Self-Normalizaton of Sums of Dependent Random Variables[1]

Olivier Wintenberger (Sorbonne, Pauli-Institut), Munyea Matsui (Nagoya), Thomas Mikosch (University of Copenhagen)

[1]Séminaire géolearning Fréjus, March 31 2025

1. Regular variation of stationary time series

[2]

(X_t) is *regularly varying with index* $\alpha > 0$ if

1. $|X|$ is regularly varying with index α: $\mathbb{P}(|X| > x) = x^{-\alpha} L(x)$.

2. There exists $(\Theta_t)_{t \geq 0}$ independent of a Pareto(α) random variable[a] Y_α such that for every $h \geq 0$ as $x \to \infty$,

$$\mathbb{P}\left(x^{-1}(X_0, \ldots, X_h) \in \cdot \,\big|\, |X_0| > x\right) \xrightarrow{w} \mathbb{P}\left(Y_\alpha \left(\Theta_0, \ldots, \Theta_h\right) \in \cdot\right).$$

[a] Pareto(α): $\mathbb{P}(Y_\alpha > x) = x^{-\alpha}$, $x > 0$

[2] Davis, Hsing (AoS 1995), Basrak, Hsing (SPA 2009)

(X_t) is ***regularly varying with index*** $\alpha > 0$ **if**

1. $|X|$ **is regularly varying with index** α**:** $\mathbb{P}(|X| > x) = x^{-\alpha} L(x)$.

2. There exists $(\Theta_t)_{t \geq 0}$ **independent of a Pareto**(α) **random variable** Y_α **such that for every** $h \geq 0$ **as** $x \to \infty$,

$$\mathbb{P}\big(x^{-1}(X_0, \ldots, X_h) \in \cdot \,\big|\, |X_0| > x\big) \xrightarrow{w} \mathbb{P}\big(Y_\alpha (\Theta_0, \ldots, \Theta_h) \in \cdot\big) .$$

- (Θ_t) **is the** ***spectral tail process*****.**

(X_t) is ***regularly varying with index*** $\alpha > 0$ **if**

1. $|X|$ **is regularly varying with index** α**:** $\mathbb{P}(|X| > x) = x^{-\alpha} L(x)$.

2. There exists $(\Theta_t)_{t \geq 0}$ **independent of a Pareto**(α) **random variable** Y_α **such that for every** $h \geq 0$ **as** $x \to \infty$,

$$\mathbb{P}\big(x^{-1}(X_0, \ldots, X_h) \in \cdot \,\big|\, |X_0| > x\big) \xrightarrow{w} \mathbb{P}\big(Y_\alpha\,(\Theta_0, \ldots, \Theta_h) \in \cdot\big)\,.$$

- (Θ_t) **is the** ***spectral tail process*** **of** (X_t).
- **It describes the propagation of large values at zero into the future.**

(X_t) is *regularly varying with index* $\alpha > 0$ if

1. $|X|$ is regularly varying with index α: $\mathbb{P}(|X| > x) = x^{-\alpha}L(x)$.

2. There exists $(\Theta_t)_{t\geq 0}$ independent of a Pareto(α) random variable Y_α such that for every $h \geq 0$ as $x \to \infty$,

$$\mathbb{P}\big(x^{-1}(X_0, \ldots, X_h) \in \cdot \,\big|\, |X_0| > x\big) \xrightarrow{w} \mathbb{P}\big(Y_\alpha\,(\Theta_0, \ldots, \Theta_h) \in \cdot\big)\,.$$

- (Θ_t) is the *spectral tail process* of (X_t).
- It describes the propagation of large values at zero into the future.
- Extremal phenomena in a time series can be described in terms of the spectral tail process.

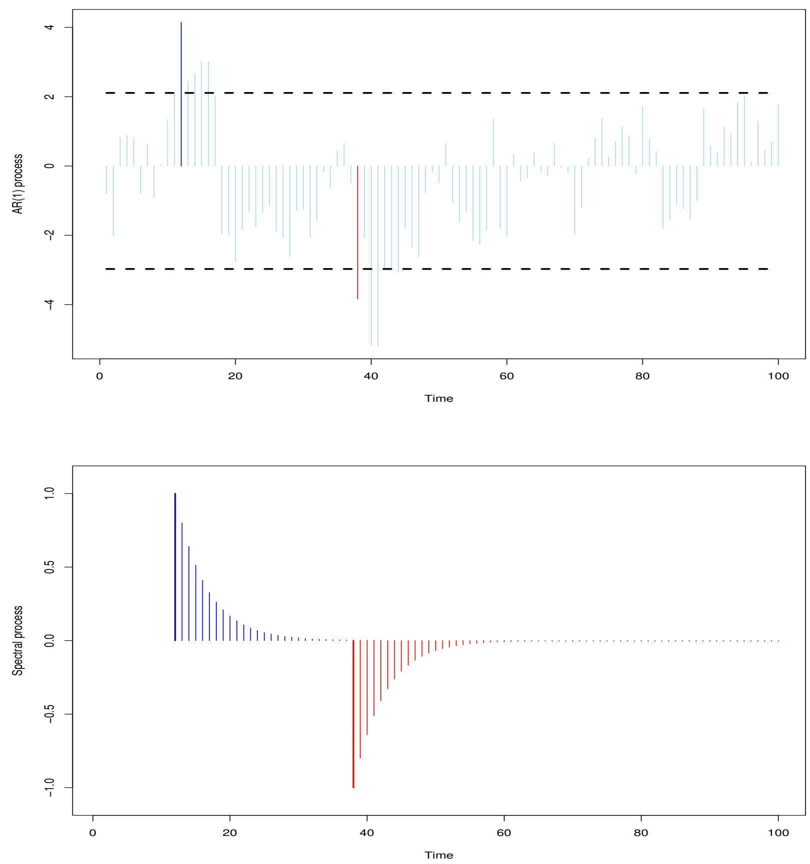


FIGURE 1. **Top.** Realization of a stationary AR(1) process $\boldsymbol{X_t = 0.8\, X_{t-1} + Z_t}$ with iid Student(5) noise $\boldsymbol{(Z_t)}$. This stationary process is regularly varying with index $\boldsymbol{\alpha = 5}$. The value $\boldsymbol{X_{12}}$ exceeds the 95%-quantile, triggering the lagged spectral tail process $\boldsymbol{0.8^{t-12}}$, $\boldsymbol{t \geq 12}$ (blue) while $\boldsymbol{X_{38}}$ falls below the 5%-quantile, triggering the lagged spectral tail process $\boldsymbol{-0.8^{t-38}}$, $\boldsymbol{t \geq 38}$ (red).

- **For non-negative X_t: consider the *maxima***

  $M_n = \max(X_1, \ldots, X_n)$ and the $(1 - 1/n)$-quantiles a_n of X.

- **For non-negative X_t: consider the *maxima***
  $M_n = \max(X_1,\ldots,X_n)$ and the $(1-1/n)$-quantiles a_n of X.
- **Under *mixing* and *anti-clustering* conditions there exists**
  $\theta_X \in (0,1]$ such that as $n \to \infty$,[3]

$$\mathbb{P}\big(a_n^{-1} M_n \le x\big) \to \Phi_\alpha^{\theta_X}(x) = \exp(-\theta_X\, x^{-\alpha})\,,\quad x > 0\,.$$

[3] $\mathbf{\Phi}_\alpha$ is the Fréchet distribution function.

- For non-negative X_t: consider the *maxima*

  $M_n = \max(X_1, \ldots, X_n)$ and the $(1 - 1/n)$-quantiles a_n of X.

- Under *mixing* and *anti-clustering* conditions there exists

  $\theta_X \in (0, 1]$ such that

$$\mathbb{P}\big(a_n^{-1} M_n \le x\big) \to \Phi_\alpha^{\theta_X}(x) = \exp(-\theta_X\, x^{-\alpha})\,, \quad x > 0\,.$$

- θ_X is the *extremal index of* (X_t).

- For non-negative X_t: consider the *maxima* $M_n = \max(X_1, \ldots, X_n)$ and the $(1-1/n)$-quantiles a_n of X.
- Under *mixing* and *anti-clustering* conditions there exists $\theta_X \in (0,1]$ such that

$$\mathbb{P}\big(a_n^{-1} M_n \le x\big) \to \Phi_\alpha^{\theta_X}(x) = \exp(-\theta_X\, x^{-\alpha})\,, \quad x > 0\,.$$

- θ_X is the *extremal index of* (X_t).
- It is interpreted as the reciprocal of the *expected extremal cluster size* above high thresholds.

- For non-negative X_t: consider the *maxima* $M_n = \max(X_1, \ldots, X_n)$ and the $(1 - 1/n)$-quantiles a_n of X.
- Under *mixing* and *anti-clustering* conditions there exists $\theta_X \in (0, 1]$ such that

$$\mathbb{P}\big(a_n^{-1} M_n \le x\big) \to \Phi_\alpha^{\theta_X}(x) = \exp(-\theta_X \, x^{-\alpha}) \,, \quad x > 0 \,.$$

- θ_X is the *extremal index of* (X_t).
- It is interpreted as the reciprocal of the *expected extremal cluster size* above high thresholds.
- The extremal index can be expressed in terms of the spectral tail process: for non-negative X_t,

$$\theta_X = \mathbb{E}\Big[\Big(1 - \sup_{t \ge 1} \Theta_t^\alpha\Big)_+\Big] \,.$$

2. Examples of regularly varying time series

- **AR(1) process:** $X_t = \varphi X_{t-1} + Z_t$, (Z_t) **iid regularly varying with index** $\alpha > 0$, $|\varphi| < 1$. **Then** (X_t) **is regularly varying with index** α **and**

$$\Theta_t = \Theta_0 \varphi^t, \qquad t \geq 0.$$

- **Affine stochastic recurrence equation:**[4] $X_t = A_t X_{t-1} + B_t$, (A_t, B_t), $t \in \mathbb{Z}$, **iid, and the equation** $\mathbb{E}[|A|^\alpha] = 1$ **has a positive solution OR** (B_t) **is regularly varying with index** α **and** $\mathbb{E}[|A|^\alpha] < 1$. **Then** (X_t) **is regularly varying with index** $\alpha > 0$ **and**

$$\Theta_t = \Theta_0 A_1 \cdots A_t, \qquad t \geq 0.$$

[4] Kesten (1973), Goldie (1991), Grincevičius (1985)

- **Stochastic volatility model:** $X_t = \sigma_t Z_t$, (σ_t) positive stationary, independent of an iid regularly varying sequence (Z_t) with index α. If $\mathbb{E}[\sigma^{\alpha+\delta}] < \infty$ for some $\delta > 0$, (X_t) is regularly varying with index α and $\Theta_t = 0$, $t \neq 0$.

  Asymptotic independence

- **GARCH(1, 1) process:**[5] $X_t = \sigma_t Z_t$, (Z_t) iid, $\mathbb{E}[Z] = 0$, $\mathbb{E}[Z^2] = 1$,

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\,\sigma_{t-1}^2\,.$$

  (σ_t^2) satisfies an affine stochastic recurrence equation.

  It is regularly varying with index $\alpha/2$ if $\mathbb{E}[(\alpha_1 Z_0^2 + \beta_1)^{\alpha/2}] = 1$ and (X_t) inherits regular variation with index α.

[5] Engle (1982), Bollerslev (1985)

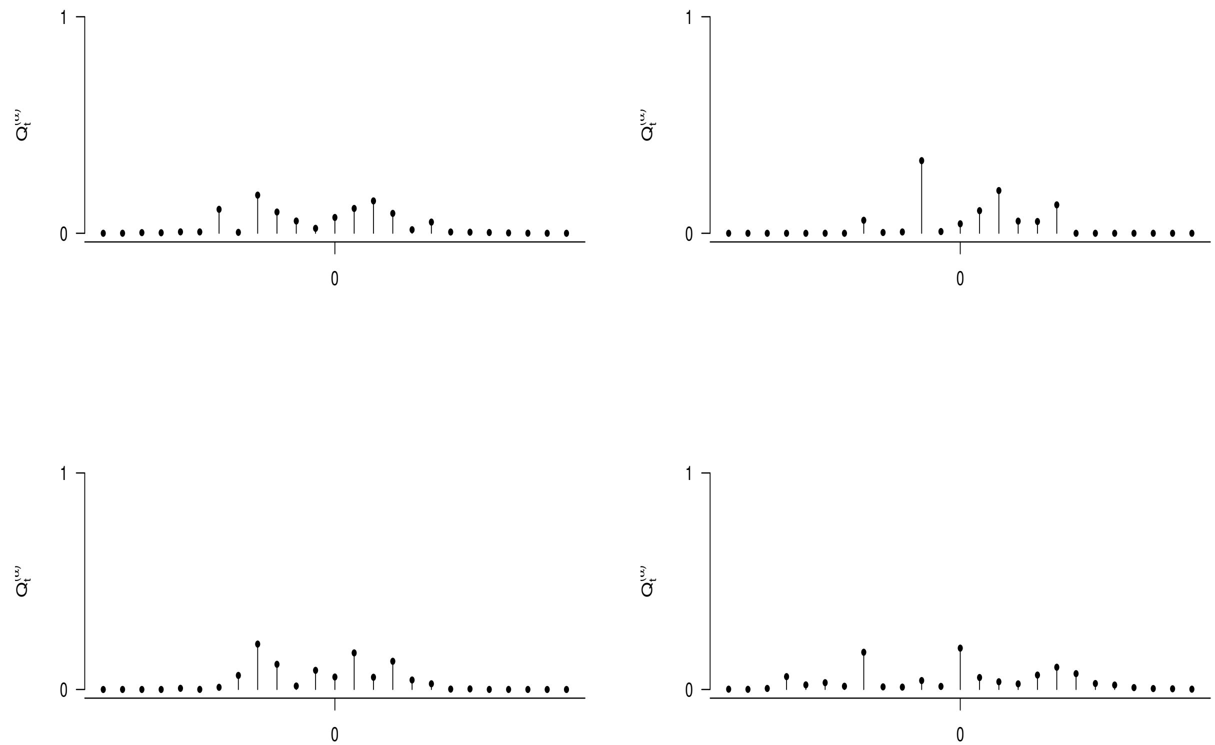


FIGURE 2. Distinct realisations of a standardized spectral tail process of a solution of a SRE.

3. Joint convergence of maxima and sums

- **Consider an iid regularly varying sequence (X_t) with index $\alpha \in (0,2)$, i.e., for a generic element X, as $x \to \infty$**

$$\mathbb{P}(|X| > x) = \frac{L(x)}{x^\alpha}, \quad \frac{\mathbb{P}(\pm X > x)}{\mathbb{P}(|X| > x)} \to p_\pm = \mathbb{P}(\Theta_0 = \pm 1)\,.$$

- **Consider an iid regularly varying sequence (X_t) with index $\alpha \in (0,2)$, i.e., for a generic element X,**

$$\mathbb{P}(|X| > x) = x^{-\alpha} L(x)\,, \quad \frac{\mathbb{P}(\pm X > x)}{\mathbb{P}(|X| > x)} \to p_\pm\,,\ x \to \infty\,.$$

- **In this case, maxima $M_n = \max(X_1, \ldots, X_n)$ and sums $S_n = X_1 + \cdots + X_n$ converge jointly:**

$$a_n^{-1}(M_n, S_n - b_n) := a_n^{-1}(\xi_\alpha, \eta_\alpha)\,, \qquad n \to \infty\,.$$

- Consider an iid regularly varying sequence (X_t) with index $\alpha \in (0,2)$, i.e., for a generic element X,

$$\mathbb{P}(|X| > x) = x^{-\alpha}\, L(x)\,, \quad \frac{\mathbb{P}(\pm X > x)}{\mathbb{P}(|X| > x)} \to p_{\pm}\,,\ x \to \infty\,.$$

- In this case, maxima $M_n = \max(X_1, \ldots, X_n)$ and sums $S_n = X_1 + \cdots + X_n$ converge jointly:

$$a_n^{-1}(M_n, S_n - b_n) := a_n^{-1}(\xi_\alpha, \eta_\alpha)\,, \qquad n \to \infty\,.$$

- ξ_α has a Fréchet Φ_α-distribution, η_α is α-stable.

- Consider an iid regularly varying sequence (X_t) with index $\alpha \in (0,2)$, i.e., for a generic element X,

$$\mathbb{P}(|X| > x) = x^{-\alpha} L(x)\,, \quad \frac{\mathbb{P}(\pm X > x)}{\mathbb{P}(|X| > x)} \to p_{\pm}\,,\ x \to \infty\,.$$

- In this case, maxima $M_n = \max(X_1, \ldots, X_n)$ and sums $S_n = X_1 + \cdots + X_n$ converge jointly:

$$a_n^{-1}(M_n, S_n - b_n) := a_n^{-1}(\xi_\alpha, \eta_\alpha)\,, \qquad n \to \infty\,.$$

- ξ_α has a Fréchet Φ_α-distribution, η_α is α-stable.
- ξ_α and η_α are dependent.

- Consider an iid regularly varying sequence (X_t) with index $\alpha \in (0,2)$, i.e., for a generic element X,

$$\mathbb{P}(|X| > x) = x^{-\alpha} L(x)\,, \quad \frac{\mathbb{P}(\pm X > x)}{\mathbb{P}(|X| > x)} \to p_\pm\,,\ x \to \infty\,.$$

- In this case, maxima $M_n = \max(X_1, \ldots, X_n)$ and sums $S_n = X_1 + \cdots + X_n$ converge jointly:

$$a_n^{-1}(M_n, S_n - b_n) := a_n^{-1}(\xi_\alpha, \eta_\alpha)\,, \qquad n \to \infty\,.$$

- ξ_α has a Fréchet Φ_α-distribution, η_α is α-stable.
- ξ_α and η_α are dependent.
- This is in stark contrast to the finite variance case $\alpha > 2$ where M_n and S_n require different normalizations a_n and $\sqrt{n}$, respectively, and the limit components are independent.

4. Self-normalized sums

- **In the iid case, sums normalized by maxima converge for $\alpha \in (0,2)$:**

$$\frac{S_n - b_n}{M_n} \stackrel{d}{\to} \frac{\xi_\alpha}{\eta_\alpha}, \qquad n \to \infty .$$

- **Goals of self-normalizations:**

1. Avoid knowledge of the quantile normalization a_n.

2. Numerator ξ_α and denominator η_a are dependent. The ratio ξ_α/η_α might have ligther tails than ξ_α (it has moments of order $\alpha - \varepsilon$ only).

3. There might be a "continuous transition" from the finite to the infinite variance case.

- **A classical result by** Logan, Mallows, Rice, Shepp (1973) **for studentized sums: for** $\alpha \in (1, 2)$ **and** $p > \alpha$

$$\frac{S_n - b_n}{\gamma_{n,p}} \xrightarrow{d} R_{\alpha,p} := \frac{\xi_\alpha}{\eta_{\alpha,p}}, \qquad n \to \infty,$$

where

$$\gamma_{n,p} = \Big(\sum_{t=1}^{n} |X_t|^p \Big)^{1/p},$$

- ξ_α **is** α**-stable,**
- $\eta_{\alpha,p}^p$ **is** α/p**-stable,**
- ξ_α **and** $\eta_{\alpha,p}$ **are dependent,**
- **The limit ratio** $R_{\alpha,p}$ **has density whose tails are asymptotically equivalent to those of the Gaussian density.**

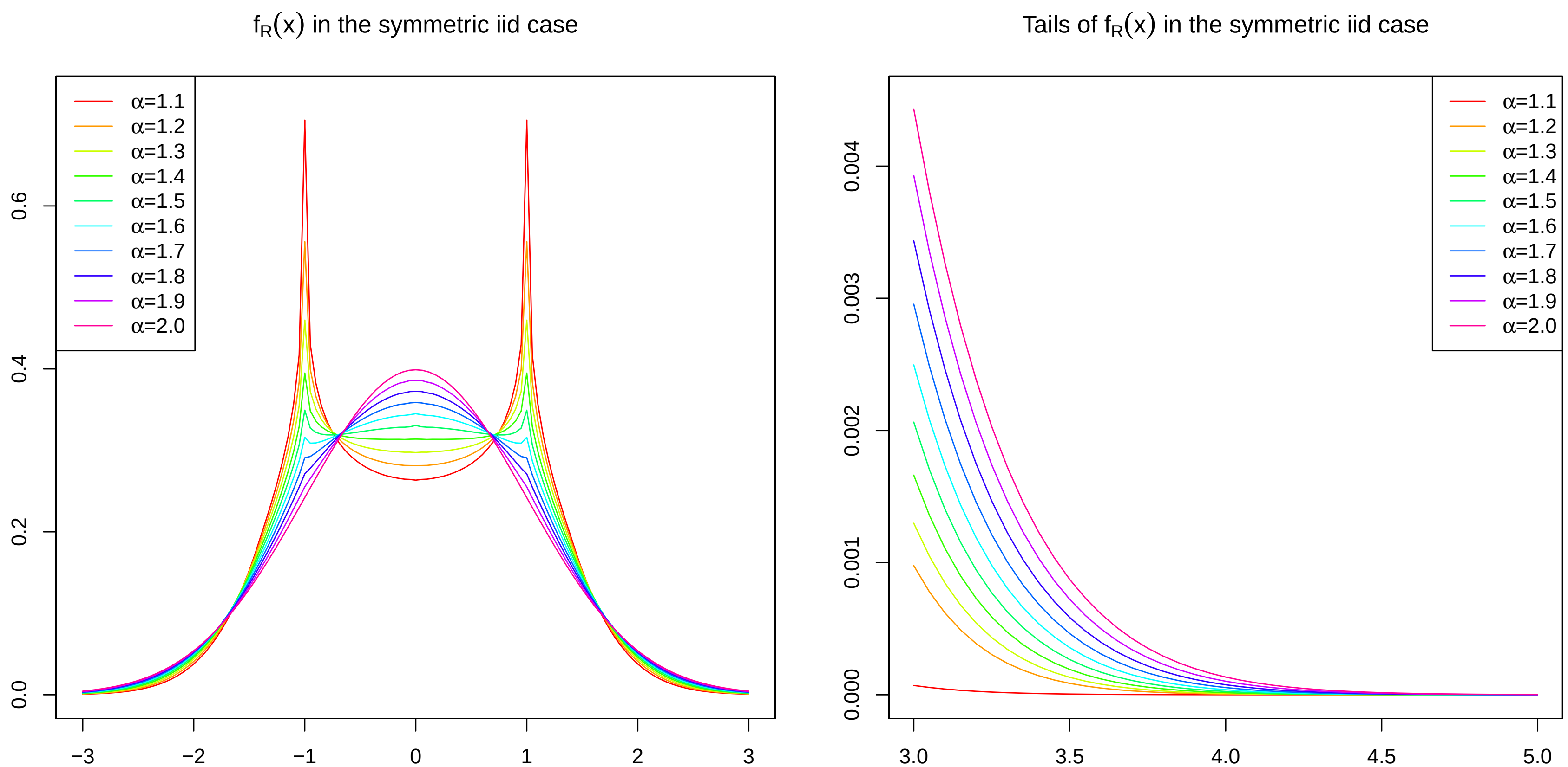


Figure 3. **Left.** Densities $f_{R^X_{\alpha,2}}$ of the studentized sums for iid symmetric X for various $\alpha \in (1,2)$. **Right.** Tail behavior of $f_{R^X_{\alpha,2}}$.

5. Self-normalizations for regularly varying stationary time series

- **Under (weak) mixing and anti-clustering conditions the limits**

$$a_n^{-1}(S_n - b_n, M_n, \gamma_{n,p}) \xrightarrow{d} (\xi_\alpha, \eta_\alpha, \eta_{\alpha,p})$$

exist for $p > \alpha$, $\alpha \in (0, 2)$ and the limiting quantities can be expressed in terms of the spectral tail process.

- **Hence**

$$\frac{S_n - b_n}{M_n} \xrightarrow{d} \frac{\xi_\alpha}{\eta_\alpha}, \qquad \frac{S_n - b_n}{\gamma_{n,p}} \xrightarrow{d} \frac{\xi_\alpha}{\eta_{\alpha,p}}$$

The good news

- **The limit ratios of self-normalized sums have the same distribution as in the iid case (modulo a change of scale)**

The good news

- The limit ratios of self-normalized sums have the same distribution as in the iid case (modulo a change of scale)
- IF and ONLY IF the spectral tail process Θ_t, $t \neq 0$, is deterministic.
- Examples.
  1. Linear processes driven by iid regularly varying noise
  2. Regularly varying stochastic volatility models
  3. Models with $\Theta_t = 0$, $t \neq 0$.

The bad news

- **If Θ_t is random for some $t \neq 0$ then the limit ratios of self-normalized sums have a very complicated random structure.**
- **In some cases one can show that all moments of the limit ratios are finite.**
- **Examples.**
  1. **Solutions to affine stochastic recurrence equations**
  2. **GARCH$(1,1)$ processes**
- **There exist examples where no even moment larger than 2 of the limit ratio is finite.**

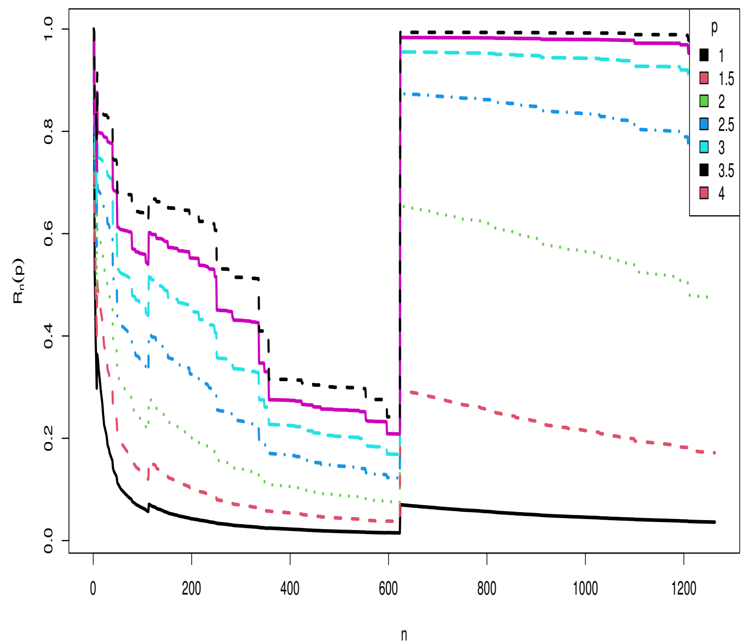
6. Self-normalized quantities can fool you

- **Assume (X_t) regularly varying stationary with index α, $p > \alpha$. Then**

$$R_n(p) = \frac{M_n^{(p)}}{S_n^{(p)}} = \frac{\max_{1\le t\le n} |X_i|^p}{|X_1|^p + \cdots + |X_n|^p} \xrightarrow{d} R(p) \le 1 \quad \text{a.s.}$$

while $R_n(p) \xrightarrow{\text{a.s.}} 0$ for $p < \alpha$.

- **For an AR(1) process $X_t = \varphi X_{t-1} + Z_t$, $|\varphi| < 1$, with iid regularly varying (Z_i) with index α, $R(p) \le 1 - |\varphi|^p$ a.s.**



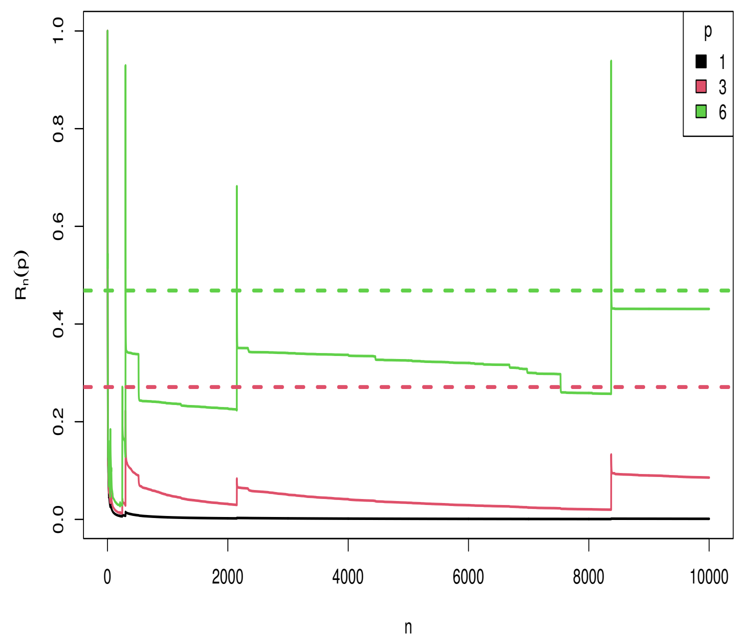


FIGURE 4. **Left.** Plots of the ratio statistics $\boldsymbol{R_n(p)}$ for an iid Pareto$(\mathbf{2})$-distributed $(\boldsymbol{X_t})$ with tail $\mathbb{P}(\boldsymbol{X_t > x}) = \boldsymbol{x^{-2}}$, $\boldsymbol{x > 1}$, hence $\boldsymbol{R_n(p)} \stackrel{\text{a.s.}}{\to} \mathbf{0}$ as $\boldsymbol{n \to \infty}$ if and only if $\boldsymbol{p < 2}$. **Right.** Plots of the ratio statistics $\boldsymbol{R_n(p)}$ for a regularly varying AR(1) process $\boldsymbol{X_t = 0.8X_{t-1} + Z_t}$ with iid Pareto(2)-distributed noise $(\boldsymbol{Z_t})$, hence $\mathbb{E}[|\boldsymbol{X}|^{\boldsymbol{p}}] = \infty$ for $\boldsymbol{p \geq 2}$. For $\boldsymbol{p > 2}$ the support of the limiting random variable $\boldsymbol{R(p)}$ is bounded by $\mathbf{1 - 0.8}^{\boldsymbol{p}}$. The stippled lines indicate this value for $\boldsymbol{p = 3}$ and $\boldsymbol{p = 6}$. One gets the wrong impression that $\boldsymbol{R_n(3)} \stackrel{\text{a.s.}}{\to} \mathbf{0}$ as $\boldsymbol{n \to \infty}$.

7. The extremogram as an alternative to the auto-correlation function

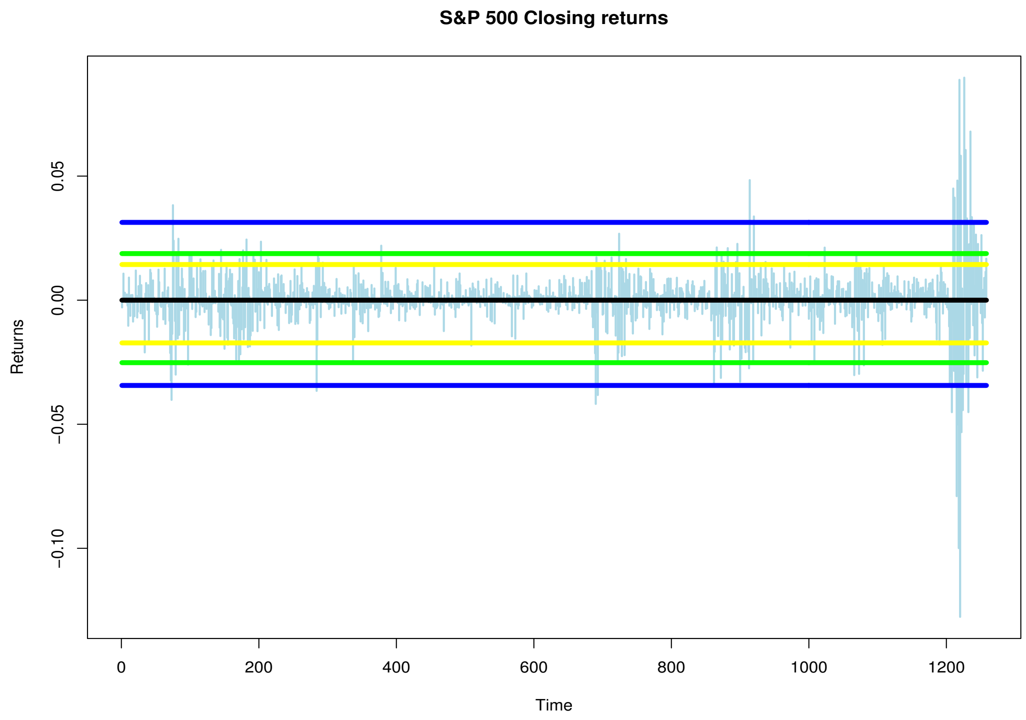


Figure 5. S&P 500 daily return series, 1 May, 2015 – 8 May, 2020. The straight lines indicate the empirical $\boldsymbol{q}$- and $\boldsymbol{(1-q)}$-quantiles of the data for $\boldsymbol{q = 0.01, 0.025, 0.05}$.

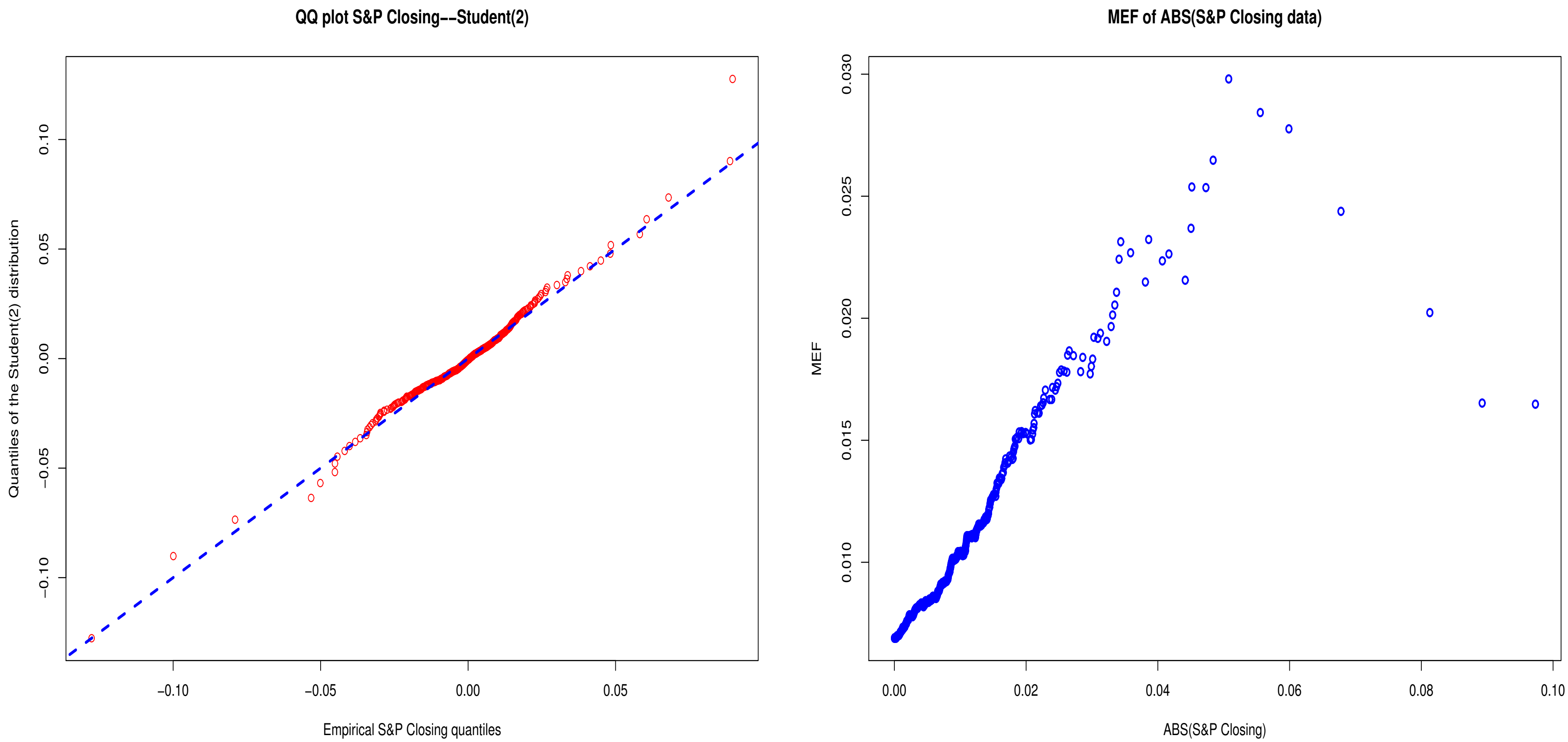


Figure 6. **Left.** *QQ-plot* of the S&P 500 daily closing log-returns against the Student(2) distribution. **Right.** *Mean Excess Plot* of the absolute values: $\mathbb{E}_{\boldsymbol{F}}[|\boldsymbol{X}| - \boldsymbol{u} \mid |\boldsymbol{X}| > \boldsymbol{u}]$, $\boldsymbol{u} > \boldsymbol{0}$.

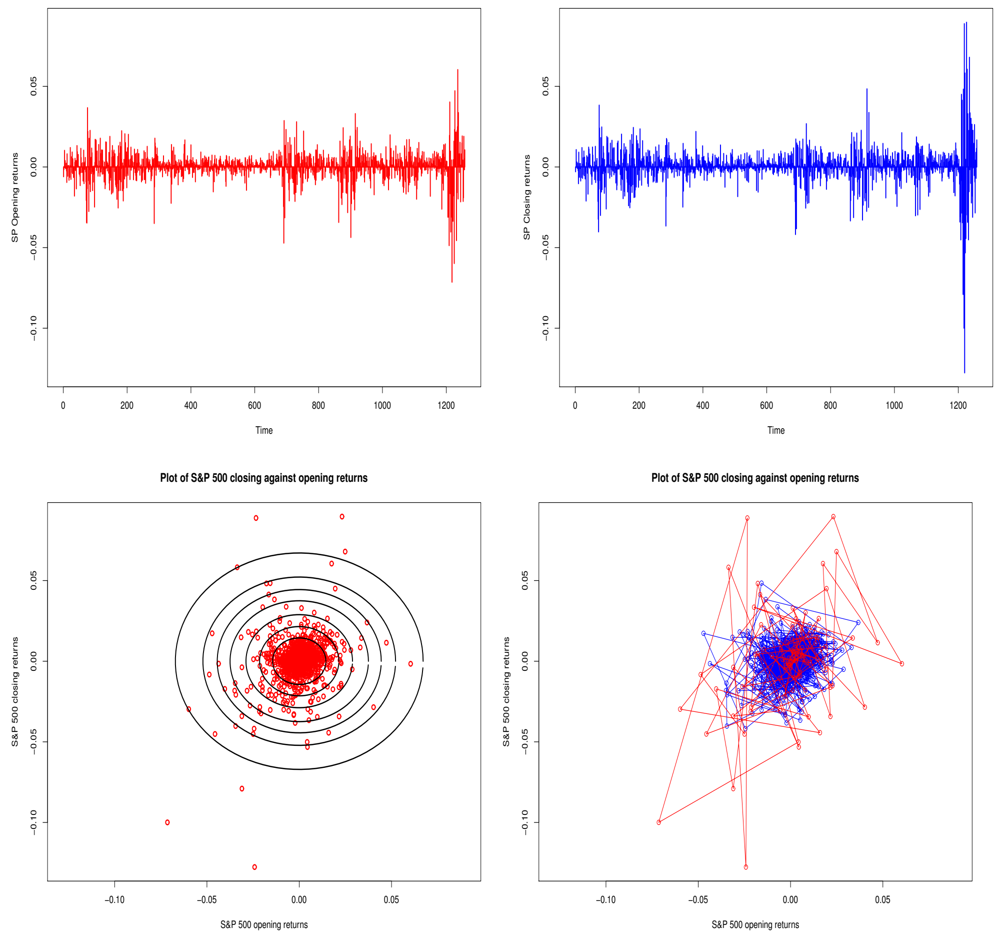


FIGURE 7. **Top:** S&P 500 daily opening (left) and closing (right) log-returns. **Bottom:** Scatterplot closing against opening. Circles indicate 80, 90, 95, 97, 98, 99, 99.5% quantiles of the distances from **0**.

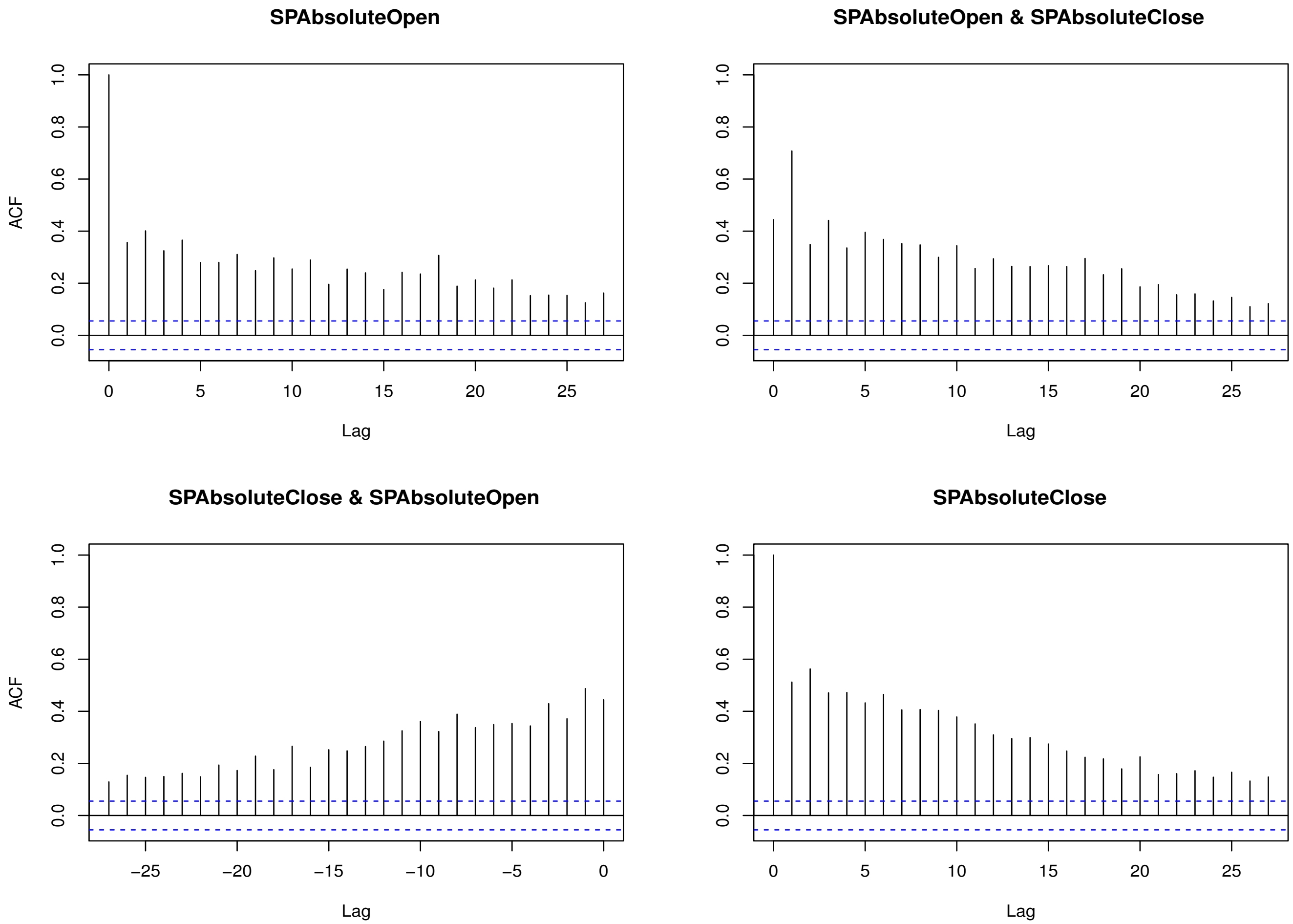


FIGURE 8. Sample auto- and cross-correlations for the corresponding absolute values.

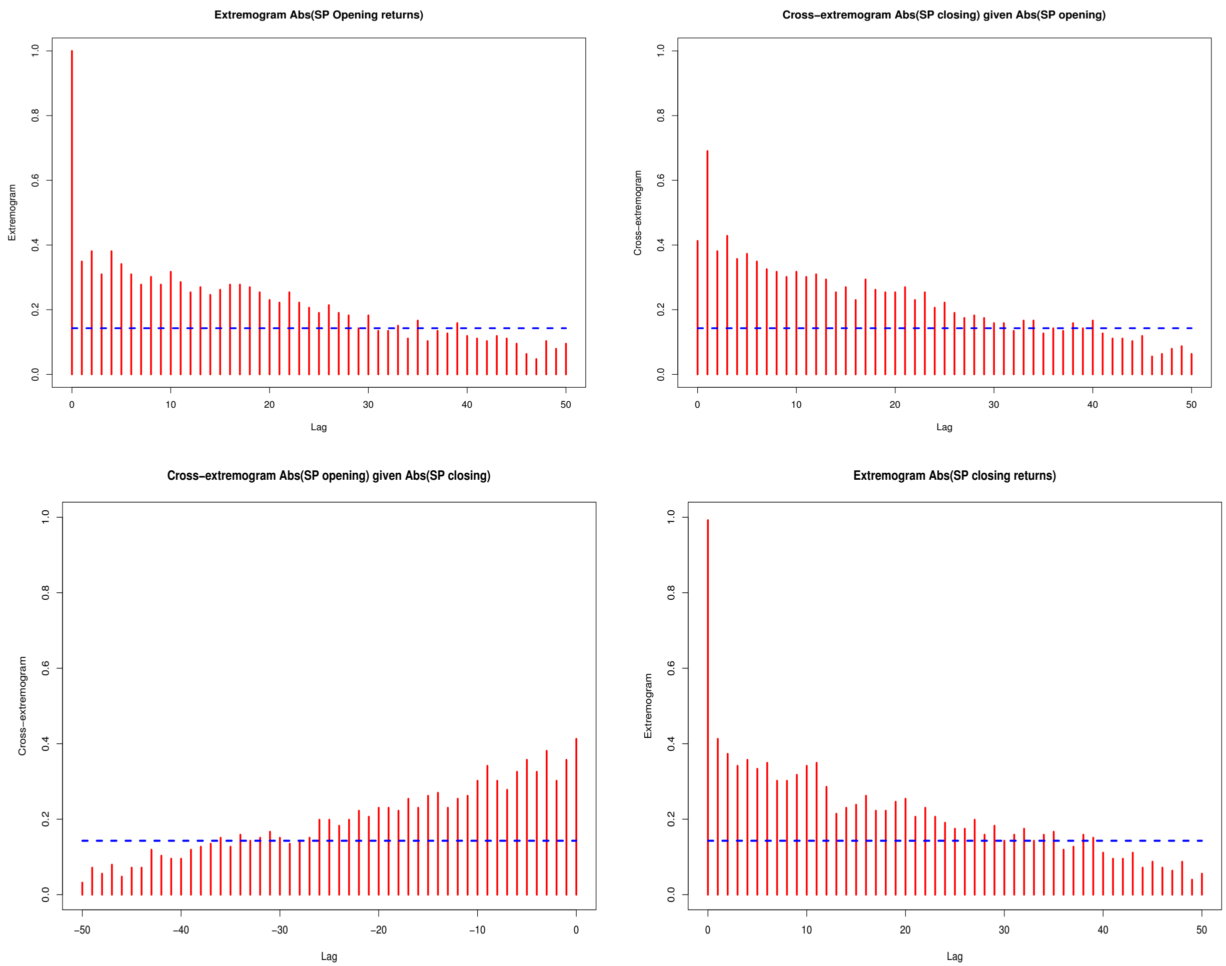


FIGURE 9. Sample extremograms and cross-extremograms of the absolute values. Thresholds are empirical 90%-quantiles.

- **The** ***extremogram*** **of a stationary sequence (X_t):**

$$\lim_{x\to\infty} \mathbb{P}\big(X_h > x \mid X_0 > x\big) = \rho_X(h)\,, \qquad h \geq 0\,.$$

- **The** ***cross-extremogram of (X_t) given Y_0*****:**

$$\lim_{x\to\infty} \mathbb{P}\big(X_h > x \mid Y_0 > x\big) = \rho_{X|Y}(h)\,, \qquad h \geq 0\,.$$

- **The extremogram is approximated by the auto-correlations of the stationary sequence $(1(X_t > x_n))$ for high quantile x_n.**

- The *extremogram* of a stationary sequence (X_t):

$$\lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x) = \rho_X(h)\,, \qquad h \ge 0\,.$$

- The *cross-extremogram of (X_t) given (Y_t)*:

$$\lim_{x\to\infty} \mathbb{P}(X_h > x \mid Y_0 > x) = \rho_{X|Y}(h)\,, \qquad h \ge 0\,.$$

- The extremogram is approximated by the auto-correlations of the stationary sequence $(1(X_t > x_n))$ for high quantile x_n.
- The extremogram is the autocorrelation function of some stationary process IF these limits exist.

- **Extremogram** for non-negative regularly varying X_t: for $h \geq 0$,

$$\lim_{x\to\infty} \mathbb{P}(x^{-1}X_h > 1 \mid X_0 > x) = \mathbb{P}(Y_\alpha \Theta_h > 1) = \mathbb{E}[\Theta_h^\alpha \wedge 1]\,.$$

$\rho_X(h) = 0$ if and only if $\Theta_h = 0$ a.s.
if and only if X_0 and X_h asymptotically independent.

Conclusion

- **Generalized regular variation is natural to deal with dependence in the extreme,**
- **Norms and tail process can be considered independently,**
- **Classical results extend to asymptotic independence,**
- **Asymptotics of standadizations depend on the distribution of the tail process in a complicated way,**
- **The asymptotic of the extremogram is safe because it involves standardized indicators.**

Thank you for your attention!