EIGENVALUES OF SAMPLE COVARIANCE MATRICES OF NON-LINEAR PROCESSES WITH INFINITE VARIANCE

RICHARD A. DAVIS AND OLIVER PFAFFEL

ABSTRACT. We study the k-largest eigenvalues of heavy-tailed sample covariance matrices of the form $\mathbf{X}\mathbf{X}^\mathsf{T}$ in an asymptotic framework, where the dimension of the data and the sample size tend to infinity. To this end, we assume that the rows of $\mathbf{X}$ are given by independent copies of some stationary process with regularly varying marginals with index $\alpha \in (0,2)$ satisfying large deviation and mixing conditions. We apply these general results to stochastic volatility and GARCH processes.

1. INTRODUCTION

In the statistical analysis of high-dimensional data one often tries to reduce its dimensionality while preserving as much of the variation in the data as possible. One important example of such an approach is the Principal Component Analysis (PCA). PCA makes a linear transformation of the data to a new set of variables, the principal components, which are ordered such that the first few retain most of the variation. Therefore one obtains a lower dimensional representation of the data by retaining only the first few principal components.

The variances of the first k principal components are given by the k-largest eigenvalues of the covariance matrix. Let us collect the samples of our multivariate data in a $p \times n$ matrix $\mathbf{X}$, where we refer to p as the dimension of the data and to n as the sample size. In practice, the true underlying covariance matrix is not available, thus one usually replaces it with the sample covariance matrix $\frac{1}{n}\mathbf{X}\mathbf{X}^\mathsf{T}$. For more details on PCA we refer the reader to one of the many textbooks available on this topic, see [2] or [13], for example.

To account for large high-dimensional data sets, we study the k-largest eigenvalues of the sample covariance matrix when both the dimension of the data p as well as the sample size n go to infinity. The field of research that investigates the spectral properties of large dimensional random matrices has become known as Random Matrix Theory (RMT). There exist several survey articles which stress the close relationship between Random Matrix Theory and multivariate statistics, including PCA, see e.g. [10] and [12]. Some authors have already employed tools from Random Matrix Theory to correct traditional tests or estimators which fail when the dimension of the data cannot be assumed to be negligible compared to the sample size. For example, Bai

et al. [3] gives corrections on some likelihood ratio tests that even fail even for moderate dimension (around 20), and El Karoui [11] consistently estimates the spectrum of a large dimensional covariance matrix using Random Matrix Theory.

Davis, Pfaffel and Stelzer [9] study the k-largest eigenvalues of a sample covariance matrix based on observations that come from a high-dimensional linear process with heavy-tailed marginals. It is often necessary to use non-linear instead of linear models to capture the complex dependence structure of the data. This is particularly true in finance where the log-returns exhibit both non-linearity and heavy-tails. The objective of this paper is to extend some of the results of [9] to a non-linear setting.

In Section 2 we study non-linear processes with regularly varying tail probabilities with index less than two which satisfy certain large deviation and mixing conditions. We then apply our results to two heavily employed models in finance; stochastic volatility models in Section 3 and GARCH(p,q) processes in Section 4. More background on these processes and other financial time series models may be found in [1], for example. We assume throughout sections 2 – 4 that the dimension $p \approx n^\beta$ with $\beta > 0$ satisfying $\beta < \frac{2-\alpha}{\alpha-1}$ if $1 < \alpha < 2$. This restriction is rather general, however, if α is close to 2, it becomes quite strict. Therefore we will present a result for the largest eigenvalue of $\mathbf{XX}^\mathsf{T}$ in Section 5 which holds for $p \geq n$ independently of the value of α as long as $0 < \alpha < 2$.

This article makes heavy use of the theory of regular variation and point processes. A good introduction may be found in Resnick [15], for example.

2. Eigenvalues of heavy-tailed sample covariance matrices of stationary processes

Throughout this section we assume that (X_t) is a strictly stationary sequence of random variables with marginals that are regularly varying with tail index smaller than two. In other words, there exist a normalizing sequence (a_n) and an $\alpha \in (0,2)$ such that, for any $x > 0$,

$$(2.1) \qquad nP(|X_0| > a_n x) \underset{n\to\infty}{\longrightarrow} x^{-\alpha}.$$

Moreover we assume that X_0 satisfies the tail balancing condition, i.e., that the limit

$$\lim_{x\to\infty} \frac{P(X_0 > x)}{P(|X_0| > x)} \quad \text{exists.}$$

Then we construct our $p \times n$ observation matrix $\mathbf{X} = (X_{it})_{it}$ as follows: for each $1 \leq i \leq p$, let $(X_{it})_{1 \leq t \leq n}$ be an independent copy of $(X_t)_{1 \leq t \leq n}$. This means that $(X_{it})_t$ and $(X_t)_t$ have the same distribution, and all rows of $\mathbf{X}$ are independent. We denote by $\lambda_1, \ldots, \lambda_p \geq 0$ the eigenvalues of $\mathbf{XX}^\mathsf{T}$ and study them via their induced point process

$$\sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i}(B) = \left| \left\{ 1 \leq i \leq p : a_{np}^{-2}\lambda_i \in B \right\} \right|, \quad B \subseteq (0, \infty).$$

Our motivation to study this problem comes from the statistical analysis of high-dimensional data. Therefore we assume in the following that $p = p_n$ is an integer valued sequence such that

$p_n \to \infty$ as $n \to \infty$.

The upcoming theorem is a combination of results from [9] where mostly linear processes are studied. This general theorem will be the basis of all our further results.

Theorem 2.1. *Assume that there exists a $b \geq 0$ such that*

$$(2.2) \qquad pP\left(\sum_{t=1}^{n} X_t^2 > a_{np}^2 x\right) \xrightarrow[n\to\infty]{} bx^{-\alpha/2} \qquad \text{for each } x > 0.$$

Suppose that $p = p_n \to \infty$ and $n \to \infty$ such that

$$(2.3) \qquad \limsup_{n\to\infty} \frac{p_n}{n^\beta} < \infty$$

for some $\beta > 0$, satisfying $\beta < \frac{2-\alpha}{\alpha-1}$ if $1 < \alpha < 2$. Then we have that

$$(2.4) \qquad \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{D} \sum_{i=1}^{\infty} \epsilon_{b^{2/\alpha}\Gamma_i^{-2/\alpha}}$$

as $n \to \infty$, where $\Gamma_i = E_1 + \ldots + E_i$ is the successive sum of independent and identically distributed (iid) Exponential random variables E_k with mean one.

The convergence in (2.4) means that, for any function $f : (0,\infty) \to (0,\infty)$ with compact support,

$$E\left(e^{-\sum_{i=1}^{p} f(a_{np}^{-2}\lambda_i)}\right) \xrightarrow[n\to\infty]{} E\left(e^{-\sum_{i=1}^{\infty} f(b^{2/\alpha}\Gamma_i^{-2/\alpha})}\right).$$

Proof. Proposition 3.3 of [9] shows that

$$(2.5) \qquad a_{np}^{-2}\left\|\mathbf{XX}^\mathsf{T} - D\right\|_2 \xrightarrow[n\to\infty]{P} 0.$$

This means that $\mathbf{XX}^\mathsf{T}$ can be approximated by its diagonal. Using Weyl's inequality ([5, Corollary III.2.6]), its eigenvalues are therefore asymptotically equal to its diagonal entries. By [15, Proposition 3.21], the large deviation result (2.2) implies that the point process of the diagonal entries of $\mathbf{XX}^\mathsf{T}$ converges to a Poisson point process,

$$(2.6) \qquad \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\sum_{t=1}^{n} X_{it}^2} \xrightarrow[n\to\infty]{D} \sum_{i=1}^{\infty} \epsilon_{b^{2/\alpha}\Gamma_i^{-2/\alpha}}.$$

Along the lines of the proof of [9, Theorem 1] it follows that this results carries over to the eigenvalues, since, as mention before, they behave like the diagonal entries. $\square$

Remark 2.2. Observe that (2.4) immediately implies the joint convergence of the k-largest eigenvalues of $\mathbf{XX}^\mathsf{T}$ in distribution. Denote by $\lambda_{(1)} \geq \ldots \geq \lambda_{(p)} \geq 0$ the eigenvalues of $\mathbf{XX}^\mathsf{T}$ in decreasing order. Then we have, for any fixed integer k, that

$$a_{np}^{-2}(\lambda_{(1)}, \ldots, \lambda_{(k)}) \xrightarrow[n\to\infty]{D} b^{2/\alpha}(\Gamma_1^{-2/\alpha}, \ldots, \Gamma_k^{-2/\alpha}).$$

If $b = 0$, then the normalized eigenvalues converge to zero in probability.

The large deviation condition (2.2) is essentially equivalent to the convergence of the point process of the partial sums of X_t^2 to a limiting point process. Instead of having a condition on the partial sums, it would be much more convenient in many cases to have a condition on the process itself. Davis and Hsing [6] give very general conditions under which the point process convergence of the sequence X_t^2 gives a large deviation result for the partial sums of X_t^2. This will be stated in our next theorem.

Theorem 2.3. *Assume that*

$$(2.7) \qquad \sum_{i=1}^{n} \epsilon_{a_n^{-2} X_i^2} \xrightarrow[n \to \infty]{D} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \epsilon_{P_i Q_{ij}},$$

where $\sum_{i=1}^{\infty} \epsilon_{P_i}$ is a Poisson process on $(0, \infty)$ with some intensity measure ν, and $(\sum_{j=1}^{\infty} \epsilon_{Q_{ij}})_i$ is a sequence of iid point processes on $[-1, 0) \cup (0, 1]$ independent of $\sum_{i=1}^{\infty} \epsilon_{P_i}$. Further assume that the sequence (X_t) is strongly mixing. If $p, n \to \infty$ such that (2.3) is satisfied, then we have (2.4) with

$$(2.8) \qquad b = \lim_{\delta \to 0} \int_0^\infty P\left(\sum_{i=1}^{\infty} u Q_{1i} \mathbf{1}_{(\delta, \infty)}(u|Q_{1i}|) > 1 \right) \nu(du) \in [0, \infty).$$

Proof. Under the above conditions, [6, Theorem 4.3] shows that

$$\lim_{n \to \infty} \frac{P\left(\sum_{t=1}^{n} X_t^2 > a_{np}^2 x \right)}{n P(X_0^2 > a_{np}^2 x)} = \lim_{\delta \to 0} \int_0^\infty P\left(\sum_{i=1}^{\infty} u Q_{1i} \mathbf{1}_{(\delta, \infty)}(u|Q_{1i}|) > 1 \right) \nu(du) =: b.$$

Since $pnP(X_0^2 > a_{np}^2 x) \to x^{-\alpha/2}$, this implies

$$pP\left(\sum_{t=1}^{n} X_t^2 > a_{np}^2 x \right) \xrightarrow[n \to \infty]{} b x^{-\alpha/2}.$$

An application of Theorem 2.1 completes the proof. $\qquad\qquad\square$

Remark 2.4. The assumption that the sequence (X_t) is strongly mixing can be replaced by the much weaker assumption [6, (2.1)].

In the remainder of this article we apply the results of this section to stochastic volatility and GARCH processes. For a stochastic volatility process the constant b in (2.4) is essentially one, see (3.4). In the case of a GARCH process the characterization of b is more involved.

3. STOCHASTIC VOLATILITY MODELS

In this section we use the previously introduced techniques to obtain results for the eigenvalues of sample covariance matrices when the observations are given by stochastic volatility models. More precisely, the rows of the observation matrix $\mathbf{X}$ are given by independent copies of a univariate stochastic volatility process $X_t = \sigma_t Z_t$, and $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of $\mathbf{X}\mathbf{X}^{\mathsf{T}}$.

Theorem 3.1. *Assume that (Z_t) is an iid sequence with regularly varying tails with index $\alpha \in (0, 2)$ and normalizing sequence (a_n). This means we have that*

$$(3.1) \qquad nP(|Z_0| > a_n x) \underset{n \to \infty}{\longrightarrow} x^{-\alpha} \quad \text{for any } x > 0.$$

Moreover we assume that Z_0 satisfies the tail balancing condition

$$(3.2) \qquad \lim_{x \to \infty} \frac{P(Z_0 > x)}{P(|Z_0| > x)} = q \in [0, 1] \quad \text{exists.}$$

Let $\sigma_t \geq 0$ be a stationary sequence with $E\sigma_0^{4\alpha} < \infty$ independent of (Z_t). Then $X_t = \sigma_t Z_t$ defines a stochastic volatility process. Suppose $p = p_n \underset{n \to \infty}{\longrightarrow} \infty$ such that

$$0 < \liminf_{n \to \infty} \frac{p_n}{n^{\beta_1}} \quad \text{and} \quad \limsup_{n \to \infty} \frac{p_n}{n^{\beta_2}} < \infty$$

for some $0 < \beta_1 < \beta_2$, where $\beta_2 < \frac{2-\alpha}{\alpha-1}$ in case $1 < \alpha < 2$. Then we have, as $n \to \infty$, that

$$(3.3) \qquad \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \overset{D}{\longrightarrow} \sum_{i=1}^{\infty} \epsilon_{(E\sigma_0^\alpha)^{2/\alpha}\Gamma_i^{-2/\alpha}},$$

where $\Gamma_i = E_1 + \ldots + E_i$ is the successive sum of iid Exponential random variables E_k with mean one.

Proof. Theorem 4.2 of [14] applied to $X_t^2 = \sigma_t^2 Z_t^2$ gives that

$$\lim_{n \to \infty} \frac{P\left(\sum_{t=1}^{n} X_t^2 > a_{np}^2 x\right)}{nP(X_0^2 > a_{np}^2 x)} = 1.$$

Since $pnP(X_0^2 > a_{np}^2 x) \to x^{-\alpha/2} E\sigma_0^\alpha$, this implies that

$$pP\left(\sum_{t=1}^{n} X_t^2 > a_{np}^2\right) \underset{n \to \infty}{\longrightarrow} x^{-\alpha/2} E\sigma_0^\alpha.$$

Theorem 2.1 concludes. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Remark 3.2. Note that the factor $(E\sigma_0^\alpha)^{2/\alpha}$ in the limiting point process in (3.3) appears due to the fact that a_n is the normalizing sequence of Z_1 and not X_1. If $\tilde{a}_n$ denotes the normalizing sequence of X_1, then we have that

$$(3.4) \qquad \sum_{i=1}^{p} \epsilon_{\tilde{a}_{np}^{-2}\lambda_i} \overset{D}{\longrightarrow} \sum_{i=1}^{\infty} \epsilon_{\Gamma_i^{-2/\alpha}}.$$

Thus we get the same result for a stochastic volatility process as we get for an iid sequence. The intuitive reason for this is that the dependence within the sequence (X_t) is inherited by the light-tailed sequence (σ_t). Hence, the extremes of (X_t) are asymptotically independent.

A combination of [14, Theorem 4.6] and Theorem 2.1 can be used to obtain analogous results for regularly varying Markov chains.

The additional assumption of $0 < \liminf_{n\to\infty} \frac{p_n}{n^{\beta_1}}$ in Theorem 3.1 is needed to apply Theorem 4.2 of [14]. We can eliminate this assumption at the expense of imposing mixing properties on the volatility sequence (σ_t). This is the content of the following theorem.

Theorem 3.3. *Let $X_t = \sigma_t Z_t$ be a stochastic volatility process such that $|Z_0|$ is regularly varying satisfying (3.1) and (3.2) with tail index $\alpha \in (0,2)$ and normalizing sequence (a_n). Assume that the volatility sequence (σ_t) is independent of (Z_t) and either*

- *a stationary sequence of m-dependent non-negative random variables such that $E\sigma_0^{\alpha+\delta} < \infty$ for some $\delta > 0$, or*
- *the exponential of a linear process with Gaussian noise, i.e.,*

$$\log \sigma_t = \sum_{k=-\infty}^{\infty} \psi_k \xi_{t-k},$$

where (ξ_t) is a sequence of iid mean-zero Gaussian random variables and (ψ_k^2) is summable.

Suppose that $p,n \to \infty$ such that (2.3) is satisfied. Then we have (3.3).

Proof. Depending on the choice of the volatility process (σ_t), Theorem 3.1. or Theorem 3.3. of Davis and Mikosch [8] yield that

$$\sum_{i=1}^{\infty} \epsilon_{a_n^{-1}X_i} \xrightarrow[n\to\infty]{D} \sum_{i=1}^{\infty} \epsilon_{\tilde{P}_i},$$

where the limiting point process is Poisson with intensity measure

$$\tilde{\nu}(dx) = \alpha E\sigma_0^{\alpha}[qx^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x) + (1-q)(-x)^{-\alpha-1}\mathbf{1}_{(-\infty,0)}(x)]dx.$$

An application of the continuous mapping theorem yields that

$$\sum_{i=1}^{\infty} \epsilon_{a_n^{-2}X_i^2} \xrightarrow[n\to\infty]{D} \sum_{i=1}^{\infty} \epsilon_{\tilde{P}_i^2} = \sum_{i=1}^{\infty} \epsilon_{P_i}.$$

The limiting point process $\sum_{i=1}^{\infty} \epsilon_{P_i}$ of the squares has intensity measure

$$\nu(dx) = \alpha E\sigma_0^{\alpha}x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x)dx$$

and representation (2.7) with $Q_{ij} = 1$ if $i = j$ and zero otherwise. For either choice of (σ_t), this sequence is strongly mixing. Davis and Mikosch [8] have shown that (X_t) is strongly mixing if (σ_t) is. Thus we can apply Theorem 2.3. Due to the simple structure of Q, we have that

$$b = \lim_{\delta\to 0} \int_0^{\infty} P\left(\sum_{i=1}^{\infty} uQ_{1i}\mathbf{1}_{(\delta,\infty)}(u|Q_{1i}|) > 1\right)\nu(du) = -\frac{\alpha}{2}E\sigma_0^{\alpha}\int_1^{\infty} u^{-\alpha/2-1}du = E\sigma_0^{\alpha}.$$

$\square$

4. GARCH PROCESSES

A GARCH(p,q) process (X_t) with $p \geq 1$ or $q \geq 1$ is given by the equations

$$(4.1) \qquad\qquad X_t = \sigma_t Z_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

where α_i and β_j are non-negative coefficients such that $\alpha_0 > 0$, $\alpha_p > 0$ if $p \geq 1$ and $\beta_q > 0$ if $q \geq 1$, and (Z_t) is an iid sequence. This is not a stochastic volatility model since (σ_t^2) is not independent of (Z_t). We will make use of the results of Section 2 to obtain the point process convergence of the eigenvalues of $\mathbf{X}\mathbf{X}^{\mathsf{T}}$ when the rows of $\mathbf{X}$ are given by iid copies of a GARCH process.

Theorem 4.1. *Suppose that (Z_t) is an iid sequence satisfying the following conditions:*

- *Z_1 is symmetric, i.e., $Z_1 \overset{D}{=} -Z_1$, and has a strictly positive density on $\mathbb{R}$.*
- *There exists an $h_0 \in (0, \infty]$ such that $E|Z_1|^h < \infty$ for $h < h_0$ and $E|Z_1|^{h_0} = \infty$.*

Assume there exists a unique strictly stationary solution (X_t) to equation (4.1). Then X_1 is regularly varying with index $\alpha > 0$ and some normalizing sequence (a_n). Suppose that $\alpha < 2$. Then we have, as $p, n \to \infty$ such that (2.3) is satisfied, that the convergence in (2.4) holds with b as given in (2.8).

The intensity measure ν in (2.8) is given by $\nu(x, \infty) = \gamma x^{-\alpha}$ for some $\gamma \in (0, 1]$, and (Q_{ij}) satisfies $\sup_j |Q_{ij}| = 1$ for all $i \geq 1$. The distribution of (Q_{ij}) is described in detail in Theorem 2.8 of Davis and Mikosch [7].

Proof. By Basrak, Davis and Mikosch [4, Corollary 3.5] we obtain that (X_t) is strongly mixing and has regularly varying tail probabilities with index $\alpha > 0$. Since we assume that $\alpha < 2$, we can apply Theorem 2.3. In [4, Theorem 2.10] it is shown, using the results from [7], that the point process $\sum_{i=1}^{n} \epsilon_{a_n^{-2} X_i^2}$ has a limiting point process with representation (2.7). For details, see also the proof of [4, Theorem 3.6]. $\qquad\square$

Remark 4.2.   (*i*) Note that $\gamma \in (0, 1]$ is the extremal index of the sequence $(|X_t|)$.

(*ii*) In the case when (X_t) is a GARCH$(1, 1)$ process, the theorem simplifies considerably. If $\alpha_0, \alpha_1, \beta_1 > 0$, then (X_t) has a unique strictly stationary solution if and only if

$$-\infty < E \log(\alpha_1 Z_1^2 + \beta_1) < 0,$$

cf. [1, p.47]. Let us further assume that $EZ_1 = 0$ and $EZ_1^2 = 1$. In this case, the index α of regular variation of X_1 is given by the unique solution to the equation

$$E[(\alpha_1 Z_1^2 + \beta_1)^\alpha] = 1.$$

To finish this section we want to mention that, using the results from [4] and [7], the same sort of argument could be applied to the class of solutions of stochastic recurrence equations.

5. The largest eigenvalue in the case when $p \geq n$

In the previous sections we assumed that (2.3) is satisfied for some $\beta > 0$ with $\beta < \frac{2-\alpha}{\alpha-1}$ in the case $1 < \alpha < 2$. This implies for $\alpha \in (1.5, 2)$ that p grows slower than n. For example, in statistical genetics one is often confronted with the case that the dimension of the data p (e.g., the number of genes) is much larges than the sample size n (e.g., the number of patients). This may also happen in finance when one considers the problem of optimizing a large portfolio. Therefore we present a result in the upcoming theorem for the largest eigenvalue of $\mathbf{XX}^\mathsf{T}$ that holds for $p \geq n$ independently of the value of α as long as $0 < \alpha < 2$. It shows that the largest eigenvalue $\lambda_{\max}$ of $\mathbf{XX}^\mathsf{T}$ is asymptotically equal to the maximum of the squares of the entries of $\mathbf{X}$.

Theorem 5.1. *Let (X_t) be a strictly stationary stochastic process which satisfies (2.1) with normalizing sequence (a_n) and tail index $\alpha \in (0, 2)$. Assume that $(|X_t|)$ has extremal index equal to one. Further let $p = p_n = n^\kappa$ for some $\kappa \geq 1$. Suppose, for any $x > 0$, that*

$$\lim_{n\to\infty} \frac{P(\sum_{t=1}^n |X_t| > a_{np}x)}{nP(|X_0| > a_{np}x)} = 1,$$

and

$$\lim_{n\to\infty} \frac{P(\sum_{t=1}^n X_t^2 > a_{np}^2 x)}{nP(X_0^2 > a_{np}^2 x)} = 1.$$

Assume the rows of $\mathbf{X}$ are given by independent copies of (X_t), and denote by $\lambda_{\max}$ the largest eigenvalue of $\mathbf{XX}^\mathsf{T}$. Then we have

$$(5.1) \qquad \frac{\lambda_{\max}}{\max_{1\leq i\leq p, 1\leq t\leq n} X_{it}^2} \xrightarrow[n\to\infty]{P} 1.$$

In particular, this implies that

$$\lim_{n\to\infty} P(\lambda_{\max} \leq a_{np}^2 x) = \exp(-x^{-\alpha/2}).$$

Proof. Since $(|X_t|)$ has extremal index equal to one, and $\max_{1\leq t\leq n} |X_t| \leq \sum_{t=1}^n |X_t|$, we have, as $n \to \infty$, that

$$1 - o(1) \leq \frac{P(\max_{1\leq t\leq n} |X_t| > a_{np} \max\{x, y\})}{nP(|X_0| > a_{np} \max\{x, y\})}$$

$$\leq \frac{P(\sum_{t=1}^n |X_t| > a_{np}x, \max_{1\leq t\leq n} |X_t| > a_{np}y)}{nP(|X_0| > a_{np} \max\{x, y\})} \leq \frac{P(\sum_{t=1}^n |X_t| > a_{np} \max\{x, y\}),}{nP(|X_0| > a_{np} \max\{x, y\})} \to 1.$$

The latter converges to one by assumption, thus

$$\frac{pP(\sum_{t=1}^n |X_t| > a_{np}x, \max_{1\leq t\leq n} |X_t| > a_{np}y)}{npP(|X_0| > a_{np} \max\{x, y\})} \to 1.$$

Hence

$$\sum_{i=1}^p \epsilon_{a_{np}^{-1}(\sum_{t=1}^n |X_{it}|, \max_{1\leq t\leq n} |X_{it}|)} \xrightarrow[n\to\infty]{D} \sum_{i=1}^\infty \epsilon_{\Gamma_i^{-1/\alpha}(1,1)}.$$

This yields

$$(5.2) \qquad \frac{\|X\|_\infty}{\max_{i,t} |X_{it}|} = \frac{\max_i \sum_{t=1}^n |X_{it}|}{\max_{i,t} |X_{it}|} \xrightarrow[n\to\infty]{P} 1.$$

Likewise, we obtain for the squares

$$(5.3) \qquad \frac{\max_i \sum_{t=1}^n X_{it}^2}{\max_{i,t} X_{it}^2} \xrightarrow[n\to\infty]{P} 1.$$

Now we proceed like in the proof of [9, Theorem 2]. Since $\max_i \sum_{t=1}^n X_{it}^2 \le \lambda_{\max} \le \|X\|_\infty^2$, we obtain by (5.2) and (5.3) that

$$\frac{\lambda_{\max}}{\max_{1\le i\le p} \sum_{t=1}^n X_{it}^2} \xrightarrow[n\to\infty]{P} 1.$$

$\square$

The above theorem applies to stochastic volatility models, as we show in the upcoming corollary. It should not be mistaken as a special case of Theorem 3.1 since the growth condition on p_n is different.

Corollary 5.2. *Let $X_t = \sigma_t Z_t$ such that (Z_t) is an iid sequence with regularly varying tails satisfying (3.1) and (3.2) with index $0 < \alpha < 2$ and normalizing sequence (a_n). Assume that $\sigma_t \ge 0$ is a stationary sequence independent of (Z_t) that is strongly mixing with rate function $r_j = O(j^{-a})$ with $a > 1$. Further assume that $E\sigma_0^{4\alpha+\delta} < \infty$ for some $\delta > 0$. Let $n \to \infty$ and suppose that $p = p_n = n^\kappa$ for some $\kappa \ge 1$. Then we have (5.1). This implies that*

$$\lim_{n\to\infty} P\left(\frac{\lambda_{\max}}{a_{np}^2 (E\sigma_0^\alpha)^{2/\alpha}} \le x \right) = \exp(-x^{-\alpha/2}).$$

Proof. Since the square of a stochastic volatility process is again a stochastic volatility process, it suffices to show that

$$\lim_{n\to\infty} \frac{P(\sum_{t=1}^n |X_t| > a_{np}x)}{nP(|X_0| > a_{np}x)} = 1.$$

For $0 < \alpha < 1$ this has been done in [14, Theorem 4.2] (note that $a_{np} = n^{\epsilon+1/\alpha}$ since $p = n^\kappa$). For $1 < \alpha < 2$ this has been shown for the centered sum. However, since $\kappa \ge 1$ and $\alpha < 2$, we have that $nE|X_0|/a_{np}$ converges to zero, therefore a centering is not needed. $\square$

References

[1] T. Andersen, R. Davis, J. Kreiß, and T. Mikosch. *Handbook of financial time series*. Springer, 2009.

[2] T. W. Anderson. *An introduction to multivariate statistical analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience, Hoboken, third edition, 2003. ISBN 0-471-36091-0.

[3] Z. Bai, D. Jiang, J.-F. Yao, and S. Zheng. Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.*, 37(6B):3822–3840, 2009. ISSN 0090-5364. doi: 10.1214/09-AOS694. URL `http://dx.doi.org/10.1214/09-AOS694`.

[4] B. Basrak, R. Davis, and T. Mikosch. Regular variation of garch processes. *Stochastic Processes and their Applications*, 99(1):95–115, 2002.

[5] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. ISBN 0-387-94846-5.

[6] R. Davis and T. Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *The Annals of Probability*, 23(2):879–917, 1995.

[7] R. Davis and T. Mikosch. The sample autocorrelations of heavy-tailed processes with applications to arch. *The Annals of Statistics*, 26(5):2049–2080, 1998.

[8] R. Davis and T. Mikosch. Point process convergence of stochastic volatility processes with application to sample autocorrelation. *Journal of Applied Probability*, pages 93–104, 2001.

[9] R. Davis, O. Pfaffel, and R. Stelzer. Limit theory for the largest eigenvalues of sample covariance matrices with heavy-tails. 2011. URL `http://arxiv.org/abs/1108.5464`.

[10] N. El Karoui. Recent results about the largest eigenvalue of random covariance matrices and statistical application. *Acta Phys. Polon. B*, 36(9):2681–2697, 2005. ISSN 0587-4254.

[11] N. El Karoui. Spectrum estimation for large dimensional covariance matrices using random matrix theory. *Ann. Statist.*, 36(6):2757–2790, 2008. ISSN 0090-5364. doi: 10.1214/07-AOS581. URL `http://dx.doi.org/10.1214/07-AOS581`.

[12] I. M. Johnstone. High dimensional statistical inference and random matrices. In *International Congress of Mathematicians. Vol. I*, pages 307–333. Eur. Math. Soc., Zürich, 2007. doi: 10.4171/022-1/13. URL `http://dx.doi.org/10.4171/022-1/13`.

[13] I. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[14] T. Mikosch and O. Wintenberger. Precise large deviations for dependent regularly varying sequences. `http://arxiv.org/pdf/1206.1395.pdf`.

[15] S. I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. ISBN 978-0-387-75952-4. Reprint of the 1987 original.

Department of Statistics, Columbia University, New York, USA
E-mail address: `rdavis@stat.columbia.edu`

TUM Institute for Advanced Study & Department of Mathematics, Technische Universität München, Germany
E-mail address: `o.pfaffel@gmx.de`