hal-00624821, version 1 - 19 Sep 2011

Estimation of limiting conditional distributions for the heavy tailed long memory stochastic volatility process

Rafał Kulik* Philippe Soulier†

Abstract

We consider Stochastic Volatility processes with heavy tails and possible long memory in volatility. We study the limiting conditional distribution of future events given that some present or past event was extreme (i.e. above a level which tends to infinity). Even though extremes of stochastic volatility processes are asymptotically independent (in the sense of extreme value theory), these limiting conditional distributions differ from the i.i.d. case. We introduce estimators of these limiting conditional distributions and study their asymptotic properties. If volatility has long memory, then the rate of convergence and the limiting distribution of the centered estimators can depend on the long memory parameter (Hurst index).

1 Introduction

One of the empirical features of financial data is that log-returns are uncorrelated, but their squares, or absolute values, are dependent, possibly with long memory. Another important feature is that log-returns are heavy-tailed. There are two common classes of processes to model such behaviour: the generalized autoregressive conditional heteroscedastic (GARCH) process and the stochastic volatility (SV) process; the latter introduced by Breidt et al. [1998] and Harvey [1998]. The former class of models rules out long memory in the squares, while the latter allows for it. We will therefore concentrate in this paper on the class of SV processes, which we define now.

Let $\{Y_j, j \in \mathbb{Z}\}$ be the observed process (e.g. log-returns of some financial time series), and assume that it can be expressed as

$$Y_j = \sigma(X_j) Z_j \; . \tag{1}$$

where σ is some (possibly unknown) positive function, $\{Z_j, j \in \mathbb{Z}\}$ is an i.i.d. sequence and $\{X_j, j \in \mathbb{Z}\}$ is a stationary Gaussian process with mean zero, unit variance, autocovariance function $\{\gamma_n\}$, and independent from the i.i.d. sequence. The sequence $\sigma(X_j)$ can be seen as a proxy for the volatility. We will assume that either $\{X_j\}$ is weakly dependent in the sense that

$$\sum_{j=1}^{\infty} |\gamma_j| < \infty \; , \tag{2}$$

---

[*] University of Ottawa

[†] Université de Paris-Ouest

hal-00624821, version 1 - 19 Sep 2011

or that it has long memory with Hurst index $H \in (1/2, 1)$, i.e.

$$\gamma_n = \mathrm{cov}(X_0, X_n) = n^{2H-2}\ell(n) \tag{3}$$

where ℓ is a slowly varying function.

Furthermore, we assume that the marginal distribution F_Z of the i.i.d. sequence $\{Z_j\}$ has a regularly varying right tail with index $\alpha > 0$, i.e., for all positive y,

$$\lim_{t\to\infty} \mathbb{P}(Z > ty \mid Z > t) = \lim_{t\to\infty} \frac{\bar{F}_Z(ty)}{\bar{F}_Z(t)} = y^{-\alpha} . \tag{4}$$

Examples of heavy tailed distributions include the stable distributions with index $\alpha \in (0, 2)$, the t distribution with α degrees of freedom, and the Pareto distribution with index α.

By Breiman's lemma Breiman [1965], Resnick [2007], if $\mathbb{E}[\sigma^{\alpha+\epsilon}(X)] < \infty$ for some $\epsilon > 0$, then the marginal distribution of $\{Y_j\}$ also has a regularly varying right tail with index α and

$$\lim_{x\to\infty} \frac{\mathbb{P}(Y > xy)}{\mathbb{P}(Z > x)} = \mathbb{E}[\sigma^\alpha(X)]y^{-\alpha} , \tag{5}$$

where X, Y and Z denote random variables with the same joint distribution as X_0, Y_0 and Z_0.

Estimation and test of the possible long memory of such processes has been studied by Hurvich et al. [2005]. Estimation of the tail of the marginal distribution by the Hill estimator has been studied in Kulik and Soulier [2011].

In this paper we are concerned with certain extremal properties of the finite dimensional joint distributions of the process $\{Y_j\}$ when Z is heavy tailed and the Gaussian process $\{X_j\}$ possibly has long memory.

From the extreme value point of view, there is a significant distinction between the GARCH and SV models. In the first one, exceedances over a large threshold are asymptotically dependent and extremes do cluster. In the SV model, exceedances are asymptotically independent. More precisely, for any positive integer m, and positive real numbers x, y,

$$\lim_{t\to\infty} t\mathbb{P}(Y_0 > a(t)x \ , \ Y_m > a(t)y) = 0 , \tag{6}$$

where $a(t) = F_Z^{\leftarrow}(1 - 1/t)$ and $F_Z^{\leftarrow}$ is the left continuous inverse of F_Z. This holds since it can be easily shown by a conditioning argument that

$$\mathbb{P}(Y_0 > t \ , \ Y_m > t) \sim \mathrm{c} \times \mathbb{P}(Y_0 > t)^2 \ , \quad t \to \infty , \tag{7}$$

for some positive constant c.

The above observations may lead to the incorrect conclusion that, for the SV process, there is no spillover from past extreme observations onto future values and from the extremal behaviour point of view we can treat the SV process as an i.i.d. sequence. However, under the assumptions stated previously, it holds that

$$\lim_{t\to\infty} \mathbb{P}(Y_m \le y \mid Y_0 > t) = \frac{\mathbb{E}[\sigma^\alpha(X_0)F_Z(y/\sigma(X_m))]}{\mathbb{E}[\sigma^\alpha(X_m)]} . \tag{8}$$

hal-00624821, version 1 - 19 Sep 2011

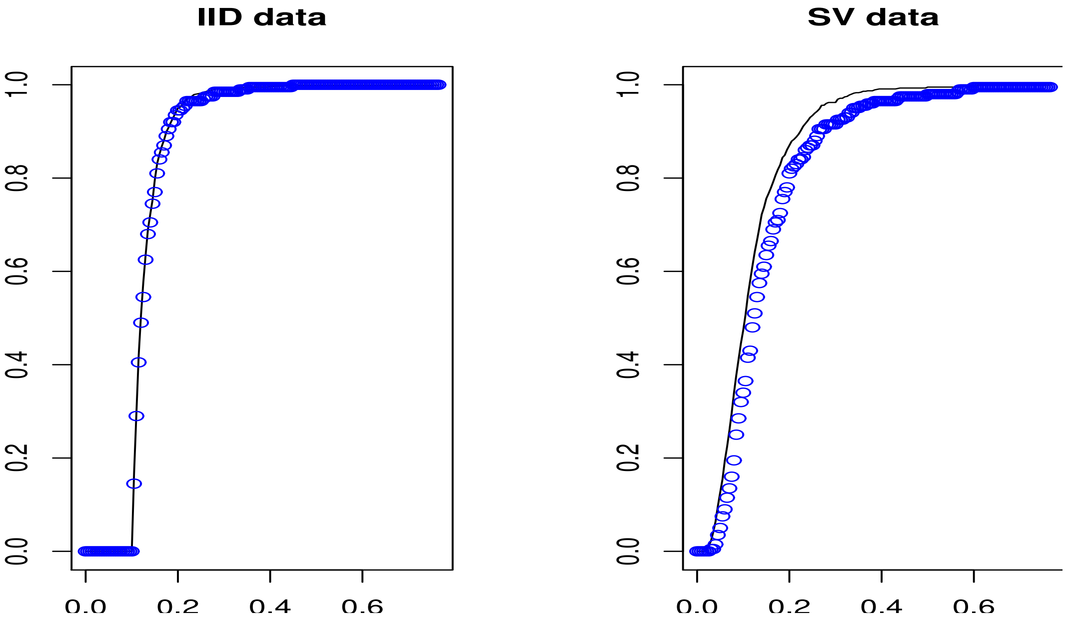

Figure 1: Empirical Conditional Distribution (points) and Empirical Distribution (solid line) for SV model (right panel) and i.i.d. data (left panel)

Therefore, the limiting conditional distribution is influenced by the dependence structure of the time series. To illustrate this, we show in Figure 1 estimates of the standard distribution function and of the conditional distribution for a simulated SV process. Clearly, the two estimated distributions are different, as suggested by (8). For a comparison, we also plot the corresponding estimates for i.i.d. data. Other kind of extremal events can be considered, for instance, we may be interested in the conditional distribution of some future values given that a linear combination (portfolio) of past values is extremely large, or that two consecutive values are large. As in Equation (8), in each of these cases, a proper limiting distribution can be obtained. To give a general framework for these conditional distributions, we introduce a modified version of the extremogram of Davis and Mikosch [2009]. For fixed positive integers $h < m$ and $h' \geq 0$, Borel sets $A \subset \mathbb{R}^h$ and $B \subset \mathbb{R}^{h'+1}$, we are interested in the limit denoted by $\rho(A, B, m)$, if it exists:

$$\rho(A, B, m) = \lim_{t\to\infty} \mathbb{P}((Y_m, \dots, Y_{m+h'}) \in B \mid (Y_1, \dots, Y_h) \in tA) \ . \tag{9}$$

The set A represents the type of events considered. For instance, if we choose $A = \{(x, y, z) \in [0, \infty)^3 \mid x + y + z > 1\}$, then for large t, $\{(Y_{-2}, Y_{-1}, Y_0) \in tA\}$ is the event that the sum of last three observations was extremely large. The set B represents the type of future events of interest.

In the original definition of the extremogram of Davis and Mikosch [2009], the set B is also dilated by t. This is well suited to the context of asymptotic dependence, as arises in GARCH processes. But in the context of asymptotic independence, this would yield a degenerate limit: if $h < m$, then for most sets A and B,

$$\lim_{t\to\infty} \mathbb{P}((Y_m, \dots, Y_{m+h'}) \in tB \mid (Y_1, \dots, Y_h) \in tA) = 0 \ .$$

hal-00624821, version 1 - 19 Sep 2011

The general aim of this paper is to investigate the existence of these limiting conditional distributions appearing in (9) and their statistical estimation. The paper is the first step towards understanding conditional laws for stochastic volatility models. Although we provide theoretical properties of estimators, their practical use should be investigated in conjunction with resampling techniques. This is a topic of authors' current research.

The paper is structured as follows. In Section 2, we present a general framework that enables to treat various examples in a unified way. In Section 3 we present the estimation procedure with appropriate limiting results.

The proofs are given in Section 4. In the Appendix we collect relevant results on second order regular variation, (long memory) Gaussian processes, and criteria for tightness.

We conclude this introduction by gathering some notation that will be used throughout the paper. We denote convergence in probability by $\to_P$, weak convergences of sequences of random variables or vectors by $\to_d$ and weak convergence in the Skorokhod space $\mathcal{D}(\mathbb{R}^q)$ of cadlag functions defined on $\mathbb{R}^q$ endowed with the J_1 topology by $\Rightarrow$.

Boldface letters denote vectors. Product of vectors and inequalities between vectors are taken componentwise: $\mathbf{u} \cdot \mathbf{v} = (u_1 v_1, \dots, u_d v_d)$; $\mathbf{x} \le \mathbf{y}$ if and only if $x_i \le y_i$ for all $i = 1, \dots, d$. The (multivariate) interval $(\boldsymbol{\infty}, \mathbf{y}]$ is defined accordingly: $(\boldsymbol{\infty}, \mathbf{y}] = \prod_{i=1}^{d} (-\infty, y_i]$.

For any univariate process $\{\xi_j\}$ and any integers $h \le h'$, let $\boldsymbol{\xi}_{h,h'}$ denote the $(h' - h + 1)$-dimensional vector $(\xi_h, \dots, \xi_{h'})$.

For $A \subset \mathbb{R}^d$ and $\mathbf{u} \in (0, \infty)^d$, $\mathbf{u}^{-1} \cdot A = \{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{u} \cdot \mathbf{x} \in A\}$.

If $\mathbf{X}$ is a random vector, we denote by $L^p(\mathbf{X})$ the set of measurable functions f such that $\mathbb{E}[|f(\mathbf{X})|^p] < \infty$.

For any univariate process $\{\xi_j\}$ and any integers $h \le h'$, let $\boldsymbol{\xi}_{h,h'}$ denote the $(h' - h + 1)$-dimensional vector $(\xi_h, \dots, \xi_{h'})$.

The σ-field generated by the process $\{X_j\}$ is denoted by $\mathcal{X}$.

2 Regular variation on subcones

Since we considered dilated sets tA, where $A \subset \mathbb{R}^h$ for some integer $h > 0$, it is natural to consider cones, that is subsets $\mathcal{C}$ of $[0, \infty]^h$ such that $tx \in \mathcal{C}$ for all $x \in \mathcal{C}$ and $t > 0$. The next definition is related to the concept of regular variation on cones of Resnick [2008]. We endow $\mathbb{R}^h$ with the topology induced by any norm and $[0, \infty]^h$ is the compactification of $[0, \infty)^h$. A subset A of $[0, \infty]^h \setminus \{0\}$ is relatively compact if its closure is compact. See Resnick [1987] for more details. We first state a general assumption and will give examples afterwards.

Assumption 1. *Let h be a fixed positive integer. Let $\mathcal{C}$ be a subcone of $[0, \infty]^h \setminus \{\mathbf{0}\}$ such that, (i) for all relatively compact subsets A of $\mathcal{C}$ and all $\mathbf{u} \in (0, \infty)^h$, $\mathbf{u}^{-1} \cdot A$ is relatively compact in $\mathcal{C}$, and (ii) there exists a function $g_{\mathcal{C}}$ and a non degenerate Radon measure $\nu_{\mathcal{C}}$ on $\mathcal{C}$ such that*

$$\lim_{t \to \infty} \frac{\mathbb{P}(\mathbf{Z}_{1,h} \in tA)}{g_{\mathcal{C}}(\bar{F}_Z(t))} = \nu_{\mathcal{C}}(A) \; . \tag{10}$$

hal-00624821, version 1 - 19 Sep 2011

Note that in the case $h = 1$, the cone $\mathcal{C} = (0, \infty)$ and Assumption 1 is nothing more than the regular variation of the tail of Z_1.

Assumption 1 implies that the function $g_{\mathcal{C}}$ is regularly varying at 0 with index $\beta_{\mathcal{C}} \in (0, \infty)$ and the measure $\nu_{\mathcal{C}}$ is homogeneous with index $-\alpha\beta_{\mathcal{C}}$. For $s \geq 1$, define

$$T_{\mathcal{C}}(s) = \lim_{t\to\infty} \frac{g_{\mathcal{C}}(\bar{F}_Z(ts))}{g_{\mathcal{C}}(\bar{F}_Z(t))} = s^{-\alpha\beta_{\mathcal{C}}} .$$

Next, Assumption 1 implies that for all $\mathbf{u} \in (0, \infty)^h$, it holds that

$$\lim_{t\to\infty} \frac{\mathbb{P}(\mathbf{u} \cdot \mathbf{Z}_{1,h} \in tA)}{g_{\mathcal{C}}(\bar{F}_Z(t))} = \nu_{\mathcal{C}}(\mathbf{u}^{-1} \cdot A) .$$

This convergence implies that there exists a function M_A such that for all $\mathbf{u} \in (0, \infty)^h$,

$$\sup_{t\geq 1} \frac{\mathbb{P}(\mathbf{u} \cdot \mathbf{Z}_{1,h} \in tA)}{g_{\mathcal{C}}(\bar{F}_Z(t))} \leq M_A(\mathbf{u}) . \tag{11}$$

Hence, if $\mathbb{E}[M_A(\boldsymbol{\sigma}(\mathbf{X}_{1,h}))] < \infty$, by bounded convergence, we have

$$\lim_{t\to\infty} \frac{\mathbb{P}(\boldsymbol{\sigma}(\mathbf{X}_{1,h}) \cdot \mathbf{Z}_{1,h} \in tA)}{g_{\mathcal{C}}(\bar{F}_Z(t))} = \mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)] .$$

For $h = 1$, and $A = (1, \infty)$, Potter's bound imply that (11) holds with $M_A(u) = Cu^{\alpha+\epsilon}$ for some constant C, i.e.

$$\sup_{t\geq 1} \frac{\mathbb{P}(uZ > t)}{\bar{F}_Z(t)} \leq Cu^{\alpha+\epsilon} . \tag{12}$$

For example, for $m > h$ and $h' \geq 0$, and for any Borel measurable set $B \subset \mathbb{R}^{h'+1}$, we have, by the same bounded convergence argument

$$\lim_{t\to\infty} \frac{\mathbb{P}(\mathbf{Y}_{1,h} \in tA \,, \mathbf{Y}_{m,m+h'} \in B)}{g_{\mathcal{C}}(\bar{F}_Z(t))} = \mathbb{E}\left[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})\right] .$$

If $\mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)] > 0$ (which in examples is seen to hold as soon as $\nu_{\mathcal{C}}(A) > 0$), we obtain that the extremogram defined in (9) can be expressed as

$$\begin{aligned} \rho(A, B, m) &= \lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathbf{Y}_{1,h} \in tA) \\ &= \frac{\mathbb{E}\left[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1}A)\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})\right]}{\mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)]} . \end{aligned} \tag{13}$$

We will consider the following type of cones. For $\mathbf{j} \in \{0, 1\}^h$, let $\mathcal{C}_{\mathbf{j}}$ denote the cone defined by

$$\mathcal{C}_{\mathbf{j}} = \{z \in [0, \infty]^h \mid \{\sum_{i:j_i=0} z_{j_i}\} \prod_{i,j_i=1} z_{j_i} > 0\} . \tag{14}$$

In words, a vector $z \in \mathcal{C}_{\mathbf{j}}$ if at least one of its entries corresponding to the components of $\mathbf{j}$ equal to zero is positive, and all of its entries corresponding to the components equal to one

hal-00624821, version 1 - 19 Sep 2011

of $\mathbf{j}$ are positive. For $h = 1$, the only cone is $(0, \infty]$ and we will denote it C_0 for consistency of the notation.

A subset A is relatively compact in $\mathcal{C}_\mathbf{j}$ if and only if there exists $\eta > 0$ such that $\sum_{i:j_i=0} z_{j_i} > \eta$ and $z_{j_i} > \eta$ for all i such that $j_i = 1$.

For example, if $h = 3$ and $j = (0, 0, 1)$, then $\mathcal{C}_\mathbf{j} = ([0, \infty] \times [0, \infty] \setminus \{(0, 0)\}) \times (0, \infty]$, and A is a relatively compact subset of $\mathcal{C}_{(0,0,1)}$ if there exists $\epsilon > 0$, such that $(z_1, z_2, z_3) \in A$ implies $z_1 > \epsilon$ or $z_2 \geq \epsilon$, and $z_3 \geq \epsilon$.

Denote $|\mathbf{j}| = j_1 + \cdots + j_h$, i.e. the number of non zero components in $\mathbf{j}$. Then, there exists a non zero Radon measure $\nu_\mathbf{j}$ on $\mathcal{C}_\mathbf{j}$ such that for each relatively compact set $A \in \mathcal{C}_\mathbf{j}$,

$$\lim_{t\to\infty} \frac{\mathbb{P}(\mathbf{Z}_{1,h} \in tA)}{\bar{F}_Z(t)^{|\mathbf{j}|+1}} = \nu_\mathbf{j}(A) \; .$$

The measure $\nu_\mathbf{j}$ can be described more precisely.

$$\nu_\mathbf{j}(\mathrm{d}\mathbf{z}) = \alpha^{|j|+1} \left\{ \sum_{i:j_i=0} z_{j_i}^{-\alpha-1} \delta_{j_i}(\mathrm{d}z_{j_i}) \right\} \prod_{i:j_i=1} z_{j_i}^{-\alpha-1} \mathrm{d}z_{j_i} \; ,$$

where δ_j is Lebesgue's's measure on the j-th coordinate axis, i.e. for any non negative measurable function ϕ,

$$\int_{[0,\infty]^h} \phi(z)\delta_j(\mathrm{d}z) = \int_0^\infty \phi(0, \ldots, z_j, \ldots, 0)\,\mathrm{d}z_j \; .$$

Moreover, for any relatively compact subset A of $\mathcal{C}_\mathbf{j}$, and for any $\epsilon > 0$, there exist $\eta > 0$ and a constant C (which both depend on A) such that, for all $\mathbf{u} \in (0, \infty)^h$,

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{u}\mathbf{Z}_{1,h} \in tA)}{\bar{F}_Z(u)^{|\mathbf{j}|+1}} &\leq \frac{\mathbb{P}\left(\cup_{i:j_i=0}\{u_{j_i} Z_{j_i} > \eta\} \cap \cap_{i:j_i=1}\{u_{j_i} Z_{j_i} > \eta\}\right)}{\bar{F}_Z(u)^{|\mathbf{j}|+1}} \\ &\leq C\eta^{-(|\mathbf{j}|+1)(\alpha+\epsilon)} \left\{ \sum_{i:j_i=0} (u_{j_i} \vee 1)^{\alpha+\epsilon} \right\} \prod_{i:j_i=1} (u_{j_i} \vee 1)^{\alpha+\epsilon} \; . \end{aligned} \tag{15}$$

Thus (11) holds and if

$$\mathbb{E}\left[\left\{ \sum_{i:j_i=0} \sigma^{\alpha+\epsilon}(X_{j_i}) \right\} \prod_{i:j_i=1} \sigma^{\alpha+\epsilon}(X_{j_i}) \right] < \infty \; , \tag{16}$$

then, cf. (13),

$$\lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathbf{Y}_{1,h} \in tA) = \frac{\mathbb{E}\left[\nu_\mathbf{j}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1}A)\,\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})\right]}{\mathbb{E}[\nu_\mathbf{j}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1}A)]} \; .$$

Remark 1. We assume that $h < m$. Otherwise, if $m < h$, then vectors $\mathbf{Y}_{m,m+h'}$ and $\mathbf{Y}_{1,h}$ may be asymptotically dependent. For example, if $\{Z_j\}$ is i.i.d with the tail distribution as in (4), then $\mathbb{P}(Z_2 + Z_3 > t \mid Z_1 + Z_2 > t) \to 1/2$. We do not think that this is of particular interest, since one is primary interested in estimating distribution of *future* vector $\mathbf{Y}_{m,m+h'}$ based on the *past* observations $\mathbf{Y}_{1,h}$.

hal-00624821, version 1 - 19 Sep 2011

Remark 2. The cones $\mathcal{C}_{\mathbf{j}}$ are the only ones such that $\mathbf{u}^{-1} \cdot A \subset \mathcal{C}$ for all $\mathbf{u} \in (0,\infty)^h$ and every $A \subset \mathcal{C}$. This assumption can be relaxed and other cones could be considered if σ is bounded above and away from zero, but this is not a desirable assumption since for instance it rules out the case $\sigma(x) = \mathrm{e}^x$.

Remark 3. Consider for example $\sigma(x) = \exp(x)$. Assumption (16) is fulfilled for arbitrary (weak and strong) dependence structure of $\{X_j\}$. The same holds for many moment assumptions which appear in the paper.

2.1 Examples

Example 1. Fix some positive integer h and consider the cone $\mathcal{C}_{\mathbf{1}} = (0,\infty)^h$. Then (10) holds with $g_h(t) = t^h$ and ν_h defined by

$$\nu_h(\mathrm{d}z_1, \dots, \mathrm{d}z_h) = \alpha^h \prod_{i=1}^{h} z_i^{-\alpha-1} \mathrm{d}z_i \; .$$

Consider the set A defined by $A = \{(z_1, \dots, z_h) \in \mathbb{R}_+^h \mid z_1 > 1, \dots, z_h > 1\}$. If

$$\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{\alpha+\epsilon}(X_i)\right] < \infty$$

for some $\epsilon > 0$, we obtain, for $m > h$, and $B \in \mathbb{R}^{h'+1}$,

$$\lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid Y_1 > t, \dots, Y_h > t) = \frac{\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{\alpha}(X_i) \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})\right]}{\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{\alpha}(X_i)\right]} \; .$$

In particular, setting $B = (-\infty, y]$ and $h' = 0$, the limiting conditional distribution of Y_m given that $Y_1, \dots, Y_h$ are simultaneously large is given by

$$\Psi_h(y) = \lim_{t\to\infty} \mathbb{P}(Y_m \le y \mid Y_1 > t, \dots, Y_h > t) = \frac{\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{\alpha}(X_i) F_Z(y/\sigma(X_m))\right]}{\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{\alpha}(X_i)\right]} \; . \tag{17}$$

Example 2. Consider again the case $\mathcal{C}_1 = (0,\infty)$. Another quantity of interest is the limiting distribution of the sum of h' consecutive values, given that past values are extreme. To keep notation simple, consider $h' = 1$ and, for $m > 1$,

$$\Psi^*(y) = \lim_{t\to\infty} \mathbb{P}(Y_m + Y_{m+1} \le y \mid Y_1 > t) = \frac{\mathbb{E}[\sigma^{\alpha}(X_1) \mathbb{P}(Y_m + Y_{m+1} \le y \mid \mathcal{X})]}{\mathbb{E}[\sigma^{\alpha}(X_1)]} \; .$$

Estimating this distribution yields for instance empirical quantiles of the sum of future returns, given the present one is large.

Example 3. Consider the cone $\mathcal{C}_{0,0} = [0,\infty) \times [0,\infty) \setminus \{\mathbf{0}\}$. Then (10) holds with $g_{0,0}(t) = t$ and $\nu_{0,0}$ defined by

$$\nu_{0,0}(\mathrm{d}z_1, \mathrm{d}z_2) = \alpha\{\delta_{(0,\infty]\times\{0\}} z_1^{-\alpha-1} \mathrm{d}z_1 + \delta_{\{0\}\times(0,\infty]} z_2^{-\alpha-1} \mathrm{d}z_2\} \; .$$

hal-00624821, version 1 - 19 Sep 2011

The bound (11) with $M_A(u,v) = C(u^{\alpha+\epsilon} + v^{\alpha+\epsilon})$ for some constant C. Consider the set A defined by $A = \{(z_1,z_2) \in \mathbb{R}_+^2 \mid z_1 + z_2 > 1\}$. If $\mathbb{E}[\sigma^{\alpha+\epsilon}(X_1)] < \infty$ for some $\epsilon > 0$, we obtain

$$\lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid Y_1 + Y_2 > t) = \frac{\mathbb{E}\left[\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})(\sigma^\alpha(X_1) + \sigma^\alpha(X_2))\right]}{\mathbb{E}[\sigma^\alpha(X_1)] + \mathbb{E}[\sigma^\alpha(X_2)]} .$$

In particular, take $B = (-\infty, y]$ and $h' = 0$. The limiting conditional distribution of Y_m given $Y_1 + Y_2$ is large is defined by

$$\Lambda(y) = \lim_{t\to\infty} \mathbb{P}(Y_m \le y \mid Y_1 + Y_2 > t) = \frac{\mathbb{E}[\{\sigma^\alpha(X_1) + \sigma^\alpha(X_2)\} F_Z(y/\sigma(X_m))]}{\mathbb{E}[\sigma^\alpha(X_1) + \sigma^\alpha(X_2)]} .$$

Example 4. We can combine the previous examples. Consider $A = \{(z_1,z_2,z_3) \in \mathbb{R}_+^3 | z_1 + z_2 > 1, z_3 > 1\}$. We may obtain for instance, for $m > 3$,

$$\begin{aligned} &\lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid Y_1 + Y_2 > t, Y_3 > t) \\ &\qquad = \frac{\mathbb{E}\left[\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B \mid \mathcal{X})\{\sigma^\alpha(X_1) + \sigma^\alpha(X_2)\}\sigma^\alpha(X_3)\right]}{\mathbb{E}[\{\sigma^\alpha(x_1) + \sigma^\alpha(X_2)\}\sigma^\alpha(X_3)]} , \end{aligned}$$

if $\mathbb{E}[\{\sigma^{\alpha+\epsilon}(X_1) + \sigma^{\alpha+\epsilon}(X_2)\}\sigma^{\alpha+\epsilon}(X_3)] < \infty$ for some $\epsilon > 0$. The relevant cone is $\mathcal{C}_{0,0,1}$, $g_{0,0,1}(t) = t^2$ and the associated measure on $\mathcal{C}_{0,0,1}$ is defined by

$$\nu_{0,0,1} = \alpha^2\{\delta_{(0,\infty]\times\{0\}} z_1^{-\alpha-1}\mathrm{d}z_1 + \delta_{\{0\}\times(0,\infty]} z_2^{-\alpha-1}\mathrm{d}z_2\} z_3^{-\alpha-1}\mathrm{d}z_3 .$$

3 Estimation

To simplify the notation, assume that we observe $Y_1, \dots, Y_{n+m+h'}$. An estimator $\hat\rho_n(A,B,m)$ is naturally defined by

$$\hat\rho_n(A,B,m) = \frac{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{j,j+h-1} \in Y_{(n:n-k)}A\}} \mathbf{1}_{\{\mathbf{Y}_{j+m,j+m+h'} \in B\}}}{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{j,j+h-1} \in Y_{(n:n-k)}A\}}} ,$$

where k is a user chosen threshold and $Y_{(n:1)} \le \dots \le Y_{(n:n)}$ are the increasing order statistics of the observations $Y_1, \dots, Y_n$. We will also consider the case $B = (-\boldsymbol{\infty}, \mathbf{y}]$, i.e. the case of the limiting conditional distribution of $\mathbf{Y}_{m,m+h'}$ given $\mathbf{Y}_{1,h} \in tA$, i.e.

$$\begin{aligned} \Psi_{A,m,h'}(\mathbf{y}) &= \lim_{t\to\infty} \mathbb{P}(\mathbf{Y}_{m,m+h'} \le \mathbf{y} \mid \mathbf{Y}_{1,h} \in tA) \\ &= \rho(A, (\boldsymbol{\infty}, \mathbf{y}], m) = \frac{\mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A) \prod_{i=1}^{h'} F(y_i/\sigma(X_{m+i}))]}{\mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)]} . \end{aligned} \tag{18}$$

An estimator $\hat\Psi_{n,A,m,h'}$ of $\Psi_{A,m,h'}$ is defined on $\mathbb{R}^{h'+1}$ by

$$\hat\Psi_{n,A,m,h'}(\mathbf{y}) = \frac{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{j,j+h-1} \in Y_{(n:n-k)}A\}} \mathbf{1}_{\{\mathbf{Y}_{j+m,j+m+h'} \le \mathbf{y}\}}}{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{j,j+h-1} \in Y_{(n:n-k)}A\}}} . \tag{19}$$

In order to obtain statistical results, we need additional assumptions. We first state two assumptions which will be needed to prove the weak convergence of a multivariate conditional empirical process.

hal-00624821, version 1 - 19 Sep 2011

Assumption 2. *For $j = 1, \dots, h$, there exist functions $\mathcal{L}_j$ such that for all $s, s' \geq 1$, $\mathbf{u}, \mathbf{v} \in (0, \infty)^h$,*

$$\lim_{t\to\infty} \frac{\mathbb{P}(\mathbf{u} \cdot \mathbf{Z}_{1,h} \in tsA, \mathbf{v} \cdot \mathbf{Z}_{j,j+h-1} \in ts'A)}{g_{\mathcal{C}}(\bar{F}_Z(t))} = \mathcal{L}_j(A, \mathbf{u}, \mathbf{v}, s, s') \ . \tag{20}$$

For $j = 1$ we only need that (20) holds with $\mathbf{u} = \mathbf{v}$. If A is a cone, then (20) holds for $j = 1$ with $\mathcal{L}_1(A, \mathbf{u}, \mathbf{u}, s, s') = T_{\mathcal{C}}(s \vee s')\nu_{\mathcal{C}}(\mathbf{u}^{-1} \cdot A)$ as an immediate consequence of Assumption 1. It may happen that $\mathcal{L}_j(A, \cdot) \equiv 0$ for $j = 2, \dots, h$. Intuitively, this happens if $\mathbf{u} \cdot \mathbf{Z}_{1,h}$ and $\mathbf{v} \cdot \mathbf{Z}_{j,j+h-1}$ belong simultaneously to tA implies that at least $h + 1$ coordinates of $\mathbf{Z}_{1,h+j-1}$ are large. This is the case for instance for Examples 1 and 4. Actually, Assumption 2 holds for the cones $\mathcal{C}_{\mathbf{j}}$, but a precise description of the functions $\mathcal{L}_j$ when they are not identically zero would be extremely involved. This will only be done for Example 3. See Section 3.3.

By Cauchy-Schwartz inequality, if Assumptions 1 and 2 hold, then, for $s, s' \geq 1$,

$$\mathcal{L}_j(\mathbf{u}, \mathbf{v}, sA, s'A) \leq \sqrt{M_A(\mathbf{u}) M_A(\mathbf{v})} \ .$$

Thus, if $\mathbb{E}[M_A(\boldsymbol{\sigma}(\mathbf{X}_{1,h}))] < \infty$, then the convergence in (20) is also in $L^1(\boldsymbol{\sigma}(\mathbf{X}_{1,h}), \boldsymbol{\sigma}(\mathbf{X}_{j,j+h-1}))$.

The next assumption is needed for the quantities (that will appear in the limiting distributions) to be well defined and to use bounded convergence arguments.

Assumption 3. $\mathbb{E}[M_A^2(\boldsymbol{\sigma}(\mathbf{X}_{1,h}))] < \infty$.

As usual, the bias of the estimators will be bounded by a second order type condition. Let k be a non decreasing sequence of integers, let F_Y denote the distribution of Y and let $u_n = (1/\bar{F}_Y)^{\leftarrow}(n/k)$. Consider the measure defined on the Borel subsets of $\mathcal{C}$ by

$$\mu_{\mathcal{C}}(A) = \frac{\mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)]}{(\mathbb{E}[\sigma^{\alpha}(X)])^{\beta_{\mathcal{C}}}} \ . \tag{21}$$

We introduce a rate of convergence:

$$v_n(A) = \mathbb{E}\left[\sup_{s\geq 1} \left| \frac{\mathbb{P}(\mathbf{Y}_{1,h} \in u_n sA \mid \mathcal{X})}{g_{\mathcal{C}}(k/n)} - T_{\mathcal{C}}(s)\mu_{\mathcal{C}}(A) \right| \right] \ . \tag{22}$$

Lemma 1. *Under Assumption 1 and 3,* $\lim_{n\to\infty} v_n(A) = 0$.

We need also the following quantities, which are well defined under Assumptions 1, 2 and 3. For $j = 2, \dots, h$ and measurable subsets B, B' of $\mathbb{R}^{h'+1}$, define

$$\begin{aligned} &\mathcal{R}_j(A, B, B') \\ &= \frac{\mathbb{E}\big[\mathcal{L}(A, \boldsymbol{\sigma}(\mathbf{X}_{1,h}), \boldsymbol{\sigma}(\mathbf{X}_{j,j+h-1}), 0, 0) \times \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B, \mathbf{Y}_{m+j-1,m+h'+j-1} \in B' \mid \mathcal{X})\big]}{(\mathbb{E}[\sigma^{\alpha}(X)])^{-\beta_{\mathcal{C}}}\, \mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)]} \\ &+ \frac{\mathbb{E}\big[\mathcal{L}(A, \boldsymbol{\sigma}(\mathbf{X}_{1,h}), \boldsymbol{\sigma}(\mathbf{X}_{j,j+h-1}), 0, 0) \times \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B', \mathbf{Y}_{m+j-1,m+h'+j-1} \in B \mid \mathcal{X})\big]}{(\mathbb{E}[\sigma^{\alpha}(X)])^{-\beta_{\mathcal{C}}}\, \mathbb{E}[\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)]} \ . \end{aligned} \tag{23}$$

For brevity, denote $\mathcal{R}_j(A, B) = \mathcal{R}_j(A, B, B)$.

hal-00624821, version 1 - 19 Sep 2011

3.1 General result: weak dependence

We can now state our main result in the weak dependence setting, i.e. when absolute summability (2) of the autocovariance function of the process $\{X_j\}$ holds.

In order to simplify the proof, we make an additional assumption.

Assumption 4. *If $s < t$ then $tA \subset sA$.*

This assumptions holds for all the examples considered here and most common examples.

Theorem 2. *Let Assumptions 1, 2, 3, 4 and the weak dependence condition (2) hold. Assume moreover that $\mu_{\mathcal{C}}(A) > 0$, $k/n \to 0$, $ng_{\mathcal{C}}(k/n) \to \infty$ and*

$$\lim_{n\to\infty} ng_{\mathcal{C}}(k/n)\, v_n(A) = 0 \; . \tag{24}$$

Then

$$\sqrt{ng_{\mathcal{C}}(k/n)\mu_{\mathcal{C}}(A)}\{\hat{\rho}_n(A,B,m) - \rho(A,B,m)\}$$

converges weakly to a centered Gaussian distribution with variance

$$\begin{aligned} &\rho(A,B,m)\{1-\rho(A,B,m)\} \\ &\quad + \sum_{j=2}^{h\wedge(m-h)} \left\{\mathcal{R}_j(A,B) - 2\rho(A,B,m)\mathcal{R}_j(A,B,\mathbb{R}^{h'+1}) + \rho^2(A,B,m)\mathcal{R}_j(A,\mathbb{R}^{h'+1})\right\} \; . \end{aligned} \tag{25}$$

Remark 4. If $h = 1$ or if the functions $\mathcal{L}_j$ defined in Assumption 2 are identically zero for $j \geq 2$, then the limiting covariance in (25) is simply $\rho(A,B,m)\{1-\rho(A,B,m)\}$.

Otherwise, the additional terms can be canceled by modifying the estimator of $\hat{\rho}_n(A,B,m)$. Assuming we have $nh + m + h' + 1$ observations, we can define

$$\tilde{\rho}_n(A,B,m) = \frac{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{(j-1)h+1,jh}\in Y_{(n:n-k)}A\}}\mathbf{1}_{\{\mathbf{Y}_{(j-1)h+m,(j-1)h+m+h'}\in B\}}}{\sum_{j=1}^n \mathbf{1}_{\{\mathbf{Y}_{(j-1)h+1,jh}\in Y_{(n:n-k)}A\}}}$$

Noting that the events $\{\mathbf{Y}_{j,j+h-1} \in A\}$ are h-dependent conditionally on $\mathcal{X}$, the proof of Theorem 2 can be easily adapted to show that the limiting variance of $\sqrt{ng_{\mathcal{C}}(k/n)}\{\tilde{\rho}_n(A,B,m) - \rho(A,B,m)\}$ is the same as in the case where $\mathcal{L}_j \equiv 0$ for $j = 2,\dots,h$. But this is of course at the cost of an increase of the asymptotic variance, due to a different sample size.

We can also obtain the functional convergence of the estimator $\hat{\Psi}_{n,A,m,h'}$ of the limiting conditional distribution function $\Psi_{A,m,h'}$, defined respectively in (19) and (18).

Corollary 3. *Under the Assumptions of Theorem 2, and if moreover the distribution $\Psi_{A,m,h'}$ is continuous, then*

$$\sqrt{ng_{\mathcal{C}}(k/n)\mu_{\mathcal{C}}(A)}\{\hat{\Psi}_{n,A,m,h'} - \Psi_{A,m,h'}\}$$

converges in $\mathcal{D}(\mathbb{R}^{h'+1})$ to a Gaussian process. If $h = 1$ or if the functions $\mathcal{L}_j$ are identically zero for $j = 2,\dots,h$, then the limiting process can be expressed as $\mathbb{B}\circ\Psi_{A,m,h'}$, where $\mathbb{B}$ is the standard Brownian bridge.

Note that a sufficient condition for $\Psi_{A,m,h'}$ to be continuous is that F_Z is continuous.

hal-00624821, version 1 - 19 Sep 2011

3.2 General result: long memory

We now state our results in the framework of long memory. This requires several additional notions, such as multivariate Hermite expansion and Hermite ranks which are recalled in Appendix B.

Define the functions G_n and G for $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^h \times \mathbb{R}^{h'+1}$ and $s \geq 1$ by

$$G_n(A, B, s, \mathbf{x}, \mathbf{x}') = \frac{\mathbb{P}(\boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{Z}_{1,h} \in u_n s A)}{g(k/n)} \, \mathbb{P}(\boldsymbol{\sigma}(\mathbf{x}') \cdot \mathbf{Z}_{m,m+h'} \in B) \tag{26}$$

$$\begin{aligned} G(A, B, \mathbf{x}, \mathbf{x}') &= \lim_{n\to\infty} G_n(A, B, 1, \mathbf{x}, \mathbf{x}') \\ &= \frac{(\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{x})^{-1} \cdot A)}{\mathbb{E}[\sigma^\alpha(X_1)])^{\beta_{\mathcal{C}}}} \, \mathbb{P}(\boldsymbol{\sigma}(\mathbf{x}') \cdot \mathbf{Z}_{m,m+h'} \in B) \, . \end{aligned} \tag{27}$$

Let $\tau_n(A, B, s)$ and $\tau(A, B)$ be the Hermite ranks with respect to $(\mathbf{X}_{1,h}, \mathbf{X}_{m,m+h'})$ of the functions $G_n(A, B, s, \cdot, \cdot)$ and $G(A, B, \cdot, \cdot)$, respectively. Define $\tau(A) = \tau(A, \mathbb{R}^d)$.

Assumption 5. *For large n, $\inf_s \tau_n(A, B, s) = \tau(A, B)$ and $\tau(A, B) \leq \tau(A)$.*

This assumption is fulfilled for example when $\sigma(x) = \exp(x)$, in which case all the considered Hermite ranks are equal to one, or if σ is an even function with Hermite rank 2 (such as $\sigma(x) = x^2$), in which case they are equal to two. The modification of Theorem 2 reads as follows.

Theorem 4. *Assume that $\{X_j\}$ is the long memory Gaussian sequence with covariance given by (3). Let Assumptions 1, 2, 3, 4 and 5 hold, $\mu_{\mathcal{C}}(A) > 0$ and $k/n \to 0$, $ng_{\mathcal{C}}(k/n) \to \infty$ and*

$$\lim_{n\to\infty} \left\{ ng_{\mathcal{C}}(k/n) \wedge \gamma_n^{-\tau(A,B)/2} \right\} v_n(A) = 0 \, . \tag{28}$$

(i) If $ng_{\mathcal{C}}(k/n)\gamma_n^{\tau(A,B)} \to 0$, then

$$\sqrt{ng_{\mathcal{C}}(k/n)\mu_{\mathcal{C}}(A)} \{\hat{\rho}_n(A, B, m) - \rho(A, B, m)\}$$

converges to a centered Gaussian distribution with variance given in (25)

(ii) If $ng_{\mathcal{C}}(k/n)\gamma_n^{\tau(A,B)} \to \infty$, then $\gamma_n^{-\tau(A,B)/2} \{\hat{\rho}_n(A, B, m) - \rho(A, B, m)\}$ converges weakly to a distribution which is non-Gaussian except if $\tau(A, B) = 1$.

The exact definition of the limiting distribution will be given in Section 4. It suffices to mention here that this distribution depends on H and $\tau(A, B)$. The meaning of the above result is the following. In the long memory setting, it is still possible to obtain the same limit as in the weakly dependent case, if k (i.e., the number of high order statistics used in the definition of the estimators) is not too large, so that both the bias and the long memory effect are canceled.

Define a new Hermite rank $\tau^*(A) = \inf_{y\in\mathbb{R}^{h'+1}} \tau(A, (\boldsymbol{\infty}, \mathbf{y}])$.

hal-00624821, version 1 - 19 Sep 2011

Corollary 5. *Under the Assumptions of Theorem 4, if the distribution function* $\Psi_{A,m,h'}$ *is continuous and if* $\tau^*(A) \leq \tau(A)$*, then*

- *If* $ng_{\mathcal{C}}(k/n)\gamma_n^{\tau^*(A)} \to 0$*, then*
$$\sqrt{ng_{\mathcal{C}}(k/n)\mu_{\mathcal{C}}(A)}\{\hat{\Psi}_{n,A,m,h'} - \Psi_{A,m,h'}\}$$
converges in $\mathcal{D}((-\infty,+\infty)^{h'+1}$ *to a Gaussian process. If* $h = 1$ *or if the functions* $\mathcal{L}_j$ *are identically zero for* $j = 2, \ldots, h$*, then the limiting process can be expressed as* $\mathbb{B} \circ \Psi_{A,m,h'}$*, where* $\mathbb{B}$ *is the standard Brownian bridge.*
- *If* $ng_{\mathcal{C}}(k/n)\gamma_n^{\tau^*(A)} \to \infty$*, then* $\gamma_n^{-\tau^*(A)/2}\{\hat{\Psi}_{n,A,m,h'} - \Psi_{A,m,h'}\}$ *converges in* $\mathcal{D}((-\infty,+\infty)^{h'+1}$ *to a process which can be expressed as* $J_{A,m,h'} \cdot \aleph$ *where* $J_{A,m,h'}$ *is a deterministic function and* $\aleph$ *is a random variable, which is non Gaussian except if* $\tau^*(A) = 1$*.*

The exact definition of the function $J_{A,m,h'}$ and of the random variable $\aleph$ will be given in Section 4. Anyhow, they are not of much practical interest. In practice, the main goal will be to choose the number k of order statistics used in the estimation procedure so that both the bias and the long memory effect are canceled, and the limiting distribution of the weakly dependent case can be used in the inference.

3.3 Examples

We now discuss the Examples introduced in Section 2.1. In order to evaluate the rate of convergence (22), it is necessary to introduce a second order regular variation condition. We follow here Drees [1998].

Assumption 6. *There exists a bounded non increasing function* η^* *on* $[0,\infty)$*, regularly varying at infinity with index* $-\alpha\zeta$ *for some* $\zeta \geq 0$*, and such that* $\lim_{t\to\infty} \eta^*(t) = 0$ *and there exists a measurable function* η *such that for* $z > 0$*,*

$$\mathbb{P}(Z > z) = cz^{-\alpha} \exp\left(\int_1^z \frac{\eta(s)}{s} \,\mathrm{d}s\right) ,$$
$$\exists C > 0 , \quad \forall s \geq 0 , \quad |\eta(s)| \leq C\eta^*(s) .$$

On account of Breiman's lemma, if the tail of Z is regularly varying with index $-\alpha$, then the same holds for $Y = \sigma(X)Z$, as long as X and Z are independent, and $\mathbb{E}[\sigma^\alpha(X)] < \infty$. Also, (SO) property is transferred from the tail of Z to Y; See [Kulik and Soulier, 2011, Proposition 2.1].

For the sake of simplicity and clarity of exposition, we will make in this section the usual assumption that $\sigma(x) = \exp(x)$, so that the Hermite rank of σ is 1. This will avoid to define many auxiliary functions and Hermite ranks. But the examples can of course be treated in a more general framework. Also, we will only state the convergence results under the conditions which imply that the limiting distribution is the same as in the weak dependence case, since this is the case of practical interest. We only treat Examples 1 and 3 since they exhibit the two different possibility for the limiting distributions. The computations for the other examples are straightforward.

hal-00624821, version 1 - 19 Sep 2011

3.3.1 Example 1 continued

Fix integers $h \geq 1$ and $m > h$. Recall the formula (17) for the conditional distribution of Y_m given that $Y_1, \dots, Y_h$ are simultaneously large. Its estimator $\hat{\Psi}_{n,h}$ is defined by

$$\hat{\Psi}_{n,h}(y) = \frac{\sum_{j=1}^{n} \mathbf{1}_{\{Y_j > Y_{(n:n-k)}, \dots, Y_{j+h-1} > Y_{(n:n-k)}, Y_{j+m} \leq y\}}}{\sum_{j=1}^{n} \mathbf{1}_{\{Y_j > Y_{(n:n-k)}, \dots, Y_{j+h-1} > Y_{(n:n-k)}\}}}$$

with a user chosen k.

Assumption 2 holds with $\mathcal{L}_j(A, \cdot) \equiv 0$, $j = 2, \dots, h$. Assumption 6 and [Kulik and Soulier, 2011, Proposition 2.8] imply that if moreover

$$\mathbb{E}\left[\prod_{i=1}^{h} \sigma^{2\alpha(\zeta+1)+\epsilon}(X_i)\right] < \infty \tag{29}$$

for some $\zeta, \epsilon > 0$, a bound for $v_n(A)$ is then given by

$$v_n(A) = O(\eta^*(u_n)) \,. \tag{30}$$

The moment restriction (29) is quite weak. In particular, it is fulfilled for $\sigma(x) = \exp(x)$; see Remark 3. Recall that in this example Assumption 1 and 2 hold and the functions $\mathcal{L}_j$ therein are vanishing for $j \geq 2$. Also, Assumption 3 is implied by (29).

Corollary 6. *Assume that $\sigma(x) = \exp(x)$. Let Assumption 6 and (29) hold. Let k be such that $k/n \to 0$, $n(k/n)^h \to \infty$, and*

$$\lim_{n\to\infty} (n(k/n)^h)^{1/2} \eta^*(u_n) = 0 \,. \tag{31}$$

In the weakly dependent case (2) or in the long memory case (3) if moreover $n(k/n)^h \gamma_n \to 0$, then

$$\sqrt{n(k/n)^h}(\hat{\Psi}_{n,h} - \Psi_h) \Rightarrow \left(\frac{\mathbb{E}[\sigma^\alpha(X_1) \cdots \sigma^\alpha(X_h)]}{\mathbb{E}^h[\sigma^\alpha(X_1)]}\right)^{-1/2} \mathbb{B} \circ \Psi_h$$

weakly in $\mathcal{D}((-\infty, \infty))$, where $\mathbb{B}$ is the standard Brownian bridge.

3.3.2 Example 3 continued

Consider the estimation of

$$\Lambda(y) = \lim_{t\to\infty} \mathbb{P}(Y_m \leq y \mid Y_1 + Y_2 > t) = \frac{\mathbb{E}[\{\sigma^\alpha(X_1) + \sigma^\alpha(X_2)\} F_Z(y/\sigma(X_m))]}{\mathbb{E}[\sigma^\alpha(X_1) + \sigma^\alpha(X_2)]} \,.$$

An estimator if defined by

$$\hat{\Lambda}_n(y) = \frac{\sum_{j=1}^{n} \mathbf{1}_{\{Y_j + Y_{j+1} > Y_{(n:n-k)}\}} \mathbf{1}_{\{Y_{j+m} \leq y\}}}{\sum_{j=1}^{n} \mathbf{1}_{\{Y_j + Y_{j+1} > Y_{(n:n-k)}\}}} \,.$$

hal-00624821, version 1 - 19 Sep 2011

We have already shown that Assumption 1 holds and Assumption 4 holds trivially. Assumption 2 holds with the function $\mathcal{L}_2$ defined by

$$\mathcal{L}_2(A, u_1, u_2, v_1, v_2, s, s') = \left(\frac{1+s}{u_2} \vee \frac{1+s'}{v_1} \right)^{-\alpha} . \tag{32}$$

If $\mathbb{E}[\sigma^{2\alpha(\zeta+1)+\epsilon}(X_1)] < \infty$, then Assumption 3 holds and applying Lemma A.1, we obtain a bound for $v_n(A)$:

$$v_n(A) = O\left(\eta^*(u_n) + u_n^{-1} \int_0^{u_n} \bar{F}_Z(s) \, \mathrm{d}s \right) . \tag{33}$$

as soon as

$$\tag{34}$$

Corollary 7. *Let Assumption 6 and (29) hold. Let k be such that $k \to \infty$, $k/n \to 0$ and*

$$\lim_{n\to\infty} k^{1/2} \left(\eta^*(u_n) + u_n^{-1} \int_0^{u_n} \bar{F}_Z(s) \, \mathrm{d}s \right) = 0 .$$

In the weakly dependent case (2) or in the long memory case (3) if moreover $k\gamma_n \to 0$, then

$$k^{1/2}(\hat{\Lambda}_n - \Lambda) \Rightarrow \left(\frac{\mathbb{E}[\sigma^\alpha(X_1) + \sigma^\alpha(X_2)]}{\mathbb{E}[\sigma^\alpha(X_1)]} \right)^{-1/2} \mathbb{W}$$

weakly in $\mathcal{D}((-\infty, \infty))$, where $\mathbb{W}$ is a Gaussian process with covariance

$$\begin{aligned} \operatorname{cov}(\mathbb{W}(y), \mathbb{W}(y')) &= \Lambda(y \wedge y') - 2\Lambda(y)\Lambda(y') \\ &+ \frac{\mathbb{E}[\sigma^\alpha(X_2)\{F_Z(y/\sigma(X_m))F_Z(y'/\sigma(X_{m+1})) + F_Z(y/\sigma(X_m))F_Z(y'/\sigma(X_{m+1}))\}]}{\mathbb{E}[\sigma^\alpha(X_1) + \sigma^\alpha(X_2)]} . \end{aligned}$$

Remark 5. If the estimator if modified by taking only every other observation, then $\sqrt{k}(\hat{\Lambda}_n - \Lambda)$ converges weakly to $2\mathbb{B} \circ \Lambda$ where $\mathbb{B}$ is the standard Brownian bridge.

4 Proofs

For clarity of notation, denote $\sigma_i = \sigma(X_i)$, $g = g_{\mathcal{C}}$, $T = T_{\mathcal{C}}$ and $\beta = \beta_{\mathcal{C}}$. Recall that F_Y denotes the distribution function of Y and $u_n = (1/\bar{F}_Y)^{\leftarrow}(n/k)$. By (4) and the regular variation of g, it holds that $\bar{F}_Y(u_n) \sim \mathbb{E}[\sigma_0^\alpha]\bar{F}_Z(u_n)$ and

$$\lim_{n\to\infty} \frac{g(k/n)}{g(\bar{F}_Z(u_n))} = (\mathbb{E}[\sigma_0^\alpha])^\beta .$$

Whenever there is no risk of confusion, we omit dependence on h, m, h' and A in the notation. For $j = 1, \ldots, n$, define the following random variables

$$W_{j,n}(s) = \mathbf{1}_{\{\mathbf{Y}_{j,j+h-1} \in u_n s A\}} , s \geq 1 , \quad V_j(B) = \mathbf{1}_{\{\mathbf{Y}_{j+m,j+m+h'} \in B\}} . \tag{35}$$

hal-00624821, version 1 - 19 Sep 2011

Assumption 1 together with the choice of u_n implies that (recall the definitions (13) and (21) of $\rho(A,B,m)$ and $\mu_{\mathcal{C}}(A)$),

$$\lim_{n\to\infty} \frac{\mathbb{E}[W_{j,n}(s)]}{g(k/n)} = T(s)\mu_{\mathcal{C}}(A) \;, \tag{36}$$

$$\lim_{n\to\infty} \frac{\mathbb{E}[W_{j,n}(s)V_j(B)]}{g(k/n)} = T(s)\mu_{\mathcal{C}}(A)\rho(A,B,m) \;. \tag{37}$$

Define, for $s \geq 1$ and $\mathbf{x} \in \mathbb{R}^h$ and $\mathbf{x}' \in \mathbb{R}^{h'+1}$, the functions L_n and G_n by

$$L_n(s,\mathbf{x}) = \frac{\mathbb{P}(\boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{Z}_{1,h} \in u_n s A)}{g(k/n)} \;, \tag{38}$$

$$G_n(s,\mathbf{x},\mathbf{x}',B) = L_n(s,\mathbf{x}) \; \mathbb{P}(\boldsymbol{\sigma}(\mathbf{x}') \cdot \mathbf{Z}_{m,m+h'} \in B) \;. \tag{39}$$

With these notations, we have,

$$L_n(s,\mathbf{X}_{j,j+h-1}) = \frac{\mathbb{E}[W_{j,n}(s) \mid \mathcal{X}]}{g(k/n)} \;,$$

$$G_n(s,\mathbf{X}_{j,j+h-1},\mathbf{X}_{j+m,j+m+h'},B) = \frac{\mathbb{E}[W_{j,n}(s)V_j(B) \mid \mathcal{X}]}{g(k/n)} \;.$$

For $\mathbf{x} \in \mathbb{R}^h$, denote

$$L(\mathbf{x}) = \frac{\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{x})^{-1} \cdot A)}{(\mathbb{E}[\sigma^\alpha(X)])^\beta} \;, \tag{40}$$

so that $\mathbb{E}[L(\mathbf{X}_{1,h})] = \mu_{\mathcal{C}}(A)$.

Proof of Lemma 1. Write

$$\begin{aligned} L_n(s,\mathbf{x}) &- T_{\mathcal{C}}(s)L(\mathbf{x}) \\ &= \left\{ \frac{g(\bar{F}_Z(u_n s))}{g(k/n)} - (\mathbb{E}[\sigma^\alpha(X)])^{-\beta}\, T_{\mathcal{C}}(s) \right\} \frac{\mathbb{P}(\sigma(\mathbf{x}) \cdot \mathbf{Z}_{1,h} \in u_n s A)}{g(\bar{F}_Z(u_n s))} \\ &\quad + (\mathbb{E}[\sigma^\alpha(X)])^{-\beta} T_{\mathcal{C}}(s) \left\{ \frac{\mathbb{P}(\sigma(\mathbf{x}) \cdot \mathbf{Z}_{1,h} \in u_n s A)}{g(\bar{F}_Z(u_n s))} - (\mathbb{E}[\sigma^\alpha(X)])^{\beta} L(\mathbf{x}) \right\} \;. \end{aligned}$$

Thus, recalling the definition of v_n from (22), we have

$$\begin{aligned} v_n(A) \leq \sup_{s\geq 1} &\left| \frac{g(\bar{F}_Z(u_n s))}{g(k/n)} - (\mathbb{E}[\sigma^\alpha(X)])^{-\beta}\, T_{\mathcal{C}}(s) \right| \mathbb{E}[M_A(\boldsymbol{\sigma}(\mathbf{X}_{1,h}))] \\ &+ (\mathbb{E}[\sigma^\alpha(X)])^{-\beta} \mathbb{E}\left[\sup_{s\geq 0} \left| \frac{\mathbb{P}(\sigma(\mathbf{X}_{1,h}) \cdot \mathbf{Z}_{1,h} \in u_n s A \mid \mathcal{X})}{g(\bar{F}_Z(u_n s))} - (\mathbb{E}[\sigma^\alpha(X)])^{\beta} L(\mathbf{X}_{1,h}) \right| \right] \end{aligned}$$

By Assumption 1, for all $\mathbf{x} \in \mathbb{R}^h$,

$$\lim_{n\to\infty} \sup_{s\geq 1} \left| \frac{\mathbb{P}(\sigma(\mathbf{x}) \cdot \mathbf{Z}_{1,h} \in u_n s A)}{g(\bar{F}_Z(u_n s))} - (\mathbb{E}[\sigma^\alpha(X)])^{\beta} L(\mathbf{x}) \right| = 0 \;.$$

hal-00624821, version 1 - 19 Sep 2011

Moreover, by (11),

$$\sup_{s\geq 1}\left|\frac{\mathbb{P}(\sigma(\mathbf{x})\cdot\mathbf{Z}_{1,h}\in u_n sA)}{g(\bar F_Z(u_n s))}-(\mathbb{E}[\sigma^\alpha(X)])^\beta L(\mathbf{x})\right|\leq 2M_A(\boldsymbol{\sigma}(\mathbf{x}))\ .$$

Thus, by Assumption 3 and bounded convergence,

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{s\geq 0}\left|\frac{\mathbb{P}(\sigma(\mathbf{X}_{1,h})\cdot\mathbf{Z}_{1,h}\in u_n sA\mid\mathcal{X})}{g(\bar F_Z(u_n s))}-(\mathbb{E}[\sigma^\alpha(X)])^\beta L(\mathbf{X}_{1,h})\right|^2\right]=0\ .$$

Since $g\circ\bar F$ is regularly varying at infinity with negative index, by [Bingham et al., 1989, Theorem 1.5.2], the convergence of $g(\bar F_Z(u_n s))/g(k/n)$ to $(\mathbb{E}[\sigma^\alpha(X)])^{-\beta}T_{\mathcal{C}}(s)$ is uniform on $[1,\infty)$. Thus we have proved that $v_n(A)\to 0$. □

Proof of Theorem 2. Define

$$K(B,s)=T_{\mathcal{C}}(s)\mu_{\mathcal{C}}(A)\rho(A,B,m)\ ,\quad \tilde K_n(B,s)=\frac{1}{ng(k/n)}\sum_{j=1}^n W_{j,n}(s)V_j(B)\ ,$$

$$\tilde e_n(s)=\tilde K_n(\mathbb{R}^{h'+1},s)=\frac{1}{ng(k/n)}\sum_{j=1}^n W_{j,n}(s)\ ,\quad \xi_n=\frac{Y_{(n:n-k)}}{u_n}\ .$$

With this notation, we have

$$\hat\rho_n(A,B,m)=\frac{\tilde K_n(B,\xi_n)}{\tilde e_n(\xi_n)}$$

Equations (37) and (36) imply, respectively, that

$$\lim_{n\to\infty}\mathbb{E}[\tilde K_n(B,s)]=K(B,s)\qquad \lim_{n\to\infty}\mathbb{E}[\tilde e_n(s)]=T(s)\mu_{\mathcal{C}}(A).$$

With this in mind, we split

$$\hat\rho_n(A,B,m)-\rho(A,B,m)=\frac{\tilde K_n(B,\xi_n)-K(B,\xi_n)}{\tilde e_n(\xi_n)}-\frac{\rho(A,B,m)}{\tilde e_n(\xi_n)}\{\tilde e_n(\xi_n)-\mu_{\mathcal{C}}(A)\,T_{\mathcal{C}}(\xi_n)\}\ . \tag{41}$$

Thus, we only need to find the correct norming sequence w_n and asymptotic distribution in $\mathcal{D}([a,b])$ for any $0<a<b$ of the sequence of processes $w_n\{\tilde K_n(B,\cdot)-K(B,\cdot)\}$. To do this, define further

$$K_n(B,s)=\mathbb{E}[\tilde K_n(B,s)]\ . \tag{42}$$

Then

$$\tilde K_n(B,s)-K(B,s)=\tilde K_n(B,s)-K_n(B,s)+K_n(B,s)-K(B,s)\ .$$

hal-00624821, version 1 - 19 Sep 2011

The term $K_n(B,s) - K(B,s)$ is a deterministic bias term that will be dealt with by the second order condition (24). Write $\tilde{K}_n - K_n = (ng(k/n))^{-1/2} E_{n,1} + E_{n,2}$ with

$$E_{n,1}(B,s) = \frac{1}{\sqrt{ng(k/n)}} \sum_{j=1}^{n} \{W_{j,n}(s)V_j(B) - \mathbb{E}[W_{j,n}(s)V_j(B) \mid \mathcal{X}]\} , \tag{43}$$

$$\begin{aligned} E_{n,2}(B,s) &= \frac{1}{ng(k/n)} \sum_{j=1}^{n} \mathbb{E}[W_{j,n}(s)V_j(B) \mid \mathcal{X}] - K_n(B,s) \\ &= \frac{1}{n} \sum_{j=1}^{n} \{G_n(s, \mathbf{X}_{j,j+h-1}, \mathbf{X}_{j+m,j+m+h'}, B) - K_n(B,s)\} . \end{aligned} \tag{44}$$

The term in (43) will be called the i.i.d. term. It is a sum of conditionally independent random variables. The term in (44) will be called the dependent term. It is a function of the dependent vectors $(\mathbf{X}_{j,j+h-1}, \mathbf{X}_{j+m,j+m+h'})$.

We now state some claims whose proofs are postponed to the end of this section. The implication of Claims 1 and 3 is, in particular, that in the weakly dependent case only the i.i.d. part contributes to the limit.

Claim 1. *The process $E_{n,1}$ converges in the sense of finite-dimensional distributions to a Gaussian process W with covariance*

$$\begin{aligned} &(\mathbb{E}[\sigma^\alpha(X_1)])^\beta \mathrm{cov}(W(B,s), W(B',s')) \\ &\quad = \mathbb{E}\Big[\mathcal{L}_1(A, \boldsymbol{\sigma}(\mathbf{X}_{1,h}), \boldsymbol{\sigma}(\mathbf{X}_{1,h}), s, s') \times \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B, \mathbf{Y}_{m,m+h'} \in B' \mid \mathcal{X})\Big] \\ &\qquad + \sum_{j=2}^{h\wedge(m-h)} \mathbb{E}\Big[\mathcal{L}_j(A, \boldsymbol{\sigma}(\mathbf{X}_{1,h}), \boldsymbol{\sigma}(\mathbf{X}_{j,j+h-1}), s, s') \\ &\qquad\qquad \times \{\mathbb{P}(\mathbf{Y}_{m,m+h'} \in B, \mathbf{Y}_{m+j-1,m+h'+j-1} \in B' \mid \mathcal{X}) \\ &\qquad\qquad\qquad + \mathbb{P}(\mathbf{Y}_{m,m+h'} \in B', \mathbf{Y}_{m+j-1,m+h'+j-1} \in B \mid \mathcal{X})\}\Big] , \end{aligned} \tag{45}$$

where the functions $\mathcal{L}_j$ are defined in Assumption 2.

Claim 2. *For each fixed B, $E_{n,1}(B,\cdot)$ is tight in $\mathcal{D}([a,b])$ for each $0 < a < b$.*

This claim is proved in Lemma C.3.

The previous two statements are valid in both weakly dependent and long memory case. The next one may not be valid in the long memory case. See Section 3.2.

Claim 3. *In the weakly dependent case $E_{n,2}(B,\cdot) = O_P(\sqrt{n})$, uniformly with respect to $s \in [a,b]$ for any $0 < a < b$.*

The next claim is proved in [Kulik and Soulier, 2011, Corollary 2.4].

Claim 4. *$\xi_n - 1 = o_P(1)$.*

The last thing we need is the negligibility of the bias term.

Claim 5. *For any $a > 0$, $\sup_{s\geq a} \sup_B |K_n(B,s) - K(B,s)| = O(v_n(A))$.*

hal-00624821, version 1 - 19 Sep 2011

Therefore if $ng(k/n) \to \infty$ and (24) holds (i.e. $ng(k/n)v_n(A) \to 0$), then

$$\sqrt{ng(k/n)}\{\tilde{K}_n(B,\cdot) - K(B,\cdot), \tilde{e}_n(\cdot) - K(\mathbb{R}^d,\cdot)\} \Rightarrow (W(B,\cdot), W(\mathbb{R}^{h'+1},\cdot)) \; .$$

This convergence and the decomposition (41) imply

$$\sqrt{ng(k/n)\mu_{\mathcal{C}}(A)}\{\hat{\rho}_n(A,B,m) - \rho(A,B,m)\} \to_d W(B,1) - \rho(A,B,m)W(\mathbb{R}^{h'+1},1) \; .$$

This distribution is Gaussian. Applying (45) and the fact that $\rho(A,\mathbb{R}^{h'+1},m) = 1$, it is easily checked that its variance is given by (25). This concludes the proof of Theorem 2. □

We now prove the claims.

Proof of Claim 1. For $j = 1,\dots,n$, denote

$$\zeta_{n,j}(B,s) = \frac{1}{\sqrt{ng(k/n)}} W_{j,n}(s) V_j(B) \; .$$

In order to prove our claim, we apply the central limit theorem for m-dependent random variables, see Orey [1958]. Let $C(B,B',s,s')$ denote the quantity in the right hand side of (45). We need to check that

$$\operatorname{cov}\left(\sum_{j=1}^n \zeta_{n,j}(B,s), \sum_{j=1}^n \zeta_{n,j}(B',s') \mid \mathcal{X}\right) \to_P C(B,B',s,s') \; , \tag{46}$$

$$\sum_{j=1}^n \mathbb{E}[\zeta_{n,j}^4(B,s) \mid \mathcal{X}] \to_P 0 \; . \tag{47}$$

By standard Lindeberg-Feller type arguments, this proves the one-dimensional convergence. The finite-dimensional convergence is proved by similar arguments and by computing the asymptotic covariances. We now prove (46) and (47).

For $u \geq 1$, $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^h$, denote

$$\mathcal{L}_{n,u}(A,\mathbf{x},\mathbf{x}',s,s') = \frac{\mathbb{P}(\boldsymbol{\sigma}(\mathbf{x})\cdot\mathbf{Z}_{1,h} \in u_n s A, \boldsymbol{\sigma}(\mathbf{x}')\cdot\mathbf{Z}_{u,u+h-1} \in u_n s' A)}{g(\bar{F}_Z(u_n))} \; .$$

For $1 \leq u \leq h$, by Assumptions 1 and 2, the functions $\mathcal{L}_{n,u}$ converge in $L^1(\mathbf{X}_{1,h}, \mathbf{X}_{u,u+h-1})$ to the functions $\mathcal{L}_u$ defined in Assumption 2. For $u > h$, $\mathbf{Z}_{1,h}$ and $\mathbf{Z}_{u,u+h-1}$ are independent, so $\mathcal{L}_{n,u}$ converges a.s. and in $L^1(\mathbf{X}_{1,h}, \mathbf{X}_{u,u+h-1})$ to 0.

The random variables $\zeta_{n,j}$ are $m + h'$ dependent. Thus,

$$\begin{aligned}\operatorname{cov}\left(\sum_{j=1}^n \zeta_{n,j}(B,s), \sum_{j=1}^n \zeta_{n,j}(B',s') \mid \mathcal{X}\right) &= \sum_{j=1}^n \operatorname{cov}(\zeta_{n,j}(B,s), \zeta_{n,j}(B',s') \mid \mathcal{X}) \\ &\quad + \sum_{j=1}^n \sum_{u=1}^{m+h'} \operatorname{cov}(\zeta_{n,j}(B,s), \zeta_{n,j+u}(B',s') \mid \mathcal{X}) \qquad (48) \\ &\quad + \sum_{j=1}^n \sum_{u=1}^{m+h'} \operatorname{cov}(\zeta_{n,j+u}(B,s), \zeta_{n,j}(B',s') \mid \mathcal{X}) \; . \qquad (49)\end{aligned}$$

hal-00624821, version 1 - 19 Sep 2011

For $u = 1, \dots, h \wedge (m-h)$ it is easily seen that

$$\begin{aligned}
&\sum_{j=1}^{n} \mathrm{cov}(\zeta_{n,j}(B,s), \zeta_{n,j+u}(B',s') \mid \mathcal{X}) \\
&\qquad \sim \frac{g(\bar{F}_Z(u_n))}{ng(k/n)} \sum_{j=1}^{n} \mathcal{L}_{n,u}(\mathbf{X}_{j,j+h-1}, \mathbf{X}_{j+u,j+u+h-1}, s, s') \\
&\qquad\qquad\qquad\qquad \times \mathbb{P}(\mathbf{Y}_{j+m,j+m+h} \in B, \mathbf{Y}_{j+u+m,j+u+m+h'} \in B' \mid \mathcal{X}) \\
&\qquad \to_P \frac{\mathbb{E}\left[\mathcal{L}_u(A, \mathbf{X}_{1,h}, \mathbf{X}_{u,h+u-1}, s, s')\mathbb{P}(\mathbf{Y}_{m,m+h} \in B, \mathbf{Y}_{u+m,u+m+h'} \in B' \mid \mathcal{X})\right]}{(\mathbb{E}[\sigma^\alpha(X)])^\beta} .
\end{aligned}$$

This yields the right-hand side of (45), so we must prove that the terms in (48) and (49) are negligible. If $h > m-h$, then for large n and $m-h < u \le h$, we have $(u_n s' A) \cap B = 0$, so, for all $j = 1 \dots, n$,

$$\begin{aligned}
&\mathbb{P}\big(\mathbf{Y}_{j,j+h-1} \in u_n s A, \mathbf{Y}_{j+u,j+u+h-1} \in u_n s' A, \\
&\qquad\qquad\qquad \mathbf{Y}_{j+m,j+m+h} \in B, \mathbf{Y}_{j+u+m,j+u+m+h'} \in B' \mid \mathcal{X}\big) = 0 .
\end{aligned}$$

For $u > h$, then as mentioned above, $\mathcal{L}_u(A, \cdot, \cdot, s, s')$ converges to 0 in $L^1(\mathbf{X}_{1,h}, \mathbf{X}_{u,u+h-1})$ so

$$\sum_{j=1}^{n} \mathrm{cov}(\zeta_{n,j}(B,s), \zeta_{n,j+u}(B',s') \mid \mathcal{X}) \to_P 0 .$$

This proves (46). Next, since $\zeta_{n,j}$ are indicators and applying (37)

$$\sum_{j=1}^{n} \mathbb{E}[\zeta_{n,j}^4(B,s)] \le C \frac{\mathbb{E}[W_{1,n}(s,A)V_1(B)]}{ng(k/n)} \to 0 .$$

This proves (47) and the weak convergence of finite dimensional distributions. □

Proof of Claim 3. By definition of the functions L_n and G_n (cf. (38) and (39)), it clearly holds that

$$|G_n(s, \mathbf{X}_{j,j+h-1}, \mathbf{X}_{j+m,j+m+h'}, B)| \le L_n(s, \mathbf{X}_{j,j+h-1}) .$$

We apply the variance inequality (B.3) in the weak dependence case to get

$$\mathrm{var}(E_{n,2}(B,s)) \le \frac{C}{n} \mathrm{var}(G_n(s, \mathbf{X}_{1,h}, \mathbf{X}_{1+m,1+m+h'}, B)) \le \frac{1}{n} \mathbb{E}[L_n^2(s, \mathbf{X}_{1,h})] .$$

By (11), $L_n(s, \mathbf{x}) \le M_A(\boldsymbol{\sigma}(\mathbf{x}))$. Thus, by Assumption 3, the right hand side is uniformly bounded, thus $\mathrm{var}(E_{n,2}(B,s)) = O(1/n)$ and for any fixed $s > 0$, $\sqrt{n}E_{n,2}(B,s) = O_P(1)$. Tightness follows from Lemma C.4, thus $E_{n,2}(B,\cdot)$ converges uniformly to 0 on any compact set of $(0,\infty]$. □

Proof of Claim 5. Consider now the bias term $K_n - K$. Recall that (see (42) and (37))

$$K_n(B,s) = \mathbb{E}[\bar{K}_n(B,s)] \to T_{\mathcal{C}}(s)\mu_{\mathcal{C}}(A)\rho(A,B,m) = K(B,s)$$

hal-00624821, version 1 - 19 Sep 2011

Therefore, $K_n(B,s)$ converges pointwise to $K(B,s)$. The goal here is to show that this convergence is uniform. Using the definition of K_n, (38) and (39) we have

$$K_n(B,s) = \mathbb{E}[G_n(s,\mathbf{X}_{1,h},\mathbf{X}_{m,m+h'},B)] = \mathbb{E}[L_n(s,\mathbf{X}_{1,h})\mathbb{P}(\sigma(\mathbf{X}_{m,m+h'})\cdot\mathbf{Z}_{m,m+h'}\in B \mid \mathcal{X})] \; .$$

Using this definition and recalling the formula for $\rho(A,B,m)$ (see (13))

$$K(B,s) = T_{\mathcal{C}}(s)\mathbb{E}[L(\mathbf{X}_{1,h})\mathbb{P}(\sigma(\mathbf{X}_{m,m+h'})\cdot\mathbf{Z}_{m,m+h'}\in B \mid \mathcal{X})] \; .$$

Therefore, recalling the definition (22) of $v_n(A)$, we obtain that

$$|K_n(B,s) - K(B,s)| \le \mathbb{E}\left[\sup_{s\ge 1}|L_n(s,\mathbf{X}_{1,h}) - T_{\mathcal{C}}(s)L(\mathbf{X}_{1,h})|\right] = v_n(A) \; .$$

□

Proof of Corollary 3. In the following, $\mathbf{y}$ stands for the set $(-\infty,\mathbf{y}]$ in the previous notation. For $\mathbf{y}\in\mathbb{R}^{h'+1}$, rewrite the decomposition (41) in the present context to get

$$\hat{\Psi}_n(\mathbf{y}) - \Psi(\mathbf{y}) = \frac{\tilde{K}_n(\mathbf{y},\xi_n) - K(\mathbf{y},\xi_n)}{\tilde{e}_n(\xi_n)} - \frac{\Psi(\mathbf{y})}{\tilde{e}_n(\xi_n)}\{\tilde{e}_n(\xi_n) - \mu_{\mathcal{C}}(A)\,T_{\mathcal{C}}(\xi_n)\} \; .$$

Thus we need only prove that the sequence of suitably normalized processes $\tilde{K}_n(s,\mathbf{y}) - K_n(\mathbf{y},s)$ converge weakly to the claimed limit. The convergence of finite dimensional distributions follows from Theorem 2 and the tightness follows from Lemmas C.3 and C.4. □

Proof of Theorem 4. Claims 1, 2, 4 and 5 hold under the assumptions of Theorem 4. Thus, the result will follow if we prove a modified version of Claim 3.

Claim 6. *If $2\tau(A,B)(1-H) < 1$, then $\gamma_n^{-\tau(A,B)/2}E_{n,2}(A,B,\cdot)$ converges weakly uniformly on compact sets of $(0,\infty]$ to a process $T_{\mathcal{C}}\cdot Z(A,B)$ where the random variable $Z(A,B)$ is in a Gaussian chaos of order $\tau(A,B)$ and its distribution depends only on the Gaussian process $\{X_n\}$.*

For any $d\in\mathbb{N}^*$, $\mathbf{q}\in\mathbb{N}^d$ and $\mathbf{x}\in\mathbb{R}^d$, denote

$$\mathbf{H}_{\mathbf{q}}(\mathbf{x}) = \prod_{i=1}^{d} H_{q_i}(x_i) \; .$$

Define $\mathbb{X}_j = (X_{j+1},\dots,X_{j+h},X_{j+m},\dots,X_{j+m+h'})$. The Hermite coefficients of $G_n(s,\cdot)$ and G with respect to $\mathbb{X}_0$ can be expressed, for $\mathbf{q}\in\mathbb{N}^{h+h'+1}$, as

$$J_n(\mathbf{q},s) = \mathbb{E}[\mathbf{H}_{\mathbf{q}}(\mathbb{X}_0)G_n(s,\mathbb{X}_0)] \; , \quad J(\mathbf{q}) = \mathbb{E}[\mathbf{H}_{\mathbf{q}}(\mathbb{X}_0)G(\mathbb{X}_0)] \; .$$

Since $G_n(s,\cdot)$ converges to $T(s)G(\cdot)$ in $L^p(\mathbb{X}_0)$ for some $p>1$, $J_n(\mathbf{q},s)$ converges to $T_{\mathcal{C}}(s)J(\mathbf{q})$. Let U be an $(h+h'+1)\times(h+h'+1)$ matrix such that UU' is equal to the inverse of the covariance matrix of $\mathbb{X}_0$. Define $J_n^*(\mathbf{q},s) = \mathbb{E}[\mathbf{H}_{\mathbf{q}}(U\mathbb{X}_0)G_n(s,U\mathbb{X}_0)]$ and $J^*(q) = \mathbb{E}[\mathbf{H}_{\mathbf{q}}(U\mathbb{X}_0)G(\mathbb{X}_0)]$. Under Assumption 5, the function G_n can be expanded for $\mathbf{x}\in\mathbb{R}^{h+h'+1}$ as

$$G_n(s,\mathbf{x}) - \mathbb{E}[G_n(s,\mathbb{X}_0)] = \sum_{|\mathbf{q}|=\tau(A,B)} \frac{J_n^*(\mathbf{q},s)}{\mathbf{q}!}\mathbf{H}_{\mathbf{q}}(U\mathbf{x}) + r_n(s,\mathbf{x}) \; ,$$

hal-00624821, version 1 - 19 Sep 2011

where r_n is implicitly defined and has Hermite rank at least $\tau(A,B)+1$ with respect to $U\mathbb{X}_0$. Denote $R_n(s) = n^{-1}\sum_{j=1}^n r_n(s,\mathbb{X}_j)$. Applying (B.3), we have

$$\mathrm{var}\,(R_n(s)) \le C\left(\gamma_n^{\tau(A,B)+1} \vee \frac{1}{n}\right)\mathrm{var}(G_n(s,\mathbb{X}_0)) \le C\left(\gamma_n^{\tau(A,B)+1} \vee \frac{1}{n}\right)\mathbb{E}[L_n^2(s,\mathbf{X}_{1,h})]\ .$$

By Assumption 3, $\mathbb{E}[L_n^2(s,\mathbf{X}_{1,h})]$ is uniformly bounded, thus $\mathrm{var}(R_n(s)) = o(\gamma_n^{\tau(A,B)})$ and $\gamma_n^{-\tau(A,B)}R_n(s)$ converges weakly to zero. The convergence is uniform by an application of Lemma C.1.

Thus, the asymptotic behaviour of $\gamma_n^{-\tau(A,B)/2}E_{n,2}$ is the same as that of

$$Z_n(s) = \sum_{|\mathbf{q}|=\tau(A,B)} \frac{J_n^*(\mathbf{q},s)n^{-1}}{\mathbf{q}!}\,\gamma_n^{-\tau(A,B)/2}\sum_{j=1}^n \mathbf{H}_{\mathbf{q}}(U\mathbb{X}_j)\ .$$

By [Arcones, 1994, Theorem 6], there exist random variables $\aleph^*(\mathbf{q})$ such that $Z_n(s)$ converges to

$$T_{\mathcal{C}}(s)\sum_{|\mathbf{q}|=\tau(A,B)}\frac{J^*(\mathbf{q})}{\mathbf{q}!}\,\aleph^*(\mathbf{q})$$

for each $s \ge 0$. To prove that the convergence is uniform, we only need to prove that $J_n^*(\mathbf{q},\cdot)$ converges uniformly to $T_{\mathcal{C}}\cdot J^*(\mathbf{q})$ for each $\mathbf{q}$ such that $|\mathbf{q}| = \tau(A)$. Since the coefficients J_n^* can be expressed linearly in terms of the coefficients J_n, it suffices to prove uniform convergence of the coefficients J_n. Applying Hölder inequality, we obtain, for $p>1$ and for any $a>0$,

$$\sup_{s\ge a}|J_n(\mathbf{q},s) - T_{\mathcal{C}}(s)J(\mathbf{q})| \le C\mathbb{E}\left[\sup_{s\ge a}|L_n(s,\mathbf{X}_{1,h}) - T_{\mathcal{C}}(s)L(\mathbf{X}_{1,h})|^p\right]\ .$$

We have already seen that this last quantity converges to 0 for $p=2$ by Assumption 3. □

Appendix

A Second order regular variation of convolutions

Denote $A \asymp B$ if there exists positive constant c_1 and c_2 such that $c_1A \le B \le c_2B$.

Lemma A.1. *Let Z_1 and Z_2 be i.i.d. non negative random variables with common distribution function F that satisfies Assumption 6. Then*

$$\left|\mathbb{P}(u_1Z_1+u_2Z_2>t) - \bar F(t/u_1) - \bar F(t/u_2)\right| \le C u_1^{\alpha+\epsilon}u_2^{\alpha+\epsilon}t^{-1}\bar F(t)\int_0^t \bar F(s)\,\mathrm{d}s\ .$$

Proof. Obviously, we have

$$\begin{aligned}\mathbb{P}(u_1Z_1+u_2Z_2>t) &= \bar F(t/u_1) + \bar F(t/u_2) - \bar F(t/u_1)\bar F(t/u_2)\\ &+\mathbb{P}(t/2<u_1Z_1\le t)\mathbb{P}(t/2<u_2Z_2\le t)\\ &+\mathbb{P}(u_1Z_1\le t/2, u_2Z_2\le t, u_1Z_1+u_2Z_2>t)\\ &+\mathbb{P}(u_2Z_2\le t/2, u_1Z_1\le t, u_1Z_1+u_2Z_2>t)\ .\end{aligned}$$

hal-00624821, version 1 - 19 Sep 2011

Consider for instance the second last term. It may be written as

$$I_1 := \mathbb{E}\left[\mathbf{1}_{\{u_1 Z_1 \le t/2\}} \left\{ \frac{\bar F(t(1-u_1Z_1/t)/u_2)}{\bar F(t/u_2)} - 1 \right\}\right] .$$

Since F satisfies Assumption 6, we have, for $u \in [1/2,1]$,

$$\begin{aligned} 0 \le \frac{\bar F(ut)}{\bar F(t)} - 1 &= u^{-\alpha} \mathrm{e}^{\int_1^u \frac{\eta(ts)}{s}\,\mathrm{d}s} - 1 = \{u^{-\alpha}-1\}\mathrm{e}^{\int_1^u \frac{\eta(ts)}{s}\,\mathrm{d}s} + \mathrm{e}^{\int_1^u \frac{\eta(ts)}{s}\,\mathrm{d}s} - 1 \\ &\le |u^{-\alpha}-1| \mathrm{e}^{\int_{1/2}^1 \frac{\eta^*(ts)}{s}\,\mathrm{d}s} + \mathrm{e}^{\int_{1/2}^1 \frac{\eta^*(ts)}{s}\,\mathrm{d}s} \int_u^1 \frac{\eta^*(ts)}{s}\,\mathrm{d}s . \end{aligned}$$

Since $\eta^*(t)$ is decreasing, we have, for all $u \in [1/2,1]$,

$$0 \le \frac{\bar F(ut)}{\bar F(t)} - 1 \le C\{|u^{-\alpha}-1| + \log(u)\} \le C(1-u) .$$

Applying this inequality with $1-u = u_1Z_1/t$ on the event $u_1Z_1 \le t/2$ yields

$$I_1 \le C u_1 t^{-1} \mathbb{E}\left[Z_1 \mathbf{1}_{\{u_1Z_1 \le t\}}\right] \le C t^{-1} \int_0^{t/u_1} \bar F(s)\,\mathrm{d}s = C t^{-1} u_1^{-1} \int_0^t \bar F(s/u_1)\,\mathrm{d}s .$$

By Potter's bounds, for any $\epsilon > 0$, there exists a constant C such for any $s,t>0$,

$$\frac{\bar F(s/u_1)}{\bar F(s)} \le C(u_1^{-1} \wedge 1)^{-\alpha-\epsilon} .$$

Applying this bound we obtain

$$I_1 \le C(u_1 \vee 1)^{\alpha+\epsilon}(u_2 \vee 1)^{\alpha+\epsilon} t^{-1} \bar F(t) \int_0^t \bar F(s)\,\mathrm{d}s .$$

To conclude, note that $\bar F^2(t) = O(t^{-1}\bar F(t)\int_0^t \bar F(s)\,\mathrm{d}s)$ if $\alpha < 1$ and $\bar F^2(t) = o(t^{-1}\bar F(t)\int_0^t \bar F(s)\,\mathrm{d}s)$ if $\alpha \ge 1$. □

Remark 6. By induction, we can obtain the bound

$$\left|\mathbb{P}(Z_1 + \cdots + Z_n > t) - n\bar F(t)\right| \le C\,t^{-1}\bar F(t) \int_0^t \bar F(s)\,\mathrm{d}s ,$$

and we can also recover a particular case of a result of Omey and Willekens [1987] in a slightly different form. For $\alpha \ge 1$ and $\mathbb{E}[Z_1] < \infty$,

$$\lim_{t\to\infty} t\left\{\frac{\mathbb{P}(Z_1 + \cdots + Z_n > t)}{\mathbb{P}(Z_1 > t)} - n\right\} = \frac{n(n-1)}{2}\mathbb{E}[Z_1] .$$

hal-00624821, version 1 - 19 Sep 2011

B Multivariate Hermite expansions and variance inequalities for Gaussian processes

Consider a multidimensional stationary centered Gaussian process $\{\mathbf{X}_n\}$ with autocovariance function $\gamma_n(i,j) = \mathbb{E}[X_0^{(i)} X_n^{(j)}]$ and assume either

$$\forall 1 \leq i,j \leq d \;, \quad \sum_{n=0}^{\infty} |\gamma_n(i,j)| < \infty \;, \tag{B.1}$$

or that there exists $H \in (1/2, 1)$ and a function ℓ slowly varying at infinity such that

$$\lim_{n\to\infty} \frac{\gamma_n(i,j)}{n^{2H-2}\ell(n)} = b_{i,j} \;, \tag{B.2}$$

and the coefficients $b_{i,j}$ are not identically zero. Then, we have the following inequality due to Arcones [1994].

For any function G such that $\mathbb{E}[G^2(\mathbb{X}_0)] < \infty$ and with Hermite rank q with respect to $\mathbf{X}_0$,

$$\mathrm{var}\left(n^{-1} \sum_{j=1}^{n} G(\mathbf{X}_j) \right) \leq C(\ell^q(n) n^{2q(H-1)}) \vee n^{-1}\, \mathrm{var}(G(\mathbf{X}_0)) \;. \tag{B.3}$$

where the constant C depends only on the Gaussian process $\{\mathbf{X}_n\}$ and not on the function G. This bound summarizes Equations 2.18, 3.10 and 2.40 in Arcones [1994]. The rate obtained is n^{-1} in the weakly dependent case where (B.1) holds and in the case where (B.2) holds and G has Hermite rank q such that $q(1-H) > 1$. Otherwise, the rate is $\ell^q(n) n^{2q(H-1)}$.

C A criterion for tightness

We state a criterion for the tightness of a sequence of random processes with path in $\mathcal{D}(\mathbb{R}^d)$, which adapts to the present context Bickel and Wichura [1971, Theorem 3] and the remarks thereafter.

Let T be a rectangle $T = T_1 \times T_d \subset \mathbb{R}^d$. A block B in T is a subset of T of the form $\prod_{i=1}^d (s_i, t_i]$ with $s_i < t_i$, $1 \leq i \leq d$. Disjoint blocks $B = \prod_{i=1}^d (s_i, t_i]$ and $B' = \prod_{i=1}^d (s'_i, t'_i]$ are neighbours if there exists $p \in \{1, \ldots, d\}$ such that $s'_p = t_p$ or $s_p = t'_p$ and $s_i = s'_i$ and $t_i = t'_i$ for $i \neq p$. (In the terminology of Bickel and Wichura [1971] the blocks B and B' are said to share a common face.) Let X be a random process indexed by T. The increment of the process X over a block $B = \prod_{i=1}^d (s_i, t_i]$ is defined by

$$X(B) = \sum_{(\epsilon_1,\ldots,\epsilon_d)\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \epsilon_i} X(s_1 + \epsilon_1(t_1 - s_1), \ldots, s_d + \epsilon_d(t_d - s_d)) \;.$$

(This is the usual d-dimensional increment of a random process X. If for instance $d = 2$, then $X(B) = X(t_1,t_2) - X(t_1,s_2) - X(s_1,t_2) + X(s_1,s_2)$). If X is an indicator, i.e. $X(\mathbf{y}) = \mathbf{1}_{\{\mathbf{Y}\leq\mathbf{y}\}}$ for some T valued random variable $\mathbf{Y}$, then $X(B) = \mathbf{1}_{\{\mathbf{Y}\in B\}}$.

hal-00624821, version 1 - 19 Sep 2011

Lemma C.1. *Let $\{\zeta_n\}$ be sequence of stochastic processes indexed by a compact rectangle $T \subset \mathbb{R}^d$. Assume that the finite dimensional marginal distributions of ζ_n converges weakly to those of a process ζ which is continuous on the upper boundary of T. Assume moreover that there exist $\gamma \geq 0$ and $\beta > 1$ such that*

$$\mathbb{P}(|\zeta_n(B)| \wedge |\zeta_n(B')| \geq \lambda) \leq C\lambda^{-\gamma}\mathbb{E}[\mu_n^\beta(B \cup B')] \tag{C.1}$$

for some sequence of random probability measures μ_n which converges weakly in probability to a (possibly random) probability measure μ with (almost surely) continuous marginals. Then the sequence of processes $\{\zeta_n\}$ is tight in $\mathcal{D}(T, \mathbb{R})$.

Sketch of proof. For f defined on $T = T_1 \times \cdots \times T_d$, $i \in \{1, \ldots, d\}$ and $t \in T_i$, define $f_t^{(i)}$ on $T_1 \times \cdots \times T_{i-1} \times T_{i+1} \times \cdots \times T_d$ by

$$f_t^{(i)}(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_d) = f(t_1, \ldots, t_{i-1}, t, t_{i+1}, \ldots, t_d)$$

and define, for $s < t \in T_i$ and $\delta > 0$,

$$\begin{aligned} w_i''(f, s, t) &= \sup_{s<u<v<w<t} \|f_u^{(i)} - f_v^{(i)}\|_\infty \wedge \|f_v^{(i)} - f_w^{(i)}\|_\infty \,, \\ w_i''(f, \delta) &= \sup_{u<v<w<u+\delta} \|f_u^{(i)} - f_v^{(i)}\|_\infty \wedge \|f_v^{(i)} - f_w^{(i)}\|_\infty \,. \end{aligned}$$

By the Corollary of Bickel and Wichura [1971], a sequence of processes $\{X_n\}$ defined on T converges weakly in $\mathcal{D}(T)$ to a process X which is continuous at the upper boundary of T with probability one, if the finite-dimensional marginal distributions of X_n converges to those of X and if, for all $\delta, \lambda > 0$, and al $i = 1, \ldots, d$,

$$\mathbb{P}(w_i''(X_n, \delta) > \lambda) \to 0 \,. \tag{C.2}$$

For any measure μ on T, define its i-th marginal $\mu^{(i)}$ by

$$\mu^{(i)}((s, t]) = \mu(T_1 \times \cdots \times T_{i-1} \times (s, t] \times T_{i+1} \times \cdots \times T_d) \,, s, t \in T_i \,.$$

As mentioned in the remarks after the proof of Bickel and Wichura [1971, Theorem 3], an easy adaptation of the proof of Billingsley [1968, Theorem 15.6] shows that (C.2) is implied by

$$\mathbb{P}(w_i''(X_n, s, t) > \lambda) \leq C\lambda^{-\gamma}\mathbb{E}[\{\mu_n^{(i)}(s, t])\}^\beta] \,, \tag{C.3}$$

where μ_n satisfies the assumptions of the Lemma. So we must show that (C.1) implies (C.3). The proof is by induction, so the first step is to prove it in the one-dimensional case, where (C.1) becomes, for $u < v < w \in T$,

$$\mathbb{P}(|\zeta_n(v) - \zeta_n(u)| \wedge |\zeta_n(w) - \zeta_n(v)| \geq \lambda) \leq C\lambda^{-\gamma}\mathbb{E}[\mu_n^\beta((u, w])] \,. \tag{C.4}$$

The proof of (C.3) under the assumption (C.4) follows the lines of the proof of [Billingsley, 1968, (15.26)] under the assumption [Billingsley, 1968, (15.21)]. The key ingredient is the maximal inequality [Billingsley, 1968, Theorem 12.5], which can be easily adapted as follows in the present context. Let $S_0, \ldots, S_n$ be random variables. Assume that there exists nonnegative random variables $u_1, \ldots, u_n$ such that

$$\mathbb{P}(|S_i - S_j| \wedge |S_k - S_j| > \lambda) \leq \lambda^{-\gamma}\mathbb{E}[(u_i + \cdots + u_k)^\beta]$$

hal-00624821, version 1 - 19 Sep 2011

for some $\beta > 1$ and $\gamma \geq 0$ and all $1 \leq i \leq j \leq k \leq n$ and, then there exists a constant C that depends only on β and γ such that

$$\mathbb{P}\left(\max_{1\leq i\leq j\leq k\leq n} |S_i - S_j| \wedge |S_k - S_j| > \lambda\right) \leq C\lambda^{-\gamma}\mathbb{E}[(u_1 + \cdots + u_n)^\beta] \; .$$

Proving by induction that (C.1) implies (C.3) in the d-dimensional case can be done exactly along the lines of Step 5 of the proof of Bickel and Wichura [1971, Theorem 1]. ☐

In order to apply this criterion to the context of empirical processes, we need the following Lemma which slightly extends the bound Billingsley [1968, (13.18)].

Lemma C.2. *Let $\{(B_i, B'_i)\}$ be a sequence of m-dependent vectors, where B_i and B'_i are Bernoulli random variables, with parameters p_i and q_i, respectively, and such that $B_iB'_i = 0$ a.s. Denote $S_n = \sum_{j=1}^n (B_j - p_j)$ and $S'_n = \sum_{j=1}^n (B'_j - q_j)$. Then, there exists a constant C which depends only on m, such that*

$$\mathbb{E}[S_n^2 {S'_n}^2] \leq C\left(\sum_{i=1}^n p_i\right)\left(\sum_{i=1}^n q_i\right) \leq C\left(\sum_{i=1}^n p_i \vee q_i\right)^2 \; . \tag{C.5}$$

Proof. We start by assuming that the pairs (B_i, B'_i) are i.i.d. and we prove (C.5) by induction. For any integrable random variable X, denote $\bar{X} = X - \mathbb{E}[X]$. For $n = 1$, since $B_1B'_1 = 0$, we obtain $\mathbb{E}[\bar{B}_i\bar{B}'_i] = -p_iq_i$ and

$$\begin{aligned}\mathbb{E}[\bar{B}_1^2 {\bar{B}'}^2_1] &= \mathbb{E}[(B_1 - 2p_1B_1 + p_1^2)(B'_1 - 2q_1B'_1 + q_1^2)] \\ &= p_1q^2 + p_1^2q_1 - 3p_1^2q^2 = p_1q_1(p_1 + q_1 - 3p_1q_1) \leq p_1q_1 \; .\end{aligned}$$

The last inequality comes from the fact that $B_1B'_1 = 0$ a.s. implies that $p_i + q_i \leq 1$, and $0 \leq p + q - 3pq \leq p + q \leq 1$ for all $p, q \geq 0$ such that $p + q \leq 1$. Assume now that (C.5) holds with $C = 3$ for some $n \geq 1$. Then, denoting $s_n = \sum_{j=1}^n p_j$ and $s'_n = \sum_{j=1}^n q_j$, we have

$$\begin{aligned}&\mathbb{E}[S_{n+1}^2 {S'_{n+1}}^2] \\ &= \mathbb{E}[S_n^2 {S'_n}^2] + \mathbb{E}[S_n^2]\mathbb{E}[{\bar{B}'}^2_{n+1}] + \mathbb{E}[{S'_n}^2]\mathbb{E}[\bar{B}_{n+1}^2] + 4\mathbb{E}[S_nS'_n]\mathbb{E}[B_{n+1}B'_{n+1}] + \mathbb{E}[\bar{B}_{n+1}^2{\bar{B}'}^2_{n+1}] \\ &\leq 3s_ns'_n + s_nq_{n+1} + s'_np_{n+1} + 4p_{n+1}q_{n+1}\sum_{i=1}^n p_iq_i + p_{n+1}q_{n+1} \\ &\leq 3s_ns'_n + 3s_nq_{n+1} + 3s'_np_{n+1} + p_{n+1}q_{n+1} \leq 3s_{n+1}s'_{n+1} \; .\end{aligned}$$

This proves that (C.5) holds for al $n \geq 1$.

We now consider the case of m-dependence. Let a_i, $1 \leq i \leq n$ be a sequence of real numbers and set $a_i = 0$ if $i > n$. Then

$$\left(\sum_{i=1}^n a_i\right)^2 = \left(\sum_{q=1}^m \sum_{j=1}^{\lceil n/m\rceil} a_{(j-1)m+q}\right)^2 \leq m\sum_{q=1}^m \left(\sum_{j=1}^{\lceil n/m\rceil} a_{(j-1)m+q}\right)^2 \; .$$

hal-00624821, version 1 - 19 Sep 2011

Applying this and the bound for the independent case (extending all sequences by zero after the index n) yields

$$\mathbb{E}[S_n^2 {S'_n}^2] \leq 3m^2 \sum_{q=1}^{m} \sum_{q'=1}^{m} \sum_{j=1}^{\lceil n/m \rceil} \sum_{j'=1}^{\lceil n/m \rceil} p_{(j-1)m+q} p_{(j'-1)m+q'} = 3m^2 s_n s'_n \; .$$

□

Let us apply this criterion in the context of section 3. Fix a cone $\mathcal{C}$ and a relatively compact subset $A \in \mathcal{C}$. Recall that $E_{n,1}$ and $E_{n,2}$ are defined in (43) and (44).

Lemma C.3. *Under the assumptions of Theorem 2 or 4, for any fixed* $B \in \mathbb{R}^{h'+1}$, $E_{n,1}(B,\cdot)$ *is tight in* $\mathcal{D}([a,b])$*, and if moreover* $\Psi_{A,m,h}$ *is continuous, then* $E_{n,1}$ *is tight in* $\mathcal{D}(\mathcal{K} \times [a,b])$ *for any* $0 < a < b$ *and any compact set* $\mathcal{K}$ *of* $\mathbb{R}^{h'+1}$.

Proof. By Assumption 4, if $s < t$, then $tA \subset sA$. Thus, a sequence of random measures $\hat{\mu}_n$ on $\mathbb{R}^d \times (0,\infty)$ can be defined by

$$\begin{aligned}\hat{\mu}_n((-\boldsymbol{\infty}, \mathbf{y}] \times (s,\infty)) &= \frac{1}{n} \sum_{j=1}^{n} \frac{\mathbb{P}(\mathbf{Y}_{j,h} \in s u_n A \mid \mathcal{X})}{g(k/n)} \mathbb{P}(\mathbf{Y}_{j+m,j+m+h'} \leq \mathbf{y} \mid \mathcal{X}) \\ &= \frac{1}{n} \sum_{j=1}^{n} G_n(s, \mathbf{X}_{j,h} \mathbf{X}_{j+m,j+m+h'}, \mathbf{y}) \; ,\end{aligned}$$

where G_n is defined in (39). Then $\hat{\mu}_n$ converges vaguely in probability to the measure μ defined by

$$\mu((-\boldsymbol{\infty}, \mathbf{y}] \times (s,\infty)) = \mu_{\mathcal{C}}(A) T(s) \Psi_{A,m,h}(\mathbf{y}) \; .$$

Then, by conditional m-dependence, for any neighbouring relatively compact blocs D, D' of $\mathbb{R}^d \times (0,\infty]$, applying Lemma C.2 yields

$$\mathbb{E}[E_{n,1}^2(D) E_{n,2}^2(D') \mid \mathcal{X}] \leq C \hat{\mu}_n(D) \hat{\mu}_n(D') \; .$$

Taking unconditional expectations then yields

$$\mathbb{E}[E_{n,2}^2(D) E_{n,2}^2(D')] \leq C \hat{\mathbb{E}}[\mu_n(D) \hat{\mu}_n(D')] \leq \mathbb{E}[\hat{\mu}_n^2(D \cup D')] \; .$$

Thus (C.1) holds with $\beta = \gamma = 2$. In the context of Theorem 2, for any fixed B, this implies that $E_{n,1}(B,\cdot)$ is fixed, since the limiting distribution is proportional to $T(s)$ which is continuous. If the distribution function Ψ is assumed to be continuous, then Lemma C.1 applies and the process $E_{n,1}$ is tight with respect to both variables. □

Lemma C.4. *Under the assumptions of Theorem 2, for any fixed* $B \in \mathbb{R}^{h'+1}$, $E_{n,2}(B,\cdot)$ *converges uniformly to zero on compact sets of* $(0,\infty]$*. Under the assumption of Corollary 3,* $E_{n,2}$ *converges uniformly to zero on compact sets of* $\mathbb{R}^{h'+1} \times (0,\infty]$.

hal-00624821, version 1 - 19 Sep 2011

Proof. We only need to prove the tightness. By the variance inequality (B.3) and Hölder's inequality, we have, for any relatively compact neighbouring blocks D, D' of $\mathbb{R}^d \times (0,\infty)$,

$$\begin{aligned}\mathbb{P}(|E_{2,n}(D)| \wedge |E_{2,n}(D')| \geq \lambda) \leq \lambda^{-2}\sqrt{\mathbb{E}[E_{2,n}^2(D)]\mathbb{E}[E_{2,n}^2(D')]} &\leq \lambda^{-2}\mathbb{E}[E_{2,n}^2(D \cup D')] \\ &\leq C\lambda^{-2}n^{-1}\mathbb{E}[\tilde{\mu}_n^2(D \cup D')]\end{aligned}$$

where $\tilde{\mu}_n$ is the random measure defined by

$$\tilde{\mu}_n(s, \mathbf{y}) = \frac{\mathbb{P}(\mathbf{Y}_{1,h} \in s u_n A \mid \mathcal{X})}{g(k/n)} \mathbb{P}(\mathbf{Y}_{m,m+h'} \leq \mathbf{y} \mid \mathcal{X}) \ .$$

Assumptions 1 and 3 imply that $\tilde{\mu}_n$ converges vaguely on $\mathbb{R}^d \times (0,\infty]$, in probability and in the mean square to the measure $\hat{\mu}$ defined by

$$\hat{\mu}((-\boldsymbol{\infty}, \mathbf{y}] \times (s,\infty]) = \frac{\nu_{\mathcal{C}}(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A)}{(\mathbb{E}[\nu_C(\boldsymbol{\sigma}(\mathbf{X}_{1,h})^{-1} \cdot A])^{\beta}} T(s) \mathbb{P}(\mathbf{Y}_{m,m+h'} \leq \mathbf{y} \mid \mathcal{X}) \ .$$

The measure $\hat{\mu}$ has continuous marginals if we consider the case of a fixed B (which takes care of Theorem 4). The marginals of $\hat{\mu}$ are almost surely continuous if F_Z is continuous, so Lemma C.1 applies. ☐

References

Miguel A. Arcones. Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability*, 22(4):2242–2274, 1994.

Peter J. Bickel and Michael J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Annals of Mathematical Statistics*, 42:1656–1670, 1971.

Patrick Billingsley. *Convergence of probability measures.* New York, Wiley, 1968.

Nicholas H. Bingham, Charles M. Goldie, and Jan L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

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.

Leo Breiman. On some limit theorems similar to the arc-sine law. *Theory of Probability and Applications*, 10:323–331, 1965.

Richard A. Davis and Thomas Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 38A:977–1009, 2009. Probability, statistics and seismology.

Holger Drees. Optimal rates of convergence for estimates of the extreme value index. *The Annals of Statistics*, 26(1):434–448, 1998.

Andrew C. Harvey. Long memory in stochastic volatility. In J. Knight and S. Satchell (eds), *Forecasting volatility in financial markets*. Butterworth-Heinemann, London, 1998.

hal-00624821, version 1 - 19 Sep 2011

Clifford M. Hurvich, Eric Moulines, and Philippe Soulier. Estimating long memory in volatility. *Econometrica*, 73(4):1283–1328, 2005.

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.

Edward Omey and Eric Willekens. Second-order behaviour of distributions subordinate to a distribution with finite mean. *Communications in Statistics. Stochastic Models*, 3(3): 311–342, 1987.

Steven Orey. A central limit theorem for m-dependent random variables. *Duke Mathematical Journal*, 25:543–546, 1958.

Sidney I. Resnick. *Extreme values, regular variation, and point processes*. Springer-Verlag, New York, 1987.

Sidney I. Resnick. *Heavy-tail phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

Sidney I. Resnick. Multivariate regular variation on cones: application to extreme values, hidden regular variation and conditioned limit laws. *Stochastics*, 80(2-3):269–298, 2008.