_j = \sum_{i=1}^{\infty} \psi_i \varepsilon_{j-i} \; . \tag{2.72}$$

The range for the differencing parameters is $0 < d < 1/2$. The closer the value of d to $1/2$, the higher the intensity of long memory displayed by the model. When $d = 0$, the classical ARMA model is recovered. For $d > 1/2$, the model is nonstationary.

2.7.3 Hermite Polynomials

In this subsection, we present some analytical and probabilistic properties of Hermite polynomials. These orthogonal polynomials are essential for the derivation of limit theorems of Gaussian long memory time series.

Definition 2.7.8. *[5, p.68]*
The k-th Hermite polynomial is defined to be

$$H_k(x) := (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}, \; x \in \mathbb{R}, \; k = 0, 1, 2, \ldots \tag{2.73}$$

Note that (2.73) is known as the Rodrigues Formula. Here are some Hermite polynomials:

$$H_0(x) = 1, H_1(x) = x, H_2(x) = x^2 - 1, H_3(x) = x^3 - 3x.$$

The Hermite polynomials satisfy a number of analytical properties that we will explore shortly.

Lemma 2.7.9. *[45] The set $(H_k)_k$ of Hermite polynomials satisfies*

$$H_{k+1}(x) = xH_k(x) - kH_{k-1}(x), \; k = 1, 2, \ldots$$

Lemma 2.7.10. *[45] The set $(H_k)_k$ of Hermite polynomials satisfies*

$$H_k(x) = k! \sum_{n=0}^{[k/2]} \frac{(-1)^n}{2^n n!(k-2n)!} x^{k-2n}, \tag{2.74}$$

where $[x]$ denotes the integer part of the real number x.

Lemma 2.7.11. *The set $(H_k)_k$ of Hermite polynomials satisfies*

$$H_k(-x) = (-1)^k H_k(x).$$

Proof. The Hermite expansion (2.74) yields for all $x \in \mathbb{R}$, that

$$H_k(-x) = k! \sum_{n=0}^{[k/2]} \frac{(-1)^n}{2^n n!(k-2n)!} (-x)^{k-2n} = k! \sum_{n=0}^{[k/2]} \frac{(-1)^{n+k-2n}}{2^n n!(k-2n)!} x^{k-2n}$$
$$= k! \sum_{n=0}^{[k/2]} \frac{(-1)^n (-1)^k (-1)^{-2n} x^{k-2n}}{2^n n!(k-2n)!} = (-1)^k k! \sum_{n=0}^{[k/2]} \frac{(-1)^n x^{k-2n}}{2^n n!(k-2n)!} = (-1)^k H_k(x).$$

$\square$

It follows from Lemma 2.7.11 that any Hermite polynomial of the form H_{2p} is even and any of the form H_{2p+1} is odd, where p is a positive integer.

Lemma 2.7.12. *[6, p.111]*
The set $(H_k)_k$ of Hermite polynomials forms an orthogonal basis in $L^2(\mathbb{R})$ with respect to the weight function

$$\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \tag{2.75}$$

Remark 2.7.13. The rationale behind Lemma 2.7.12 is that

$$\langle H_j(x), H_k(x)\rangle = \int_{-\infty}^{\infty} H_j(x)H_k(x)\phi(x)dx = j!\delta_{jk}, \tag{2.76}$$

where δ_{jk} stands for the Kronecker's symbol. Recall that $L^2(\mathbb{R})$ endowed with the scalar product $\langle\cdot,\cdot\rangle$ above forms a Hilbert space.

We now turn our attention to some probabilistic properties of Hermite polynomials. Most importantly, we introduce the concept of Hermite rank in Definition 2.7.15, the cornerstone for limit theorems for Hermite polynomials Theorems 2.7.22 and 2.7.23.

Corollary 2.7.14. *[6, p.113, p.22]*
If X is a standard normal random variable, then for $k > 0$

$$E\left(H_k(X)\right) = 0\ , \quad Var\left(H_k(X)\right) = k!\ , \quad Cov\left(H_j(X), H_k(X)\right) = j!\delta_{jk}. \tag{2.77}$$

If (X_1, X_2) is a pair of standard normal random variables with $Cov(X_1, X_2) = \rho$, then

$$Cov(H_j(X_1), H_k(X_2)) = \begin{cases} k!\rho^k & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{2.78}$$

Definition 2.7.15. *[6, p.112]*
Let $X \sim N(0,1)$ and G be a measurable function such that:

$$E\left(G(X)\right) = 0 \text{ and } E\left(G^2(X)\right) < \infty. \tag{2.79}$$

The Hermite rank m of G is defined to be

$$m = \inf\{k \in \mathbb{N}^* : E\left(H_k(X)G(X)\right) \neq 0\}, \tag{2.80}$$

where $\mathbb{N}^* = \{1, 2, 3, \ldots\}$.

Lemma 2.7.16. *[6, p.111]*
Let X be a standard normal random variable and G be a measurable function such that (2.79) holds. Then $G(X)$ is uniquely represented in $L^2(\mathbb{R})$ by:

$$G(X) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k!} H_k(X). \tag{2.81}$$

The sequence, $(\mu(k))_k$, of Hermite coefficients, is defined by

$$\mu(k) := \langle H_k(X), G(X) \rangle = E\left(H_k(X)G(X)\right). \tag{2.82}$$

Remark 2.7.17. If the Hermite rank of G is m, then the Hermite series of G is

$$G(X) = \sum_{k=m}^{\infty} \frac{\mu(k)}{k!} H_k(X). \tag{2.83}$$

Example 2.7.18. To illustrate the concept of Hermite rank, we first recall that if $X \sim N(0,1)$, then $E(G(X)) = 0$, for any odd measurable function G. Moreover, the k-th moment of X satisfies for all $k, p \in \mathbb{N}$,

$$E\left(X^k\right) = \begin{cases} \frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} & \text{if } k = 2p, \\ 0 & \text{if } k = 2p+1. \end{cases} \tag{2.84}$$

We consider the following examples:

1. Let $X \sim N(0,1)$. Take $G(X) = H_1(X)$. It follows that $E\left(G(X)\right) = 0$. Since

   $$\mu(1) = E\left(H_1(X)G(X)\right) = E\left(X^2\right) = 1 \neq 0,$$

   then, we conclude that G is of Hermite rank $m = 1$.

2. Let $X \sim N(0,1)$. Take $G(X) = H_2(X)$. It follows that $E\left(G(X)\right) = 0$. We have

   $$\mu(1) = E\left(H_1(X)G(X)\right) = E\left(X(X^2-1)\right) = E\left(X^3\right) - E(X) = 0.$$

   Moreover, we have

   $$\mu(2) = E\left(H_2(X)G(X)\right) = E(X^4) - 2E(X^2) + 1 = 2 \neq 0.$$

   Thus, G is of Hermite rank $m = 2$.

3. Since Hermite polynomials are orthogonal, it is easily seen that the Hermite rank of H_m is m.

4. Let $X \sim N(0,1)$. Take $G(X) = H_m(X) + H_k(X)$, with $1 < m < k$. By (2.77),
$$E(G(X)) = 0 \text{ and } \mu(k) = 0, \ 1 \leq k \leq m-1.$$
Consequently, we obtain
$$\mu(m) = E(H_m(X)G(X)) = E\left(H_m^2(X)\right) + E\left(H_m(X)H_k(X)\right) = m! \neq 0.$$
Thus, G is of Hermite rank m.

5. Let $X \sim N(0,1)$. Recall that the moment generating M_X satisfies for all $t \in \mathbb{R}$,
$$M_X(t) = e^{\frac{t^2}{2}} \text{ and } M_X'(t) = E\left(Xe^{tX}\right) = te^{\frac{t^2}{2}}.$$
Take $G(X) = e^X - E\left(e^X\right)$. Trivially, $E(G(X)) = 0$. We have
$$\mu(1) = E\left(H_1(X)G(X)\right) = E\left(Xe^X\right) - E\left(XE\left(e^X\right)\right) = M_X'(1) - E(X)E\left(e^X\right) = \sqrt{e}.$$
Thus, G is of Hermite rank $m = 1$.

The Hermite expansion of transformations of Gaussian random variables makes feasible the computation of their moments.

Corollary 2.7.19. *[6, p. 113]*
Let X be a standard normal random variable and G be a measurable function such that (2.79) holds. If G is of Hermite rank m, then
$$Var(G(X)) = E\left(G^2(X)\right) = \sum_{k=m}^{\infty} \frac{(\mu(k))^2}{k!}. \tag{2.85}$$
If (X_1, X_2) is a pair of standard normal random variables with $Cov(X_1, X_2) = \rho$, then
$$Cov(G(X_1), G(X_2)) = E\left(G(X_1)G(X_2)\right) = \sum_{k=1}^{\infty} \frac{(\mu(k))^2}{k!}\rho^k. \tag{2.86}$$

Proof. Let $X \sim N(0,1)$ and G be a measurable function of Hermite rank m. We
$$Var(G(X)) = Var\left(\sum_{k=m}^{\infty} \frac{\mu(k)}{k!} H_k(X)\right) = \sum_{k=m}^{\infty} \left(\frac{\mu(k)}{k!}\right)^2 Var(H_k(X)) = \sum_{k=m}^{\infty} \frac{(\mu(k))^2}{k!}.$$

Now assuming that X_1 and X_2 are standard normal, we obtain

$$E\left[G(X_1)G(X_2)\right] = \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \frac{\mu_1(i)}{i!}\frac{\mu_2(k)}{k!} E\left[H_i(X_1)H_k(X_2)\right] = \sum_{k=1}^{\infty} \frac{(\mu(k))^2}{k!}\rho^k.$$

Note that the last line holds by virtue of (2.78). $\square$

Definition 2.7.20. *[6, p.194]*
Let $B(\cdot)$ denote a standard Brownian motion on $\mathbb{R}$, m a positive integer and $h \in \mathbb{R}_+$ be such that $1 - 1/2m < h < 1$. A Hermite-Rosenblatt process is a stochastic process $(\xi_{m,h}(u))$ defined for all $u \geq 0$ by,

$$\xi_{m,h}(u) = \omega(m,h) \int_{-\infty}^{\infty}\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_{m-1}} \left(\int_0^u \prod_{j=1}^{m} (s - x_i)_+^{h-\frac{3}{2}} ds \right) dB(x_m) \cdots dB(x_1),$$

where $x_+ := \max(0, x)$ and $\omega(m,h) > 0$ satisfies

$$\omega^2(m,h) = \frac{m!(2m(h-1)+1)(m(h-1)+1)}{\left(\int_0^\infty [x(x+1)]^{h-\frac{3}{2}} dx\right)^m}.$$

Remark 2.7.21. Note that the choice of the constant $\omega(m,h)$ assures that $E\left(\xi^2_{m,h}(1)\right) = 1$. Due to symmetry of the following function,

$$(x_1 \cdots x_m) \mapsto \int_0^u \prod_{j=1}^{m} (s - x_i)_+^{h-\frac{3}{2}} ds,$$

the Hermite-Rosenblatt process can be alternatively represented as follows:

$$\xi_{m,h}(u) = \frac{\omega(m,h)}{m!} \int_{\mathbb{R}^m} \left(\int_0^u \prod_{j=1}^{m} (s - x_i)_+^{h-\frac{3}{2}} ds \right) dB(x_m) \cdots dB(x_1).$$

If $m = 1$, then $\xi_{1,h}$ is a fractional Brownian motion with the Hurst parameter $\frac{1}{2} < h < 1$.

2.7.4 Limit Theorems for Gaussian Long Memory Processes

We present limit theorems for partial sums associated with a Hermite transformation of a second-order stationary long memory Gaussian sequences.

Theorem 2.7.22. *[6, p.228], [17]*
Let $(Y_j)_j$ be a stationary sequence of standard normal random variables with autocovariance function $\gamma(k) \sim k^{2d-1}\ell_\gamma(k)$, where $0 < d < 1/2$ and ℓ_γ is a slowly varying function at infinity. Denote by H_m be the m-th Hermite polynomial.

- *If* $m(1-2d) < 1$, *then for* $u \in [0,1]$,

$$\frac{1}{n^{1-m(\frac{1}{2}-d)}\sqrt{m!C_m\ell_\gamma^m(n)}}\sum_{j=1}^{[nu]} H_m(Y_j) \xrightarrow[n\to\infty]{d} \xi_{m,h}(u) \text{ in } D[0,1], \tag{2.87}$$

  where $\xi_{m,h}(\cdot)$ *is Hermite-Rosenblatt with* $h = d + 1/2$ *and*

$$C_m = \frac{2}{[1-m(1-2d)][2-m(1-2d)]}. \tag{2.88}$$

- *If* $m(1-2d) > 1$, *then*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} H_m(Y_j) \xrightarrow[n\to\infty]{d} \sigma_m\mathcal{N}\ , \tag{2.89}$$

  where $\mathcal{N}$ *stands for the standard random variable and*

$$\sigma_m^2 = \lim_{n\to\infty}\frac{1}{n}\mathrm{Var}\left(\sum_{j=1}^{n} H_m(Y_j)\right) \in (0,\infty)\ . \tag{2.90}$$

We extend the previous result to an arbitrary measurable transformation of a Gaussian long memory sequence.

Theorem 2.7.23. *[6, p. 223, 229], [17]*
Let $\{Y_j, j \in \mathbb{Z}\}$ *be a stationary sequence of standard normal random variables with autocovariance function* $\gamma(k) \sim k^{2d-1}\ell_\gamma(k)$, *where* $0 < d < 1/2$ *and* ℓ_γ *is a slowly varying function at infinity. Let* G *be a measurable function of Hermite rank* m.

- *If* $m(1-2d) < 1$, *then for* $u \in [0,1]$,

$$\frac{1}{n^{1-m(\frac{1}{2}-d)}\sqrt{m!C_m\ell_\gamma^m(n)}}\sum_{j=1}^{[nu]} G(Y_j) \xrightarrow[n\to\infty]{d} \frac{\mu(m)}{m!}\xi_{m,h}(u) \text{ in } D[0,1]. \tag{2.91}$$

  where C_m *is defined in* (2.88) *and* $\mu(m) = E\left(H_m(Y_0)G(Y_0)\right)$. *The limiting process* $\xi_{m,h}(\cdot)$ *is Hermite-Rosenblatt with* $h = d + 1/2$.

- *If* $m(1-2d) > 1$, *then*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} G(Y_j) \xrightarrow[n\to\infty]{d} \sigma\mathcal{N}\ , \tag{2.92}$$

where $\mathcal{N}$ stands for the standard normal random variable and

$$\sigma^2 = \sum_{k=m}^{\infty} \frac{\mu^2(k)}{k!}\sigma_k^2 < \infty \tag{2.93}$$

and σ_k^2 is defined in (2.90).

Remark 2.7.24. Here are the striking facts about limit theorems under long memory:

- In the case of scenarios (2.89) and (2.92), where the long memory (as measured by the parameter d), is not very strong, then the scaling for partial sums remains the same as in the Central Limit Theorem, that is $n^{1/2}$. We will call the regime (2.92) *weak long memory.*

- However in the case of scenarios (2.87) and (2.91), where correlations decay very slowly, the scaling of partial sums departs from $n^{1/2}$. Actually, it is of the form n^b, with $b = 1 - m(\frac{1}{2} - d) > 1/2$. In a nutshell, in contrast with the weak long memory regime, the long memory effect leads to nonstandard limit theorems. In addition, under long memory one dimensional transformations of the form $G(Y_j)$ contribute to the limiting distribution through the Hermite rank of G. We will call the regime (2.91) *strong long memory.*

- The boundary case $m(1 - 2d) = 1$ is more delicate and will not be discussed. In fact, the limiting random variable is a linear combination of $\xi_{m,h}$ and $\mathcal{N}$. These random variables are uncorrelated but not independent. See [17].

2.8 Leverage Effect

In the financial literature, volatility depicts the magnitude of the price fluctuations during a specified period of time. In other words, it is a measure of how much the price of an asset moves each day (or week or month, and so on). See e.g. [19]. In general, higher volatility is synonymous with higher profit or loss risk.

Now we turn our attention to another empirical behaviour of financial data: the leverage effect. This phenomenon is understood as an asymmetric behaviour of stock prices and is extensively discussed in [20]. The leverage effect refers to the observed tendency of an asset's volatility to be negatively correlated with the asset's returns. Typically, rising asset prices are accompanied by declining volatility, and vice versa. See

[3]. Informally speaking (see [52, p. 167]), if X_j is the return at period j and σ_j^2 is the return volatility at period j, then the leverage effect is a negative relationship between $E\left(\ln \sigma_{j+1}^2 \| X_j\right)$ and X_j.

It follows from this informal definition that previous returns of a portfolio are negatively correlated to future volatility. Note that future volatility is the predicted or expected price fluctuation of a period of time until the option has expired. See [19].

The original modelling approach to the leverage is due to [37]. The authors consider the following model:

$$X_j = \sigma_j Z_j \ , \quad \ln \sigma_j^2 = Y_j \ , \quad Y_j = \rho Y_{j-1} + \epsilon_{j-1} \ ,$$

where $|\rho| < 1$, $\{(\epsilon_j, Z_j)\}$ are i.i.d. normal vectors with mean zero, unit variance and correlation ω. Under this set-up, we have the following representation

$$Y_j = \sum_{i=1}^{\infty} \rho^i \epsilon_{j-i}.$$

Notice that the logarithm of the volatility can be written as

$$Y_{j+1} = \rho Y_j + \omega \sigma_j^{-1} X_j + (\epsilon_j - \omega Z_j) \ .$$

Since (ϵ_j, Z_j) are multivariate normal, then $(\epsilon_j, Z_j) \stackrel{d}{=} (\omega Z_j + W_j, Z_j)$, where W_j is a standard normal random variable. Therefore, we have

$$\begin{aligned} E\left(\ln \sigma_{j+1}^2 \| X_j, \sigma_j\right) &= E\left(\rho Y_j + \omega \sigma_j^{-1} X_j \| X_j, \sigma_j\right) + E(\epsilon_j - \omega Z_j \| X_j, \sigma_j) \\ &= \rho Y_j + \omega \sigma_j^{-1} X_j = \rho \ln \sigma_j^2 + \omega \sigma_j^{-1} X_j. \end{aligned}$$

In fact, we have

$$\begin{aligned} E\left(\epsilon_j - \omega Z_j \| X_j, \sigma_j\right) &= E(\epsilon_j - \omega Z_j \| Z_j, \sigma_j) = E(\epsilon_j - \omega Z_j \| Z_j) \\ &= E\left(\epsilon_j \| Z_j\right) - \omega Z_j = \omega Z_j - \omega Z_j = 0. \end{aligned}$$

Thus, the expected log-volatility is a linear function of X_j whenever $\omega \neq 0$, that is

$$E\left(\ln \sigma_{j+1}^2 \| X_j\right) = \mu + \nu X_j, \tag{2.94}$$

where μ, ν are constants. This is a linear function in X_j. Hence, if $\omega < 0$ and everything else is held constant, a fall in the stock price (return) leads to an increase of future expected log-volatility. As such, (2.94) captures the leverage effect in the model that is *volatility tends to rise in response to bad news but to fall in response to good news.* Other modeling approaches to leverage can be found e.g. in [52] or [16]. The only difference lies in alternative specifications in the equation for $\ln \sigma_j^2$, allowing for an additional random term.

2.9 Concluding Remarks

The quantitative analysis of time series requires mathematical tools such as those presented in this chapter: regular variation, second-order regular variation, second-order stationary processes, long memory processes, Hermite polynomials, as well as leverage.

Since any mathematical model approximating the evolution of asset price should be able to generate long memory, heavy tails and the leverage effect, then in the next chapter we discuss heteroscedastic processes such as stochastic volatility models. In particular, we are going to focus on the class of long memory stochastic volatility models with leverage. The mathematical foundations explored throughout this chapter will help handling their limiting behaviour.

Chapter 3

Long Memory Stochastic Volatility Model with Leverage

3.1 Introductory Comments

Long memory time series have become increasingly popular since the pioneering work of Granger et al. [35]. These models find various applications in hydrology, financial risk management etc. Financial data such as return on investments exhibit three widely accepted features:

1. Returns are uncorrelated, but their squares, or absolute values, are (highly) correlated. Such behaviour is known as *long range dependence or long memory*. We refer to [26] for a detailed discussion about the long memory property of stock market returns.

2. Log-returns are *heavy tailed*, that is - some moments of the log-returns are infinite.

3. Previous returns are negatively correlated with future volatility. Such behaviour is referred to as the *leverage effect*. This means that rising asset prices are accompanied by declining volatility, and vice versa.

These empirical findings have opened the door to stochastic volatility models. In this class of stochastic processes, log-returns $\{X_j\}$ are modeled as follows

$$X_j = \sigma_j Z_j,$$

where $\{Z_j\}$ is an i.i.d. sequence and $\{\sigma_j^2\}$ is the conditional variance or, more generally, a certain process which stands as a proxy for the volatility. In such a process, long memory can only be modelled through the sequence $\{\sigma_j\}$, while the tails can be modelled either through the sequence $\{Z_j\}$ or through $\{\sigma_j\}$, or both. The well known GARCH processes belong to this class of models. The volatility sequence $\{\sigma_j\}$ is heavy tailed unless the distribution of Z_0 has finite support, and leverage can be present. However, long memory of squares cannot be modelled by the GARCH process.

In order to capture this feature, the so-called long memory stochastic volatility (LMSV) model was introduced in [15]. An overview of stochastic volatility models with long range dependence and their basic properties is given in [24] and in [25]. In the classical LMSV model, $\{Z_j\}$ is a sequence of i.i.d. standard normal random variables, independent of the volatility sequence $\{\sigma_j\}$, assumed to be of the form $\sigma_j = \exp(Y_j)$, where $\{Y_j\}$ is a long memory Gaussian sequence. However, the independence assumption excludes the possibility of modelling leverage effects. For this, in the next section we introduce the long memory stochastic volatility model with leverage.

We structure this chapter as follows. In section 3.2, we define the long memory stochastic volatility model and provide its main assumptions. In section 3.3, we examine in what way the heavy-tail and long memory assumptions on Z_j and Y_j, respectively, get transferred to X_j. In section 3.4, we discuss no-bias conditions which will play a major role in subsequent arguments. Finally, we give two examples in section 3.5.

3.2 Model: Description and Assumptions

The long memory stochastic volatility model with leverage is defined by

$$X_j = \phi(Y_j)Z_j \ , \quad j \in \mathbb{Z}. \tag{3.1}$$

A(i) The sequence $\{Y_j\}$ is strictly stationary and ergodic long memory Gaussian,

$$Y_j = \sum_{i=1}^{\infty} a_i \epsilon_{j-i} \ ,$$

where $\{\epsilon_j\}$ is a sequence of i.i.d. standard normal random variables and

$$a_j = j^{d-1}\ell_a(j), \ \ \sum_{j=1}^{\infty} a_j^2 = 1 \ .$$

As a consequence,

$$\gamma_Y(j) = Cov(Y_0, Y_j) \sim j^{2d-1}\ell_Y(j).$$

Note that ℓ_a and ℓ_Y are slowly varying functions at infinity such that:

$$\ell_Y(j) = \ell_a^2(j)B(1-2d, d),$$

where $B(a,b)$ denotes the Beta function and $0 < d < \frac{1}{2}$ is referred to as the long memory parameter (for details, see [6]). Furthermore, $\{(\epsilon_j, Z_j)\}$ is a sequence of i.i.d. random vectors. Let $\{\mathcal{G}_j\}$ be the minimal filtration generated by the random vectors $\{(\epsilon_j, Z_j)\}$, that is

$$\mathcal{G}_j := \sigma\left(\{(\epsilon_k, Z_k) : k \leq j\}\right),\ j \in \mathbb{Z}\,. \tag{3.2}$$

For each fixed j, ϵ_j and Z_j may be dependent, but due to the construction above, the random variables Y_j and Z_j are independent. However, there can be dependence between the sequences $\{Z_j\}$ and $\{Y_j\}$, allowing for *leverage* in the model.

A(ii) The random variables Z_j are i.i.d. and the tail distribution function $\bar{F}_Z = 1 - F_Z \in 2RV_\infty(-\alpha, -\kappa, \eta^*)$, with $\alpha, \kappa > 0$. By (2.31), this means that for all $x > 0$,

$$\bar{F}_Z(x) = c^* x^{-\alpha} \exp\left(\int_1^x \frac{\eta^*(u)}{u}\,du\right) = x^{-\alpha}\ell^*(x), \tag{3.3}$$

where $c^* > 0$ and η^* is either nonnegative or nonpositive, regularly varying with index $-\kappa$ and bounded - that is, there exists $\beta > 0$ such that for all $x > 0$,

$$|\eta^*(x)| \leq \beta. \tag{3.4}$$

A(iii) The function ϕ is a nonnegative measurable function and $\phi(Y_0)$ is not equal to 0 with probability one. In addition, denote the Hermite rank of ϕ^α by m.

A(iv) Let $k_n \to \infty$ be an increasing sequence of positive integers such that $k_n/n \to 0$ and let u_n be defined by $u_n = \bar{F}_X^{\leftarrow}(k_n/n)$ where $\bar{F}_X^{\leftarrow}$ is the inverse function of the tail distribution function $\bar{F}_X$ of X. (As will be argued below, $\bar{F}_X$ is continuous). For ease of notation, we suppress dependence of k_n on n, which is the standard practice in the extreme value literature. For all $n \geq 1$, let $\{a_{n,m}\}$ and $\{b_{n,m}\}$ be such that:

$$a_{n,m} := \left(\sqrt{n\bar{F}_Z(u_n)} + \frac{n}{b_{n,m}}\right)\mathbb{1}_{\{m(1-2d)<1\}} + \sqrt{n}\mathbb{1}_{\{m(1-2d)>1\}}\,, \tag{3.5a}$$

$$b_{n,m} := n^{1-m(\frac{1}{2}-d)}\sqrt{\frac{2m!(\ell_Y(n))^m}{[(2d-1)m+1][(2d-1)m+2]}}\,. \tag{3.5b}$$

$$a_{n,m}\eta^*\left(\bar{F}_X^{\leftarrow}(k/n)\right) = a_{n,m}\eta^*(u_n) \underset{n\to\infty}{\to} 0 \; . \tag{3.6}$$

A(v) For all $\epsilon > 0$, and $\alpha, \beta, \kappa > 0$ as above,

$$E\left((\phi(Y))^{2\alpha+2\beta}\right) + E\left((\phi(Y))^{2\alpha-2\beta}\right) < \infty, \tag{3.7a}$$

$$E\left((\phi(Y))^{\alpha+\kappa+\epsilon}\right) + E\left((\phi(Y))^{\alpha+\kappa-\epsilon}\right) < \infty. \tag{3.7b}$$

Remark 3.2.1. The assumptions A(i)and A(iii) model long memory and leverage, while A(ii) determines the tail behaviour. The remaining assumptions have to do with technicalities. **For clarity throughout the remainder of the thesis, when referring to the long memory stochastic volatility model with leverage, we suppose that all the assumptions from A(i) to A(v) are satisfied. However, some results do not require all the assumptions.**

3.3 Transfer Theorems and Technicalities

In this section, we consider the long memory stochastic volatility model with leverage as in (3.1). We examine in what way assumptions made on Z and Y transfer to X.

Under the assumption A(i), X_j is $\mathcal{G}_j$-adapted, Y_j is $\mathcal{G}_{j-1}$-measurable. Therefore,

$$E\left(\mathbb{1}_{\{X_j>x\}}\|\mathcal{G}_{j-1}\right) = E\left(\mathbb{1}_{\{\phi(Y_j)Z_j>x\}}\|\mathcal{G}_{j-1}\right) = \bar{F}_Z(x/\phi(Y_j)), \quad x > 0 \; . \tag{3.8}$$

This formula will play a crucial role in subsequent proofs.

Transfer of dependence. The next lemma deals with transfer of properties such as stationarity, ergodicity and long range dependence.

Lemma 3.3.1. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then $\{X_j\}$ is a strictly stationary and ergodic sequence. Further, if $Var(Z_j) < \infty$, then*

$$Var\left(X_j\|\mathcal{G}_{j-1}\right) = \left(\phi(Y_j)\right)^2 Var(Z_j) \; , \tag{3.9a}$$

$$Cov(X_0, X_j) = (E(Z_0))^2 Cov(\phi(Y_0), \phi(Y_j)) =: \gamma_X(j) \; , \tag{3.9b}$$

$$Cov\left(X_0^2, X_j^2\right) = (E(Z_0^2))^2 Cov(\phi^2(Y_0), \phi^2(Y_j)) =: \gamma_{X^2}(j) \; . \tag{3.9c}$$

Proof. Stationarity and ergodicity of X_j follows from Theorem 2.6.4. Since Y_j is $\mathcal{G}_{j-1}$-measurable and Z_j is independent from $\mathcal{G}_{j-1}$, then

$$\begin{aligned} Var\left(X_j \| \mathcal{G}_{j-1}\right) &= E\left(\phi^2(Y_j)Z_j^2 \| \mathcal{G}_{j-1}\right) - \left(E\left(\phi(Y_j)Z_j \| \mathcal{G}_{j-1}\right)\right)^2 \\ &= (\phi(Y_j))^2 E(Z_j^2) - (\phi(Y_j))^2 (E(Z_j))^2 \\ &= (\phi(Y_j))^2 \operatorname{Var}(Z_j) \; . \end{aligned}$$

Furthermore,

$$\begin{aligned} \gamma_X(j) &= Cov(X_0, X_j) = Cov\left(\phi(Y_0)Z_0, \phi(Y_j)Z_j\right) \\ &= E\left(\phi(Y_0)\phi(Y_j)Z_0Z_j\right) - E\left(\phi(Y_0)\right)E(Z_0)E\left(\phi(Y_j)\right)E(Z_j) \\ &= E\left(\phi(Y_0)\phi(Y_j)\right)E(Z_0)E(Z_j) - E\left(\phi(Y_0)\right)E(Z_0)E\left(\phi(Y_j)\right)E(Z_j) \\ &= (E(Z_0))^2 Cov(\phi(Y_0), \phi(Y_j)). \end{aligned}$$

The calculation for (3.9c) is similar. □

Remark 3.3.2. It turns out that the conditional variance of this stochastic process is random, which is referred to as stochastic volatility. Under the assumptions of Lemma 3.3.1, $\{X_j\}$ may inherit long memory behaviour from $\{Y_j\}$.

- If $E(Z_0) = 0$, then the random variables $\{X_j\}$ are always uncorrelated, regardless the memory of $\{Y_j\}$;
- If $E(Z_0) \neq 0$, then long memory properties of $\{Y_j\}$ are transferred to $\{X_j\}$ via the covariance function $\gamma_{\phi(Y)}(j) := Cov(\phi(Y_0), \phi(Y_j))$. The behaviour of the latter was studied in Corollary 2.7.19.
- Long memory properties of $\{Y_j\}$ are always transferred to $\{X_j^2\}$ via the covariance function $\gamma_{\phi^2(Y)}(j) := Cov(\phi^2(Y_0), \phi^2(Y_j))$. The behaviour of the latter was studied in Corollary 2.7.19.

Transfer of regular variation. The elementary consequence of A(ii) is that the second-order regular variation of Z also implies that $\bar{F}_Z$ is regularly varying at infinity with index $-\alpha$, that is

$$J_x(t) := \frac{\bar{F}_Z(xt)}{\bar{F}_Z(x)} \underset{x\to\infty}{\to} T(t) =: t^{-\alpha} \; , \tag{3.10}$$

uniformly on compact subsets of $(0,\infty)$. Furthermore, since (3.7a) holds, then Breiman's Lemma 2.3.20 yields that

$$\frac{\bar{F}_X(x)}{\bar{F}_Z(x)} \underset{x\to\infty}{\to} E(\phi^\alpha(Y)). \tag{3.11}$$

This means that the random variables Z and X are *tail equivalent*. In other words, the tail distribution of X is regularly varying with index $-\alpha$ as well - that is $\bar{F}_X \in RV_\infty(-\alpha)$. Therefore,

$$T_x(t) := \frac{\bar{F}_X(xt)}{\bar{F}_X(x)} \underset{x\to\infty}{\to} T(t) = t^{-\alpha} \,. \tag{3.12}$$

Moreover, $\left(\bar{F}_Z(u_n)/\bar{F}_X(u_n)\right)_n$ is a strictly positive sequence converging to a positive limit. So, it is bounded away from zero and infinity, that is, there exists $K>0$ such that for all $n \geq 1$,

$$1/\lambda_0 < \bar{F}_Z(u_n)/\bar{F}_X(u_n) < \lambda_0. \tag{3.13}$$

Furthermore, A(ii) implies that $\bar{F}_Z$ is continuous. So is $\bar{F}_X$, by (3.11).

Regular variation of $\bar{F}_Z$ and $\bar{F}_X$ yields different versions of Potter's bounds (cf. [50], [14, p. 25]). We state them the way they are used in this thesis. First, for all $\epsilon > 0$, there exists $\omega(\epsilon) > 1$ such that $\forall\, x \geq 1$, $t > 0$,

$$J_x(t) \leq \omega(\epsilon) \max\left(1, t^{-(\alpha+\epsilon)}\right) \,. \tag{3.14}$$

Further $\forall\, C > 1$, $\epsilon > 0$, there exists $\delta = \delta(C;\epsilon) \geq 0$ such that for $x \geq \delta$, $t > 0$,

$$T_x(t) \leq C\left(t^{-(\alpha+\varepsilon)} \vee t^{-(\alpha-\varepsilon)}\right). \tag{3.15}$$

Notice that (3.13) in conjunction with (3.15) yields

$$\frac{\bar{F}_Z(xt)}{\bar{F}_X(x)} \leq \lambda C\left(t^{-(\alpha+\varepsilon)} \vee t^{-(\alpha-\varepsilon)}\right). \tag{3.16}$$

The next result precisely characterizes the tail distribution of X.

Lemma 3.3.3. *[39]*
Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). The tail distribution function of X is regularly varying with index $-\alpha$, that is

$$\bar{F}_X(x) = x^{-\alpha} E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)\right], \tag{3.17}$$

where $\tilde{\ell}(x) := E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)\right]$ is slowly varying at infinity and ℓ^* as in (3.3).

Proof. The law of total expectation and (3.8) ensure that

$$\bar{F}_X(x) = P(\phi(Y)Z > x) = E\left(\bar{F}_Z\left(x/\phi(Y)\right)\right).$$

Since $\bar{F}_Z \in RV_\infty(-\alpha)$, then $\bar{F}_Z(x) = x^{-\alpha}\ell^*(x)$, for some $\ell^* \in SV_\infty$, by (2.17). Therefore,

$$\bar{F}_X(x) = E\left[\bar{F}_Z\left(x/\phi(Y)\right)\right] = x^{-\alpha}E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)\right] =: x^{-\alpha}\tilde{\ell}(x).$$

Theorem 2.3.10 requires that it remains to show $\tilde{\ell} \in SV_\infty$ - that is for all $t > 0$,

$$\lim_{x\to\infty}\frac{\tilde{\ell}(xt)}{\tilde{\ell}(x)} = \lim_{x\to\infty}\frac{E\left[\phi^\alpha(Y)\ell^*\left(xt/\phi(Y)\right)/\ell^*(x)\right]}{E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)/\ell^*(x)\right]} = 1.$$

Since ℓ^* is bounded away from 0 and ∞ on every compact subset of $(0,\infty)$, then Theorem 2.3.7 ensures that for all $\epsilon > 0$, there exists $C = C(\epsilon) > 1$, such that:

$$\begin{aligned}
\phi^\alpha(Y)\frac{\ell^*\left(xt/\phi(Y)\right)}{\ell^*(x)} &\le C(t^\epsilon \vee t^{-\epsilon})\left(\phi^{\alpha-\epsilon}(Y) + \phi^{\alpha+\epsilon}(Y)\right), \\
\phi^\alpha(Y)\frac{\ell^*\left(x/\phi(Y)\right)}{\ell^*(x)} &\le C\left(\phi^{\alpha-\epsilon}(Y) + \phi^{\alpha+\epsilon}(Y)\right).
\end{aligned}$$

As x goes to ∞, $\phi^\alpha(Y)\ell^*\left(xt/\phi(Y)\right)/\ell^*(x)$ and $\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)/\ell^*(x)$ converge with probability one to $\phi^\alpha(Y)$. Hence, (3.7a) and the Dominated Convergence Theorem yield

$$\begin{aligned}
E\left[\phi^\alpha(Y)\ell^*\left(xt/\phi(Y)\right)/\ell^*(x)\right] &\underset{x\to\infty}{\to} E\left[\phi^\alpha(Y)\right] \\
E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)/\ell^*(x)\right] &\underset{x\to\infty}{\to} E\left[\phi^\alpha(Y)\right].
\end{aligned}$$

□

Differentiability of J_x. In the next two lemmas, we look into differentiability of J_x. This is an ingredient for differentiability of T_x.

Lemma 3.3.4. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $x \ge 1$, $t > 0$,*

$$J_x'(t) \to T'(t) = -\alpha t^{-\alpha-1}. \tag{3.19}$$

Further, there exist $\beta, M_{\alpha,\beta} > 0$ such that for all $x \ge 1$, $t > 0$,

$$|J_x'(t)| \le M_{\alpha,\beta}\left(t^{-(\alpha+\beta+1)} \vee t^{-(\alpha-\beta+1)}\right). \tag{3.20}$$

Proof. Assume that (3.10) holds. It follows from (3.3) that if $x \geq 1$, then

$$J_x(t) = \frac{\bar{F}_Z(xt)}{\bar{F}_Z(x)} = t^{-\alpha} \exp\left(\int_x^{xt} \frac{\eta^*(u)}{u}\, du\right) = t^{-\alpha} \exp\left(\int_1^t \frac{\eta^*(xv)}{v}\, dv\right),$$

where the last equality holds by setting $v = u/x$. Therefore, for $x \geq 1$ and $t \geq 1$,

$$\begin{aligned} J_x'(t) &= -\alpha t^{-\alpha-1} \exp\left(\int_1^t \frac{\eta^*(xv)}{v}\, dv\right) + t^{-\alpha-1}\eta^*(xt) \exp\left(\int_1^t \frac{\eta^*(xv)}{v}\, dv\right) \\ &= t^{-1}\left(-\alpha + \eta^*(xt)\right) J_x(t). \end{aligned}$$

Similarly, for $x \geq 1$ and $0 < t \leq 1$, $J_x'(t) = t^{-1}\left(-\alpha - \eta^*(xt)\right) J_x(t)$. Altogether $\forall\, t > 0$,

$$J_x'(t) \underset{x\to\infty}{\to} -\alpha t^{-\alpha-1},$$

since $\eta^* \in RV_\infty(-\kappa)$ and (3.10) holds. Hence (3.19) is proven.

On the other hand, since η^* is bounded, then there exists $\beta > 0$ such that $|\eta^*(xv)| \leq \beta$. Moreover by (2.13), $J_x(t) \leq C(\beta)\left(t^{-(\alpha-\beta)} \vee t^{-(\alpha+\beta)}\right)$. Therefore, for all $x, t > 0$,

$$\begin{aligned} |J_x'(t)| &\leq t^{-1}\left(\alpha + |\eta^*(xt)|\right) J_x(t) \\ &\leq C(\beta)\left(\alpha+\beta\right)\left(t^{-(\alpha+\beta+1)} \vee t^{-(\alpha-\beta+1)}\right), \end{aligned}$$

Setting $M_{\alpha,\beta} = C(\beta)(\alpha+\beta) > 0$, (3.20) is satisfied. □

Lemma 3.3.5. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $x \geq 1$, $t > 0$,*

$$\frac{d}{dt}E(J_x(t/\phi(Y))) = E\left(\frac{d}{dt}J_x(t/\phi(Y))\right) = E\left((1/\phi(Y))J_x'(t/\phi(Y))\right). \tag{3.21}$$

Proof. By setting $g(t) = E\left(J_x(t/\phi(Y))\right)$, it follows that for all $h > 0$, we have

$$\begin{aligned} &\frac{g(t+h) - g(t)}{h} - E\left((1/\phi(Y))J_x'(t/\phi(Y))\right) \\ &= E\left(\frac{J_x(t+h/\phi(Y)) - J_x(t/\phi(Y))}{h} - \frac{J_x'(t/\phi(Y)}{\phi(Y)}\right) \\ &= E\left(\left(\frac{J_x(t+h/\phi(Y)) - J_x(t/\phi(Y))}{h/\phi(Y)} - J_x'(t/\phi(Y))\right) 1/\phi(Y)\right) \underset{h\to 0}{\to} 0, \end{aligned}$$

by the Dominated Convergence Theorem. In fact since $t \mapsto J_x(t)$ is differentiable, then

$$\frac{J_x(t+h/\phi(Y)) - J_x(t/\phi(Y))}{h/\phi(Y)} - J_x'(t/\phi(Y)) \underset{h\to 0}{\to} 0.$$

In addition, by the mean value theorem, there exists $t < \tau < t + h$ such that

$$\begin{aligned}
&\left|\frac{J_x(t+h/\phi(Y)) - J_x(t/\phi(Y))}{h/\phi(Y)} - J_x'(t/\phi(Y))\right| 1/\phi(Y) \\
&\leq \left(\left|\frac{J_x(t+h/\phi(Y)) - J_x(t/\phi(Y))}{h/\phi(Y)}\right| + |J_x'(t/\phi(Y))|\right) 1/\phi(Y) \\
&= \left(|1/\phi(Y) J_x'(\tau/\phi(Y))| + J_x'(\tau/\phi(Y))\right) 1/\phi(Y) \\
&\leq M^*_{\alpha,\beta}\left(h_1(\tau)(\phi^{\alpha+\beta-1}(Y) + \phi^{\alpha-\beta-1}(Y))\right) + M_{\alpha,\beta}\left(h_2(t)(\phi^{\alpha+\beta}(Y) + \phi^{\alpha-\beta}(Y))\right),
\end{aligned}$$

where the functions h_1 and h_2 are respectively defined by $h_1(\tau) = \tau^{-(\alpha+\beta+1)} \vee \tau^{-(\alpha-\beta+1)}$ and $h_2(t) = t^{-(\alpha+\beta+1)} \vee t^{-(\alpha-\beta+1)}$. Note that (3.20) justifies the last inequality. □

Transfer of differentiability. The next lemma deals with the transfer of differentiability from Z to X. For instance, the convergence in (3.19) is transferred to T_x' and the bound in (3.20) is transferred to T_x.

Lemma 3.3.6. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $x \geq 1$, $t > 0$,*

$$T_x'(t) \to T'(t)\ , \ \text{as } x \to \infty, \tag{3.22}$$

$$|T_x'(t)| \leq K_0\left(t^{-(\alpha+\beta+1)} \vee t^{-(\alpha-\beta+1)}\right), \tag{3.23}$$

where K_0 is a constant that does not depend on either t or x.

Proof. By virtue of (3.10), it holds that for all $t > 0$, $x \geq 1$,

$$T_x(t) = \frac{\bar{F}_Z(x)}{\bar{F}_X(x)} E\left(\frac{\bar{F}_Z(xt/\phi(Y))}{\bar{F}_Z(x)}\right) = \frac{\bar{F}_Z(x) E\left(J_x\left(t/\phi(Y)\right)\right)}{\bar{F}_X(x)}.$$

Taking the derivative with respect to t and applying (3.21), we obtain

$$T_x'(t) = \frac{\bar{F}_Z(x)}{\bar{F}_X(x)} E\left(\frac{d}{dt} J_x\left(t/\phi(Y)\right)\right) = \frac{\bar{F}_Z(x) E\left((1/\phi(Y)) J_x'\left(t/\phi(Y)\right)\right)}{\bar{F}_X(x)}. \tag{3.24}$$

The interchange of the integral and the derivative is allowed since (3.21) holds. Furthermore, by (3.20) and (3.13), it holds that

$$\begin{aligned}
|T_x'(t)| &\leq \lambda_0 E\left((1/\phi(Y))\, |J_x'\left(t/\phi(Y)\right)|\right) \\
&\leq \lambda_0 M E\left((1/\phi(Y))\left((t/\phi(Y))^{-(\alpha+\beta+1)} \vee (t/\phi(Y))^{-(\alpha-\beta+1)}\right)\right) \\
&\leq \lambda_0 M\left(t^{-(\alpha+\beta+1)} \vee t^{-(\alpha-\beta+1)}\right)\left(E\left([\phi(Y)]^{\alpha+\beta}\right) + E\left([\phi(Y)]^{\alpha-\beta}\right)\right).
\end{aligned}$$

Thus, setting $\delta_0 = \lambda_0 M \left(E\left([\phi(Y)]^{\alpha+\beta}\right) + E\left([\phi(Y)]^{\alpha-\beta}\right)\right)$, leads to the desired bound in (3.22). The constant K_0 is finite by (3.7a). To end the proof, notice that the dominated convergence theorem and (3.19) yield

$$E\left(1/\phi(Y)\right) J_x'\left(t/\phi(Y)\right) \underset{x\to\infty}{\to} T'(t)E\left(\phi^\alpha(Y)\right).$$

Hence, convergence of T_x' to T', as $x \to \infty$ follows from (3.24) and (3.11). □

Transfer of second-order regular variation. Does second-order regular variation property on Z transfer to X? This question was examined in [39, p. 117]. It turns out the answer is affirmative. Note that the present proof is adapted to Definition 2.4.16. We provide a slight improvement of the rate function of $\bar{F}_X$. Note that the rate function of $\bar{F}_X$, say $\tilde{\eta}$, is asymptotically proportional to η^*.

Lemma 3.3.7. *[39, p. 117]*
Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). The tail distribution function of X is second-order regularly varying with indices $-\alpha$ and κ and rate function $\tilde{\eta}$, that is in short $\bar{F}_X \in 2RV_\infty(-\alpha, -\kappa, \tilde{\eta})$ with

$$\tilde{\eta}(x) = \frac{E\left(\phi^\alpha(Y)\eta^*(x/\phi(Y))\ell^*(x/\phi(Y))\right)}{E\left(\phi^\alpha(Y)\ell^*(x/\phi(Y))\right)} \underset{x\to\infty}{\sim} \frac{E\left(\phi^{\alpha+\kappa}(Y)\right)}{E\left(\phi^\alpha(Y)\right)}\eta^*(x). \tag{3.25}$$

The rate function η^* as well as the slowly varying function ℓ^* are defined in (3.3).

Proof. Considering (3.17) in conjunction with (2.28) yield that

$$\begin{aligned}\bar{F}_X(x) &= x^{-\alpha}E\left(\phi^\alpha(Y)\left(\ell^*(1) + \int_1^{x/\phi(Y)} \frac{\eta^*(u)\ell^*(u)}{u}\,du\right)\right)\\ &= x^{-\alpha}E\left(\phi^\alpha(Y)\left(\ell^*(1) + \int_{\phi(Y)}^{x} \frac{\eta^*(v/\phi(Y))\ell(v/\phi(Y))}{v}\,dv\right)\right),\end{aligned}$$

where the last equality holds thanks to the change of variables $v = u\phi(Y)$. Therefore,

$$\begin{aligned}\bar{F}_X(x) &= x^{-\alpha}E\left(\phi^\alpha(Y)\left(\ell^*(1) - \int_1^{\phi(Y)} \frac{\eta^*(v/\phi(Y))\ell^*(v/\phi(Y))}{v}\,dv\right)\right)\\ &\quad + x^{-\alpha}E\left(\phi^\alpha(Y)\int_1^x \frac{\eta^*(v/\phi(Y))\ell^*(v/\phi(Y))}{v}\,dv\right).\end{aligned}$$

Now, the change of variables $t = v/\phi(Y)$ yields that

$$\int_1^{\phi(Y)} \frac{\eta^*(v/\phi(Y))\ell^*(v/\phi(Y))}{v}\,dv = -\int_1^{1/\phi(Y)} \frac{\eta^*(t)\ell^*(t)}{t}\,dt.$$

As a consequence, the tail distribution of X becomes

$$\begin{aligned}
\bar{F}_X(x) &= x^{-\alpha}\left(E\left(\phi^\alpha(Y)\ell^*(1/\phi(Y))\right) + \int_1^x \frac{E\left(\phi^\alpha(Y)\eta^*(v/\phi(Y))\ell^*(v/\phi(Y))\right)}{v}\,dv\right)\\
&= x^{-\alpha}\left(\tilde{\ell}(1) + \int_1^x \frac{E\left(\phi^\alpha(Y)\eta^*(v/\phi(Y))\ell^*(v/\phi(Y))\right)}{E\left(\phi^\alpha(Y)\ell^*(v/\phi(Y))\right)} E\left(\phi^\alpha(Y)\ell^*(v/\phi(Y))\right)\frac{dv}{v}\right)\\
&= x^{-\alpha}\left(\tilde{\ell}(1) + \int_1^x \frac{\tilde{\eta}(v)\tilde{\ell}(v)}{v}\,dv\right),
\end{aligned}$$

where $\tilde{\eta}$ is defined in (3.25) and $\tilde{\ell}(x) = E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)\right]$. Notice that $\tilde{\eta}$ is well defined. In fact, for any $\epsilon > 0$ such that $P(Y \geq \epsilon) = \delta > 0$,

$$E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)/\ell^*(x)\right] = \frac{P\left(X > x\right)}{P(Z > x)} \geq \frac{P(Y \geq \epsilon)P(Z > x/\epsilon)}{P(Z > x)} \geq \delta \inf_{x>0} \frac{P(Z > x/\epsilon)}{P(Z > x)}.$$

It remains to show that $\tilde{\eta}$ is a bounded, nonnegative or nonpositive regularly varying function at infinity with index $-\kappa$. Clearly $\tilde{\eta}$ is of the same sign as η^* since ϕ, ℓ^* are nonnegative. Further, for all $x > 0$, $|\tilde{\eta}(x)| \leq |\eta^*(x/\phi(Y))|$. Since η^* is bounded, then is $\tilde{\eta}$. Also, note that $\tilde{\eta} \in RV_\infty(-\kappa)$. In fact, for all $t > 0$,

$$\frac{\tilde{\eta}(xt)}{\tilde{\eta}(x)} = \left\{\frac{E\left(\phi^\alpha(Y)\eta^*(xt/\phi(Y))\ell^*(xt/\phi(Y))\right)}{E\left(\phi^\alpha(Y)\eta^*(x/\phi(Y))\ell^*(x/\phi(Y))\right)}\right\}\left\{\frac{E\left(\phi^\alpha(Y)\ell^*(x/\phi(Y))\right)}{E\left(\phi^\alpha(Y)\ell^*(xt/\phi(Y))\right)}\right\}.$$

Since $x \mapsto E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)\right] \in SV_\infty$, by (3.17), then the second factor in the right hand side tends to 1, as x goes to infinity. On the other hand, the first factor of the right hand side equals to

$$\frac{E\left(\phi^\alpha(Y)[\eta^*(xt/\phi(Y))/\eta^*(x)][\ell^*(xt/\phi(Y))/\ell^*(x)]\right)}{E\left(\phi^\alpha(Y)[\eta^*(x/\phi(Y))/\eta^*(x)][\ell^*(x/\phi(Y))/\ell^*(x)]\right)}.$$

By Theorem 2.3.13 we have

$$\begin{aligned}
&\phi^\alpha(Y)\left(\frac{\eta^*(xt/\phi(Y))}{\eta^*(x)}\right)\left(\frac{\ell^*(xt/\phi(Y))}{\ell^*(x)}\right)\\
&\leq B(\epsilon)f(t,\epsilon,\kappa)(\phi^{\alpha+\kappa-2\epsilon}(Y) + 2\phi^{\alpha+\kappa}(Y) + \phi^{\alpha+\kappa+2\epsilon}(Y)),
\end{aligned}$$

where $f(t,\epsilon,\kappa) = (t^{-\kappa+\epsilon} \vee t^{-\kappa-\epsilon})(t^{-\epsilon} \vee t^{\epsilon})$. We also have

$$\phi^\alpha(Y)\left(\frac{\eta^*(xt/\phi(Y))}{\eta^*(x)}\right)\left(\frac{\ell^*(xt/\phi(Y))}{\ell^*(x)}\right) \underset{x\to\infty}{\to} t^{-\kappa}\phi^{\alpha+\kappa}(Y), \text{ w.p.1.}$$

Therefore, the Dominated Convergence Theorem yields that

$$\frac{E\left(\phi^\alpha(Y)\eta^*(xt/\phi(Y))\ell^*(xt/\phi(Y))\right)}{E\left(\phi^\alpha(Y)\eta^*(x/\phi(Y))\ell^*(x/\phi(Y))\right)} \underset{x\to\infty}{\to} t^{-\kappa}.$$

We wrap up this proof by showing asymptotic equivalence of $\tilde{\eta}$ and η^*:

$$\frac{\tilde{\eta}(x)}{\eta^*(x)} = \frac{E\left[\phi^\alpha(Y)\left\{\eta^*\left(x/\phi(Y)\right)/\eta^*(x)\right\}\left\{\ell^*\left(x/\phi(Y)\right)/\ell^*(x)\right\}\right]}{E\left[\phi^\alpha(Y)\ell^*\left(x/\phi(Y)\right)/\ell^*(x)\right]} \underset{x\to\infty}{\to} \frac{E(\phi^{\alpha+\kappa}(Y))}{E(\phi^\alpha)(Y)}.$$

□

Remark 3.3.8. It follows from (2.32) and Lemma 2.4.23 that

$$\frac{\bar{F}_X(tx)/\bar{F}_X(x) - t^{-\alpha}}{\widetilde{\eta}(x)} \underset{x\to\infty}{\to} t^{-\alpha}\frac{1-t^{-\kappa}}{\kappa} \tag{3.26a}$$

$$\frac{U_X(xt)/U_X(x) - t^{1/\alpha}}{\widetilde{\eta}^\dagger(x)} \underset{x\to\infty}{\to} t^{1/\alpha}\frac{1-t^{-\kappa/\alpha}}{\kappa/\alpha}\ , \tag{3.26b}$$

where $\widetilde{\eta}$ is defined in (3.25) and

$$\widetilde{\eta}^\dagger(x) = \alpha^{-2}\widetilde{\eta}\left(\bar{F}_X^{\leftarrow}(x)\right)\ .$$

3.4 No-bias Conditions

Replacing T_x with its limit T induces bias in statistical inference. This bias is controlled with the help of second-order regular variation. From [39] we have that for all $\epsilon > 0$, there exists $C(\epsilon) > 0$, $C_1(\epsilon) > 0$ such that for all $x \geq 1$, $t > 0$,

$$|J_x(t) - T(t)| \leq C(\epsilon)\left(t^{-(\alpha+\kappa+\epsilon)} \vee t^{-(\alpha+\kappa-\epsilon)}\right)|\eta^*(x)|\ , \tag{3.27a}$$

$$|T_x(t) - T(t)| \leq C_1(\epsilon)\left(t^{-(\alpha+\kappa+\epsilon)} \vee t^{-(\alpha+\kappa-\epsilon)}\right)|\tilde{\eta}(x)|\ . \tag{3.27b}$$

The last bound follows from (3.25). Recall that m is the Hermite rank of ϕ^α, d the long memory parameter and

$$a_{n,m} = \left(\sqrt{n\bar{F}_Z(u_n)} + \frac{n}{b_{n,m}}\right)\mathbb{1}_{\{m(1-2d)<1\}} + \sqrt{n}\mathbb{1}_{\{m(1-2d)>1\}}\ .$$

As a result, (3.6) implies that for all $\tau_0 > 0$,

$$a_{n,m}\sup_{t>\tau_0}|J_{u_n}(t) - T(t)| \underset{n\to\infty}{\to} 0 \tag{3.28a}$$

$$a_{n,m}\sup_{t>\tau_0}|T_{u_n}(t) - T(t)| \underset{n\to\infty}{\to} 0. \tag{3.28b}$$

Subsequently, we shall refer to (3.6) as the **no-bias condition**.

Since $\bar{F}_X$ is regularly varying at infinity with index $-\alpha$, then $U_X(t) = Q_X(t) = F_X^{\leftarrow}(1-1/t)$ is regularly varying with index $1/\alpha$. How then is Q_X affected by the no-bias condition? In the next two Lemmas 3.4.1 and 3.4.2, we study some implications of (3.6) on the quantile Q_X. These implications are essential for derivation of limit theorems of measures of financial risk. This will be done in Chapter 5.

Lemma 3.4.1. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Let $\gamma = 1/\alpha$. Assume that $p = p_n \to 0$, $n/k \to \infty$ and $k/(np) \to \infty$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \left\{ \left(\frac{k}{np}\right)^{-\gamma} \frac{Q_X(1/p)}{Q_X(n/k)} - 1 \right\} = 0\ . \tag{3.29}$$

Proof. Since F_X is continuous and strictly increasing, then $U_X = Q_X$. Using Lemma 2.4.21 with $x = n/k$, $t = k/(np)$ yields that

$$\begin{aligned}
&\left| \frac{Q_X(1/p)/Q_X(n/k) - (k/(np))^{\gamma}}{D_0(n/k)} - \left(\frac{k}{np}\right)^{\gamma} \frac{1-(k/np)^{-\rho^{\dagger}}}{\rho^{\dagger}} \right| \\
&\le \epsilon \left(\frac{k}{np}\right)^{\gamma-\rho^{\dagger}} \left\{ \left(\frac{k}{np}\right)^{\delta} \vee \left(\frac{k}{np}\right)^{-\delta} \right\} ,
\end{aligned}$$

where

$$D_0(x) = \rho^{\dagger} \left\{ \frac{x^{\gamma}}{Q_X(x)} \lim_{s\to\infty} \frac{Q_X(s)}{s^{\gamma}} - 1 \right\} .$$

Choose $\delta < \rho^{\dagger}$. Divide both sides by $(k/(np))^{\gamma}$ to get

$$\begin{aligned}
&\left| \frac{(k/(np))^{-\gamma} Q_X(1/p)/Q_X(n/k) - 1}{D_0(n/k)} - \frac{(k/np)^{-\rho^{\dagger}} - 1}{-\rho^{\dagger}} \right| \\
&\le \epsilon \left(\frac{k}{np}\right)^{-\rho^{\dagger}} \left\{ \left(\frac{k}{np}\right)^{\delta} \vee \left(\frac{k}{np}\right)^{-\delta} \right\} .
\end{aligned}$$

Since $k/(np) \to \infty$, then the right hand side converges to 0. Therefore, we have

$$\lim_{n\to\infty} \left(\frac{k}{np}\right)^{-\gamma} \frac{Q_X(1/p)/Q_X(n/k) - 1}{D_0(n/k)} = \frac{1}{\rho^{\dagger}}. \tag{3.30}$$

Since $D_0(n/k) \to 0$, as $n \to \infty$, then we conclude that

$$\lim_{n\to\infty} \left\{ \left(\frac{k}{np}\right)^{-\gamma} \frac{Q_X(1/p)}{Q_X(n/k)} - 1 \right\} = 0\ .$$

$\square$

Lemma 3.4.2. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Assume that $p = p_n \to 0$, $n/k \to \infty$ and $k/(np) \to \infty$ as $n \to \infty$. If (3.6) holds, then*

$$\limsup_{n\to\infty} a_{n,m} \left\{ \left(\frac{k}{np}\right)^{-\gamma} \frac{Q_X(1/p)}{Q_X(n/k)} - 1 \right\} = 0 \,. \tag{3.31}$$

Proof. Since F_X is continuous and strictly increasing, $U_X = Q_X$. Applying Lemma 2.4.21 with $x = n/k$, $t = k/(np)$ ensures that

$$\begin{aligned}
&\left| a_{n,m} \frac{Q_X(1/p)/Q_X(n/k) - (k/(np))^{\gamma}}{a_{n,m} D_0(n/k)} - \left(\frac{k}{np}\right)^{\gamma} \frac{1-(k/np)^{-\rho^\dagger}}{\rho^\dagger} \right| \\
&\leq \epsilon \left(\frac{k}{np}\right)^{\gamma-\rho^\dagger} \left\{ \left(\frac{k}{np}\right)^{\delta} \vee \left(\frac{k}{np}\right)^{-\delta} \right\} ,
\end{aligned}$$

where

$$D_0(x) = \rho^\dagger \left\{ \frac{x^\gamma}{Q_X(x)} \lim_{s\to\infty} \frac{Q_X(s)}{s^\gamma} - 1 \right\} .$$

Choose $\delta < \rho^\dagger$. Divide both sides by $(k/(np))^\gamma$ to get

$$\begin{aligned}
&\left| \frac{a_{n,m} \left((k/(np))^{-\gamma} Q_X(1/p)/Q_X(n/k) - 1\right)}{a_{n,m} D_0(n/k)} - \frac{1-(k/np)^{-\rho^\dagger}}{\rho^\dagger} \right| \\
&\leq \epsilon \left(\frac{k}{np}\right)^{-\rho^\dagger} \left\{ \left(\frac{k}{np}\right)^{\delta} \vee \left(\frac{k}{np}\right)^{-\delta} \right\} .
\end{aligned}$$

Since $k/(np) \to \infty$, then the right hand side converges to 0. Therefore, we have

$$\lim_{n\to\infty} a_{n,m} \frac{(k/(np))^{-\gamma} Q_X(1/p)/Q_X(n/k) - 1}{a_{n,m} D_0(n/k)} = \frac{1}{\rho^\dagger} . \tag{3.32}$$

Note that (3.6) implies $a_{n,m} D_0(n/k) \to 0$ as $n \to \infty$. Hence (3.31) holds. □

3.5 Examples

In the examples below, we illustrate the various assumptions.

Example 3.5.1. Let the tail distribution function, $\bar{F}_Z$, be of the form:

$$\bar{F}_Z(x) = \begin{cases} \frac{1}{2}\left(x^{-\alpha} + x^{-\alpha\gamma}\right) & x \geq 1 \\ 1 & 0 < x < 1. \end{cases}$$

1. Notice that $\bar{F}_Z$ fulfills (3.3). In fact, for all $x \geq 1$,
$$\bar{F}_Z(x) = x^{-\alpha} \exp\left(\int_1^x \frac{\alpha(\gamma-1)t^{-\alpha(\gamma-1)-1}dt}{1+t^{-\alpha(\gamma-1)}}\right).$$
Therefore, $\bar{F}_Z \in 2RV_\infty(-\alpha, -\alpha(\gamma-1), \eta^*)$, where the rate function is defined by
$$\eta^*(x) = x^{-\alpha(\gamma-1)} \frac{\alpha(\gamma-1)}{1+x^{-\alpha(\gamma-1)}} = \frac{\alpha(\gamma-1)}{1+x^{\alpha(\gamma-1)}} \underset{x\to\infty}{\sim} x^{\alpha(1-\gamma)}.$$
Note that η^* is nonnegative, bounded by $\alpha(\gamma-1)$. Also, $\eta^* \in RV_\infty(-\alpha(\gamma-1))$.

2. We have for all $x \geq 1$ and $xt \geq 1$,
$$J_x(t) = \frac{(xt)^{-\alpha} + (xt)^{-\alpha\gamma}}{x^{-\alpha} + x^{-\alpha\gamma}} = \frac{t^{-\alpha}}{1+x^{-\alpha(\gamma-1)}} + \frac{t^{-\alpha\gamma}}{1+x^{-\alpha(1-\gamma)}} \underset{x\to\infty}{\to} T(t) =: t^{-\alpha}.$$
Therefore, (3.10) is proven.

3. Furthermore, we have for all $x \geq 1$ and $xt \geq 1$,
$$J'_x(t) = \frac{x^{-\alpha}(-\alpha t^{-\alpha-1})}{x^{-\alpha}+x^{-\alpha\gamma}} + \frac{x^{-\alpha\gamma}(-\alpha\gamma t^{-\alpha\gamma-1})}{x^{-\alpha}+x^{-\alpha\gamma}}.$$
If $x \geq 1$ and $0 < xt < 1$, then $J'_x(t) = 0$. Thus, for $x > 1$ and $t > 0$,
$$\begin{aligned} |J'_x(t)| &\leq \left(\frac{x^{-\alpha}}{x^{-\alpha}+x^{-\alpha\gamma}}\right)\alpha t^{-\alpha-1} + \left(\frac{x^{-\alpha\gamma}}{x^{-\alpha}+x^{-\alpha\gamma}}\right)\alpha\gamma t^{-\alpha\gamma-1} \\ &\leq \alpha t^{-\alpha-1}\left(1+\gamma t^{\alpha(1-\gamma)}\right) \\ &= \alpha t^{-\alpha-1} + \alpha\gamma t^{-\alpha+\alpha(1-\gamma)-1}. \end{aligned}$$
Hence (3.20) holds with $M = \alpha(1+\gamma)$ and $\beta = \alpha(\gamma-1)$.

4. Finally to get a sense of when (3.6) is satisfied, notice that $\bar{F}_Z(x) \underset{x\to\infty}{\sim} x^{-\alpha}$. As a result, $\bar{F}_Z^{\leftarrow}(x) \underset{x\to\infty}{\sim} x^{-1/\alpha}$. Using the fact Z and X are tail equivalent (3.11), we conclude that
$$\bar{F}_Z(u_n) \sim k/n \text{ and } u_n \sim \bar{F}_Z^{\leftarrow}(k/n) = (n/k)^{1/\alpha}.$$
Recall (3.5a) and note that $n/b_{n,m} \approx n^{m(1/2-d)}$. So, it follows that
   - If $m(1-2d) < 1$, then $\sqrt{n\bar{F}_Z(u_n)}|\eta^*(u_n)| \to 0$ as $n \to \infty$ if and only if
$$k^{\gamma-1/2}/n^{\gamma-1} \to 0, \text{ as } n \to \infty.$$
On the other hand, $\frac{n}{b_{n,m}}|\eta^*(u_n)| \to 0$ as $n \to \infty$ if and only if
$$n^{m(1/2-d)}(k/n)^{\gamma-1} \to 0, \text{ as } n \to \infty.$$

- If $m(1-2d) > 1$, then $\sqrt{n}|\eta^*(u_n)| \to 0$ if and only if

$$n^{3/2-\gamma}k^{\gamma-1} \to 0, \text{ as } n \to \infty.$$

Example 3.5.2 (Student-t distribution)**.** Assume that a random variable X has a Student-t density with ν degrees of freedom, that is the density is given by

$$f(x) = c_\nu \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2} , \quad x \in \mathbb{R} ,$$

for some explicit constant c_ν. Then (cf. [23, Exercise 2.15])

$$\bar{F}(x) = c_\nu \nu^{\nu/2} x^{-\nu} \left\{1 + d_\nu x^{-2} + o(x^{-4})\right\} , \quad \text{as} \quad x \to \infty ,$$

where d_ν is another explicit constant. This implies that the tail index is $\alpha = \nu$, while the second order index is 2.

Thus, the no-bias condition (3.6) is satisfied in the following situations:

- If $m(1-2d) < 1$, then $\sqrt{n\bar{F}_Z(u_n)}|\eta^*(u_n)| \to 0$, as $n \to \infty$ if and only if

$$k^{1/2+2/\nu}n^{-2/\nu} \to 0, \text{ as } n \to \infty.$$

  On the other hand, $\frac{n}{b_{n,m}}|\eta^*(u_n)| \to 0$, as $n \to \infty$ if and only if

$$n^{m(1/2-d)}(k/n)^{2/\nu} \to 0, \text{ as } n \to \infty.$$

- If $m(1-2d) > 1$, then $\sqrt{n}|\eta^*(u_n)| \to 0$ if and only if

$$n^{1/2}\left(k/n\right)^{2/\nu} \to 0, \text{ as } n \to \infty.$$

3.6 Concluding Remarks

In this chapter, we have introduced the long memory stochastic volatility model with leverage (3.1). We have described the model and stated its relevant assumptions. The major contributions of this chapter are the so-called transfer theorems. Some existing results such as Lemma 3.3.7 have been adapted to the second-order regular variation framework of this thesis. We illustrate these assumptions via two simple examples (Section 3.5).

To estimate extremal values of the interest in this thesis, it is important to study the limiting behaviour of the so-called *tail empirical process*, a variation of the classical empirical process that takes into account only large values. This is an important tool used in nonparametric estimation of extremal quantities. The mathematical theory of the tail empirical process is much more involved than that of the usual empirical process and has only been studied since the beginning of the 21st century.

The transfer theorems as well as the no-bias conditions discussed in the current chapter will play a major role in the analysis of the limiting behaviour of the tail empirical process associated with long memory stochastic volatility sequences with leverage. This is the subject of the next chapter.

Chapter 4

Limit Theorems for the Tail Empirical Processes

4.1 Introductory Comments

The tail empirical process (in short TEP) is an important tool used in nonparametric estimation of extremal quantities, like the Hill estimator of the index of regular variation or various risk measures.

Our goal is to study weak convergence for the tail empirical processes associated with the LMSV with leverage. These results are not only of theoretical interest, but are applicable to different statistical procedures based on intermediate extremes. A similar problem was studied in the case of independent, identically distributed random variables in [34], or for weakly dependent sequences in [30], [29], [48], [42]. In [39] the authors considered heavy tailed, long memory stochastic volatility models and obtained asymptotic results for tail empirical processes. This was extended later on to the multiparameter situation in [40]. However, in the latter two articles leverage was excluded, greatly simplifying theoretical considerations. As evidenced in [41], the presence of long memory, heavy tails and leverage may affect the limiting behaviour of relevant statistics.

It turns out that in the present setting, leverage does not affect the limiting behaviour of the tail empirical process, and hence the results are comparable to those in [39] where leverage is not present. The limiting behaviour depends only on the interplay between the tail index α and the strength of long memory. However, it should be pointed out clearly that the extension from models without leverage to those with leverage is highly nontrivial from a theoretical point of view. In [39] the authors were able to exploit the conditional independence of the sequence $\{X_j\}$ given $\{Y_j\}$. Here this approach is not applicable and instead we introduce a martingale-long memory decomposition of the tail empirical process. This makes the proof of tightness technically very involved.

Furthermore, as in [39], for applications we must replace the unobservable sequence u_n with appropriate order statistics. It turns out that the limiting behaviour of the resulting *TEP with random levels* is not affected by either long memory or leverage. This, through integral functionals, allows us to obtain limiting results for different estimators of the tail index, including the classical Hill estimator (see [23] for results in the i.i.d. case) or the more general class of harmonic mean estimators (see [7] again for results in the i.i.d. case).

In summary, our contribution in this chapter is twofold. From a theoretical point of view, our most important contribution is the proof of weak convergence of the tail empirical process (with fixed and random levels) in the presence of heavy tails, long memory and leverage. Due to the complicated dependence structure of the process, the proof is not at all straightforward. From a practical point of view, the key result is that the asymptotic behaviour of the tail empirical process with random levels is unaffected by the presence of long memory and/or leverage in the model, and so in applications log returns $\{X_j\}$ may be handled exactly as if they were i.i.d. heavy-tailed random variables. This greatly enhances the utility of the LMSV model with leverage considered here.

The rest of the chapter is organized as follows. Throughout Section 4.2 up to the first sub-section of Section 4.3, we establish central and non-central limit theorems for the TEP via a "martingale-long memory decompostion". In Section 4.2.4 we state our main result on convergence of the tail empirical process with fixed levels (Theorem 4.2.18). This theorem is complemented in Section 4.3.1 by the corresponding result for random levels (Theorem 4.3.4). In Section 4.3.2 we prove weak convergence of integral functionals of the tail empirical process, which provides a unified approach to central limit theorems for estimators of the tail index (Theorem 4.3.16) in Section 4.3.3. We end with a brief conclusion in Section 4.4.

The content of this chapter is published in the *Electronic Journal of Statistics* [8].

4.2 Deterministic Levels

Consider the long memory stochastic volatility model with leverage defined in (3.1). Recall that $k_n \to \infty$ is an increasing sequence of positive integers such that $k_n/n \to 0$ and u_n is defined by $u_n = \bar{F}_X^{-1}(k_n/n)$. As a consequence

$$u_n \to \infty \text{ and } n\bar{F}_Z(u_n) \to \infty. \tag{4.1}$$

This section is devoted to the study of weak convergence of the tail empirical process with deterministic levels u_n. We respectively establish central and non-central limit theorems for the TEP via a "martingale-long memory decomposition" in Section 4.2.2 and Section 4.2.3.

Definition 4.2.1. *The empirical tail distribution function of $\{X_j\}$ is defined as:*

$$\widetilde{T}_n(t) := \frac{1}{k}\sum_{j=1}^{n} \mathbb{1}_{\{X_j > u_n t\}} = \frac{1}{k}\sum_{j=1}^{n} \mathbb{1}_{\{\phi(Y_j)Z_j > u_n t\}}\ , \quad t > 0. \tag{4.2}$$

As opposed to the ordinary empirical distribution function which deals with the entire distribution function, the tail empirical distribution function deals with extremes. We define its inverse to be

$$\widetilde{T}_n^{\leftarrow}(y) := \inf\{x \in \mathbb{R} : \widetilde{T}_n(x) \le y\}. \tag{4.3}$$

In the sequel, for notational convenience, let

$$V_{j,n}(t) := \mathbb{1}_{\{X_j > u_n t\}} = \mathbb{1}_{\{\phi(Y_j)Z_j > u_n t\}}\ , \quad T_n(t) = E\left(\widetilde{T}_n(t)\right) = \frac{\bar{F}_X(u_n t)}{\bar{F}_X(u_n)}. \tag{4.4}$$

Definition 4.2.2. *The tail empirical process of $\{X_j\}$ is defined to be*

$$\widetilde{S}_n(t) := k\left(\widetilde{T}_n(t) - T_{u_n}(t)\right) = \sum_{j=1}^{n}\left(V_{j,n}(t) - E\left(V_{j,n}(t)\right)\right),\ t > 0. \tag{4.5}$$

In this chapter, **our goal is to determine the asymptotic behaviour** of $\widetilde{S}_n$ under suitable normalizations. To fulfill it, we take the following approach.

4.2.1 Methodology

The structure of the model considered in (3.1) suggests the following "**martingale-long memory Doob decomposition**":

$$\widetilde{S}_n(t) := M_n(t) + L_n(t),\ t > 0, \tag{4.6}$$

where the summands M_n and L_n are respectively defined as follows:

$$M_n(t) := \sum_{j=1}^{n} \left[V_{j,n}(t) - E\left(V_{j,n}(t) \| \mathcal{G}_{j-1}\right)\right], \tag{4.7a}$$

$$L_n(t) := \sum_{j=1}^{n} \left[E\left(V_{j,n}(t) \| \mathcal{G}_{j-1}\right) - E\left(V_{j,n}(t)\right)\right]. \tag{4.7b}$$

We will call M_n the **martingale part** and L_n the **long memory part**. To establish weak convergence of $\widetilde{S}_n$ under suitable normalizations, we will establish weak convergence for M_n and L_n, suitably normalized. This will then determine the appropriate normalization for $\widetilde{S}_n$. The process M_n will be handled with a classical martingale Central Limit Theorem. Subsequently, the process L_n will be handled with a limit theorem for Hermite polynomials (cf. Theorem 2.7.23).

4.2.2 Weak Convergence of the Martingale Part

We consider in this section the process M_n defined in (4.7a) by

$$M_n(t) = \sum_{j=1}^{n} \triangle_j M_n(t), \; t > 0, \tag{4.8}$$

where the summands $\triangle_j M_n$'s are defined as follows:

$$\triangle_j M_n(t) := V_{j,n}(t) - E(V_{j,n}(t) \| \mathcal{G}_{j-1}). \tag{4.9}$$

We claim that $\left(M_n(t)\right)_n$ is a $\mathcal{G}$-martingale. In fact:

- $\left(\triangle_j M_n(t)\right)_j$ is adapted to the filtration $\mathcal{G}$ and $\triangle_j M_n(t)$ is integrable.
- $\left(\triangle_j M_n(t)\right)_j$ is a martingale difference since:

$$E(\triangle_j M_n(t) \| \mathcal{G}_{j-1}) = E(V_{j,n}(t) \| \mathcal{G}_{j-1}) - E(V_{j,n}(t) \| \mathcal{G}_{j-1}) = 0, \text{ w.p.1.}$$

Theorem 4.2.3. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then,*

$$\frac{M_n(\cdot)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} (B \circ T)(\cdot), \tag{4.10}$$

in $D(0,\infty)$ equipped with the Skorokhod J_1 topology. Note that $B(\cdot)$ is a standard Brownian motion on $(0,\infty)$, that is, a continuous path process with stationary independent increments, $B(0) = 0$ and $B(1) \sim \mathcal{N}$, with $\mathcal{N}$ being a standard normal random variable.

Proof. In Proposition 4.2.7, we prove convergence of the finite-dimensional distributions of properly scaled $M_n(\cdot)$, while in Proposition 4.2.12, we show its tightness. Propositions 4.2.7 and 4.2.12 complete the proof of Theorem 4.2.3. $\square$

Convergence of variance.

Lemma 4.2.4. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then,*

$$\operatorname{Var}\left(\frac{M_n(t)}{\sqrt{k}}\right) \underset{n\to\infty}{\to} t^{-\alpha}. \tag{4.11}$$

Proof. We start the proof by observing that for all $t > 0$, $E\left(\triangle_j M_n(t)\right) = 0$. We have

$$\begin{aligned}
\operatorname{Var}(\triangle_j M_n(t)) &= E\left((\triangle_j M_n(t))^2\right) - \left(E(\triangle_j M_n(t))\right)^2 \\
&= E\left((V_{j,n}(t) - E(V_{j,n}(t)\|\mathcal{G}_{j-1}))^2\right) \\
&= E(V_{j,n}(t)) - E\left((E(V_{j,n}(t)\|\mathcal{G}_{j-1}))^2\right).
\end{aligned}$$

Since Y_j and Z_j are independent, by (3.11),

$$\lim_{n\to\infty} \frac{E(V_{j,n}(t))}{\bar{F}_Z(u_n)} = \lim_{n\to\infty} \frac{\bar{F}_X(u_n t)}{\bar{F}_Z(u_n)} = \lim_{n\to\infty} \frac{\bar{F}_X(u_n t)}{\bar{F}_X(u_n)} \frac{\bar{F}_X(u_n)}{\bar{F}_Z(u_n)} = t^{-\alpha} E(\phi^\alpha(Y_j)),$$

which is equivalent to writing $E(V_{j,n}(t)) \sim t^{-\alpha} E(\phi^\alpha(Y_j))\bar{F}_Z(u_n)$ as $n \to \infty$. On the other hand, recalling that Y_j is $\mathcal{G}_{j-1}$-measurable, we have

$$\lim_{n\to\infty} \frac{E\left((E(V_{j,n}(t)\|\mathcal{G}_{j-1}))^2\right)}{(\bar{F}_Z(u_n))^2} = \lim_{n\to\infty} E\left(\left(\frac{E(V_{j,n}(t)\|\mathcal{G}_{j-1})}{\bar{F}_Z(u_n)}\right)^2\right) = \lim_{n\to\infty} E\left((J_{u_n}\left(t/\phi(Y)\right))^2\right).$$

Due to the regular variation of $(Z_j)_j$, we have for all $t > 0, \epsilon > 0$,

$$\begin{aligned}
&J_{u_n}\left(t/\phi(Y)\right) \xrightarrow[n\to\infty]{a.s} t^{-\alpha}\phi^\alpha(Y) \\
&J_{u_n}\left(t/\phi(Y)\right) \le C(\epsilon) t^{-(\alpha+\epsilon)} \max\left(1, \phi^{\alpha+\epsilon}(Y)\right),
\end{aligned}$$

where $C(\epsilon)$ is a constant depending on ϵ, but not on n. The latter bound follows from Potter's bound (cf. (3.14)). Therefore, the dominated convergence theorem ensures that

$$\lim_{n\to\infty} \frac{E\left((E\left(V_{j,n}(t)\|\mathcal{G}_{j-1}\right))^2\right)}{(\bar{F}_Z(u_n))^2} = t^{-2\alpha} E(\phi^{2\alpha}(Y_j)),$$

where we used (3.7a). Equivalently, $E\left((E(V_{j,n}(t)\|\mathcal{G}_{j-1}))^2\right) \sim t^{-2\alpha}E(\phi^{2\alpha}(Y_j))(\bar{F}_Z(u_n))^2$, as $n \to \infty$. Thus, as $n \to \infty$,

$$\begin{aligned}\mathrm{Var}(\triangle_j M_n(t)) &\sim t^{-\alpha}E\left(\phi^{\alpha}(Y_j)\right)\bar{F}_Z(u_n) - t^{-2\alpha}E\left(\phi^{2\alpha}(Y_j)\right)\left(\bar{F}_Z(u_n)\right)^2 \\ &\sim t^{-\alpha}E(\phi^{\alpha}(Y_j))\bar{F}_Z(u_n) \sim t^{-\alpha}\bar{F}_X(u_n),\end{aligned}$$

by Breiman's Lemma (3.11) and since the second term of the difference above is dominated by the first one. Also, we proved that

$$E\left[(\triangle_j M_n(t))^2\right] \sim t^{-\alpha}\bar{F}_X(u_n). \tag{4.12}$$

Hence, using the fact that $\left(\triangle_j M_n(t)\right)_j$ is a stationary martingale difference sequence,

$$\begin{aligned}\mathrm{Var}(M_n(t)) = \mathrm{Var}\left(\sum_{j=1}^{n}\triangle_j M_n(t)\right) &= \sum_{j=1}^{n}\mathrm{Var}(\triangle_j M_n(t)) \\ &= n\mathrm{Var}(\triangle_1 M_n(t)) \sim nt^{-\alpha}\bar{F}_X(u_n) \text{ as } n \to \infty.\end{aligned}$$

Hence, the desired result follows. □

Remark 4.2.5. We note that the above result is valid under the assumption

$$E(\phi^{2\alpha+\epsilon}(Y)) < \infty$$

instead of (3.7a).

Convergence of the marginal distributions.

Proposition 4.2.6. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $t > 0$,*

$$\frac{M_n(t)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} t^{-\frac{\alpha}{2}}\mathcal{N}, \tag{4.13}$$

where $\mathcal{N}$ stands for a standard normal random variable.

This limiting distribution coincides with the one defined in Theorem 4.2.3, that is,

$$t^{-\frac{\alpha}{2}}\mathcal{N} \stackrel{d}{=} (B \circ T)(t). \tag{4.14}$$

Note that $\stackrel{d}{=}$ refers to equality in distribution.

Proof. We recall that $(\triangle_j M_n)_j$ is a martingale difference. We define

$$M_n^*(t) := \frac{M_n(t)}{\sqrt{k}} \tag{4.15}$$

$$\triangle_j M_n^*(t) := \frac{\triangle_j M_n(t)}{\sqrt{k}} = \frac{V_{j,n}(t) - E(V_{j,n}(t) \| \mathcal{G}_{j-1})}{\sqrt{k}}. \tag{4.16}$$

To prove Proposition 4.2.6, it is sufficient to show that $\forall\, t, \epsilon > 0$, (cf. Theorem 2.6.9)

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t))^2 \| \mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} t^{-\alpha}, \tag{4.17a}$$

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t))^2\, \mathbb{1}_{\{|\triangle_j M_n^*(t)|>\epsilon\}} \| \mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} 0. \tag{4.17b}$$

- We begin by proving (4.17a). By definition, we have for all $t > 0$,

$$\begin{aligned} E\left((\triangle_j M_n^*(t))^2 \| \mathcal{G}_{j-1}\right) &= \frac{Var(V_{j,n}(t) \| \mathcal{G}_{j-1})}{n\bar{F}_X(u_n)} \\ &= \frac{E\left((V_{j,n}(t))^2 \| \mathcal{G}_{j-1}\right) - \left(E(V_{j,n}(t) \| \mathcal{G}_{j-1})\right)^2}{n\bar{F}_X(u_n)}. \end{aligned}$$

So, to show (4.17a), it suffices to prove the following set of conditions:

$$\frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} E\left((V_{j,n}(t))^2 \| \mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} t^{-\alpha} \tag{4.18a}$$

$$\frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left(E(V_{j,n}(t) \| \mathcal{G}_{j-1})\right)^2 \xrightarrow[n\to\infty]{p} 0. \tag{4.18b}$$

- First of all, let us prove (4.18a). We have

$$\begin{aligned} \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} E\left((V_{j,n}(t))^2 \| \mathcal{G}_{j-1}\right) &= \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} E\left(\mathbb{1}_{\{\phi(Y_j)Z_j > u_n t\}} \| \mathcal{G}_{j-1}\right) \\ = \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \bar{F}_Z\left(u_n t/\phi(Y_j)\right) &= \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \frac{\bar{F}_Z\left(u_n t/\phi(Y_j)\right)}{\bar{F}_Z(u_n)} \\ = \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} &\left(J_{u_n}\left(t/\phi(Y_j)\right) - t^{-\alpha}\phi^{\alpha}(Y_j) + t^{-\alpha}\phi^{\alpha}(Y_j)\right) \\ = \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} &\left(J_{u_n}\left(t/\phi(Y_j)\right) - t^{-\alpha}\phi^{\alpha}(Y_j)\right) + \frac{\bar{F}_Z(u_n)}{nt^{\alpha}\bar{F}_X(u_n)} \sum_{j=1}^{n} \phi^{\alpha}(Y_j). \end{aligned}$$

Then, (3.11), ergodicity and Slutsky's theorem yield w.p.1 that

$$\frac{\bar{F}_Z(u_n)}{nt^\alpha \bar{F}_X(u_n)} \sum_{j=1}^n \phi^\alpha(Y_j) \underset{n\to\infty}{\to} t^{-\alpha}.$$

We are left to show negligibility of the second term. It is sufficient to prove

$$I_n := \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^n \left(J_{u_n}\left(t/\phi(Y_j)\right) - t^{-\alpha}\phi^\alpha(Y_j)\right) = o_p(1).$$

In fact, by stationarity of $(Y_j)_j$, second order regular variation and (3.27a), we have

$$\begin{aligned} E(|I_n|) &\leq KC|\eta^*(u_n)|E\left(\left(t[\phi(Y)]^{-1}\right)^{-\kappa-\alpha}\left(t[\phi(Y)]^{-1} \vee t^{-1}\phi(Y)\right)^\epsilon\right) \\ &\leq KC|\eta^*(u_n)|\left(t^{-\kappa-\alpha+\epsilon}E(\phi^{\alpha+\kappa-\epsilon}(Y)) + t^{-\kappa-\alpha-\epsilon}E(\phi^{\alpha+\kappa+\epsilon}(Y))\right) \\ &\leq KC\left(t^{-(\alpha+\kappa+\epsilon)} \vee t^{-(\alpha+\kappa-\epsilon)}\right)[E(\phi^{\alpha+\kappa-\epsilon}(Y)) + E(\phi^{\alpha+\kappa+\epsilon}(Y))]|\eta^*(u_n)|. \end{aligned}$$

The second inequality holds by using the fact that for any nonnegative random variables X, Y, we have $X \vee Y \leq X + Y$.
Moreover, the sum of expectations in the last inequality is finite by (3.7b).
Therefore, $E(|I_n|) \underset{n\to\infty}{\to} 0$. Consequently, $n \to \infty$, $I_n \xrightarrow{p} 0$. Hence, (4.18a) is proven.

- Second of all, we prove (4.18b). Potter's bounds (cf. (3.14)), yields $\forall t > 0$, $\delta > 0$,

$$\begin{aligned} &\frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^n E\left((V_{j,n}(t))^2 \| \mathcal{G}_{j-1}\right) = \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^n \left(\bar{F}_Z\left(u_n t/\phi(Y_j)\right)\right)^2 \\ &= \frac{\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^n \left(J_{u_n}\left(t/\phi(Y_j)\right)\right)^2 \leq \frac{C\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^n \left(1 \vee t^{-1}\phi(Y_j)\right)^{2(\delta+\alpha)} \\ &\leq \frac{\bar{F}_Z^2(u_n)}{\bar{F}_X(u_n)} \left(C + C(t,\delta)\left(\frac{1}{n}\sum_{j=1}^n \phi^{2(\delta+\alpha)}(Y_j)\right)\right) \xrightarrow[n\to\infty]{p} 0. \end{aligned}$$

  It is important to note that $C(t, \delta)$ is a constant depending on t and δ but not on n. Note that the last inequality holds by ergodicity and Slutsky's Theorem. Thus, (4.18b) is proven and hence, so is (4.17a).

- To prove the Lindeberg condition (4.17b), we observe that

$$|\triangle_j M_n^*(t)| = \frac{1}{\sqrt{k}} |V_{j,n}(t) - E\left(V_{j,n}(t) \| \mathcal{G}_{j-1}\right)| \leq \frac{1}{\sqrt{k}} \underset{n\to\infty}{\to} 0.$$

  So, for arbitrary $\epsilon > 0$, $\mathbb{1}_{\{|\triangle_j M_n^*(t)| > \epsilon\}} = 0$, for all n sufficiently large. Thus, (4.17b) is trivially fulfilled.

□

Finite dimensional convergence of the martingale part

Proposition 4.2.7. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for any set of points $t_1, \ldots, t_k > 0$,*

$$\left(\frac{M_n(t_i)}{\sqrt{k}}\right)_{1 \le i \le k} \xrightarrow[n\to\infty]{d} (B \circ T(t_i))_{1 \le i \le k}\,, \tag{4.19}$$

where W denotes a standard Brownian motion. The covariance matrix of this limiting Gaussian process is therefore

$$\Sigma = \left((t_i \vee t_m)^{-\alpha}\right)_{i,m=1}^{k} \ . \tag{4.20}$$

In Lemma 4.2.11, we use the **Cramer-Wold device** (cf. Theorem 2.5.6) for the proof. In turn, the proof of Lemma 4.2.11 is supported by Lemma 4.2.8 and Lemma 4.2.10, which serve to check the set of conditions (2.62) of Theorem 2.6.9 for

$$\sum_{i=1}^{k} \frac{a_i M_n(t_i)}{\sqrt{k}}. \tag{4.21}$$

Lemma 4.2.8. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $t > 0$,*

$$\mathrm{Cov}\left(\frac{M_n(s)}{\sqrt{k}}, \frac{M_n(t)}{\sqrt{k}}\right) \underset{n\to\infty}{\rightarrow} (s \vee t)^{-\alpha}. \tag{4.22}$$

Remark 4.2.9. It will be clear from the proof that the lemma holds under the assumption $E(\phi^{\alpha+\epsilon}(Y)) < \infty$ instead of (3.7a).

Proof. We consider the sequence $(M_n)_n$ as defined in (4.8). Since

$$\begin{aligned}
&\mathrm{Cov}\left(M_n(s), M_n(t)\right) = \mathrm{Cov}\left(\sum_{j=1}^{n} \triangle_j M_n(s), \sum_{p=1}^{n} \triangle_p M_n(t)\right) \\
&= \sum_{\substack{j,p=1 \\ j\ne p}}^{n} \mathrm{Cov}\left(\triangle_j M_n(s), \triangle_p M_n(t)\right) + \sum_{j=1}^{n} \mathrm{Cov}\left(\triangle_j M_n(s), \triangle_j M_n(t)\right) \\
&= 0 + \sum_{j=1}^{n} \mathrm{Cov}\left(\triangle_j M_n(s), \triangle_j M_n(t)\right) \ .
\end{aligned}$$

It follows that

$$\operatorname{Cov}\left(M_n(s), M_n(t)\right) = \sum_{j=1}^{n} \operatorname{Cov}\left(V_{j,n}(s) - E(V_{j,n}(s)\|\mathcal{G}_{j-1}), V_{j,n}(t) - E(V_{j,n}(t)\|\mathcal{G}_{j-1})\right)$$

$$= \sum_{j=1}^{n} \left(\operatorname{Cov}(V_{j,n}(s), V_{j,n}(t)) - \operatorname{Cov}[V_{j,n}(s), E(V_{j,n}(t)\|\mathcal{G}_{j-1}))\right)$$

$$- \sum_{j=1}^{n} \operatorname{Cov}(E(V_{j,n}(s)\|\mathcal{G}_{j-1}), V_{j,n}(t)) + \sum_{j=1}^{n} \left(\operatorname{Cov}(E(V_{j,n}(s)\|\mathcal{G}_{j-1}), E(V_{j,n}(t)\|\mathcal{G}_{j-1}))\right)$$

$$= \sum_{j=1}^{n} \left(\operatorname{Cov}(V_{j,n}(s), V_{j,n}(t)) - \operatorname{Cov}(E(V_{j,n}(s)\|\mathcal{G}_{j-1}), E(V_{j,n}(t)\|\mathcal{G}_{j-1}))\right),$$

where the last equality holds by the following fact:

$$\operatorname{Cov}(X, E(Y\|\mathcal{F})) = \operatorname{Cov}(E(X\|\mathcal{F}), Y) = \operatorname{Cov}(E(X\|\mathcal{F}), E(Y\|\mathcal{F})).$$

Therefore,

$$\operatorname{Cov}\left(\frac{M_n(s)}{\sqrt{k}}, \frac{M_n(t)}{\sqrt{k}}\right) = \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \operatorname{Cov}\left(V_{j,n}(s), V_{j,n}(t)\right)$$

$$- \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \operatorname{Cov}\left(E(V_{j,n}(s)\|\mathcal{G}_{j-1}), E(V_{j,n}(t)\|\mathcal{G}_{j-1})\right).$$

Note that

$$\begin{aligned}
\operatorname{Cov}(V_{j,n}(s), V_{j,n}(t)) &= E(V_{j,n}(s)V_{j,n}(t)) - E(V_{j,n}(s))E(V_{j,n}(t)) \\
&= E\left(\mathbb{1}_{\{X_j > u_n s\}}\mathbb{1}_{\{X_j > u_n t\}}\right) - E\left(\mathbb{1}_{\{X_j > u_n s\}}\right) E\left(\mathbb{1}_{\{X_j > u_n t\}}\right) \\
&= E\left(\mathbb{1}_{\{X_j > u_n (s\vee t)\}}\right) - \bar{F}_X(u_n s)\bar{F}_X(u_n t) \\
&= \bar{F}_X\left(u_n(s\vee t)\right) - \bar{F}_X(u_n s)\bar{F}_X(u_n t),
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname{Cov}(E(V_{j,n}(s)\|\mathcal{G}_{j-1}), E(V_{j,n}(t)\|\mathcal{G}_{j-1})) \\
&= E\left(E(V_{j,n}(s)\|\mathcal{G}_{j-1})E(V_{j,n}(t)\|\mathcal{G}_{j-1})\right) - E(V_{j,n}(s))E(V_{j,n}(t)) \\
&= E\left(\bar{F}_Z\left(u_n[\phi(Y_j)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y_j)]^{-1}t\right)\right) - \bar{F}_X(u_n s)\bar{F}_X(u_n t).
\end{aligned}$$

Thus, by the stationarity of $(Y_j)_j$ and regular variation of $\{X_j\}$, we have

$$\begin{aligned}
&\operatorname{Cov}\left(\frac{M_n(s)}{\sqrt{k}},\frac{M_n(t)}{\sqrt{k}}\right)\\
&= \frac{1}{n}\sum_{j=1}^{n}\frac{\bar{F}_X\left(u_n(s\vee t)\right)-E\left(\bar{F}_Z\left(u_n[\phi(Y_j)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y_j)]^{-1}t\right)\right)}{\bar{F}_X(u_n)}\\
&= \frac{\bar{F}_X\left(u_n(s\vee t)\right)}{\bar{F}_X(u_n)}-\frac{E\left(\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)\right)}{\bar{F}_X(u_n)}\\
&= (s\vee t)^{-\alpha}(1+o(1))+O\left(\frac{E\left(\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)\right)}{\bar{F}_X(u_n)}\right).
\end{aligned}$$

On the other hand, by Breiman's Lemma (3.11), we have w.p.1,

$$\lim_{n\to\infty}\frac{\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)}{\bar{F}_X(u_n)}=\lim_{n\to\infty}\frac{\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)}{E(\phi^{\alpha}(Y))\bar{F}_Z(u_n)}=0.$$

Moreover, by Potter's bounds (cf. (3.14)), we have

$$\frac{\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)}{\bar{F}_Z(u_n)}\leq C(\epsilon)s^{-(\alpha+\epsilon)}\max\left(1,\phi^{\alpha+\epsilon}(Y)\right).$$

where $C(\epsilon)$ is a constant depending on ϵ, but not on n.
Furthermore, since for all $t>0$, $\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)\leq 1$, then

$$\frac{\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)}{E(\phi^{\alpha}(Y))\bar{F}_Z(u_n)}\leq K(s,\epsilon)\max\left(1,\phi^{\alpha+\epsilon}(Y)\right),$$

where $K(s,\epsilon)$ is a constant depending on s and ϵ, but not on n. So, the dominated convergence theorem guarantees the following:

$$\frac{E\left(\bar{F}_Z\left(u_n[\phi(Y)]^{-1}s\right)\bar{F}_Z\left(u_n[\phi(Y)]^{-1}t\right)\right)}{\bar{F}_X(u_n)}\underset{n\to\infty}{\to}0. \tag{4.23}$$

Thus, (4.22) holds. □

Lemma 4.2.10. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for any set of points $a_1,\ldots,a_k\in\mathbb{R}$ and $t_1,\ldots,t_k>0$,*

$$\operatorname{Var}\left(\sum_{i=1}^{k}\frac{a_iM_n(t_i)}{\sqrt{k}}\right)\underset{n\to\infty}{\to}\sum_{i=1}^{k}a_i^2(t_i)^{-\alpha}+2\sum_{1\leq i<m\leq k}a_ia_m(t_i\vee t_m)^{-\alpha}. \tag{4.24}$$

Proof. By definition, we have for all $a_1, \ldots, a_k \in \mathbb{R}$ and $t_1, \ldots, t_k > 0$,

$$\mathrm{Var}\left(\sum_{i=1}^{k} \frac{a_i M_n(t_i)}{\sqrt{k}}\right) = \sum_{i=1}^{k} a_i^2 \mathrm{Var}\left(\frac{M_n(t_i)}{\sqrt{k}}\right) + 2\sum_{i<m} \mathrm{Cov}\left(\frac{M_n(t_i)}{\sqrt{k}}, \frac{M_n(t_m)}{\sqrt{k}}\right).$$

Furthermore, by Lemma 4.2.4, we have for all $t_i > 0$:

$$\mathrm{Var}\left(\frac{M_n(t_i)}{\sqrt{k}}\right) \underset{n\to\infty}{\to} (t_i)^{-\alpha}.$$

By Lemma 4.2.8, we have for all $t_i, t_m > 0$:

$$\mathrm{Cov}\left(\frac{M_n(t_i)}{\sqrt{k}}, \frac{M_n(t_m)}{\sqrt{k}}\right) \underset{n\to\infty}{\to} (t_i \vee t_m)^{-\alpha}.$$

Thus, the result follows. □

To complete the proof of Proposition 4.2.7, we use the **Cramer-Wold device** which boils our task down to proving one dimensional weak convergence, that is (4.25).

Lemma 4.2.11. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for any set of points $a_1, \ldots, a_k \in \mathbb{R}$ and $t_1, \ldots, t_k > 0$,*

$$\sum_{i=1}^{k} \frac{a_i M_n(t_i)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} \sum_{i=1}^{k} a_i B \circ T(t_i). \tag{4.25}$$

Proof. We have for all $a_1, \ldots, a_k \in \mathbb{R}$ and $t_1, \ldots, t_k > 0$,

$$\sum_{i=1}^{k} a_i M_n(t_i) = \sum_{i=1}^{k}\sum_{j=1}^{n} a_i \triangle_j M_n(t_i) = \sum_{j=1}^{n}\sum_{i=1}^{k} a_i \triangle_j M_n(t_i) =: \sum_{j=1}^{n} \triangle_j M_n(t_1, \ldots, t_k),$$

where

$$\triangle_j M_n(t_1, \ldots, t_k) := \sum_{i=1}^{k} a_i \triangle_j M_n(t_i), \tag{4.26a}$$

$$\triangle_j M_n^*(t_1, \ldots, t_k) := \sum_{i=1}^{k} a_i \triangle_j M_n^*(t_i). \tag{4.26b}$$

We recall that the summands $\triangle_j M_n$ and $\triangle_j M_n^*$ are respectively defined in (4.9) and (4.16). We claim that $(\triangle_j M_n)_j$ is a martingale difference. In fact, for all $t_1, \ldots, t_k > 0$,

- $E\left(|\triangle_j M_n(t_1, \ldots, t_k)|\right) = E\left(\left|\sum_{i=1}^{k} a_i \triangle_j M_n(t_i)\right|\right) \le \sum_{i=1}^{k} |a_i| E(|\triangle_j M_n(t_i)|) < \infty.$

- $E\left(\triangle_j M_n(t_1,\ldots,t_k)\|\mathcal{G}_{j-1}\right) = \sum_{i=1}^{k} a_i E\left(\triangle_j M_n(t_i)\|\mathcal{G}_{j-1}\right) = 0, \text{ w.p.1.}$

Since $\left(\triangle_j M_n(t_1,\ldots,t_k)\right)_j$ is a martingale difference sequence, then by Theorem 2.6.9, proving Lemma 4.2.11 is equivalent to showing for all $t_1,\ldots,t_k,\epsilon > 0$,

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t_1,\ldots,t_k))^2\|\mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} \sum_{i=1}^{k} a_i^2 t_i^{-\alpha} + 2\sum_{i<m} a_i a_m (t_i \vee t_m)^{-\alpha}, \tag{4.27a}$$

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t_1,\ldots,t_k))^2 \mathbb{1}_{\{|\triangle_j M_n^*(t_1,\ldots,t_k)|>\epsilon\|\mathcal{G}_{j-1}\}}\right) \xrightarrow[n\to\infty]{p} 0. \tag{4.27b}$$

We begin by proving (4.27a). By definition, we have

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t_1,\ldots,t_k))^2\|\mathcal{G}_{j-1}\right) = \sum_{j=1}^{n} E\left(\left(\sum_{i=1}^{k} a_i \triangle_j M_n^*(t_i)\right)^2 \|\mathcal{G}_{j-1}\right)$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{k} a_i^2 E\left((\triangle_j M_n^*(t_i))^2\|\mathcal{G}_{j-1}\right) + 2\sum_{j=1}^{n}\sum_{i<m} a_i a_m E\left(\triangle_j M_n^*(t_i)\triangle_j M_n^*(t_m)\|\mathcal{G}_{j-1}\right)$$

$$= \sum_{i=1}^{k} a_i^2 \sum_{j=1}^{n} E\left((\triangle_j M_n^*(t_i))^2\|\mathcal{G}_{j-1}\right) + 2\sum_{i<m} a_i a_m \sum_{j=1}^{n} E\left(\triangle_j M_n^*(t_i)\triangle_j M_n^*(t_m)\|\mathcal{G}_{j-1}\right).$$

From (4.18a), we have that for $t_i > 0$,

$$\sum_{j=1}^{n} E\left((\triangle_j M_n^*(t_i))^2\|\mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} t_i^{-\alpha}.$$

In addition, we claim that for all t_i, $t_m > 0$,

$$\sum_{j=1}^{n} E\left(\triangle_j M_n^*(t_i)\triangle_j M_n^*(t_m)\|\mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} (t_i \vee t_m)^{-\alpha}. \tag{4.28}$$

In fact, we have by definition

$$E\left(\triangle_j M_n^*(t_i)\triangle_j M_n^*(t_m)\|\mathcal{G}_{j-1}\right) = \frac{1}{n\bar{F}_X(u_n)}\mathrm{Cov}\left(V_{j,n}(t_i), V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right)$$

$$= \frac{E\left(V_{j,n}(t_i)V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right)}{n\bar{F}_X(u_n)} - \frac{E\left(V_{j,n}(t_i)\|\mathcal{G}_{j-1}\right)E\left(V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right)}{n\bar{F}_X(u_n)}.$$

It follows that establishing (4.28) is sufficient to proving

$$\frac{1}{n\bar{F}_X(u_n)}\sum_{j=1}^{n} E\left(V_{j,n}(t_i)V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} (t_i \vee t_m)^{-\alpha}, \tag{4.29a}$$

$$\frac{1}{n\bar{F}_X(u_n)}\sum_{j=1}^{n} E\left(V_{j,n}(t_i)\|\mathcal{G}_{j-1}\right)E\left(V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} 0. \tag{4.29b}$$

Let us start with (4.29a). By definition, We have

$$\frac{1}{n\bar{F}_X(u_n)}\sum_{j=1}^{n} E\left(V_{j,n}(t_i)V_{j,n}(t_m)\|\mathcal{G}_{j-1}\right) = \frac{1}{n\bar{F}_X(u_n)}\sum_{j=1}^{n}\bar{F}_Z\left(u_n[\phi(Y_j)]^{-1}(t_i\vee t_m)\right)$$
$$= \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)}\sum_{j=1}^{n}\left(J_{u_n}\left(t_i\vee t_m/\phi(Y_j)\right) - (t_i\vee t_m)^{-\alpha}\phi^{\alpha}(Y_j) + (t_i\vee t_m)^{-\alpha}\phi^{\alpha}(Y_j)\right)$$
$$= A_n + B_n.$$

The summands A_n and B_n are respectively defined as follows:

$$\begin{aligned}
A_n : &= \frac{\bar{F}_Z(u_n)}{n(t_i\vee t_m)^{\alpha}\bar{F}_X(u_n)}\sum_{j=1}^{n}\phi^{\alpha}(Y_j),\\
B_n : &= \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)}\sum_{j=1}^{n}\left(J_{u_n}\left(t_i\vee t_m/\phi(Y_j)\right) - (t_i\vee t_m)^{-\alpha}\phi^{\alpha}(Y_j)\right)\\
&= \frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)}\sum_{j=1}^{n}\left(J_{u_n}\left(t_i\vee t_m/\phi(Y_j)\right) - T\left(t_i\vee t_m/\phi(Y_j)\right)\right).
\end{aligned}$$

First of all, consider A_n. By ergodicity, we have w.p.1,

$$\frac{1}{n}\sum_{j=1}^{n}\phi^{\alpha}(Y_j) \underset{n\to\infty}{\to} E\left(\phi^{\alpha}(Y_j)\right).$$

Slutsky's Theorem and (3.11) allow us to conclude w.p.1,

$$A_n \to (t_i\vee t_m)^{-\alpha}, \text{ as } n\to\infty.$$

Next, consider B_n. Showing $B_n \xrightarrow[n\to\infty]{p} 0$ is equivalent to proving

$$B_n^* := \frac{1}{n}\sum_{j=1}^{n}\left(J_{u_n}\left(t_i\vee t_m/\phi(Y_j)\right) - T\left(t_i\vee t_m/\phi(Y_j)\right)\right) \xrightarrow[n\to\infty]{p} 0.$$

Stationarity of $\{Y_j\}$, the second order regular variation and (3.27a) yield

$$\begin{aligned}
E(|B_n^*|) &\le E\left(|J_{u_n}\left(t_i\vee t_m/\phi(Y)\right) - T\left(t_i\vee t_m/\phi(Y_j)\right)|\right)\\
&\le C(\epsilon)E\left(\left(\frac{t_i\vee t_m}{\phi(Y)}\right)^{-(\kappa+\alpha+\epsilon)}\vee\left(\frac{t_i\vee t_m}{\phi(Y)}\right)^{-(\kappa+\alpha-\epsilon)}\right)|\eta^*(u_n)|\\
&\le K(\epsilon)(t_i\vee t_m)^{-(\kappa+\alpha+\epsilon)}\vee(t_i\vee t_m)^{-(\kappa+\alpha-\epsilon)}|\eta^*(u_n)| \underset{n\to\infty}{\to} 0,
\end{aligned}$$

where $K(\epsilon) = C(\epsilon)\left(E\left((\phi(Y))^{\kappa+\alpha+\epsilon}\right) + E\left((\phi(Y))^{\kappa+\alpha-\epsilon}\right)\right)$ is a constant depending on ϵ but not on n. Consequently, as $n \to \infty$, $B_n^* \xrightarrow{p} 0$.
This finishes the proof of (4.29a).

We continue with (4.29b). By Potter's bounds, we have $\forall\ \delta > 0$,

$$\begin{aligned}
&\frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} E\left(V_{j,n}(t_i) \| \mathcal{G}_{j-1}\right) E\left(V_{j,n}(t_m) \| \mathcal{G}_{j-1}\right) \\
&= \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \bar{F}_Z\left(u_n[\phi(Y_j)]^{-1} t_i\right) \bar{F}_Z\left(u_n[\phi(Y_j)]^{-1} t_m\right) \\
&= \frac{\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} J_{u_n}\left(t_i/\phi(Y_j)\right) J_{u_n}\left(t_m/\phi(Y_j)\right) \\
&\leq \frac{C(\delta)B(\delta)\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left(\left(1 \vee t_i^{-1}\phi(Y_j)\right)\left(1 \vee t_m^{-1}\phi(Y_j)\right)\right)^{(\delta+\alpha)} \\
&= \frac{C(\delta)B(\delta)\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left(1 \vee t_i^{-(\delta+\alpha)}\phi^{(\delta+\alpha)}(Y_j)\right)\left(1 \vee t_m^{-(\delta+\alpha)}\phi^{(\delta+\alpha)}(Y_j)\right) \\
&\leq \frac{C(\delta)B(\delta)\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left(1 + \left(t_i^{-1}\phi(Y_j)\right)^{\alpha+\delta}\right)\left(1 + \left(t_m^{-1}\phi(Y_j)\right)^{\alpha+\delta}\right) \\
&= \frac{C(\delta)B(\delta)\bar{F}_Z^2(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left(1 + \left(t_i^{-(\alpha+\delta)} + t_m^{-(\alpha+\delta)}\right)\phi^{(\alpha+\delta)}(Y_j) + (t_i t_m)^{-(\alpha+\delta)}\phi^{2(\alpha+\delta)}(Y_j)\right) \\
&= \frac{\bar{F}_Z^2(u_n)}{\bar{F}_X(u_n)} \left(R(\delta) + \frac{K(t_i,\delta) + G(t_m,\delta)}{n} \sum_{j=1}^{n} \phi^{(\delta+\alpha)}(Y_j) + \frac{I(t_i,t_m,\delta)}{n} \sum_{j=1}^{n} \phi^{2(\delta+\alpha)}(Y_j)\right),
\end{aligned}$$

where $C(\delta)$, $B(\delta)$, $R(\delta)$, $K(t_i,\delta)$, $G(t_m,\delta)$ and $I(t_i,t_m,\delta)$ are constants depending on t_i, t_m and δ but not on n. Note that the last inequality holds due to the fact that $X \vee Y \leq X + Y$, for nonnegative random variables X and Y. Regular variation of $(X_j)_j$ and $(Z_j)_j$, ergodicity and Slutsky's Theorem yield

$$\frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} E\left(V_{j,n}(t_i) \| \mathcal{G}_{j-1}\right) E\left(V_{j,n}(t_m) \| \mathcal{G}_{j-1}\right) \xrightarrow[n\to\infty]{p} 0.$$

Hence Slutsky's Theorem completes the proof of (4.29b) and hence the proof of (4.27a).

To prove the Lindeberg condition (4.27b), we observe that

$$|\triangle_j M_n^*(t)| = \frac{1}{\sqrt{k}} \left|V_{j,n}(t) - E\left(V_{j,n}(t) \| \mathcal{G}_{j-1}\right)\right| \leq \frac{1}{\sqrt{k}} \underset{n\to\infty}{\rightarrow} 0.$$

So, for arbitrary $\epsilon > 0, \mathbb{1}_{\{|\triangle_j M_n^*(t)|>\epsilon\}} = 0$, for all n sufficiently large. Thus, (4.27b) is trivially fulfilled. $\square$

Tightness of the martingale part

This section is entirely devoted to the study of tightness of M_n properly scaled as it appears in Proposition 4.2.12. The proof of tightness requires tedious and technical calculations. In the sequel, we consider for ease of notation the following setup: for all $0 < s < t < \infty$, let

$$V_{j,n}(s,t) := V_{j,n}(s) - V_{j,n}(t) = \mathbb{1}_{\{u_n s < \phi(Y_j) Z_j < u_n t\}}, \tag{4.30a}$$

$$\triangle_j M_n(s,t) := V_{j,n}(s,t) - E\left(V_{j,n}(s,t) \| \mathcal{G}_{j-1}\right), \tag{4.30b}$$

$$M_n(s,t) := M_n(s) - M_n(t) = \sum_{j=1}^{n} \triangle_j M_n(s,t), \tag{4.30c}$$

$$\triangle_j M_n^*(s,t) := \frac{\triangle_j M_n(s,t)}{\sqrt{k}} \text{ and } M_n^*(s,t) := \frac{M_n(s,t)}{\sqrt{k}}. \tag{4.30d}$$

Proposition 4.2.12. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then the process $M_n/\sqrt{k}$ is tight.*

The proof consists of checking the tightness conditions (2.55)-(2.57) of Theorem 2.5.14. We state and prove Lemma 4.2.13 and Lemma 4.2.14 as prerequisite for Lemma 4.2.15, which corresponds to (2.55). Next, we establish Lemma 4.2.16, which corresponds to (2.57) in Theorem 2.5.14. We wrap up the proof of Proposition 4.2.12 at the end of the section.

Lemma 4.2.13. *If X is a nonnegative random variable, then*

$$E\left((X - E\left(X\|\mathcal{G}\right))^4\right) \leq 8E\left(X^4\right). \tag{4.31}$$

Proof. Assume that X is a nonnegative random variable. We have

$$\begin{aligned} E\left((X - E\left(X\|\mathcal{G}\right))^4\right) &= E\left(X^4\right) - 4E\left(X^3 E\left(X\|\mathcal{G}\right)\right) + 6E\left(X^2\left(E\left(X\|\mathcal{G}\right)\right)^2\right) \\ &\quad -4E\left(X^3\left(E\left(X\|\mathcal{G}\right)^3\right)\right) + E\left(\left(E\left(X\|\mathcal{G}\right)\right)^4\right) \\ &\leq E\left(X^4\right) + 6E\left(X^2\left(E\left(X\|\mathcal{G}\right)\right)^2\right) + E\left(\left(E\left(X\|\mathcal{G}\right)\right)^4\right) \\ &\leq 2E\left(X^4\right) + 6E\left(X^2\left(E\left(X^2\|\mathcal{G}\right)\right)\right), \end{aligned}$$

where the last inequality holds by the tower property and Jensen's inequality.

$$E\left(X^2\left(E\left(X^2\|\mathcal{G}\right)\right)\right) = E\left(E\left(X^2 E\left(X^2\|\mathcal{G}\right)\|\mathcal{G}\right)\right) = E\left(\left(E\left(X^2\|\mathcal{G}\right)\right)^2\right) \leq E\left(X^4\right),$$

where the first equality holds by the tower property, the second, by measurability, and finally, the last inequality, by Jensen's inequality. Thus, the desired result holds. □

Lemma 4.2.14. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then,*

$$E\left((M_n^*(s,t))^4\right) \leq 2C_4K^2E\left((J_{u_n}(s/\phi(Y)) - J_{u_n}(t/\phi(Y)))^2\right) + \frac{16C_4}{n\bar{F}_X(u_n)}\left|T_{u_n}(s) - T_{u_n}(t)\right|,$$

where C_4 stands for Rosenthal's constant (cf. Theorem 2.6.8) and K comes from (3.13).

Proof. By Rosenthal's inequality, it holds that (cf. Theorem 2.6.8)

$$E\left((M_n^*(s,t))^4\right) \leq C_4\left(\sum_{j=1}^n E\left((\triangle_j M_n^*(s,t))^4\right) + E\left(\left(\sum_{j=1}^n E\left((\triangle_j M_n^*(s,t))^2 \|\mathcal{G}_{j-1}\right)\right)^2\right)\right).$$

- By stationarity of $\left(\triangle_j M_n^*(s,t)\right)_j$ and Lemma 4.2.13, we obtain that

$$\begin{aligned}\sum_{j=1}^n E\left((\triangle_j M_n^*(s,t))^4\right) &= nE\left((\triangle_1 M_n^*(s,t))^4\right) = \frac{E\left((\triangle_1 M_n(s,t))^4\right)}{n\left(\bar{F}_X(u_n)\right)^2}\\ &\leq \frac{8E\left((V_{1,n}(t,s))^4\right)}{n\left(\bar{F}_X(u_n)\right)^2} = \frac{8E\left(V_{1,n}(t,s)\right)}{n\left(\bar{F}_X(u_n)\right)^2} = \frac{8\left(\bar{F}_X(u_ns) - \bar{F}_X(u_nt)\right)}{n\left(\bar{F}_X(u_n)\right)^2}.\end{aligned}$$

  So, it follows that

$$\sum_{j=1}^n E\left((\triangle_j M_n^*(s,t))^4\right) \leq \frac{8}{n\bar{F}_X(u_n)}\left|T_{u_n}(s) - T_{u_n}(t)\right| \cdot \tag{4.32}$$

- On the other hand, we have

$$\begin{aligned}E\left(\left(\sum_{j=1}^n E\left((\triangle_j M_n^*(s,t))^2 \|\mathcal{G}_{j-1}\right)\right)^2\right) &= E\left(\sum_{j=1}^n \left(E\left((\triangle_j M_n^*(s,t))^2 \|\mathcal{G}_{j-1}\right)\right)^2\right)\\ &+ 2E\left(\sum_{i<j}^n E\left((\triangle_i M_n^*(s,t))^2 \|\mathcal{G}_{i-1}\right)E\left((\triangle_j M_n^*(s,t))^2 \|\mathcal{G}_{j-1}\right)\right).\end{aligned}$$

  i) Stationarity of $\left(\triangle_j M_n^*(s,t)\right)_j$, Jensen's inequality, and Lemma 4.2.13 yield

$$\begin{aligned}E\left(\sum_{j=1}^n \left(E\left((\triangle_j M_n^*(s,t))^2 \|\mathcal{G}_{j-1}\right)\right)^2\right) &= nE\left(\left(E\left((\triangle_1 M_n^*(t,s))^2 \|\mathcal{G}_0\right)\right)^2\right)\\ &\leq nE\left((\triangle_1 M_n^*(s,t))^4\right) \leq \frac{8E\left(V_{1,n}(t,s)\right)}{n\left(\bar{F}_X(u_n)\right)^2} = \frac{8}{n\bar{F}_X(u_n)}\left|T_{u_n}(s) - T_{u_n}(t)\right|.\end{aligned}$$

ii) Cauchy-Schwartz's inequality and stationarity of $\left(\triangle_j M_n^*(s,t)\right)_j$ ensure that

$$
\begin{aligned}
&2E\left(\sum_{i<j}^{n} E\left(\left(\triangle_i M_n^*(s,t)\right)^2 \|\mathcal{G}_{i-1}\right) E\left(\left(\triangle_j M_n^*(s,t)\right)^2 \|\mathcal{G}_{j-1}\right)\right)\\
&\leq \; 2n(n-1)E\left(\left(E\left(\left(\triangle_1 M_n^*(s,t)\right)^2 \|\mathcal{G}_0\right)\right)^2\right)\\
&= \; \frac{2n(n-1)}{\left(n\bar{F}_X(u_n)\right)^2}E\left(\left(E\left(\left(\triangle_1 M_n(s,t)\right)^2 \|\mathcal{G}_0\right)\right)^2\right)\\
&\leq \; \frac{2}{\left(\bar{F}_X(u_n)\right)^2}E\left(\left(E\left(V_{1,n}(s,t)\|\mathcal{G}_0\right)\right)^2\right)\\
&\leq \; 2\left(\frac{\bar{F}_Z(u_n)}{\bar{F}_X(u_n)}\right)^2 E\left(\left(\frac{\bar{F}_Z\left(u_n s/\phi(Y)\right)}{\bar{F}_Z(u_n)} - \frac{\bar{F}_Z\left(u_n t/\phi(Y)\right)}{\bar{F}_Z(u_n)}\right)^2\right).
\end{aligned}
$$

Therefore, by the same argument as in (3.13), it follows that for $n \geq 1$,

$$
\begin{aligned}
E\left(\left(\sum_{j=1}^{n} E\left(\left(\triangle_j M_n^*(s,t)\right)^2 \|\mathcal{G}_{j-1}\right)\right)^2\right) \leq \frac{8}{n\bar{F}_X(u_n)}\left|T_{u_n}(s) - T_{u_n}(t)\right| \\
+2K^2 E\left(\left(J_{u_n}\left(s/\phi(Y)\right) - J_{u_n}\left(t/\phi(Y)\right)\right)^2\right).
\end{aligned} \tag{4.33}
$$

Thus, (4.32) and (4.33) imply that for all $n \geq 1$,

$$
\begin{aligned}
E\left(\left(M_n^*(s,t)\right)^4\right) \;\leq\; & 2C_4 K^2 E\left(\left(J_{u_n}\left(s/\phi(Y)\right) - J_{u_n}\left(t/\phi(Y)\right)\right)^2\right)\\
& + \frac{16C_4}{n\bar{F}_X(u_n)}\left|T_{u_n}(s) - T_{u_n}(t)\right|.
\end{aligned} \tag{4.34}
$$

□

Lemma 4.2.15. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then for all $0 < a \leq s, t \leq b < \infty$, there exists $C_{a,b,4} > 0$ such that*

$$
E\left(\left(M_n^*(s,t)\right)^4\right) \leq C_{a,b,4}\left(\frac{|s-t|}{n\bar{F}_X(u_n)} + (s-t)^2\right). \tag{4.35}
$$

Proof. By the mean value theorem and (3.20), there exists $\tau = \tau(\omega) \in (s,t)$, such that:

$$
\begin{aligned}
&E\left(\left(J_{u_n}\left(s/\phi(Y)\right) - J_{u_n}\left(t/\phi(Y)\right)\right)^2\right) = (s-t)^2 E\left(\left(J'_{u_n}\left([\phi(Y)]^{-1}\tau\right)[\phi(Y)]^{-1}\right)^2\right)\\
&\leq \left(M(s-t)\right)^2 E\left(\left(\left([\phi(Y)]^{-1}\tau\right)^{-(\alpha+\beta+1)} \vee \left([\phi(Y)]^{-1}\tau\right)^{-(\alpha-\beta+1)}[\phi(Y)]^{-1}\right)^2\right)\\
&\leq \left(M(s-t)\right)^2 E\left(\left(\frac{[\phi(Y)]^{\alpha+\beta}}{\tau^{\alpha+\beta+1}} \vee \frac{[\phi(Y)]^{\alpha-\beta}}{\tau^{\alpha-\beta+1}}\right)^2\right) \leq C_{a,b}\left(s-t\right)^2 \;,
\end{aligned}
$$

where $C_{a,b} = \left(M \max\left(a^{-\alpha-1}\left(a^{-\beta} \vee b^{\beta}\right)\right)\right)^2 E\left(\left([\phi(Y)]^{\alpha+\beta} + [\phi(Y)]^{\alpha-\beta}\right)^2\right)$. The constant is finite by (3.7a). Hence, (4.34) becomes:

$$E\left((M_n^*(s,t))^4\right) \;\leq\; \frac{16C_4}{n\bar{F}_X(u_n)} \left|T_{u_n}(s) - T_{u_n}(t)\right| + 2C_4 C_{a,b} K^2 (s-t)^2 .$$

Again by the mean value theorem, there exists $\tau^* \in (s,t)$ such that:

$$E\left((M_n^*(s,t))^4\right) \leq \frac{16C_4}{n\bar{F}_X(u_n)} |s-t| T'_{u_n}(\tau^*) + 2K^2 C_4 C_{a,b}(s-t)^2.$$

Therefore (3.21) allows us to use (3.23) in order to obtain

$$E\left((M_n^*(s,t))^4\right) \leq \frac{16C_4K_0}{n\bar{F}_X(u_n)} \left(a^{-\alpha-1}\left(a^{-\beta} \vee b^{\beta}\right)\right) |s-t| + 2K^2 C_4 C_{a,b}(s-t)^2.$$

Thus, the desired result holds by taking

$$C_{a,b,4} = \max\left(16C_4K_0\left(a^{-\alpha-1}\left(a^{-\beta} \vee b^{\beta}\right)\right), 2K^2C_4C_{a,b}\right) .$$

□

Lemma 4.2.16. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then the process M_n^* can be decomposed as a difference of two non-increasing processes M_n° and $M_n^{\circ\circ}$, that is*

$$M_n^*(t) = M_n^{\circ}(t) - M_n^{\circ\circ}(t), \tag{4.36}$$

where the summand processes M_n° and $M_n^{\circ\circ}$ are respectively defined by for all $t > 0$:

$$M_n^{\circ}(t) = \frac{1}{\sqrt{k}} \sum_{j=1}^{n} V_{j,n}(t) \tag{4.37a}$$

$$M_n^{\circ\circ}(t) = \frac{1}{\sqrt{k}} \sum_{j=1}^{n} \bar{F}_Z\left(u_n t/\phi(Y_j)\right). \tag{4.37b}$$

Furthermore,

$$\max_{0 \leq i \leq l_n} \left|M_n^{\circ\circ}(a + t_{i+1,n}) - M_n^{\circ\circ}(a + t_{i,n})\right| \xrightarrow[n\to\infty]{p} 0 , \tag{4.38}$$

where

$$t_{i,n} := \frac{i}{n\bar{F}_X(u_n)} , \;\; l_n = \left[(b-a)n\bar{F}_X(u_n)\right] , \;\; t_{l_n+1} := b-a,$$

with $[\cdot]$ denoting the integer part.

Proof. Recall that (3.10) yields $J_{u_n}(t) = \bar{F}_Z(u_n t)/\bar{F}_Z(u_n)$. The decomposition in (4.36) is straightforward from (4.8). We have

$$\begin{aligned}
M_n^{\circ\circ}(t) &:= \frac{1}{\sqrt{k}} \sum_{j=1}^{n} \bar{F}_Z\left(u_n t/\phi(Y_j)\right) \\
&= \sqrt{k}\frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} J_{u_n}\left(t/\phi(Y_j)\right).
\end{aligned}$$

It follows that for $0 \le i \le l_{n-1}$,

$$\begin{aligned}
&\left|M_n^{\circ\circ}\left(a+t_{i+1,n}\right) - M_n^{\circ\circ}\left(a+t_{i,n}\right)\right| \\
&= \left|M_n^{\circ\circ}\left(a + \frac{i+1}{n\bar{F}_X(u_n)}\right) - M_n^{\circ\circ}\left(a + \frac{i}{n\bar{F}_X(u_n)}\right)\right| \\
&= \sqrt{k}\frac{\bar{F}_Z(u_n)}{n\bar{F}_X(u_n)} \sum_{j=1}^{n} \left|J_{u_n}\left(\frac{a+t_{i+1,n}}{\phi(Y_j)}\right) - J_{u_n}\left(\frac{a+t_{i,n}}{\phi(Y_j)}\right)\right| \\
&\le Kn^{-1}\sqrt{k} \sum_{j=1}^{n} \left|J_{u_n}\left(\frac{a+t_{i+1,n}}{\phi(Y_j)}\right) - J_{u_n}\left(\frac{a+t_{i,n}}{\phi(Y_j)}\right)\right|,
\end{aligned}$$

where the last inequality holds by (3.13). The mean value theorem and (3.20) yield that there exists $\tau_{i,n,j} = \tau_{i,n,j}(\omega) \in (t_{i,n}, t_{i,n+1})$, which depends on $\phi(Y_j)$, such that if $i \le l_{n-1}$,

$$\begin{aligned}
&\left|M_n^{\circ\circ}\left(a+t_{i+1,n}\right) - M_n^{\circ\circ}\left(a+t_{i,n}\right)\right| \\
&\le K\sqrt{k} \sum_{j=1}^{n} \frac{1}{n\bar{F}_X(u_n)\phi(Y_j)} \left|J'_{u_n}\left(\frac{a+\tau_{i,n,j}}{\phi(Y_j)}\right)\right| \\
&\le \frac{MK}{n\sqrt{k}} \sum_{j=1}^{n} \frac{1}{\phi(Y_j)} \left(\left(\frac{a+\tau_{i,n,j}}{\phi(Y_j)}\right)^{-\alpha-\beta-1} \vee \left(\frac{a+\tau_{i,n,j}}{\phi(Y_j)}\right)^{-\alpha+\beta-1}\right) \\
&\le \frac{MK}{\sqrt{k}} \max\left(a^{-\alpha-1}\left(a^{-\beta} \vee b^{\beta}\right)\right) \frac{1}{n} \sum_{j=1}^{n} \left(\phi^{\alpha+\beta}(Y_j) + \phi^{\alpha-\beta}(Y_j)\right) .
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\max_{0 \le i \le l_{n-1}} \left|M_n^{\circ\circ}\left(a+t_{i+1,n}\right) - M_n^{\circ\circ}\left(a+t_{i,n}\right)\right| \\
&\le \frac{K}{\sqrt{k}} \left(\frac{1}{n} \sum_{j=1}^{n} \left(\phi^{\alpha+\beta}(Y_j) + \phi^{\alpha-\beta}(Y_j)\right)\right) \xrightarrow[n\to\infty]{p} 0.
\end{aligned}$$

The last convergence holds by the Law of Large Numbers and the assumption (3.7a).

For $i = l_n$, since $M_n^{\circ\circ}$ is monotone and $b < a + \frac{l_n+1}{n\bar{F}_X(u_n)}$, we obtain

$$\begin{aligned}\left|M_n^{\circ\circ}\left(a + t_{l_{n+1},n}\right) - M_n^{\circ\circ}\left(a + t_{l_n,n}\right)\right| &= \left|M_n^{\circ\circ}(b) - M_n^{\circ\circ}\left(a + t_{l_n,n}\right)\right| \\ &\leq \left|M_n^{\circ\circ}\left(a + \frac{l_n+1}{n\bar{F}_X(u_n)}\right) - M_n^{\circ\circ}\left(a + \frac{l_n}{n\bar{F}_X(u_n)}\right)\right| .\end{aligned}$$

By the same argument as above the last term converges to zero in probability.

□

Proof of Proposition 4.2.12.

Proof. We recall the notation of Theorem 2.5.14. Let

$$\gamma = 4,\ \ \delta = 2,\ \ c_n = 1/n\bar{F}_X(u_n) \text{ and } \xi_n = M_n^*.$$

Lemma 4.2.16 yields (2.57) on the interval $[a, b]$ (cf. the second remark following Theorem 2.5.14).

Letting $s = a$ and $t \to \infty$ in the statement of Lemma 4.2.14 we obtain via (3.14)

$$\begin{aligned}E\left((M_n^*(a))^4\right) &\leq 2C_4K^2E\left((J_{u_n}(a/\phi(Y)))^2\right) + \frac{16C_4}{n\bar{F}_X(u_n)}T_{u_n}(a) \\ &\leq 2C_4K^2C^2(\beta)\left(1 + a^{-2(\alpha+\beta)}E\big(\,\phi^{2(\alpha+\beta)}(Y)\,\big)\right) + o(1)\ ,\end{aligned}$$

hence the fourth moment is bounded proving the first part of (2.55).

From Lemma 4.2.15, if moreover $|s - t| \geq c_n = 1/(n\bar{F}_X(u_n))$, we obtain

$$E\left((M_n^*(s,t))^4\right) \leq C_{a,b,4}\left(\frac{|s-t|}{n\bar{F}_X(u_n)} + (s-t)^2\right) \leq 2C_{a,b,4}(s-t)^2.$$

This completes the proof of (2.55). Hence, by Theorem 2.5.14 and the subsequent remark we conclude that the process M_n^* is tight in $D[a, b]$, where $0 < a < b < \infty$. Since a, b are arbitrary this implies tightness on $D(0, \infty)$.

□

4.2.3 Weak Convergence of the Long Memory Part

In this section we consider the process L_n defined in (4.7b). Our goal is to determine the asymptotic behaviour of L_n suitably normalized. Recall (2.88), that is

$$C_m = \frac{2}{[1 - m(1-2d)][2 - m(1-2d)]}.$$

Theorem 4.2.17. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1).*

- *If $m(1-2d) < 1$, then*
$$\frac{L_n(t)}{kb_{n,m}/n} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)T(t)}{m!E\left(\phi^\alpha(Y)\right)}\xi_{m,d+1/2}(1) \text{ in } D(0,\infty). \tag{4.39}$$
The limiting random variable, $\xi_{m,d+1/2}(1)$, is Hermite-Rosenblatt.

- *If $m(1-2d) > 1$, then*
$$\frac{\sqrt{n}L_n(t)}{k} \xrightarrow[n\to\infty]{d} t^{-\alpha}\frac{\sigma\mathcal{N}}{E\left(\phi^\alpha(Y)\right)} \text{ in } D(0,\infty), \tag{4.40}$$
where σ^2 is defined in (2.93) and $\mathcal{N}$ denotes a standard normal random variable.

Proof. By (4.7b), (3.10) and the fact that $E(V_{j,n}(t) = E\left(E(V_{j,n}(t))||\mathcal{G}_{j-1}\right)$, we have
$$\begin{aligned}
\frac{L_n(t)}{\bar{F}_Z(u_n)} &= \sum_{j=1}^n \left(\frac{\bar{F}_Z\left(u_n t/\phi(Y_j)\right)}{\bar{F}_Z(u_n)} - E\left(\frac{\bar{F}_Z\left(u_n t/\phi(Y_j)\right)}{\bar{F}_Z(u_n)}\right)\right) \\
&= \sum_{j=1}^n \left(J_{u_n}\left(t/\phi(Y_j)\right) - E\left(J_{u_n}\left(t/\phi(Y_j)\right)\right)\right).
\end{aligned}$$
By regular variation of $(Z_j)_j$, it holds w.p.1 that
$$J_{u_n}\left(t/\phi(Y_j)\right) \underset{n\to\infty}{\to} T\left(t/\phi(Y_j)\right) := t^{-\alpha}\phi^\alpha(Y_j).$$
This motivates the following decomposition:
$$\frac{L_n(t)}{\bar{F}_Z(u_n)} = \sum_{j=1}^3 L_{n,j}(t), \tag{4.41}$$
where the summands $L_{n,j}$'s are respectively defined as follows:
$$L_{n,1}(t) := \sum_{j=1}^n \left[J_{u_n}\left(t/\phi(Y_j)\right) - T\left(t/\phi(Y_j)\right)\right], \tag{4.42a}$$
$$L_{n,2}(t) := \sum_{j=1}^n \left[T\left(t/\phi(Y_j)\right) - E\left(T\left(t/\phi(Y_j)\right)\right)\right], \tag{4.42b}$$
$$L_{n,3}(t) := \sum_{j=1}^n \left[E\left(T\left(t/\phi(Y_j)\right)\right) - E\left(J_{u_n}\left(t/\phi(Y_j)\right)\right)\right]. \tag{4.42c}$$

We start by establishing weak convergence of $L_{n,2}$. Notice that

$$L_{n,2}(t) = t^{-\alpha} \sum_{j=1}^{n} \left(\phi^\alpha(Y_j) - E(\phi^\alpha(Y_j))\right) = t^{-\alpha} \sum_{j=1}^{n} G_\alpha(Y_j),$$

where $G_\alpha(\cdot) = \phi^\alpha(\cdot) - E\left(\phi^\alpha(\cdot)\right)$. Then, if m denotes the Hermite rank of G_α, we obtain:

$$\mu_{\phi,\alpha}(m) = E\left(H_m(Y)G_\alpha(Y)\right) = E\left(H_m(Y)\phi^\alpha(Y)\right).$$

Assume first that $m(1-2d) < 1$. Therefore, by Theorem 2.7.23, if $m(1-2d) < 1$, then for $t > t_0 > 0$,

$$\frac{L_{n,2}(t)}{b_{n,m}} = \frac{t^{-\alpha}}{b_{n,m}} \sum_{j=1}^{n} G_\alpha(Y_j) \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)T(t)}{m!} \xi_{m,d+1/2}(1),$$

in the uniform topology on every compact subset of $(0,\infty)$. It remains to show that $L_{n,1} + L_{n,3}$ is negligible, when divided by $b_{n,m}$. By stationarity of $\{Y_j\}$, (3.27a), (3.6) and (3.7b), we have for every $t_0 > 0$,

$$\begin{aligned}
&\frac{1}{b_{n,m}} E\left(\sup_{t>t_0} |L_{n,1}(t)|\right) \le \frac{n}{b_{n,m}} E\left(\sup_{t>t_0} \left|J_{u_n}\left(t/\phi(Y)\right) - T\left(t/\phi(Y)\right)\right|\right) \\
&\le C(\epsilon) E\left(\sup_{t>t_0} \left((t/\phi(Y))^{-(\kappa+\alpha+\epsilon)} \vee (t/\phi(Y))^{-(\kappa+\alpha-\epsilon)} \right)\right) \frac{n}{b_{n,m}} |\eta^*(u_n)| \\
&\le K(\epsilon) \sup_{t>t_0} \left(t^{-(\kappa+\alpha+\epsilon)} \vee t^{-(\kappa+\alpha-\epsilon)}\right) \frac{n}{b_{n,m}} |\eta^*(u_n)| \underset{n\to\infty}{\to} 0,
\end{aligned}$$

where $K(\epsilon) = C(\epsilon)\left(E\left(\phi^{\alpha+\kappa+\epsilon}(Y)\right) + E\left(\phi^{\alpha+\kappa-\epsilon}(Y)\right)\right)$ is a constant depending on ϵ but not on n. This allows us to conclude that

$$\frac{1}{b_{n,m}} E\left(\sup_{t>t_0} |L_{n,1}(t)|\right) = o\left(1\right). \tag{4.43}$$

Finally, we consider the process $L_{n,3}$. Since

$$\frac{1}{b_{n,m}} \sup_{t>t_0} |L_{n,3}(t)| \le \frac{1}{b_{n,m}} E\left(\sup_{t>t_0} |L_{n,1}(t)|\right) \underset{n\to\infty}{\to} 0.$$

Then, we conclude that

$$\frac{1}{b_{n,m}} \sup_{t>t_0} |L_{n,3}(t)| = o_P\left(1\right). \tag{4.44}$$

Altogether, $(L_{n,1} + L_{n,3})/b_{n,m}$ is negligible on compact subsets of $(0,\infty)$. Therefore,

$$\frac{L_n(t)}{b_{n,m} \bar{F}_Z(u_n) E\left(\phi^\alpha(Y)\right)} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)T(t)}{m! E\left(\phi^\alpha(Y)\right)} \xi_{m,d+1/2}(1) \text{ in } D(0,\infty).$$

Thus (3.11) and Slutsky's Theorem end the proof of (4.39) for the case $m(1-2d) < 1$.

Now, assume that $m(1-2d) > 1$. We keep the same notation and the decompositions as for the previous case. By Theorem 2.7.23, if $m(1-2d) > 1$, we have

$$\frac{L_{n,2}(t)}{\sqrt{n}} \xrightarrow[n\to\infty]{d} \sigma t^{-\alpha}\mathcal{N}\ , \text{ in } D(0,\infty),$$

where σ^2 is defined in (2.93). Moreover, by (3.6),

$$\frac{1}{\sqrt{n}}E\left(\sup_{t>t_0}|L_{n,1}(t)|\right) \le K(\epsilon)\frac{n}{\sqrt{n}}|\eta^*(u_n)|\sup_{t>t_0}\left(t^{-(\kappa+\alpha+\epsilon)} \vee t^{-(\kappa+\alpha-\epsilon)}\right) \underset{n\to\infty}{\rightarrow} 0.$$

The corresponding argument applies to $L_{n,3}$. Thus,

$$\frac{L_n(t)}{\bar{F}_Z(u_n)E\left(\phi^\alpha(Y)\right)\sqrt{n}} \xrightarrow[n\to\infty]{d} t^{-\alpha}\frac{\sigma\mathcal{N}}{E\left(\phi^\alpha(Y)\right)} \text{ in } D(0,\infty).$$

Again, Breiman's Lemma and Slutsky's Theorem finish the proof for $m(1-2d) > 1$. □

4.2.4 Weak convergence of the TEP with Deterministic Levels

The main result of this section is Theorem 4.2.18. It essentially pertains to the different regimes to which the appropriately scaled process $\widetilde{S}_n$ defined in (4.5) converges weakly. In other words, depending on the interplay between the rates of convergence in Theorem 4.2.3 and Theorem 4.2.17, the asymptotic behaviour of $\widetilde{S}_n$ is dominated either by the martingale part M_n or the long memory part L_n, defined in (4.7a) and (4.7b), respectively.

Theorem 4.2.18. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1).*

- *If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to 0$, as $n \to \infty$, then*

$$\frac{\widetilde{S}_n(t)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} (B\circ T)(t) \text{ in } D(0,\infty). \tag{4.45}$$

- *If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, as $n \to \infty$, then*

$$\frac{\widetilde{S}_n(t)}{kb_{n,m}/n} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)T(t)}{m!E(\phi^\alpha(Y))}\xi_{m,d+1/2}(1) \text{ in } D(0,\infty). \tag{4.46}$$

- *If $m(1-2d) > 1$, then (4.45) holds.*

Proof. Let $m(1-2d)<1$. It follows from the decomposition in (4.6) that

$$\frac{\widetilde{S}_n(t)}{\sqrt{k}}=\frac{M_n(t)}{\sqrt{k}}+\frac{L_n(t)}{b_{n,m}\bar{F}_X(u_n)}\frac{b_{n,m}}{n}\sqrt{k}.$$

Assume that Theorems 4.2.3 and 4.2.17 hold. If $\frac{b_{n,m}}{n}\sqrt{k}\to 0$, as $n\to\infty$, then

$$\frac{L_n(t)}{b_{n,m}\bar{F}_X(u_n)}\frac{b_{n,m}}{n}\sqrt{k}=o_P(1),$$

uniformly in t on all compact subsets of $(0,\infty)$. Thus, (4.45) follows. Again, the martingale-long memory decomposition (cf. (4.6)) yields that

$$\frac{\widetilde{S}_n(t)}{kb_{n,m}/n}=\frac{M_n(t)}{\sqrt{k}}\frac{1}{b_{n,m}}\sqrt{\frac{n}{\bar{F}_X(u_n)}}+\frac{L_n(t)}{b_{n,m}\bar{F}_X(u_n)}.$$

Analogously, assume that Theorems 4.2.3 and 4.2.17 hold. If $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, as $n\to\infty$, then

$$\frac{M_n(t)}{\sqrt{k}}\frac{1}{b_{n,m}}\sqrt{\frac{n}{\bar{F}_X(u_n)}}=o_P(1).$$

Thus, (4.46) follows by Slutsky's Theorem. Now if $m(1-2d)>1$, then

$$\frac{\widetilde{S}_n(t)}{\sqrt{k}}=\frac{M_n(t)}{\sqrt{k}}+\frac{L_n(t)}{\sqrt{n}\bar{F}_X(u_n)}\frac{\bar{F}_X(u_n)}{\sqrt{\bar{F}_X(u_n)}}.$$

By (4.40) and since $\bar{F}_X(u_n)\to 0$ as $n\to\infty$, (4.45) follows. □

Remark 4.2.19. We note that leverage has no effect on the limiting distribution. Long memory affects the limiting behaviour. We have a dichotomous behaviour, according to strength of long memory (that is, the value of the parameter d that appears explicitly in the definition of $b_{n,m}$). In the long memory case, the limiting random variable, $\xi_{m,d+1/2}(1)$, is Hermite-Rosenblatt. It is non-Gaussian unless $m=1$.

Remark 4.2.20. If $\frac{b_{n,m}}{n}\sqrt{k}\to c\in(0,\infty)$, then $\widetilde{S}_n(t)/\sqrt{k}$ converges to a linear combination of the processes that appear on the right-hand sides of (4.45) and (4.46). However, the dependence structure is unclear.

Remark 4.2.21. Since the distribution F_X is not known, the upper quantiles u_n cannot be observed. Therefore, in the following section we replace u_n with an appropriate order statistic.

4.3 Random Levels

The tail empirical process defined in (4.5) is unobservable because $\bar{F}_X(u_n)$ is unknown. Therefore, the choice of a suitable level u_n is conditioned on the knowledge of $\bar{F}_X$. So, any result based on such a tail empirical process is purely theoretical. This motivates the introduction of a data based tail empirical process (see Definition 4.3.1).

Let $X_1, \ldots, X_n$ be a sample from a stochastic volatility model defined in (3.1). Let $X_{(1)} \le \cdots \le X_{(i)} \le \cdots \le X_{(n)}$ be their corresponding order statistics. Let F_X denote their distribution function. The quantile function $\bar{F}_X^{\leftarrow}$ is defined by

$$\bar{F}_X^{\leftarrow}(p) = \inf\{x \in \mathbb{R} : \bar{F}_X(x) \le p\}, \quad 0 < p < 1.$$

Choose a sequence of integers $(k_n)_n$ such that as $n \to \infty$,

$$k := k_n \to \infty \text{ and } k/n \to 0. \tag{4.47}$$

Let $F_{n,X}$ denote the usual empirical distribution function defined by

$$F_{n,X}(x) := \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{X_j \le x\}}\ , t > 0.$$

Consequently, $\bar{F}_{n,X}(x) = 1 - F_{n,X}(x)$. Let $u_n = \bar{F}_X^{\leftarrow}(n/k)$. By continuity of $\bar{F}_X$, $\{u_n\}$ satisfies $\bar{F}_X(u_n) = k/n$. So, $u_n \to \infty$ and $n\bar{F}_X(u_n) \to \infty$. Let

$$\bar{F}_{n,X}^{\leftarrow}(p) = \inf\{x \in \mathbb{R} : \bar{F}_{n,X}(x) \le p\}, \quad 0 < p < 1.$$

Then $\bar{F}_{n,X}^{\leftarrow}(k/n) = X_{(n-k)}$. Thus, it is natural to approximate u_n with $X_{(n-k)}$.

4.3.1 Weak Convergence of the TEP with Random Levels

This section is devoted to the study of weak convergence of the tail empirical process with random levels defined below in (4.49).

Definition 4.3.1. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1).*

- *The empirical tail distribution function with random levels of $\{X_j\}$ is*

$$\widehat{T}_n(t) := \frac{1}{k}\sum_{j=1}^{n} \mathbb{1}_{\{X_j > X_{(n-k)}t\}}\ , \quad t > 0. \tag{4.48}$$

- *The tail empirical process with random levels of* $\{X_j\}$ *is*

$$\widehat{S}_n(t) := k\left(\widehat{T}_n(t) - T(t)\right). \tag{4.49}$$

Our goal is to determine the limiting behaviour of the process $\widehat{S}_n$.

Remark 4.3.2. The centering T_{u_n} used in (4.5) is not desirable since it depends on n. So, for statistical purposes, we will need to replace T_{u_n} by its limit T. For this, we require the no-bias condition (3.6). This yields the introduction of the following TEP:

$$\widetilde{S}_n^*(t) : = k\left(\widetilde{T}_n(t) - T(t)\right) = n\bar{F}_X(u_n)\left(\widetilde{T}_n(t) - T(t)\right),\ t > 0.$$

Under the no-bias condition (3.6) we have (3.28b), and hence the results for Theorem 4.2.18 remain valid for the process $\widetilde{S}_n^*$. For notational convenience, we introduce the sequence $(\rho_n)_n$, defined by

$$\rho_n := \frac{X_{(n-k)}}{u_n}. \tag{4.50}$$

The next result is about convergence in distribution of $(\rho_n)_n$ jointly with $\widetilde{S}_n^*$. This will serve as an ingredient for establishing weak convergence of $\widehat{S}_n$.

Lemma 4.3.3. *Let* $\{X_j\}$ *be the long memory stochastic volatility model with leverage as in* (3.1).

- *If* $m(1-2d) < 1$ *and* $\frac{b_{n,m}}{n}\sqrt{k} \to 0$, *or* $m(1-2d) > 1$, *then as* $n \to \infty$,

$$\left(\frac{\widetilde{S}_n^*(t)}{\sqrt{k}}\,,\sqrt{k}\left(\frac{X_{(n-k)}}{u_n} - 1\right)\right) \xrightarrow[n\to\infty]{d} \left((B\circ T)(t)\,,\frac{B(1)}{\alpha}\right). \tag{4.51}$$

- *If* $m(1-2d) < 1$ *and* $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$ *then as* $n \to \infty$,

$$\begin{aligned}&\left(\frac{n}{kb_{n,m}}\widetilde{S}_n^*(t)\,,\frac{n}{b_{n,m}}\left(\frac{X_{(n-k)}}{u_n} - 1\right)\right)\\ &\xrightarrow[n\to\infty]{d} \left(\frac{\mu_{\phi,\alpha}(m)}{m!E(\phi^\alpha(Y))}T(t)\xi_{m,d+1/2}(1), \frac{\mu_{\phi,\alpha}(m)}{\alpha m!E(\phi^\alpha(Y))}\xi_{m,d+1/2}(1)\right).\end{aligned} \tag{4.52}$$

These two joint weak convergences hold in $D(0,\infty)\times\mathbb{R}$.

Proof. For conciseness, we prove (4.51); the proof of (4.52) is analogous. Recall that $D_1(0,\infty)$ and $C_1(0,\infty)$ are respectively the set of non-increasing functions in $D(0,\infty)$ and the set of continuous non-increasing and positive functions in $C(0,\infty)$. We have

$$\begin{cases} \widetilde{S}_n^*/\sqrt{k} \in D(0,\infty)\ ,\ B\circ T \in C(0,\infty), \\ \qquad T^{\leftarrow} \in D_1(0,\infty)\ ,\ T^{\leftarrow} \in C_1(0,\infty)\ . \end{cases}$$

Therefore, Corollary 2.5.19 yields

$$\frac{\widetilde{S}_n^*}{\sqrt{k}} \circ T^{\leftarrow} \xrightarrow[n\to\infty]{d} B \text{ in } D(0,\infty).$$

This is equivalent to writing

$$\sqrt{k}\left\{\widetilde{T}_{u_n} \circ T^{\leftarrow}(t) - t\right\} \xrightarrow[n\to\infty]{d} W(t).$$

By Skorokhod's representation Theorem 2.5.5, there exist a probability space and processes $Z_n^\dagger$, $B^\dagger$ and $\widetilde{T}_n^\dagger$ such that:

$$Z_n^\dagger \stackrel{d}{=} \frac{\widetilde{S}_n^*}{\sqrt{k}} \circ T^{\leftarrow}\ , \quad B^\dagger \stackrel{d}{=} B\ , \quad Z_n^\dagger \xrightarrow[n\to\infty]{a.s} B^\dagger\ ,$$
$$T_{u_n}^\dagger(\cdot) := \frac{Z_n^\dagger(T(\cdot))}{\sqrt{k}} + T(\cdot) \stackrel{d}{=} \widetilde{T}_{u_n}(\cdot).$$

We note that almost sure convergence of $Z_n^\dagger$ to $B^\dagger$ is uniform on compact subsets of $(0,\infty)$. Recall that $T^{\leftarrow}$ is non-increasing. Moreover, $T_{u_n}^\dagger$ is almost surely non-increasing since $T_{u_n}^\dagger \stackrel{d}{=} \widetilde{T}_{u_n}$. So, $T \circ T_n^{\dagger\leftarrow}$ is almost surely nondecreasing. Vervaat's Lemma 2.2.5 holds and yields

$$\sqrt{k}\left\{T \circ T_n^{\dagger\leftarrow}(t) - t\right\} \xrightarrow[n\to\infty]{a.s} -B^\dagger(t) \text{ in } D(0,\infty).$$

In particular, for $t = 1$, we get

$$\sqrt{k}\left\{T \circ T_n^{\dagger\leftarrow}(1) - 1\right\} \xrightarrow[n\to\infty]{a.s} -B^\dagger(1).$$

But Taylor's expansion yields

$$\begin{aligned} \left\{T \circ T_n^{\dagger\leftarrow}(1) - 1\right\} &= \left\{T[T_n^{\dagger\leftarrow}(1)] - T\left[T^{\leftarrow}(1)\right]\right\} \\ &= T'[T^{\leftarrow}(1)]\left(T_n^{\dagger\leftarrow}(1) - 1\right)\left(1 + o_{a.s}\left(T_n^{\dagger\leftarrow}(1) - 1\right)\right). \end{aligned}$$

It follows that

$$\sqrt{k}T'(1)\left(T_n^{\dagger\leftarrow}(1)-1\right) \xrightarrow[n\to\infty]{a.s} -B^{\dagger}(1).$$

As a result, this convergence holds in distribution as well. Going back to the initial probability space, since $T_{u_n}^{\dagger} \stackrel{d}{=} \widetilde{T}_{u_n}$ and $T_n^{\dagger\leftarrow}(1) \stackrel{d}{=} \widetilde{T}_n^{\leftarrow}(1) = \rho_n$, we obtain

$$\sqrt{k}\left(\frac{X_{(n-k)}}{u_n}-1\right) \xrightarrow[n\to\infty]{d} \frac{B(1)}{\alpha}.$$

$\square$

We are now ready to establish weak convergence for the tail empirical process $\widehat{S}_n$. Notice that by introducing random levels, the tail empirical process vanishes at 1 and ∞, which forces the limiting process to be of a bridge type. More surprisingly, the introduction of random levels causes the effect of long memory to disappear. The reason for this, as will be seen in the proof, is that the limiting behaviour of $\widehat{S}_n$ follows informally from the continuous mapping theorem applied to $\widetilde{S}_n^*$ and $\frac{X_{(n-k)}}{u_n}$. Thanks to the degenerate structure of the limiting process for $\widetilde{S}_n^*$ (that is, a random variable scaled by a deterministic function), the long memory effect cancels out. Once again, the presence of leverage does not affect the limit.

Theorem 4.3.4. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then,*

$$\frac{\widehat{S}_n(t)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B\left(T(t)\right)-T(t)B(1), \tag{4.53}$$

in $D(0,\infty)$ equipped with the Skorokhod J_1 topology. The limiting process $B\left(T(\cdot)\right) - T(\cdot)B(1)$ is a centered time-changed Brownian bridge on $[1,\infty)$. Further, for each $t>0$,

$$B\left(T(t)\right)-T(t)B(1) \quad \stackrel{d}{=} \quad \left(t^{-\alpha}+t^{-2\alpha}-2\left(t^{-\alpha}\wedge t^{-2\alpha}\right)\right)^{1/2}\mathcal{N}\,, \tag{4.54}$$

where $\mathcal{N}$ stands for a standard normal random variable.

Proof. The process $\widehat{S}_n$ defined in (4.49) can be decomposed as follows:

$$\widehat{S}_n(t) = \widehat{S}_{n,1}(t)+\widehat{S}_{n,2}(t)+\widehat{S}_{n,3}(t),$$

where the summands $\widehat{S}_{n,j}$'s are respectively defined as follows:

$$\widehat{S}_{n,1}(t) = n\bar{F}_X(u_n)\left(\widetilde{T}_n\left(\rho_n t\right)-T_{u_n}\left(\rho_n t\right)\right), \tag{4.55}$$

$$\widehat{S}_{n,2}(t) = n\bar{F}_X(u_n)\left(T_{u_n}\left(\rho_n t\right)-T\left(\rho_n t\right)\right), \tag{4.56}$$

$$\widehat{S}_{n,3}(t) = n\bar{F}_X(u_n)\left(T\left(\rho_n t\right)-T(t)\right). \tag{4.57}$$

Since $\widehat{S}_{n,1}(t) = \widetilde{S}_n(\rho_n t)$, $\widetilde{T}_n(\rho_n) = 1$ and $T(\rho_n t) = T(\rho_n)T(t)$, then

$$\begin{aligned}
\widehat{S}_n(t) &= \left(\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)\right) + T(t)\widetilde{S}_n(\rho_n) + \widehat{S}_{n,2}(t) + \widehat{S}_{n,3}(t) \\
&= \left(\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)\right) + n\bar{F}_X(u_n)T(t)\left(\widetilde{T}_n(\rho_n) - T_{u_n}(\rho_n)\right) + \widehat{S}_{n,2}(t) + \widehat{S}_{n,3}(t) \\
&= \left(\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)\right) + n\bar{F}_X(u_n)T(t)\left(1 - T_{u_n}(\rho_n)\right) + \widehat{S}_{n,3}(t) + \widehat{S}_{n,2}(t) \\
&= \left(\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)\right) - n\bar{F}_X(u_n)T(t)\left(T_{u_n}(\rho_n) - T(\rho_n)\right) + \widehat{S}_{n,2}(t).
\end{aligned} \tag{4.58}$$

Notice that the last two terms combined in (4.58) are negligible thanks to the no-bias condition (3.6) (cf. (3.28b)) and the fact that $\rho_n = 1 + o_P(1)$, that is for all $t_0 > 0$,

$$\sup_{t \geq t_0} \frac{\left| -n\bar{F}_X(u_n)T(t)\left(T_{u_n}(\rho_n) - T(\rho_n)\right) + \widehat{S}_{n,2}(t)\right|}{\sqrt{k}} = o_P(1).$$

Hence, (4.53) holds if yielded by the first term in (4.58). This will be proven in Proposition 4.3.5. □

Proposition 4.3.5. *Under the conditions of Theorem 4.3.4, we have*

$$\frac{\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B(T(t)) - T(t)B(1), \tag{4.59}$$

in $D(0,\infty)$ equipped with the Skorokhod J_1 topology.

Proof. The "martingale-long memory decomposition" in (4.6) yields

$$\frac{\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)}{\sqrt{k}} = \left(\frac{M_n(\rho_n t) - T(t)M_n(\rho_n)}{\sqrt{k}}\right) + \left(\frac{L_n(\rho_n t) - T(t)L_n(\rho_n)}{\sqrt{k}}\right). \tag{4.60}$$

The idea is that in Lemma 4.3.6, we will prove that the first term in (4.60) converges to the right hand side of (4.59) while in Lemma 4.3.7, we will prove that the second term in (4.60) is negligible. □

Lemma 4.3.6. *Under the conditions of Theorem 4.3.4, we have*

$$\frac{M_n(\rho_n t) - T(t)M_n(\rho_n)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B(T(t)) - T(t)B(1), \tag{4.61}$$

in $D(0,\infty)$ equipped with the Skorokhod J_1 topology.

Proof. Since weak convergence to a continuous limit implies uniform convergence on compact sets, then by virtue of (4.10) and (4.51), we conclude that

$$\frac{M_n(\rho_n t) - T(t)M_n(\rho_n)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B\left(T(t)\right) - T(t)B(1),$$

in $D(0,\infty)$. In fact, the following hold

$$\frac{M_n(\rho_n t)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B\left(T(t)\right) \text{ in } D(0,\infty) \text{ and } \frac{M_n(\rho_n)}{\sqrt{k}} \xrightarrow[n\to\infty]{d} B\left(T(1)\right).$$

In addition, taking into account (4.19), we obtain for all $t > 0$,

$$\begin{aligned} Var\left(B(T(t)) - T(t)B(1)\right) &= Var\left(B(T(t))\right) + T^2(t)Var\left(B(1)\right) \\ &\quad -2T(t)Cov\left(B(T(t)), B(1)\right) \\ &= t^{-\alpha} + t^{-2\alpha} - 2(t^{-\alpha} \wedge t^{-2\alpha}). \end{aligned}$$

Since the limiting process $B\left(T(t)\right) - T(t)B(1)$ is a centered time-change Brownian bridge, therefore Gaussian, we conclude (4.54). □

Lemma 4.3.7. *Under the conditions of Theorem 4.3.4, we have for each* $t_0 > 0$,

$$\sup_{t\geq t_0}\left(\frac{L_n(\rho_n t) - T(t)L_n(\rho_n)}{\sqrt{k}}\right) = o_P(1).$$

Proof. It follows from (4.41) that

$$\frac{L_n(\rho_n t) - T(t)L_n(\rho_n)}{\sqrt{k}} = \frac{\bar{F}_Z(u_n)}{\sqrt{k}} \sum_{j=1}^{3} \left(L_{n,j}(\rho_n t) - T(t)L_{n,j}(\rho_n)\right),$$

where the summands $L_{n,j}$ are respectively defined in (4.42a)-(4.42c). Recall the proof of Theorem 4.2.17. There, the term $L_{n,2}$ yielded the long memory limit. The main point of the proof of the present lemma is that the term $L_{n,2}(\rho_n t) - T(t)L_{n,2}(\rho_n)$ vanishes. As such, the long memory part does not contribute to the limiting behaviour of the tail empirical process with random levels. For sake of conciseness, let for $t > 0$,

$$\begin{aligned} &\lambda(t) := t^{-(\alpha+\kappa+\epsilon)} \vee t^{-(\alpha+\kappa-\epsilon)} + t^{-\alpha} \\ &\Theta_{n,j} := \frac{1}{\sqrt{k}} \sup_{t\geq t_0}\left(\bar{F}_Z(u_n)\left|L_{n,j}(\rho_n t) - T(t)L_{n,j}(\rho_n)\right|\right),\ j = 1,3 \\ &\zeta_n := \rho_n^{-(\alpha+\kappa+\epsilon)}\left(\frac{1}{n}\sum_{j=1}^{n}\phi^{-(\alpha+\kappa+\epsilon)}(Y_j)\right) + \rho_n^{-(\alpha+\kappa-\epsilon)}\left(\frac{1}{n}\sum_{j=1}^{n}\phi^{\alpha+\kappa-\epsilon}(Y_j)\right). \end{aligned}$$

i) First, we have by (3.27a)

$$\begin{aligned}
&|L_{n,1}(\rho_n t) - T(t)L_{n,1}(\rho_n)| \leq \sum_{j=1}^{n} |J_{u_n}(\rho_n t/\phi(Y_j)) - T(\rho_n t/\phi(Y_j))| \\
&+ T(t)\sum_{j=1}^{n} |J_{u_n}(\rho_n/\phi(Y_j)) - T(\rho_n/\phi(Y_j))| \\
&\leq C(\epsilon)|\eta^*(u_n)| \sum_{j=1}^{n} \left((\rho_n t/\phi(Y_j))^{-(\alpha+\kappa+\epsilon)} \vee (\rho_n t/\phi(Y_j))^{-(\alpha+\kappa-\epsilon)} \right) \\
&+ C(\epsilon)T(t)|\eta^*(u_n)| \sum_{j=1}^{n} \left((\rho_n/\phi(Y_j))^{-(\alpha+\kappa+\epsilon)} \vee (\rho_n/\phi(Y_j))^{-(\alpha+\kappa-\epsilon)} \right) \\
&\leq nC(\epsilon)\lambda(t)|\eta^*(u_n)|\zeta_n.
\end{aligned} \tag{4.62}$$

As a consequence, we obtain for all $t_0 > 0$,

$$\varTheta_{n,1} \quad \leq \quad C(\epsilon) \sup_{t \geq t_0} \lambda(t) \sqrt{\frac{\bar{F}_Z(u_n)}{\bar{F}_X(u_n)}} \sqrt{n\bar{F}_Z(u_n)} |\eta^*(u_n)| \zeta_n = o_P(1).$$

In fact, by ergodicity, for $\delta = \alpha + \kappa - \epsilon$ or $\delta = -(\alpha + \kappa + \epsilon)$, with w.p.1,

$$\frac{1}{n} \sum_{j=1}^{n} \phi^{\delta}(Y_j) \underset{n \to \infty}{\to} E\left(\phi^{\delta}(Y_1)\right).$$

Moreover, since $\rho_n = 1 + o_P(1)$, so are $\rho_n^{-(\alpha+\kappa+\epsilon)}$ and $\rho_n^{-(\alpha+\kappa-\epsilon)}$, by the continuous mapping theorem. Hence, $\zeta_n = O_P(1)$. On account of (3.11) and (3.6),

$$\sqrt{\frac{\bar{F}_Z(u_n)}{\bar{F}_X(u_n)}} \sqrt{n\bar{F}_Z(u_n)} |\eta^*(u_n)| = o(1).$$

ii) Second, since $L_{n,2}(\rho_n t) = (\rho_n t)^{-\alpha} \sum_{j=1}^{n} (\phi^{\alpha}(Y_j) - E(\phi^{\alpha}(Y_j)))$ and $T(t)L_{n,2}(\rho_n) = t^{-\alpha}\rho_n^{-\alpha} \sum_{j=1}^{n} (\phi^{\alpha}(Y_j) - E(\phi^{\alpha}(Y_j)))$, it follows that

$$L_{n,2}(\rho_n t) - T(t)L_{n,2}(\rho_n) = 0. \tag{4.63}$$

iii) As for the third term, we recall again from (3.27a)

$$|J_{u_n}(t/\phi(Y_1)) - T(t/\phi(Y_1))| \leq C(\epsilon) \left((t/\phi(Y_1))^{-(\alpha+\kappa+\epsilon)} \vee (t/\phi(Y_1))^{-(\alpha+\kappa-\epsilon)} \right) |\eta^*(u_n)|.$$

Applying Jensen's inequality, we have

$$\begin{aligned}
|L_{n,3}(t)| &= n|E(J_{u_n}(t/\phi(Y_1))) - E(T(t/\phi(Y_1)))| \\
&\leq nE(|J_{u_n}(t/\phi(Y_1)) - T(t/\phi(Y_1))|) \\
&\leq nBC|\eta^*(u_n)| \left(t^{-(\alpha+\kappa+\epsilon)} \vee t^{-(\alpha+\kappa-\epsilon)}\right) .
\end{aligned}$$

As a consequence, we obtain

$$|L_{n,3}(\rho_n t) - T(t)L_{n,3}(\rho_n)| \leq nBC|\eta^*(u_n)|\Lambda(t) \left(\rho_n^{-(\alpha+\kappa+\epsilon)} \vee \rho_n^{-(\alpha+\kappa-\epsilon)}\right) . \tag{4.64}$$

The same argument as the one used for $\Theta_{n,1}$ yields for all $t_0 > 0$, $\Theta_{n,3} = o_P(1)$.

□

Remark 4.3.8. We note that to get the convergence (4.61) of the martingale part with random levels, we used Theorem 4.2.3. However, to prove that the long memory part with random levels is negligible, we did not use Theorem 4.2.17. The reason for this is that the long memory part with random levels vanishes due to the no-bias condition. As such, we do not need to consider cases $m(1-2d) < 1$ or $m(1-2d) > 1$ as we did in Theorem 4.2.17.

4.3.2 Weak Convergence of Integral Functionals

The power of weak convergence theory comes from the fact that many diverse results emerge as corollaries of a basic convergence theorem. As we shall see in Theorem 4.3.9, our main convergence Theorem 4.3.4 can be extended to integral functionals of the TEP. This in turn yields a unified approach to establishing weak convergence of estimators of the tail index (see Section 4.3.3) and makes estimation of risk measures feasible as will be seen in Chapter 5. In what follows, r denotes a nonnegative integer.

Theorem 4.3.9. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). If $\alpha > 2(1-r)$, then*

$$\frac{1}{\sqrt{k}} \int_1^\infty \frac{\widehat{S}_n(t)}{t^r}\, dt \xrightarrow[n\to\infty]{d} \int_1^\infty \frac{B\left(T(t)\right) - B(1)T(t)}{t^r}\, dt. \tag{4.65}$$

Proof. We first show finiteness of the limiting variance of the random variable in the right-hand side of (4.65). This implies that the limiting process in (4.65) is Gaussian since it is a continuous linear functional of a Gaussian process. Since $\alpha > 2(1-r)$, we have

$$\begin{aligned}
& Var\left(\int_1^\infty \frac{B\left(T(t)\right) - B(1)T(t)}{t^r}\, dt\right) \\
&= Var\left(\int_1^\infty \frac{B(T(t))}{t^r}\, dt - B(1)\int_1^\infty t^{-\alpha-r}\, dt\right) \\
&= Var\left(\int_1^\infty \frac{B(T(t))}{t^r}\, dt\right) + \frac{Var\left(B(1)\right)}{(\alpha+r-1)^2} \\
&\quad -\frac{2}{\alpha+r-1} Cov\left(B(1), \int_1^\infty \frac{B(T(t))}{t^r}\, dt\right) \\
&= \frac{\alpha}{(\alpha+r-1)^2(\alpha+2r-2)} < \infty. \qquad (4.66)
\end{aligned}$$

In fact, by Fubini's theorem and (4.14), we have

$$\begin{aligned}
Var\left(\int_1^\infty \frac{B(T(t))}{t^r}\, dt\right) &= E\left(\int_1^\infty \frac{B(T(t))}{t^r}\, dt \int_1^\infty \frac{B(T(s))}{s^r}\, ds\right) \\
&= 2\int_1^\infty \left(\int_s^\infty \frac{E\left(B^2(T(t))\right)}{t^r}\, dt\right) \frac{ds}{s^r} \\
&= \frac{2}{(\alpha+r-1)(\alpha+2r-2)}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
Cov\left(B(1), \int_1^\infty \frac{B(T(t))}{t^r}\, dt\right) &= E\left(B(1)\int_1^\infty \frac{B(T(t))}{t^r}\, dt\right) \\
&= \int_1^\infty \frac{E\left(B(T(t))B(1)\right)}{t^r}\, dt \\
&= \frac{1}{\alpha+r-1}. \qquad (4.67)
\end{aligned}$$

The calculations above justify the limiting variance. In what follows, we will check the assumptions of Theorem 2.5.20 to establish weak convergence of the process in the left-hand side of (4.65). To do so, let $M \geq 1$ and decompose

$$\frac{1}{\sqrt{k}}\int_1^\infty \frac{\widehat{S}_n(t)}{t^r}\, dt = \frac{1}{\sqrt{k}}\int_1^M \frac{\widehat{S}_n(t)}{t^r}\, dt + \frac{1}{\sqrt{k}}\int_M^\infty \frac{\widehat{S}_n(t)}{t^r}\, dt\ .$$

Since the integral functionals are continuous only over compact intervals, then the continuous mapping theorem and (4.53) yield

$$\frac{1}{\sqrt{k}}\int_1^M \frac{\widehat{S}_n(t)}{t^r}\, dt \xrightarrow[n\to\infty]{d} \int_1^M \frac{B\left(T(t)\right) - B(1)T(t)}{t^r}\, dt.$$

Also, by (4.66), we have

$$\lim_{M\to\infty} Var\left(\int_M^\infty \frac{B\left(T(t)\right)-B(1)T(t)}{t^r}\,dt\right)=0.$$

Hence,

$$\int_1^M \frac{B\left(T(t)\right)-B(1)T(t)}{t^r}\,dt \xrightarrow[M\to\infty]{d} \int_1^\infty \frac{B\left(T(t)\right)-B(1)T(t)}{t^r}\,dt.$$

Thus, (4.65) holds if we establish (2.60c), that is, $\forall\ \epsilon\geq 0$,

$$\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\frac{1}{\sqrt{k}}\int_M^\infty \frac{\widehat{S}_n(t)}{t^r}\,dt\right|\geq\epsilon\right)=0.$$

This will be proven in Proposition 4.3.10.

□

Proposition 4.3.10. *Under the conditions of Theorem 4.3.9, we have for all* $\epsilon\geq 0$*,*

$$\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\frac{1}{\sqrt{k}}\int_M^\infty \frac{\widehat{S}_n(t)}{t^r}\,dt\right|\geq\epsilon\right)=0. \tag{4.68}$$

Proof. Recall the definition of ρ_n in (4.50) and the decomposition (4.58), that is:

$$\widehat{S}_n(t)=\left(\widetilde{S}_n\left(\rho_n t\right)-T(t)\widetilde{S}_n\left(\rho_n\right)\right)-n\bar{F}_X(u_n)T(t)\left(T_{u_n}\left(\rho_n\right)-T\left(\rho_n\right)\right)+\widehat{S}_{n,2}(t).$$

Since it is hard to work with random levels, we go back to the process with the deterministic threshold. We have for all $\epsilon\geq 0$,

$$\begin{aligned}
&\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{\widehat{S}_n(t)}{t^r\sqrt{k}}\,dt\right|\geq\frac{\epsilon}{3}\right)\\
\leq\ &\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{\widetilde{S}_n\left(\rho_n t\right)-T(t)\widetilde{S}_n\left(\rho_n\right)}{t^r\sqrt{k}}\,dt\right|\geq\frac{\epsilon}{3}\right)\\
&+\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{n\bar{F}_X(u_n)T(t)\left(T_{u_n}\left(\rho_n\right)-T\left(\rho_n\right)\right)}{t^r\sqrt{k}}\,dt\right|\geq\frac{\epsilon}{3}\right)\\
&+\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{\widehat{S}_{n,2}(t)}{t^r\sqrt{k}}\,dt\right|\geq\frac{\epsilon}{3}\right).
\end{aligned}\tag{4.69}$$

Therefore, (4.68) holds if these three upper bounds in (4.69) vanish. We will show this in Lemmas 4.3.11 to 4.3.13, respectively. □

Lemma 4.3.11. *Under the conditions of Theorem 4.3.9, we have for each* $\epsilon \geq 0$*,*

$$\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)}{t^r\sqrt{k}}\,dt\right| \geq \epsilon\right) = 0. \tag{4.70}$$

Proof. For ease of notation, let

$$A_n^M := P\left(\left|\int_M^\infty \frac{\widetilde{S}_n(\rho_n t) - T(t)\widetilde{S}_n(\rho_n)}{t^r\sqrt{k}}\,dt\right| \geq \epsilon\right).$$

The change of variable $v = \rho_n t$ yields that

$$A_n^M = P\left(|\rho_n|^{m-1}\left|\int_{M\rho_n}^\infty \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \epsilon\right).$$

Since $\rho_n = 1 + o_P(1)$, then it suffices to deal with

$$\widetilde{A}_n^M = P\left(\left|\int_{M\rho_n}^\infty \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \epsilon\right).$$

We have

$$\widetilde{A}_n^M \leq \widetilde{A}_n^{M,1} + \widetilde{A}_n^{M,2},$$

where

$$\widetilde{A}_n^{M,1} = P\left(\left|\int_{M\rho_n}^M \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \frac{\epsilon}{2}\right),$$

$$\widetilde{A}_n^{M,2} = P\left(\left|\int_M^\infty \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \frac{\epsilon}{2}\right).$$

In what follows, we establish negligibility of both terms. For the first one, we will apply directly weak convergence result for the process:

$$\frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}.$$

For the second term, we will proceed with the martingale-long memory decomposition.

1. We establish negligibility of $\widetilde{A}_n^{M,1}$. Let $\delta \geq 0$. We have

$$\begin{aligned}
\widetilde{A}_n^{M,1} \leq\ & P\left(\left|\int_{M\rho_n}^M \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \frac{\epsilon}{2},\ |\rho_n - 1| < \delta\right) \\
& + P\left(\left|\int_{M\rho_n}^M \frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \geq \frac{\epsilon}{2},\ |\rho_n - 1| \geq \delta\right) \\
\leq\ & P\left(\int_{M(1-\delta)}^{M(1+\delta)} \left|\frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\right|\,dv \geq \frac{\epsilon}{2}\right) + P\left(|\rho_n - 1| \geq \delta\right).
\end{aligned}$$

- $\limsup_{n\to\infty} P\left(|\rho_n - 1| \geq \delta\right) = 0$, since $\rho_n = 1 + o_P(1)$.
- By weak convergence of $\tilde{S}_n$, $\rho_n = 1 + o_P(1)$ and Markov's inequality, we get

$$\begin{aligned}
&\limsup_{n\to\infty} P\left(\int_{M(1-\delta)}^{M(1+\delta)} \left|\frac{\widetilde{S}_n(v) - T\left(v/\rho_n\right)\widetilde{S}_n(\rho_n)}{v^r\sqrt{k}}\right| dv \geq \frac{\epsilon}{2}\right) \\
&\leq \frac{2}{\epsilon} E\left(\int_{M(1-\delta)}^{M(1+\delta)} \frac{|B(T(v)) - T(v)B(1)|}{v^r}\, dv\right) \\
&\leq \frac{2}{\epsilon}\int_{M(1-\delta)}^{M(1+\delta)} \frac{E\left(|B(T(v)) - T(v)B(1)|\right)}{v^r}\, dv\ .
\end{aligned}$$

Let $\mathcal{N}$ denote a standard normal random variable. Recall that

$$B(T(v)) \stackrel{d}{=} v^{-\alpha/2}\mathcal{N} \text{ and } E|\mathcal{N}| = \sqrt{2/\pi}\ .$$

So, it follows that

$$\begin{aligned}
&\lim_{M\to\infty} \int_{M(1-\delta)}^{M(1+\delta)} \frac{E\left(|B(T(v)) - T(v)B(1)|\right)}{v^r}\, dv \\
&\leq \lim_{M\to\infty} \sqrt{2/\pi} \int_{M(1-\delta)}^{M(1+\delta)} \left(v^{-(\alpha/2+r)} + v^{-(\alpha+r)}\right)\, dv = 0,
\end{aligned}$$

as long as $\alpha > 2(1-r)$. Thus, the term $\widetilde{A}_n^{M,1}$ is negligible.

2. Now, we establish negligibility of $\widetilde{A}_n^{M,2}$. By virtue of the decomposition (4.6),

$$\begin{aligned}
\widetilde{A}_n^{M,2} \leq\ & P\left(\left|\int_M^\infty \frac{M_n(v) - T\left(v/\rho_n\right)M_n(\rho_n)}{v^r\sqrt{k}}\, dv\right| \geq \frac{\epsilon}{4}\right) \\
&+ P\left(\left|\int_M^\infty \frac{L_n(v) - T\left(v/\rho_n\right)L_n(\rho_n)}{v^r\sqrt{k}}\, dv\right| \geq \frac{\epsilon}{4}\right).
\end{aligned}$$

- We deal with the martingale part first. To do so, set

$$\begin{aligned}
B_n^{M,1} &= P\left(\left|\int_M^\infty \frac{M_n(v)}{v^r\sqrt{k}}\, dv\right| \geq \frac{\epsilon}{8}\right)\ , \\
B_n^{M,2} &= P\left(\left|\frac{M_n(\rho_n)}{\sqrt{k}}\right| \rho_n^\alpha \int_M^\infty \frac{dv}{v^{\alpha+r}} \geq \frac{\epsilon}{8}\right)\ .
\end{aligned}$$

It follows that

$$P\left(\left|\int_M^\infty \frac{M_n(v) - T\left(v/\rho_n\right)M_n(\rho_n)}{v^r\sqrt{k}}\, dv\right| \geq \frac{\epsilon}{4}\right) \leq B_n^{M,1} + B_n^{M,2}\ .$$

Since $\{\Delta_j M_n\}$ is a stationary martingale difference sequence, then

$$B_n^{M,1} \le \frac{64}{\epsilon^2 k}\sum_{j=1}^{n} Var\left(\int_M^\infty \frac{\Delta_j M_n(v)}{v^r}\,dv\right) \le \frac{64n}{\epsilon^2 k} E\left[\left(\int_M^\infty \frac{\Delta_1 M_n(v)}{v^r}\,dv\right)^2\right]$$
$$\le \frac{64n}{\epsilon^2 k} E\left[\left(\int_M^\infty \frac{V_1(v)}{v^r}\,dv\right)^2\right] \le \frac{128n}{\epsilon^2 k} E\left[\left(\int_M^\infty \int_s^\infty \frac{V_1(t)}{t^r}\,dt\right)\frac{ds}{s^r}\right].$$

Notice that $E\left[V_1(t)\right] = E[\bar{F}_Z(u_n t/\phi(Y_1))]$. Furthermore, by (3.13) and (3.15),

$$\frac{\bar{F}_Z(u_n t/\phi(Y_1))}{\bar{F}_X(u_n)} \le \lambda C\left(t^{-\alpha+\epsilon}\phi^{\alpha-\epsilon}(Y_1) \vee t^{-\alpha-\epsilon}\phi^{\alpha+\epsilon}(Y_1)\right).$$

Therefore, since $M \ge 1$, we have

$$B_n^{M,1} \le \frac{128\lambda k C}{k\epsilon^2} E\left(\int_M^\infty \int_s^\infty \frac{t^{-\alpha+\epsilon}\phi^{\alpha-\epsilon}(Y_1) \vee t^{-\alpha-\epsilon}\phi^{\alpha+\epsilon}(Y_1)}{t^r}\,dt\right)\frac{ds}{s^r}$$
$$\le \frac{128\lambda C}{\epsilon^2} E\left(\phi^{\alpha-\epsilon}(Y_1) + \phi^{\alpha+\epsilon}(Y_1)\right)\left(\int_M^\infty \int_s^\infty \frac{t^{-\alpha+\epsilon}}{t^r}\,dt\right)\frac{ds}{s^r} = \frac{1}{\epsilon^2} O(M^{-\alpha+\epsilon-2r+2}).$$

As $M \to \infty$, the latter expression vanishes whenever $\alpha > 2(1-r)$ and $\epsilon < \alpha - 2(1-r)$.
Since $B_n^{M,1}$ decreases as ϵ increases, it is negligible for all $\epsilon > 0$.
We can omit ρ_n^α when dealing with $B_n^{M,2}$ since $\rho_n = 1 + o_P(1)$. We have

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\left|\frac{M_n(\rho_n)}{\sqrt{k}}\right| > (\alpha+r-1)M^{\alpha+r-1}\epsilon/8\right)$$
$$= \lim_{M\to\infty} P\left(|W(1)| > (\alpha+r-1)M^{\alpha+r-1}\epsilon/8\right) = 0,$$

since $\alpha > 2(1-r) > 0$. Therefore, the term $B_n^{M,2}$ is negligible.

In summary, the martingale part is negligible.

- We deal with the long memory part. Recall (4.42a)-(4.42c). Then

$$P\left(\left|\int_M^\infty \frac{L_n(v) - T\left(v/\rho_n\right)L_n(\rho_n)}{v^r\sqrt{k}}\,dv\right| \ge \frac{\epsilon}{4}\right)$$
$$\le \sum_{j=1}^{3} \underbrace{P\left(\bar{F}_Z(u_n)\int_M^\infty \frac{|L_{n,j}(v) - T\left(v/\rho_n\right)L_{n,j}(\rho_n)|}{v^r\sqrt{k}}\,dv \ge \frac{\epsilon}{12}\right)}_{I_{n,j}^M}.$$

Recalling from (4.63) that $|L_{n,2}(\rho_n t) - T(t)L_{n,2}(\rho_n)| = 0$, then $I_{n,2}^M = 0$. So, it remains only to deal with $I_{n,1}^M$ and $I_{n,3}^M$. By (4.62), we have

$$I_{n,1}^M \le P\left(\frac{C(\epsilon)n\bar{F}_Z(u_n)|\eta^*(u_n)|}{\sqrt{k}}\zeta_n \int_M^\infty \frac{\Lambda(v/\rho_n)}{v^r}\,dv \ge \frac{\epsilon}{12}\right),$$

where $\Lambda(t) = t^{-\alpha} + \left(t^{-(\kappa+\alpha+\epsilon)} \vee t^{-(\kappa+\alpha+\epsilon)}\right)$. We have

$$\begin{aligned}\int_M^\infty \frac{\Lambda(v/\rho_n)}{v^r}\, dv &\le \rho_n^\alpha \int_M^\infty v^{-(\alpha+r)}\, dv + \rho_n^{\kappa+\alpha+\epsilon} \int_M^\infty v^{-(\kappa+\alpha+\epsilon+r)}\, dv \\ &\quad + \rho_n^{\kappa+\alpha-\epsilon} \int_M^\infty v^{-(\kappa+\alpha-\epsilon+r)}\, dv \\ &= O_P(1) O(M^{-(\alpha+r)+\epsilon+1})\,.\end{aligned}$$

By ergodicity, $\zeta_n = O_P(1)$. Since $\sqrt{k}|\eta^*(u_n)| \to 0$, as $n \to \infty$, $n\bar{F}_X(u_n) = k$ and $\bar{F}_X(u_n) \sim E(\phi^\alpha(Y))\bar{F}_Z(u_n)$, as $n \to \infty$, we have that

$$\frac{C(\epsilon) n \bar{F}_Z(u_n)|\eta^*(u_n)|}{\sqrt{k}} \left(\int_M^\infty \frac{\Lambda(v/\rho_n)}{v^r}\, dv\right) \zeta_n \xrightarrow[n\to\infty]{p} 0.$$

This shows that $\lim_{n\to\infty} I_{n,1}^M = 0$, for all $M \ge 1$. On the other hand, by (4.64),

$$\begin{aligned}&\bar{F}_Z(u_n)\left|L_{n,3}(v) - T\left(v/\rho_n\right) L_{n,3}(\rho_n)\right| \\ &\quad \le \; nBC|\eta^*(u_n)|\bar{F}_Z(u_n) \left(\rho_n^{-(\kappa+\alpha+\epsilon)} \vee \rho_n^{-(\kappa+\alpha-\epsilon)}\right) \lambda\left(v/\rho_n\right).\end{aligned}$$

It follows that

$$\begin{aligned}&P\left(\bar{F}_Z(u_n) \int_M^\infty \frac{|L_{n,3}(v) - T\left(v/\rho_n\right) L_{n,3}(\rho_n)|}{v^r k} \ge \frac{\epsilon}{12}\right) \\ &\le \; P\left(\frac{CBn\bar{F}_Z(u_n)|\eta^*(u_n)|}{\sqrt{k}} \left(\rho_n^{-(\kappa+\alpha+\epsilon)} \vee \rho_n^{-(\kappa+\alpha-\epsilon)}\right) \int_M^\infty \frac{\Lambda(v/\rho_n)}{v^r}\, dv \ge \frac{\epsilon}{12}\right).\end{aligned}$$

Hence, the term is negligible by the same argument as for $I_{n,1}^M$. Thus, we get

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{L_n(v) - T\left(v/\rho_n\right) L_n(\rho_n)}{v^r\sqrt{k}}\, dv\right| \ge \frac{\epsilon}{12}\right) = 0\,.$$

Therefore, $\widetilde{A}_{n,M}^2$ is negligible.

This finishes the proof of (4.70). □

Lemma 4.3.12. *Under the conditions of Theorem 4.3.9, we have for each* $\epsilon \ge 0$,

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{T(t)\sqrt{k}\left(T_{u_n}(\rho_n) - T(\rho_n)\right)}{t^r}\, dt\right| \ge \epsilon\right) = 0. \tag{4.71}$$

Proof. Notice that $\sqrt{k}\left\{T_{u_n}(\rho_n) - T(\rho_n)\right\} = o_P(1)$ by (3.28b). Since $\int_M^\infty t^{-(\alpha+r)}dt < \infty$, then (4.71) holds. □

Lemma 4.3.13. *Under the conditions of Theorem 4.3.9, we have for all* $\epsilon \geq 0$*,*

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\left|\int_M^\infty \frac{\widehat{S}_{n,2}(t)}{t^r\sqrt{k}}\, dt\right| \geq \epsilon\right) = 0. \tag{4.72}$$

Proof. The definition of $\widehat{S}_{n,2}$ in (4.56), the change of variable $z = \rho_n t$ and (3.27b) yield

$$\begin{aligned}
\rho_n^{r-1}\sqrt{k}\int_{\rho_n M}^\infty \left|\frac{T_{u_n}(z) - T(z)}{z^r}\right| dz \quad &\leq \quad C\rho_n^{r-1}\sqrt{k}|\eta^*(u_n)| \int_{\rho_n M}^\infty \left(z^{-\alpha-\kappa-r+\epsilon} \vee z^{-\alpha-\kappa-r-\epsilon}\right) dv \\
&= \quad C\rho_n^{r-1}\sqrt{k}|\eta^*(u_n)| \frac{(\rho_n M)^{-(\alpha+\kappa+r-\epsilon)+1}}{\alpha+\kappa+r-\epsilon-1}.
\end{aligned}$$

Since $\alpha > 2(1-r)$, we also have $\alpha + \kappa + r - 1 > 0$. Therefore, the integral is negligible since $\rho_n \to 1$ in probability and $\sqrt{k}|\eta^*(u_n)| \to 0$, as $n \to \infty$. □

Remark 4.3.14. The limiting process in (4.65) is the same as in the i.i.d. case. The long memory and leverage effect do not influence the limiting behaviour of such integral functionals. Furthermore, we have

$$\int_1^\infty \frac{B\left(T(t)\right) - B(1)T(t)}{t^r}\, dt \stackrel{d}{=} \frac{\alpha^{1/2}\mathcal{N}}{(\alpha+r-1)(\alpha+2r-2)^{1/2}}\,, \tag{4.73}$$

where $\mathcal{N}$ is a standard normal random variable. It should be noted that

$$\begin{aligned}
&Cov\left(B(1), \int_1^\infty \frac{B(T(t)) - B(1)T(t)}{t^r}\, dt\right) \\
&= \quad E\left(B(1)\int_1^\infty \frac{B(T(t)) - B(1)T(t)}{t^r}\, dt\right) \\
&= \quad \int_1^\infty \frac{E\left(B(T(t))B(1)\right)}{t^r}\, dt - \int_1^\infty \frac{T(t)E\left(B^2(1)\right)}{t^r}\, dt \\
&= \quad \frac{1}{\alpha+r-1} - \frac{1}{\alpha+r-1} = 0,
\end{aligned} \tag{4.74}$$

where the last equality holds by virtue of (4.67).

4.3.3 Tail Index Estimation - Harmonic Moment Estimators

We consider the long memory stochastic volatility model with leverage defined in (3.1). Since the tail distribution of X is regularly varying with index $-\alpha$, then this raises the question of estimating the index of regular variation α. To answer this question, we restrict our attention to the *Harmonic Moment Estimators (HME)* $\widehat{\gamma}_{r,k}$ of order r

of $\gamma := 1/\alpha$. This class of estimators were studied in [7]. We aim at studying their asymptotic normality. We get started with the construction of such estimators. It follows from (3.10) that for all $r \geq 0$,

$$\zeta_r := \int_1^\infty \frac{T(t)}{t^r}\, dt = \frac{1}{\alpha + r - 1}\cdot \tag{4.75}$$

If $\widehat{\zeta}_{r,k}$ denotes an estimator of ζ_r, then the plug-in method together with (4.48) yield

$$\widehat{\zeta}_{r,k} = \int_1^\infty \frac{\widehat{T}_n(t)}{t^r}\, dt = \int_1^\infty \frac{1}{k}\sum_{j=1}^n \mathbb{1}_{\{X_{(j)} > X_{(n-k)} t\}} \frac{dt}{t^r} = \frac{1}{k}\sum_{j=1}^n \int_1^\infty \mathbb{1}_{\left\{\frac{X_{(j)}}{X_{(n-k)}} > t\right\}} \frac{dt}{t^r}\cdot$$

Furthermore, since $t \geq 1$, then we have

$$\begin{aligned}
\widehat{\zeta}_{r,k} &= \frac{1}{k}\sum_{j=1}^k \int_1^{\frac{X_{(n-j+1)}}{X_{(n-k)}}} \frac{dt}{t^r} \\
&= \begin{cases} \frac{1}{r-1}\left[1 - \frac{1}{k}\sum_{j=1}^k \left(\frac{X_{(n-k)}}{X_{(n-j+1)}}\right)^{r-1}\right] & \text{if } r \neq 1, \\ \frac{1}{k}\sum_{j=1}^k \ln\left(\frac{X_{(n-j+1)}}{X_{(n-k)}}\right) & \text{if } r = 1. \end{cases}
\end{aligned}$$

To derive the estimators of $\gamma = 1/\alpha$, we solve for $1/\alpha$ in (4.75) and obtain

$$\zeta_r = \frac{1}{\alpha + r - 1} \Rightarrow \frac{1}{\alpha} = \frac{\zeta_r}{1 + (1-r)\zeta_r}\cdot$$

Thanks to the plug-in method, we derive the HMEs below:

$$\widehat{\gamma}_{r,k} = \frac{\widehat{\zeta}_{r,k}}{1 + (1-r)\widehat{\zeta}_{r,k}} \tag{4.76}$$

$$= \begin{cases} \frac{1}{r-1}\left[\left(\frac{1}{k}\sum_{j=1}^k \left(\frac{X_{(n-k)}}{X_{(n-j+1)}}\right)^{r-1}\right)^{-1} - 1\right] & \text{if } r \neq 1\ , \\ \frac{1}{k}\sum_{j=1}^k \ln\left(\frac{X_{(n-j+1)}}{X_{(n-k)}}\right) & \text{if } r = 1\ . \end{cases} \tag{4.77}$$

- The HME that corresponds to $r = 1$ is the Hill estimator of $\gamma = 1/\alpha$.
- The HME that corresponds to $r = 2$ is the t-Hill estimator of γ, that is

$$\widehat{\gamma}_{2,k} = \left(\frac{1}{k}\sum_{j=1}^k \frac{X_{(n-k)}}{X_{(n-j+1)}}\right)^{-1} - 1\ . \tag{4.78}$$

From now on, we are interested in studying the limiting behaviour of these Harmonic Moment Estimators. The next result will serve as a building block for the asymptotic normality of $\widehat{\gamma}_{r,k}$.

Theorem 4.3.15. *Under the assumptions of Theorem 4.3.9,*

$$\sqrt{k}\left(\widehat{\zeta}_{r,k}-\zeta_r\right)\xrightarrow[n\to\infty]{d}\int_1^{\infty}\frac{B\left(T(t)\right)-B(1)T(t)}{t^r}\,dt. \tag{4.79}$$

Proof. We observe that

$$\sqrt{k}\left(\widehat{\zeta}_{r,k}-\zeta_r\right)=\sqrt{k}\left(\int_1^{\infty}\frac{\widehat{T}_n(t)}{t^r}\,dt-\int_1^{\infty}\frac{T(t)}{t^r}\,dt\right)=\frac{1}{\sqrt{k}}\int_1^{\infty}\frac{\widehat{S}_n(t)}{t^r}\,dt\ ,$$

As a consequence, (4.79) follows from Theorem 4.3.9. □

We are now ready to deal with asymptotic normality of the Harmonic Moment Estimators of order r of $1/\alpha$. The next result provides a unified approach to central limit theorems for estimators of the tail index α.

Theorem 4.3.16. *Under the assumptions of Theorem 4.3.9 and if $\alpha>2(1-r)$,*

$$\sqrt{k}\left(\widehat{\gamma}_{r,k}-\gamma\right)\xrightarrow[n\to\infty]{d}\frac{(\alpha+r-1)}{\sqrt{\alpha^3(\alpha+2r-2)}}\mathcal{N}, \tag{4.80}$$

where $\mathcal{N}$ is a standard random variable.

Proof. We observe from (4.76) that $\widehat{\gamma}_{r,k}=g\left(\widehat{\zeta}_{r,k}\right)$, where g is defined by

$$g(x)=\frac{x}{1+(1-r)x}\cdot$$

Therefore (4.80) follows from (4.79) in conjunction with the δ-method. □

Remark 4.3.17. The striking fact about (4.80) is that the asymptotic behaviour of Harmonic Moment Estimators is unaffected either by long memory or leverage.

- the Hill and t-Hill estimators are asymptotically normal (AN), that is

$$\widehat{\gamma}_{1,k}\sim AN\left(\gamma,\frac{1}{k\alpha^2}\right)\text{ and }\widehat{\gamma}_{2,k}\sim AN\left(\gamma,\frac{(\alpha+1)^2}{k\alpha^3\left(\alpha+2\right)}\right), \tag{4.81}$$

- The limiting variance of the HMEs $\widehat{\gamma}_{r,k}$ is

$$v^2(r) = \frac{(\alpha + r - 1)^2}{\alpha^3(\alpha + 2r - 2)}.$$

Consequently, if $\alpha > 2(1-r)$, then the minimal limiting variance of $\widehat{\gamma}_{r,k}$: is attained at $r = 1$. In fact,

$$\frac{dv^2(r)}{dr} = \frac{2(\alpha + r - 1)(r - 1)}{\alpha^3(\alpha + 2r - 2)^2} \text{ and } \frac{d^2 v_r^2}{dr^2} \;|_{r=1} = \frac{2}{\alpha^{10}} > 0.$$

4.4 Concluding Remarks

In this chapter, we have considered the heavy-tailed long memory stochastic volatility model with leverage given in (3.1). We have studied the limiting behaviour of the tail empirical process with both fixed and random levels (Theorems 4.2.18 and 4.3.4). We have shown a dichotomous behaviour for the tail empirical process with fixed levels, according to the interplay between the long memory parameter d and the tail index α; leverage does not play a role in the limiting results, but makes proofs technically involved. On the other hand, the tail empirical process with random levels is unaffected by either long memory or leverage. Further, we have proven the weak convergence of integral functionals (Theorem 4.3.9). The tail empirical process with random levels is used to construct a family of estimators of the tail index, including the famous Hill estimator and harmonic mean estimators. Consequently, all HMEs of the tail index of $\{X_j\}$ remain valid for this model and have the same asymptotic behaviour as in the case of i.i.d. observations (Theorem 4.3.16).

In the next chapter, we consider the asymptotic behaviour of estimators of risk measures in the context of the long memory stochastic volatility models with leverage.

Chapter 5

Estimation of Financial Risk Measures

5.1 Introductory Comments

This chapter deals with statistics of financial risk of a portfolio. By financial risk, we refer to the prospect of financial loss. Financial institutions are exposed to various forms of financial risk during the course of their transactions. Here are some common types of financial risk:

- market risk (the risk of loss or gain arising from unexpected changes in market prices, such as security prices) or market rates (e.g., such as interest or exchange rates). The existing forms of market risk are discussed in detail in [19],

- credit risk (the risk of loss arising from the failure of a counterparty to make a promised payment),

- operational risk (the risk of loss arising from the failures of internal systems (accidents) or the people who operate in them (fraud, ethics, etc)).

- liquidity risk (the risk that assets cannot be sold or bought as and when required).

Financial risk management is essential in order to make these institutions resilient to future events with adverse effects. It is mainly concerned with quantifying market risk (measurement or estimation) and decision making. Three approaches are usually considered when it comes to financial risk measurement or estimation.

The first approach stems from portfolio-theory, in which an investor relies heavily on the performance of the expected return and the magnitude of the standard deviation. In this mean-variance approach, returns are assumed to be normally distributed. It turns out that the standard deviation of a return is interpreted as a financial risk measure. The higher the variance, the riskier the asset. However, this measure of risk is of limited use because it captures only volatility and fails to quantify likely losses of financial institutions. The normality assumption is also very questionable.

The second approach is based on Value-at-Risk (VaR), a measure of financial risk that captures not only the volatility of assets of a portfolio but also the maximum of the likely losses of a portfolio. Unlike the mean-variance approach, returns are assumed to follow an arbitrary distribution. VaR has a number of serious drawbacks. It captures only risks that are quantifiable. It fails for instance to capture either operational risks or liquidity risks. In the presence of very non-normal distributions, VaR turns out not to be a reliable (and perhaps not even useful) risk measure. We refer to [27] and [1] for a detailed discussion about the limitations of Value-at-Risk.

The third approach is based on coherent risk measures, which capture the size of a potential loss of a financial institution (see Definition 5.2.1). These financial risk measures are alternatives to Value-at-Risk and were introduced in [4] in the late nineties. See also [1]. One such coherent risk measure is the so-called Expected Shortfall (cf. [1])

Our goal in this chapter is to estimate financial risk measures associated with the long memory stochastic volatility model with leverage under the second and third approaches only. To do so, we use the tail empirical process. We use the results of Chapter 4 to investigate the asymptotic behaviour of the estimators of Value-at-Risk and Expected Shortfall.

This chapter is organized as follows. Coherent risk measures are introduced in section 5.2. In section 5.3, we give an overview of Value-at-Risk and study the asymptotic behaviour of estimators of VaR. In section 5.4, we present Expected Shortfall, a natural coherent alternative to VaR. We study the asymptotic behaviours of its estimators.

5.2 Coherent Financial Risk Measures

In this section we present the defining properties of coherent financial risk measures. These properties are known as axioms of coherence and were introduced in [4].

Definition 5.2.1. *[4]*
Let $\mathcal{L}$ be the set of all real valued random variables (future values of portfolios). A financial risk measure $\varrho : \mathcal{L} \to \mathbb{R}$ is said to be coherent if it is

C(i) *normalized, that is $\varrho(0) = 0$.*

C(ii) *monotonically decreasing, that is, for all $L_1,\ L_2 \in \mathcal{L}$,*

$$L_1 \leq L_2 \Rightarrow \varrho\left(L_2\right) \leq \varrho\left(L_1\right).$$

C(iii) *subadditive, that is for all $L_1,\ L_2 \in \mathcal{L}$,*

$$\varrho\left(L_1 + L_2\right) \leq \varrho\left(L_1\right) + \varrho\left(L_2\right).$$

C(iv) *positively homogeneous, that is for all $L \in \mathcal{L}$, $a > 0$,*

$$\varrho\left(aL\right) = a\varrho(L).$$

C(v) *transitionally invariant that is, for all $L \in \mathcal{L}$, $b \in \mathbb{R}$,*

$$\varrho\left(L + b\right) = \varrho\left(L\right) - b.$$

Remark 5.2.2. What do these axioms of coherence mean concretely? The rationale behind C(i) is that the risk of holding no assets is zero. The rationale behind C(ii) is that a portfolio with greater future returns has less risk. The meaning of C(iii) is that the aggregate portfolio risk is less than or equal to the sum of the individual risks of the assets that compose this portfolio. Positive homogeneity C(iv) implies that the risk of a position is proportional to its size. In financial risk management, the implication of axiom C(v) is that addition of a sure amount of capital reduces the risk by the same amount.

Thanks to these axioms of coherence, financial risks can be effectively regulated and managed. For instance, when it comes to decision making, coherent financial risk measures are more reliable than the other existing traditional financial risk measures (standard deviation, Value-at-Risk). Note that subadditivity reflects an expectation that when we aggregate individual risks, they diversify or, at worst, do not increase. It means that aggregating risks does not increase overall risk [27].

5.3 Value-at-Risk (VaR)

In this section, we briefly present Value-at-Risk, derive its estimators and establish their limit theorems. This statistic represents the maximum likely loss over some target period - the most we expect to lose over that period, at a specified probability level. The different methodologies to compute VaR are discussed in [19].

Definition 5.3.1. *[1],[27]*
Let L be a continuous random variable with a distribution function F_L and $p \in (0,1)$. The Value-at-Risk (or more precisely, p-VaR) is defined to be

$$\mathrm{VaR}_p(L) = F_L^{\leftarrow}(1-p) = Q_L(1/p) \,, \tag{5.1}$$

where $F_L^{\leftarrow}$ is the left-continuous inverse of F_L and $Q_L(t) = F_L^{\leftarrow}(1-1/t)\,, \quad t>1$.

5.3.1 Estimation of VaR

Value-at-Risk fails to be a coherent risk measure. In fact, VaR is not subadditive [1]. Why is its estimation still so important? In fact, the asymptotic behaviours of many estimators of coherent risk measures depend on those of VaR. This is illustrated in Section 5.4.1, where estimation of expected shortfall depends on estimation of VaR.

Let $(X_1,\ldots,X_n)$ be a sample from the long memory stochastic volatility model with leverage considered in (3.1). Let $X_{(1)} \le \cdots \le X_{(i)} \le \cdots \le X_{(n)}$ be the corresponding order statistics. Recall that F_X denotes the common distribution function of the X_i's.

Our goal is to estimate $Q_X(1/p)$ when p is very small. Note that

$$E\left(\sum_{j=1}^{n} \mathbb{1}_{\{X_j > Q_X(1/p)\}}\right) = np.$$

This means that the expected number of observations above $Q_X(1/p)$, np, is also very small. We assume that p depends on n and $p=p_n \to 0$, as $n\to\infty$. If F_n denotes the empirical distribution function of the sample $(X_1,\ldots,X_n)$, then the empirical estimate of $Q_X(1/p)$ is

$$F_n^{\leftarrow}(1-p) = X_{(n-[np])} \,. \tag{5.2}$$

However, for a very small value of p, this procedure is not very reliable since for $p_1 \neq p_2$ such that $[np_1] = [np_2]$, the values $Q_X(1/p_1)$ and $Q_X(1/p_2)$ may differ significantly, but both values will be estimated by the same $X_{(n-[np_1])}$. In particular, for all $p < 1/n$, $Q_X(1/p)$ will be always estimated by $X_{(n)}$.

To address this, let k be an intermediate sequence, that is $k = k_n \to \infty$ and $k/n \to 0$. Notice that if $\bar{F}_X(x) = x^{-\alpha}$, then $F_X^{\leftarrow}(1-p) = p^{-\frac{1}{\alpha}}$. So recalling that $\gamma = 1/\alpha$, we have

$$Q_X(1/p) = p^{-\gamma} \text{ and } Q_X(n/k) = (k/n)^{-\gamma} \ .$$

Therefore, we obtain

$$\frac{Q_X(1/p)}{Q_X(n/k)} = \left(\frac{k}{np}\right)^{\gamma} \ .$$

Since the tail distribution of X is regularly varying with index $-\alpha$ (cf. (3.17)), then

$$\frac{Q_X(1/p)}{Q_X(n/k)} = \left(\frac{k}{np}\right)^{\gamma}(1+o(1)) \tag{5.3}$$

as long as (k/np) is bounded away from 0 and ∞ or $k/(np) \to \infty$. The latter case is what we are really interested in. By (5.2), $Q_X(n/k)$ can be estimated by $X_{(n-k)}$. This suggests the following estimators of $Q_X(1/p)$:

$$\widetilde{Q}_X(1/p) = X_{(n-k)}\left(\frac{k}{np}\right)^{\gamma}, \quad \text{if } \alpha \text{ known} \tag{5.4a}$$

$$\widehat{Q}_X(1/p) = X_{(n-k)}\left(\frac{k}{np}\right)^{\widehat{\gamma}}, \quad \text{otherwise.} \tag{5.4b}$$

Note that $\widehat{\gamma}$ denotes an estimator of $\gamma = 1/\alpha$. To study the asymptotic behaviour of VaR, limit theorems for the intermediate order statistics $X_{(n-k)}$ and those for estimators of the tail index α are required. Recall that these limit theorems were already studied in Lemma 4.3.3 and Theorem 4.3.16, respectively.

5.3.2 Limit Theorems for VaR

In light of these estimators of Value-at-Risk, we are in position to investigate throughout this subsection their limiting behaviours under the assumptions of the long memory stochastic volatility model with leverage studied in Chapters 2 and 3. We start in Proposition 5.3.2 with the class of estimators (5.4a) and then wrap up in Theorems 5.3.3 and 5.3.5 with the class of estimators (5.4b).

Proposition 5.3.2. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then we have the following:*

- *If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ or $m(1-2d)>1$, then*

$$\sqrt{k}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}\xrightarrow[n\to\infty]{d}\frac{B(1)}{\alpha}. \tag{5.5}$$

- *If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, then*

$$\frac{n}{b_{n,m}}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}\xrightarrow[n\to\infty]{d}\frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m!E(\phi^{\alpha}(Y))}. \tag{5.6}$$

Proof. Since $k=n\bar{F}_X(u_n)$, then $u_n=F_X^{\leftarrow}(1-k/n)=Q_X(n/k)$. Thus,

$$\begin{aligned}
\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\} &= \left\{\frac{X_{(n-k)}}{u_n}\left(\frac{k}{np}\right)^{\gamma}\frac{Q_X(n/k)}{Q_X(1/p)}-1\right\}\\
&= \left(\frac{k}{np}\right)^{\gamma}\frac{Q_X(n/k)}{Q_X(1/p)}\left\{\frac{X_{(n-k)}}{u_n}-1\right\}+\left\{\left(\frac{k}{np}\right)^{\gamma}\frac{Q_X(n/k)}{Q_X(1/p)}-1\right\}\\
&= \left(\frac{k}{np}\right)^{\gamma}\frac{Q_X(n/k)}{Q_X(1/p)}\left\{\left(\frac{X_{(n-k)}}{u_n}-1\right)-\left(\left(\frac{k}{np}\right)^{-\gamma}\frac{Q_X(1/p)}{Q_X(n/k)}-1\right)\right\}.
\end{aligned} \tag{5.7}$$

The result follows by Lemma 4.3.3, Lemma 3.4.1 and Lemma 3.4.2.

□

We continue with the situation where α is unknown and investigate the asymptotic behaviour of the Value-at-Risk estimator $\widehat{Q}_X(1/p)$. In what follows, we replace the unknown value of $\gamma=1/\alpha$ with the harmonic moment estimator $\widehat{\gamma}_{r,k}$ of order r, as defined in (4.76).

In the following theorem, we consider two limiting schemes, when $k/(np)\to\infty$ (relevant in practice) and $k/(np)\to\nu\in(0,\infty)$. The latter case is not relevant in practice, but may explain the results of some finite sample simulation studies.

Theorem 5.3.3. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Let $\mathcal{N}$ denote a standard normal variable.*

1. *Assume that $k/np\to\nu$, $0<\nu<\infty$, as $n\to\infty$. If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ or $m(1-2d)>1$, then*

$$\sqrt{k}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\}\xrightarrow[n\to\infty]{d}\mathcal{N}\sqrt{\frac{1}{\alpha^2}+\frac{(\alpha+r-1)^2\ln^2\nu}{\alpha^3(\alpha+2r-2)}}\,. \tag{5.8}$$

2. *Assume that* $k/np \to \nu$, $0 < \nu < \infty$, *as* $n \to \infty$. *If* $m(1-2d) < 1$ *and* $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, *then*

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m! E(\phi^\alpha(Y))}. \tag{5.9}$$

3. *Assume that* $k/np \to \infty$, *as* $n \to \infty$ *and moreover if* $m(1-2d) < 1$ *and* $\frac{b_{n,m}}{n}\sqrt{k} \to 0$ *or* $m(1-2d) > 1$, *then*

$$\frac{\sqrt{k}}{\ln(k/np)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{(\alpha + r - 1)}{\sqrt{\alpha^3(\alpha + 2r - 2)}}\mathcal{N}. \tag{5.10}$$

4. *Assume that* $k/np \to \infty$, *as* $n \to \infty$ *If* $m(1-2d) < 1$ *and* $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, *then*

$$\frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \xrightarrow[n\to\infty]{d} 0. \tag{5.11}$$

Proof. In what follows, for simplicity we write $\widehat{\gamma}$ for $\widehat{\gamma}_{r,k}$. Further, notice that

$$\frac{n}{b_{n,m}\sqrt{k}} = \frac{1}{b_{n,m}}\sqrt{\frac{n}{\bar{F}_X(u_n)}}\,. \tag{5.12}$$

Write the following decomposition

$$\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1 \;=\; \left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} + \frac{\widehat{Q}_X(1/p) - \widetilde{Q}_X(1/p)}{Q_X(1/p)}.$$

Bearing in mind $u_n = Q_X(n/k)$, the first-order Taylor expansion applied to $\widehat{\gamma} \mapsto (k/np)^{\widehat{\gamma}}$ around γ yields

$$\begin{aligned}
\frac{\widehat{Q}_X(1/p) - \widetilde{Q}_X(1/p)}{Q_X(1/p)} &= \frac{X_{(n-k)}}{u_n}\frac{Q_X(n/k)}{Q_X(1/p)}\left\{\left(\frac{k}{np}\right)^{\widehat{\gamma}} - \left(\frac{k}{np}\right)^{\gamma}\right\} \\
&= \frac{X_{(n-k)}}{u_n}\frac{Q_X(n/k)}{Q_X(1/p)}\left\{\left(\frac{k}{np}\right)^{\gamma}\ln\left(\frac{k}{np}\right)(\widehat{\gamma} - \gamma) + R_n\right\} \\
&= C_n\left\{\ln\left(\frac{k}{np}\right)(\widehat{\gamma} - \gamma) + \left(\frac{k}{np}\right)^{-\gamma} R_n\right\},
\end{aligned}$$

where R_n is the Lagrange form of the remainder term, that is

$$R_n = \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}}\ln^2\left(\frac{k}{np}\right)(\widehat{\gamma} - \gamma)^2,$$

for some random variable $\tilde{\gamma}$ such that $|\tilde{\gamma} - \gamma| \leq |\widehat{\gamma} - \gamma|$ and C_n is defined as follows:

$$C_n := \frac{X_{(n-k)}}{u_n} \left(\frac{k}{np}\right)^{\gamma} \frac{Q_X(n/k)}{Q_X(1/p)}.$$

In light of Lemma 4.3.3 and (5.3), $C_n \xrightarrow[n\to\infty]{p} 1$. Therefore, we have

$$\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1 = \left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} + C_n \left\{\ln\left(\frac{k}{np}\right)(\widehat{\gamma} - \gamma) + \left(\frac{k}{np}\right)^{-\gamma} R_n\right\}.$$

1. Assume that $k/np \to \nu$, $0 < \nu < \infty$, $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to 0$ or $m(1-2d) > 1$. We have

$$\sqrt{k}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} = \sqrt{k}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} + C_n \ln\left(\frac{k}{np}\right)\sqrt{k}\,(\widehat{\gamma} - \gamma) + \left(\frac{k}{np}\right)^{-\gamma} C_n \sqrt{k} R_n. \tag{5.13}$$

Since $\sqrt{k}\,(\widehat{\gamma} - \gamma) = O_p(1)$, by (4.80), then

$$\left(\frac{k}{np}\right)^{-\gamma} \sqrt{k} R_n = \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma} \frac{\ln^2(k/np)}{\sqrt{k}} \left\{\sqrt{k}\,(\widehat{\gamma} - \gamma)\right\}^2 = o_p(1). \tag{5.14}$$

As noted above, $C_n \xrightarrow{\mathrm{p}} 1$. Thus, using (5.5) and (4.80), we conclude that

$$\sqrt{k}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{B(1)}{\alpha} + \frac{(\alpha + r - 1)\ln\nu}{\sqrt{\alpha^3(\alpha + 2r - 2)}}\mathcal{N}.$$

Recalling that $B(1)$ is a standard normal random variable and taking into account (4.74) and the joint convergence, we have (5.8) (in other words, estimators of the tail index and order statistics are asymptotically uncorrelated).

2. Assume that $k/np \to \nu$, $0 < \nu < \infty$, as $n \to \infty$. If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, then

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} = \frac{n}{b_{n,m}}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} + \ln\left(\frac{k}{np}\right) C_n \frac{n}{b_{n,m}\sqrt{k}}\sqrt{k}\,(\widehat{\gamma} - \gamma) + C_n \left(\frac{k}{np}\right)^{-\gamma} \frac{n}{b_{n,m}} R_n.$$

Since $\sqrt{k}\,(\widehat{\gamma}-\gamma)=O_p(1)$, by (4.80), then

$$\begin{aligned}\left(\frac{k}{np}\right)^{-\gamma}\frac{n}{b_{n,m}}R_n &= \frac{1}{2k}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma}\frac{n}{kb_{n,m}}\ln^2\left(\frac{k}{np}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2\\ &= \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma}\left(\frac{n}{b_{n,m}\sqrt{k}}\right)\left(\frac{\ln^2(k/np)}{\sqrt{k}}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2\\ &= o_p(1).\end{aligned}$$

We already know that $C_n\overset{\mathrm{P}}{\to}1$. Taking into account (5.6), (5.9) follows.

3. Now, assume that $k/np\to\infty$, as $n\to\infty$. If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ or $m(1-2d)>1$, then

$$\begin{aligned}\frac{\sqrt{k}}{\ln(k/np)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\} &= \underbrace{\frac{1}{\ln(k/np)}\sqrt{k}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}}_{=:I_n}\\ &+C_n\sqrt{k}\,(\widehat{\gamma}-\gamma)+C_n\left(\frac{k}{np}\right)^{-\gamma}\frac{\sqrt{k}}{\ln(k/np)}R_n\,.\end{aligned}$$

By (5.5) the term I_n converges in probability to zero. We already know that $C_n\overset{\mathrm{P}}{\to}1$. Since $\sqrt{k}\,(\widehat{\gamma}-\gamma)=O_p(1)$ (by (4.80)), we have

$$(k/np)^{\tilde{\gamma}-\gamma}=e^{\sqrt{k}(\tilde{\gamma}-\gamma)\frac{\ln(k/np)}{\sqrt{k}}}=O_p(1), \tag{5.15}$$

and we obtain

$$\begin{aligned}\left(\frac{k}{np}\right)^{-\gamma}\frac{\sqrt{k}}{\ln(k/np)}R_n &= \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma}\frac{\sqrt{k}}{\ln(k/np)}\ln^2(k/np)\,(\widehat{\gamma}-\gamma)^2\\ &= \frac{1}{2}O_P(1)\left(\frac{\ln(k/np)}{\sqrt{k}}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2\\ &= o_p(1).\end{aligned}$$

By virtue of (4.80), we get (5.10).

4. Assume that $k/np\to\infty$, as $n\to\infty$. If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, then

$$\begin{aligned}\frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\} &= \underbrace{\frac{1}{\ln(k/np)}\frac{n}{b_{n,m}}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}}_{=:J_n}\\ &+C_n\frac{n}{b_{n,m}\sqrt{k}}\sqrt{k}\,(\widehat{\gamma}-\gamma)+C_n\left(\frac{k}{np}\right)^{-\gamma}\frac{n}{b_{n,m}\ln(k/np)}R_n.\end{aligned}$$

Again, as before $J_n = o_P(1)$. Since $\sqrt{k}\,(\widehat{\gamma} - \gamma) = O_p(1)$, then by (5.15),

$$\begin{aligned}\left(\frac{k}{np}\right)^{-\gamma} \frac{n}{b_{n,m}\ln(k/np)} R_n &= \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma} \frac{n}{b_{n,m}\ln(k/np)} \ln^2(k/np)\,(\widehat{\gamma}-\gamma)^2 \\ &= \frac{1}{2} e^{\sqrt{k}(\tilde{\gamma}-\gamma)\ln(k/np)/\sqrt{k}} \left(\frac{n}{b_{n,m}\sqrt{k}}\right)\left(\frac{\ln(k/np)}{\sqrt{k}}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2 = o_p(1).\end{aligned}$$

As a consequence, assuming that Lemma 3.4.1, (5.6) and (4.80) hold, we obtain (5.11).

□

Remark 5.3.4. Generally speaking, when returns are assumed to exhibit heavy-tails, long range dependence and leverage, estimation of VaR is influenced in a dichotomous way either by heavy-tails or long memory. The striking fact about this estimation is that the leverage effect does not contribute at all. This is demonstrated in the following scenarios:

1. In case of (5.8) the limiting behaviour is affected by both order statistics and estimation of the tail index.
2. In case of (5.9) the limiting behaviour is affected by limiting behaviour of order statistics only. Estimation of the tail index does not play any role.
3. In case of (5.10) the limiting behaviour is affected by estimation of the tail index only.
4. In case of (5.11) we end up with a degenerate limit. This is addressed in the following result by imposing more detailed conditions.

Theorem 5.3.5. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Assume that $k/np \to \infty$. If $m(1-2d) < 1$ and*

$$\frac{b_{n,m}}{n}\sqrt{k}/\ln(k/np) \to \infty, \tag{5.16}$$

then

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m! E(\phi^{\alpha}(Y))}. \tag{5.17}$$

Proof. We have

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\}=\frac{n}{b_{n,m}}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}$$
$$+\ln\left(\frac{k}{np}\right)C_n\frac{n}{b_{n,m}\sqrt{k}}\sqrt{k}\,(\widehat{\gamma}-\gamma)+C_n\left(\frac{k}{np}\right)^{-\gamma}\frac{n}{b_{n,m}}R_n.$$

We already know that $C_n\overset{\mathrm{p}}{\to}1$. Since $\sqrt{k}\,(\widehat{\gamma}-\gamma)=O_p(1)$ by (4.80), then under assumption (5.16), the second term on the right hand side is $o_p(1)$. Finally, we have

$$\begin{aligned}\left(\frac{k}{np}\right)^{-\gamma}\frac{n}{b_{n,m}}R_n &= \frac{1}{2k}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma}\frac{n}{b_{n,m}}\ln^2\left(\frac{k}{np}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2\\ &= \frac{1}{2}\left(\frac{k}{np}\right)^{\tilde{\gamma}-\gamma}\left(\frac{n}{b_{n,m}\sqrt{k}}\right)\left(\frac{\ln^2(k/np)}{\sqrt{k}}\right)\left\{\sqrt{k}\,(\widehat{\gamma}-\gamma)\right\}^2\\ &= o_p(1).\end{aligned}$$

Taking into account (5.6), (5.9) follows. □

Remark 5.3.6. We note that (5.16) automatically implies $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$. However, we do not know what happens if

$$\frac{b_{n,m}}{n}\sqrt{k}\to\infty \text{ but } \frac{b_{n,m}}{n}\sqrt{k}/\ln(k/np)\to 0.$$

We now turn our attention to Expected Shortfall.

5.4 Expected Shortfall (ES)

As previously discussed, a number of deficiencies make VaR unsatisfactory as a financial risk measure. Since VaR is simply a quantile, it lacks subadditivity, a serious limitation. To get around the limitations of VaR, in [4] the authors introduced coherent measures as alternatives to VaR. It turns out that Expected Shortfall (ES) is the natural coherent alternative to VaR [1]. The coherence of expected shortfall is extensively discussed in [2]. ES can be simply viewed as the excess mean function - that is the average value of all values exceeding a certain threshold of the VaR. The ES is also called the Conditional Tail Expectation (CTE) when the distribution of returns is continuous.

Our goal in this section is to establish limit theorems for estimators of ES. Before doing so, we define Expected Shortfall and derive its estimators.

Definition 5.4.1. *[1]*
Let L be the profit-loss of a portfolio on a specified time horizon and let $p \in (0,1)$ be some specified probability level. The expected shortfall of the portfolio is then defined as

$$ES_p(L) = E\left(L \| L > Q_L(1/p)\right) =: \theta_L(p). \tag{5.18}$$

5.4.1 Estimation of Expected Shortfall

Recall that $\gamma = 1/\alpha$. Notice that if L is standard Pareto distributed and $\alpha > 1$, then

$$\theta_L(p) = E\left(L \| L > Q_L(1/p)\right) = \frac{1}{(1-\gamma)p^\gamma}. \tag{5.19}$$

Lemma 5.4.2. *Let L be a random variable whose tail distribution $\bar{F}_L$ is regularly varying at infinity with index $-\alpha$ where $\alpha > 1$. Then as $p \to 0$,*

$$\theta_L(p) \sim \frac{1}{1-\gamma} Q_L(1/p). \tag{5.20}$$

Proof. By definition, we have

$$\theta_L(p) = \frac{E\left(L \mathbb{1}_{\{L > Q_L(1/p)\}}\right)}{\bar{F}_L\left(Q_L(1/p)\right)} = -\frac{\displaystyle\int_{Q_L(1/p)}^{\infty} u \bar{F}_L(du)}{\bar{F}_L\left(Q_L(1/p)\right)}.$$

By integrating by parts the right hand side and applying (2.17), we obtain

$$\begin{aligned} E\left(L \mathbb{1}_{\{L > Q_L(1/p)\}}\right) &= -\left[u\bar{F}_L(u)\right]_{Q_L(1/p)}^{\infty} + \int_{Q_L(1/p)}^{\infty} u^{-\alpha} \ell_L(u)\, du \\ &\underset{Q_L(1/p)\to\infty}{\sim} Q_L(1/p)\bar{F}_L\left(Q_L(1/p)\right) + \frac{Q_L^{1-\alpha}(1/p)}{\alpha - 1} \ell_L\left(Q_L(1/p)\right), \end{aligned}$$

by Theorem 2.3.8. Notice that $Q_L(1/p) \to \infty$ if and only if $p \to 0$. Therefore,

$$\begin{aligned} E\left(L \| L > Q_L(1/p)\right) &\underset{p\to 0}{\sim} Q_L(1/p) + \frac{\left(Q_L(1/p)\right)^{1-\alpha} \ell_L\left(Q_L(1/p)\right)}{(\alpha-1)\bar{F}_L\left(Q_L(1/p)\right)} \\ &\underset{p\to 0}{\sim} Q_L(1/p) + \frac{Q_L(1/p)}{\alpha - 1} \\ &\underset{p\to 0}{\sim} \frac{\alpha}{\alpha - 1} Q_L(1/p). \end{aligned}$$

□

Remark 5.4.3. It follows from Lemma 5.4.2 that estimators of Expected Shortfall can be defined as:

$$\tilde{\theta}_L(p) = \frac{1}{1-\gamma}\widetilde{Q}_L(1/p), \quad \text{if } \alpha \text{ known} \tag{5.21a}$$

$$\widehat{\theta}_L(p) = \frac{1}{1-\widehat{\gamma}}\widehat{Q}_L(1/p), \quad \text{otherwise.} \tag{5.21b}$$

The bottom line is that estimation of Expected Shortfall relies on estimation of VaR. Notice that Expected Shortfall blows up for values of $\alpha \leq 1$. Recall that it is exactly under these circumstances that the mean of a Pareto distributed random variable fails to exist. That is why this scenario is excluded in Lemma 5.4.2 and it must be assumed that the tail index $\alpha > 1$ when dealing with estimation of Expected Shortfall.

5.4.2 Limit Theorems for Expected Shortfall

In this subsection, we investigate the limiting behaviours of estimators of Expected Shortfall. This is done in Proposition 5.4.5 and Theorems 5.4.6 and 5.4.8, respectively. We get started with Lemma 5.4.4, an ingredient for proofs of the above results.

Lemma 5.4.4. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then if $\alpha > 1$,*

$$\lim_{n\to\infty} a_{n,m}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\} = 0\ . \tag{5.22}$$

Proof. We have

$$\begin{aligned}
a_{n,m}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\} &= a_{n,m}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - \left(\frac{k}{np}\right)^{-\gamma}\frac{Q_X(1/p)}{Q_X(n/k)}\right\} \\
&+ a_{n,m}\left\{\left(\frac{k}{np}\right)^{-\gamma}\frac{Q_X(1/p)}{Q_X(n/k)} - 1\right\} \\
&\underset{p\to 0}{\sim} a_{n,m}\left\{1 - \left(\frac{k}{np}\right)^{-\gamma}\frac{Q_X(1/p)}{Q_X(n/k)}\right\} + a_{n,m}\left\{\left(\frac{k}{np}\right)^{-\gamma}\frac{Q_X(1/p)}{Q_X(n/k)} - 1\right\},
\end{aligned}$$

on account of (5.20). Therefore as $n \to \infty$, Lemma 3.4.2 yields (5.22). $\square$

Proposition 5.4.5. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1). Then if $\alpha > 1$, we have the following:*

- *If* $m(1-2d)<1$ *and* $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ *or* $m(1-2d)>1$, *then*

$$\sqrt{k}\left\{\frac{\widetilde{\theta}_X(p)}{\theta_X(p)}-1\right\}\xrightarrow[n\to\infty]{d}\frac{B(1)}{\alpha}. \tag{5.23}$$

- *If* $m(1-2d)<1$ *and* $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, *then*

$$\frac{n}{b_{n,m}}\left\{\frac{\widetilde{\theta}_X(p)}{\theta_X(p)}-1\right\}\xrightarrow[n\to\infty]{d}\frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m!E(\phi^\alpha(Y))}. \tag{5.24}$$

Proof. We have the following decomposition:

$$\begin{aligned}\left\{\frac{\widetilde{\theta}_X(p)}{\theta_X(p)}-1\right\} &= \left\{\frac{\widetilde{Q}_X(1/p)}{(1-\gamma)\theta_X(p)}-1\right\}\\ &= \left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}+\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)}-1\right\}\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}.\end{aligned}$$

If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ or $m(1-2d)>1$, then (5.5) and Lemma 5.4.4 yield (5.23).
If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, then (5.24) follows by (5.6) and Lemma 5.4.4.

□

We continue with the situation where α is unknown and investigate the asymptotic behaviour of the Expected Shortfall estimator $\widehat{\theta}_X(1/p)$. In what follows, we replace the unknown value of $\gamma=1/\alpha$ with the harmonic moment estimator $\widehat{\gamma}=\widehat{\gamma}_{r,k}$ of order r, as defined in (4.76).

Theorem 5.4.6. *Let* $\{X_j\}$ *be the long memory stochastic volatility model with leverage as in* (3.1) *with* $\alpha>1$. *Let* $\mathcal{N}$ *denote a standard normal variable.*

1. *Assume that* $k/np\to\nu$, $0<\nu<\infty$, *as* $n\to\infty$. *If* $m(1-2d)<1$ *and* $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ *or* $m(1-2d)>1$, *then*

$$\sqrt{k}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\}\xrightarrow[n\to\infty]{d}\mathcal{N}\sqrt{\frac{1}{\alpha^2}+\frac{(\alpha+r-1)^2}{\alpha^3(\alpha+2r-2)}\left\{\frac{\alpha}{\alpha-1}+\ln\nu\right\}^2}. \tag{5.25}$$

2. *Assume that* $k/np\to\nu$, $0<\nu<\infty$, *as* $n\to\infty$. *If* $m(1-2d)<1$ *and* $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, *then*

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\}\xrightarrow[n\to\infty]{d}\frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m!E(\phi^\alpha(Y))}. \tag{5.26}$$

3. *Assume that $k/np \to \infty$, as $n \to \infty$. If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to 0$ or $m(1-2d) > 1$, then*

$$\frac{\sqrt{k}}{\ln(k/np)}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{(\alpha+r-1)}{\sqrt{\alpha^3(\alpha+2r-2)}}\mathcal{N}. \tag{5.27}$$

4. *Assume that $k/np \to \infty$, as $n \to \infty$. If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, then*

$$\frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} \xrightarrow[n\to\infty]{d} 0. \tag{5.28}$$

Proof. Similarly to the proof of Proposition 5.4.5, we have the following decomposition

$$\begin{aligned}
\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} &= \left\{\frac{\widehat{Q}_X(1/p)}{(1-\widehat{\gamma})\theta_X(p)} - 1\right\} \\
&= \left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} + \underbrace{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}}_{=:A_n}\underbrace{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)}}_{=:a_n}\underbrace{\frac{1}{(1-\widehat{\gamma})}}_{=:B_n}\{\widehat{\gamma}-\gamma\} \\
&\quad + \left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\}\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}.
\end{aligned}$$

Theorem 5.3.3 implies that under any of its conditions, A_n converges in probability to 1. Also, Theorem 4.3.16 gives that B_n converges in probability to $(1-\gamma)^{-1}$. Finally, a_n converges to 1 by Lemma 5.4.4. Thus, by Slutsky's theorem, A_n and a_n can be replaced by 1, while B_n can be replaced by $(1-\gamma)^{-1}$ and we will denote this change by $\approx$ in the calculations that follow.

1. Assume that $k/np \to \nu$, $0 < \nu < \infty$, as $n \to \infty$. If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to 0$ or $m(1-2d) > 1$, then

$$\begin{aligned}
\sqrt{k}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} &\approx \sqrt{k}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\} \\
&+ \frac{1}{(1-\gamma)}\sqrt{k}\,(\widehat{\gamma}-\gamma) + \sqrt{k}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\}.
\end{aligned}$$

Since the limiting behaviour of $\widehat{Q}_X(1/p)$ involves the limiting behaviour of the intermediate order statistics and $\widehat{\gamma}$, we need to further decompose it using (5.13):

$$\sqrt{k}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\} \approx \sqrt{k}\left\{\frac{\widetilde{Q}_X(1/p)}{Q_X(1/p)}-1\right\}$$
$$+\left\{\frac{1}{1-\gamma}+\ln\left(\frac{k}{np}\right)\right\}\sqrt{k}\left\{\widehat{\gamma}-\gamma\right\}+\left(\frac{k}{np}\right)^{-\gamma}\sqrt{k}R_n.$$
$$+\sqrt{k}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)}-1\right\}.$$

By Lemma 5.4.4, the last term in the line above is $o(1)$. Using (5.14), Theorems 4.3.16 and 5.3.3, we obtain

$$\sqrt{k}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\} \xrightarrow[n\to\infty]{d} \frac{B(1)}{\alpha}+\frac{(\alpha+r-1)}{\sqrt{\alpha^3(\alpha+2r-2)}}\left\{\frac{1}{1-\gamma}+\ln\nu\right\}\mathcal{N}\,.$$

Thus (5.25) follows since estimators of the tail index and order statistics are asymptotically uncorrelated by (4.74).

2. Assume that $k/np \to \nu$, $0<\nu<\infty$, as $n\to\infty$. If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to\infty$, then

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\} \approx \frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\}$$
$$+\frac{1}{(1-\gamma)}\frac{n}{b_{n,m\sqrt{k}}}\sqrt{k}\left(\widehat{\gamma}-\gamma\right)+\frac{n}{b_{n,m}}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)}-1\right\}.$$

Theorems 4.3.16 and 5.3.3 and Lemma 5.4.4 yield (5.26).

3. Assume that $k/np\to\infty$, as $n\to\infty$. If $m(1-2d)<1$ and $\frac{b_{n,m}}{n}\sqrt{k}\to 0$ or $m(1-2d)>1$, then

$$\frac{\sqrt{k}}{\ln\left(k/np\right)}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)}-1\right\} \approx \frac{\sqrt{k}}{\ln\left(k/np\right)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)}-1\right\}$$
$$+\frac{1}{\ln\left(k/np\right)}\frac{\sqrt{k}\{\widehat{\gamma}-\gamma\}}{1-\gamma}+\frac{\sqrt{k}}{\ln\left(k/np\right)}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)}-1\right\}.$$

Theorems 4.3.16 and 5.3.3 and Lemma 5.4.4 yield (5.27).

4. Assume that $k/np \to \infty$, as $n \to \infty$. If $m(1-2d) < 1$ and $\frac{b_{n,m}}{n}\sqrt{k} \to \infty$, then

$$\frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} \approx \frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\}$$
$$+\frac{1}{(1-\gamma)}\frac{n}{b_{n,m}\sqrt{k}\ln(k/np)}\sqrt{k}\,(\widehat{\gamma}-\gamma) + \frac{n}{b_{n,m}\ln(k/np)}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\}.$$

Theorems 4.3.16 and 5.3.3 and Lemma 5.4.4 yield (5.28).

□

Remark 5.4.7. In general, as was the case with VaR, when returns of a portfolio are assumed to exhibit heavy-tails, long range dependence and leverage, estimation of ES is influenced in a dichotomous way either by heavy-tails or long memory, while leverage does not contribute at all. As before, this is demonstrated in the following scenarios:

1. In the case scenario (5.25) the limiting behaviour is affected by both order statistics and estimation of the tail index.

2. In the case scenario (5.26) the limiting behaviour is affected by limiting behaviour of order statistics only. Estimation of the tail index does not play any role.

3. In the case scenario (5.27) the limiting behaviour is affected by estimation of the tail index only.

4. In the case scenario (5.28) we end up with a degenerate limit. This is addressed in the result by imposing more detailed conditions.

The next result addresses the degeneracy observed in (5.28).

Theorem 5.4.8. *Let $\{X_j\}$ be the long memory stochastic volatility model with leverage as in (3.1) with $\alpha > 1$. Assume that $k/np \to \infty$. If $m(1-2d) < 1$ and (5.16) holds, then*

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} \xrightarrow[n\to\infty]{d} \frac{\mu_{\phi,\alpha}(m)\xi_{m,d+1/2}(1)}{\alpha m! E(\phi^\alpha(Y))}. \tag{5.29}$$

Proof. We have

$$\frac{n}{b_{n,m}}\left\{\frac{\widehat{\theta}_X(p)}{\theta_X(p)} - 1\right\} \approx \frac{n}{b_{n,m}}\left\{\frac{\widehat{Q}_X(1/p)}{Q_X(1/p)} - 1\right\}$$
$$+\frac{1}{(1-\gamma)}\frac{n}{b_{n,m}\sqrt{k}}\sqrt{k}\,(\widehat{\gamma}-\gamma) + \frac{n}{b_{n,m}}\left\{\frac{Q_X(1/p)}{(1-\gamma)\theta_X(p)} - 1\right\}.$$

Theorems 4.3.16 and 5.3.5 and Lemma 5.4.4 yield (5.29). □

5.5 Concluding Remarks

Throughout this chapter, we have discussed estimation of financial risk using two approaches: Value-at-Risk and Expected Shortfall under the assumptions that returns of a portfolio are heavy-tailed long memory sequences with leverage (3.1).

We have derived estimators of VaR in (5.4a)-(5.4b) and subsequently studied their asymptotic behaviour in Proposition 5.3.2 and Theorems 5.3.3 and 5.3.5. We have shown that these estimators are consistent and are functions of both intermediate order statistics and Harmonic Moment Estimators of the index of regular variation. Therefore to prove limit theorems for estimators of VaR, we have used convergence of intermediate order statistics and asymptotic normality of HMEs.

On the other hand, we have derived estimators of ES and investigated their limiting behaviour in Proposition 5.4.5 and Theorems 5.4.6 and 5.4.8. We have shown that the estimators of ES are consistent and depend not only on those of VaR but also on HMEs (5.21a)-(5.21b). This justifies the importance of VaR despite its deficiencies and confirms the fact ES is a natural coherent alternative to VaR. To establish the asymptotic behaviour of the estimators of ES, we have used the asymptotic normality of HMEs and limit theorems for estimators of VaR.

The conclusion we reach is that when returns of a portfolio are assumed to exhibit heavy-tails, long memory and leverage, estimators of VaR and ES are not affected by the leverage effect at all, while heavy tails and long memory influence the estimators in a dichotomous way.

Chapter 6

Simulation Studies

6.1 Introductory Comments

In this chapter, we perform some numerical studies to illustrate the following theoretical results obtained in the preceding chapters:

1. tail index estimation using the Hill estimator,

2. estimation of Value-at-Risk,

3. and estimation of Expected Shortfall.

For this, three types of data are considered:

- iid data (Pareto or the absolute value of Student-t);

- the long memory stochastic volatility model without leverage, with $\{Z_j, j = 1, \ldots, n\}$ either Pareto or the absolute value of Student-t;

- the long memory stochastic volatility model with leverage.

For the Student-t distribution, recall that the tail index α is equal to the number of degrees of freedom. For the LMSV models (without and with leverage) we choose the long memory parameter d as $d = 0.1$ or $d = 0.4$.

In what follows, the term *Hill plot* refers to the type of plot where an estimator (of the tail index, Value-at-Risk, Expected Shortfall, etc.) is plotted against the number of order statistics used in the evaluation of the estimator. From a practical perspective, the Hill plot is used to choose k, an appropriate number of order statistics, from a stable region of the plot.

This chapter is structured as follows. In Sections 6.2 and 6.3, we do experiments aiming at graphing Hill plots of the following estimators: Hill estimator, Value-at-Risk and Expected Shortfall using iid data and the LMSV model, respectively. In Sections 6.4 and 6.5 we perform Monte Carlo simulation studies for the aforementioned estimators. In Section 6.6 we perform Monte Carlo simulations for the Hill estimator in the case of the LMSV model with leverage. All figures appear in Section 6.7.

Our simulations illustrate the following:

- Estimation in the case of Student-t noise is usually very biased, regardless of the presence of long memory or leverage.
- Long range dependence **does not affect the behaviour of the Hill estimator**, as indicated by our theory.
- Long range dependence **affects the behaviour of the estimators of VaR and ES**, as indicated by our theory.
- Leverage **does not affect the behaviour of the Hill estimator**, as indicated by our theory.

The main message of these experiments is that the quality of the estimators depends largely on the underlying distribution of the noise, with a lot of bias and instability in the case of the Student-t distribution. Long memory does not play any role in the case of the Hill estimator, but influences the asymptotics of the estimators of Value-At-Risk and Expected Shortfall. Leverage does not play any role in the case of the Hill estimator.

6.2 Simulations for i.i.d data

6.2.1 Experiment 1: Hill Plots for Hill Estimator

In this experiment, we assume that X_j are independent and identically Pareto or Student distributed with parameter α. We simulated 1000 observations from these two distributions with the following choices for the parameter: $\alpha = 2$ and $\alpha = 4$, respectively. On the left panel of Figure 6.1, we plotted estimates of the reciprocal of the tail index $1/\alpha$ of the

Pareto distribution using the Hill estimator. On the right panel of Figure 6.1 we plotted the corresponding estimates for the Student-t distribution. We note the instability of the Hill estimator when dealing with the Student-t distribution.

6.2.2 Experiment 2: Hill Plots for Value-at-Risk

In this experiment, we assume that X_j are independent and identically Pareto or Student distributed with parameter α. We simulated 1000 observations from these two distributions with the following choices for the parameter $\alpha = 2$ and $\alpha = 4$, respectively. Value-at-Risk, $Q_X(1/p)$, is estimated by $\widehat{Q}_X(1/p)$ defined in (5.4b). On the left panel of Figure 6.2, we plotted estimates of the ratio $\widehat{Q}_X(1/p)/Q_X(1/p)$ for the level $p = 0.05$ using the Pareto distribution. On the right panel of Figure 6.2 we plotted those for the Student-t distribution. Recall that the quantile function $Q_X(1/p)$ is

$$Q_X(1/p) = 2(1/2 - p)\sqrt{\frac{1}{2p(1-p)}}\ , \quad \alpha = 2\ ,$$

$$Q_X(1/p) = \operatorname{sign}(1-p)2\sqrt{\frac{\cos(\frac{1}{3}\arccos(\sqrt{4p(1-p)}))}{\sqrt{4p(1-p)}} - 1}\ , \quad \alpha = 4\ .$$

We note some consistency in the Pareto case but again instability and bias in the Student case.

6.2.3 Experiment 3: Hill Plots for Expected Shortfall

In this experiment, we assume that X_j are independent and identically Pareto or Student distributed with parameter α. We simulated 1000 observations from these two distributions with the following choices for the parameter $\alpha = 2$ and $\alpha = 4$, respectively. Expected Shortfall $\theta_X(p)$ is estimated by $\widehat{\theta}_X(p)$ defined in (5.21b). On the left panel of Figure 6.3, we plotted estimates of the ratio $\widehat{\theta}_X(p)/\theta_X(p)$ with the level $p = 0.05$ using the Pareto distribution. On the right panel of Figure 6.3 we plotted those for the Student-t distribution. Again, we observe relatively good behaviour for the Pareto distribution and bias in the Student case.

6.3 Simulations for the LMSV Model

We simulate 1000 observations from the long memory stochastic volatility model:

$$X_j = \exp(\sigma Y_j))Z_j \ , j = 1, \dots, 1000 \tag{6.1}$$

where

1) Z_j is a regularly varying sequence of random variables with index $-\alpha$ (again, either Pareto or Student).

2) Y_j is a fractional Gaussian noise sequence, that is, $\mathrm{Cov}(Y_0, Y_j) \sim j^{2d-1}$, with the long memory parameter $d \in (0, 1/2)$. This is simulated using the `R` command `fracdiff`.

3) $\sigma > 0$. Throughout this section, the variability parameter is $\sigma = 0.1$.

6.3.1 Experiment 1: Hill Plots for Hill Estimator

On the left panel of Figure 6.4, the long memory parameter d is chosen to be 0.1. We plotted estimates of the reciprocal of the tail index $1/\alpha$ of the Pareto distribution using the Hill estimator for $\alpha = 2$ and $\alpha = 4$, respectively.

On the right panel of Figure 6.4, the long memory parameter d is chosen to be 0.4. The quality of the estimator does not depend on the long memory parameter.

As for the Student case, we proceed similarly with the long memory parameter d being 0.1 and 0.4 on the left and right panel of Figure 6.5, respectively. We observe again instability of the Hill estimator regardless the memory parameter.

In summary, the tail index estimation using the Hill estimator is unaffected by long memory.

6.3.2 Experiment 2: Hill Plots for Value-at-Risk

In this experiment, we plotted estimates of the ratio $\widehat{Q}_X(1/p)/Q_X(1/p)$ for the level $p = 0.05$ using the LMSV model as in (6.1). In Figure 6.6, we used as noise the Pareto distribution for $\alpha = 2$ and 4, respectively. On the left panel, the long memory parameter d is chosen to be 0.1 while on its right panel it is 0.4. Figure 6.7 is similar with the Student-t distribution as noise. The estimation of Value-at-Risk is good in the case of Pareto noise for $\alpha = 2, d = 0.1$ and $\alpha = 4, d = 0.4$. In the case of Student noise, it is good only for $\alpha = 2$.

6.3.3 Experiment 3: Hill Plots for Expected Shortfall

We plotted estimates of the ratio $\widehat{\theta}_X(p)/\theta_X(p)$ for the level $p = 0.05$ using the LMSV model as in (6.1). In Figure 6.8, we considered as noise the Pareto distribution with $\alpha = 2$ and 4, respectively. On the left panel, the long memory parameter d is chosen to be 0.1 while on its right panel it is 0.4. Figure 6.9 is analogous for Student-t noise. When $d = 0.4$, the estimation of Expected Shortfall is good in the case of Pareto noise. In the Student case, it is only good for $d = 0.4$ but $\alpha = 2$.

6.4 Monte Carlo Simulations for i.i.d data

6.4.1 Experiment 1: Boxplots for Hill Estimator

In this experiment we produced boxplots for the estimates of the reciprocal of the tail index $1/\alpha$ using the Hill estimator. For this, we simulated $n = 1000$ observations from either the Pareto or Student-t distributions with $\alpha = 2$ and $\alpha = 4$, respectively. In Figures 6.10 and 6.11, we deal with the Pareto distribution while in Figures 6.12 and 6.13, we consider the Student-t distribution. These estimates are obtained from 1000 Monte Carlo simulations, for the following number of order statistics in both figures: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel). We note that

- for the Pareto case, the estimator is unbiased;
- for the Student case, the bias increases with k;
- the larger k is, the smaller the variability is, as suggested by the asymptotic theory.

6.4.2 Experiment 2: Boxplots for Value-at-Risk

In this experiment, we display boxplots for the ratio $\widehat{Q}_X(1/p)/Q_X(1/p)$ for the level $p = 0.05$. To do so, we simulated $n = 1000$ observations from the Pareto distribution for $\alpha = 2$ and $\alpha = 4$, in Figures 6.14 and 6.15 respectively. These estimates are obtained from 1000 Monte Carlo iterations, for the following number of order statistics in both figures: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel). In Figures 6.16 and 6.17, the same procedure is repeated with data being generated with Student noise. We note that

- variability decreases with k, as suggested by the asymptotic theory;
- for both Pareto and Student noise, we have several estimates that differ significantly from the true value;
- estimates for the Student case are biased, but there is less bias for $\alpha = 4$ than for $\alpha = 2$.

6.4.3 Experiment 3: Boxplots for Expected Shortfall

In this experiment, we display boxplots of the ratio $\widehat{\theta}_X(p)/\theta_X(p)$ for the level $p = 0.05$. To do so, we simulated $n = 1000$ observations from the Pareto distribution for $\alpha = 2$ and $\alpha = 4$, in Figures 6.18 and 6.19 respectively. These estimates are obtained from 1000 Monte Carlo iterations, for the following selected choice of the number of order statistics in both figures: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel). In Figures 6.20 and 6.21, the same procedure is repeated with data being generated from the Student distribution.

- variability decreases with k, as suggested by the asymptotic theory;
- for both Pareto and Student noise, we have several estimates that differ significantly from the true value;
- estimates for the Student case are very biased for $\alpha = 2$, but there is a little bias for $\alpha = 4$.

6.5 Monte Carlo Simulations for LMSV data

6.5.1 Experiment 1: Boxplots for Hill Estimator

In this experiment we produce boxplots for the estimates $1/\hat{\alpha}_k$ of the reciprocal of the tail index $1/\alpha$ using the Hill estimator. For this, we simulated $n = 1000$ observations from the LMSV model defined in (6.1) with $d = 0.1$ and $d = 0.4$. In Figures 6.22 and 6.23, we consider the Pareto distribution while in Figures 6.24 and 6.25, we consider the Student-t distribution with $\alpha = 2, 4$ in both cases. These estimates are obtained from 1000 Monte Carlo simulations, for the following selected choice of the number of order statistics in both figures: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

- There is no difference between boxplots for both long memory parameters, as suggested by the theory.

6.5.2 Experiment 2: Boxplots for Value-at-Risk

In this experiment, we produced boxplots for the ratio $\widehat{Q}_X(1/p)/Q_X(1/p)$ for the level $p = 0.05$ using the LMSV model as in (6.1) with long memory parameters $d = 0.1$ and $d = 0.4$. In Figures 6.26 and 6.27, we considered the Pareto distribution while in Figures 6.28 and 6.29, we considered the Student-t distribution with $\alpha = 2, 4$ in both cases. These estimates are obtained from 1000 Monte Carlo simulations, for the following selected choice of the number of order statistics in both figures: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

- We note that variability for $d = 0.4$ is bigger than for $d = 0.1$ (unlike in the Hill estimator case), as suggested by the asymptotic theory, since long memory dominates.

6.5.3 Experiment 3: Boxplots for Expected Shortfall

In this experiment, we produced boxplots for the ratio $\widehat{\theta}_X(p)/\theta_X(p)$ for the level $p = 0.05$ using the LMSV model as in (6.1) with long memory parameters $d = 0.1$ and $d = 0.4$. In Figures 6.30 and 6.31, we considered the Pareto distribution while in Figures 6.32 and 6.33, we considered the Student-t distribution with $\alpha = 2, 4$ in both cases. These estimates are obtained from 1000 Monte Carlo simulations, for the following selected choice of the number of order statistics in both figures: $k = 0.05*n$ (top panel), $k = 0.1*n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

- We note that variability for $d = 0.4$ is bigger than for $d = 0.1$ (similarly as for Value-at-Risk, but unlike the Hill estimator), as suggested by the asymptotic theory, since long memory dominates.

- There is a lot of bias in the case $\alpha = 2$ for both Pareto and Student noise (the ratios in Figures 6.30 and 6.32 should be centered around 1). Again, this is not related to long memory or leverage.

6.6 Simulations for the LMSV Model with Leverage

In this section we illustrate the lack of influence of leverage on the asymptotic distribution of the Hill estimator, as indicated by our theory. The simulation set-up is as follows:

- The sequence $\{Z_j, j = 1, \ldots, n\}$ is simulated from a Pareto distribution with tail index $\alpha = 2$ or $\alpha = 4$.
- The long memory sequence $\{Y_j, j = 1, \ldots, n\}$ is simulated in two ways using `fracdiff.sim` function from `R`:
  - First, we simulate `innov` sequence using `rnorm` command, independent from everything else. This `innov` sequence is fed into the `fracdiff.sim`. The result of that function is a long memory Gaussian sequence having (approximately) the same finite dimensional distributions as $Y_j = \sum_{i=1}^{\infty} a_i \epsilon_{j-i}$. Then we set $X = \exp(0.1Y)Z$. In this case there is no leverage.
  - Second, we obtain the `innov` sequence via the following procedure:

$$\text{innov}_j = \rho \Phi^{-1}(F_Z(Z_{j+1})) + \sqrt{1-\rho^2}\mathcal{N}_j \ ,$$

    where $\rho \in (0,1)$, F_Z is the distribution function of Z_0, Φ^{-1} is the quantile function for the standard normal distribution and $\mathcal{N}_j$ are i.i.d. standard normal random variables. This way, the innovation sequence is standard normal and is *not* independent from $\{Z_j, j = 1, \ldots, n\}$, allowing for the leverage effect.

The results are presented in Figures 6.34-6.37 for $\alpha = 2, 4$ and $d = 0.1, 0.4$. We compare the Monte Carlo simulations for the first (no-leverage) and the second (leverage) simulation methods. The following numbers of order statistics are used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

- We note very little difference between the performance of the Hill estimator in the two cases.

6.7 Figures

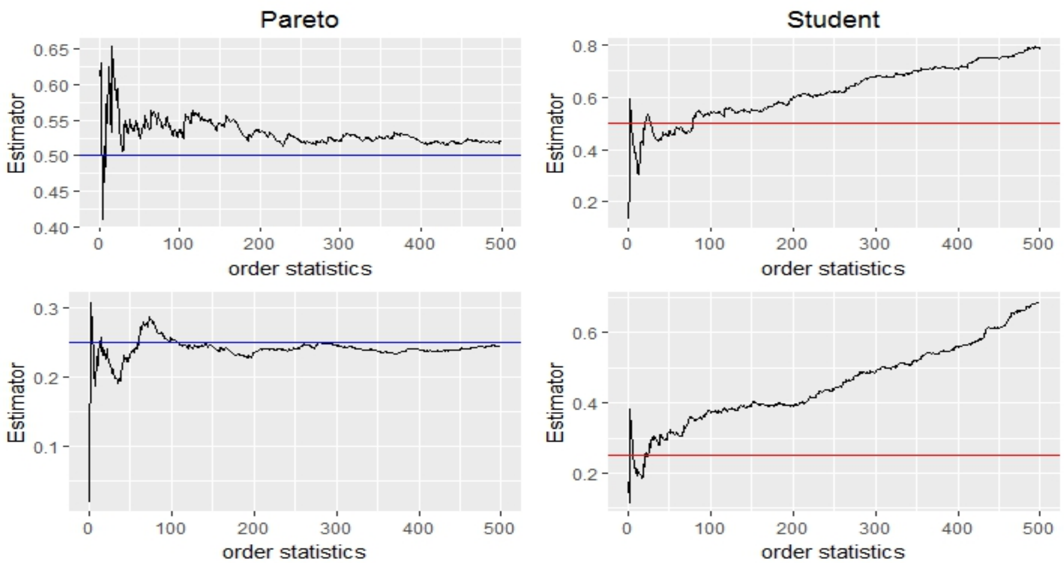

Figure 6.1: Estimation of the reciprocal of the tail index using the Hill estimator (iid case). The horizontal lines indicate the true values of $1/\alpha$. Left panel: Pareto; Right panel: Student-t. Top panel: $\alpha = 2$; Bottom panel: $\alpha = 4$.

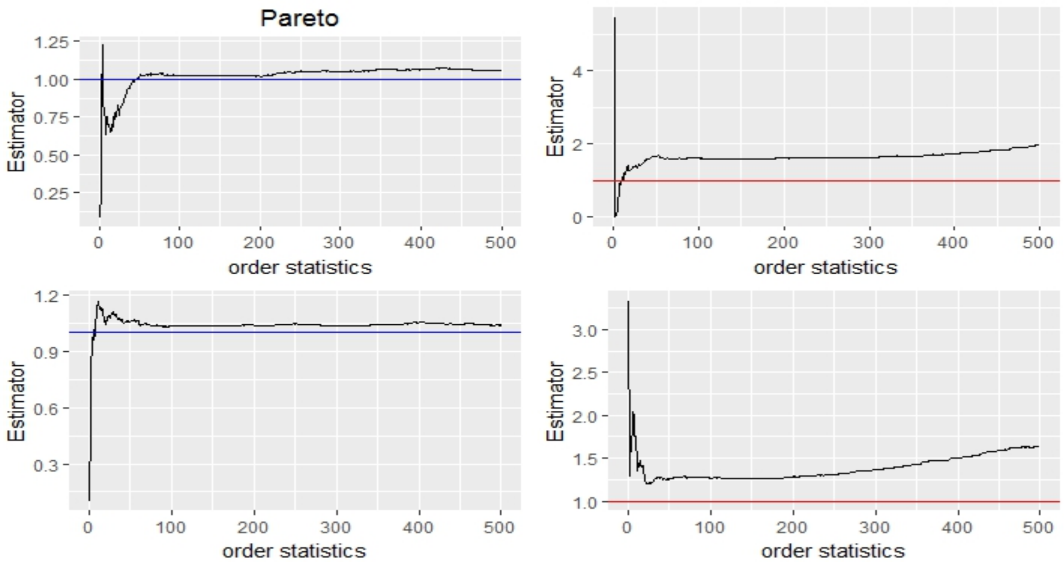

Figure 6.2: Estimation of Value-at-Risk (iid case). The horizontal lines indicate the true ratio 1. Left panel: Pareto; Right panel: Student-t. Top panel: $\alpha = 2$; Bottom panel: $\alpha = 4$.

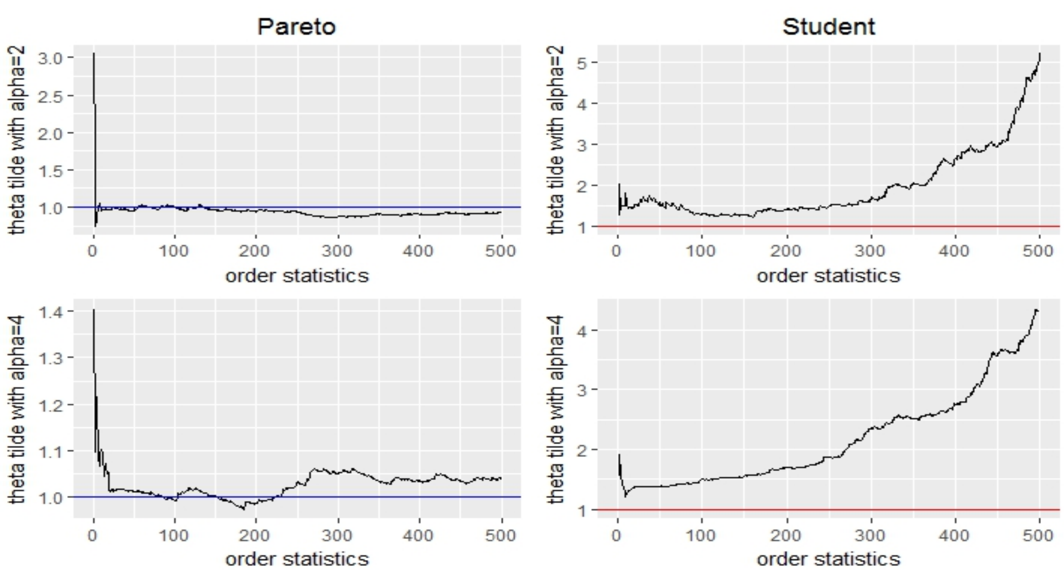

Figure 6.3: Estimation of Expected Shortfall (iid case). The horizontal lines indicate the true ratio 1. Left panel: Pareto; Right panel: Student-t. Top panel: $\alpha = 2$; Bottom panel: $\alpha = 4$.

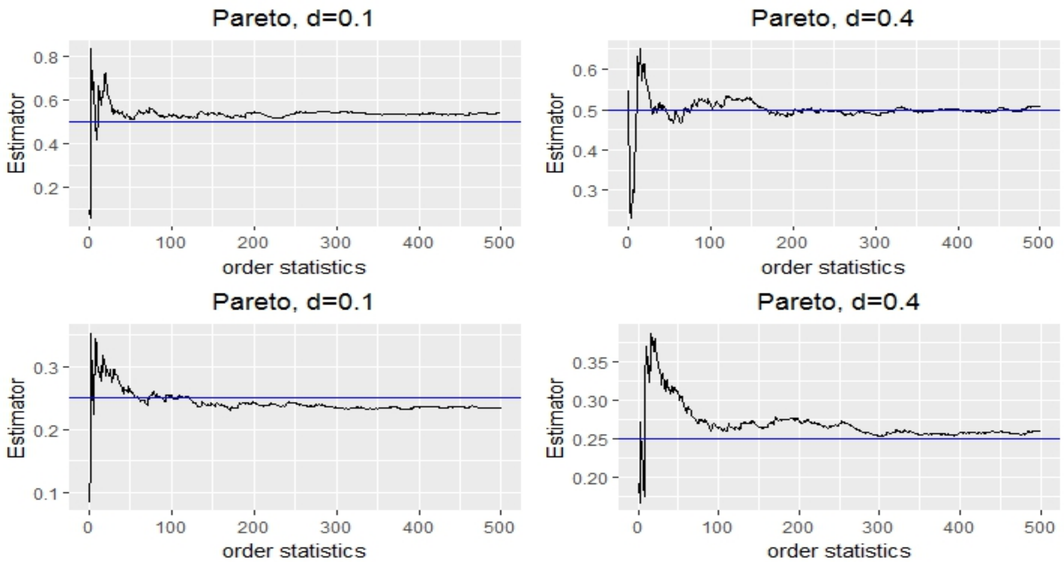

Figure 6.4: Estimation of the reciprocal of the tail index using the Hill estimator (LMSV case). The horizontal lines indicate the true values of $1/\alpha$. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Pareto with $\alpha = 2$; Bottom panel: Pareto with $\alpha = 4$.

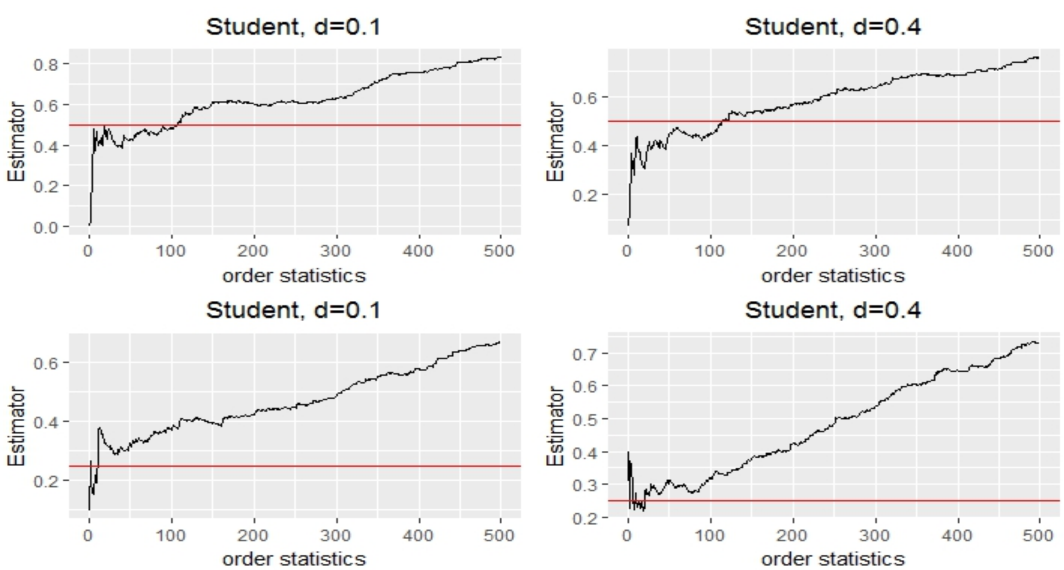

Figure 6.5: Estimation of the reciprocal of the tail index using the Hill estimator (LMSV case). The horizontal lines indicate the true values of $1/\alpha$. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Student-t with $\alpha = 2$; Bottom panel: Student-t with $\alpha = 4$.

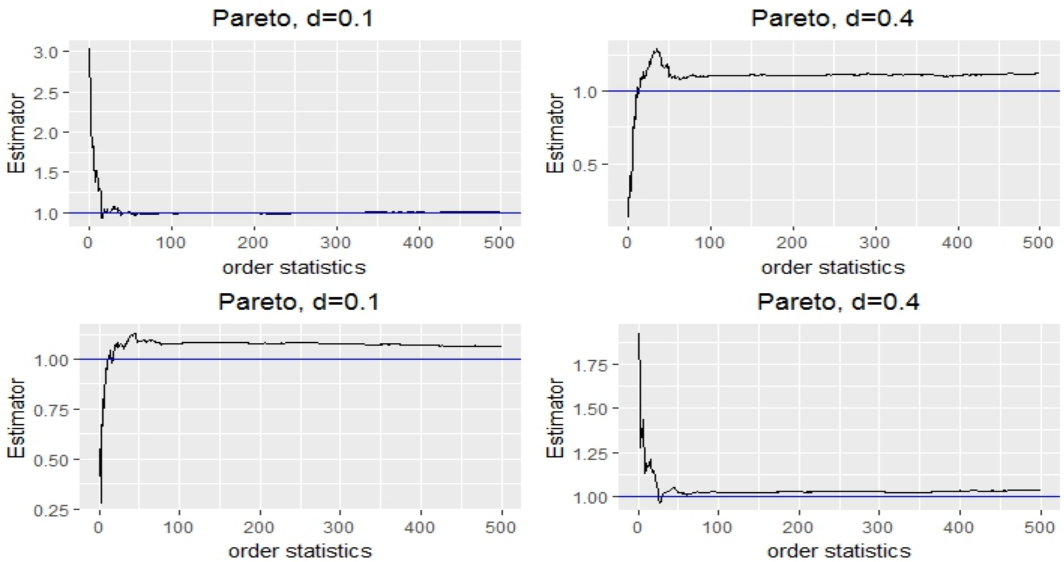

Figure 6.6: Estimation of the Value-at-Risk (LMSV case). The horizontal lines indicate the true ratio 1. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Pareto with $\alpha = 2$; Bottom panel: Pareto with $\alpha = 4$.

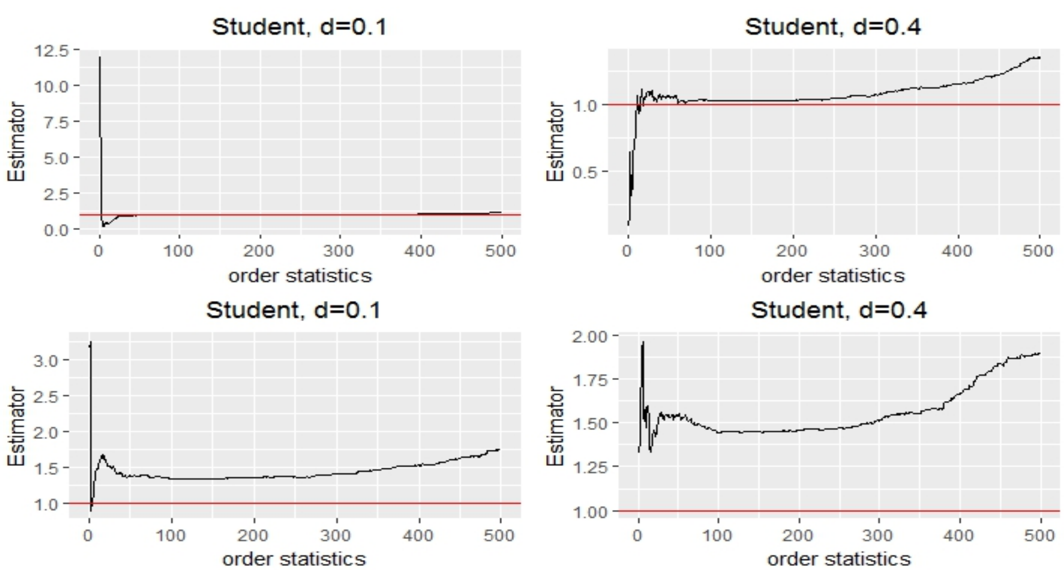

Figure 6.7: Estimation of the Value-at-Risk (LMSV case). The horizontal lines indicate the true ratio 1. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Student-t with $\alpha = 2$; Bottom panel: Student-t with $\alpha = 4$.

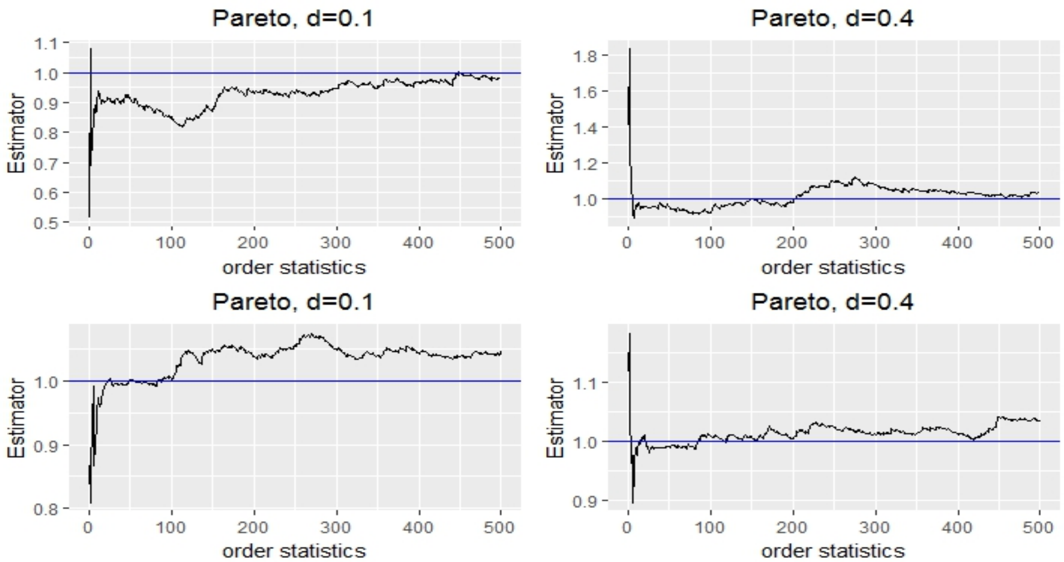

Figure 6.8: Estimation of the Expected Shortfall (LMSV case). The horizontal lines indicate the true ratio 1. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Pareto with $\alpha = 2$; Bottom panel: Pareto with $\alpha = 4$.

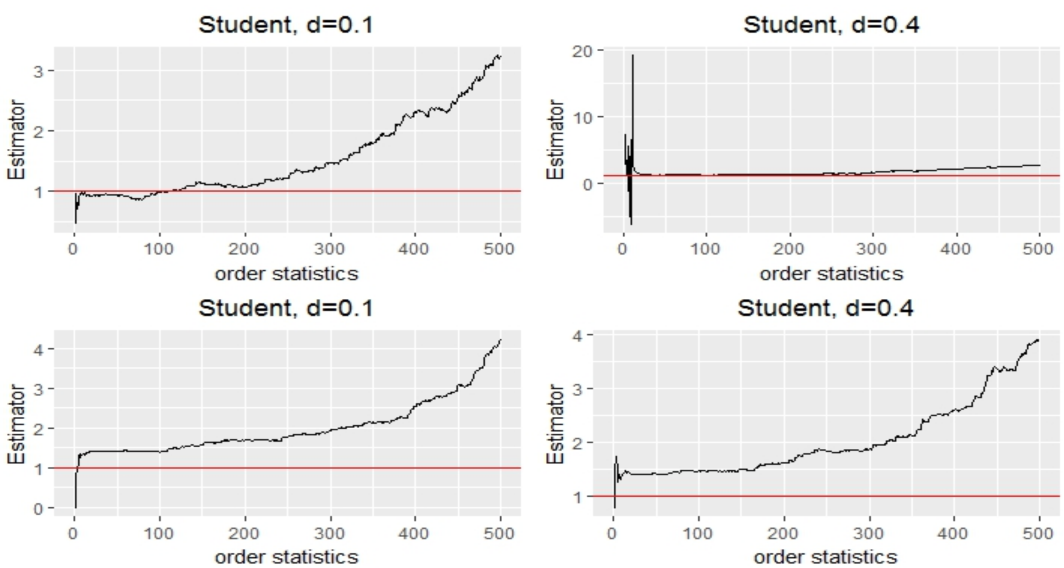

Figure 6.9: Estimation of the Expected Shortfall (LMSV case). The horizontal lines indicate the true ratio 1. Left panel: $d = 0.1$; Right panel: $d = 0.4$. Top panel: Student-t with $\alpha = 2$; Bottom panel: Student-t with $\alpha = 4$.

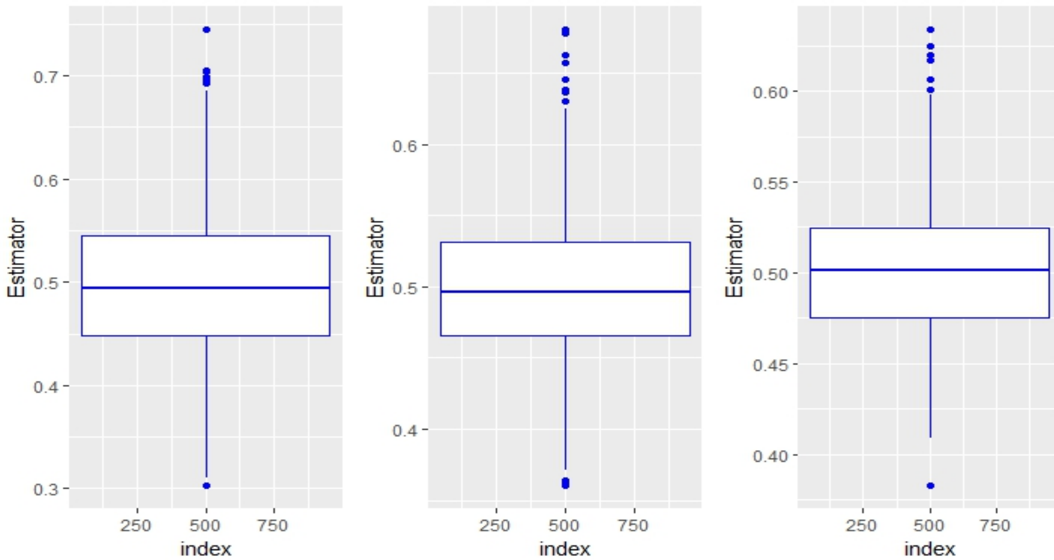

Figure 6.10: Boxplots for Hill estimator using Pareto distribution (iid case). The true value of $1/\alpha$ is $1/2$. Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

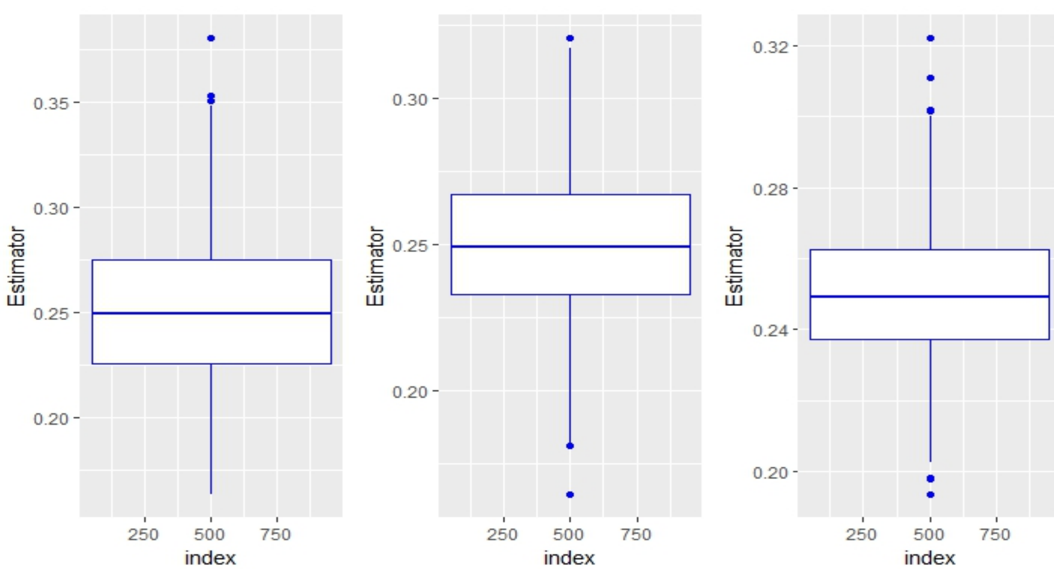

Figure 6.11: Boxplots for Hill estimator using Pareto distribution (iid case). The true value of $1/\alpha$ is $1/4$. Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

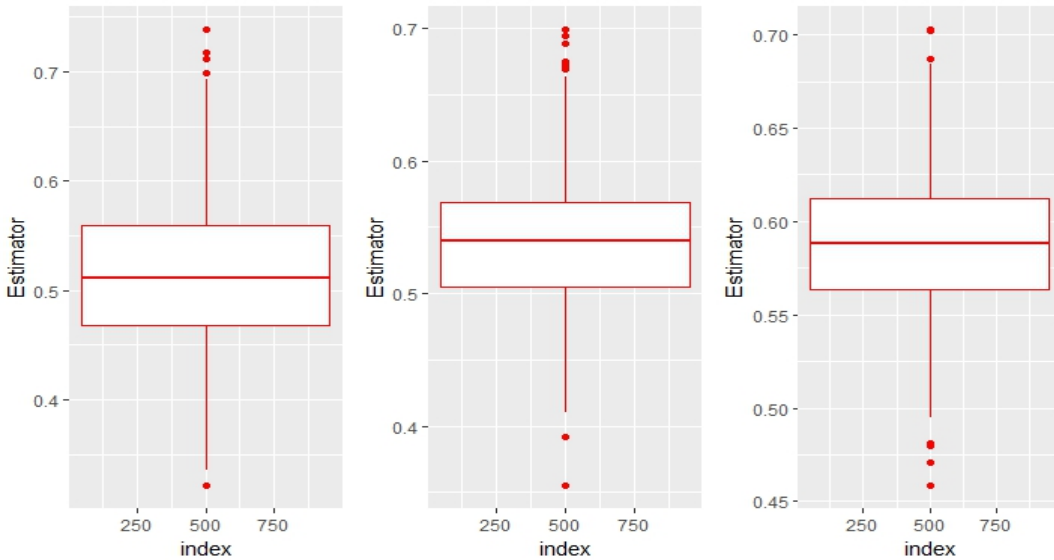

Figure 6.12: Boxplots for Hill estimator using Student distribution (iid case). The true value of $1/\alpha$ is $1/2$. Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

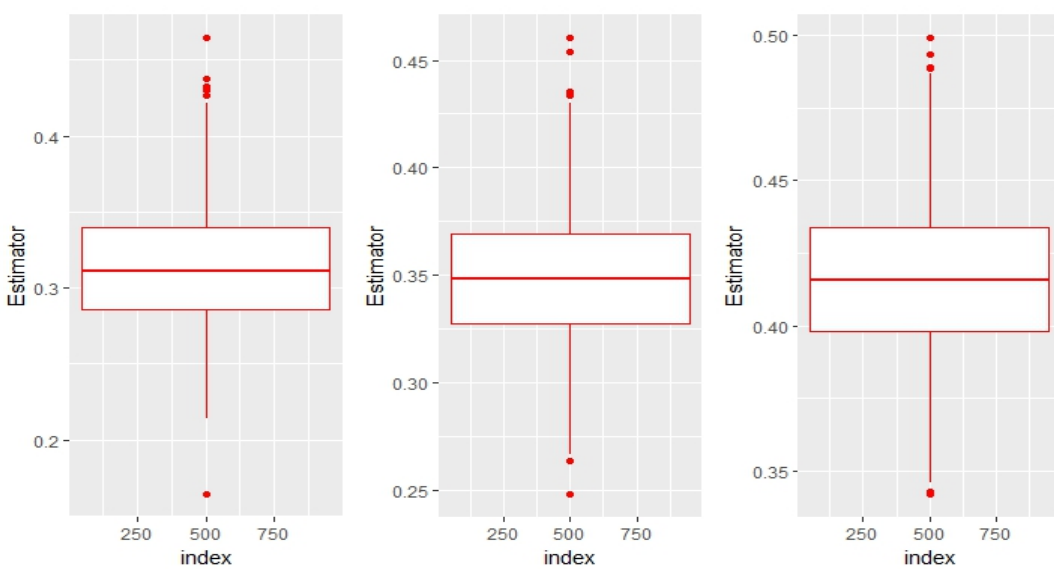

Figure 6.13: Boxplots for Hill estimator using Student distribution (iid case). The true value of $1/\alpha$ is $1/4$. Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

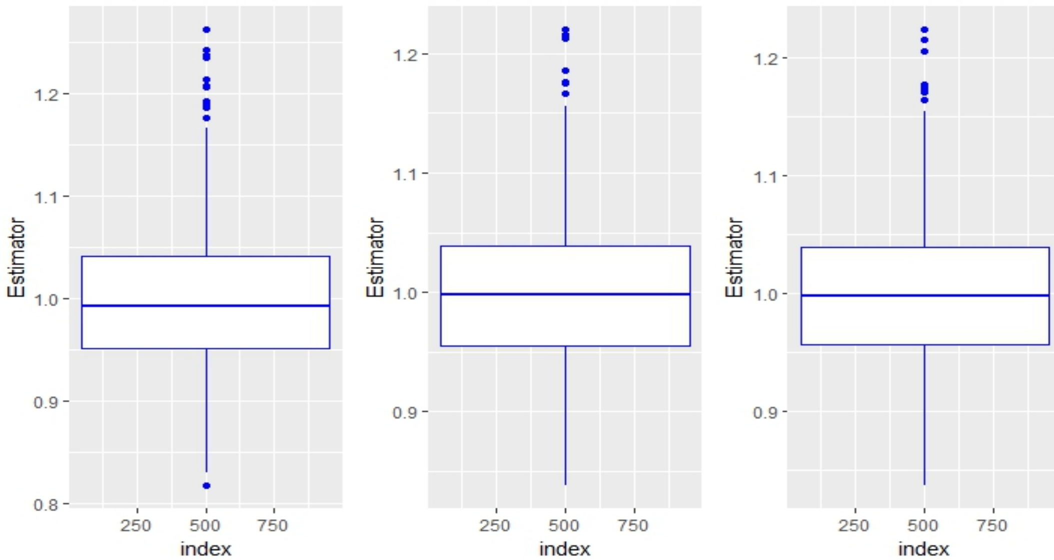

Figure 6.14: Boxplots for Value-at-Risk (iid Pareto case, $\alpha = 2$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

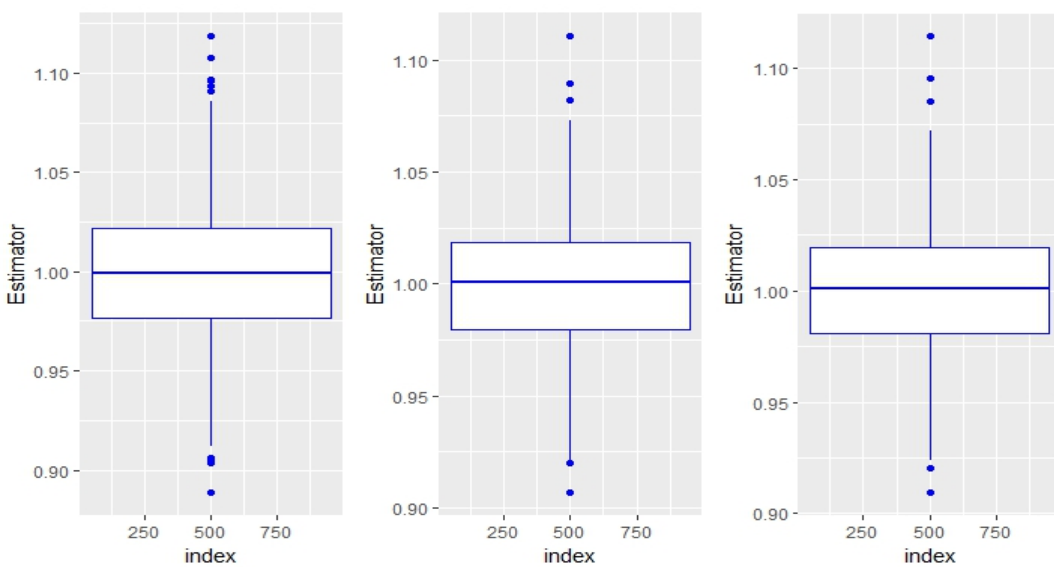

Figure 6.15: Boxplots for Value-at-Risk (iid Pareto case, $\alpha = 4$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

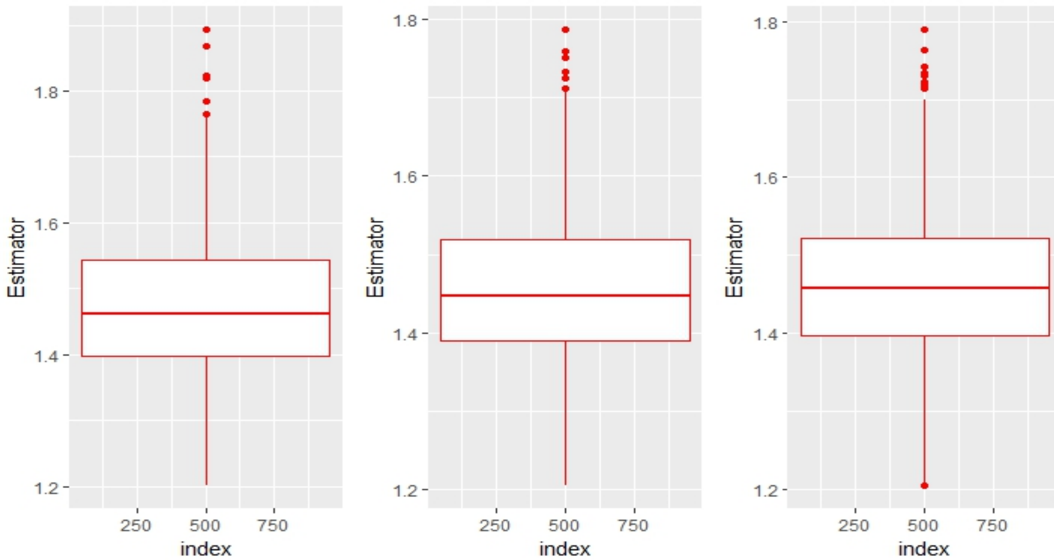

Figure 6.16: Boxplots for Value-at-Risk (iid Student case, $\alpha = 2$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

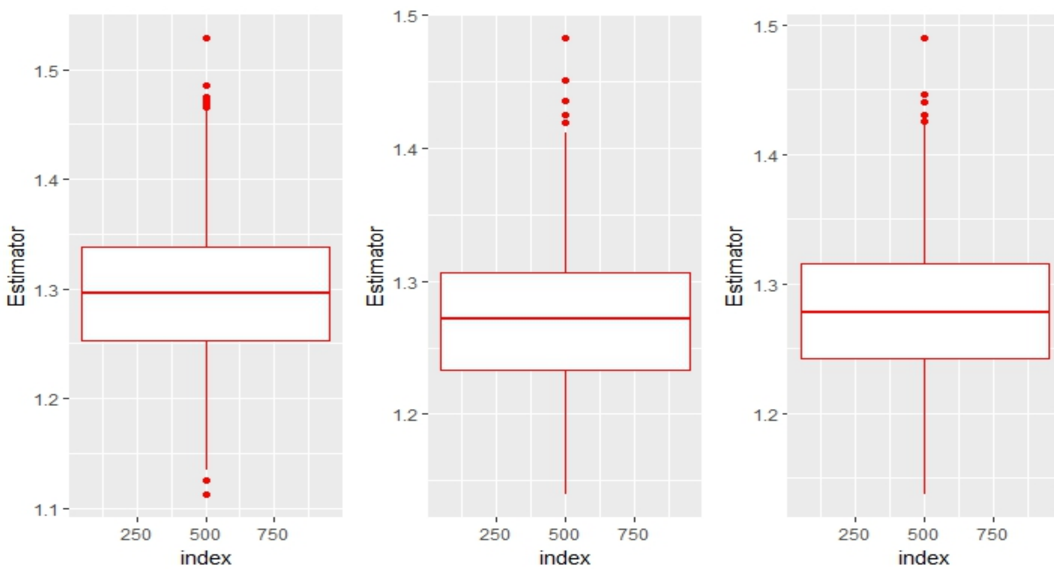

Figure 6.17: Boxplots for Value-at-Risk (iid Student case, $\alpha = 4$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

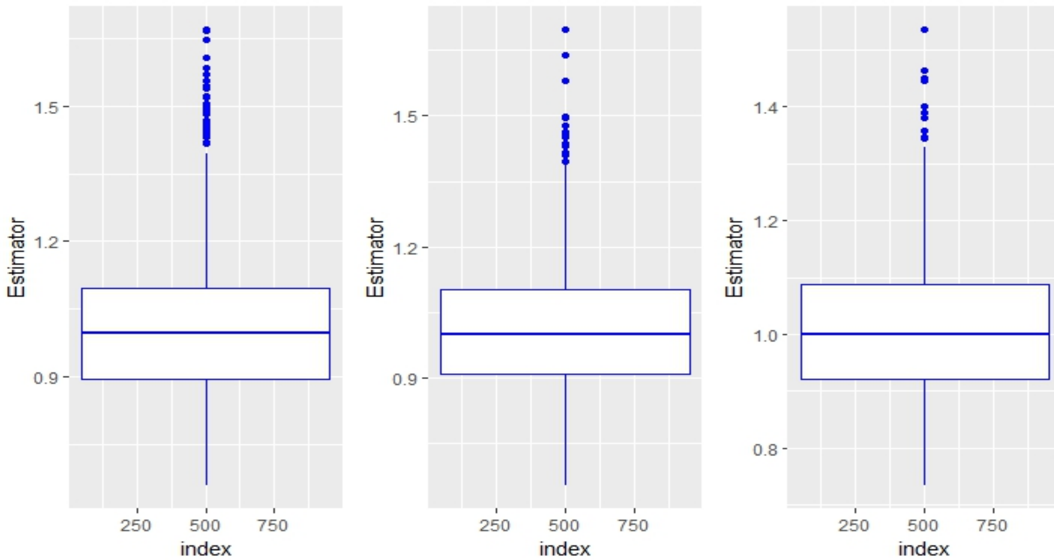

Figure 6.18: Boxplots for Expected Shortfall (iid Pareto case, $\alpha = 2$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

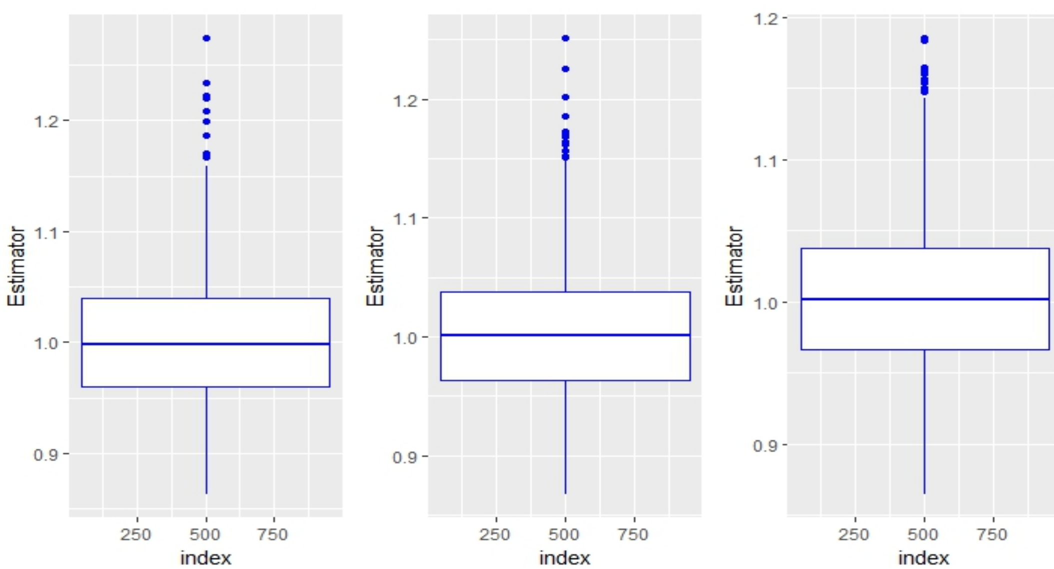

Figure 6.19: Boxplots for Expected Shortfall (iid Pareto case, $\alpha = 4$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

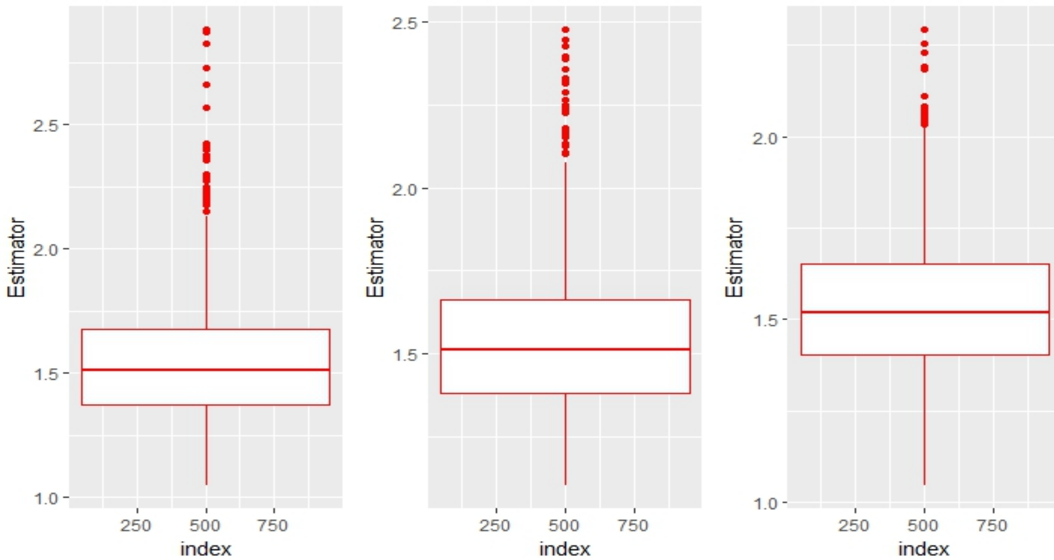

Figure 6.20: Boxplots for Expected Shortfall (iid Student case, $\alpha = 2$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

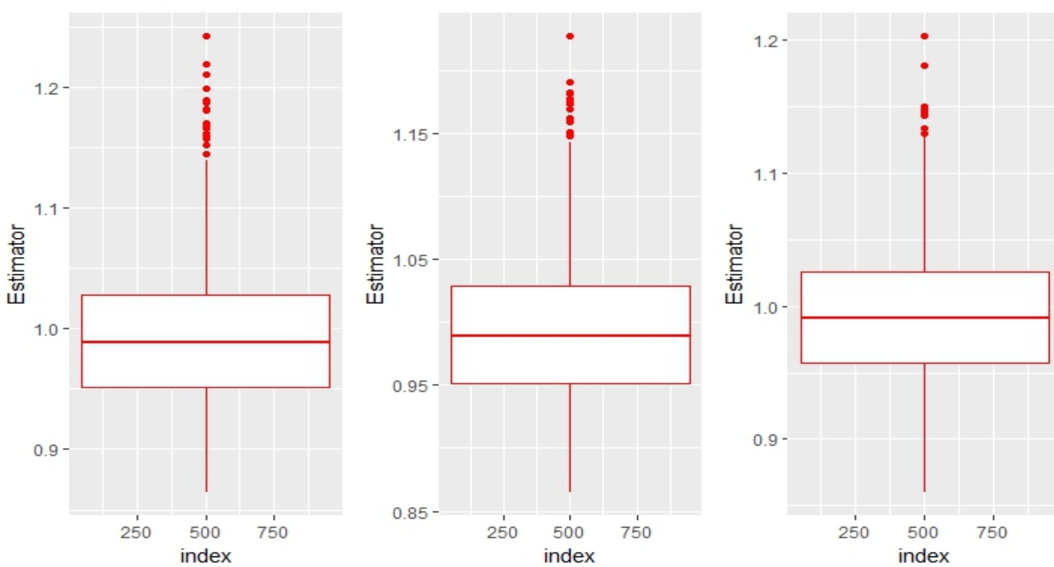

Figure 6.21: Boxplots for Expected Shortfall (iid Student case, $\alpha = 4$). Number of order statistics used: $k = 0.05 * n$ (left panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (right panel).

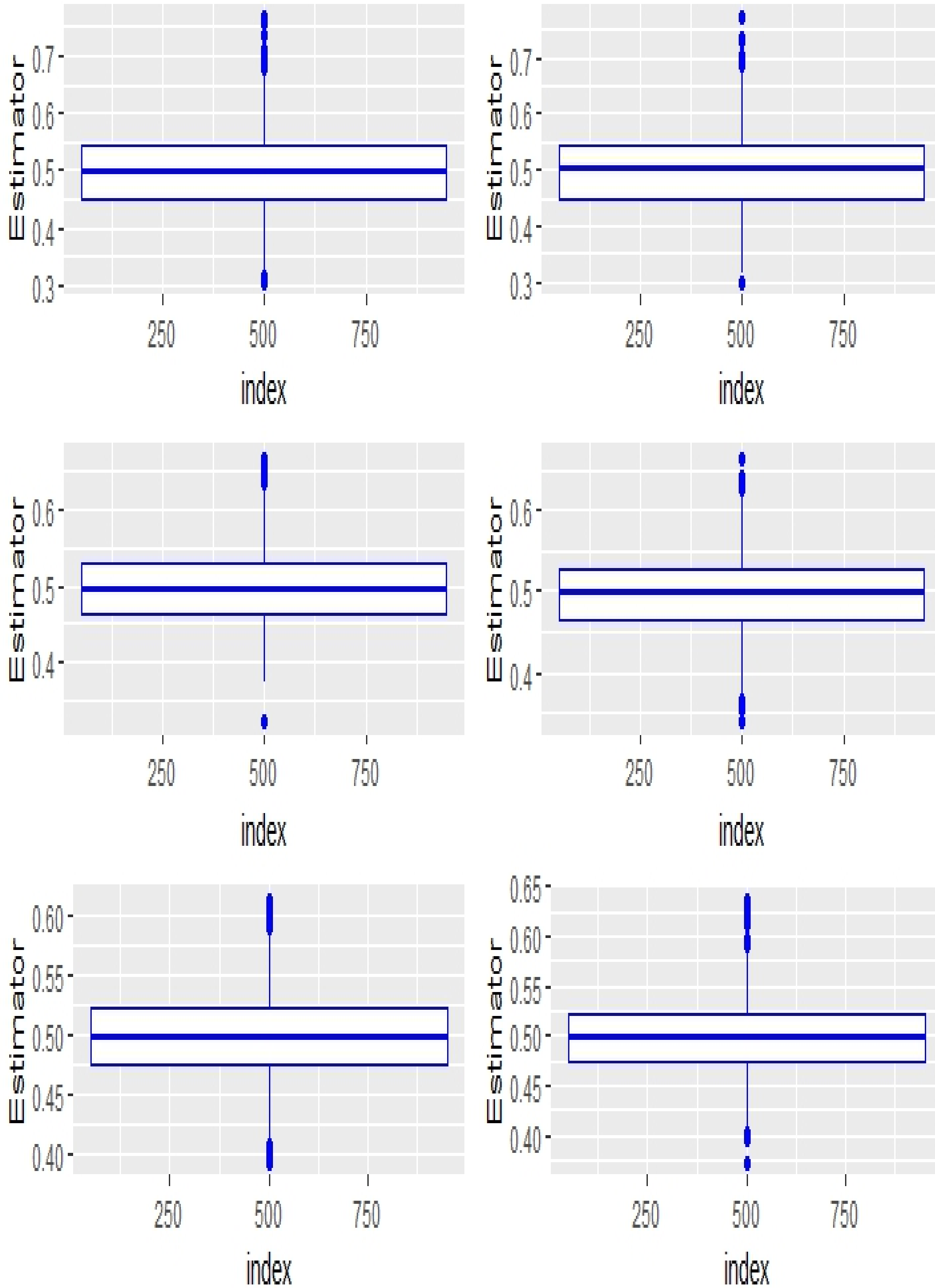

Figure 6.22: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). The true value of $1/\alpha$ is $1/2$. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

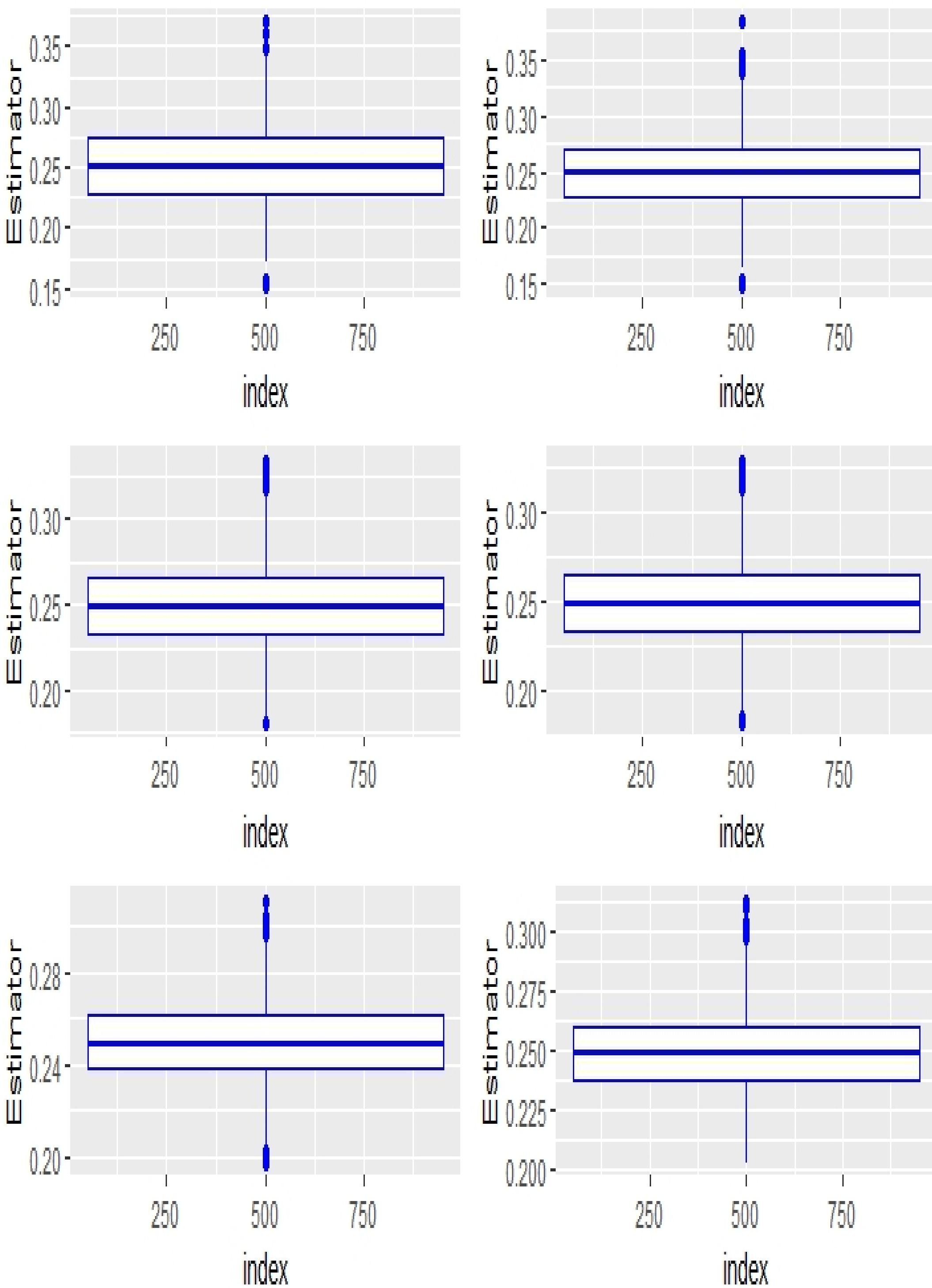

Figure 6.23: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). The true value of $1/\alpha$ is $1/4$. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

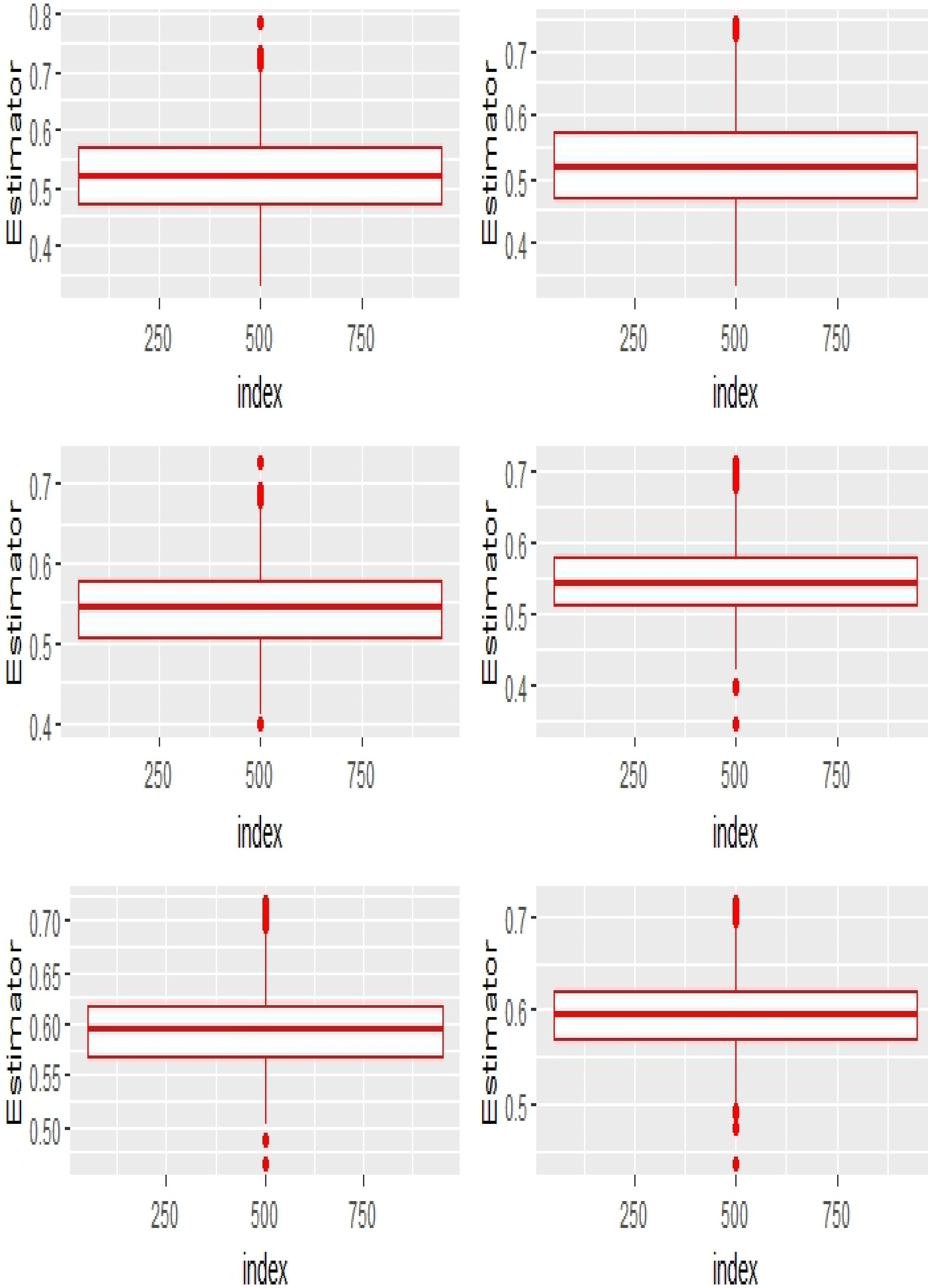

Figure 6.24: Boxplots for Hill estimator for LMSV model with Student noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). The true value of $1/\alpha$ is $1/2$. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

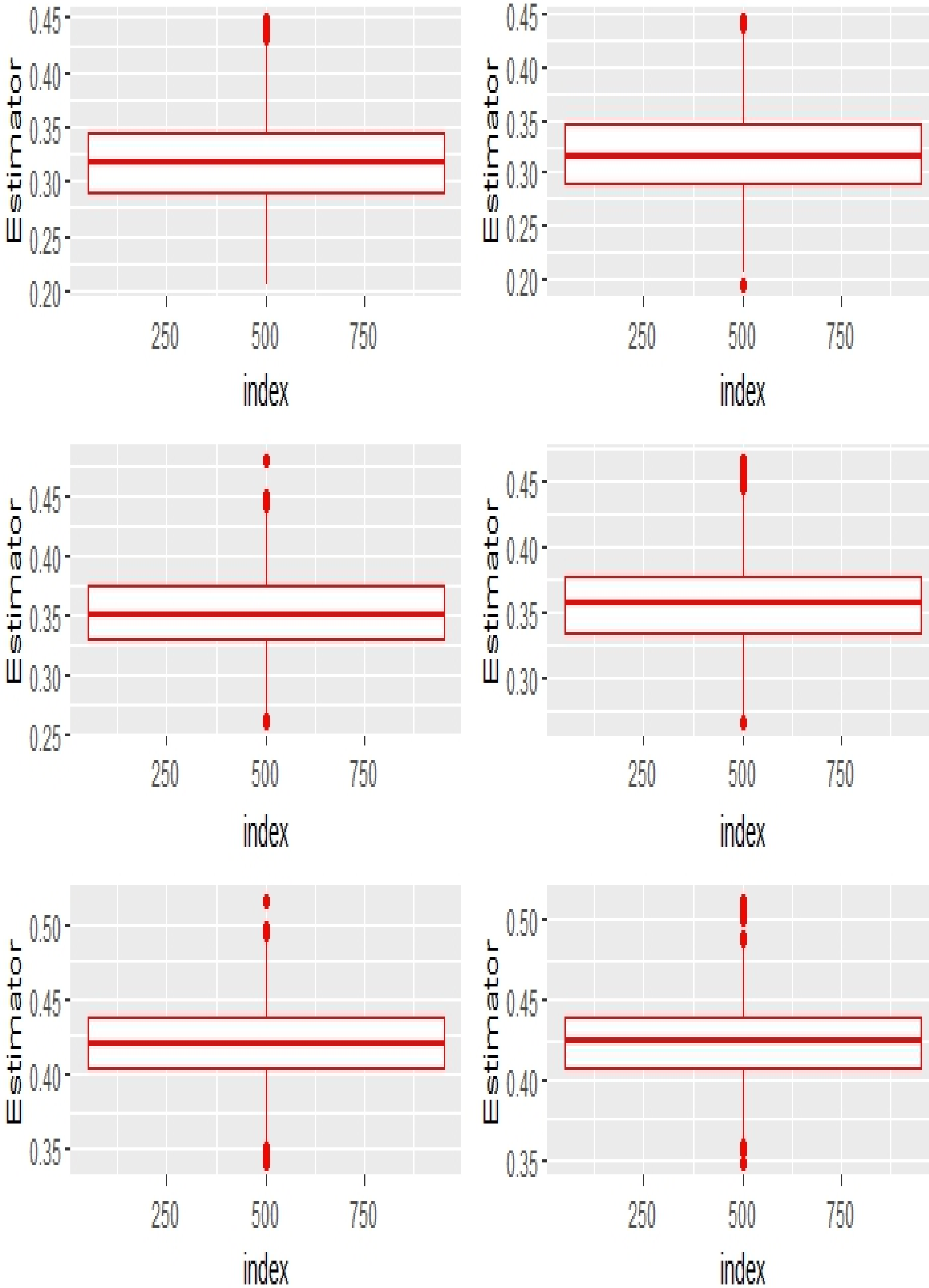

Figure 6.25: Boxplots for Hill estimator for LMSV model with Student noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). The true value of $1/\alpha$ is $1/4$. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

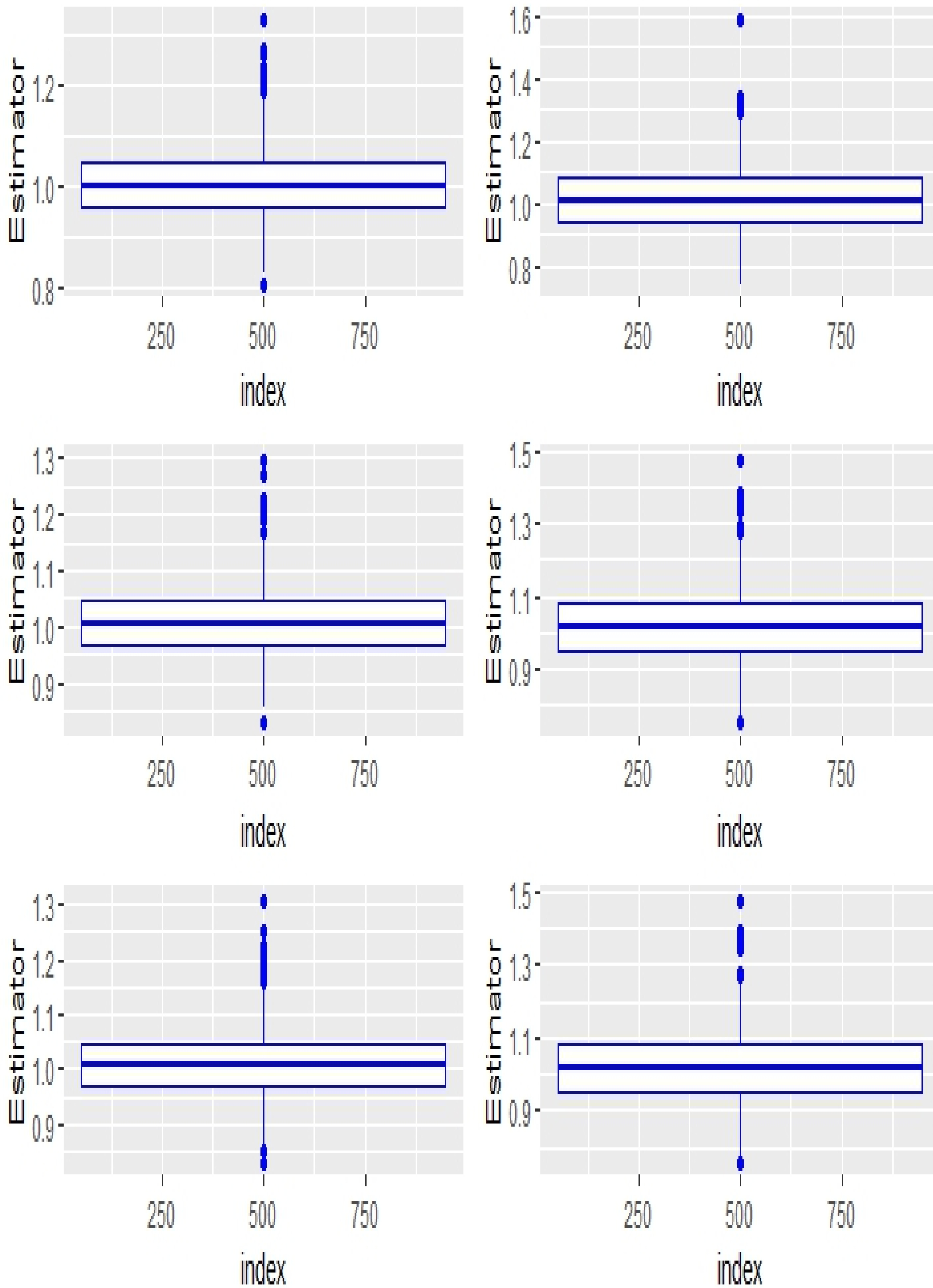

Figure 6.26: Boxplots for Value-At-Risk for LMSV model with Pareto noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05{*}n$ (top panel), $k = 0.1{*}n$ (middle panel) and $k = 0.2{*}n$ (bottom panel).

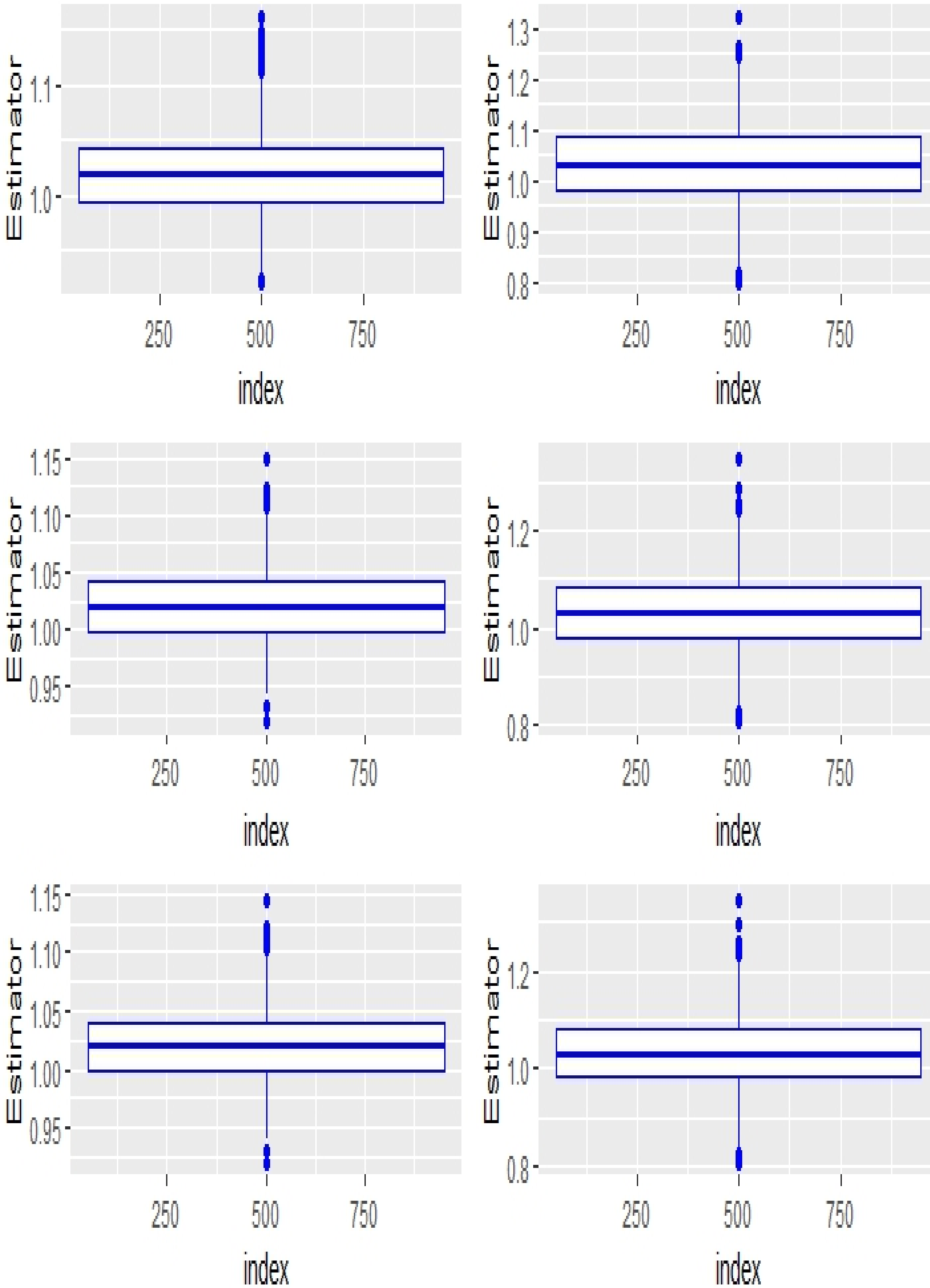

Figure 6.27: Boxplots for Value-At-Risk for LMSV model with Pareto noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05{*}n$ (top panel), $k = 0.1{*}n$ (middle panel) and $k = 0.2{*}n$ (bottom panel).

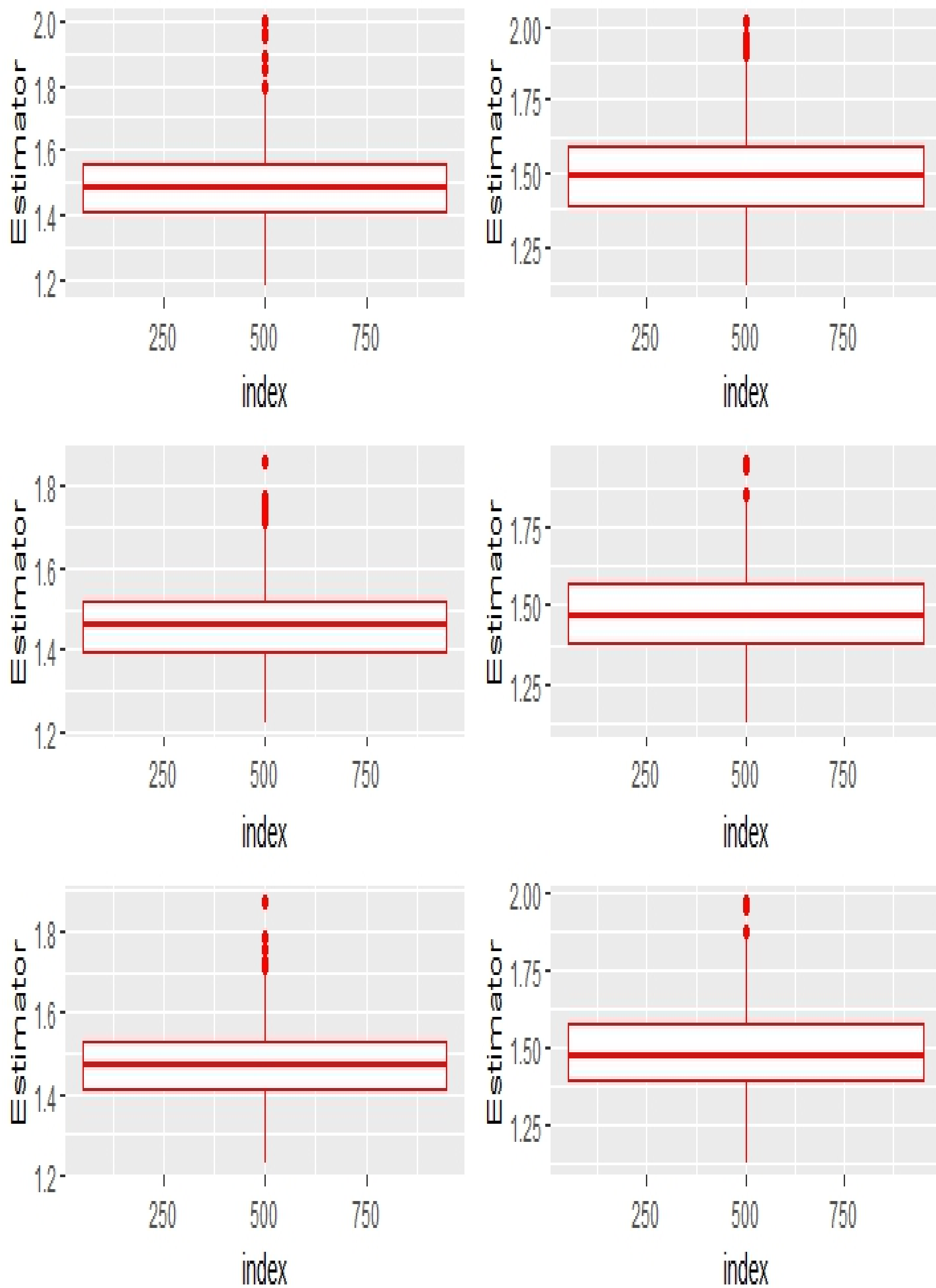

Figure 6.28: Boxplots for Value-At-Risk for LMSV model with Student noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05{*}n$ (top panel), $k = 0.1{*}n$ (middle panel) and $k = 0.2{*}n$ (bottom panel).

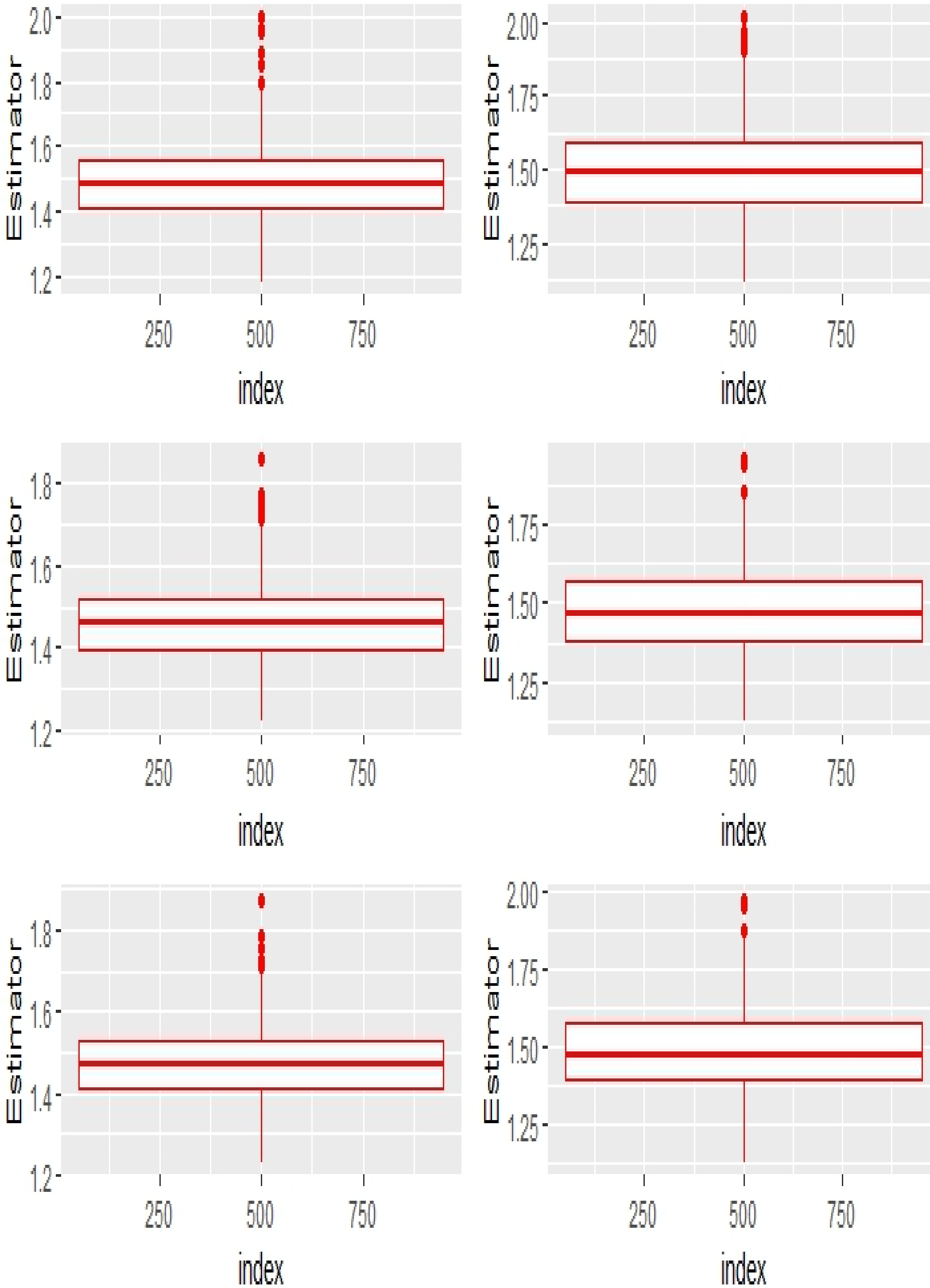

Figure 6.29: Boxplots for Value-At-Risk for LMSV model with Student noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05{*}n$ (top panel), $k = 0.1{*}n$ (middle panel) and $k = 0.2{*}n$ (bottom panel).

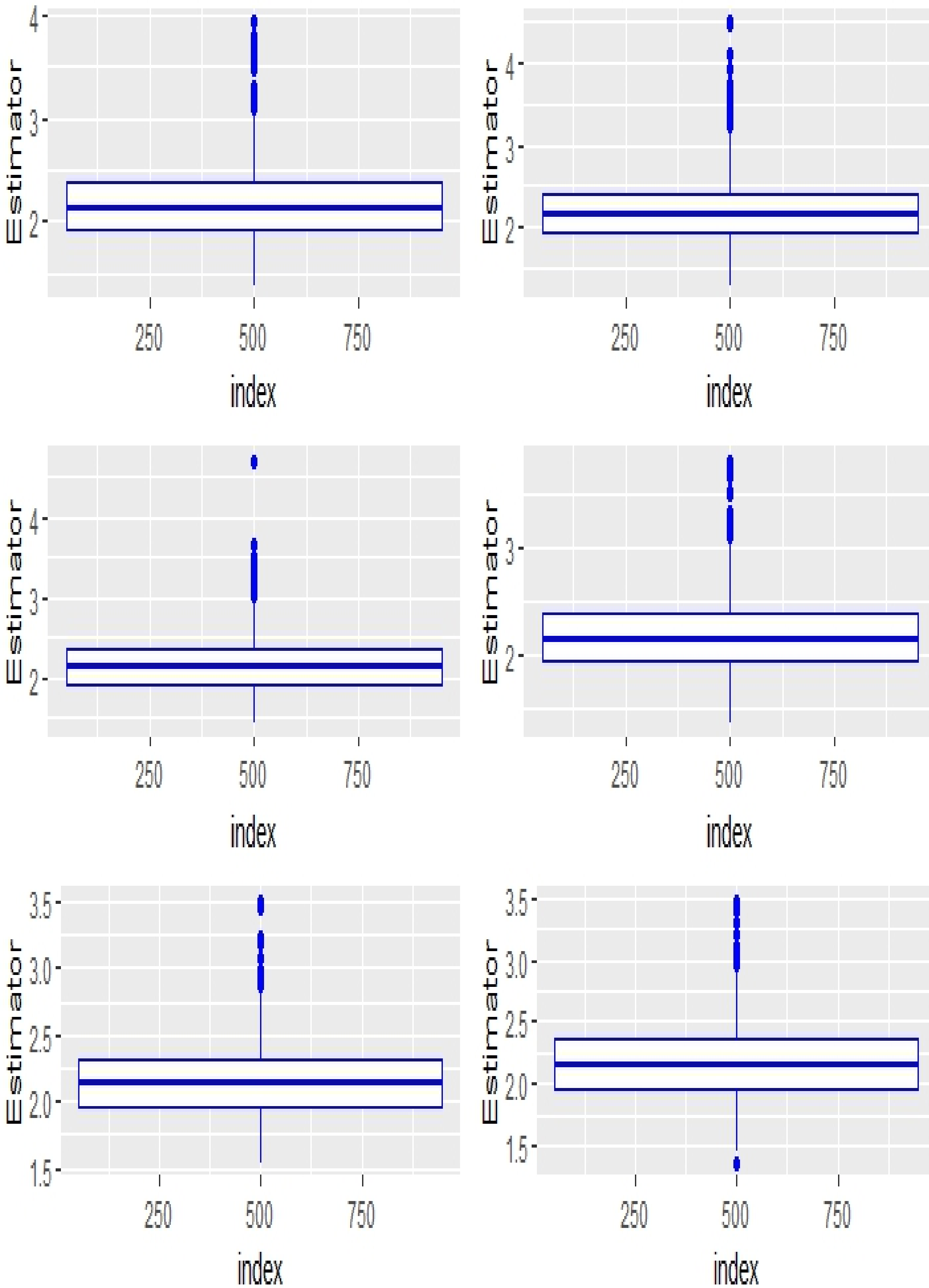

Figure 6.30: Boxplots for Expected Shortfall for LMSV model with Pareto noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

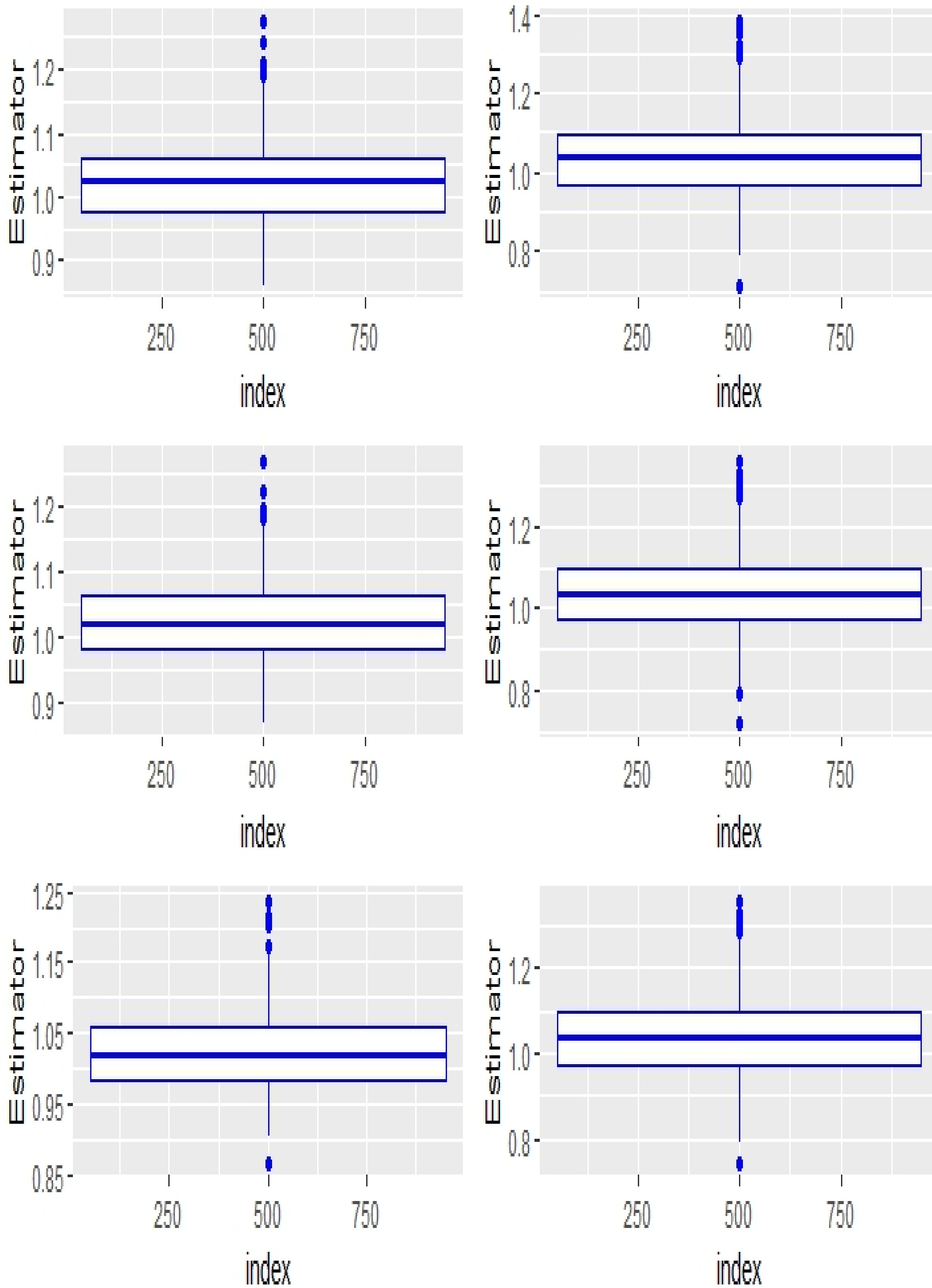

Figure 6.31: Boxplots for Expected Shortfall for LMSV model with Pareto noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

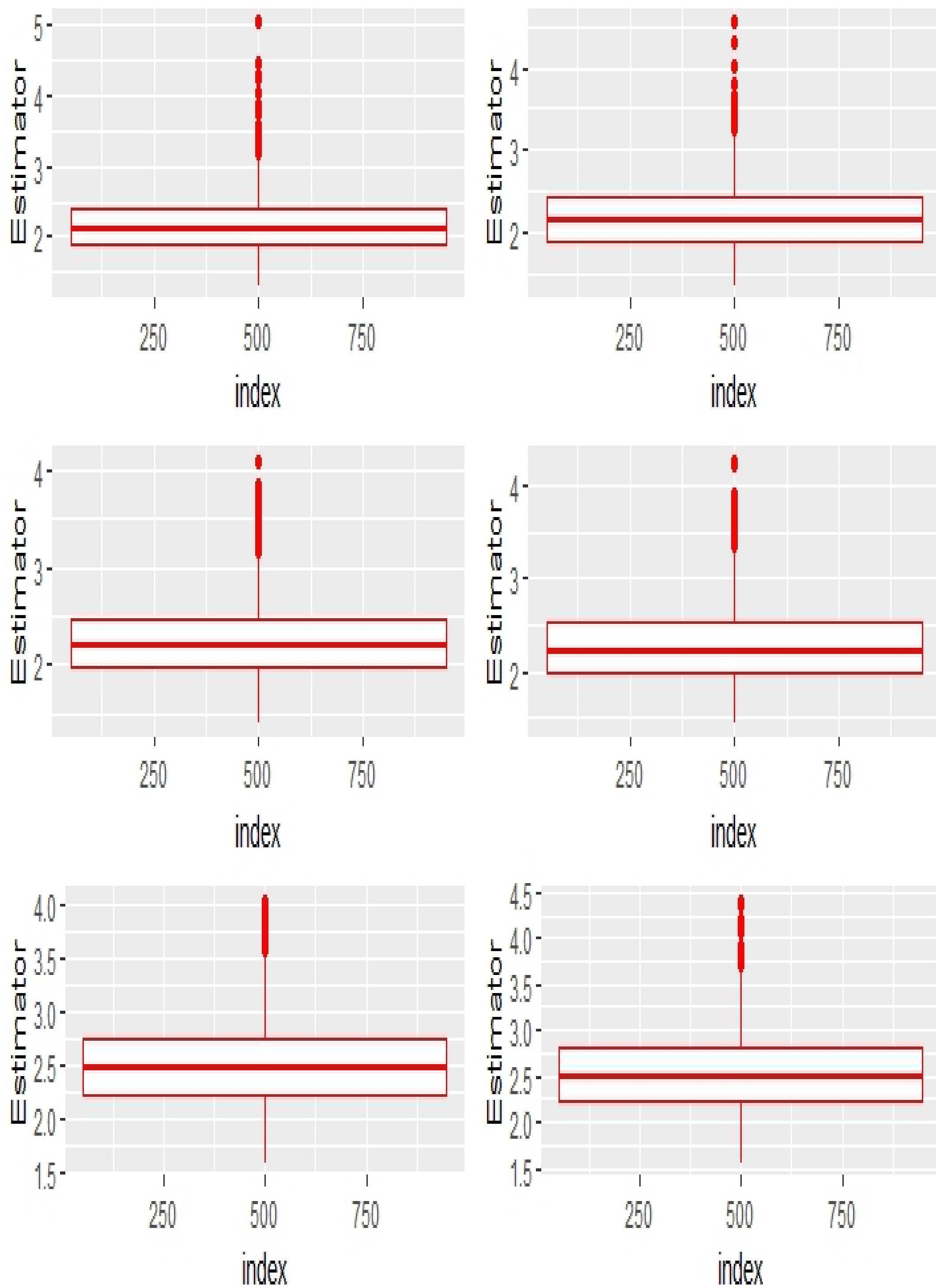

Figure 6.32: Boxplots for Expected Shortfall for LMSV model with Student noise with index $\alpha = 2$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

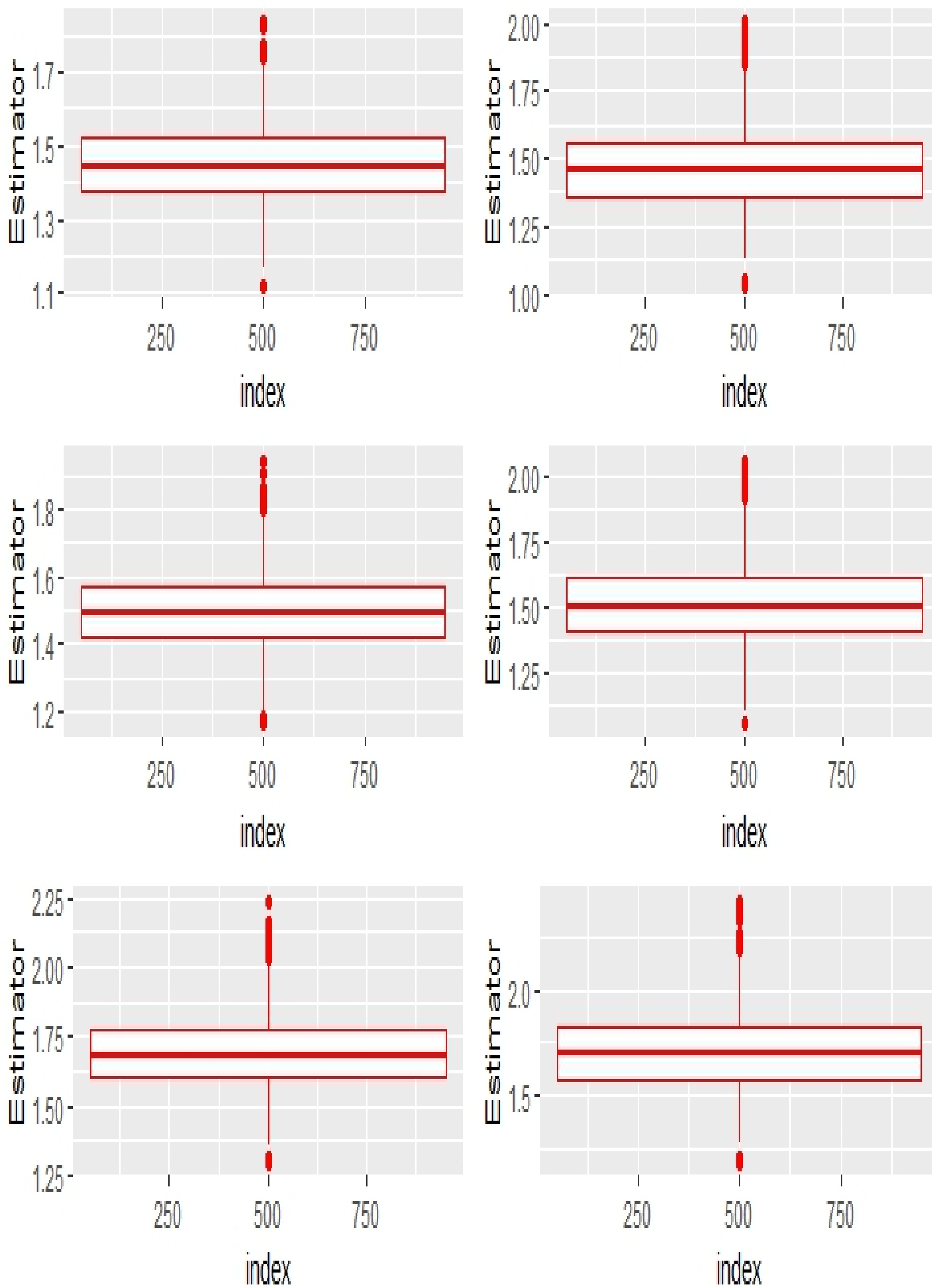

Figure 6.33: Boxplots for Expected Shortfall for LMSV model with Student noise with index $\alpha = 4$. The memory parameter is $d = 0.1$ (left panel) and $d = 0.4$ (right panel). Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

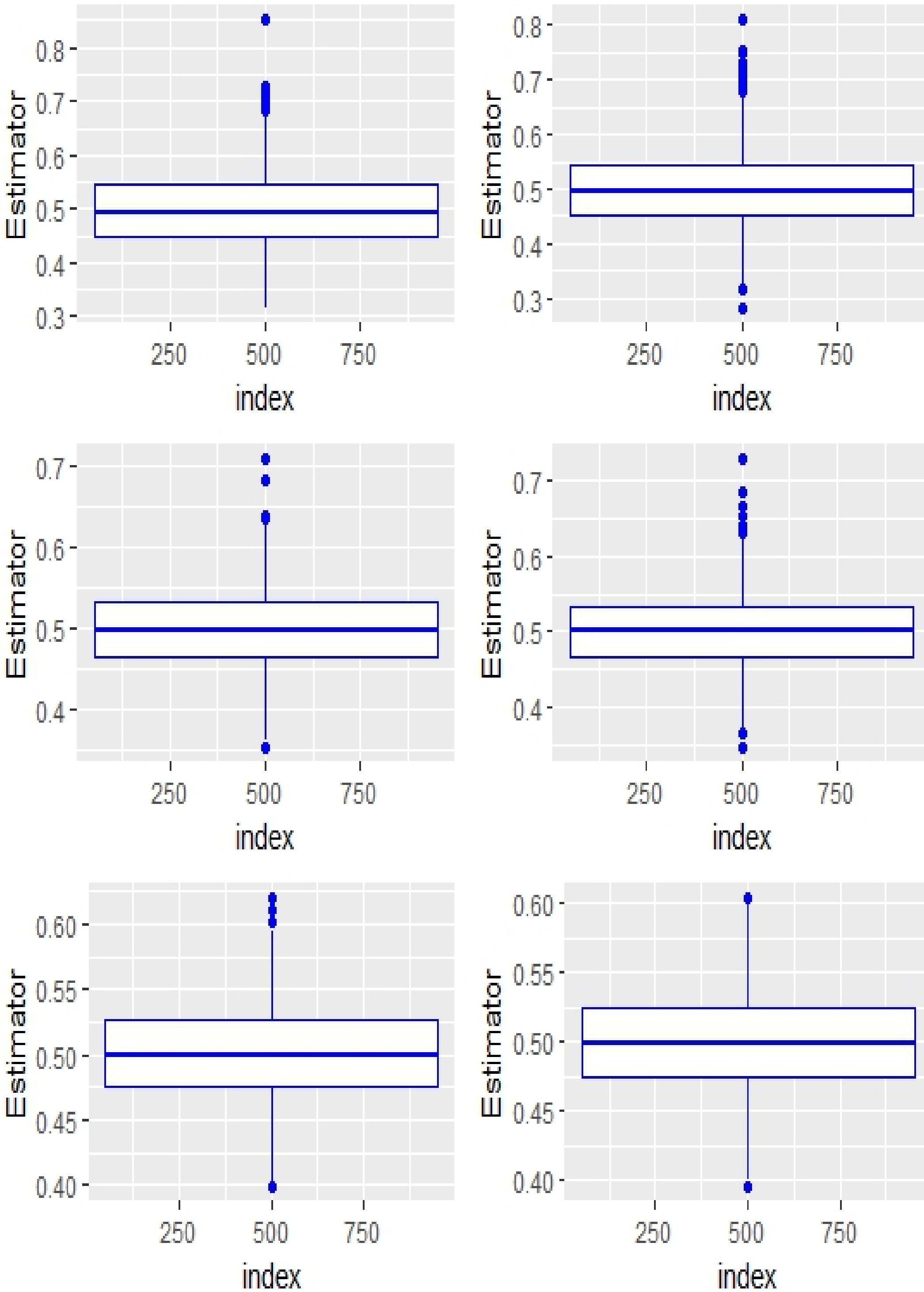

Figure 6.34: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 2$ and the memory parameter $d = 0.1$. Left panel: model without leverage; Right panel: model with leverage. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

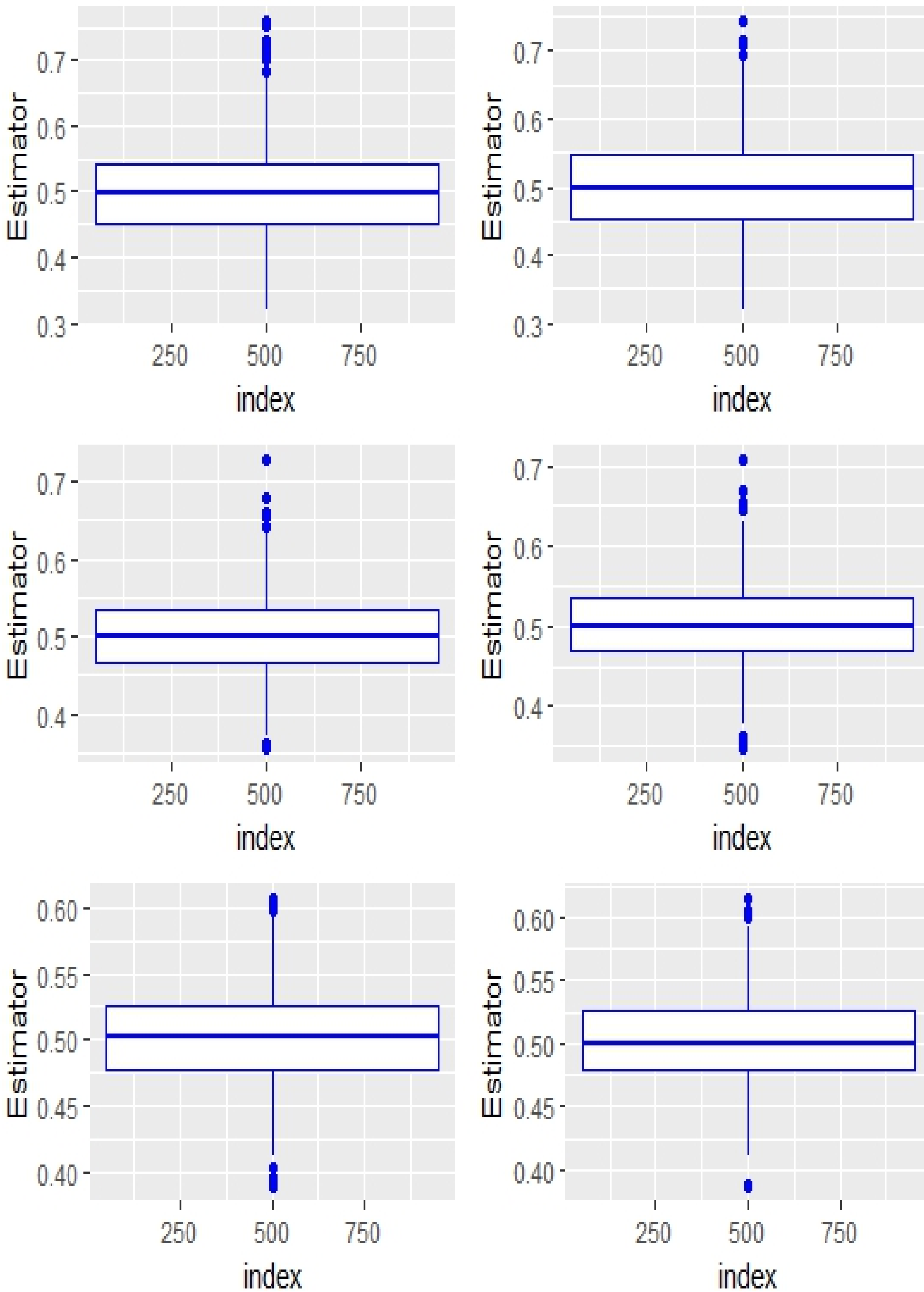

Figure 6.35: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 2$ and the memory parameter $d = 0.4$. Left panel: model without leverage; Right panel: model with leverage. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

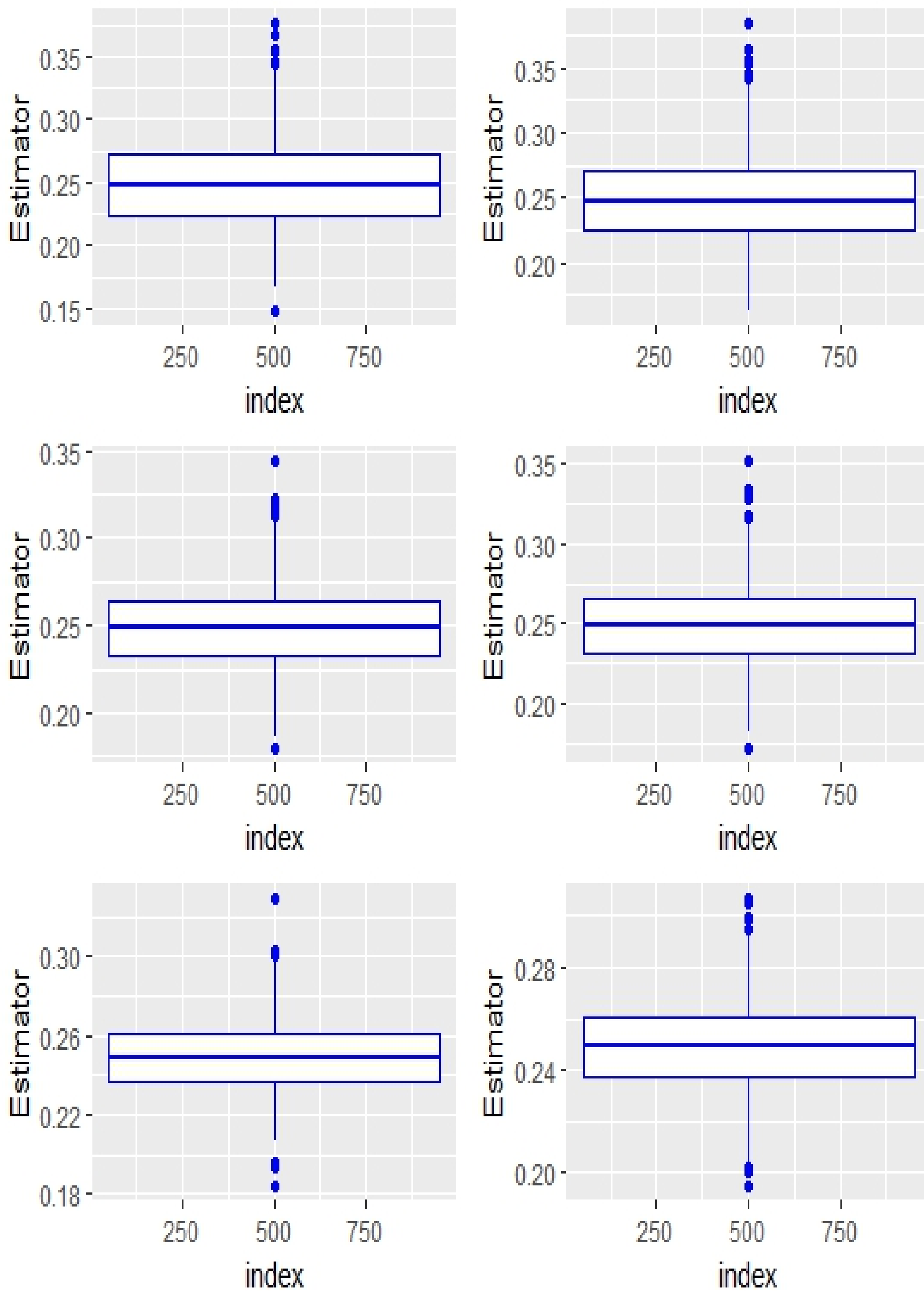

Figure 6.36: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 4$ and the memory parameter $d = 0.1$. Left panel: model without leverage; Right panel: model with leverage. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

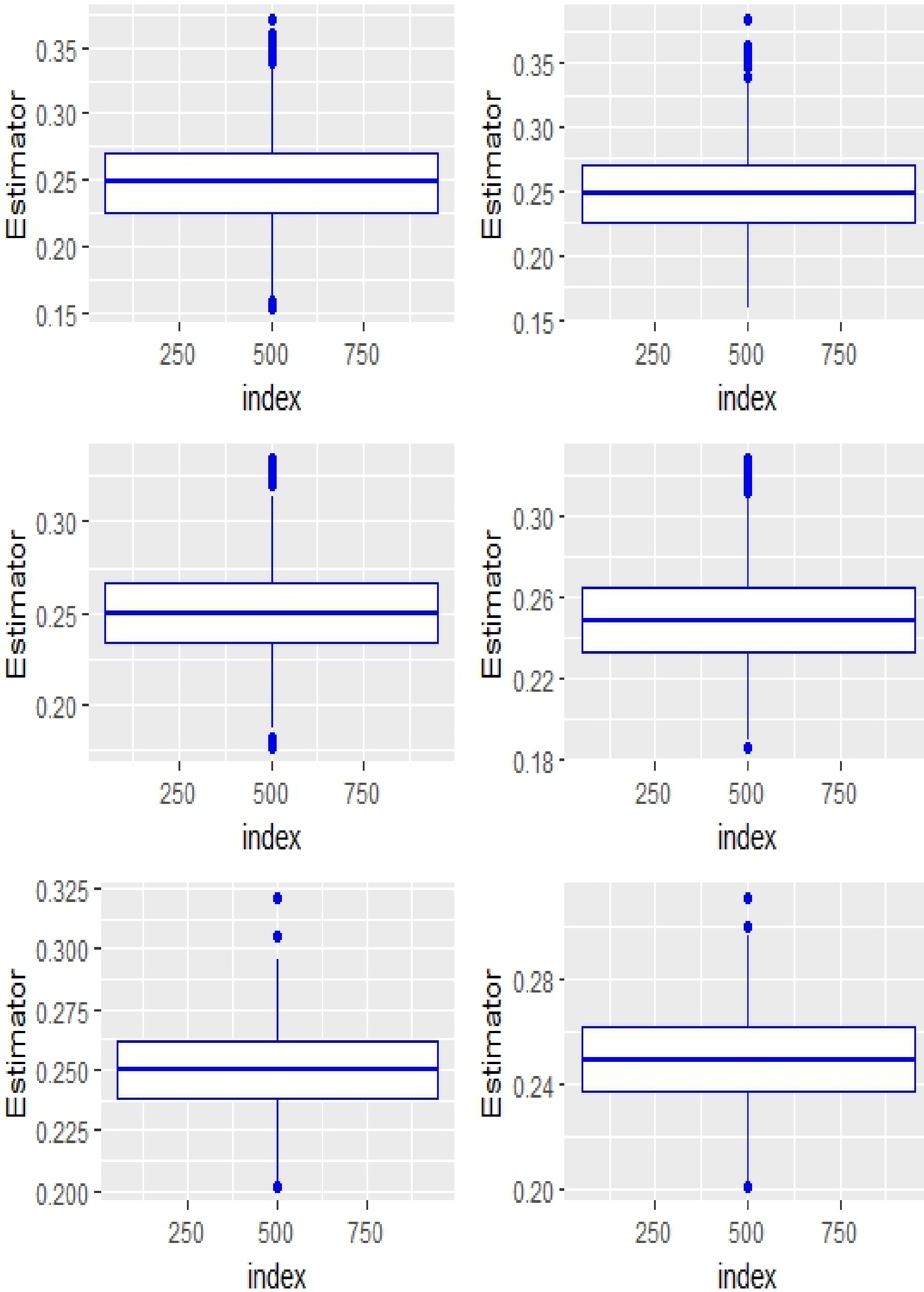

Figure 6.37: Boxplots for Hill estimator for LMSV model with Pareto noise with index $\alpha = 4$ and the memory parameter $d = 0.4$. Left panel: model without leverage; Right panel: model with leverage. Number of order statistics used: $k = 0.05 * n$ (top panel), $k = 0.1 * n$ (middle panel) and $k = 0.2 * n$ (bottom panel).

Chapter 7

Conclusion

There are various forms of risk (operational risk, liquidity risk, credit risk, etc.) that financial institutions deal with on a regular basis. In response to extreme events with adverse effects, financial institutions use risk management to ensure their resilience and solvency. In this regard, in this thesis we have focused on some quantitative aspects of risk management (Chapter 5). With the help of the asymptotic theory for tail empirical processes (Chapter 4), we have studied estimation of Value-at-Risk and Expected Shortfall under the assumptions that returns of a portfolio are heavy-tailed, long memory sequences with leverage as defined in (3.1). While estimation of both VaR and ES is unaffected by leverage, heavy-tails and long range dependence do influence the limiting behaviour in a dichotomous manner. These theoretical results have been illustrated in the simulation studies done in Chapter 6.

There are a number of questions about the assumptions of the LMSV model with leverage and some technicalities that we discuss below in section 7.1. Then, we wrap with some future research directions in section 7.2.

7.1 Comments on Assumptions and Technicalities

- The Gaussian assumption on Y_j can be easily replaced with Y_j being an infinite order moving average process. Instead of using Hermite polynomials, convergence of the long memory part can be concluded using tools such as Appell polynomials or a version of martingale approximation. See [6, Section 4.2.5].

- Second-order regular variation is needed to handle the bias induced by convergence in (3.10) and (3.12). In the i.i.d. case, instead of (3.6) it suffices to assume that $\sqrt{n\bar{F}_Z(u_n)}|\eta^*(u_n)| \to 0$ (cf. Theorem 3.3.5 in [23]; note that the bias condition (3.3.9) there is written in a different form). Here, due to dependence, we have the additional restriction.

Through the relationships $k_n = n\bar{F}_X(u_n)$ and (3.11), the bias assumption restricts the number of order statistics that can be used in the construction of the Hill estimator.

- We excluded the case of $d = 0$ which yields short memory. It is justified in [21] that in the case of short memory, the stochastic volatility sequence $\{X_j\}$ is mixing and limiting results for tail empirical processes can be concluded from [48].

- For clarity, throughout Chapters 3 to 5 we work under all assumptions introduced in Section 3.2. For some partial results not all the assumptions are needed. Indeed, for the tail empirical process with deterministic levels, instead of A(ii), only regular variation is needed, while the moment conditions (3.7a)-(3.7a) can be replaced with a weaker assumption, $E\left((\phi(Y))^{\alpha+\epsilon}\right) < \infty$ for some $\epsilon > 0$, in order to guarantee that the tail distribution $\bar{F}_X$ is regularly varying. For the tail empirical process with random levels and for the Hill estimator, a version of second-order regular variation is needed.

  In our method of proof, we utilize second-order regular variation for the tail empirical process with deterministic levels. Possibly, with another method of proof, this could be avoided.

- More specifically, for finite dimensional convergence of the martingale part (Proposition 4.2.7) the moment condition (3.7a) is not needed, but second-order regular variation plays a crucial role in the proof. The no-bias condition (3.6) is not used.

- Lemma 4.2.14 does not require any distributional assumption on Z. Also, the moment conditions are not needed. Lemma 4.2.15 requires the moment assumption (3.7a). Only regular variation of Z is needed; second-order regular variation is not required. Lemma 4.2.16 again requires only (3.7a). In summary, the proof of tightness of the martingale part (Proposition 4.2.12) requires only regular variation and the moment condition (3.7a).

- Thus, weak convergence of the martingale part requires all assumptions except for (3.6).

- Weak convergence of the long memory part (Theorem 4.2.17) needs (3.7b) and second order regular variation with (3.6).

7.2 Further Research Directions

There are several open questions that we intend to pursue:

- The model should be extended to multivariate data. We expect that the techniques developed in the thesis should be applicable.
- Bootstrap techniques, in particular for estimation of risk measures, should be developed. There are challenges stemming both from long memory and the delicate structure of the tail empirical process.
- The volatility sequence $\{\sigma_j\}$ could be extended to include heavy tails. This will require new methodology.
- Our methodology can be applied to inference problems for other families of risk measures such as convex risk measures.
- Techniques should be developed for the detection and estimation of a change point in the model.

Bibliography

[1] Carlo Acerbi and Dirk Tasche. Expected shortfall: A natural coherent alternative to value at risk. *Economic Notes*, 31(2):379–388, 2002.

[2] Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002.

[3] Yacine Aït-Sahalia, Jianqing Fan, and Yingying Li. The leverage effect puzzle: Disentangling sources of bias at high frequency. *Journal of Financial Economics*, 109(1):224–249, 2013.

[4] Philippe Artzner, Freddy Delbaen, JeanMarc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[5] Jan Beran. *Statistics for Long-Memory Processes*. Chapman and Hall, 1994.

[6] Jan Beran, Yuanhua Feng, Sucharita Ghosh, and Rafał Kulik. *Long Memory Processes: Probabilistic Properties and Statistical Methods*. Springer, 2013.

[7] Jan Beran, Dieter Schell, and Milan Stehlík. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1):193–220, 2014.

[8] Clémonell Bilayi-Biakana, Gail Ivanoff, and Rafał Kulik. The tail empirical process for long memory stochastic volatility models with leverage. *Electronic Journal of Statistics*, 2019. Accepted.

[9] Clémonell Bilayi-Biakana, Rafał Kulik, and Philippe Soulier. Statistical inference for heavy tailed series with extremal independence. arXiv:1804.10948, 2018.

[10] Patrick Billingsley. *Convergence of probability measures*. New York, Wiley, 1968.

[11] Patrick Billingsley. *Probability and Measure*. Wiley-Interscience, third edition, 1995.

[12] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

[13] Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels. *Regular variation*. Press Syndicate of the University of Cambridge, first edition, 1987.

[14] Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels. *Regular variation*. Cambridge University Press, Cambridge, 1989.

[15] F. Jay Breidt, Nuno Crato, and Pedro de Lima. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1-2):325–348, 1998.

[16] Carles Bretó. On idiosyncratic stochasticity of financial leverage effects. *Statistics & Probability Letters*, 91:20 – 26, 2014.

[17] Peter Breuer and Péter Major. Central limit theorems for nonlinear functionals of Gaussian fields. *Journal of Multivariate Analysis*, 13(3):425–441, 1983.

[18] Dariusz Buraczewski, Ewa Damek, and Thomas Mikosch. *Stochastic models with power-law tails: the equation $X = AX + B$*. Springer Series in Operations Research and Financial Engineering. Springer, 2016.

[19] Moorad Choudhry. *An Introduction to Value-at-Risk*. Wiley, 2013.

[20] Andrew A. Christie. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, 10(4):407–432, 1982.

[21] Richard A. Davis and Thomas Mikosch. Point process convergence of stochastic volatility processes with application to sample autocorrelation. *Journal of Applied Probability.*, 38A(93-104):93–104, 2001.

[22] Youri Davydov and Ricardas Zitikis. On weak convergence of random fields. *Annals of the Institute of Statistical Mathematics*, 60:345–365, 2008.

[23] Laurens de Haan and Ana Ferreira. *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. An introduction.

[24] Rohit Deo, Meng-Chen Hsieh, Clifford M. Hurvich, and Philippe Soulier. Long memory in nonlinear processes. In *Dependence in probability and statistics*, volume 187 of *Lecture Notes in Statist.*, pages 221–244. Springer, New York, 2006.

[25] Rohit Deo, Clifford M Hurvich, Philippe Soulier, and Yi Wang. Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory*, 25(3):764–792, 2009.

[26] Zhuanxin Ding, Clive W.J. Granger, and Robert F. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106, 1993.

[27] Kevin Dowd. *Measuring Market Risk*. Wiley, 2005.

[28] Holger Drees. Optimal rates of convergence for estimates of the extreme value index. *The Annals of Statistics*, 26(1):434–448, 1998.

[29] Holger Drees. Weighted approximations of tail processes for β-mixing random variables. *The Annals of Applied Probability*, 10(4):1274–1301, 2000.

[30] Holger Drees. Tail empirical processes under mixing conditions. In *Empirical process techniques for dependent data*, pages 325–342. Birkhäuser Boston, Boston, MA, 2002.

[31] Holger Drees. Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(4):617–657, 2003.

[32] Holger Drees and Holger Rootzén. Limit theorems for empirical processes of cluster functionals. *Annals of Statistics*, 38(4):2145–2186, 2010.

[33] Richard Durrett. *Probability Theory and Examples*. Cambridge University Press, 2010.

[34] John H. J. Einmahl. Limit theorems for tail processes with application to intermediate quantile estimation. *Journal of Statistical Planning and Inference*, 32(1):137–145, 1992.

[35] Clive W. J. Granger and Roselyne Joyeux. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29, 1980.

[36] Peter Hall and Chris C. Heyde. *Martingale limit theory and its application.* Academic Press, New York, 1980.

[37] Andrew C. Harvey and Neil Shephard. Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434, 1996.

[38] Inge S. Helland. Central limit theorems for martingales with discrete or continuous time. *Scandinavian Journal of Statistics*, 9:79–94, 1982.

[39] Rafał Kulik and Philippe Soulier. The tail empirical process for long memory stochastic volatility sequences. *Stochastic Processes and their Applications*, 121(1):109 – 134, 2011.

[40] Rafał Kulik and Philippe Soulier. Estimation of limiting conditional distributions for the heavy tailed long memory stochastic volatility process. *Extremes*, 16(2):203 – 239, 2013.

[41] Rafał Kulik and Philippe Soulier. Heavy tailed time series with extremal independence. *Extremes*, 18:273–299, 2015.

[42] Rafał Kulik, Philippe Soulier, and Olivier Wintenberger. The tail empirical process of regularly varying functions of geometrically ergodic markov chains. *Stochastic Processes and their Applications*, 2018. Accepted. DOI: j.spa.2018.11.014.

[43] Hicham Loukrati. Tail empirical processes: Limit theorems and bootstrap techniques, with applications to risk measures. https://ruor.uottawa.ca/handle/10393/37594, 2018. PhD thesis.

[44] Thomas Mikosch. Regular variation, subexponentiality and their applications in probability theory. *Technical Report*, 99-013, 1999.

[45] Frank W. J. Olver. *Handbook of Mathematical Functions.* Cambridge University Press : NIST, 2010.

[46] Wilfredo Palma. *Long-Memory Time Series : Theory and Methods.* Wiley series in probability and statistics. John Wiley & Sons, Inc, Hoboken, 2007.

[47] Sidney I. Resnick. *Heavy-Tail Phenomena:Probabilistic and Statistical Modelling.* Springer, 2007.

[48] Holger Rootzén. Weak convergence of the tail empirical process for dependent sequences. *Stochastic Processes and their Applications*, 119(2):468–490, 2009.

[49] Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications: With R Examples.* Springer, fourth edition, 2017.

[50] Philippe Soulier. Some applications of regular variation in probability and statistics. *XXII Escuela Venezolana de Mathematicas*, 2009.

[51] Ward Whitt. *Stochastic-process limits.* Springer-Verlag, New York, 2002.

[52] Jun Yu. On leverage in a stochastic volatility model. *Journal of Econometrics*, 127(2):165 – 178, 2005.