Second order properties of distribution tails and estimation of tail exponents in random difference equations

Article *in* Extremes · December 2009

DOI: 10.1007/s10687-009-0082-x

Herwig Wendt
University of Toulouse

Patrice Abry
Ecole normale supérieure de Lyon

Extremes
DOI 10.1007/s10687-009-0082-x

Second order properties of distribution tails and estimation of tail exponents in random difference equations

Changryong Baek · Vladas Pipiras ·
Herwig Wendt · Patrice Abry

Received: 11 August 2008 / Revised: 16 February 2009 / Accepted: 16 February 2009

Abstract According to a celebrated result of Kesten (Acta Math 131:207–248, 1973), random difference equations have a power-law distribution tail in the asymptotic sense. Empirical evidence shows that classical estimators of tail exponent of random difference equations, such as Hill estimator, are extremely biased for larger values of tail exponents. It is argued in this work that the bias occurs because the power-tail region is too far in the tail from a practical perspective. This is supported by analysis of a few examples where a stationary distribution of random difference equation is known explicitly, and by proving a weaker form of the so-called second order regular variation of distribution tails of random difference equations, which measures deviations from the asymptotic power tail. The latter, in particular, suggests a specific second order term for a distribution tail. Estimation of tail exponents can be adapted by taking this second order term into account. One such method

The second author is supported in part by the NSF grant DMS-0505628.

C. Baek · V. Pipiras (✉)
Department of Statistics and Operations Research, UNC at Chapel Hill,
CB#3260, Smith Bldg. Chapel Hill, NC 27599, USA
e-mail: pipiras@email.unc.edu

C. Baek
e-mail: crbaek@email.unc.edu

H. Wendt · P. Abry
Laboratoire de Physique, Ecole Normale Supérieure de Lyon,
46, allée d'Italie 69364 Lyon cedex 7, France

H. Wendt
e-mail: herwig.wendt@ens-lyon.fr

P. Abry
e-mail: patrice.abry@ens-lyon.fr

Published online: 17 March 2009

available in the literature is examined, and a new, simple, regression type estimator is proposed. Simulation study shows that the proposed estimator works very well. ARCH models of interest in Finance and multiplicative cascades used in Physics are considered as motivating examples throughout the work. Extension to multidimensional random difference equations with nonnegative entries is also considered.

Keywords Random difference equations · Tail exponent and its estimation · Second order regular variation · ARCH models · Multiplicative cascades

AMS 2000 Subject Classifications Primary—60G70 · 60H25

1 Introduction

We are interested in tail exponents of random difference equations (RDEs, in short), also known as random recurrence equations, autoregressive models with random coefficients. In one dimension, RDE is given by

$$X_n = A_n X_{n-1} + B_n, \quad n \geq 1, \tag{1.1}$$

where (A_n, B_n) are typically assumed to be i.i.d. vectors and X_0 is some starting position. Important examples of RDEs include autoregressive conditionally heteroscedastic (ARCH) processes used in Finance or multiplicative cascades of interest in Physics. Several examples are introduced in detail in Section 2. We shall focus throughout on one dimensional RDE in Eq. 1.1 though the multidimensional case will also be considered (see Section 5.3 below).

Under mild assumptions, the series $\{X_n\}$ in Eq. 1.1 has a stationary solution X satisfying the equation (which we also call RDE)

$$X \stackrel{d}{=} AX + B, \tag{1.2}$$

where $(A, B) =_d (A_1, B_1)$ independent of X, and the tail distribution of X has a power tail. This result was first shown by Kesten (1973) and studied further by many authors, for example, Grintsyavichyus (1981), Goldie (1991) to name a few. It is stated in the following theorem (analogous result for the multidimensional case is given in Theorem 5.1).

Theorem 1.1 (Kesten 1973, Theorem 5) *Let $\{X_n\}_{n\geq 1}$ be defined by Eq. 1.1. Suppose that (A_n, B_n), $n \geq 1$, are i.i.d. random vectors such that*

$$E \log |A_1| < 0, \tag{1.3}$$

and that, for some $\alpha > 0$,

$$E|A_1|^\alpha = 1, \tag{1.4}$$

$$E|A_1|^\alpha \log^+ |A_1| < \infty, \quad 0 < E|B_1|^\alpha < \infty. \tag{1.5}$$

If, in addition, $\log|A_1|$ *does not have a lattice distribution and* B_1 *is not a constant times* $(1 - X_1)$*, then*

$$X_n \xrightarrow{d} X, \tag{1.6}$$

where

$$X \stackrel{d}{=} \sum_{k=1}^{\infty} A_1 \dots A_{k-1} B_k, \tag{1.7}$$

and the series on the right-hand side of Eq. 1.7 converges a.s. Moreover,

$$P(X < -x) \sim c_- x^{-\alpha} \text{ and } P(X > x) \sim c_+ x^{-\alpha}, \quad \text{as } x \to \infty, \tag{1.8}$$

where at least one of c_- *and* c_+ *is nonzero.*

We are interested in questions concerning estimation of the tail exponent α appearing in Eq. 1.8. A common estimation method is based on a Hill estimator (see, for example, Embrechts et al. 1997). If $Y_1, \dots, Y_n$ are n given observations with a common distribution of Y (independent or not, depending on the context) and

$$Y_{(1)} \geq Y_{(2)} \geq \dots \geq Y_{(n)}$$

are the observations in the decreasing order, the Hill estimator is defined as

$$\widehat{\alpha}_H = \left(\frac{1}{k} \sum_{i=1}^{k} (\log Y_{(i)} - \log Y_{(k+1)}) \right)^{-1}, \tag{1.9}$$

where k is a threshold. If the underlying distribution of Y has a power tail

$$P(Y > y) \sim c y^{-\alpha}, \quad \text{as } y \to \infty, \ \alpha > 0, \tag{1.10}$$

the Hill estimator in Eq. 1.9 of α is known to have nice theoretical properties such as consistency, asymptotic normality under fairly mild assumptions. In practice, the presence of heavy tails is assessed by examining the so-called Hill plot. This plot is produced by plotting $\widehat{\alpha}_H$ as a function of threshold k (from the

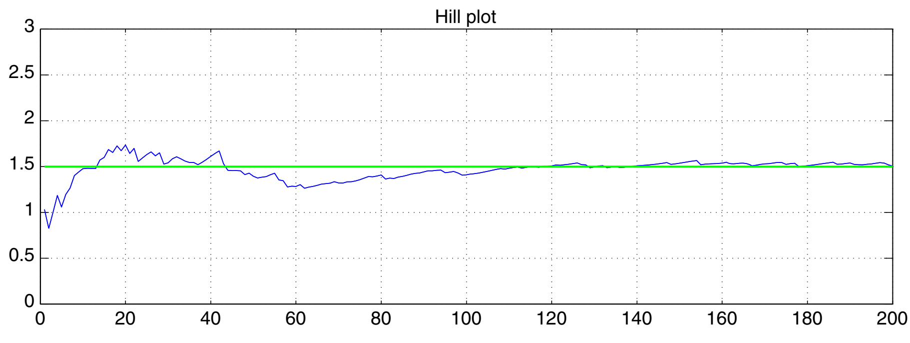


Fig. 1 Hill estimator from Pareto distribution with $\alpha = 1.5$.

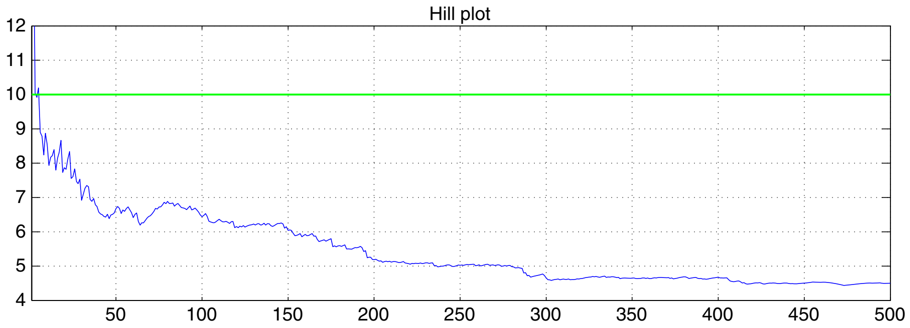


Fig. 2 Hill plot for 5000 independent realizations of ARCH(1) model with $\alpha = 10$.

smallest k to larger k). An example is given in Fig. 1 based on i.i.d. observations from the Pareto distribution with $\alpha = 1.5$. If the distribution tail has power law, as in Fig. 1, the Hill plot levels off in a region of small k and that level is taken as the Hill estimate for the power-tail exponent.

Statement of the problem When tail exponent α is *large*, Hill plots for RDEs show tail exponent estimates surprisingly biased. For example, Fig. 2 shows a typical Hill plot for an ARCH(1) model (see Section 2.1) having tail exponent $\alpha = 10$, based on 5,000 independent realizations of the process at a chosen, fixed time. Observe from the figure how far the Hill plot is from the true value of α. Perhaps even more surprising is that estimation improves only slightly by taking any reasonably larger sample size. For example, Fig. 3 also shows the Hill plot for a million independent realizations. The basic goal of this work is to understand why above estimation of tail exponents fails in RDEs for larger values of exponents and how this can be remedied.

Here are some further important comments about the above problem.

- *Why do we consider RDEs?* The problem described above seems to be characteristic to all RDEs. We illustrate this in Section 2 through

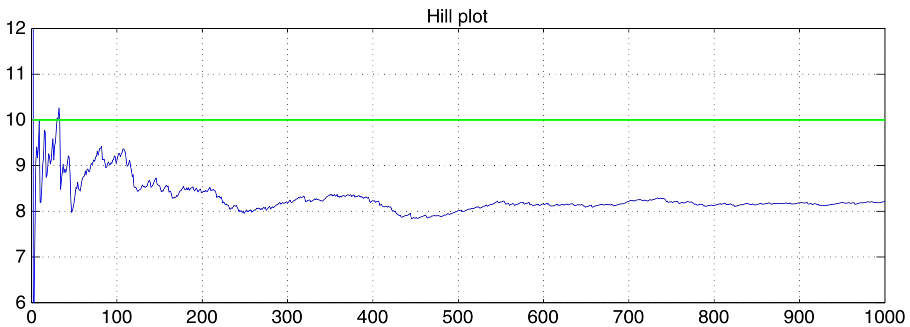


Fig. 3 Hill plot for a million independent realizations of ARCH(1) model with $\alpha = 10$.

simulations in a number of different RDE models. One of our main goals is also to explain this in theory.

- *Why should one care about larger tail exponent?* The problem described above and supporting simulations involve larger values of tail exponent α. It is important to ask then why one should care about larger α. Several points should be made in this regard. First, in some applications of RDE models, larger values of α are, in fact, expected. This is the case, for example, with multiplicative cascades and other so-called multifractal models of interest in Physics. Second, observed bias in estimation of α becomes larger with increasing α and is still present (though smaller) for smaller α. If one believes that RDE models are appropriate for data at hand, this should be taken into account for either larger or smaller tail exponent α. Moreover, it would be desirable to have an estimation method that takes into account the possibilities of bias, and which performs well for both larger and smaller values of α. Proposed estimation methods will be discussed in Section 4.
- *Independent versus dependent observations.* Two types of observations can be considered in regard to the problem stated above. First, one may suppose that observations are obtained from the RDE in Eq. 1.1 and hence dependent in time. Second, one may suppose given independent copies of X_N for large fixed N (which can be thought as independent copies of the stationary solution X). For simplicity, we shall focus throughout on the second case. Perhaps surprising but this case is also relevant in practice (for example, in the context of multiplicative cascades) and the problem stated above is as equally relevant. Moreover, for dependent data given by Eq. 1.1, tail exponent estimation problems are known to get only worse. See Section 5.2 for related discussion.

Possible explanations for the problem Since we have removed temporal dependence from observations, two explanations seem plausible for the above problem:

1. Convergence to stationary solution in Eq. 1.2 is so slow that the observation X_N is still far from the stationary solution X.
2. The result in Eq. 1.8 is asymptotic in nature. It can happen that the region where Eq. 1.8 actually happens is too far in the tail to be observed for practical purposes. In other words, even with a huge number of data points, there are significant deviations from the Pareto tail in practice.

In fact, in the context of RDEs, using their Markov structure, one expects that underlying measure P_N induced by RDE in Eq. 1.1 converges to its invariant measure P_∞ induced by Eq. 1.2 at a geometrically fast rate. The latter fact is known as geometric ergodicity. Basrak et al. (2002b) and Stelzer (2009) show detailed description of geometric ergodicity of the RDE $\{X_n\}$. We also summarize their result in Theorem 5.3 in Appendix A to the reader's convenience.

Theoretical properties of tail distribution We therefore suspect that the tail exponent is not observed because of the second explanation above. As clearly pointed out in Resnick (1997), all nice theoretical properties of Hill and other related estimators are valid only when the underlying distribution is close to Pareto distribution. If the underlying distribution deviates from Pareto distribution, bias is inevitable.

There is a general theoretical framework, called second order regular variation or 2RV, in short (see, for example, de Haan and Ferreira 2006), that allows one to study bias in first order regular variation such as Eq. 1.8. The tail distribution $\overline{F}(x) = P(X > x)$ is second order regularly varying with first order parameter $\alpha > 0$ and second order parameter $\rho < 0$ (denoted as $\overline{F} \in 2RV(-\alpha, \rho)$) if there exists a function $G(x) \to 0$ as $x \to \infty$ which ultimately has constant sign such that, for any $a > 0$,

$$\lim_{x\to\infty} \frac{\frac{\overline{F}(ax)}{\overline{F}(x)} - a^{-\alpha}}{G(x)} = ca^{-\alpha}\frac{a^{\rho} - 1}{\rho}, \tag{1.11}$$

for some constant $c \neq 0$. Second order regular variation can be thought as

$$\overline{F}(x) - c_1 x^{-\alpha} \sim l_2(x) x^{-\alpha+\rho}, \tag{1.12}$$

where $l_2(x)$ is a slowly varying function at infinity (Bingham et al. 1989, Thoerem 3.6.6, p. 158). The right-hand side of Eq. 1.12 is thought as a bias. For practical (estimation) questions, the slowly varying function in Eq. 1.12 is taken as

$$l_2(x) = c_2$$

for some constant c_2, that is,

$$\overline{F}(x) - c_1 x^{-\alpha} \sim c_2 x^{-\alpha+\rho}. \tag{1.13}$$

Remark It is important to note that 2RV is asymptotic in nature. Even if proved for RDEs, it does not yield the exact region where Eq. 1.12 holds. Hence, without further analysis, establishing 2RV, in principle, does not completely address the problem raised in this work. Despite these limitations, 2RV at least indicates that there exists a bias and that it should be taken into account, for example, in questions of estimation.

To the best of our knowledge and understanding, 2RV is still an open and difficult problem for any larger class of RDEs. Instead of trying to prove 2RV, we shall focus on its weaker forms by considering the asymptotics of

$$P(X > x) - P(AX > x), \tag{1.14}$$

where A is the multiplier appearing in Eq. 1.2, and more specifically that of

$$\int_x^{\infty} (P(X > u) - P(AX > u))\, du. \tag{1.15}$$

We will show under mild assumptions that, as $x \to \infty$,

$$\int_x^\infty (P(X > u) - P(AX > u))\,du \sim cx^{-\alpha}, \tag{1.16}$$

which also suggests that, as $x \to \infty$,

$$P(X > x) - P(AX > x) \sim c\alpha x^{-\alpha-1}. \tag{1.17}$$

The expressions in Eqs. 1.14 and 1.15 are much easier to consider than Eq. 1.11 because AX can be related back to X by using the RDE in Eq. 1.2. In fact, as seen from Section 3, Eq. 1.16 follows just by using the RDE structure in Eq. 1.2 and the Kesten's result itself. A particularly simple case of RDE is considered in the beginning of Section 3.

How is Eq. 1.17 related to 2RV in Eq. 1.11? Observe that with $g(x) = -x^\alpha G(x)\overline{F}(x)$, Eq. 1.11 can be rewritten as

$$\lim_{x\to\infty} \frac{x^\alpha P(X > x) - (ax)^\alpha P(X > ax)}{g(x)} = c\frac{a^\rho - 1}{\rho}. \tag{1.18}$$

Equation 1.17, on the other hand, can be expressed as

$$\lim_{x\to\infty} \frac{x^\alpha P(X > x) - x^\alpha (EA^\alpha)^{-1} P(X > A^{-1}x)}{x^{-1}} = c\alpha, \tag{1.19}$$

since $EA^\alpha = 1$ by Eq. 1.4. Hence, Eq. 1.19 can be viewed as 2RV in Eq. 1.18 at random $a = 1/A$.

From a practical perspective, Eq. 1.17 says that there is a bias in Eq. 1.8 (if there is no bias in Eq. 1.8, then $P(X > x) - P(AX > x) = 0$). Moreover, if one believes that $\overline{F}(x)$ satisfies Eq. 1.13, then necessarily $\rho = -1$ and

$$P(X > x) - c_1 x^{-\alpha} \sim c_2 x^{-\alpha-1}, \tag{1.20}$$

as $x \to \infty$ (Proposition 3.1 below).

Discussion on estimation If one believes in RDE model and that the model has 2RV, it is natural to estimate tail exponent by taking 2RV into account. Tail exponent estimation based on 2RV has been studied by a number of authors. In particular, Peng (1998) shows asymptotic bias of Hill estimator under 2RV and proposes linear estimator considering second order parameter to adjust for the asymptotic bias. More recently, Gomes and Rodrigues (2008) consider weighted Hill estimator where the weights are determined by 2RV parameters. Feuerverger and Hall (1999) utilize normalized log-spacings of order-statistics

$$i(\log Y_{(i)} - \log Y_{(i+1)}),$$

which are known to follow Exponential distribution with mean 1 by Rényi's representation theorem for order statistics. Under the relation in Eq. 1.13, these authors derive the maximum likelihood estimators of parameters α, ρ.

We will apply the approach of Feuerverger and Hall (1999) to estimate tail exponent in RDE in Section 4. In that section, we also propose a simple estimator of linear regression type based on Eq. 1.20. By taking log-transformation, observe that Eq. 1.20 becomes

$$\log\left(\overline{F}(x)\right) \approx \log c_1 - \alpha \log x + \log\left(1 + c_2/c_1 x^{-1}\right) \approx \log c_1 - \alpha \log x + c_2/c_1 x^{-1}. \tag{1.21}$$

Equation 1.21 suggests that regressing the logarithm of empirical tail distribution on $(1, \log x, 1/x)$ gives the least squares estimator of tail exponent α. This is a generalization of the regression

$$\log\left(\overline{F}(x)\right) \approx \log c_1 - \alpha \log x,$$

which considers only the first order regular variation. Simulation study shows that our proposed estimator performs very well. Section 4 only touches upon estimation questions under the framework in Eq. 1.20. A much more detailed study of these questions can be found in Baek and Pipiras (2009).

The rest of this work is organized in the following way. Several examples of RDEs and some simulation studies with unobservable exponents for RDEs are introduced in Section 2. In Section 3, we prove the weaker form in Eq. 1.16 of second order regular variation in RDEs. In Section 4, we introduce simple least squares estimator of tail exponent and study performance through simulations. All these sections concern one dimensional RDEs. Some further issues including a multidimensional extension are discussed in Section 5.

2 Examples of RDEs and simulation study

In this section, several examples of one dimensional RDEs in Eq. 1.1 are given. These include autoregressive conditionally heteroscedastic processes of order 1, an example with an explicit stationary distribution and multiplicative cascades. We also report here further simulation study supporting the statement of the problem discussed in Section 1. The simulations are to show that the problem seems prevalent for all RDEs.

2.1 ARCH(1) models

A particular example of RDEs is a popular autoregressive conditionally heteroscedastic (ARCH(1)) model of order 1, defined by

$$\xi_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \beta + \lambda \xi_{t-1}^2, \tag{2.1}$$

where $\{\epsilon_t\}$ are i.i.d. $\mathcal{N}(0, \sigma^2)$ random variables and coefficients β and λ are strictly positive. The squares of ARCH(1) series can be written as

$$\xi_t^2 = \lambda \epsilon_t^2 \xi_{t-1}^2 + \beta \epsilon_t^2, \tag{2.2}$$

which is the RDE in Eq. 1.1 with

$$X_t = \xi_t^2, \quad A_t = \lambda \epsilon_t^2, \quad B_t = \beta \epsilon_t^2. \tag{2.3}$$

By Theorem 1.1, the tail exponent $\kappa = \alpha/2$ of ξ^2 is the solution of

$$\Gamma(\kappa + 1/2) = \sqrt{\pi}(2\sigma^2\lambda)^{-\kappa}. \tag{2.4}$$

Equation 2.4 does not have a closed-form solution. For example, if $\sigma^2 = 1$, numerical calculations yield:

κ	2	3	4	5	6	7	8	9	10
λ	.577	.406	.312	.254	.214	.185	.163	.145	.105.

Note also that, by symmetry, the tail exponent of ξ is $\alpha = 2\kappa$ because

$$P(\xi > x) = 1/2P(\xi^2 > x^2) \sim c/2(x^2)^{-\kappa} = c/2x^{-2\kappa}.$$

Consider the ARCH(1) series in Eq. 2.1 with $\lambda = .254$, $\beta = 1$, $\epsilon_t =_d \mathcal{N}(0, 1)$. From the above table, the tail exponent is

$$\alpha = 10.$$

In this simulation, we generated $R = 5,000$ independent samples with $N = 5,000$ iterations of ARCH(1) series. Figure 4 shows tail exponent estimation, and again it does not find the true tail exponent $\alpha = 10$.

Remark Empirical observations for ARCH(1) models similar to those above can also be found in Beirlant et al. (1999) (see Figure 12 on p. 195). Though this was the only place in the literature that we found to make such observations.

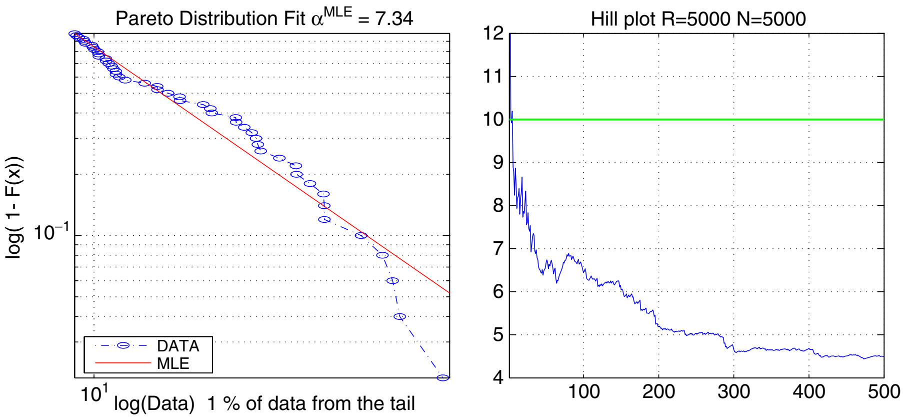


Fig. 4 ARCH(1) with $\alpha = 10$.

2.2 Examples of RDEs with explicit power-tail distributions

The following appears to be the only known family of RDEs for which a stationary solution has a power-tail distribution in closed form. Consider the so-called beta prime distribution $\beta(a, b)$ given by the density

$$\frac{1}{B(a,b)} x^{a-1}(1+x)^{-a-b} 1_{\{x>0\}}, \quad a, b > 0. \tag{2.5}$$

Simple algebra shows that

$$\beta(a,b) \stackrel{d}{=} \frac{1}{Z} - 1, \tag{2.6}$$

where $Z =_d B(b, a)$ follows the beta distribution. Now fix $k \in \mathbb{N}$ and let $a_1, \ldots, a_k, b$ be positive reals. Denote $a_{k+1} = a_1$ and set

$$A = Y_1 \ldots Y_k, \quad B = Y_1 \ldots Y_k + \ldots + Y_{k-1}Y_k + Y_k, \tag{2.7}$$

where $Y_k =_d \beta(a_{j+1}, a_j + b)$ for $j = 1, \ldots, k$. Then, as shown in Goldie (1991), Chamayou and Letac (1991),

$$X \stackrel{d}{=} \beta(a_1, b) \tag{2.8}$$

satisfies RDE in Eq. 1.2 with A, B in Eq. 2.7.

Observe that the distribution tail of random variable X in Eq. 2.8 satisfies

$$\begin{aligned} P(X > x) &= \int_x^\infty \frac{1}{B(a_1, b)} u^{a_1-1}(1+u)^{-a_1-b} du \\ &= \frac{1}{B(a_1, b)} \int_x^\infty u^{-b-1} \left(1 + \frac{1}{u}\right)^{-a_1-b} du \\ &= \frac{1}{B(a_1, b)} \int_x^\infty u^{-b-1} \left(1 - (a_1+b)u^{-1} + o(u^{-1})\right) du. \end{aligned} \tag{2.9}$$

Therefore, we have

$$P(X > x) - \frac{1}{B(a_1, b)} \frac{1}{b} x^{-b} \sim \frac{-(a_1+b)}{B(a_1, b)} \int_x^\infty u^{-b-2} du = \frac{-(a_1+b)}{B(a_1, b)(b+1)} x^{-b-1}, \tag{2.10}$$

as $x \to \infty$, that is, the tail exponent for X is b and the second order term has the exponent $b + 1$.

Consider the above example with $k = 2$, $a_1 = a_2 = 1$ and $b = 9$. The tail exponent in this case is $b = 9$. The simulations here are based on 5,000 independent realizations of X_N with $N = 100$ iterations. Figure 5 shows tail exponent estimation with asymptotic 95% confidence interval of Hill estimator in dotted line. In addition, Fig. 6 shows tails of empirical and theoretical distributions over the range of data. The theoretical power tail is also plotted and it is seen from Fig. 6 that the theoretical tail is not yet in the asymptotic range in Eq. 1.8.

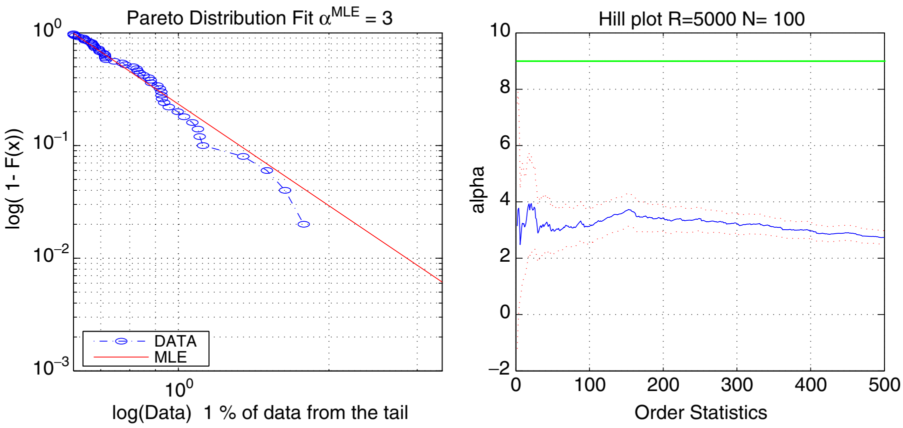


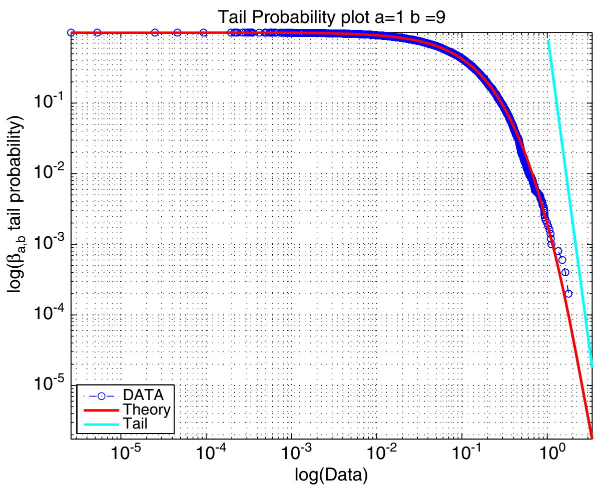
Fig. 5 Explicit example of Section 2.2 with $a_1 = a_2 = 1$ and $b = 9$. True exponent $\alpha = 9$.

The fact that the asymptotic range Eq. 1.8 is not observed here with data, can be explained in theory because the corresponding distribution has a closed form. Using Eq. 2.10 and comparing the ratio of distribution tail and the power law x^{-b}, suppose that

$$\left| \frac{P(X > x)}{x^{-b}/bB(a_1, b)} - 1 \right| = \left| \frac{P(Z < 1/(1+x))}{x^{-b}/bB(a_1, b)} - 1 \right| \leq \epsilon, \qquad (2.11)$$

where $Z =_d B(b, a_1)$ from the relation in Eq. 2.6. For example, with $a_1 = 1, b = 9$ and $\epsilon = .5$, numerical computations show that Eq. 2.11 holds for $x > 12.49$. The probability of having observation greater than 12.49 is 6.76×10^{-11}.

Fig. 6 Explicit example of Section 2.2 with $a_1 = a_2 = 1$ and $b = 9$. Empirical tail distribution stays far below the asymptotic tail behavior.



2.3 Multiplicative cascades

Let $T = [0, 1)$. For $k_i \in \{0, 1\}$, $i \geq 1$, denote

$$I_{k_1,\ldots,k_n} = \left[\frac{l}{2^n}, \frac{l+1}{2^n}\right), \quad l = k_1 2^0 + \ldots + k_n 2^{n-1}, \tag{2.12}$$

subintervals of T obtained by splitting in a dyadic fashion. Let also $\{W_{k_1,\ldots,k_i}, k_i \in \{0, 1\}, i \geq 1\}$ be a family of i.i.d., nonnegative, mean 1 random variables, called multipliers. Define a random measure λ_n on $\mathcal{B}(T)$ by

$$\lambda_n(E) = \int_E f_n(t)dt, \quad \text{with } f_n(t) = \sum_{k_i \in \{0,1\}} \left(\prod_{i=1}^n W_{k_1,\ldots,k_i}\right) 1_{I_{k_1,\ldots,k_n}}(t). \tag{2.13}$$

Note, in particular, that

$$\lambda_n(I_{k_1,\ldots,k_n}) = 2^{-n} \prod_{i=1}^n W_{k_1,\ldots,k_i}. \tag{2.14}$$

(For example, $\lambda_1[0, 1/2) = 2^{-1} W_0$, $\lambda_3([1/8, 2/8)) = 2^{-3} W_0 W_{0,0} W_{0,0,1}$ and so on.) Provided $E(W \log_2 W) < 1$, one can show that the sequence λ_n converges weakly to a random measure λ_∞ on $\mathcal{B}(T)$ almost surely, that is,

$$\lambda_n \Rightarrow \lambda, \text{ on } \mathcal{B}(T) \text{ a.s.} \tag{2.15}$$

where $\Rightarrow$ indicates weak convergence. The limiting random measure λ_∞ is known as a multiplicative cascade (MC, in short). See, for example, Mandelbort (1974), Ossiander and Waymire (2000). The following theorem is a well-known fact about the existence of moments of λ_∞ and related results. Let

$$\chi_2(h) = \log_2 E(W^h 1_{\{W>0\}}) - (h-1) \tag{2.16}$$

be the so-called structure function associated with a multiplier W.

Theorem 2.1 (Kahane and Peyrière 1976, Guivarc'h 1990) *The following statements hold:*

i) $E\lambda_\infty(T) = 1$ *iff* $\chi_2'(1-) < 0$.
ii) $E(\lambda_\infty(T))^h < \infty$ *for* $0 \leq h \leq 1$ *and if*

$$\alpha := \sup\{h \geq 1 : \chi_2(h) \leq 0\} > 1, \tag{2.17}$$

then $E\lambda_\infty^h(T) < \infty$ *for* $1 < h < \alpha$.
iii) *Furthermore, if the cascade (multiplier) is non-lattice, then*

$$P(\lambda_\infty(T) > x) \sim cx^{-\alpha}, \quad \text{as } x \to \infty. \tag{2.18}$$

The tail behavior in Eq. 2.18 of interest here can be proved by using Theorem 1.1 in the following way. Denote

$$M_n = \lambda_n[0, 1), \quad M = \lambda_\infty[0, 1). \tag{2.19}$$

By "separating" the multipliers W_1 and W_2 at the first generation, one can see that

$$M_n \stackrel{d}{=} \frac{W_1}{2} M_{n-1}^{(1)} + \frac{W_2}{2} M_{n-1}^{(2)}, \quad n \geq 1, \tag{2.20}$$

where W_1, W_2, $M_{n-1}^{(1)}$, $M_{n-1}^{(2)}$ are all independent, $M_{n-1}^{(1)} =_d M_{n-1}^{(2)} =_d M_{n-1}$, $W_1 =_d W_2 =_d W$ (the general multipliers in Eq. 2.14), and by convention, $M_0 \equiv 1$. Similarly, the limiting measure satisfies the equation

$$M \stackrel{d}{=} \frac{W_1}{2} M^{(1)} + \frac{W_2}{2} M^{(2)}. \tag{2.21}$$

Equation 2.20 resembles the RDE in Eq. 1.1 when considered in distribution by setting $A_n = W_1/2$, $B_n = W_2 M_{(n-1)}^{(2)}/2$. The key difference is that in Eq. 1.1, it is supposed that (A_n, B_n) are i.i.d. vectors which is not the case for MC because the distribution of $B_n = W_2 M_{n-1}^{(2)}/2$ depends on n. Equation 2.21, on the other hand, can be thought as a special case of RDE in Eq. 1.2.

Though Eq. 2.21 is RDE, Theorem 1.1 cannot be applied directly to it. Indeed, in view of Eqs. 2.21 and 1.2, supposing $A =_d W_1/2$, $B =_d MW_2/2$, the assumption in Eq. 1.5 requires that

$$E|B|^\alpha = E\left|\frac{W}{2}\right|^\alpha E|M|^\alpha < \infty. \tag{2.22}$$

But one expects M to have the tail exponent α and hence one cannot expect that Eq. 2.22 is satisfied. Despite this, however, there is still a way that Theorem 1.1 can be applied to obtain the tail behavior of M. The trick can be found in Guivarc'h (1990), Liu (2000) and others (though, seems to be originally due to Guivarc'h 1990).

The basic idea is as follows. Let $\widetilde{M}$ be a random variable with distribution $P_{\widetilde{M}}(dx) = xP_M(dx)$. Note that $EM = 1$, so $xP_M(dx)$ is a probability measure. Equation 2.21 can be rewritten in terms of characteristic functions as

$$\phi(t) = E(e^{itM}) = E\left(e^{it(A_1 M^{(1)} + A_2 M^{(2)})}\right) = \left(E\left(\phi(A_1 t)\right)\right)^2 \tag{2.23}$$

(for the shortness of notation, we denote $A_i = W_i/2, i = 1, 2$). Consider the random vector $(\widetilde{A}, \widetilde{B})$, independent of $\widetilde{M}$, with the distribution given by

$$\begin{aligned} Eh(\widetilde{A}, \widetilde{B}) &= E\left(A_1 h(A_1, A_2 M^{(2)}) + A_2 h(A_2, A_1 M^{(1)})\right) \\ &= 2E\left(A_1 h(A_1, A_2 M)\right). \end{aligned} \tag{2.24}$$

Note also that the characteristic function of $\widetilde{M}$ is

$$\widetilde{\phi}(t) = E(e^{it\widetilde{M}}) = \int e^{itx} x P_M(dx) = E(Me^{itM}) = -i\phi'(t), \tag{2.25}$$

where $\phi'(t)$ is the derivative of $\phi(t)$.

Observe that

$$E\left(e^{it(\widetilde{A}\widetilde{M}+\widetilde{B})}\right) = E\left(e^{it\widetilde{B}}e^{it\widetilde{A}\widetilde{M}}\right) = E\left(e^{it\widetilde{B}}\widetilde{\phi}(\widetilde{A}t)\right).$$

By applying Eq. 2.24 with $h(a,b) = e^{itb}\widetilde{\phi}(at)$ and Eq. 2.25, it becomes

$$\begin{aligned} 2E\left(A_1 e^{itA_2M}\widetilde{\phi}(A_1t)\right) &= 2E\left(A_1\widetilde{\phi}(A_1t)\right)E\left(e^{itA_2M}\right) \\ &= -2iE\left(A_1\phi'(A_1t)\right)E\left(\phi(A_2t)\right). \end{aligned} \tag{2.26}$$

Note that differentiating the right-hand side of Eq. 2.23 gives

$$\phi'(t) = 2E\left(\phi'(A_1t)A_1\right)E\left(\phi(A_1t)\right).$$

Therefore, one can conclude that

$$E(e^{it\widetilde{M}}) = E\left(e^{it(\widetilde{A}\widetilde{M}+\widetilde{B})}\right)$$

or

$$\widetilde{M} \overset{d}{=} \widetilde{A}\widetilde{M} + \widetilde{B}, \tag{2.27}$$

where $(\widetilde{A}, \widetilde{B})$ is independent of $\widetilde{M}$.

Now, consider the solution $\widetilde{M}$ of Eq. 2.27 instead of M of Eq. 2.21. The conditions in Theorem 1.1 become

$$\begin{aligned} E|\widetilde{A}|^{p-1} &= 2E|A_1|^p = 2^{\chi_2(p)}, \\ E|\widetilde{B}|^{p-1} &= 2EA_1|A_2M|^{p-1} = 2EA_1E|A_2|^{p-1}E|M|^{p-1}. \end{aligned}$$

Since $\chi_2(\alpha) = 0$ by Eq. 2.17 and $E|M|^{\alpha-1} < \infty$ is expected, RDE in Eq. 2.27 should now have a stationary solution with the tail exponent $(\alpha - 1)$.

Establishing the tail behavior of $\widetilde{M}$ leads naturally to that of M. Observe that

$$P(M > x) = \int_x^\infty P_M(dy) = \int_x^\infty \frac{1}{y} y P_M(dy) = \int_x^\infty \frac{1}{y} P_{\widetilde{M}}(dy).$$

Integration by part gives,

$$P(M > x) = -\frac{1}{x}P_{\widetilde{M}}(x) + \int_x^\infty y^{-2}P_{\widetilde{M}}(y)dy,$$

where $P_{\widetilde{M}}(y) = P(\widetilde{M} \le y)$, or

$$xP(M > x) = P(\widetilde{M} > x) - x\int_x^\infty y^{-2}P(\widetilde{M} > y)dy. \tag{2.28}$$

Rewriting Eq. 2.28 gives

$$\frac{x^{\alpha} P(M > x)}{x^{\alpha-1} P(\widetilde{M} > x)} = 1 - \frac{x^{\alpha} \int_x^{\infty} y^{-2} P(\widetilde{M} > y) dy}{x^{\alpha-1} P(\widetilde{M} > x)}. \tag{2.29}$$

Since the tail of random variable $\widetilde{M}$ is expected as

$$P(\widetilde{M} > x) \sim \widetilde{c} x^{-(\alpha-1)},$$

for some positive constant $\widetilde{c}$, the right-hand side of Eq. 2.29 converges to $1 - 1/\alpha$, as $x \to \infty$. This yields

$$P(M > x) \sim c x^{\alpha},$$

where $c = \widetilde{c}(\alpha - 1)/\alpha$.

To illustrate our problem through simulations, consider the case of multiplicative cascade with log-normal multipliers $LN(-\sigma^2/2, \sigma^2)$ where the latter choice of parameters ensures mean 1. Simulations are based on i.i.d. copies of M_N^r, $r = 1, \ldots, R$, where $N = 13$ and $R = 1,000$. The parameter is taken as $\sigma^2 = .2 \log 2$. According to a small calculation found in Appendix B, the corresponding tail exponent is given by

$$\alpha = \frac{2 \log 2}{\sigma^2} = \frac{2 \log 2}{.2 \log 2} = 10. \tag{2.30}$$

Figure 7 shows the corresponding tail distribution plot with Pareto distribution fit (left) and Hill plot (right). The tail appears power-law from the tail plot. However, as seen from the Hill plot, theoretical tail exponent in Eq. 2.30 is far from any reasonable estimate of the tail.

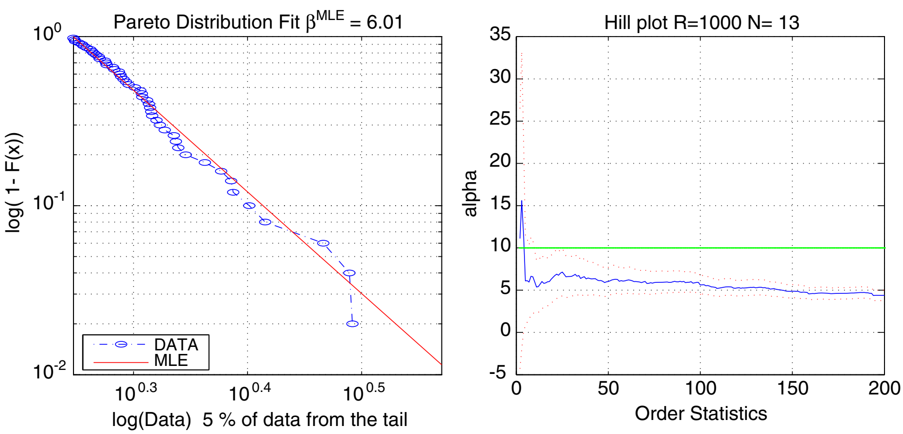


Fig. 7 Multiplicative cascade with log-normal multipliers. Theoretical tail exponent 10 is far from the estimated tail exponent.

3 Second order properties of distribution tails of RDEs

In this section, we show that weaker form in Eq. 1.16 of second order regular variation holds for RDEs, that is,

$$\int_x^\infty (P(X > u) - P(AX > u))\, du \sim cx^{-\alpha}, \tag{3.1}$$

as $x \to \infty$. We first illustrate Eq. 3.1 in a simple example, and then extend our proof to more general RDEs.

Example 3.1 Consider RDE defined as

$$X \overset{d}{=} AX + 1.$$

Then,

$$\begin{aligned}\int_x^\infty (P(X > u) - P(AX > u))\, du &= \int_x^\infty (P(X > u) - P(X > u + 1))\, du \\ &= \int_x^\infty P(X > u)du - \int_{x+1}^\infty P(X > u)du \\ &= \int_x^{x+1} P(X > u)du.\end{aligned}$$

Since $P(X > u)$ is monotone decreasing, we have

$$P(X > x + 1) \le \int_x^{x+1} P(X > u)du \le P(X > x). \tag{3.2}$$

By Theorem 1.1, both sides of Eq. 3.2 behave as $c_+x^{-\alpha}$. Therefore, as $x \to \infty$,

$$\int_x^\infty (P(X > u) - P(AX > u))\, du \sim c_+x^{-\alpha}.$$

As the following theorem shows, the relation in Eq. 3.1 holds for a large class of RDEs. We first consider the case when A, B and X in Eq. 1.2 are all nonnegative. The general case is considered later in the section.

Theorem 3.1 *Let $A \ge 0$, $B \ge 0$ and $X \ge 0$ a.s. and (A, B) be independent of X. Suppose that the assumptions of Theorem 1.1 hold. In addition, if*

$$EX < \infty, \quad EA^\alpha B < \infty, \tag{3.3}$$

$$x^\alpha EA1_{\{B>Cx\}} \to 0, \quad \textit{for any } C > 0, \ \textit{as } x \to \infty, \tag{3.4}$$

$$x^\alpha EA1_{\{A>Cx\}} \to 0, \quad \textit{for any } C > 0, \ \textit{as } x \to \infty, \tag{3.5}$$

$$x^\alpha \int_x^\infty P(B > z)dz \to C_+, \quad \textit{as } x \to \infty, \tag{3.6}$$

where $C_+ \geq 0$, then RDE X in Eq. 1.2 satisfies

$$\int_x^\infty (P(X > u) - P(AX > u))\, du \sim x^{-\alpha}\,(C_+ + c_+ E(A^\alpha B))\,, \tag{3.7}$$

with positive constant c_+ defined in Eq. 1.8.

Proof Observe that

$$\begin{aligned}
&\int_x^\infty (P(X > u) - P(AX > u))\, du \\
&\quad = \int_x^\infty \int_0^\infty \int_0^\infty (P\,(aX > u - b) - P\,(aX > u))\, F_{A,B}(da, db) du \\
&\quad = \int_0^\infty \int_0^\infty \int_{x-b}^x P(aX > u) du F_{A,B}(da, db).
\end{aligned} \tag{3.8}$$

For fixed $c < 1$, we can further rewrite Eq. 3.8 by splitting the range of b into $(0, cx)$ and (cx, ∞)

$$\begin{aligned}
&\int_{cx}^\infty \int_0^\infty \int_{x-b}^x P(aX > u) du F_{A,B}(da, db) \\
&\quad + \int_0^{cx} \int_0^\infty \int_{x-b}^x P(aX > u) du F_{A,B}(da, db) =: I + L.
\end{aligned} \tag{3.9}$$

For the integral I, write $I = I_1 + I_2$, where

$$I_1 = \int_{cx}^x \int_0^\infty \int_{x-b}^x P(aX > u) du F_{A,B}(da, db),$$

$$I_2 = \int_x^\infty \int_0^\infty \int_{x-b}^x P(aX > u) du F_{A,B}(da, db).$$

The integral I_1 can be bounded as

$$I_1 \leq \int_{cx}^\infty \int_0^\infty E(aX) F_{A,B}(da, db) = EXEA1_{\{B > cx\}},$$

and by the assumptions in Eqs. 3.3 and 3.4, we have

$$x^\alpha I_1 \to 0. \tag{3.10}$$

Observe for I_2 that

$$I_2 = \int_x^\infty \int_0^\infty \left(\int_{x-b}^0 P(aX > u)du + \int_0^x P(aX > u)du \right) F_{A,B}(da, db)$$
$$= \int_x^\infty \int_0^\infty \left((b-x) + \int_0^x P(aX > u)du \right) F_{A,B}(da, db)$$
$$= \int_x^\infty \int_0^\infty (b-x) F_{A,B}(da, db)$$
$$+ \int_x^\infty \int_0^\infty \int_0^x P(aX > u)du F_{A,B}(da, db) =: I_{2,1} + I_{2,2}.$$

Note that, by Eq. 3.6,

$$x^\alpha I_{2,1} = x^\alpha E(B-x)_+ = x^\alpha \int_0^\infty P((B-x)_+ > y)dy = x^\alpha \int_x^\infty P(B > z)dz \to C_+. \tag{3.11}$$

By the assumptions in Eqs. 3.3 and 3.4 with $C = 1$, we have

$$x^\alpha I_{2,2} \le x^\alpha EXEA1_{\{B \ge x\}} \to 0. \tag{3.12}$$

Combining Eqs. 3.10, 3.11 and 3.12 yields

$$x^\alpha I \to C_+. \tag{3.13}$$

We now turn to the integral L in Eq. 3.9. By Theorem 1.1, we can select x_0 such that, for any given $\epsilon > 0$,

$$|x^\alpha P(X > x) - c_+| \le \epsilon, \text{ for all } x \ge x_0, \tag{3.14}$$

where c_+ is a constant described in Eq. 1.8. For such x_0, write the integral L as

$$L = \int_0^{cx} \int_{(x-b)/x_0}^\infty \int_{x-b}^x P(aX > u)du F_{A,B}(da, db)$$
$$+ \int_0^{cx} \int_0^{(x-b)/x_0} \int_{x-b}^x P(aX > u)du F_{A,B}(da, db) =: J + K. \tag{3.15}$$

For fixed x_0, the integral J can be bounded as

$$J \le EXE\big(A1_{\{0<B<cx\}} 1_{\{A>(x-B)/x_0\}}\big) \le EXE\left(A1_{\{A>(1-c)x/x_0\}}\right),$$

since $B \in (0, cx)$ implies $(x-B)/x_0 \ge (1-c)x/x_0$. Hence, Eq. 3.5 implies that

$$x^\alpha J \to 0. \tag{3.16}$$

Consider now the integral K in Eq. 3.15. Since $P(aX > u)$ is monotone decreasing, the integral K satisfies,

$$x^{\alpha} K_1 := x^{\alpha} \int_0^{cx} \int_0^{(x-b)/x_0} bP(aX > x)\, F_{A,B}(da, db) \leq x^{\alpha} K$$
$$\leq x^{\alpha} \int_0^{cx} \int_0^{(x-b)/x_0} bP(aX > x - b)\, F_{A,B}(da, db) =: x^{\alpha} K_2. \tag{3.17}$$

Write the integral $x^{\alpha} K_2$ as

$$c_+ \int_0^{cx} \int_0^{(x-b)/x_0} b a^{\alpha} \left(1 - \frac{b}{x}\right)^{-\alpha} \frac{P\left(X > \frac{x-b}{a}\right)}{c_+ \left(\frac{x-b}{a}\right)^{-\alpha}} F_{A,B}(da, db).$$

Since $b \in (0, cx)$ and $a \in (0, (x-b)/x_0)$,

$$\left(1 - \frac{b}{x}\right)^{-\alpha} \frac{P\left(X > \frac{x-b}{a}\right)}{c_+ \left(\frac{x-b}{a}\right)^{-\alpha}} \leq (1-c)^{-\alpha} D_1,$$

where D_1 is some constant determined by Eq. 3.14. The assumption in Eq. 3.3 and the dominated convergence theorem yield

$$x^{\alpha} \int_0^{cx} \int_0^{(x-b)/x_0} bP(aX > x - b)\, F_{A,B}(da, db) \to c_+ E(A^{\alpha} B). \tag{3.18}$$

Similarly, $x^{\alpha} K_1$ becomes

$$c_+ \int_0^{cx} \int_0^{(x-b)/x_0} b a^{\alpha} \frac{P\left(X > \frac{x}{a}\right)}{c_+ \left(\frac{x}{a}\right)^{-\alpha}} F_{A,B}(da, db)$$

and $a \leq (x-b)/x_0$ implies

$$\frac{P\left(X > \frac{x}{a}\right)}{c_+ \left(\frac{x}{a}\right)^{-\alpha}} \leq D_2,$$

for some constant D_2. Again, by the assumption in Eq. 3.3 and the dominated convergence theorem,

$$\int_0^{cx} \int_0^{(x-b)/x_0} bP(aX > x)\, F_{A,B}(da, db) \to c_+ E(A^{\alpha} B). \tag{3.19}$$

Hence, Eqs. 3.18 and 3.19 imply

$$x^{\alpha} K \to c_+ E(A^{\alpha} B). \tag{3.20}$$

Finally, combining Eqs. 3.13, 3.16 and 3.20 yields Eq. 3.7. □

Remarks

1. Equation 1.16 implies Eq. 1.17, for example, when $P(X > u) - P(AX > u)$ is ultimately monotone (see, for example, Bingham et al.

1989, p. 39). Whether the latter monotonicity holds is still an open question.

2. Note that, for $\delta > 0$,

$$EA1_{\{B>Cx\}} \leq x^{-\alpha-\delta}C^{-\alpha-\delta}EAB^{\alpha+\delta}.$$

Hence, if

$$EAB^{\alpha+\delta} < \infty \tag{3.21}$$

for some $\delta > 0$, then Eq. 3.4 is satisfied. Similarly, the conditions in Eqs. 3.5 and 3.6 hold if

$$EA^{\alpha+1+\delta} < \infty \text{ and } EB^{\alpha+1+\delta} < \infty, \tag{3.22}$$

for some $\delta > 0$, respectively. In particular, $EB^{\alpha+1+\delta} < \infty$ implies $C_+ = 0$.

Example 3.2 (ARCH(1) model) Recall the discussion on ARCH(1) model found in Section 2.1. The model satisfies the assumptions of Theorem 3.1 for $\kappa = \alpha/2 > 1$. Indeed, for such κ, $E\xi_t^2 < \infty$ and obviously $EA_t^\kappa B_t = E(\lambda\epsilon_t^2)^\kappa(\beta\epsilon_t^2) < \infty$ so that Eq. 3.3 holds. The conditions in Eqs. 3.21–3.22 hold (with $C_+ = 0$) in the second remark above because $A_t = \lambda\epsilon_t^2$, $B_t = \beta\epsilon_t^2$ have all their moments finite for normal error terms ϵ_t. Hence, by Theorem 3.1, if $\kappa = \alpha/2 > 1$,

$$\int_x^\infty \big(P(\xi^2 > u) - P(\lambda\epsilon^2\xi^2 > u)\big)\,du \sim c_+\lambda^{2\alpha}\beta E\big(\epsilon^{2\alpha+2}\big)x^{-\alpha/2}$$

or, by symmetry and a change of variables,

$$\int_x^\infty \Big(P(\xi > v) - P(\sqrt{\lambda}\epsilon\xi > v)\Big)\,vdv \sim \frac{c_+}{4}\lambda^{2\alpha}\beta E(\epsilon^{2\alpha+2})x^{-\alpha}.$$

Example 3.3 (Multiplicative cascades with lognormal multipliers) Consider RDE in Eq. 2.27 with tail exponent $\alpha - 1$ and log-normal multipliers. If $\alpha > 2$ or $\alpha - 1 > 1$, then the condition in Eq. 3.3 in Theorem 3.1 is satisfied because

$$E\widetilde{M} = EM^2 < \infty, \quad E\widetilde{A}^{\alpha-1}\widetilde{B} = 2EA_1^\alpha EA_2EM = \frac{1}{2} < \infty$$

since $EM = 1$ by Theorem 2.1. Condition in Eq. 3.4 can be easily checked by observing that

$$x^{\alpha-1}E\widetilde{A}1_{\{\widetilde{B}>Cx\}} = x^{\alpha-1}2E\left(A_1^2 1_{\{A_2M>Cx\}}\right) = x^{\alpha-1}2EA_1^2P(A_2M > Cx) \to 0,$$

since Breiman's theorem (Breiman 1965) implies $P(A_2M>Cx)\sim 1/2c_+(Cx)^{-\alpha}$. The condition in Eq. 3.5 can be verified through the first condition in Eq. 3.22 with $\delta = 1$,

$$E\widetilde{A}^{\alpha+1} = 2EA_1^{\alpha+2} < \infty,$$

since log-normal distribution has all its moments finite. Note next that the condition in Eq. 3.6 becomes

$$x^{\alpha-1}\int_x^\infty P(\widetilde{B}>z)dz = x^{\alpha-1}\int_x^\infty P(A_2M>z)\,dz. \tag{3.23}$$

Applying Breiman's theorem again, we have

$$P(A_2M>z)\sim\frac{1}{2}c_+z^{-\alpha},$$

and for sufficiently large x, Eq. 3.23 leads to

$$x^{\alpha-1}\int_x^\infty P(A_2M>z)\,dz\sim x^{\alpha-1}\int_x^\infty \frac{1}{2}c_+z^{-\alpha}dz=\frac{c_+}{2(\alpha-1)}.$$

Hence, if $\alpha>2$, by Theorem 3.1, MC with log-normal multiplier satisfies the relation

$$x^{\alpha-1}\int_x^\infty \left(P(\widetilde{M}>u)-P(\widetilde{A}\widetilde{M}>u)\right)du\sim\frac{c_+(\alpha+1)}{2(\alpha-1)}. \tag{3.24}$$

By using the relationship between $\widetilde{M}$, $\widetilde{A}$ and M, A found in Section 2.3, the relation in Eq. 3.24 can be rewritten as

$$x^{\alpha-1}\int_x^\infty \left(\int_u^\infty yP_M(dy)-E\int_{u/\widetilde{A}}^\infty yP_M(dy)\right)du\sim\frac{c_+(\alpha+1)}{2(\alpha-1)}. \tag{3.25}$$

Furthermore, if Eq. 3.24 implies

$$x^{\alpha}\left(P(\widetilde{M}>x)-P(\widetilde{A}\widetilde{M}>x)\right)\sim\frac{c_+(\alpha+1)}{2}, \tag{3.26}$$

then this could be translated back to M, A as follows. Similarly to Eq. 2.28 one has

$$xP(AM>x)=\frac{1}{2}P(\widetilde{A}\widetilde{M}>x)-\frac{x}{2}\int_x^\infty y^{-2}P(\widetilde{A}\widetilde{M}>y)dy. \tag{3.27}$$

By Eqs. 2.28 and 3.27,

$$\frac{x\left(P(M>x)-2P(AM>x)\right)}{P(\widetilde{M}>x)-P(\widetilde{A}\widetilde{M}>x)}=1-\frac{x\int_x^\infty y^{-2}\left(P(\widetilde{M}>y)-P(\widetilde{A}\widetilde{M}>y)\right)dy}{P(\widetilde{M}>x)-P(\widetilde{A}\widetilde{M}>x)}. \tag{3.28}$$

The right-hand side of Eq. 3.28 converges to $1-1/(\alpha+1)=\alpha/(\alpha+1)$, as $x\to\infty$. Therefore, using Eq. 3.26,

$$x^{\alpha+1}\left(P(M>x)-2P(AM>x)\right)\sim\frac{c_+\alpha}{2}. \tag{3.29}$$

Theorem 3.2 below generalizes second order properties to real-valued X, A and B satisfying RDE in Eq. 1.2.

Theorem 3.2 *Suppose that the assumptions of Theorem* 1.1 *hold. If*

$$E|X| < \infty, \quad E|A|^\alpha |B| < \infty, \tag{3.30}$$

$$x^\alpha E|A|1_{\{|A|>Cx\}} \to 0, \quad \text{for any } C > 0, \text{ as } x \to \infty, \tag{3.31}$$

$$x^\alpha E|A|1_{\{|B|>Cx\}} \to 0, \quad \text{for any } C > 0, \text{ as } x \to \infty, \tag{3.32}$$

$$x^\alpha E(B-x)_+ 1_{\{A>0\}} \to C_+^1, \quad x^\alpha E(B-x)_+ 1_{\{A<0\}} \to C_+^2, \quad \text{as } x \to \infty, \tag{3.33}$$

where C_+^1 *and* C_+^2 *are nonnegative constants. Then, RDE X in Eq.* 1.2 *satisfies*

$$x^\alpha \int_x^\infty (P(X > u) - P(AX > u))\, du \sim C_+^1 + C_+^2 + c_+ E(A_+^\alpha B) + c_- E(A_-^\alpha B), \tag{3.34}$$

where constants c_+ *and* c_- *are defined in Eq.* 1.8.

Proof We sketch the proof as it is similar to that of Theorem 3.1. Split the integral in Eq. 3.34 according to the sign of random variables A and B as

$$\begin{aligned}
&\int_x^\infty (P(X > u) - P(AX > u))\, du \\
&\quad = \Big(\int_0^\infty \int_0^\infty \int_x^\infty + \int_0^\infty \int_{-\infty}^0 \int_x^\infty \\
&\qquad + \int_{-\infty}^0 \int_0^\infty \int_x^\infty + \int_{-\infty}^0 \int_{-\infty}^0 \int_x^\infty \Big) (P(aX + b > u) \\
&\qquad - P(aX > u)) du F_{A,B}(da, db) =: J_1 + J_2 + J_3 + J_4.
\end{aligned}$$

Applying the proof of Theorem 3.1 with A_+ and B_+ gives

$$x^\alpha J_1 \to C_+^1 + c_+ E(A_+^\alpha B_+).$$

Note that J_2 can be related to J_1 by rewriting it as

$$J_2 = \int_0^\infty \int_0^\infty \int_x^\infty (P(aX < -(u-b)) - P(aX < -u))\, du F_{-A,B}(da, db).$$

Hence,

$$x^\alpha J_2 \to C_+^2 + c_- E A_-^\alpha B_+.$$

From the Kesten's result, there is x_0 such that, for all $x > x_0$ and given $\epsilon > 0$,

$$|x^\alpha P(X > x) - c_+| \le \epsilon. \tag{3.35}$$

For such x_0, J_3 can be rewritten as

$$\begin{aligned}
J_3 &= \int_{-\infty}^{0}\int_{0}^{\infty}\int_{x-b}^{x} P(aX>u)duF_{A,B}(da,db)\\
&= \int_{-\infty}^{0}\int_{0}^{x/x_0}\int_{x-b}^{x} P(aX>u)duF_{A,B}(da,db)\\
&\quad + \int_{-\infty}^{0}\int_{x/x_0}^{\infty}\int_{x-b}^{x} P(aX>u)duF_{A,B}(da,db) =: J_{3,1}+J_{3,2}.
\end{aligned}$$

Second term $J_{3,2}$ does not contribute to the asymptotics because Eqs. 3.30 and 3.31 imply

$$\begin{aligned}
x^{\alpha}|J_{3,2}| \le x^{\alpha}\int_{-\infty}^{0}\int_{x/x_0}^{\infty}\int_{0}^{\infty} P(aX>u)duF_{A,B}(da,db)\\
= x^{\alpha}EX_{+}EA1_{\{A>x/x_0\}}1_{\{B<0\}} \to 0.
\end{aligned}$$

First term $J_{3,1}$ satisfies

$$\begin{aligned}
x^{\alpha}\int_{-\infty}^{0}\int_{0}^{x/x_0} bP\left(aX>x\right)F_{A,B}(da,db) \le x^{\alpha}J_{3,1}\\
\le x^{\alpha}\int_{-\infty}^{0}\int_{0}^{x/x_0} bP\left(aX>x-b\right)F_{A,B}(da,db). \qquad (3.36)
\end{aligned}$$

The right-hand side of Eq. 3.36 can be rewritten as

$$c_{+}\int_{-\infty}^{0}\int_{0}^{x/x_0} ba^{\alpha}\left(1-\frac{b}{x}\right)^{-\alpha}\frac{P\left(X>\frac{x-b}{a}\right)}{c_{+}\left(\frac{x-b}{a}\right)^{-\alpha}}F_{A,B}(da,db).$$

Relation in Eq. 3.35 and $1-b/x>1$ imply further that

$$\left(1-\frac{b}{x}\right)^{-\alpha}\frac{P\left(X>\frac{x-b}{a}\right)}{c_{+}\left(\frac{x-b}{a}\right)^{-\alpha}} \le D_1,$$

for some constant D_1. Arguing similarly for the left-hand side of Eq. 3.36 and applying the dominated convergence theorem lead to

$$x^{\alpha}J_{3,1} \to -c_{+}EA_{+}^{\alpha}B_{-}.$$

Observe for J_4 that

$$J_4 = \int_{0}^{\infty}\int_{0}^{\infty}\int_{x}^{\infty}\left(P\left(aX<-(u+b)\right)-P\left(aX<-u\right)\right)duF_{-A,-B}(da,db).$$

As for J_3, this leads to

$$x^{\alpha}J_4 \to -c_{-}E(A_{-}^{\alpha}B_{-}).$$

Gathering the results for J_1, J_2, J_3 and J_4 leads to the desired result. □

Finally, we show that if distribution tail behaves as in Eq. 1.13 and satisfies weaker form of 2RV in Eq. 1.16, then $\rho = -1$. For simplicity, we only consider the case of nonnegative A, B and X. We need the following lemma relating distribution tails $P(X > x)$ and $P(AX > x)$.

Lemma 3.1 *Suppose $A \geq 0$ a.s., $EA^{\alpha-\rho} < \infty$ for $\rho < 0$ and*

$$P(X > x) - c_1 x^{-\alpha} \sim c_2 x^{-\alpha+\rho}. \tag{3.37}$$

Then

$$P(AX > x) - c_1 x^{-\alpha} \sim c_2 EA^{\alpha-\rho} x^{-\alpha+\rho}. \tag{3.38}$$

Proof Observe that

$$P(AX > x) - c_1 x^{-\alpha} = \int_0^\infty \left(P\left(X > \frac{x}{a}\right) - c_1 \left(\frac{x}{a}\right)^{-\alpha} \right) F_A(da), \tag{3.39}$$

since $EA^\alpha = 1$. Relation in Eq. 3.37 implies that for any $\epsilon > 0$, there is x_0 such that

$$|x^{\alpha-\rho}(P(X > x) - c_1 x^{-\alpha}) - c_2| \leq \epsilon, \text{ for all } x > x_0. \tag{3.40}$$

For such chosen x_0, we have

$$\frac{P(AX > x) - c_1 x^{-\alpha}}{c_1 EA^{\alpha-\rho} x^{-\alpha+\rho}} = \frac{\int_0^{x/x_0} \left(P\left(X > \frac{x}{a}\right) - c_1 \left(\frac{x}{a}\right)^{-\alpha} \right) F_A(da)}{c_1 EA^{\alpha-\rho} x^{-\alpha+\rho}} + \frac{\int_{x/x_0}^\infty \left(P\left(X > \frac{x}{a}\right) - c_1 \left(\frac{x}{a}\right)^{-\alpha} \right) F_A(da)}{c_1 EA^{\alpha-\rho} x^{-\alpha+\rho}} =: I + J. \tag{3.41}$$

Note first that the second term J in Eq. 3.41 does not contribute to the asymptotics. Indeed,

$$\left| \int_{x/x_0}^\infty \left(P\left(X > \frac{x}{a}\right) - c_1 \left(\frac{x}{a}\right)^{-\alpha} \right) F_A(da) \right| \leq \int_{x/x_0}^\infty \left(P\left(X > \frac{x}{a}\right) + c_1 \left(\frac{x}{a}\right)^{-\alpha} \right) F_A(da).$$

$$\leq E1_{\{A \geq x/x_0\}} + c_1 x^{-\alpha} E(A^\alpha 1_{\{A > x/x_0\}})$$

$$\leq E\left(\left(\frac{A}{x/x_0}\right)^{\alpha-\rho} 1_{\{A \geq x/x_0\}} \right) + c_1 x^{-\alpha} E\left(A^\alpha \left(\frac{A}{x/x_0}\right)^{-\rho} 1_{\{A > x/x_0\}} \right)$$

yields

$$|J| \le \frac{x^{-\alpha+\rho}x_0^{\alpha-\rho}E\left(A^{\alpha-\rho}1_{\{A>x/x_0\}}\right) + c_1 x^{-\alpha+\rho}x_0^{-\rho}E(A^{\alpha-\rho}1_{\{A>x/x_0\}})}{c_1 EA^{\alpha-\rho}x^{-\alpha+\rho}} \to 0,$$

as $x \to \infty$ since $EA^{\alpha-\rho} < \infty$.

The first term I, on the other hand, can be bounded using relation in Eq. 3.40 as

$$\frac{\int_0^{x/x_0}(c_2-\epsilon)(x/a)^{-\alpha+\rho}F_A(da)}{c_2 EA^{\alpha-\rho}x^{-\alpha+\rho}} \le I \le \frac{\int_0^{x/x_0}(c_2+\epsilon)(x/a)^{-\alpha+\rho}F_A(da)}{c_2 EA^{\alpha-\rho}x^{-\alpha+\rho}}.$$

By taking limit $x \to \infty$ and $\epsilon \downarrow 0$, dominated convergence theorem implies

$$I \to 1,$$

since $EA^{\alpha-\rho} < \infty$. □

Proposition 3.1 *If $A \ge 0$ a.s., $EA^{\alpha-\rho} < \infty$ for some $\rho < 0$, and*

$$P(X > x) - c_1 x^{-\alpha} \sim c_2 x^{-\alpha+\rho}, \tag{3.42}$$

$$\int_x^\infty (P(X > u) - P(AX > u))\,du \sim cx^{-\alpha}, \tag{3.43}$$

then $\rho = -1$ and $c = c_2(1 - EA^{\alpha-\rho})/\alpha$.

Proof Lemma 3.1 implies that

$$P(AX > x) - c_1 x^{-\alpha} \sim c_2 EA^{\alpha-\rho}x^{-\alpha+\rho}. \tag{3.44}$$

Therefore, we have

$$\begin{aligned}\int_x^\infty (P(X>u) - P(AX>u))\,du &= \int_x^\infty \left(c_2(1-EA^{\alpha-\rho})u^{-\alpha+\rho} + o(u^{-\alpha+\rho})\right)du\\ &= \frac{-c_2(1-EA^{\alpha-\rho})}{-\alpha+\rho+1}x^{-\alpha+\rho+1} + \int_x^\infty o(u^{-\alpha+\rho})du.\\ &= \frac{-c_2(1-EA^{\alpha-\rho})}{-\alpha+\rho+1}x^{-\alpha+\rho+1} + o(x^{-\alpha+\rho+1}).\end{aligned}$$

Finally, assumption in Eq. 3.43 gives $\rho = -1$ and $c = c_2(1 - EA^{\alpha-\rho})/\alpha$. □

Remark Our results suggest for RDE that

$$P(X > x) = c_1 x^{-\alpha} + c_2 x^{-\alpha-1} + o(x^{-\alpha-1}). \tag{3.45}$$

This imposes conditions on c_1 and c_2 in the following sense. One expects that

$$\begin{aligned}P(X>x) &= P(AX+B>x) = \int_0^\infty\int_0^\infty P\left(X>\frac{x-b}{a}\right)F_{A,B}(da,db)\\ &= \int_0^\infty\int_0^\infty \left\{c_1\left(\frac{x-b}{a}\right)^{-\alpha}+c_2\left(\frac{x-b}{a}\right)^{-\alpha-1}\right.\\ &\qquad \left. +o\left(\left(\frac{x-b}{a}\right)^{-\alpha-1}\right)\right\}F_{A,B}(da,db)\\ &= \int_0^\infty\int_0^\infty \left(c_1a^\alpha x^{-\alpha}+c_1a^\alpha\left((x-b)^{-\alpha}-x^{-\alpha}\right)+c_2a^{\alpha+1}x^{-\alpha-1}\right.\\ &\qquad \left.+o\left(x^{-\alpha-1}\right)\right)F_{A,B}(da,db)\\ &= c_1x^{-\alpha}+\left(c_1\alpha E(A^\alpha B)+c_2EA^{\alpha+1}\right)x^{-\alpha-1}+o\left(x^{-\alpha-1}\right), \end{aligned} \tag{3.46}$$

by using the relation

$$(x-b)^{-\alpha}-x^{-\alpha}=x^{-\alpha}\left(\left(1-\frac{b}{x}\right)^{-\alpha}-1\right)=b\alpha x^{-\alpha-1}+o\left(x^{-\alpha-1}\right).$$

Therefore, Eqs. 3.45 and 3.46 are consistent only when

$$c_2=\frac{c_1\alpha E(A^\alpha B)}{1-EA^{\alpha+1}}. \tag{3.47}$$

4 Estimating tail exponent under second order regular variation

Theorem 3.1 and explicit example considered in Section 2.2 suggest to introduce the second order term when estimating tail exponent in RDE. In this section, we propose a simple tail exponent estimator of regression type and compare its performance to that of Feuerverger and Hall (1999) estimator through a small simulation study. More comprehensive discussion including theoretical properties of proposed estimator can be found in Baek and Pipiras (2009).

4.1 Several estimation methods

Suppose that the distribution tail behaves as

$$\overline{F}(x)=P(X>x)=c_1x^{-\alpha}+c_2x^{-\alpha-1}+o\left(x^{-1}\right)=c_1x^{-\alpha}\left(1+\frac{c_2}{c_1}x^{-1}\right)+o\left(x^{-1}\right), \tag{4.1}$$

as $x\to\infty$. By taking the log-transformation, we have, as $x\to\infty$,

$$\log\left(\overline{F}(x)\right)\approx\log c_1-\alpha\log x+\log\left(1+\frac{c_2}{c_1}x^{-1}\right)\approx\log c_1-\alpha\log x+\frac{c_2}{c_1}x^{-1}. \tag{4.2}$$

Estimator based on linear regression According to Eq. 4.2, tail exponent α could be estimated by a linear regression of the logarithm of empirical distribution tail of X on $(1, \log x, 1/x)$. This approach generalizes the least squares tail estimator based on the first order asymptotics, namely,

$$\log\left(\overline{F}(x)\right) \approx \log c_1 - \alpha \log x$$

(see, for example, Resnick 1997). More precisely, let $X^{(1)} \geq X^{(2)} \geq \ldots \geq X^{(r)}$ be the order statistics from $R \geq r$ i.i.d. copies of X. Minimizing the sum of squared errors

$$S(\beta_0, \alpha, \beta_1) = \sum_{i=1}^{r} \left(\log(i/R) - \beta_0 + \alpha \log X^{(i)} - \beta_1/X^{(i)}\right)^2$$

gives the least squares estimator of α as

$$\widehat{\alpha}_{LSE2} = e'(\mathbb{W}'\mathbb{W})^{-1}\mathbb{W}'\mathbb{F}, \tag{4.3}$$

where $e' = (0, -1, 0)$, $\mathbb{W} = (\mathbf{1}\ \ \log \mathbf{X}\ \ 1/\mathbf{X})$, $\mathbf{X} = (X^{(1)} X^{(2)} \ldots X^{(r)})'$, $\mathbf{1}/X = (1/X^{(1)} 1/X^{(2)} \ldots 1/X^{(r)})'$ and $\mathbb{F} = (\log 1/R\ \ \log 2/R\ \ \ldots\ \ \log r/R)'$. The parameter r plays the role of a threshold. We refer to the corresponding estimator of α as LSE2 estimator.

Feuerverger and Hall (1999) estimator Another estimator of tail exponent under second order regular variation was suggested by Feuerverger and Hall (1999). These authors use normalized log-spacings of order statistics and approximate them by a normalized Exponential distribution. Under the same setting as above, denote

$$U_i = i(\log X^{(i)} - \log X^{(i+1)})$$

and set $\delta(x) = -\alpha^{-1} c_1^{-(\alpha^{-1}+1)} c_2 x^{1/\alpha} =: D x^{1/\alpha}$. Then, one expects that

$$U_i \approx Z_i \alpha^{-1}(1 + \delta(i/R)) \approx Z_i \alpha^{-1} \exp\left(\delta(i/R)\right),$$

where Z_i's are independent Exponential random variables. This suggests the maximum likelihood estimator of α based on minimizing

$$L(D, \alpha) = D r^{-1} \sum_{i=1}^{r} (i/R)^{\alpha^{-1}} + \log\left(r^{-1} \sum_{i=1}^{r} U_i \exp\left\{-D(i/R)^{\alpha^{-1}}\right\}\right).$$

We refer to the corresponding estimator of α as FH estimator.

4.2 Simulation study

In this section, we present a simulation study based on the discussion above. In all cases considered, selecting threshold r is quite essential. We adopt here the same approach as in Hill plot. The tail exponent estimator should remain stable for a range of small choices of thresholds r. We replicated the simulations 100

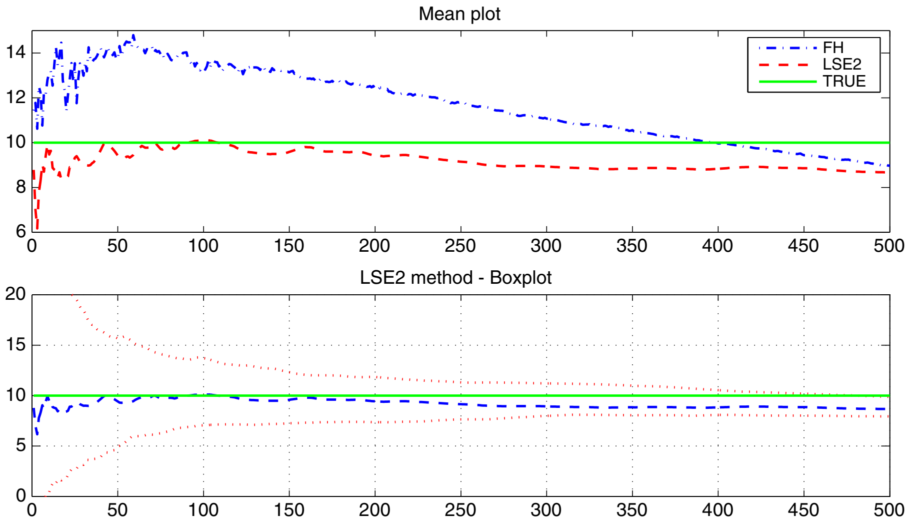


Fig. 8 Multiplicative cascade with log-normal multipliers, $\alpha = 10$ and $R = 1{,}000$. LSE2 estimator performs well.

times to produce a boxplot. Also, we have used $N = 100$ iterations in RDE in Eq. 1.1.

Figure 8 shows tail exponent estimator based on LSE2 and FH for multiplicative cascade model. LSE2 estimator performs quite well, especially for the thresholds between 50 to 150, while FH estimator overestimates parameter and

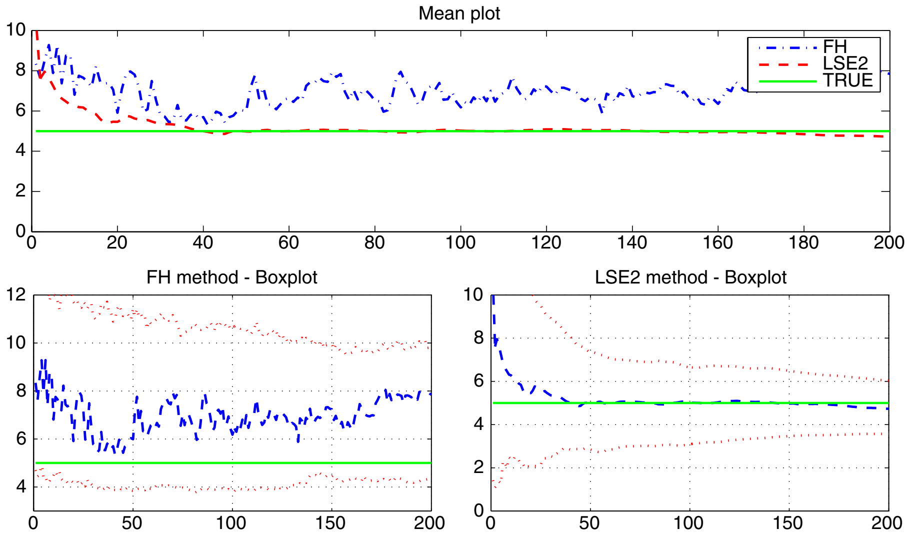


Fig. 9 Explicit example in Section 2.2 with $a_1 = 9$, $b = 5$ and $R = 2{,}000$.

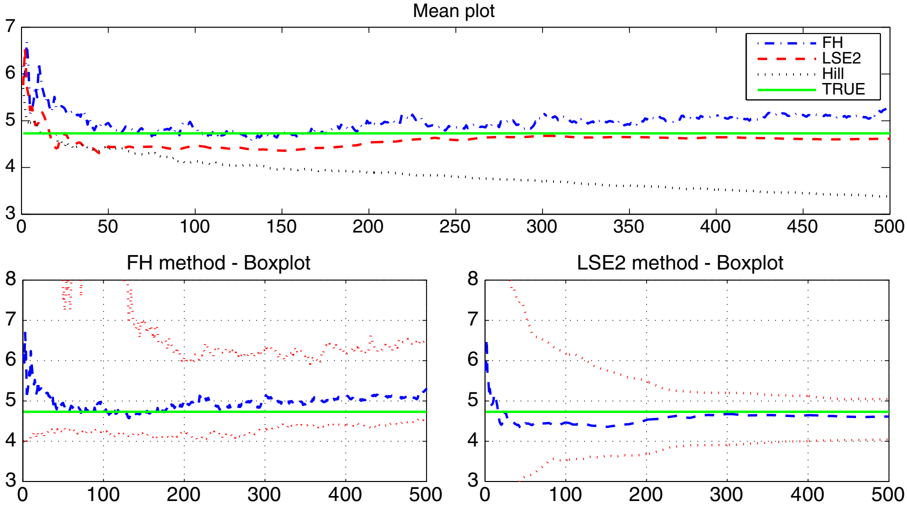


Fig. 10 ARCH(1) series with $\alpha = 4.73$ and $R = 5{,}000$.

is relatively unstable. Figure 9 shows estimation based on the explicit example of Section 2.2 with $a_1 = 9$ and $b = 5$. LSE2 estimator works quite well for the threshold greater than 40, while FH method overestimates the parameter and inter-quartile range is wider than for LSE2 method.

For ARCH(1) model, we have chosen parameters closer to those found in practice. For many economic time series representing volatility, one typically observes tail exponents between 3 and 4. Taking $\sigma^2 = 1$, $\lambda = .5$ and $\beta = 1$, we have tail exponent $\alpha = 4.73$. Figure 10 shows three estimators, FH, LSE2 and Hill estimators. All three estimators seem to work in some sense. Hill

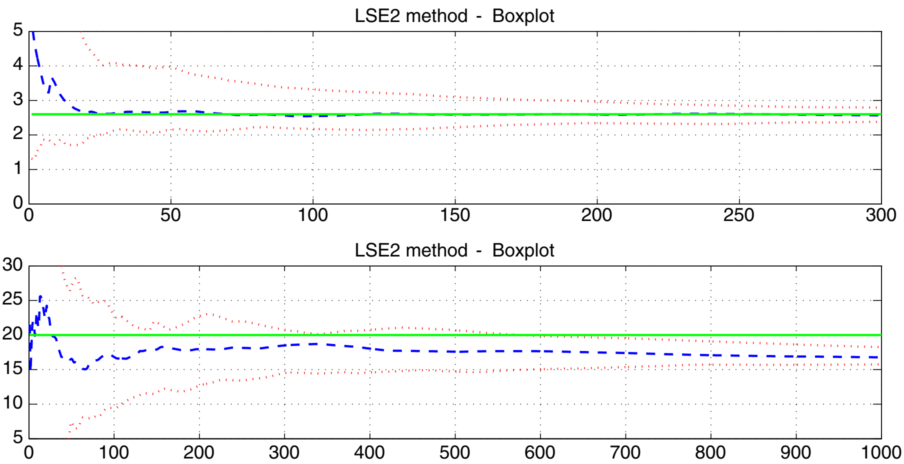


Fig. 11 ARCH(1) series with $\alpha = 2.6$, $R = 1{,}000$ (*top*) and $\alpha = 20$, $R = 100{,}000$ (*bottom*).

estimator works for few order statistics from the tail. Variants of Hill plot such as altHill, smooth Hill plot may be used to improve the estimation, but we will not pursue these approaches here. FH estimator is closer to the true value for the thresholds between 100 to 200, but IQR is rather wide. LSE2 method works well especially for large number of order statistics used (between 250 to 500) and accordingly it provides smaller variability.

It is also of interest to examine the performance of LSE2 estimator with respect to sample size R and tail exponent α. Simulation study shows that LSE2 estimator works fine even for moderate sample size when α is small. However, when tail exponent is really large, LSE2 method does not work for reasonable sample size. Figure 11 illustrates LSE2 method for ARCH(1) model with $\alpha = 2.6$, $R = 1000$ (top) and $\alpha = 20$, $R = 100{,}000$ (bottom). As seen form the figure, LSE2 method works poorly even for 100,000 independent samples when $\alpha = 20$.

5 Other issues

In this section, we discuss generalized Pareto distribution and the effect of temporal dependence on the tail exponent estimation of RDEs. We also extend our results to multidimensional RDEs.

5.1 Generalized Pareto distribution

One other popular family of distributions for power-law tail behavior consists of Generalized Pareto distributions (GPDs). Parametrized by the parameters $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$, GPD(α, μ, σ) has distribution tail given by

$$\overline{F}(x) = \left(1 + \frac{x-\mu}{\alpha\sigma}\right)^{-\alpha}, \quad x > \mu. \tag{5.1}$$

It has the tail exponent α. In view of our results for RDEs, it is interesting to ask whether GPD shares the same problems for larger values of α. If this is the case, then fitted values of large α should be interpreted with care. For example, if data were generated by exact GPD, the values of α fitted by MLE and that from the Hill plot would be quite different.

Similar to Section 2.2, consider deviations from the true Pareto tail as

$$\left|\frac{\overline{F}(x)}{((x-\mu)/\alpha\sigma)^{-\alpha}} - 1\right| = \left|\left(1 - \frac{\alpha\sigma}{x-\mu+\alpha\sigma}\right)^{\alpha} - 1\right| \le \epsilon. \tag{5.2}$$

For example, taking $\alpha = 10$, $\sigma = 2$, $\mu = 1$ and $\epsilon = .5$ gives numerical solution to Eq. 5.2 as $x > 279.05$. The probability of having GPD observations in this range is approximately 1.85×10^{-12}. If we increase σ to $\sigma = 3$, then $x > 418.92$ and the corresponding probability is approximately 3.95×10^{-14}. This shows that, for larger values of α, GPD has the Pareto-like region too far in the tail

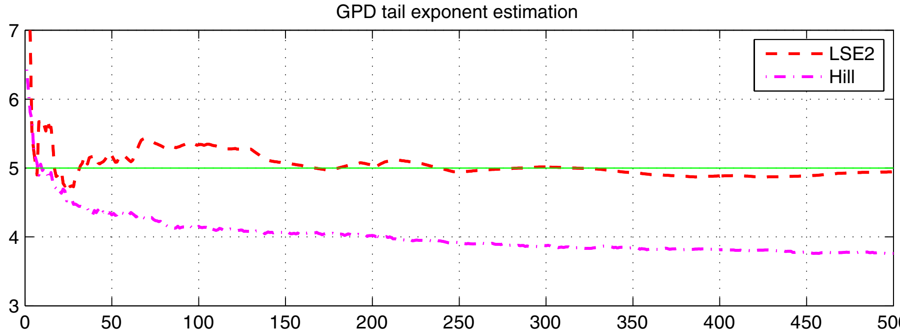


Fig. 12 GPD (5, 1, 2) tail exponent estimation with $R = 200{,}000$ observations.

for practical purposes as well. Note also that, as in the case of RDEs, GPD has similar second order term, namely,

$$\begin{aligned}\log \overline{F}(x) &= -\alpha \log\left(1 + \frac{x-\mu}{\alpha\sigma}\right)\\ &= -\alpha \log\left(\frac{x-\mu}{\alpha\sigma}\right) - \frac{\alpha^2\sigma}{x-\mu} + o\left(x^{-1}\right)\\ &= \alpha \log(\alpha\sigma) - \alpha \log x - \frac{\alpha^2\sigma - \alpha\mu}{x} + o\left(x^{-1}\right). \qquad (5.3)\end{aligned}$$

Figure 12 presents a Hill plot for $R = 200{,}000$ independent observations from GPD(5, 1, 2), and a similar, superimposed plot based on least squares (LSE2) taking the second term in Eq. 5.3 into account. Note that, even for such large sample size, Hill estimate is very biased. The LSE2 performs much better.

5.2 Temporal dependence

In this subsection, we study the effect of temporal dependence in estimating tail exponent. In brief, with temporal dependence, estimation is worse than that for independent observations. Figure 13 shows tail exponent estimation in ARCH(1) model with $\alpha = 8$. We generated 20,000 and 50,000 dependent ARCH(1) observations from Eq. 2.2 with first 200 observations disregarded for convergence. For comparison, we also generated 20,000 independent observations. Note from the figure that LSE2 estimator works well in the independent case, and its performance is worse for dependent observations. Increasing the sample size to 50,000, the dependent case resembles that with 20,000 independent observations. Note also that the simple Hill estimation is poor in both independent and dependent cases, the dependent case being worse.

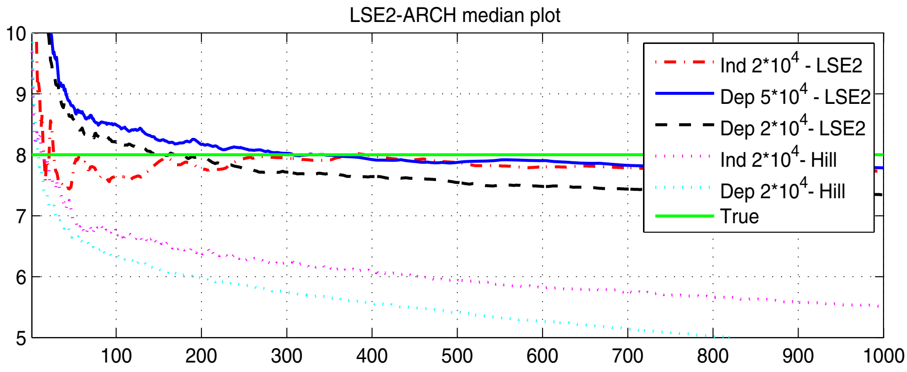


Fig. 13 Tail exponent estimations for dependent and independent ARCH(1) series.

5.3 Multidimensional extension

In this subsection, we extend our results to multidimensional RDEs,

$$\mathbf{X}_n = \mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{B}_n, \quad n \in \mathbb{Z}, \tag{5.4}$$

where $(\mathbf{A}_n, \mathbf{B}_n)$ is an i.i.d. sequence of $d \times d$ random matrices $\mathbf{A}_n$ and d-dimensional random vectors $\mathbf{B}_n$. We consider only the case when the entries of $\mathbf{A}_n, \mathbf{B}_n$ are nonnegative. Under mild conditions, multidimensional RDE has a stationary solution,

$$\mathbf{X} \stackrel{d}{=} \mathbf{A}\mathbf{X} + \mathbf{B}, \tag{5.5}$$

where $(\mathbf{A}, \mathbf{B}) =_d (\mathbf{A}_1, \mathbf{B}_1)$ is independent of $\mathbf{X}$. We recall next the result of Kesten (1973) for multidimensional RDEs. (Generalizations of Kesten's result can be found in Basrak et al. 2002a, de Saporta et al. 2004, Klüppelberg and Pergamenchtchikov 2004, Guivarc'h 2006 to name but a few.) Denote the Euclidean norm as $\|\cdot\|$ and the operator norm as $\|\cdot\|_{op}$, namely,

$$\|\mathbf{A}\|_{op} = \sup_{\|\mathbf{y}\|=1} \|\mathbf{A}\mathbf{y}\|.$$

Let also $S_+ = \{\mathbf{z} \in \mathbb{R}^d : \|\mathbf{z}\| = 1, \mathbf{z} > 0\}$ and, for $\mathbf{w} \in \mathbb{R}^d$,

$$\mathbf{w}^{\#} = \frac{\mathbf{w}}{\|\mathbf{w}\|}.$$

In particular, for $\mathbf{w} \in \mathbb{R}^d$ with nonnegative entries, note that $\mathbf{w}^{\#} \in S_+$.

Theorem 5.1 (Kesten 1973, Theorems 3 and 4) *Let $(\mathbf{A}_n, \mathbf{B}_n)$ be a sequence of i.i.d. $d \times d$ matrices $\mathbf{A}_n$ and $d \times 1$ vectors $\mathbf{B}_n$ with nonnegative entries. Assume that the following conditions hold:*

(A1) *For some $\epsilon > 0$, $E\|\mathbf{A}_1\|_{op}^{\epsilon} < 1$.*
(A2) *$\mathbf{A}_1$ has no zeros rows a.s.*

(A3) *The group generated by*

$$\{\log \rho(\pi) : \pi = \mathbf{a}_n \ldots \mathbf{a}_1 > 0, n \geq 1, \mathbf{a}_n \ldots \mathbf{a}_1 \in \textit{support of } \mathbf{A}_1\}$$

is dense in $\mathbb{R}$*, where* $\rho(\pi)$ *denotes the largest positive eigenvalue, known as Frobenius eigenvalue and* $\pi > 0$ *means that all entries of this matrix are positive.*

(A4) *There exists* $\kappa_0 > 0$ *such that*

$$E\left(\min_{i=1,\ldots,d} \sum_{j=1}^{d} A_1(i,j)\right)^{\kappa_0} \geq d^{\kappa_0/2},$$

where $A_1(i,j)$ *is a* (i,j) *entry of matrix* $\mathbf{A}_1$ *and*

$$E\left(\|\mathbf{A}_1\|_{op}^{\kappa_0} \log^+ \|\mathbf{A}_1\|_{op}\right) < \infty.$$

Then the following statements hold:

(R1) *There exists a unique solution* $\alpha \in (0, \kappa_0]$ *to the equation*

$$0 = \lim_{n\to\infty} \frac{1}{n} E \log \|\mathbf{A}_n \ldots \mathbf{A}_1\|_{op}^{\alpha}.$$

(R2) *There exists a unique solution* $\mathbf{X}$ *to the RDE in Eq. 5.5.*

(R3) *If* $E\|\mathbf{B}_1\|^{\alpha} < \infty$*, then, for all* $\mathbf{z} \in S_+$*,*

$$P(\mathbf{z}'\mathbf{X} > x) \sim x^{-\alpha} c_+ r(\mathbf{z}), \quad \textit{as } x \to \infty, \tag{5.6}$$

where c_+ *is a positive constant and* $r(\mathbf{z})$ *is a continuous and strictly positive function on* S_+ *satisfying*

$$r(\mathbf{z}) = E\|\mathbf{z}'\mathbf{A}\|^{\alpha} r(\mathbf{z}'\mathbf{A}^{\#}). \tag{5.7}$$

The following result extends Theorem 3.1 to the multidimensional case. The proof is similar to that of Theorem 3.1 (with an additional technical difficulty reflected by assumptions in Eqs. 5.12 and 5.13 below). We denote by μ a Haar measure on S_+.

Theorem 5.2 *Suppose that the assumptions of Theorem* 5.1 *hold and let* $(\mathbf{A}, \mathbf{B}) =_d (\mathbf{A}_1, \mathbf{B}_1)$*. In addition, suppose that for* $\mathbf{z} \in S_+$*,*

$$E\mathbf{X} < \infty, \quad E\|\mathbf{z}'\mathbf{A}\|^{\alpha}\mathbf{z}'\mathbf{B} < \infty, \tag{5.8}$$

$$x^{\alpha} E(\mathbf{z}'\mathbf{A} 1_{\{\mathbf{z}'\mathbf{B} > Cx\}}) \to \mathbf{0}, \quad \textit{for any } C > 0, \quad \textit{as } x \to \infty, \tag{5.9}$$

$$x^{\alpha} E(\mathbf{z}'\mathbf{A} 1_{\{\|\mathbf{z}'\mathbf{A}\| > Cx\}}) \to \mathbf{0}, \quad \textit{for any } C > 0, \quad \textit{as } x \to \infty, \tag{5.10}$$

$$x^{\alpha} \int_x^{\infty} P(\mathbf{z}'\mathbf{B} > u) du \to C_+(\mathbf{z}), \tag{5.11}$$

for some function $C_+(\mathbf{z}) \geq 0$. *Assume also that either*

$$\mathbf{z}'\mathbf{A} \text{ is a discrete, finite random vector, or} \tag{5.12}$$

$$P\left((\mathbf{z}'\mathbf{A}^{\#}) \in E\right) \to 0, \ \text{as } \mu(E) \to 0. \tag{5.13}$$

Then, the stationary solution $\mathbf{X}$ *of Eq.* 5.5 *satisfies*

$$\int_x^{\infty} \left(P(\mathbf{z}'\mathbf{X} > u) - P(\mathbf{z}'\mathbf{A}\mathbf{X} > u)\right) du \sim x^{-\alpha} \left(C_+(\mathbf{z}) + c_+ E\left(r(\mathbf{z}'\mathbf{A}^{\#})\|\mathbf{z}'\mathbf{A}\|^{\alpha}\mathbf{z}'\mathbf{B}\right)\right). \tag{5.14}$$

Proof Observe that

$$\begin{aligned}
&\int_x^{\infty} \left(P(\mathbf{z}'\mathbf{X} > u) - P(\mathbf{z}'\mathbf{A}\mathbf{X} > u)\right) du \\
&\quad = \int_x^{\infty} \int_{\mathbf{a},\mathbf{b}} \left(P\big((\mathbf{z}'\mathbf{a})\mathbf{X} + \mathbf{z}'\mathbf{b} > u\big) - P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big)\right) F_{\mathbf{A},\mathbf{B}}(d\mathbf{a}, d\mathbf{b}) \\
&\quad = \int_{\mathbf{a},\mathbf{b}} \int_{x-\mathbf{z}'\mathbf{b}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\mathbf{B}}(d\mathbf{a}, d\mathbf{b}) \\
&\quad = \int_{\mathbf{a},\overline{\mathbf{b}}} \int_{x-\overline{\mathbf{b}}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}),
\end{aligned}$$

where $\overline{\mathbf{b}} = \mathbf{z}'\mathbf{b}$ is a scalar and $\overline{\mathbf{B}} = \mathbf{z}'\mathbf{B}$. By splitting the range of $\overline{\mathbf{b}}$ with fixed $c < 1$, this can further be written as

$$\begin{aligned}
&\int_{cx}^{\infty} \int_{\mathbf{a}} \int_{x-\overline{\mathbf{b}}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}) \\
&\qquad + \int_0^{cx} \int_{\mathbf{a}} \int_{x-\overline{\mathbf{b}}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}) =: I + L.
\end{aligned}$$

For the integral I, write it as $I = I_1 + I_2$, where

$$\begin{aligned}
I_1 &= \int_{cx}^{x} \int_{\mathbf{a}} \int_{x-\overline{\mathbf{b}}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}), \\
I_2 &= \int_{x}^{\infty} \int_{\mathbf{a}} \int_{x-\overline{\mathbf{b}}}^{x} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big) du F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}).
\end{aligned}$$

For the integral I_1, we have

$$x^{\alpha} I_1 \leq x^{\alpha} \int_{cx}^{x} \int_{\mathbf{a}} E\big(\mathbf{z}'\mathbf{a}\mathbf{X}\big) F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}) \leq x^{\alpha} E\left(\mathbf{z}'\mathbf{A} 1_{\{\overline{\mathbf{B}} > cx\}}\right) E\mathbf{X} \to 0, \tag{5.15}$$

by Eq. 5.9.

Observe for I_2 that

$$I_2 = \int_x^\infty \int_{\mathbf{a}} \left\{ \int_{x-\bar{\mathbf{b}}}^{0} P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big)du + \int_0^x P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big)du \right\} F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big)$$

$$= \int_x^\infty \int_{\mathbf{a}} \left\{ \big(\bar{\mathbf{b}} - x\big) + \int_0^x P\big((\mathbf{z}'\mathbf{a})\mathbf{X} > u\big)du \right\} F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big) =: I_{2,1} + I_{2,2}.$$

Note that, by Eq. 5.11,

$$x^\alpha I_{2,1} = x^\alpha E(\mathbf{z}'\mathbf{B} - x)_+ = x^\alpha \int_x^\infty P(\mathbf{z}'\mathbf{B} > u)du \to C_+(\mathbf{z}). \tag{5.16}$$

Also note from Eq. 5.9 with $C = 1$ that

$$x^\alpha I_{2,2} \le x^\alpha E\left(\mathbf{z}'\mathbf{A}1_{\{\mathbf{z}'\mathbf{B}>x\}}\right) E\mathbf{X} \to 0. \tag{5.17}$$

Combining Eqs. 5.15, 5.16 and 5.17 yields

$$x^\alpha I \to C_+(\mathbf{z}). \tag{5.18}$$

For the integral L, for some x_0 to be determined later, write it as

$$L = \int_0^{cx} \int_{\mathbf{a}} \int_{x-\bar{\mathbf{b}}}^{x} P\big(\|\mathbf{z}'\mathbf{a}\|\big(\mathbf{z}'\mathbf{a}^{\#}\big)\mathbf{X} > u\big)du F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big)$$

$$= \int_0^{cx} \int_{\|\mathbf{z}'\mathbf{a}\|>\left(x-\bar{\mathbf{b}}\right)/x_0} \int_{x-\bar{\mathbf{b}}}^{x} P\big(\|\mathbf{z}'\mathbf{a}\|\big(\mathbf{z}'\mathbf{a}^{\#}\big)\mathbf{X} > u\big)du F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big)$$

$$+ \int_0^{cx} \int_{\|\mathbf{z}'\mathbf{a}\|\le\left(x-\bar{\mathbf{b}}\right)/x_0} \int_{x-\bar{\mathbf{b}}}^{x} P\big(\|\mathbf{z}'\mathbf{a}\|\big(\mathbf{z}'\mathbf{a}^{\#}\big)\mathbf{X} > u\big)du F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big) =: J + K.$$

For fixed x_0, observe that

$$x^\alpha J \le x^\alpha E\left(\mathbf{z}'\mathbf{A}1_{\{\|\mathbf{z}'\mathbf{A}\|>(1-c)x/x_0\}}\right) E\mathbf{X} \to 0, \tag{5.19}$$

using Eq. 5.10, since $\bar{\mathbf{b}} \in (0, cx)$ implies $(x - \bar{\mathbf{b}})/x_0 > (1 - c)x/x_0$.

We now show how one can deal with the integral K. We consider only the case in Eq. 5.13. (The case in Eq. 5.12 is easier and can be proved as below.) Let μ denote a Haar measure on S_+ as in the statement of the theorem. Since $\mu(S_+) < \infty$, Theorem 5.1 and Egoroff's theorem imply that, for any $\epsilon > 0$, there is $E_\epsilon \subset S_+$ such that $\mu(E_\epsilon) < \epsilon$ and

$$\sup_{\mathbf{w}\in S_+\setminus E_\epsilon} \left| x^\alpha P(\mathbf{w}'\mathbf{X} > x) - c_+ r(\mathbf{w}) \right| \to 0, \tag{5.20}$$

as $x \to \infty$. Now write the integral K as

$$K = \int_0^{cx} \int_{\|\mathbf{z}'\mathbf{a}\|\le(x-\bar{\mathbf{b}})/x_0} \int_{x-\bar{\mathbf{b}}}^{x} P\big(\|\mathbf{z}'\mathbf{a}\|\big(\mathbf{z}'\mathbf{a}^{\#}\big)\mathbf{X} > u\big)du F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\bar{\mathbf{b}}\big)$$

$$\cdot\left(1_{\{(\mathbf{z}'\mathbf{A}^{\#})\in E_\epsilon\}} + 1_{\{(\mathbf{z}'\mathbf{A}^{\#})\in S_+\setminus E_\epsilon\}}\right) =: K_1 + K_2. \tag{5.21}$$

For K_2, observe that

$$x^\alpha K_{2,1} := x^\alpha \int_0^{cx} \int_F \overline{\mathbf{b}} P\big(\|\mathbf{z}'\mathbf{a}\| \big(\mathbf{z}'\mathbf{a}^\#\big)\mathbf{X} > x\big) F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\overline{\mathbf{b}}\big) \le x^\alpha K_2$$
$$\le x^\alpha \int_0^{cx} \int_F \overline{\mathbf{b}} P\big(\|\mathbf{z}'\mathbf{a}\| \big(\mathbf{z}'\mathbf{a}^\#\big)\mathbf{X} > x - \overline{\mathbf{b}}\big) F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\overline{\mathbf{b}}\big) =: x^\alpha K_{2,2},$$

where $F = \{\mathbf{a} : \|\mathbf{z}'\mathbf{a}\| \le (x - \overline{\mathbf{b}})/x_0, (\mathbf{z}'\mathbf{a}^\#) \in S_+ \backslash E_\epsilon\}$. Write the integral $x^\alpha K_{2,2}$ as

$$x^\alpha K_{2,2} = \int_0^{cx} \int_F \overline{\mathbf{b}} c_+ r(\mathbf{z}'\mathbf{a}^\#) \|\mathbf{z}'\mathbf{a}\|^\alpha \left(1 - \frac{\overline{\mathbf{b}}}{x}\right)^{-\alpha} \frac{P\left((\mathbf{z}'\mathbf{a}^\#)\mathbf{X} \ge \frac{x-\overline{\mathbf{b}}}{\|\mathbf{z}'\mathbf{a}\|}\right)}{c_+ r(\mathbf{z}'\mathbf{a}^\#)\left(\frac{x-\overline{\mathbf{b}}}{\|\mathbf{z}'\mathbf{a}\|}\right)^{-\alpha}} F_{\mathbf{A},\overline{\mathbf{B}}}(d\mathbf{a}, d\overline{\mathbf{b}}).$$

By condition in Eq. 5.20, the term

$$\left(1 - \frac{\overline{\mathbf{b}}}{x}\right)^{-\alpha} \frac{P\left((\mathbf{z}'\mathbf{a}^\#)\mathbf{X} \ge \frac{x-\overline{\mathbf{b}}}{\|\mathbf{z}'\mathbf{a}\|}\right)}{c_+ r(\mathbf{z}'\mathbf{a}^\#)\left(\frac{x-\overline{\mathbf{b}}}{\|\mathbf{z}'\mathbf{a}\|}\right)^{-\alpha}}$$

is bounded on F for large enough x_0, and converges to 1 as $x \to \infty$. The dominated convergence theorem implies that

$$x^\alpha K_{2,2} \to c_+ Er(\mathbf{z}'\mathbf{A}^\#) \|\mathbf{z}'\mathbf{A}\|^\alpha 1_{\{(\mathbf{z}'\mathbf{A}^\#) \in S_+ \backslash E_\epsilon\}} \mathbf{z}'\mathbf{B}.$$

The same asymptotics holds for $x^\alpha K_{2,1}$, and we can conclude that

$$x^\alpha K_2 \to c_+ Er(\mathbf{z}'\mathbf{A}^\#) \|\mathbf{z}'\mathbf{A}\|^\alpha 1_{\{(\mathbf{z}'\mathbf{A}^\#) \in S_+ \backslash E_\epsilon\}} \mathbf{z}'\mathbf{B}. \tag{5.22}$$

For K_1, observe that

$$x^\alpha K_1 \le x^\alpha \int_0^{cx} \int_{\|\mathbf{z}'\mathbf{A}\| \le (x-\overline{\mathbf{b}})/x_0, (\mathbf{z}'\mathbf{a}^\#) \in E_\epsilon} \overline{\mathbf{b}} P\left(\|\mathbf{z}'\mathbf{a}\| \sqrt{d} \mathbf{z}_0' \mathbf{X} > x - \overline{\mathbf{b}}\right) F_{\mathbf{A},\overline{\mathbf{B}}}\big(d\mathbf{a}, d\overline{\mathbf{b}}\big)$$
$$=: x^\alpha K_{1,1}, \tag{5.23}$$

where $\mathbf{z}_0 = (1, \ldots, 1)/\sqrt{d} \in S_+$. The argument as above yields in the same way that

$$x^\alpha K_{1,1} \to r(\mathbf{z}_0) d^{\alpha/2} E\|\mathbf{z}'\mathbf{A}\|^\alpha 1_{\{(\mathbf{z}'\mathbf{A}^\#) \in E_\epsilon\}} \mathbf{z}'\mathbf{B}. \tag{5.24}$$

Using assumption in Eq. 5.13, since ϵ is arbitrarily small, we conclude from Eqs. 5.22, 5.23 and 5.24 that

$$x^\alpha K \to c_+ Er(\mathbf{z}'\mathbf{A}^\#) \|\mathbf{z}'\mathbf{A}\|^\alpha \mathbf{z}'\mathbf{B}. \tag{5.25}$$

The conclusion follows from Eqs. 5.18, 5.19 and 5.25. □

Example 5.1 Generalized autoregressive conditionally heteroscedastic process $\{\xi_t\}_{t\in\mathbb{Z}}$ of order (p,q) with $p,q \geq 0$ (GARCH(p,q)) is given as

$$\xi_t = \sigma_t \epsilon_t, \tag{5.26}$$

$$\sigma_t^2 = \beta + \sum_{i=1}^{p} \lambda_i \xi_{t-i}^2 + \sum_{j=1}^{q} \phi_j \sigma_{t-j}^2, \tag{5.27}$$

where $\{\epsilon_t\}$ are i.i.d. normal random variables, and $\beta > 0$, $\lambda_i \geq 0$, $\phi_i \geq 0$, with the convention that $\lambda_p > 0$ if $p \geq 1$ and $\phi_q > 0$ if $q \geq 1$. (See, for example, Bollerslev 1986 and Embrechts et al. 1997). The squares of the GARCH model can be expressed as a multidimensional RDE

$$\mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t, \tag{5.28}$$

where

$$\mathbf{X}_t = \left(\xi_t^2, \xi_{t-1}^2, \ldots, \xi_{t-p+2}^2, \xi_{t-p+1}^2, \sigma_t^2, \sigma_{t-1}^2, \ldots, \sigma_{t-q+2}^2, \sigma_{t-q+1}^2\right)',$$

$$\mathbf{A}_t = \begin{pmatrix} \lambda_1\epsilon_t^2 & \lambda_2\epsilon_t^2 & \ldots & \lambda_{p-1}\epsilon_t^2 & \lambda_p\epsilon_t^2 & \phi_1\epsilon_t^2 & \phi_2\epsilon_t^2 & \ldots & \phi_{q-1}\epsilon_t^2 & \phi_q\epsilon_t^2 \\ 1 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \ldots & 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_{p-1} & \lambda_p & \phi_1 & \phi_2 & \ldots & \phi_{q-1} & \phi_q \\ 0 & 0 & \ldots & 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix},$$

$$\mathbf{B}_t = \left(\beta\epsilon_t^2, 0, \ldots, 0, 0, \beta, 0, \ldots, 0, 0\right)'.$$

(The matrix $\mathbf{A}_t$ has to be interpreted with care when either p or q is zero. In this case, one should take $p = 1$ and $\lambda_1 = 0$ or $q = 1$ and $\phi_1 = 0$ respectively.) It can be seen that assumption in Eq. 5.12 or 5.13 holds for the squares of a GARCH process with continuous innovations ϵ_t.

Acknowledgements The authors would like to thank an anonymous referee for useful comments and suggestions. The authors also benefited from comments of and thank Prof. Laurens de Haan and Prof. Sid Resnick.

Appendix

A Geometric ergodicity of RDE

Let $\{X_n\}$ be given by RDE in Eq. 1.1 where (A_n, B_n) are i.i.d. vectors. Then $\{X_n\}$ is a Markov chain and we recall here that it is geometrically ergodic. The following definitions will be used.

Markov chain $\{Y_n\}$ in a general state space S equipped with σ-field $\mathcal{S}$ is called $\mu-$irreducible for some non-degenerate measure μ on $(S, \mathcal{S})$, if $\mu(A) > 0$ implies

$$\sum_{N=1}^{\infty} p_N(y, A) > 0 \quad \text{for all } y \in S,$$

where $p_N(y, A)$ is the $N-$step transition probability of Markov chain starting from y to A. The Markov chain $\{Y_n\}$ is said to be geometrically ergodic if there is $\rho \in (0, 1)$ and constant C_y for each $y \in S$ such that,

$$\|p_N(y, \cdot) - \pi(\cdot)\| := \sup_{A \in \mathcal{S}} \{|p_N(x, A) - \pi(A)|\} \leq C_y \rho^n, \tag{5.29}$$

where $\pi(\cdot)$ denotes the invariant measure of the Markov chain.

Theorem 5.3 (Basrak et al. 2002b, Stelzer 2009) *Suppose there exists an $\epsilon > 0$ such that $E|A_1|^\epsilon < 1$ and $E|B_1|^\epsilon < \infty$. If the Markov chain $\{X_n\}$ is μ-irreducible, then it is geometrically ergodic.*

The condition of $\mu-$irreducibility is satisfied for most models of practical interest. For example, for the squares of ARCH(1) series $X_t = \xi_t^2 = \lambda\epsilon_t^2\xi_{t-1}^2 + \beta\epsilon_t^2$, with $S = (0, \infty)$, $\mu =$ Lebesgue measure and $y > 0$, one obviously already has $p_1(y, A) = P((\lambda y + \beta)\epsilon_t^2 \in A)$ whenever $\mu(A) > 0$ and ϵ_t^2 has a density on $(0, \infty)$. For the existence of ϵ such that $E|A_1|^\epsilon < 1$, consider $h(p) =$

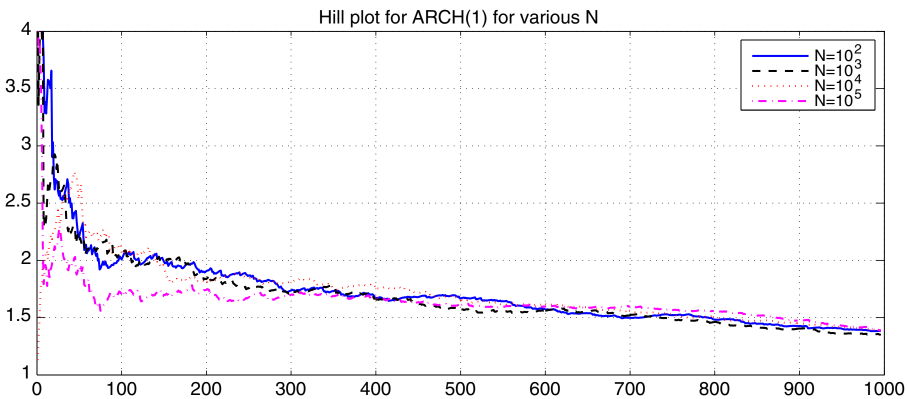


Fig. 14 Hill plot for ARCH(1) series with $\alpha = 10$ and $R = 5{,}000$ for various N.

$E|A_1|^p$. Note that $h(0) = 1$ and $h'(0) = E\log|A_1| < 0$ assuming conditions of Theorem 1.1. Hence, there is ϵ such that $E|A_1|^\epsilon < 1$. Observe that, for $\epsilon < \alpha$,

$$E|B_1|^\epsilon \le (E|B_1|^\alpha)^{\epsilon/\alpha} < \infty,$$

from the assumption of Theorem 1.1. Therefore, as a result of Theorem 5.3, P_N converges to P_∞ at an exponentially fast rate. Figure 14 shows Hill plot based on 5,000 independent observations of ARCH(1) model with various choices of N. There is no difference in Hill plot, which supports the claim that the convergence is quite fast.

B Tail exponent for multiplicative cascades with log-normal multipliers

For a log-normal multiplier $W = LN(-\sigma^2/2, \sigma^2)$,

$$\chi_2(h) = \log_2 EW^h - (h-1) = \log_2 \exp\left(-\frac{h\sigma^2}{2} + \frac{h^2\sigma^2}{2}\right) - (h-1)$$
$$= \frac{1}{\log 2}\left(-\frac{h\sigma^2}{2} + \frac{h^2\sigma^2}{2}\right) - (h-1).$$

To find the tail exponent, we need to set

$$\chi_2(h) = 0$$

and look for solution $h > 1$. This yields

$$(\sigma^2 h - 2\log 2)(h-1) = 0$$

or

$$\alpha = 2\log 2/\sigma^2 > 1. \tag{5.30}$$

References

Baek, C., Pipiras, V.: Estimation of parameters in heavy-tailed distribution when its second order tail exponent is known. Preprint. Available at http://www.stat.unc.edu/faculty/pipiras (2009)

Basrak, B., Davis, R.A., Mikosch, T.: A characterization of multivariate regular variation. Ann. Appl. Probab. **12**(3), 908–920 (2002a)

Basrak, B., Davis, R.A., Mikosch, T.: Regular variation of GARCH processes. Stoch. Process. Appl. **99**(1), 95–115 (2002b)

Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G.: Tail index estimation and an exponential regression model. Extremes **2**(2), 177–200 (1999)

Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation. In: Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge (1989)

Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. J. Econom. **31**(3), 307–327 (1986)

Breiman, L.: On some limit theorems similar to the arc-sin law. Theory Probab. Appl. **10**, 351–360 (1965)

Chamayou, J.-F., Letac, G.: Explicit stationary distributions for compositions of random functions and products of random matrices. J. Theoret. Probab. **4**(1), 3–36 (1991) de Haan, L., Ferreira, A.: Extreme Value Theory. Springer Series in Operations Research and Financial Engineering. Springer, New York (2006)

de Saporta, B., Guivarc'h, Y., Le Page, E.: On the multidimensional stochastic equation $Y_{n+1} = A_n Y_n + B_n$. Comptes Rendus Mathématique. Académie des Sciences. Paris **339**(7), 499–502 (2004)

Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extremal Events. Springer, Berlin (1997)

Feuerverger, A., Hall, P.: Estimating a tail exponent by modelling departure from a Pareto distribution. Ann. Stat. **27**(2), 760–781 (1999)

Goldie, C.M.: Implicit renewal theory and tails of solutions of random equations. Ann. Appl. Probab. **1**(1), 126–166 (1991)

Gomes, M., Ivette, d. H.L., Rodrigues, L.H.: Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. J. R. Stat. Soc. Ser. B **70**(Part 1), 31–52 (2008)

Grintsyavichyus, A.: On a random difference equation. Lith. Math. J. **21**(4), 302–305 (1981)

Guivarc'h, Y.: Sur une extension de la notion de loi semi-stable. Ann. Inst. Henri Poincaré B Calc. Probab. Stat. **26**(2), 261–285 (1990)

Guivarc'h, Y.: Heavy tail properties of stationary solutions of multidimensional stochastic recursions. In: IMS Lecture Notes Monogr. Ser., vol. 48, pp. 85–99. Institute of Mathematical Statistics, Beachwood (2006)

Kahane, J.-P., Peyrière, J.: Sur certaines martingales de Benoit Mandelbrot. Adv. Math. **22**(2), 131–145 (1976)

Kesten, H.: Random difference equations and renewal theory for products of random matrices. Acta Math. **131**, 207–248 (1973)

Klüppelberg, C., Pergamenchtchikov, S.: The tail of the stationary distribution of a random coefficient AR(q) model. Ann. Appl. Probab. **14**(2), 971–1005 (2004)

Liu, Q.: On generalized multiplicative cascades. Stoch. Process. Appl. **86**(2), 263–286 (2000)

Mandelbort, B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. J. Fluid Mech. **62**(2), 331–358 (1974)

Ossiander, M., Waymire, E.C.: Statistical estimation for multiplicative cascades. Ann. Stat. **28**(6), 1533–1560 (2000)

Peng, L.: Asymptotically unbiased estimators for the extreme-value index. Stat. Probab. Lett. **38**(2), 107–115 (1998)

Resnick, S.I.: Heavy tail modeling and teletraffic data. Ann. Stat. **25**(5), 1805–1869 (1997) (with discussion and a rejoinder by the author)

Stelzer, R.: On Markov-switching ARMA processes—stationarity, existence of moments and geometric ergodicity. Econom. Theory **25**, 43–62 (2009)