

# LIMIT THEOREMS FOR LONG-MEMORY STOCHASTIC VOLATILITY MODELS WITH INFINITE VARIANCE: PARTIAL SUMS AND SAMPLE COVARIANCES

RAFAŁ KULIK,\* *University of Ottawa*

PHILIPPE SOULIER,\*\* *Université Paris Ouest-Nanterre*

## Abstract

In this paper we extend the existing literature on the asymptotic behavior of the partial sums and the sample covariances of long-memory stochastic volatility models in the case of infinite variance. We also consider models with leverage, for which our results are entirely new in the infinite-variance case. Depending on the interplay between the tail behavior and the intensity of dependence, two types of convergence rates and limiting distributions can arise. In particular, we show that the asymptotic behavior of partial sums is the same for both long memory in stochastic volatility and models with leverage, whereas there is a crucial difference when sample covariances are considered.

*Keywords:* Heavy tail; long-range dependence; sample autocovariance; stochastic volatility

2010 Mathematics Subject Classification: Primary 60G55

Secondary 60F05; 62M10; 62P05

## 1. Introduction

One of the standardized features of financial data is that returns are uncorrelated, but their squares, or absolute values, are (highly) correlated, a property referred to as long memory (which will be defined precisely later). A second commonly accepted feature is that log-returns are heavy tailed, in the sense that some moment of the log-returns is infinite. The last feature we want to mention is leverage. In the financial time series context, leverage is understood to mean negative dependence between previous returns and future volatility (i.e. a large negative return will be followed by a high volatility). Motivated by these empirical findings, one of the common modeling approaches is to represent log-returns  $\{Y_i\}$  as a stochastic volatility sequence  $Y_i = Z_i\sigma_i$ , where  $\{Z_i\}$  is an independent and identically distributed (i.i.d.) sequence and  $\{\sigma_i^2\}$  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 modeled through the sequence  $\{\sigma_i\}$ , and the tails can be modeled either through the sequence  $\{Z_i\}$  or through  $\{\sigma_i\}$ , or both. The well-known GARCH processes belong to this class of models. The volatility sequence  $\{\sigma_i\}$  is heavy tailed, unless the distribution of  $Z_0$  has finite support, and leverage can be present. But, long memory in squares cannot be modeled by GARCH processes. The FIGARCH process was introduced in [3] for this purpose, but it is not known if it really has a long-memory property; see, e.g. [15].

---

Received 27 September 2011; revision received 27 April 2012.

\* Postal address: Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa ON, K1N 6N5, Canada. Email address: rkulik@uottawa.ca

\*\* Postal address: Département de Mathématiques, Université Paris Ouest-Nanterre, 200 Avenue de la République, 92000 Nanterre Cedex, France. Email address: philippe.soulier@u-paris10.fr

To model long memory in squares, the so-called *long memory in stochastic volatility* (LMSV) process was introduced in [8], generalizing the earlier short-memory version of this model. In this model, the sequences  $\{Z_i\}$  and  $\{\sigma_i\}$  are fully independent, and  $\{\sigma_i\}$  is the exponential of a Gaussian long-memory process. Tails and long memory are easily modeled in this way, but leverage is absent. Throughout the paper, we will refer to this process as LMSV, even though we do not rule out the short-memory case.

In order to model leverage, Nelson [26] introduced the EGARCH model (where E stands for exponential), later extended by Bollerslev and Mikkelsen [6] to the FIEGARCH model (where FI stands for fractionally integrated) in order to also model long memory. In these models,  $\{Z_i\}$  is a Gaussian white noise, and  $\{\sigma_i\}$  is the exponential of a linear process with respect to a function of the Gaussian sequence  $\{Z_i\}$ . Surgailis and Viano [32] extended the type of dependence between the sequences  $\{Z_i\}$  and  $\{X_i\}$  and relaxed the Gaussian assumption for both sequences, but assumed finite moments of all orders. Thus, long memory and leverage are possibly present in these models, but heavy tails are excluded.

A number of other models have been introduced, e.g. the models of Robinson and Zaffaroni [29], [30] and their further extensions in [28]; LARCH( $\infty$ ) processes [19] and their bilinear extensions [18], and LARCH<sub>+</sub>( $\infty$ ) [31]; to mention a few. All of these models have long memory and some have leverage and allow for heavy tails. The theory for these models is usually extremely involved, and only the asymptotic properties of partial sums are known in certain cases. We will not consider these models here. In [20] the leverage effect and long-memory property of a LARCH( $\infty$ ) model was studied thoroughly.

The theoretical effect of long memory is that the covariance of absolute powers of the returns  $\{Y_i\}$  is slowly decaying and nonsummable. This induces nonstandard limit theorems, such as convergence of the partial sum process to the fractional Brownian motion or finite-variance non-Gaussian processes or even Lévy processes. In practice, long memory is often evidenced by sample covariance plots, showing an apparent slow decay of the covariance function. Therefore, it is of interest to investigate the asymptotic behavior of the sample mean or of the partial sum process, and of the sample variance and covariances.

In the case where  $\sigma_i = \sigma(X_i)$ ,  $\{X_i\}$  is a stationary Gaussian process with summable covariances and  $\sigma(x) = \exp(x)$ , the asymptotic theory for the sample mean of LMSV processes with infinite variance is a straightforward consequence of a point process convergence result in [11]. The limit is a Lévy stable process. Surgailis and Viano [32] considered the convergence of the partial sum process of absolute powers of generalized EGARCH processes with finite moments of all orders and showed convergence to the fractional Brownian motion. To the best of the authors' knowledge, the partial sum process of absolute powers has never been studied in the context of heavy tails and long memory and possible leverage for a general function  $\sigma$ .

The asymptotic theory for sample covariances of weakly dependent stationary processes with finite moments dates back to Anderson; see [1]. The case of linear processes with regularly varying innovations was studied in [13] and [14] for infinite-variance innovation and for innovations with finite variance but infinite fourth moment, respectively. The limiting distribution of the sample covariances (suitably centered and normalized) is then a stable law. These results were obtained under conditions that rule out long memory. For infinite-variance innovation with tail index  $\alpha \in (1, 2)$ , these results were extended to long-memory linear processes in [24]. The limiting distributions of the sample covariances are again stable laws. However, if  $\alpha \in (2, 4)$ , Giraitis and Surgailis [21] showed that, as for partial sums, a dichotomy appears: the limiting distribution and the rate of convergence depend on an interplay between a memory parameter and the tail index  $\alpha$ . The limit is either stable (as in the weakly dependent

or i.i.d. case) or, if the memory is strong enough, the limiting distribution is non-Gaussian but with finite variance (the so-called Hermite–Rosenblatt distributions). If the fourth moment is finite then the dichotomy is between Gaussian or finite-variance non-Gaussian distributions (again of Hermite–Rosenblatt type); see [21, Theorem 3.3], [22], and [34].

The asymptotic properties of sample autocovariances of GARCH processes have been studied in [4]. Stable limits arise as soon as the marginal distribution has an infinite fourth moment. Davis and Mikosch [11] studied the sample covariance of a zero-mean stochastic volatility process, under implicit conditions that rule out long memory, and also found stable limits. McElroy and Politis [25] (generalized by Jach *et al.* [23]) studied partial sums and sample variance of a possibly nonzero-mean stochastic volatility process with infinite variance and where the volatility is a Gaussian long-memory process (in which case it is not positive, but this is not important for the theoretical results). They obtained a dichotomy between stable and finite-variance non-Gaussian limits, and also the surprising result that when the sample mean has a long-memory-type limit, then the studentized sample mean converges in probability to 0.

The first aim of this paper is to study asymptotic properties of partial sums, sample variance, and covariances of stochastic volatility processes where the volatility is an arbitrary function of a Gaussian, possibly long-memory process  $\{X_i\}$  independent of the sequence  $\{Z_i\}$ , which is a heavy-tailed i.i.d. sequence. We refer to these processes as LMSV processes. The interest of considering other functions than the exponential function is that it allows us to have other distributions than the log-normal for the volatility, while keeping the convenience of Gaussian processes, without which dealing with long-memory processes rapidly becomes extremely involved or even intractable. The results we obtain extend in various aspects all the previous literature in this domain.

Another important aim of the paper is to consider models with possible leverage. To this end, we need to give precise assumptions on the nature of the dependence between the sequences  $\{Z_i\}$  and  $\{X_i\}$ , and since they are related in the process  $\{Y_i\}$  through the function  $\sigma$ , these assumptions also involve the function  $\sigma$ . We have not looked for the widest generality, but the functions  $\sigma$  that we consider include the exponential functions and all symmetric polynomials with positive coefficients. This is not a severe restriction since the function  $\sigma$  must be nonnegative. Whereas the asymptotic theory for the partial sums is entirely similar to the case of the LMSV process without leverage, asymptotic properties of sample autocovariances may be very different in the presence of leverage. Due to the dependence between the two sequences, the rates of convergence and asymptotic distribution may be entirely different when not stable.

The paper is organized as follows. In Section 2 we formulate proper assumptions, as well as prove some preliminary results on the marginal and multivariate tail behavior of the sequence  $\{Y_i\}$ . In Section 3 we establish the limit theory for a point process based on the rescaled sequence  $\{Y_i\}$ . This methodology was first used in this context by Davis and Mikosch [11] and our proofs are closely related to those in this reference. In Section 4 we apply these results to obtain the functional asymptotic behavior of the partial sum process of the sequences  $\{Y_i\}$  and of its powers. In Section 5 the limiting behavior of the sample covariances and autocorrelation of the process  $\{Y_i\}$  and of its powers is investigated. Proofs are given in Section 6. In Appendix A we recall some results on multivariate Gaussian processes with long memory.

### 1.1. A note on the terminology

We consider in this paper sequences  $\{Y_i\}$  which can be expressed as  $Y_i = Z_i\sigma(X_i) = Z_i\sigma_i$ , where  $\{Z_i\}$  is an i.i.d. sequence and  $Z_i$  is independent of  $X_i$  for each  $i$ . Originally, SV and LMSV processes referred to processes where the sequences  $\{Z_i\}$  and  $\{\sigma_i\}$  are fully independent,

$\sigma_i = \sigma(X_i)$ ,  $\{X_i\}$  is a Gaussian process, and  $\sigma(x) = \exp(x)$ ; see, e.g. [8], [7], and [11]. The names EGARCH and FIEGARCH, respectively introduced in [26] and [6], refer to the case where  $\sigma(x) = \exp(x)$  and where  $\{X_i\}$  is a non-Gaussian process which admits a linear representation with respect to an instantaneous function of the Gaussian i.i.d. sequence  $\{Z_i\}$ , with dependence between the sequences  $\{Z_i\}$  and  $\{X_i\}$ . Surgailis and Viano [32] considered the case  $\sigma(x) = \exp(x)$ , but relaxed the assumptions on  $\{Z_i\}$  and  $\{X_i\}$ , and retained the name EGARCH. The LMSV processes can be seen as border cases of EGARCH-type processes, where the dependence between the sequences  $\{Z_i\}$  and  $\{X_i\}$  vanishes. In this paper we consider both LMSV models, and models with leverage which generalize the EGARCH models as defined in [32]. In order to refer to the latter models, we have chosen not to use the acronym EGARCH or FIEGARCH, since these models were defined with very precise specifications and this could create some confusion, nor to create a new one such as GEGARCH (with G standing twice for generalized, which seems a bit too much) or (IV)LMSVwL (for (possibly) infinite-variance long-memory stochastic volatility with leverage). Considering that the main feature which distinguishes these two classes of models is the presence or absence of leverage, we decided to refer to LMSV models when leverage is excluded, and to models with leverage when we include the possibility thereof.

**2. Model description, assumptions, and tail behaviour**

Let  $\{Z_i, i \in \mathbb{Z}\}$  be an i.i.d. sequence whose marginal distribution has regularly varying tails:

$$\lim_{x \rightarrow +\infty} \frac{P(Z_0 > x)}{x^{-\alpha} L(x)} = \beta, \quad \lim_{x \rightarrow +\infty} \frac{P(Z_0 < -x)}{x^{-\alpha} L(x)} = 1 - \beta. \tag{2.1}$$

Here  $\alpha > 0$ ,  $L$  is slowly varying at  $\infty$ , and  $\beta \in [0, 1]$ . Condition (2.1) is referred to as the balanced tail condition. It is equivalent to assuming that  $P(|Z_0| > x) = x^{-\alpha} L(x)$  and

$$\beta = \lim_{x \rightarrow +\infty} \frac{P(Z_0 > x)}{P(|Z_0| > x)} = 1 - \lim_{x \rightarrow +\infty} \frac{P(Z_0 < -x)}{P(|Z_0| > x)}.$$

We will say that two random variables  $Y$  and  $Z$  are right-tail equivalent if there exists  $c \in (0, \infty)$  such that

$$\lim_{x \rightarrow +\infty} \frac{P(Y > x)}{P(Z > x)} = c.$$

If one of the random variables has a regularly varying right tail then so has the other, with the same tail index. The converse is false, i.e. two random variables can have the same tail index without being tail equivalent. Two random variables  $Y$  and  $Z$  are said to be left-tail equivalent if  $-Y$  and  $-Z$  are right-tail equivalent, and they are said to be tail equivalent if they are both left- and right-tail equivalent.

Under (2.1), if, moreover,  $E[|Z_0|^\alpha] = \infty$  then  $Z_1 Z_2$  is regularly varying and (see, e.g. [14, Equation (3.5)])

$$\lim_{x \rightarrow +\infty} \frac{P(Z_0 > x)}{P(Z_0 Z_1 > x)} = 0, \quad \lim_{x \rightarrow +\infty} \frac{P(Z_1 Z_2 > x)}{P(|Z_1 Z_2| > x)} = \beta^2 + (1 - \beta)^2.$$

For example, if (2.1) holds and the tail of  $|Z_0|$  has Pareto-type tails, i.e.  $P(|Z_0| > x) \sim cx^{-\alpha}$  as  $x \rightarrow +\infty$  for some  $c > 0$ , then  $E[|Z_0|^\alpha] = \infty$ . We will further assume that  $\{X_i\}$  is a stationary zero-mean, unit-variance Gaussian process which admits a linear representation with respect

to an i.i.d. Gaussian white noise  $\{\eta_i\}$  with zero mean and unit variance, i.e.

$$X_i = \sum_{j=1}^{\infty} c_j \eta_{i-j} \tag{2.2}$$

with  $\sum_{j=1}^{\infty} c_j^2 = 1$ . We assume that the process  $\{X_i\}$  either has short memory, in the sense that its covariance function is absolutely summable, or exhibits long memory with Hurst index  $H \in (\frac{1}{2}, 1)$ , i.e. its covariance function  $\{\rho_n\}$  satisfies

$$\rho_n = \text{cov}(X_0, X_n) = \sum_{j=1}^{\infty} c_j c_{j+n} = n^{2H-2} \ell(n),$$

where  $\ell$  is a slowly varying function.

Let  $\sigma$  be a deterministic, nonnegative, and continuous function defined on  $\mathbb{R}$ . Define  $\sigma_i = \sigma(X_i)$  and the stochastic volatility process  $\{Y_i\}$  by

$$Y_i = \sigma_i Z_i = \sigma(X_i) Z_i.$$

At this point we do not assume independence of  $\{\eta_i\}$  and  $\{Z_i\}$ . We will deal with the following two special cases.

- The LMSV model, where  $\{\eta_i\}$  and  $\{Z_i\}$  are independent.
- The model with leverage, where  $\{(\eta_i, Z_i)\}$  is a sequence of i.i.d. random vectors. For fixed  $i$ ,  $Z_i$  and  $X_i$  are independent, but  $X_i$  may not be independent of the past  $\{Z_j, j < i\}$ .

Both cases are encompassed in the following assumption which will be in force throughout the paper.

**Assumption 2.1.** *The stochastic volatility process  $\{Y_i\}$  is defined by*

$$Y_i = \sigma_i Z_i,$$

where  $\sigma_i = \sigma(X_i)$ ,  $\{X_i\}$  is a Gaussian linear process with respect to the i.i.d. sequence  $\{\eta_i\}$  of standard Gaussian random variables such that (2.2) holds,  $\sigma$  is a nonnegative function such that  $P(\sigma(a\eta_0) > 0) = 1$  for all  $a \neq 0$ ,  $\{(Z_i, \eta_i)\}$  is an i.i.d. sequence, and  $Z_0$  satisfies the balanced tail condition (2.1) with  $E[|Z_0|^\alpha] = \infty$ .

Let  $\mathcal{F}_i$  be the sigma-field generated by  $(\eta_j, Z_j)$ ,  $j \leq i$ . Then the following properties hold.

- $Z_i$  is  $\mathcal{F}_i$ -measurable and independent of  $\mathcal{F}_{i-1}$ .
- $X_i$  and  $\sigma_i$  are  $\mathcal{F}_{i-1}$ -measurable.

We will also impose the following condition on the continuous function  $\sigma$ . There exists  $q > 0$  such that

$$\sup_{0 \leq \gamma \leq 1} E[\sigma^q(\gamma X_0)] < \infty. \tag{2.3}$$

It is clearly fulfilled for all  $q$  and  $q'$  if  $\sigma$  is a polynomial or  $\sigma(x) = \exp(x)$  and  $X_0$  is a standard Gaussian random variable. Note that if (2.3) holds for some  $q > 0$  then, for  $q' \leq q/2$ , it holds that

$$\sup_{0 \leq \gamma \leq 1} E[\sigma^{q'}(\gamma X_0) \sigma^{q'}(\gamma X_s)] < \infty, \quad s = 1, 2, \dots$$

**2.1. Marginal tail behavior**

If (2.3) holds then, clearly,  $E[\sigma^q(X_0)] < \infty$ . If, moreover,  $q > \alpha$ , since  $X_i$  and  $Z_i$  are independent for fixed  $i$ , it follows from Breiman’s lemma (see, e.g. [27, Proposition 7.5]) that the distribution of  $Y_0$  is regularly varying and

$$\lim_{x \rightarrow +\infty} \frac{P(Y_0 > x)}{P(Z_0 > x)} = \lim_{x \rightarrow +\infty} \frac{P(Y_0 < -x)}{P(Z_0 < -x)} = E[\sigma^\alpha(X_0)].$$

Thus, we see that there is no effect of leverage on marginal tails. Define

$$a_n = \inf \left\{ x : P(|Y_0| > x) < \frac{1}{n} \right\}. \tag{2.4}$$

Then the sequence  $a_n$  is regularly varying at  $\infty$  with index  $1/\alpha$ . Moreover, since  $\sigma$  is nonnegative,  $Z_0$  and  $Y_0$  have the same skewness, i.e.

$$\lim_{n \rightarrow +\infty} n P(Y_0 > a_n) = 1 - \lim_{n \rightarrow +\infty} n P(Y_0 < -a_n) = \beta.$$

**2.2. Joint exceedances**

One of the properties of heavy-tailed stochastic volatility models is that large values do not cluster. Mathematically, for all  $h > 0$ ,

$$P(|Y_0| > x, |Y_h| > x) = o(P(|Y_0| > x)). \tag{2.5}$$

For the LMSV model, conditioning on  $\sigma_0$  and  $\sigma_h$  yields

$$\lim_{x \rightarrow +\infty} \frac{P(|Y_0| > x, |Y_h| > x)}{P^2(|Z_0| > x)} = E[(\sigma_0 \sigma_h)^\alpha],$$

if (2.3) holds for some  $q > 2\alpha$ . Property (2.5) still holds when leverage is present. Indeed, let  $F_Z$  denote the distribution function of  $Z_0$  and let  $\bar{F}_Z = 1 - F_Z$ . Recall that  $\mathcal{F}_{h-1}$  is the sigma-field generated by  $(\eta_j, Z_j)$ ,  $j \leq h - 1$ . Thus,  $Y_0$  and  $X_h$  are measurable with respect to  $\mathcal{F}_{h-1}$ , and  $Z_h$  is independent of  $\mathcal{F}_{h-1}$ . Conditioning on  $\mathcal{F}_{h-1}$  yields

$$P(Y_0 > x, Y_h > x) = E \left[ \bar{F}_Z \left( \frac{x}{\sigma_h} \right) \mathbf{1}_{\{Y_0 > x\}} \right].$$

Next, fix some  $\varepsilon > 0$ . Applying Lemma 6.2, there exists a constant  $C$  such that, for all  $x \geq 1$ ,

$$\frac{P(Y_0 > x, Y_h > x)}{P(Z_0 > x)} = E \left[ \frac{\bar{F}_Z(x/\sigma_h)}{\bar{F}_Z(x)} \mathbf{1}_{\{Y_0 > x\}} \right] \leq C E[(1 \vee \sigma_h)^{\alpha+\varepsilon} \mathbf{1}_{\{Y_0 > x\}}].$$

If (2.3) holds for some  $q > \alpha$ , and  $\varepsilon$  is chosen small enough so that  $\alpha + \varepsilon < q$ , then, by bounded convergence, the latter expression is finite and converges to 0 as  $x \rightarrow +\infty$ .

**2.3. Products**

For the LMSV model, another application of Breiman’s lemma reveals that  $Y_0 Y_h$  is regularly varying for all  $h$ . If (2.3) holds for some  $q > 2\alpha$  then

$$\lim_{x \rightarrow +\infty} \frac{P(Y_0 Y_h > x)}{P(Z_0 Z_1 > x)} = E[(\sigma_0 \sigma_h)^\alpha], \quad \lim_{x \rightarrow +\infty} \frac{P(Y_0 Y_h < -x)}{P(Z_0 Z_1 < -x)} = E[(\sigma_0 \sigma_h)^\alpha].$$

For further reference, we gather in a lemma some properties of the products in the LMSV case, some of which are mentioned in [11] in the case  $\sigma(x) = \exp(x)$ .

**Lemma 2.1.** *Let Assumption 2.1 hold, and let the sequences  $\{\eta_i\}$  and  $\{Z_i\}$  be mutually independent. Assume that (2.3) holds with  $q > 2\alpha$ . Then  $Y_0Y_1$  is tail equivalent to  $Z_0Z_1$  and has regularly varying and balanced tails with index  $\alpha$ . Moreover, for all  $h \geq 1$ , there exist real numbers  $d_+(h)$  and  $d_-(h)$  such that*

$$\lim_{x \rightarrow \infty} \frac{P(Y_0Y_h > x)}{P(|Y_0Y_1| > x)} = d_+(h), \quad \lim_{x \rightarrow \infty} \frac{P(Y_0Y_h < -x)}{P(|Y_0Y_1| > x)} = d_-(h). \tag{2.6}$$

Let  $b_n$  be defined by

$$b_n = \inf \left\{ x : P(|Y_0Y_1| > x) \leq \frac{1}{n} \right\}. \tag{2.7}$$

The sequence  $\{b_n\}$  is regularly varying with index  $1/\alpha$  and

$$a_n = o(b_n). \tag{2.8}$$

For all  $i \neq j > 0$ , it holds that

$$\lim_{n \rightarrow \infty} n P(|Y_0| > a_n x, |Y_0Y_j| > b_n x) = 0, \tag{2.9}$$

$$\lim_{n \rightarrow \infty} n P(|Y_0Y_i| > b_n x, |Y_0Y_j| > b_n x) = 0. \tag{2.10}$$

The quantities  $d_+(h)$  and  $d_-(h)$  can be easily computed in the LMSV case:

$$d_+(h) = \{\beta^2 + (1 - \beta)^2\} \frac{E[\sigma^\alpha(X_0)\sigma^\alpha(X_h)]}{E[\sigma^\alpha(X_0)\sigma^\alpha(X_1)]}, \quad d_-(h) = 2\beta(1 - \beta) \frac{E[\sigma^\alpha(X_0)\sigma^\alpha(X_h)]}{E[\sigma^\alpha(X_0)\sigma^\alpha(X_1)]}.$$

When leverage is present, many different situations can occur, obviously depending on the type of dependence between  $Z_0$  and  $\eta_0$ , and also on the function  $\sigma$ . We consider the exponential function  $\sigma(x) = \exp(x)$ , and a class of subadditive functions. In each case we give an assumption on the type of dependence between  $Z_0$  and  $\eta_0$  that will allow us to prove our results. Examples are given after the lemmas.

**Lemma 2.2.** *Assume that  $\sigma(x) = \exp(x)$  and  $\exp(k\eta_0)Z_0$  is tail equivalent to  $Z_0$  for all  $k \in \mathbb{R}$ . Then all the conclusions of Lemma 2.1 hold.*

**Lemma 2.3.** *Assume that the function  $\sigma$  is subadditive, i.e. there exists a constant  $C > 0$  such that, for all  $x, y \in \mathbb{R}$ ,  $\sigma(x + y) \leq C\{\sigma(x) + \sigma(y)\}$ . Assume that, for any  $a, b > 0$ ,  $\sigma(a\xi + b\eta_0)Z_0$  is tail equivalent to  $Z_0$ , where  $\xi$  is a standard Gaussian random variable independent of  $\eta_0$ , and  $\sigma(b\eta_0)Z_0$  is either tail equivalent to  $Z_0$  or  $E\{\sigma(b\eta_0)|Z_0\}^q < \infty$  for some  $q > \alpha$ . Then all the conclusions of Lemma 2.1 hold.*

**Example 2.1.** Assume that  $Z_0 = |\eta_0|^{-1/\alpha}U_0$  with  $\alpha > 0$ , where  $U_0$  is independent of  $\eta_0$  and  $E[|U_0|^q] < \infty$  for some  $q > \alpha$ . Then  $Z_0$  is regularly varying with index  $-\alpha$ .

Case 1:  $\sigma(x) = \exp(x)$ . For each  $c > 0$ ,  $Z_0 \exp(c\eta_0)$  is tail equivalent to  $Z_0$ . See Lemma 6.1 below for a proof of this fact.

Case 2:  $\sigma(x) = x^2$ . Let  $q' \in (\alpha, q \wedge \{\alpha/(1 - 2\alpha)_+\})$ . Then

$$E[\sigma^{q'}(b\eta_0)|Z_0|^{q'}] = b^{2q'} E[|\eta_0|^{q'(2-1/\alpha)}|U_0|^{q'}] < \infty.$$

Furthermore, let  $\xi$  be a standard Gaussian random variable independent of  $\eta_0$  and  $Z_0$ . Then

$$\sigma(a\xi + b\eta_0)Z_0 = a^2\xi^2Z_0 + 2ab\xi \operatorname{sgn}(\eta_0)|\eta_0|^{1-1/\alpha}U_0 + b^2|\eta_0|^{2-1/\alpha}U_0.$$

Since  $\xi$  is independent of  $Z_0$  and Gaussian, by Breiman’s lemma, the first term on the right-hand side of the above equation is tail equivalent to  $Z_0$ . The last two terms have finite moments of order  $q'$  for some  $q' > \alpha$  and do not contribute to the tail. Thus, the assumptions of Lemma 2.3 are satisfied.

**Example 2.2.** Let  $Z'_0$  have regularly varying balanced tails with index  $-\alpha$ , independent of  $\eta_0$ . Let  $\Psi_1(\cdot)$  and  $\Psi_2(\cdot)$  be polynomials, and define  $Z_0 = Z'_0\Psi_1(\eta_0) + \Psi_2(\eta_0)$ . Then, by Breiman’s lemma,  $Z_0$  is tail equivalent to  $Z'_0$ , and it is easily checked that the assumptions of Lemma 2.2 are satisfied and that the assumptions of Lemma 2.3 are satisfied with  $\sigma$  being any symmetric polynomial with positive coefficients. We omit the details.

### 3. Point process convergence

For  $s = 0, \dots, h$ , define a Radon measure  $\lambda_s$  on  $[-\infty, \infty] \setminus \{0\}$  by

$$\begin{aligned} \lambda_0(dx) &= \alpha\{\beta x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x) + (1-\beta)(-x)^{-\alpha-1}\mathbf{1}_{(-\infty,0)}(x)\} dx, \\ \lambda_s(dx) &= \alpha\{d_+(s)x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x) + d_-(s)(-x)^{-\alpha-1}\mathbf{1}_{(-\infty,0)}(x)\} dx, \end{aligned}$$

where  $d_{\pm}(s)$  are defined in (2.6). For  $s = 0, \dots, h$ , define the Radon measure  $\nu_s$  on  $[0, 1] \times [-\infty, \infty] \setminus \{0\}$  by

$$\nu_s(dt, dx) = dt\lambda_s(dx).$$

Set  $\mathbf{Y}_{n,i} = (a_n^{-1}Y_i, b_n^{-1}Y_iY_{i+1}, \dots, b_n^{-1}Y_iY_{i+h})$ , where  $a_n$  and  $b_n$  are defined in (2.4) and (2.7), respectively, and let  $N_n$  be the point process defined on  $[0, 1] \times ([-\infty, \infty]^{h+1} \setminus \{\mathbf{0}\})$  by

$$N_n = \sum_{i=1}^n \delta_{(i/n, \mathbf{Y}_{n,i})},$$

where  $\delta_x$  denotes the Dirac measure at  $x$ .

Our first result is that, for the usual univariate point process of exceedances, there is no effect of leverage. This is a consequence of the asymptotic independence (2.5).

**Proposition 3.1.** *Let Assumption 2.1 hold, and assume that  $\sigma$  is a continuous function such that (2.3) holds with  $q > \alpha$ . Then  $\sum_{i=1}^n \delta_{(i/n, Y_i/a_n)}$  converges weakly to a Poisson point process with mean measure  $\nu_0$ .*

For the multivariate point process  $N_n$ , we consider first LMSV models and then models with leverage.

#### 3.1. Point process convergence: LMSV case

**Proposition 3.2.** *Let Assumption 2.1 hold, and assume that the sequences  $\{\eta_i\}$  and  $\{Z_i\}$  are independent. Assume that the continuous volatility function  $\sigma$  satisfies (2.3) for some  $q > 2\alpha$ . Then*

$$N_n \Rightarrow \sum_{i=0}^h \sum_{k=1}^{\infty} \delta_{(t_k, j_{k,i} \mathbf{e}_i)}, \tag{3.1}$$

where  $\sum_{k=1}^{\infty} \delta_{(t_k, j_{k,0})}, \dots, \sum_{k=1}^{\infty} \delta_{(t_k, j_{k,h})}$  are independent Poisson processes with mean measures  $\nu_0, \dots, \nu_h$ , and  $\mathbf{e}_i \in \mathbb{R}^{h+1}$  is the  $i$ th basis component. Here ‘ $\Rightarrow$ ’ denotes convergence in distribution in the space of Radon point measures on  $(0, 1] \times [-\infty, \infty]^{h+1} \setminus \{\mathbf{0}\}$  equipped with the vague topology.



**3.2. Point process convergence: case of leverage**

**Proposition 3.3.** *Let Assumption 2.1 hold. Assume that  $\sigma(x) = \exp(x)$  and  $Z_0 \exp(c\eta_0)$  is tail equivalent to  $Z_0$  for all  $c$ . Then convergence (3.1) holds.*

**Proposition 3.4.** *Let Assumption 2.1 hold. Assume that the distribution of  $(Z_0, \eta_0)$  and the function  $\sigma$  satisfy the assumptions of Lemma 2.3, and, moreover, that*

$$|\sigma(x + y) - \sigma(x + z)| \leq C(\sigma(x) \vee 1)\{(\sigma(y) \vee 1) + (\sigma(z) \vee 1)\}|y - z|. \tag{3.2}$$

*Assume that condition (2.3) holds for some  $q > 2\alpha$ . Then convergence (3.1) holds.*

Condition (3.2) is an ad-hoc condition which is needed for a truncation argument used in the proof. It is satisfied by all symmetric polynomials with positive coefficients. (The proof would not be simplified by considering polynomials rather than functions satisfying this assumption.)

**4. Partial sums**

Define

$$S_n(t) = \sum_{i=1}^{[nt]} Y_i, \quad S_{p,n}(t) = \sum_{i=1}^{[nt]} |Y_i|^p.$$

For any function  $g$  such that  $E[g^2(\eta_0)] < \infty$  and any integer  $q \geq 1$ , define

$$J_q(g) = E[H_q(\eta_0)g(\eta_0)],$$

where  $H_q$  is the  $q$ th Hermite polynomial. The Hermite rank  $\tau(g)$  of the function  $g$  is the smallest positive integer  $\tau$  such that  $J_\tau(g) \neq 0$ . Let  $R_{\tau,H}$  be the so-called Hermite process of order  $\tau$  with self-similarity index  $1 - \tau(1 - H)$ . See [2] or Appendix A for more details. Let ‘ $\xrightarrow{\mathcal{D}}$ ’ denote convergence in the Skorokhod space  $\mathcal{D}([0, 1], \mathbb{R})$  of real-valued, right-continuous functions with left limits, endowed with the  $J_1$  topology; cf. [33].

**Theorem 4.1.** *Let Assumption 2.1 hold, and assume that the function  $\sigma$  is continuous and that (2.3) holds for some  $q > 2\alpha$ .*

- (i) *If  $1 < \alpha < 2$  and  $E[Z_0] = 0$ , then  $a_n^{-1}S_n$  converges weakly in the space  $\mathcal{D}([0, 1], \mathbb{R})$  endowed with Skorokhod’s  $J_1$  topology to an  $\alpha$ -stable Lévy process with skewness  $2\beta - 1$ .*

Let  $\tau_p = \tau(\sigma^p)$  be the Hermite rank of the function  $\sigma^p$ .

- (ii) *If  $p < \alpha < 2p$  and  $1 - \tau_p(1 - H) < p/\alpha$ , then*

$$a_n^{-p}(S_{p,n} - n E[|Y_0|^p]) \xrightarrow{\mathcal{D}} L_{\alpha/p},$$

where  $L_{\alpha/p}$  is a totally skewed to the right  $\alpha/p$ -stable Lévy process.

- (iii) *If  $p < \alpha < 2p$  and  $1 - \tau_p(1 - H) > p/\alpha$ , then*

$$n^{-1} \rho_n^{-\tau_p/2} (S_{p,n} - n E[|Y_0|^p]) \xrightarrow{\mathcal{D}} \frac{J_{\tau_p}(\sigma^p) E[|Z_1|^p]}{\tau_p!} R_{\tau_p,H}.$$

- (iv) *If  $p > \alpha$  then  $a_n^{-p}S_{p,n} \xrightarrow{\mathcal{D}} L_{\alpha/p}$ , where  $L_{\alpha/p}$  is a positive  $\alpha/p$ -stable Lévy process.*

Note that there is no effect of leverage. The situation will be different for the sample covariances. The fact that when the marginal distribution has infinite mean, long memory does not play any role and only a stable limit can arise was observed in a different context in [9].

### 5. Sample covariances

In order to explain more clearly the nature of the results and the problems that arise, we start by considering the sample covariances of the sequence  $\{Y_i\}$ , without assuming that  $E[Z_0] = 0$ . For notational simplicity, assume that we observe a sample of length  $n + h$ . Assume that  $\alpha > 1$ . Let  $\bar{Y}_n = n^{-1} \sum_{j=1}^n Y_j$  denote the sample mean,  $m = E[Z_0]$ , and  $\mu_Y = E[Y_0] = m E[\sigma_0]$ , and define the sample covariances by

$$\hat{\gamma}_n(s) = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)(Y_{i+s} - \bar{Y}_n), \quad 0 \leq s \leq h.$$

For simplicity, we have defined all the sample covariances as sums with the same range of indices  $1, \dots, n$ . This obviously does not affect the asymptotic theory. For  $s = 0, \dots, h$ , we furthermore define

$$C_n(s) = \frac{1}{n} \sum_{i=1}^n Y_i Y_{i+s}.$$

Then, defining  $\gamma(s) = \text{cov}(Y_0, Y_s)$ , we have, for  $s = 0, \dots, h$ ,

$$\hat{\gamma}_n(s) - \gamma(s) = C_n(s) - E[Y_0 Y_s] + \mu_Y^2 - \bar{Y}_n^2 + O_P\left(\frac{1}{n}\right).$$

Under the assumptions of Theorem 4.1,  $\bar{Y}_n^2 - \mu_Y^2 = O_P(a_n)$ . This term never contributes to the limit. Consider now  $C_n(s)$ . Recall that  $\mathcal{F}_i$  is the sigma-field generated by  $(\eta_j, Z_j)$ ,  $j \leq i$ , and define

$$\hat{X}_{i,s} = \frac{E[X_{i+s} \mid \mathcal{F}_{i-1}]}{\text{var}(E[X_{i+s} \mid \mathcal{F}_{i-1}])} = \varsigma_s^{-1} \sum_{j=s+1}^{\infty} c_j \eta_{i+s-j},$$

with  $\varsigma_s^2 = \sum_{j=s+1}^{\infty} c_j^2$ . Let  $K$  be the function defined on  $\mathbb{R}^2$  by

$$K(x, \hat{x}) = E[Z_s] E\left[ Z_0 \sigma(x) \sigma\left(\sum_{j=1}^s c_j \eta_{s-j} + \varsigma_s \hat{x}\right) \right] - E[Y_0 Y_s].$$

Then, for each  $i \geq 0$ , it holds that

$$E[Y_i Y_{i+s} \mid \mathcal{F}_{i-1}] - E[Y_0 Y_s] = K(X_i, \hat{X}_{i,s}).$$

We see that if  $m = E[Z_s] = 0$  then the function  $K$  is identically vanishing. We next write

$$\begin{aligned} C_n(s) - E[Y_0 Y_s] &= \frac{1}{n} \sum_{i=1}^n \{Y_i Y_{i+s} - E[Y_i Y_{i+s} \mid \mathcal{F}_{i-1}]\} + \frac{1}{n} \sum_{i=1}^n K(X_i, \hat{X}_{i,s}) \\ &= \frac{1}{n} M_{n,s} + \frac{1}{n} T_{n,s}. \end{aligned}$$

The point process convergence results of the previous section will allow us to prove that  $b_n^{-1} M_{n,s}$  has a stable limit. If  $m = E[Z] = 0$  then this will be the limit of  $b_n^{-1}(C_n(s) - E[Y_0 Y_s])$ , regardless of the presence of leverage. We can thus state a first result. Let ‘ $\xrightarrow{D}$ ’ denote weak convergence of sequences of finite-dimensional random vectors.

**Theorem 5.1.** Assume that  $\alpha \in (1, 2)$  and  $E[Z_0] = 0$ . Under the assumptions of Proposition 3.2, 3.3, or 3.4,

$$nb_n^{-1}(\hat{\gamma}_n(1) - \gamma(1), \dots, \hat{\gamma}_n(h) - \gamma(h)) \xrightarrow{D} (\mathcal{L}_1, \dots, \mathcal{L}_h),$$

where  $\mathcal{L}_1, \dots, \mathcal{L}_h$  are independent  $\alpha$ -stable random variables.

This result was obtained in [11] in the (LM)SV case for the function  $\sigma(x) = \exp(x)$  and under implicit conditions that rule out long memory.

We continue the discussion under the assumption that  $m \neq 0$ . Then the term  $T_{n,s}$  is the partial sum of a sequence which is a function of a bivariate Gaussian sequence. It can be treated by applying the results in [2]. Its rate of convergence and limiting distribution will depend on the Hermite rank of the function  $K$  with respect to the bivariate Gaussian vector  $(X_0, \hat{X}_{0,s})$ , which is fully characterized by the covariance between  $X_0$  and  $\hat{X}_{0,s}$ :

$$\text{cov}(X_0, \hat{X}_{0,s}) = \zeta_s^{-1} \sum_{j=1}^{\infty} c_j c_{j+s} = \zeta_s^{-1} \rho_s.$$

*LMSV case.* Since in this context the noise sequence  $\{Z_i\}$  and the volatility sequence  $\{\sigma_i\}$  are independent, we easily compute that

$$K(x, y) = m^2 \sigma(x) E[\sigma(\kappa_s \zeta + c_s \eta_0 + \zeta_s y)] - m^2 E[\sigma(X_0)\sigma(X_s)],$$

where  $\kappa_s^2 = \sum_{j=1}^{s-1} c_j^2$  and  $\zeta$  is a standard Gaussian random variable, independent of  $\eta_0$ . Thus, the Hermite rank of the function  $K$  depends only on the function  $\sigma$  (but is not necessarily equal to the Hermite rank of  $\sigma$ ).

*Case of leverage.* In this case, the dependence between  $\eta_0$  and  $Z_0$  comes into play. We now have

$$K(x, y) = m \sigma(x) E[\sigma(\kappa_s \zeta + c_s \eta_0 + \zeta_s y) Z_0] - m E[\sigma(X_0)\sigma(X_s) Z_0],$$

and now the Hermite rank of  $K$  also depends on  $Z_0$ . Different situations can occur. We give two examples.

**Example 5.1.** Consider the case  $\sigma(x) = \exp(x)$ . Then

$$\begin{aligned} E[Y_0 Y_s \mid \mathcal{F}_{-1}] &= E[Z_0 Z_s \exp(X_0) \exp(X_s) \mid \mathcal{F}_{-1}] \\ &= m E[Z_0 \exp(c_s \eta_0)] E\left[\exp\left(\sum_{j=1}^{s-1} c_j \eta_{s-j}\right)\right] \exp(X_0 + \zeta_s \hat{X}_{0,s}). \end{aligned}$$

Define  $\tilde{m} = E[Z_0 \exp(c_s \eta_0)]$ , and note that  $E[\exp(\sum_{j=1}^{s-1} c_j \eta_{s-j})] = \exp(\kappa_s^2/2)$ . Thus,

$$K(x, y) = m \tilde{m} \exp\left(\frac{\kappa_s^2}{2}\right) \{\exp(x + \zeta_s y) - E[\exp(X_0 + \zeta_s \hat{X}_{0,s})]\}.$$

If  $E[Z_0] = 0$  or  $E[Z_0 \exp(c_s \eta_0)] = 0$ , then the function  $K$  is identically vanishing and  $T_{n,s} = 0$ . Otherwise, the Hermite rank of  $K$  with respect to  $(X_0, \hat{X}_{0,s})$  is 1. Thus, applying [2, Theorem 6] (in the one-dimensional case) reveals that  $n^{-1} \rho_n^{-1/2} T_{n,s}$  converges weakly to a zero-mean Gaussian distribution. The rate of convergence is the same as in the LMSV case, but the asymptotic variance is different unless  $E[Z_0 \exp(c_s \eta_0)] = E[Z_0] E[\exp(c_s \eta_0)]$ .

**Example 5.2.** Consider  $\sigma(x) = x^2$ . Define

$$\check{X}_{i,s} = \kappa_s^{-1} \sum_{j=1}^{s-1} c_j \eta_{i+s-j}.$$

Then

$$\begin{aligned} E[Y_0 Y_s \mid \mathcal{F}_{-1}] &= E[Z_0 Z_s X_0^2 (\kappa_s \check{X}_{0,s} + \zeta_s \hat{X}_{0,s} + c_s \eta_0)^2 \mid \mathcal{F}_{-1}] \\ &= m X_0^2 \{ \kappa_s^2 m + c_s E[Z_0 \eta_0^2] + \zeta_s m (\hat{X}_{0,s})^2 + 2 \zeta_s c_s E[Z_0 \eta_0] \hat{X}_{0,s} \}. \end{aligned}$$

Thus,

$$\begin{aligned} K(x, y) &= \zeta_s m^2 (x^2 y^2 - E[X_0^2 (\hat{X}_{0,s})^2]) + 2 \zeta_s c_s m E[Z_0 \eta_0] \{ x^2 y - E[X_0^2 \hat{X}_{0,s}] \} \\ &\quad + (\kappa_s^2 m^2 + c_s m E[Z_0 \eta_0^2]) (x^2 - 1), \end{aligned}$$

and it can be verified that the Hermite rank of  $K$  with respect to  $(X_0, \hat{X}_0^{(s)})$  is 1, except if  $E[Z_0 \eta_0] = 0$ , which holds in the LMSV case. Thus, we see that the rate of convergence of  $T_{n,s}$  depends on the presence or absence of leverage. See Example 5.4 below for details.

Let us now introduce the notation that will be used to deal with sample covariances of powers. For  $p > 0$ , define  $m_p = E[|Z_0|^p]$ . If  $p \in (\alpha, 2\alpha)$  and assumption (2.1) holds,  $m_p$  is finite and  $E[|Z_0|^{2p}] = \infty$ . Moreover, under the assumptions of Lemma 2.1 or 2.2, for  $s > 0$ ,  $E[|Y_0 Y_s|^p] < \infty$  and  $E[|Y_0 Y_s|^{2p}] = \infty$  for  $p \in (\alpha/2, \alpha)$ . Thus, the autocovariance  $\gamma_p(s) = \text{cov}(|Y_0|^p, |Y_s|^p)$  is well defined. Furthermore, define  $\bar{Y}_{p,n} = n^{-1} \sum_{i=1}^n |Y_i|^p$  and

$$\hat{\gamma}_{p,n}(s) = \frac{1}{n} \sum_{i=1}^n (|Y_i|^p - \bar{Y}_{p,n})(|Y_{i+s}|^p - \bar{Y}_{p,n}).$$

Define the functions  $K_{p,s}^*$  (LMSV case) and  $K_{p,s}^\dagger$  (case with leverage) by

$$K_{p,s}^*(x, y) = m_p^2 \sigma^p(x) E[\sigma^p(\kappa_s \zeta + c_s \eta_0 + \zeta_s y)] - m_p^2 E[\sigma^p(X_0) \sigma^p(X_s)], \tag{5.1}$$

$$K_{p,s}^\dagger(x, y) = m_p \sigma^p(x) E[\sigma^p(\kappa_s \zeta + c_s \eta_0 + \zeta_s y) |Z_0|^p] - m_p E[\sigma^p(X_0) \sigma^p(X_s) |Z_0|^p]. \tag{5.2}$$

**5.1. Convergence of the sample covariance of powers: LMSV case**

**Theorem 5.2.** *Let Assumption 2.1 hold, and assume that the sequences  $\{\eta_i\}$  and  $\{Z_i\}$  are independent. Let the function  $\sigma$  be continuous and satisfy (2.3) with  $q > 4\alpha$ . For a fixed integer  $s \geq 1$ , let  $\tau_p^*(s)$  be the Hermite rank of the bivariate function  $K_{p,s}^*$  defined by (5.1), with respect to a bivariate Gaussian vector with standard marginal distributions and correlation  $\zeta_s^{-1} \gamma_s$ .*

- If  $p < \alpha < 2p$  and  $1 - \tau_p^*(s)(1 - H) < p/\alpha$ , then

$$n b_n^{-p} (\hat{\gamma}_{p,n}(s) - \gamma_p(s)) \xrightarrow{D} \mathcal{L}_s,$$

where  $\mathcal{L}_s$  is an  $\alpha/p$ -stable random variable.

- If  $p < \alpha < 2p$  and  $1 - \tau_p^*(s)(1 - H) > p/\alpha$ , then

$$\rho_n^{-\tau_p^*(s)/2} (\hat{\gamma}_{p,n}(s) - \gamma_p(s)) \xrightarrow{D} G_s^*,$$

where the random variable  $G_s^*$  is Gaussian if  $\tau_p^*(s) = 1$ .

For different values  $s = 1, \dots, h$ , the Hermite ranks  $\tau_p^*(s)$  of the functions  $K_{p,s}^*$  may be different. Therefore, in order to consider the joint autocovariances at lags  $s = 1, \dots, h$ , we define

$$\tau_p^* = \min\{\tau_p^*(1), \dots, \tau_p^*(h)\}.$$

**Corollary 5.1.** *Under the assumptions of Theorem 5.2,*

- if  $1 - \tau_p^*(1 - H) < p/\alpha$  then

$$nb_n^{-p}(\hat{\gamma}_{p,n}(1) - \gamma_p(1), \dots, \hat{\gamma}_{p,n}(h) - \gamma_p(h)) \xrightarrow{D} (\mathcal{L}_1, \dots, \mathcal{L}_h),$$

where  $\mathcal{L}_1, \dots, \mathcal{L}_p$  are independent  $\alpha/p$ -stable random variables;

- if  $1 - \tau_p^*(1 - H) > p/\alpha$  then

$$\rho_n^{-\tau_p^*/2}(\hat{\gamma}_{p,n}(1) - \gamma_p(1), \dots, \hat{\gamma}_{p,n}(h) - \gamma_p(h)) \xrightarrow{D} (\tilde{G}_1^*, \dots, \tilde{G}_h^*),$$

where  $\tilde{G}_s^* = G_s^*$  if  $\tau_p^*(s) = \tau_p^*$  and  $\tilde{G}_s^* = 0$  otherwise.

We see that the joint limiting vector  $(\tilde{G}_1^*, \dots, \tilde{G}_h^*)$  may have certain zero components if there exist indices  $s$  such that  $\tau_p^*(s) > \tau_p^*$ . However, for standard choices of the function  $\sigma$ , the Hermite rank  $\tau_p^*(s)$  does not depend on  $s$ . For instance, for  $\sigma(x) = \exp(x)$ ,  $\tau_p^*(s) = 1$  for all  $s$  and, for  $\sigma(x) = x^2$ ,  $\tau_p^*(s) = 2$  for all  $s$ .

**5.2. Convergence of the sample covariance of powers: case of leverage**

**Theorem 5.3.** *Let the assumptions of Proposition 3.3 or 3.4 hold, and assume that (2.3) holds for some  $q > 4\alpha$ . Let  $\tau_p^\dagger(s)$  be the Hermite rank of the bivariate function  $K_{p,s}^\dagger$  defined by (5.2), with respect to a bivariate Gaussian vector with standard marginal distributions and correlation  $\zeta_s^{-1}\gamma_s$ .*

- If  $p < \alpha < 2p$  and  $1 - \tau_p^\dagger(s)(1 - H) < p/\alpha$ , then

$$nb_n^{-p}(\hat{\gamma}_{p,n}(s) - \gamma_p(s)) \xrightarrow{D} \mathcal{L}_s,$$

where  $\mathcal{L}_s$  is an  $\alpha/p$ -stable random variable.

- If  $p < \alpha < 2p$  and  $1 - \tau_p^\dagger(s)(1 - H) > p/\alpha$ , then

$$\rho_n^{-\tau_p^\dagger(s)/2}(\hat{\gamma}_{p,n}(s) - \gamma_p(s)) \xrightarrow{D} G_s^\dagger,$$

where the random vector  $G_s^\dagger$  is Gaussian if  $\tau_p^\dagger(s) = 1$ .

Again, as in the previous case, in order to formulate the multivariate result, we further define

$$\tau_p^\dagger = \min\{\tau_p^\dagger(1), \dots, \tau_p^\dagger(h)\}.$$

**Corollary 5.2.** *Under the assumptions of Theorem 5.3,*

- if  $1 - \tau_p^\dagger(1 - H) < p/\alpha$  then

$$nb_n^{-p}(\hat{\gamma}_{p,n}(1) - \gamma_p(1), \dots, \hat{\gamma}_{p,n}(h) - \gamma_p(h)) \xrightarrow{D} (\mathcal{L}_1, \dots, \mathcal{L}_h),$$

where  $\mathcal{L}_1, \dots, \mathcal{L}_p$  are independent  $\alpha/p$ -stable random variables;

- if  $1 - \tau_p^\dagger(1 - H) > p/\alpha$  then

$$\rho_n^{-\tau_p^\dagger/2}(\hat{\gamma}_{p,n}(1) - \gamma_p(1), \dots, \hat{\gamma}_{p,n}(h) - \gamma_p(h)) \xrightarrow{D} (\tilde{G}_1^\dagger, \dots, \tilde{G}_h^\dagger),$$

where  $\tilde{G}_s^\dagger = G_s^\dagger$  if  $\tau_p^\dagger(s) = \tau_p^\dagger$  and  $\tilde{G}_s^\dagger = 0$  otherwise.

The main difference between Theorems 5.2 and 5.3 (or Corollaries 5.1 and 5.2) is the Hermite rank considered. Under the conditions that ensure convergence to a stable limit, the rates of convergence and the limits are the same in both theorems. Otherwise, the rates and the limits may be different.

**Example 5.3.** Consider the case  $\sigma(x) = \exp(x)$ . For all  $s \geq 1$ , we have  $\tau_p^\dagger = \tau_p^\dagger(s) = 1$ . Thus, under the assumptions of Theorem 5.3,

- if  $H < p/\alpha$  then  $nb_n^{-1}\{\hat{\gamma}_{p,n}(s) - \gamma_p(s)\}$  converges weakly to a stable law;
- if  $H > p/\alpha$  then  $\rho_n^{-1/2}\{\hat{\gamma}_{p,n}(s) - \gamma_p(s)\}$  converges weakly to a zero-mean Gaussian distribution.

The dichotomy is the same as in the LMSV case, but the variance of the limiting distribution in the case  $H > p/\alpha$  is different except if  $E[Z_0 \exp(c_s \eta_0)] = E[Z_0] E[\exp(c_s \eta_0)]$ .

**Example 5.4.** Consider the case  $\sigma(x) = x^2$  and  $p = 1$ . Assume that  $E[\eta_1|Z_1] \neq 0$ . Then, for each  $s \geq 1$ ,  $\tau_1^\dagger = \tau_1^\dagger(s) = 1$ , whereas  $\tau_p^* = \tau_p^*(s) = 2$ ; thus, the dichotomy is not the same as in the LMSV case and the rate of convergence differs in the case  $H > 1/\alpha$ .

- If  $H < 1/\alpha$  then  $nb_n^{-1}\{\hat{\gamma}_{n,1}(s) - \gamma_1(s)\}$  converges weakly to a stable law.
- If  $H > 1/\alpha$  then  $\rho_n^{-1/2}\{\hat{\gamma}_{n,1}(s) - \gamma_1(s)\}$  converges weakly to a zero-mean Gaussian distribution.

If we assume now that  $E[\eta_1|Z_1] = 0$  then  $\tau_1^\dagger = \tau_p^* = 2$ . Thus, the dichotomy is the same as in the LMSV case, but the limiting distribution in the nonstable case can be different from the one in the LMSV case.

- If  $2H - 1 < 1/\alpha$  then  $nb_n^{-1}\{\hat{\gamma}_{1,n}(s) - \gamma_1(s)\}$  converges weakly to a stable law.
- If  $2H - 1 > 1/\alpha$  then  $\rho_n^{-1}\{\hat{\gamma}_{1,n}(s) - \gamma_1(s)\}$  converges weakly to a zero-mean non-Gaussian distribution.

If, moreover,  $E[H_2(\eta_1)|Z_1] = 0$  then, for each  $s$ , the functions  $K_{p,s}^*$  and  $K_{p,s}^\dagger$  are equal, and, thus, the limiting distribution is the same as in the LMSV case.

### 6. Proofs

**Lemma 6.1.** Let  $Z$  be a nonnegative random variable with a regularly varying right tail with index  $-\alpha$ ,  $\alpha > 0$ . Let  $g$  be a bounded function on  $[0, \infty)$  such that  $\lim_{x \rightarrow +\infty} g(x) = c_g \in (0, \infty)$ . Then  $Zg(Z)$  is tail equivalent to  $Z$ :

$$\lim_{x \rightarrow +\infty} \frac{P(Zg(Z) > x)}{P(Z > x)} = c_g^\alpha.$$

*Proof.* Fix some  $\varepsilon \in (0, 1)$ , and let  $x_0$  be large enough so that  $|g(x) - c_g|/c_g < \varepsilon$  for all  $x > x_0$ . The function  $g$  is bounded; thus,  $zg(z) > x$  implies that  $z > x/\|g\|_\infty$  and if  $x > x_0\|g\|_\infty$ , we have

$$\begin{aligned} P(Zg(Z) > x) &= P\left(Zg(Z) > x, Z > \frac{x}{\|g\|_\infty}\right) \\ &\leq P\left(Zc_g(1 + \varepsilon) > x, Z > \frac{x}{\|g\|_\infty}\right) \\ &\leq P(Zc_g(1 + \varepsilon) > x). \end{aligned}$$

This yields the upper bound

$$\limsup_{x \rightarrow +\infty} \frac{P(Zg(Z) > x)}{P(Z > x)} \leq \limsup_{x \rightarrow +\infty} \frac{P(Zc_g(1 + \varepsilon) > x)}{P(Z > x)} = c_g^\alpha(1 + \varepsilon)^\alpha.$$

Conversely, we have

$$\begin{aligned} P(Zg(Z) > x) &= P\left(Zg(Z) > x, Z > \frac{x}{\|g\|_\infty}\right) \\ &\geq P\left(Zc_g(1 - \varepsilon) > x, Z > \frac{x}{\|g\|_\infty}\right) \\ &= P\left(Z > x \max\left\{\frac{1}{c_g(1 - \varepsilon)}, \frac{1}{\|g\|_\infty}\right\}\right) \\ &= P\left(Z > \frac{x}{c_g(1 - \varepsilon)}\right), \end{aligned}$$

where the last equality comes from the fact that  $(1 - \varepsilon)c_g \leq c_g = \lim_{z \rightarrow +\infty} g(z) \leq \|g\|_\infty$ . Thus,

$$\liminf_{x \rightarrow +\infty} \frac{P(Zg(Z) > x)}{P(Z > x)} \geq \liminf_{x \rightarrow +\infty} \frac{P(Zc_g(1 - \varepsilon) > x)}{P(Z > x)} = c_g^\alpha(1 - \varepsilon)^\alpha.$$

Since  $\varepsilon$  is arbitrary, we obtain the desired limit.

**Lemma 6.2.** *Let  $Z$  be a nonnegative random variable with a regularly varying right tail with index  $-\alpha$ ,  $\alpha > 0$ . For each  $\varepsilon > 0$ , there exists a constant  $C$  such that, for all  $x \geq 1$  and all  $y > 0$ ,*

$$\frac{P(yZ > x)}{P(Z > x)} \leq C(y \vee 1)^{\alpha+\varepsilon}. \tag{6.1}$$

*Proof.* If  $y \leq 1$  then  $P(yZ > x) \leq P(Z > x)$ , so the requested bound holds trivially with  $C = 1$ . Assume now that  $y \geq 1$ . Then, by Markov’s inequality,

$$P(yZ > x) \leq P(Z > x) + P\left(Z\mathbf{1}_{\{Z \leq x\}} > \frac{x}{y}\right) \leq P(Z > x) + x^{-\alpha-\varepsilon}y^{\alpha+\varepsilon} E[Z^{\alpha+\varepsilon}\mathbf{1}_{\{Z \leq x\}}]. \tag{6.2}$$

Next, by [5, Theorem 8.1.2] or [17, Theorem VIII.9.2],

$$\lim_{x \rightarrow +\infty} \frac{E[Z^{\alpha+\varepsilon}\mathbf{1}_{\{Z \leq x\}}]}{x^{\alpha+\varepsilon} P(Z > x)} = \frac{\alpha}{\varepsilon}.$$

Moreover, the function  $x \rightarrow P(Z > x)$  is decreasing on  $[0, \infty)$ , and, hence, bounded away from 0 on compact sets of  $[0, \infty)$ . Thus, there exists a constant  $C$  (that does not depend on  $x$  and  $y$ ) such that, for all  $x \geq 1$ ,

$$\frac{E[Z^{\alpha+\varepsilon} \mathbf{1}_{\{Z \leq x\}}]}{P(Z > x)} \leq Cx^{\alpha+\varepsilon}. \tag{6.3}$$

Substituting (6.3) into (6.2) yields, for all  $x, y \geq 1$ ,

$$\frac{P(yZ > x)}{P(Z > x)} = 1 + Cy^{\alpha+\varepsilon}.$$

This concludes the proof of (6.1).

*Proof of Lemma 2.1.* Under the assumption of independence between the sequences  $\{Z_i\}$  and  $\{\eta_i\}$ , as already mentioned,  $Y_0$  is tail equivalent to  $Z_0$  and  $Y_0Y_h$  is tail equivalent to  $Z_0Z_1$  for all  $h$ . Properties (2.6), (2.7), and (2.8) are straightforward. We need to prove (2.9) and (2.10). Since  $Z_0$  is independent of  $\sigma_j$  and  $Z_j$ , by conditioning we have

$$n P(|Y_0| > a_n x, |Y_0 Y_j| > b_n x) = E \left[ n \bar{F}_{|Z|} \left( \frac{a_n x}{\sigma_0} \vee \frac{b_n x}{\sigma_0 \sigma_j |Z_j|} \right) \right]$$

with  $F_{|Z|}$  the distribution function of  $|Z_0|$ . Since  $a_n/b_n \rightarrow 0$ , for any  $y > 0$ , it holds that  $\lim_{n \rightarrow +\infty} n \bar{F}_{|Z|}(b_n y) = 0$ . Thus,

$$n \bar{F}_{|Z|} \left( \frac{a_n x}{\sigma_0} \vee \frac{b_n x}{\sigma_0 \sigma_j |Z_j|} \right) \leq n \bar{F}_{|Z|} \left( \frac{b_n x}{\sigma_0 \sigma_j |Z_j|} \right) \rightarrow 0 \text{ almost surely.}$$

Moreover, by Lemma 6.2 and the definition of  $a_n$ , for any  $\varepsilon > 0$ , there exists a constant  $C$  such that

$$n \bar{F}_{|Z|} \left( \frac{a_n x}{\sigma_0} \vee \frac{b_n x}{\sigma_0 \sigma_j |Z_j|} \right) \leq n \bar{F}_{|Z|} \left( \frac{a_n x}{\sigma_0} \right) \leq Cx^{-\alpha-\varepsilon} \sigma_0^{\alpha+\varepsilon}.$$

By assumption, (2.3) holds for some  $q > \alpha$ . Thus, choosing  $\varepsilon$  small enough allows us to apply the bounded convergence theorem and this proves (2.9). Next, to prove (2.10), note that  $|Y_i| \wedge |Y_j| \leq (\sigma_i \vee \sigma_j)(|Z_i| \wedge |Z_j|)$ . Thus, applying Lemma 6.2, we have

$$\begin{aligned} P(|Y_0 Y_i| > x, |Y_0 Y_j| > x) &= P(|Z_0| \sigma_0 (\sigma_i |Z_i| \wedge \sigma_j |Z_j|) > x) \\ &\leq C P(|Z_0| > x) E[\sigma_0^{\alpha+\varepsilon} (\sigma_i \vee \sigma_j)^{\alpha+\varepsilon}] E[(|Z_i| \wedge |Z_j|)^{\alpha+\varepsilon}]. \end{aligned}$$

The expectation  $E[\sigma_0^{\alpha+\varepsilon} (\sigma_i \vee \sigma_j)^{\alpha+\varepsilon}]$  is finite for small enough  $\varepsilon$ , since assumption (2.3) holds with  $q > 2\alpha$ . Since  $P(|Z_0| > x) = o(P(|Z_1 Z_2| > x))$ , this yields (2.10) in the LMSV case.

*Proof of Lemma 2.2.* It suffices to prove the lemma when the random variables  $Z_i$  are nonnegative. Under the assumption of the lemma,  $\exp(c_h \eta_0) Z_0$  is tail equivalent to  $Z_0$ . Thus, by the corollary in [16, p. 245],  $Z_0 \exp(c_h \eta_0) Z_h$  is regularly varying with index  $\alpha$  and tail equivalent to  $Z_0 Z_h$ . Since  $E[Z_0^\alpha] = \infty$ , it also holds that  $P(Z_0 > x) = o(P(\exp(c_h \eta_0) Z_0 Z_1 > x))$ ; cf. [14, Equation (3.5)].

Define  $\hat{X}_h = \sum_{k=1, k \neq h}^\infty c_k \eta_{h-k}$ . Then  $\hat{X}_h$  is independent of  $Z_0, \eta_0$ , and  $Z_h$ . Since  $Y_0 Y_h = \exp(X_0 + \hat{X}_h) Z_0 \exp(c_h \eta_0) Z_h$ , we can apply Breiman's lemma to show that  $Y_0 Y_h$  is tail equivalent to  $Z_0 \exp(c_h \eta_0) Z_h$ , and, hence, to  $Z_0 Z_1$ . Thus, (2.6) and (2.8) hold with

$$d_+(h) = \tilde{\beta} \frac{E[\exp(\alpha(X_0 + \hat{X}_h))]}{E[\exp(\alpha(X_0 + \hat{X}_1))]}, \quad d_-(h) = (1 - \tilde{\beta}) \frac{E[\exp(\alpha(X_0 + \hat{X}_h))]}{E[\exp(\alpha(X_0 + \hat{X}_1))]},$$

where  $\tilde{\beta}$  is the skewness parameter of  $Z_0 \exp(c_h \eta_0) Z_h$ .



We now prove (2.10). For fixed  $i$  and  $j$  such that  $0 < i < j$ , define

$$\hat{\sigma}_i = \sigma(\hat{X}_i) = \exp\left(\sum_{k=1, k \neq i}^{\infty} c_k \eta_{i-k}\right), \quad \check{\sigma}_{i,j} = \sigma(\check{X}_{i,j}) = \exp\left(\sum_{k=1, k \neq j, j-i}^{\infty} c_k \eta_{j-k}\right).$$

Define  $\tilde{Z}_0^{(k)} = Z_0 \exp(c_k \eta_0)$  and  $V_i = \exp(c_{j-i} \eta_i)$ . Then

$$\begin{aligned} P(Y_0 Y_i > x, Y_0 Y_j > x) &= P(\sigma_0 \hat{\sigma}_i \tilde{Z}_0^{(i)} Z_i > x, \sigma_0 \check{\sigma}_{i,j} \tilde{Z}_0^{(i)} \exp(c_{j-i} \eta_i) Z_j > x) \\ &\leq P(\sigma_0 (\hat{\sigma}_i \vee \check{\sigma}_{i,j}) (\tilde{Z}_0^{(i)} + \tilde{Z}_0^{(j)}) (Z_i \wedge V_i Z_j) > x). \end{aligned}$$

Now,  $(Z_i \wedge V_i Z_j)$  is independent of  $\sigma_0 (\hat{\sigma}_i \vee \check{\sigma}_{i,j}) (\tilde{Z}_0^{(i)} + \tilde{Z}_0^{(j)})$ , which is tail equivalent to  $Z_0$  by assumption and Breiman’s lemma. Thus, in order to prove (2.10), we only need to show that, for some  $\delta > \alpha$ ,  $E[(Z_i \wedge V_i Z_j)^\delta] < \infty$ . This is true. Indeed, since  $E[V_i^q] < \infty$  for all  $q > 1$ , we can apply Hölder’s inequality with  $q$  arbitrarily close to 1. This yields, for  $p^{-1} + q^{-1} = 1$ ,

$$E[(Z_i \wedge V_i Z_j)^\delta] \leq E[(1 \vee V_i)^\delta (Z_i \wedge Z_j)^\delta] \leq E^{1/p}[(1 \vee V_i)^{p\delta}] E^{1/q}[(Z_i \wedge Z_j)^{q\delta}].$$

The tail index of  $(Z_i \wedge Z_j)$  is  $2\alpha$ , and, thus,  $E^{1/q}[(Z_i \wedge Z_j)^{q\delta}] < \infty$  for any  $q$  and  $\delta$  such that  $q\delta < 2\alpha$ . Thus,  $E[(Z_i \wedge V_i Z_j)^\delta] < \infty$  for any  $\delta \in (\alpha, 2\alpha)$  and (2.10) holds. The proof of (2.9) is similar.

*Proof of Lemma 2.3.* We omit the proof of the regular variation and the tail equivalence between  $Y_0 Y_h$  and  $Z_0 Z_1$  which is a straightforward consequence of the assumption. We prove (2.10). Using the notation in the proof of Lemma 2.2, by the subadditivity property of  $\sigma$ , we have, for  $j > i > 0$  and some constant  $C$ ,

$$\begin{aligned} P(Y_0 Y_i > x, Y_0 Y_j > x) &= P(\sigma_0 \sigma(\hat{X}_i + c_i \eta_0) Z_0 Z_i > x, \sigma_0 \sigma(\check{X}_{i,j} + c_j \eta_0 + c_{j-i} \eta_i) Z_0 Z_j > x) \\ &\leq P(C\sigma_0 | Z_0 | \{\sigma(\hat{X}_i) + \sigma(c_i \eta_0)\} \{\sigma(\check{X}_{i,j}) + \sigma(c_j \eta_0) + \sigma(c_{j-i} \eta_i)\} (|Z_i| \wedge |Z_j|) > x) \\ &\leq P(C\sigma_0 | Z_0 | \sigma(\hat{X}_i) \sigma(\check{X}_{i,j}) (|Z_i| \wedge |Z_j|) > x) \\ &\quad + P(C\sigma_0 | Z_0 | \sigma(\hat{X}_i) \sigma(c_j \eta_0) (|Z_i| \wedge |Z_j|) > x) \\ &\quad + P(C\sigma_0 | Z_0 | \sigma(\hat{X}_i) \sigma(c_{j-i} \eta_i) (|Z_i| \wedge |Z_j|) > x) \\ &\quad + P(C\sigma_0 | Z_0 | \sigma(c_i \eta_0) \sigma(\check{X}_{i,j}) (|Z_i| \wedge |Z_j|) > x) \\ &\quad + P(C\sigma_0 | Z_0 | \sigma(c_i \eta_0) \sigma(c_j \eta_0) (|Z_i| \wedge |Z_j|) > x) \\ &\quad + P(C\sigma_0 | Z_0 | \sigma(c_i \eta_0) \sigma(c_{j-i} \eta_i) (|Z_i| \wedge |Z_j|) > x). \end{aligned}$$

Now, under the assumptions of the lemma, each of the last six probabilities can be expressed as  $P(\tilde{Z}U > x)$ , where  $\tilde{Z}$  is tail equivalent to  $Z_0$  and  $U$  is independent of  $\tilde{Z}$  and  $E[|U|^q] < \infty$  for some  $q > \alpha$ . Thus, by Breiman’s lemma,  $\tilde{Z}U$  is also tail equivalent to  $Z_0$ , and, thus,  $P(Y_0 Y_i > x, Y_0 Y_j > x) = O(P(|Z_0| > x)) = o(P(|Y_0 Y_1| > x))$ , which proves (2.10).

**6.1. Proof of Propositions 3.1, 3.2, 3.3, and 3.4**

We omit some details of the proof, since it is a slight modification of the proofs of Theorems 3.1 and 3.2 of [11], adapted to a general stochastic volatility with possible leverage and long memory. Note that the proof of [11, Theorem 3.2] refers to the proof of Theorem 2.4 of [12]. The latter proof uses Condition (2.6) of [12], which rules out long memory.

The proof is in two steps. In the first step we consider an  $m$ -dependent approximation  $X^{(m)}$  of the Gaussian process and prove point process convergence for the corresponding stochastic volatility process  $Y^{(m)}$  for each fixed  $m$ . The second step naturally consists in proving that the limits for the  $m$ -dependent approximations converge when  $m$  tends to  $\infty$ , and that this limit is indeed the limit of the original sequence.

*Step 1.* Let  $X_i^{(m)} = \sum_{k=1}^m c_k \eta_{i-k}$  and  $Y_i^{(m)} = \sigma(X_i^{(m)})Z_i$ , and accordingly define  $\mathbf{Y}_{n,i}^{(m)}$ . Note that the tail properties of the process  $\{Y_i^{(m)}\}$  are the same as those of the process  $\{Y_i\}$ , since the latter are proved without any particular assumptions on the coefficients  $c_j$  of expansion (2.2) apart from square summability. In order to prove the desired point process convergence, as in the proof of [11, Theorem 3.1], we must check the following two conditions (which are Equations (3.3) and (3.4) of [11]):

$$P(\mathbf{Y}_{n,1}^{(m)} \in \cdot) \xrightarrow{v} \mathbf{v}_m, \tag{6.4}$$

$$\lim_{k \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \sum_{i=2}^{[n/k]} E[g(\mathbf{Y}_{n,1}^{(m)})g(\mathbf{Y}_{n,i}^{(m)})] = 0. \tag{6.5}$$

Here  $\mathbf{v}_m$  is the mean measure of the limiting point process and (6.5) must hold for any continuous bounded function  $g$ , compactly supported on  $[0, 1] \times [-\infty, \infty]^h \setminus \{0\}$ .

Convergence (6.4) is a straightforward consequence of the joint regular variation and the asymptotic independence properties (2.9) and (2.10) of  $Y_0, Y_0Y_1, \dots, Y_0Y_h$ . Let us now prove (6.5). Note first that, because of asymptotic independence, for any fixed  $i$ ,

$$\lim_{n \rightarrow +\infty} n E[g(\mathbf{Y}_{n,1}^{(m)})g(\mathbf{Y}_{n,i}^{(m)})] = 0.$$

Next, by  $m$ -dependence, for each  $k$ , as  $n \rightarrow +\infty$ , we have

$$\begin{aligned} n \sum_{i=2+m+h}^{[n/k]} E[g(\mathbf{Y}_{n,1}^{(m)})g(\mathbf{Y}_{n,i}^{(m)})] &= n \sum_{i=2+m+h}^{[n/k]} E[g(\mathbf{Y}_{n,1}^{(m)})]E[g(\mathbf{Y}_{n,i}^{(m)})] \\ &\sim \frac{1}{k} (n E[g(\mathbf{Y}_{n,1}^{(m)})])^2 \\ &\rightarrow \frac{1}{k} \left( \int g \, d\mathbf{v}_m \right)^2. \end{aligned}$$

This yields (6.5). Thus, we obtain

$$\sum_{i=1}^n \delta_{(i/n, \mathbf{Y}_{n,i}^{(m)})} \Rightarrow \sum_{l=1}^h \sum_{k=1}^{\infty} \delta_{(t_{k,l}, j_{k,l}^{(m)}) \mathbf{e}_l},$$

where  $\sum_{k=1}^{\infty} \delta_{(t_{k,j}, j_{k,0}^{(m)})}, \dots, \sum_{k=1}^{\infty} \delta_{(t_{k,j}, j_{k,h}^{(m)})}$  are independent Poisson processes with respective mean measures

$$\lambda_{0,m}(dx) = \alpha \{ \beta_m x^{-\alpha-1} \mathbf{1}_{(0,\infty)}(x) + (1 - \beta_m) (-x)^{-\alpha-1} \mathbf{1}_{(-\infty,0)}(x) \} dx, \tag{6.6}$$

$$\lambda_{s,m}(dx) = \alpha \{ d_+^{(m)}(s) x^{-\alpha-1} \mathbf{1}_{(0,\infty)}(x) + d_-^{(m)}(s) (-x)^{-\alpha-1} \mathbf{1}_{(-\infty,0)}(x) \} dx, \tag{6.7}$$

where the values of  $d_+^{(m)}(s)$  and  $d_-^{(m)}(s)$  depend on the process considered and

$$\beta_m = \frac{\beta E[\sigma^\alpha(X^{(m)})]}{E[\sigma^\alpha(X)]}.$$

Step 2. We must now prove that

$$N_m \Rightarrow N \tag{6.8}$$

as  $m \rightarrow +\infty$  and that, for all  $\eta > 0$ ,

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} P(\varrho(N_n, N_n^{(m)}) > \eta) = 0, \tag{6.9}$$

where  $\varrho$  is the metric inducing the vague topology; cf. Equations (3.13) and (3.14) of [11]. To prove (6.8), it suffices to prove that

$$\lim_{m \rightarrow +\infty} \beta_m = \beta, \tag{6.10}$$

$$\lim_{m \rightarrow +\infty} d_+^{(m)}(s) = d_+(s), \quad \lim_{m \rightarrow +\infty} d_-^{(m)}(s) = d_-(s). \tag{6.11}$$

To prove (6.9), as in the proof of [11, Theorem 3.3], it suffices to show that, for all  $\varepsilon > 0$ ,

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} n P(a_n^{-1} |Y_0 - Y_0^{(m)}| > \varepsilon) = 0, \tag{6.12}$$

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} n P(b_n^{-1} |Y_0 Y_s - Y_0^{(m)} Y_s^{(m)}| > \varepsilon) = 0. \tag{6.13}$$

If (2.3) holds for some  $q > \alpha$  and if  $\sigma$  is continuous, then (6.10) holds by bounded convergence, in both the LMSV case and the case of leverage. We now prove (6.12). Since  $Y_0$  and  $Z_0$  are tail equivalent, by Breiman’s lemma we have

$$\limsup_{n \rightarrow +\infty} n P(a_n^{-1} |Y_0 - Y_0^{(m)}| > \varepsilon) \leq C \varepsilon^{-\alpha} E[|\sigma(X_0^{(m)}) - \sigma(X_0)|^\alpha].$$

The continuity of  $\sigma$ , assumption (2.3) with  $q > \alpha$ , and the bounded convergence theorem imply that  $\lim_{m \rightarrow +\infty} E[|\sigma(X_0^{(m)}) - \sigma(X_0)|^\alpha] = 0$ . This proves (6.12) in both the LMSV case and the case of leverage. We now split the proofs of (6.11) and (6.13) between the LMSV and leverage cases.

*LMSV case.* In this case we have

$$d_+^{(m)}(s) = d_+(s) \frac{E[\sigma^\alpha(X_0^{(m)})\sigma^\alpha(X_s^{(m)})]}{E[\sigma^\alpha(X_0)\sigma^\alpha(X_s)]}, \quad d_-^{(m)}(s) = d_-(s) \frac{E[\sigma^\alpha(X_0^{(m)})\sigma^\alpha(X_s^{(m)})]}{E[\sigma^\alpha(X_0)\sigma^\alpha(X_s)]}.$$

For  $s = 1, \dots, h$ , define

$$W_{m,s} = \sigma(X_0^{(m)})\sigma(X_s^{(m)}) - \sigma(X_1)\sigma(X_{1+s}).$$

The continuity of  $\sigma$  implies that  $W_{m,s} \xrightarrow{P} 0$  as  $m \rightarrow +\infty$ . Under the Gaussian assumption,  $X^{(m)} \stackrel{D}{=} u_m X$  for some  $u_m \in (0, 1)$ . Thus, if (2.3) holds for some  $q' > \alpha$  then it also holds that

$$\sup_{m \geq 1} E[\sigma^{q'}(X^{(m)})] < \infty;$$

hence  $W_m$  converges to 0 in  $L^q$  for any  $q < q'$ . Likewise, since assumption (2.3) holds for some  $q' > 2\alpha$ ,  $W_{m,s}$  converges to 0 in  $L^q$  for any  $q < q'$ . Since  $|W_m|$  and  $|W_{m,s}|$  converge to 0 in  $L^\alpha$ ,  $d_+^{(m)}(s)$  and  $d_-^{(m)}(s)$  converge to the required limits. We now prove (6.13). Since  $Z_0 Z_s$  is

tail equivalent to  $Y_0Y_1$ , by another application of Breiman’s lemma, we obtain, for  $s = 1, \dots, h$  and  $\varepsilon > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \mathbb{P}(b_n^{-1} |Y_0Y_s - Y_0^{(m)}Y_s^{(m)}| > \varepsilon) &\leq \limsup_{n \rightarrow +\infty} n \mathbb{P}(b_n^{-1} |Z_0Z_s| |W_{m,s}| > C\varepsilon) \\ &\leq C^{-\alpha} \varepsilon^{-\alpha} \mathbb{E}[|W_{m,s}|^\alpha], \end{aligned}$$

which converges to 0 as  $m \rightarrow +\infty$ . This concludes the proof of (6.13) in the LMSV case.

To prove (6.13) in the case of leverage, we further split the proof between the cases  $\sigma(x) = \exp(x)$  and  $\sigma$  subadditive.

*Case of leverage with  $\sigma(x) = \exp(x)$ .* Define

$$\hat{X}_s = \sum_{j=1, j \neq s}^{\infty} c_j \eta_{s-j}, \quad \hat{X}_s^{(m)} = \sum_{j=1, j \neq s}^m c_j \eta_{s-j},$$

and

$$\tilde{W}_{m,s} = |\exp(X_0 + \hat{X}_s) - \exp(X_0^{(m)} + \hat{X}_s^{(m)})|.$$

As previously, we see that  $\tilde{W}_{m,s}$  converges to 0 in  $L^q$  for some  $q > \alpha$ . Thus, we obtain

$$\sum_{i=1}^n \delta_{(i/n, Y_{n,i}^{(m)})} \Rightarrow \sum_{s=0}^h \sum_{k=1}^{\infty} \delta_{(t_k, j_{k,s}^{(m)} e_s)} \quad \text{as } n \rightarrow +\infty,$$

where  $\sum_{k=1}^{\infty} \delta_{(t_k, j_{k,0}^{(m)})}, \dots, \sum_{k=1}^{\infty} \delta_{(t_k, j_{k,h}^{(m)})}$  are independent Poisson processes with respective mean measures  $\lambda_{s,m}(dx)$ ,  $s = 0, \dots, h$ , defined in (6.6)–(6.7) with the constants  $d_+^{(m)}(s)$  and  $d_-^{(m)}(s)$  that appear therein given by

$$d_+^{(m)}(s) = d_+(s) \frac{\mathbb{E}[\exp(\alpha(X_0^{(m)} + \hat{X}_s^{(m)}))]}{\mathbb{E}[\exp(\alpha(X_0 + \hat{X}_s))]}, \quad d_-^{(m)}(s) = d_-(s) \frac{\mathbb{E}[\exp(\alpha(X_0^{(m)} + \hat{X}_s^{(m)}))]}{\mathbb{E}[\exp(\alpha(X_0 + \hat{X}_s))]}.$$

Since  $|\tilde{W}_{m,s}|$  converges to 0 in  $L^q$ , we obtain

$$\sum_{k=1}^{\infty} \delta_{(t_k, j_{k,s}^{(m)})} \Rightarrow \sum_{k=1}^{\infty} \delta_{(t_k, j_{k,s})} \quad \text{as } m \rightarrow +\infty, \quad s = 0, \dots, h.$$

Then, for  $s = 1, \dots, h$ , we obtain, with  $\tilde{Z}_0^{(s)} = Z_0 \exp(c_s \eta_0)$ , for  $\varepsilon > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow +\infty} n \mathbb{P}(b_n^{-1} |Y_0Y_s - Y_0^{(m)}Y_s^{(m)}| > \varepsilon) &= \limsup_{n \rightarrow +\infty} n \mathbb{P}(b_n^{-1} |Z_0\tilde{Z}_0^{(s)}| |\tilde{W}_{m,s}| > \varepsilon) \\ &\leq C\varepsilon^{-\alpha} \mathbb{E}[|\tilde{W}_{m,s}|^\alpha], \end{aligned}$$

which converges to 0 as  $m \rightarrow +\infty$ . This proves (6.13) and concludes the proof in the case of leverage with  $\sigma(x) = \exp(x)$ .

*Case of leverage with  $\sigma$  subadditive.* We have to bound

$$n \mathbb{P}(|Z_0Z_s| |\sigma(X_0)\sigma(X_s) - \sigma(X_0^{(m)})\sigma(X_s^{(m)})| > \varepsilon b_n).$$

It suffices to bound two terms:

$$\begin{aligned} I_1(n, m) &= n \mathbb{P}(|Z_0Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| |\sigma(X_s^{(m)})| > \varepsilon b_n), \\ I_2(n, m) &= n \mathbb{P}(|Z_0Z_s| |\sigma(X_0)| |\sigma(X_s) - \sigma(X_s^{(m)})| > \varepsilon b_n). \end{aligned}$$

Recall that  $X_s^{(m)} = \hat{X}_s^{(m)} + c_s \eta_0$  and  $X_s = \hat{X}_s + c_s \eta_0$ . By the subadditivity of  $\sigma$ , we have, for some constant  $\delta$ ,

$$I_1(n, m) \leq n \mathbb{P}(|Z_0 Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| \sigma(\hat{X}_s^{(m)}) > C \varepsilon b_n) + n \mathbb{P}(|Z_0 Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| \sigma(c_s \eta_0) > \delta \varepsilon b_n).$$

The product  $Z_0 Z_s$  is independent of  $|\sigma(X_0) - \sigma(X_0^{(m)})| \sigma(\hat{X}_s^{(m)})$  and tail equivalent to  $Y_0 Y_1$ ; thus, we obtain

$$\limsup_{n \rightarrow +\infty} n \mathbb{P}(|Z_0 Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| \sigma(\hat{X}_s^{(m)}) > \delta \varepsilon b_n) \leq C \varepsilon^{-\alpha} \mathbb{E}[|\sigma(X_0) - \sigma(X_0^{(m)})|^\alpha \sigma^\alpha(\hat{X}_s^{(m)})].$$

We have already seen that  $\sigma(X_0^{(m)})$  converges to  $\sigma(X_0)$  in  $L^\alpha$ ; thus, the latter expression converges to 0 as  $m \rightarrow +\infty$ . By assumption,  $\sigma(c_s \eta_0) |Z_0 Z_s|$  is either tail equivalent to  $|Z_0 Z_s|$  or  $\mathbb{E}[\sigma^q(c_s \eta_0) |Z_0 Z_s|^q] < \infty$  for some  $q > \alpha$ , and since it is independent of  $|\sigma(X_0) - \sigma(X_0^{(m)})|$ , we obtain

$$\limsup_{n \rightarrow +\infty} n \mathbb{P}(\sigma(c_s \eta_0) |Z_0 Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| > \varepsilon b_n) \leq C \varepsilon^{-\alpha} \mathbb{E}[|\sigma(X_0) - \sigma(X_0^{(m)})|^\alpha],$$

where  $C = 0$  in the latter case. In both cases, this yields

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} n \mathbb{P}(\sigma(c_s \eta_0) |Z_0 Z_s| |\sigma(X_0) - \sigma(X_0^{(m)})| > \varepsilon b_n) = 0.$$

Thus, we have obtained  $\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} I_1(n, m) = 0$ .

For the term  $I_2(n, m)$ , we use assumption (3.2) with  $x = c_s \eta_0$ ,  $y = \hat{X}_s$ , and  $z = \hat{X}_s^{(m)}$ . Thus,

$$I_2(n, m) \leq n \mathbb{P}(|Z_0 Z_s| (\sigma(c_s \eta_0) \vee 1) \tilde{W}_{m,s} > \varepsilon b_n),$$

with

$$\tilde{W}_{m,s} = \sigma(X_0) \{(\sigma(\hat{X}_s) \vee 1) + (\sigma(\hat{X}_s^{(m)}) \vee 1)\} |\hat{X}_s - \hat{X}_s^{(m)}|.$$

Note that  $\tilde{W}_{m,s}$  is independent of  $|Z_0 Z_s| (\sigma(c_s \eta_0) \vee 1)$  and  $\tilde{W}_{m,s}$  converges to 0 when  $m \rightarrow +\infty$  in  $L^q$  for some  $q > \alpha$ . Since  $|Z_0 Z_s| \sigma(c_s \eta_0)$  is tail equivalent to  $|Y_0 Y_1|$  or has a finite moment of order  $q'$  for some  $q' > \alpha$ , we have

$$\limsup_{n \rightarrow +\infty} n \mathbb{P}(|Z_0 Z_s| (\sigma(c_s \eta_0) \vee 1) \tilde{W}_{m,s} > \varepsilon b_n) \leq C \mathbb{E}[\tilde{W}_{m,s}^\alpha],$$

where the constant  $C$  can be 0 in the latter case. In both cases, we conclude that

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} n \mathbb{P}(|Z_0 Z_s| (\sigma(c_s \eta_0) \vee 1) \tilde{W}_{m,s} > \varepsilon b_n) = 0.$$

### 6.2. Proof of Theorem 4.1

We start by studying  $S_{p,n}$ . Write

$$\begin{aligned} \sum_{i=1}^{[nt]} (|Y_i|^p - \mathbb{E}[|Y_0|^p]) &= \sum_{i=1}^{[nt]} (|Y_i|^p - \mathbb{E}[|Y_i|^p | \mathcal{F}_{i-1}]) + \sum_{i=1}^{[nt]} (\mathbb{E}[|Y_i|^p | \mathcal{F}_{i-1}] - \mathbb{E}[|Y_0|^p]) \\ &=: M_n(t) + R_n(t). \end{aligned}$$

Note that  $E[|Y_i|^p \mid \mathcal{F}_{i-1}] = E[|Z_0|^p] \sigma^p(X_i)$  is a function of  $X_i$  and does not depend on  $Z_i$ . Then, by [2, Theorem 6], for  $\tau_p(1 - H) < \frac{1}{2}$ , we have

$$n^{-1} \rho_n^{-\tau_p/2} R_n \xrightarrow{\mathcal{D}} \frac{J_{\tau_p}(\sigma^p) E[|Z_1|^p]}{\tau_p!} R_{\tau_p, H}. \tag{6.14}$$

If  $\tau_p(1 - H) > \frac{1}{2}$  then by [2, Theorem 4] we obtain

$$n^{-1/2} R_n \xrightarrow{\mathcal{D}} \zeta E[|Z_0|^p] B, \tag{6.15}$$

where  $B$  is the standard Brownian motion and  $\zeta^2 = \text{var}(\sigma^p(X_0)) + 2 \sum_{i=1}^{\infty} \text{cov}(\sigma^p(X_0), \sigma^p(X_i))$ . We will show that, under the assumptions of Theorem 4.1, we have

$$a_n^{-p} M_n \xrightarrow{\mathcal{D}} L_{\alpha/p}. \tag{6.16}$$

Convergences (6.14), (6.15), and (6.16) conclude the proof of the theorem. We now prove (6.16). The proof is very similar to the proof of the convergence of the partial sum of an i.i.d. sequence in the domain of attraction of a stable law to a Lévy stable process. The differences are some additional technicalities. See, e.g. [27, Proof of Theorem 7.1] for more details. For  $0 < \varepsilon < 1$ , decompose it further as

$$\begin{aligned} M_n(t) &= \sum_{i=1}^{[nt]} \{ |Y_i|^p \mathbf{1}_{\{|Y_i| < \varepsilon a_n\}} - E[|Y_i|^p \mathbf{1}_{\{|Y_i| < \varepsilon a_n\}} \mid \mathcal{F}_{i-1}] \} \\ &\quad + \sum_{i=1}^{[nt]} \{ |Y_i|^p \mathbf{1}_{\{|Y_i| > \varepsilon a_n\}} - E[|Y_i|^p \mathbf{1}_{\{|Y_i| > \varepsilon a_n\}} \mid \mathcal{F}_{i-1}] \} \\ &=: M_n^{(\varepsilon)}(t) + \tilde{M}_n^{(\varepsilon)}(t). \end{aligned}$$

The term  $\tilde{M}_n^{(\varepsilon)}(\cdot)$  is treated using the point process convergence. Since, for any  $\varepsilon > 0$ , the summation functional is almost surely continuous from the set of Radon measures on  $[0, 1] \times [\varepsilon, \infty)$  onto  $\mathcal{D}([0, 1], \mathbb{R})$  with respect to the distribution of the Poisson point process with mean measure  $\nu_0$  (see, e.g. [27, p. 215]), from Proposition 3.1 we conclude that

$$a_n^{-p} \sum_{i=1}^{[n \cdot]} |Y_i|^p \mathbf{1}_{\{|Y_i| > \varepsilon a_n\}} \xrightarrow{\mathcal{D}} \sum_{t_k \leq \cdot} |j_k|^p \mathbf{1}_{\{|j_k| > \varepsilon\}}. \tag{6.17}$$

Taking the expectation in (6.17) we obtain

$$\lim_{n \rightarrow +\infty} [nt] a_n^{-p} E[|Y_0|^p \mathbf{1}_{\{|Y_1| > \varepsilon a_n\}}] = t \int_{\{x: |x| > \varepsilon\}} |x|^p \lambda_0(dx) \tag{6.18}$$

uniformly with respect to  $t \in [0, 1]$  since it is a sequence of increasing functions with a continuous limit. Furthermore, we claim that

$$a_n^{-p} \left| \sum_{i=1}^{[nt]} \{ E[|Y_0|^p \mathbf{1}_{\{|Y_1| > \varepsilon a_n\}}] - E[|Y_i|^p \mathbf{1}_{\{|Y_i| > \varepsilon a_n\}} \mid \mathcal{F}_{i-1}] \} \right| \xrightarrow{\mathbb{P}} 0 \tag{6.19}$$

uniformly in  $t \in [0, 1]$ . We use the variance inequality (A.3) below to bound the variance of

the last expression by

$$a_n^{-2p} [nt]^2 \rho_{[nt]} \text{var}(\mathbb{E}[|Y_1|^p \mathbf{1}_{\{|Y_1| > \varepsilon a_n\}} \mid \mathcal{F}_0]) \leq a_n^{-2p} [nt]^2 \rho_{[nt]} \mathbb{E}[(\mathbb{E}[|Y_1|^p \mathbf{1}_{\{|Y_1| > \varepsilon a_n\}} \mid \mathcal{F}_0])^2].$$

If  $p < \alpha < 2p$ , by Karamata’s theorem (see [27, p. 25]) and Potter’s bound,

$$\mathbb{E}[\sigma^p(x) | Z_1|^p \mathbf{1}_{\{|\sigma(x)Z_1| > \varepsilon a_n\}}] \leq Cn^{-1} a_n^p \frac{\bar{F}_Z(\varepsilon a_n / \sigma(x))}{\bar{F}_Z(a_n)} \leq Cn^{-1} a_n^p \sigma^{\alpha+\varepsilon}(x).$$

Since, by assumption,  $\mathbb{E}[\sigma^{2\alpha+2\varepsilon}(X_0)] < \infty$  for some  $\varepsilon > 0$ , for each  $t$ , we have

$$\begin{aligned} \text{var}\left(a_n^{-p} \sum_{i=1}^{[nt]} \{\mathbb{E}[|Y_0|^p \mathbf{1}_{\{|Y_0| > \varepsilon a_n\}}] - \mathbb{E}[|Y_i|^p \mathbf{1}_{\{|Y_i| > \varepsilon a_n\}} \mid \mathcal{F}_{i-1}]\}\right) &\leq Cn^{-2} [nt]^2 \rho_{[nt]} \\ &\leq Cn^{2H-2+\varepsilon} t^{2H-\varepsilon}, \end{aligned}$$

where the last bound is obtained for some  $\varepsilon > 0$  by Potter’s bound. This proves convergence of the finite-dimensional distribution to 0 and tightness in  $\mathcal{D}([0, 1], \mathbb{R})$ . As in [27, p. 216], we now argue that (6.17), (6.18), and (6.19) imply that

$$a_n^{-p} \tilde{M}_n^{(\varepsilon)} \xrightarrow{\mathcal{D}} L_{\alpha/p},$$

and it also holds that  $L_{\alpha/p}^{(\varepsilon)} \xrightarrow{\mathcal{D}} L_{\alpha/p}$  as  $\varepsilon \rightarrow 0$ . Therefore, to show (6.16), it suffices to show the negligibility of  $a_n^{-p} M_n^{(\varepsilon)}$ . By Doob’s martingale inequality we obtain

$$\begin{aligned} \mathbb{E}\left[\left(\sup_{t \in [0, 1]} a_n^{-p} \sum_{i=1}^{[nt]} \{\mathbb{E}[|Y_i|^p \mathbf{1}_{\{|Y_i| < \varepsilon a_n\}}] - \mathbb{E}[|Y_i|^p \mathbf{1}_{\{|Y_i| < \varepsilon a_n\}} \mid \mathcal{F}_{i-1}]\}\right)^2\right] \\ \leq Cna_n^{-2p} \mathbb{E}[(|Y_1|^p \mathbf{1}_{\{|Y_1| < \varepsilon a_n\}} - \mathbb{E}[|Y_1|^p \mathbf{1}_{\{|Y_1| < \varepsilon a_n\}} \mid \mathcal{F}_0])^2] \\ \leq 4Cna_n^{-2p} \mathbb{E}[|Y_1|^{2p} \mathbf{1}_{\{|Y_1| < \varepsilon a_n\}}]. \end{aligned}$$

Recall that  $\alpha < 2p$ . By Karamata’s theorem (see [27, p. 25]),

$$\mathbb{E}[|Y_1|^{2p} \mathbf{1}_{\{|Y_1| < \varepsilon a_n\}}] \sim \frac{2\alpha}{2p - \alpha} (\varepsilon a_n)^{2p} \bar{F}_Y(\varepsilon a_n) \sim \frac{2\alpha}{2p - \alpha} \varepsilon^{2p-\alpha} a_n^{2p} n^{-1}.$$

Applying this and letting  $\varepsilon \rightarrow 0$  we conclude that  $a_n^{-p} M_n^{(\varepsilon)}$  is uniformly negligible in  $L^2$  and so in probability, and, thus, we conclude that  $a_n^{-p} M_n \xrightarrow{\mathcal{D}} L_{\alpha/p}$ .

For  $p > \alpha$ ,  $\mathbb{E}[|Y_0|^p] = \infty$ . In that case it is well known (see, e.g. [10, Theorem 3.1]) that the convergence of  $a_n^{-p} S_{p,n}$  to an  $\alpha/p$ -stable Lévy process follows directly from the convergence of the point process  $\sum_{i=1}^n \delta_{Y_i/a_n}$  to a Poisson point process, and that no centering is needed. In the present context, this entirely dispenses with the conditioning argument and the long-memory part does not appear. Therefore, convergence to a stable Lévy process always holds.

As for the sum  $S_n$ , since  $\mathbb{E}[Y_0] = \mathbb{E}[Z_0] = 0$ , the long-memory part  $R_n$  is identically vanishing, and, thus, in this case also, only the stable limit arises.

**6.3. Proof of Theorem 5.2**

Let  $U_i = |Y_i Y_{i+s}|$ . We now write

$$\begin{aligned} \sum_{i=1}^n (U_i^p - E[U_0^p]) &= \sum_{i=1}^n (U_i^p - E[U_i^p \mid \mathcal{F}_{i-1}]) + \sum_{i=1}^n (E[U_i^p \mid \mathcal{F}_{i-1}] - E[U_0^p]) \\ &= M_{n,s} + \sum_{i=1}^n K_p^*(X_i, \hat{X}_{i,s}) \\ &= M_{n,s} + T_{n,s}. \end{aligned}$$

As mentioned above, the second part is the partial sum of a sequence of a function of the bivariate Gaussian sequence  $(X_i, \hat{X}_{i,s})$ . The proof of the convergence to a stable law mimics the proof of Theorem 4.1. We split  $M_{n,s}$  between big jumps and small jumps. Write  $M_{n,s}^{(\varepsilon)} + \tilde{M}_{n,s}^{(\varepsilon)}$ , with

$$M_{n,s}^{(\varepsilon)} = \sum_{i=1}^n (U_i^p \mathbf{1}_{\{U_i \leq b_n \varepsilon\}} - E[U_i^p \mathbf{1}_{\{U_i \leq b_n \varepsilon\}} \mid \mathcal{F}_{i-1}]).$$

The point process convergence yields the convergence of the big jumps part by the same argument as in the proof of Theorem 4.1. In order to prove the asymptotic negligibility of the small jumps part, the only change that has to be made comes from the observation that  $\tilde{M}_{n,s}^{(\varepsilon)}$  is no longer a martingale. However, assuming for simplicity that we have  $(s + 1)n$  observations  $Y_i$ , we write, with  $U_{i,k} = U_{(s+1)i-k} = |Y_{(s+1)i-k} Y_{(s+1)i+s-k}|$ ,

$$M_{n,s}^{(\varepsilon)} = \sum_{k=0}^s \sum_{i=1}^n \{U_{i,k}^p \mathbf{1}_{\{U_{i,k} \leq b_n \varepsilon\}} - E[U_{i,k}^p \mathbf{1}_{\{U_{i,k} \leq b_n \varepsilon\}} \mid \mathcal{F}_{(s+1)i-k-1}]\} =: \sum_{k=0}^s M_{n,s,k}^{(\varepsilon)}.$$

Clearly, each  $M_{n,s,k}^{(\varepsilon)}$ ,  $k = 0, \dots, s$ , is a martingale with respect to the filtration  $\{\mathcal{F}_{i(s+1)}, 1 \leq i \leq n\}$ ; therefore, we can apply Doob’s inequality and conclude the proof with the same arguments as previously.

**6.4. Proof of Theorem 5.3**

Again, we mimic the proof of Theorem 4.1; however, some technical modifications are needed. We use the decomposition between small jumps and big jumps. To prove negligibility of the small jumps, we use the same splitting technique as in the proof of Theorem 5.2. To deal with the big jumps, the only adaptation needed is to obtain a bound for the quantity

$$b_n^{-2p} n^2 \rho_n E[(E[|Y_0 Y_s|^p \mathbf{1}_{\{|Y_0 Y_s| > \varepsilon b_n\}} \mid \mathcal{F}_{-1}])^2]. \tag{6.20}$$

To show that (6.19) still holds in the present context, we must prove that the expectation in (6.20) is of the order  $n^{-2} b_n^{2p}$ . The rest of the arguments to prove the convergence of the big jumps part remains unchanged. Note that  $E[|Y_0 Y_s|^p \mathbf{1}_{\{|Y_0 Y_s| > \varepsilon b_n\}} \mid \mathcal{F}_{-1}] = G(X_0, \hat{X}_{0,s})$ ; thus, we need an estimate for the bivariate function

$$G(x, y) = \sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(c_s \eta_0 + \zeta_s \zeta + y) \mathbf{1}_{\{|Z_0 Z_s| \sigma(c_s \eta_0 + \zeta_s \zeta + y) > \varepsilon b_n\}}],$$

where  $\zeta$  is a standard Gaussian random variable, independent of  $Z_0, \eta_0$ , and  $Z_s$ . We obtain this estimate first in the case  $\sigma(x) = \exp(x)$  and then for subadditive functions.



Let  $\sigma(x) = \exp(x)$ . As in the proof of point process convergence, we write

$$Y_0 Y_s = Z_0 Z_s \exp(c_s \eta_0) \exp(X_0 + \hat{X}_s).$$

By Lemma 2.2,  $Z_0 Z_s \exp(c_s \eta_0)$  is regularly varying and tail equivalent to  $Z_0 Z_s$ . Since  $\exp(p_{\zeta_s} \zeta)$  is independent of  $Z_0 Z_s \exp(c_s \eta_0)$  and has finite moments of all orders, we find that  $Z_0 Z_s \exp(c_s \eta_0) \exp(p_{\zeta_s} \zeta)$  is also tail equivalent to  $Z_0 Z_s$ , and, hence, to  $Y_0 Y_1$ . Thus, by Karamata's theorem and Potter's bounds, we obtain, for some  $\delta > 0$ ,

$$\begin{aligned} G(x, y) &= \exp(p(x + y)) \\ &\quad \times E[|Z_0 Z_s|^p \exp(p c_s \eta_0) \exp(p_{\zeta_s} \zeta) \mathbf{1}_{\{|Z_0 Z_s| \exp(p c_s \eta_0) \exp(p_{\zeta_s} \zeta) > \varepsilon b_n \exp(-y)\}}] \\ &\leq C n^{-1} b_n^p \exp(p x) \exp((p - \alpha + \delta)(y \vee 0)). \end{aligned}$$

Since the log-normal distribution has finite moments of all orders, we obtain  $E[G^2(X_0, \hat{X}_{0,s})] = O(n^{-2} b_n^{2p})$ , which is the required bound. This concludes the proof in the case  $\sigma(x) = \exp(x)$ .

Now let the assumptions of Proposition 3.4 be in force. Using the subadditivity of  $\sigma^p$ , we obtain  $G(x, y) \leq \sum_{i=1}^4 I_i(x, y)$  with

$$\begin{aligned} I_1(x, y) &= \sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(\vartheta_s) \mathbf{1}_{\{|Z_0 Z_s| \sigma(\vartheta_s) > \varepsilon b_n\}}], \\ I_2(x, y) &= \sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(y) \mathbf{1}_{\{|Z_0 Z_s| \sigma(y) > \varepsilon b_n\}}], \\ I_3(x, y) &= \sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(\vartheta_s) \mathbf{1}_{\{|Z_0 Z_s| \sigma(\vartheta_s) > \varepsilon b_n\}}], \\ I_4(x, y) &= \sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(y) \mathbf{1}_{\{|Z_0 Z_s| \sigma(\vartheta_s) > \varepsilon b_n\}}], \end{aligned}$$

where for brevity we have defined  $\vartheta_s = c_s \eta_0 + \zeta_s \zeta$ . We now give the bounds for  $E[I_j^2(X_0, \hat{X}_{0,s})]$ ,  $j = 1, 2, 3, 4$ . Since, by the assumptions,  $|Z_0 Z_s| \sigma(\vartheta_s)$  is tail equivalent to  $|Z_0 Z_s|$ , Karamata's theorem yields

$$\sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(\vartheta_s) \mathbf{1}_{\{|Z_0 Z_s| \sigma(\vartheta_s) > \varepsilon b_n\}}] \leq C n^{-1} b_n^p \sigma^p(x),$$

and since  $E[\sigma^{2p}(X_0)] < \infty$  by assumption, we obtain, by integrating,  $E[I_1^2(X_0, \hat{X}_{0,s})] = O(n^{-2} b_n^{2p})$ . For  $I_2$ , again using Karamata's theorem and Potter's bound, we obtain, for some  $\delta > 0$ ,

$$\sigma^p(x) E[|Z_0 Z_s|^p \sigma^p(y) \mathbf{1}_{\{|Z_0 Z_s| \sigma(y) > \varepsilon b_n\}}] \leq C n^{-1} b_n^p \sigma^p(x) (\sigma(y) \vee 1)^{p-\alpha+\delta}.$$

Since  $|Z_0| \sigma(\vartheta_s)$  is tail equivalent to  $|Z_0|$  and  $Z_s$  is independent of  $Z_0 \sigma(\vartheta_s)$ , we easily obtain a bound for the tail of  $|Z_0 Z_s| (\sigma(\vartheta_s) \vee 1)$ :

$$P(|Z_0 Z_s| (\sigma(\vartheta_s) \vee 1) > x) \leq P(|Z_0 Z_s| \sigma(\vartheta_s) > x) + P(|Z_0 Z_s| > x) \leq C P(Z_0 Z_s > x)$$

for large  $x$ . Thus, applying Karamata's theorem and Potter's bound to  $|Z_0 Z_s|$  yields, for some arbitrarily small  $\delta > 0$ ,

$$I_3(x, y) \leq C \sigma^p(x) E[|Z_0 Z_s|^p \mathbf{1}_{\{\sigma(y) |Z_0 Z_s| > \varepsilon b_n\}}] \leq C n^{-1} b_n^p \sigma^p(x) (\sigma(y) \vee 1)^{\alpha+\delta},$$

and, thus, we conclude that  $E[I_3^2(X_0, \hat{X}_{0,s})] = O(n^{-2} b_n^{2p})$ . Finally, we write,

$$I_4(x, y) \leq \sigma^p(x) \sigma^p(y) E[|Z_0 Z_s|^p (\sigma^p(\vartheta_s) \vee 1) \mathbf{1}_{\{|Z_0 Z_s| (\sigma(\vartheta_s) \vee 1) > \varepsilon b_n\}}]$$

and by the same argument as for  $I_3$  we obtain  $E[I_4^2(X_0, \hat{X}_{0,s})] = O(n^{-2} b_n^{2p})$ .

### Appendix A. Gaussian long-memory sequences

For the sake of completeness, we recall in this appendix the main definitions and results pertaining to Hermite coefficients and expansions of square-integrable functions with respect to a possibly nonstandard multivariate Gaussian distribution. Expansions with respect to the multivariate standard Gaussian distribution are easy to obtain and describe. The theory for nonstandard Gaussian vectors is more cumbersome. The main reference is [2].

#### A.1. Hermite coefficients and rank

Let  $G$  be a function defined on  $\mathbb{R}^k$ , and let  $\mathbf{X} = (X^{(1)}, \dots, X^{(k)})$  be a  $k$ -dimensional centered Gaussian vector with covariance matrix  $\Gamma$ . The Hermite coefficients of  $G$  with respect to  $\mathbf{X}$  are defined as

$$J(G, \mathbf{X}, \mathbf{q}) = \mathbb{E} \left[ G(\mathbf{X}) \prod_{j=1}^k H_{q_j}(X^{(j)}) \right],$$

where  $\mathbf{q} = (q_1, \dots, q_k) \in \mathbb{N}^k$ . If  $\Gamma$  is the  $k \times k$  identity matrix (denoted by  $I_k$ ), i.e. the components of  $\mathbf{X}$  are i.i.d. standard Gaussian, then the corresponding Hermite coefficients are denoted by  $J^*(G, \mathbf{q})$ . The Hermite rank of  $G$  with respect to  $\mathbf{X}$  is the smallest integer  $\tau$  such that

$$J(G, \mathbf{X}, \mathbf{q}) = 0 \quad \text{for all } \mathbf{q} \text{ such that } 0 < |q_1 + \dots + q_k| < \tau.$$

#### A.2. Variance inequalities

Now consider a  $k$ -dimensional stationary centered Gaussian process  $\{X_i, i \geq 0\}$  with covariance function  $\rho_n(i, j) = \mathbb{E}[X_0^{(i)} X_n^{(j)}]$ , and assume that either

$$\sum_{n=0}^{\infty} |\rho_n(i, j)| < \infty \quad \text{for all } 1 \leq i, j \leq k \tag{A.1}$$

or there exist  $H \in (\frac{1}{2}, 1)$  and a function  $\ell$  slowly varying at  $\infty$  such that

$$\lim_{n \rightarrow +\infty} \frac{\rho_n(i, j)}{n^{2H-2}\ell(n)} = b_{i,j}, \tag{A.2}$$

where the  $b_{i,j}$ s are not identically 0. Define  $\rho_n = n^{2H-2}\ell(n)$ . Then, we have the following cases.

- If (A.2) holds and  $2\tau(1 - H) < 1$ , then, for any function  $G$  with Hermite rank  $\tau$  with respect to  $\mathbf{X}_0$ ,

$$\text{var} \left( n^{-1} \sum_{j=1}^n G(\mathbf{X}_j) \right) \leq C \rho_n^\tau \text{var}(G(\mathbf{X}_0)). \tag{A.3}$$

- If (A.2) holds and  $2\tau(1 - H) > 1$ , then, for any function  $G$  with Hermite rank  $\tau$  with respect to  $\mathbf{X}_0$ ,

$$\text{var} \left( \sum_{j=1}^n G(\mathbf{X}_j) \right) \leq C n \text{var}(G(\mathbf{X}_0)). \tag{A.4}$$

- If (A.1) holds then (A.4) still holds.

In all these cases, the constant  $C$  depends only on the Gaussian process  $\{X_i\}$  and not on the function  $G$ . Bounds (A.3) and (A.4) are Equations 3.10 and 2.40 of [2], respectively. Bound (A.4), under assumption (A.1), is a consequence of Equation 2.18 of [2, Theorem 2].

**A.3. Limit theorems**

We now recall [2, Theorem 6]. Again, let  $\{X_i\}$  be a stationary sequence of  $k$ -dimensional Gaussian vectors with covariance matrix  $G$  and such that (A.2) holds, and let  $\tau$  be the Hermite rank of  $G$  with respect to  $X_0$ . If  $\tau(1 - H) < \frac{1}{2}$ , there exists a process  $R_{G,\tau,H}$  such that

$$\frac{1}{n\rho_n^{\tau/2}} \sum_{i=1}^{[n\cdot]} (G(X_i) - E[G(X_0)]) \xrightarrow{\mathcal{D}} R_{G,\tau,H}.$$

In particular, if  $k = 1$  then

$$\frac{1}{n\rho_n^{\tau/2}} \sum_{i=1}^{[n\cdot]} \{G(X_i) - E[G(X_0)]\} \xrightarrow{\mathcal{D}} \frac{J_\tau(G)}{\tau!} R_{\tau,H},$$

where  $J_\tau(G) = E[G(X_1)H_\tau(X_1)]$  and  $R_{\tau,H}$  is the so-called Hermite or Rosenblatt process of order  $\tau$ , defined as a  $\tau$ -fold stochastic integral

$$R_{\tau,H}(t) = K_1(\tau, H) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\exp(it(x_1 + \dots + x_\tau)) - 1}{x_1 + \dots + x_\tau} \times \prod_{i=1}^{\tau} x_i^{-H+1/2} W(dx_1) \dots W(dx_\tau),$$

where  $W$  is an independently scattered Gaussian random measure with Lebesgue control measure and

$$K_1^2(\tau, H) = \frac{(\tau(H - 1) + 1)(2\tau(H - 1) + 1)}{\tau! \{2\Gamma(2 - 2H) \sin \pi(H - 1/2)\}^\tau}.$$

In particular, for  $\tau = 1$ , then the limiting process is the fractional Brownian motion, which is a Gaussian process, so

$$\frac{1}{n\rho_n^{1/2}} \sum_{i=1}^n \{G(X_i) - E[G(X_0)]\} \xrightarrow{\mathcal{D}} N\left(0, \frac{J(1)}{H(2H - 1)}\right).$$

On the other hand, if  $1 - \tau(1 - H) < \frac{1}{2}$  then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[n\cdot]} \{G(X_i) - E[G(X_0)]\} \xrightarrow{\mathcal{D}} \zeta B,$$

where  $B$  is the standard Brownian motion and  $\zeta^2 = \text{var}(G(X_0)) + 2\sum_{j=1}^{\infty} \text{cov}(G(X_0), G(X_j))$ , the latter series being absolutely summable.

**Acknowledgements**

We are grateful to an anonymous referee whose remarks led to a substantial improvement of our paper. The research of the first author was supported by an NSERC grant. The research of the second author was partially supported by the ANR grant ANR-08-BLAN-0314-02.

## References

- [1] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley, New York.
- [2] ARCONES, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Prob.* **22**, 2242–2274.
- [3] BAILLIE, R. T., BOLLERSLEV, T. AND MIKKELSEN, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **74**, 3–30.
- [4] BASRAK, B., DAVIS, R. A. AND MIKOSCH, T. (2002). Regular variation of GARCH processes. *Stoch. Process. Appl.* **99**, 95–115.
- [5] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1989). *Regular Variation* (Encyclopedia Math. Appl. **27**). Cambridge University Press.
- [6] BOLLERSLEV, T. AND MIKKELSEN, H. O. (1996). Modeling and pricing long memory in stock market volatility. *J. Econometrics* **73**, 151–184.
- [7] BREIDT, F. J. AND DAVIS, R. A. (1998). Extremes of stochastic volatility models. *Ann. Appl. Prob.* **8**, 664–675.
- [8] BREIDT, F. J., CRATO, N. AND DE LIMA, P. (1998). The detection and estimation of long memory in stochastic volatility. *J. Econometrics* **83**, 325–348.
- [9] DAVIS, R. A. (1983). Stable limits for partial sums of dependent random variables. *Ann. Prob.* **11**, 262–269.
- [10] DAVIS, R. A. AND HSING, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23**, 879–917.
- [11] DAVIS, R. A. AND MIKOSCH, T. (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation. In *Probability, Statistics and Seismology* (J. Appl. Prob. Spec. Vol. **38A**), ed. D. J. Daley, Applied Probability Trust, Sheffield, pp. 93–104.
- [12] DAVIS, R. AND RESNICK, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Prob.* **13**, 179–195.
- [13] DAVIS, R. AND RESNICK, S. (1985). More limit theory for the sample correlation function of moving averages. *Stoch. Process. Appl.* **20**, 257–279.
- [14] DAVIS, R. AND RESNICK, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**, 533–558.
- [15] DOUC, R., ROUEFF, F. AND SOULIER, P. (2008). On the existence of some ARCH( $\infty$ ) processes. *Stoch. Process. Appl.* **118**, 755–761.
- [16] EMBRECHTS, P. AND GOLDIE, C. M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc. Ser. A* **29**, 243–256.
- [17] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
- [18] GIRAITIS, L. AND SURGAILIS, D. (2002). ARCH-type bilinear models with double long memory. *Stoch. Process. Appl.* **100**, 275–300.
- [19] GIRAITIS, L., ROBINSON, P. M. AND SURGAILIS, D. (2000). A model for long memory conditional heteroscedasticity. *Ann. Appl. Prob.* **10**, 1002–1024.
- [20] GIRAITIS, L., LEIPUS, R., ROBINSON, P. M. AND SURGAILIS, D. (2004). LARCH, leverage, and long memory. *J. Financial Econometrics* **2**, 177–210.
- [21] HORVÁTH, L. AND KOKOSZKA, P. (2008). Sample autocovariances of long-memory time series. *Bernoulli* **14**, 405–418.
- [22] HOSKING, J. R. M. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econometrics* **73**, 261–284.
- [23] JACH, A., MCELROY, T. AND POLITIS, D. N. (2012). Subsampling inference for the mean of heavy-tailed long-memory time series. *J. Time Ser. Anal.* **33**, 96–111.
- [24] KOKOSZKA, P. S. AND TAQQU, M. S. (1996). Parameter estimation for infinite variance fractional ARIMA. *Ann. Statist.* **24**, 1880–1913.
- [25] MCELROY, T. AND POLITIS, D. (2007). Self-normalization for heavy-tailed time series with long memory. *Statistica Sinica* **17**, 199–220.
- [26] NELSON, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**, 347–370.
- [27] RESNICK, S. I. (2007). *Heavy-Tail Phenomena*. Springer, New York.
- [28] ROBINSON, P. M. (2001). The memory of stochastic volatility models. *J. Econometrics* **101**, 195–218.
- [29] ROBINSON, P. M. AND ZAFFARONI, P. (1997). Modelling nonlinearity and long memory in time series. In *Nonlinear Dynamics and Time Series* (Montreal, PQ, 1995; Fields Inst. Commun. **11**), American Mathematical Society, Providence, RI, pp. 161–170.
- [30] ROBINSON, P. M. AND ZAFFARONI, P. (1998). Nonlinear time series with long memory: a model for stochastic volatility. *J. Statist. Planning. Infer.* **68**, 359–371.

- [31] SURGAILIS, D. (2008). A quadratic ARCH( $\infty$ ) model with long memory and Lévy stable behavior of squares. *Adv. Appl. Prob.* **40**, 1198–1222.
- [32] SURGAILIS, D. AND VIANO, M.-C. (2002). Long memory properties and covariance structure of the EGARCH model. *ESAIM Prob. Statist.* **6**, 311–329 (electronic).
- [33] WHITT, W. (2002). *Stochastic-Process Limits*. Springer, New York.
- [34] WU, W. B., HUANG, Y. AND ZHENG, W. (2010). Covariances estimation for long-memory processes. *Adv. Appl. Prob.* **42**, 137–157.