Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Econometrics 129 (2005) 299–327
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Renewal regime switching and stable limit laws

Remigijus Leipus[a,b,*], Vygantas Paulauskas[a,b], Donatas Surgailis[b]

[a] Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, 03225 Vilnius, Lithuania
[b] Vilnius Institute of Mathematics and Informatics, Vilnius, Lithuania

Abstract

The paper discusses long-memory properties and large sample behavior of partial sums in a general renewal regime switching scheme. The linear model $X_t = \mu_t + a_t X_{t-1} + \sigma_t \varepsilon_t$ with renewal switching in levels, slope or volatility and general (possibly heavy-tailed) i.i.d. noise ε_t is discussed in detail. Conditions on the tail behavior of interrenewal distribution and the tail index $\alpha \in (0, 2]$ of ε_t are obtained, in order that the partial sums process of X_t is asymptotically λ-stable with index $\lambda < \alpha$.

JEL classification: C22

Keywords: Regime switching; Renewal reward process; Stable Lévy process

1. Introduction

A widely used explanation of the long-memory phenomenon in economic and financial data is regime switching, where the duration of regime has a heavy tailed distribution. Empirical evidence of heavy tailed regime durations is discussed in Jensen and Liu (2001) (lengths of the US business cycle's), Chow and Liu (1999) (dividend series from the CRSP data), Liu (2000) (daily S&P composite price index).

*Corresponding author. Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, 03225 Vilnius, Lithuania. Tel.: 370 5 2336017.

E-mail address: remigijus.leipus@maf.vu.lt (R. Leipus).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2004.09.010

Jensen and Liu (2001), Gourieroux and Jasiak (2001) argue that regime switching with heavy tails may lead to a new forecasting methodology, as an alternative to ARFIMA forecasting. Various regime switching models leading to the long-memory property and related econometrical issues were discussed in Parke (1999), Granger and Hyung (2004), Diebold and Inoue (2001), Liu (2000), Jensen and Liu (2001), Gourieroux and Jasiak (2001), Leipus and Viano (2003). In particular, Liu (2000) noted that the Markov regime switching model of Hamilton (1989) with finite number of states has short memory. On the other hand, simple models with heavy tailed regime switching of mean are known to exhibit *covariance long memory*, in the sense that their autocovariance decays slowly with the lag as $t^{-(1-2d)}$, with some $0<d<1/2$, see Taqqu and Levy (1986), Liu (2000), Jensen and Liu (2001), Davidson and Sibbertsen (2002), Mikosch et al. (2002). Leipus and Surgailis (2003a) established a similar long-memory behavior of autocovariance of random coefficient AR(1) equation

$$X_t = a_t X_{t-1} + \varepsilon_t \tag{1.1}$$

with slope a_t performing a heavy tailed regime switching in the interval $[0, 1]$, including the unit root.

However, autocovariances may carry very limited information for statistical analysis, especially for hypotheses testing and estimation, which usually require an asymptotic theory for distributions. Furthermore, an approach based solely on autocovariances may lead to spurious inferences (Lobato and Savin, 1998). *Long-range dependence* (long memory) is often defined to be persistent in the distributional limit: a stationary time series X_t is said to be long-range dependent if its partial sums process, when suitably normalized, converges (in the sense of distribution) to some random process with dependent increments, see e.g., Cox (1984, p. 59), Dehling and Philipp (2002, p. 78). The main conclusion of the present paper is that *a large class of stationary models with heavy tailed regime switching exhibit an increase of variability and do not exhibit long memory in the distributional limit*. Namely, the limit of partial sums of X_t is a stable Lévy process $W_\lambda(\tau)$ which has infinite variance while X_t itself can have finite variance, and the stability index $0<\lambda<2$ of the limit process is strictly less than the tail index α of innovations ε_t in (1.3), see Theorems 2.1, 5.2 and 5.3. The limit process $W_\lambda(\tau)$ also has independent increments, which means that the long memory in X_t does not persist in the distributional limit. This fact should be contrasted with persistent long memory in d-integrated $(0<d<1/2)$ stationary processes, whose partial sums converge to a d-fractional Brownian motion with dependent increments (Davydov, 1970). The econometric implication of our result is that temporal aggregation of models with heavy tailed regime durations can lead to nonpersistent, although highly leptokurtic, behavior. A similar lack of persistency of long memory seems characteristic also to some other econometric models, in particular, to Parke's (1999) error duration model (see Davidson and Sibbertsen, 2002; Hsieh et al., 2003). See also Davydov (1973) for early probabilistic example of such behavior.

The class of regime switching models which exhibit the above behavior of partial sums seems to be very general. The main idea of our approach is the following. Let

$\cdots < S_{j-1} < S_j < \cdots$ be consecutive moments of regime switches, which follow a renewal process with a possibly heavy tailed but finite mean interrenewal distribution U. Under mild conditions, partial sums of the regime switching process X_t can be shown to behave similarly as partial sums of the aggregated process

$$Y_j = \sum_{S_{j-1} < t \leqslant S_j} X_t, \tag{1.2}$$

due to the fact that the number of renewal points in large interval $[1, n]$ is asymptotically proportional to n/μ, where $\mu = \mathrm{E}U$. We assume that the Y_j's can be written in the product form

$$Y_j = \Phi_j Z_j,$$

characteristic to stochastic volatility models, where $\Phi_j > 0$ ('aggregate volatility') depends on the current regime variables (duration, type) and the previous history, while Z_j ('aggregate innovation') is determined by the 'dynamics' of X_t between the regime changes. The precise assumptions on Φ_j, Z_j are given in Section 2 (Assumptions A_1–A_5). Intuitively, these assumptions say that, as the interval length $U_j = S_j - S_{j-1}$ increases, the r.v.'s Φ_j, Z_j become *independent* and tend in some sense to (independent) r.v.'s Φ_j^0, Z_j^0, respectively, where Φ_j^0 has a heavy tail with some $\lambda \in (0, 2)$, and Z_j^0 has a tail lighter than λ (in many cases, Z_j^0 is a standard normal variable). By the well-known Breiman's lemma (Breiman, 1965), these assumptions imply heavy tailedness of the product $\Phi_j^0 Z_j^0$, and a λ-stable limit distribution of the partial sums process.

The above set up is illustrated by considering particular cases of the autoregressive equation

$$X_t = \mu_t + a_t X_{t-1} + \sigma_t \varepsilon_t \tag{1.3}$$

with renewal switching in levels (μ_t), slope (a_t) and/or volatility (σ_t). The main attention is given to the changes in slope, or the model (1.1). Here, we extend the results of Leipus and Surgailis (2003a), by considering (i) more general (in particular, heavy tailed) noise ε_t, and (ii) a_t switching between 0 and some value $A > 1$. The regime corresponding to $a_t = A > 1$ can be characterized as exponential growth (or $\mathrm{I}(\infty)$ regime in the terminology of Granger (2000)) after which the process drops back into i.i.d. regime, so that a stationary solution of (1.1) may still exist. Such random coefficient AR(1) equation can describe periodically collapsible and restarting bubbles with variance which diverges to infinity exponentially in corresponding random intervals. The collapsible bubbles' model was first introduced in Blanchard (1979) and Blanchard and Watson (1982) for i.i.d. a_t taking two values 0 and $A > 1$. Tail behavior in this model was studied in Lux and Sornette (2002). Empirical evidence from the US and Hong Kong stock index data and testing procedures for the existence of bubbles are discussed in Wu and Xiao (2002).

Let us finally note that in the main Theorem 2.1 and its applications to (1.3), the 'switching mechanism', or duration distribution U, may have heavy tails, as in the case of slope a_t switching between 0 and $A = 1$, but also may have light (exponential) tails, as in the case of a_t switching between 0 and $A > 1$. The heavy λ-tails ($0 < \lambda < 2$)

in the partial sums limit arise essentially from λ-tails of 'aggregate volatility' Φ_j alone, which does not contradict condition $\mathrm{E}U < \infty$.

The plan of the paper is the following. In Section 2, we define renewal regime switching process and formulate the general result (Theorem 2.1) about λ-stable limit of partial sums. Sections 3 and 4 discuss application of Theorem 2.1 to renewal switching in levels and volatility, respectively. In particular, our regime switching volatility model is compared to Liu's (2000) model. Renewal regime switching in slope is discussed in detail in Section 5. Proofs are collected in Appendices A and B.

2. Renewal regime switching and a generalization of Breiman's lemma

By *regime switching process* we mean a stochastic process X_t ($t = 0, 1, \ldots$) whose evolution (conditional probability) at time t is determined by its past $X_0, \ldots, X_{t-1}$ and the value R_t of some (vector-valued) process, which is called *regime* and which changes its value at random moments. The times and values of regime switches can occur independently of the process X_t (such mechanism is considered in our paper), or can be dependent on past values of X_t (as in threshold models). A rather general class of regime switching processes is given by recurrent equations $X_t = f_t(X_0, \ldots, X_{t-1}; R_t; \varepsilon_t)$, where $f_t(x_0, \ldots, x_{t-1}; y; z)$ are some functions, and ε_t is a noise process; in the sequel, unless specified otherwise, ε_t will stand for i.i.d. noise independent of regime process. The econometric literature on regime switching models and their inference is quite large, see, e.g., Granger and Teräsvirta (1993); Tong (1990). For various regime switching specifications of model (1.3), see Franses and van Dijk (2000) and the references therein. A widely used regime switching scheme is the Markov switching model of Hamilton (1989), in which regime durations have light (exponential) tails. Some recent models involving heavy tailed switching mechanism, with applications to econometrics, are mentioned in Section 1.

Let us define more precisely a class of regime processes related to a renewal process. These are processes R_t, taking values in p-dimensional Euclidean space $\mathbb{R}^p$, which change their value randomly and independently at random times S_j of a given renewal process and then keep the value constant until the next renewal time. To give a formal definition, let $(U_1, \zeta_1), (U_2, \zeta_2), \ldots$ be a sequence of independent vectors, where $U_j = 1, 2, \ldots$ is the duration and $\zeta_j \in \mathbb{R}^p$ is the value of the jth subsequent regime. Moreover, we assume that random vectors (U_j, ζ_j), $j = 2, 3, \ldots$ follow a common distribution (U, ζ) with $\mu = \mathrm{E}U < \infty$. Let $S_0 = 0, S_j = U_1 + \cdots + U_j$.

Definition 2.1. We call a *renewal regime process* a stochastic process R_t, $t = 1, 2, \ldots$ such that $R_t = \zeta_j$ for $t \in (S_{j-1}, S_j]$, $j = 1, 2, \ldots$.

According to the above definition, regime switch times S_j constitute an integer-valued renewal process with interrenewal distribution U and initial distribution U_1. The distribution of U_1 is generally different from U; in the case of *stationary* renewal process it is given by

$$\mathrm{P}[U_1 = u] = \mu^{-1}\mathrm{P}[U \geqslant u], \quad u = 1, 2, \ldots .$$

The assumption of independence of the sequence (U_j, ζ_j) helps to avoid mathematical technicalities, although most of the results below are expected to hold under suitable weak dependence conditions on the sequence (U_j, ζ_j) as well. On the other hand, our assumptions allow for dependence between components U_j and ζ_j.

A regime switching process X_t, $t = 0, 1, \ldots$ corresponding to a renewal regime process R_t will be called a *renewal regime switching process*. Let $\mathscr{F}_t$ be the history σ-field which contains all information about X_s and R_s up to time $s = t$, and let $\mathscr{G}_{j-1} = \mathscr{F}_{S_{j-1}}$ be the history until the last regime change at random time $S_{j-1} + 1$. Denote $N_n = \max\{j : S_j \leqslant n\}$ the number of renewal points S_j in the interval $[1, n]$. Let Y_j be the sum of X_t's in the interval $(S_{j-1}, S_j]$ as defined in (1.2). All relations below, involving random variables, conditional probabilities and expectations, are supposed to hold almost surely (a.s.) and uniformly in $j \geqslant 1$.

Assumption $\mathbf{A_1}$. The sum Y_j in (1.2) can be represented as the product of two random variables:

$$Y_j = \Phi_j Z_j, \tag{2.1}$$

where $\Phi_j > 0$ is a function of the current regime variables (U_j, ζ_j) and the past history $\mathscr{G}_{j-1}$ (in other words, Φ_j is measurable w.r.t. the σ-field $\sigma\{U_j, \zeta_j, \mathscr{G}_{j-1}\}$).

The representation (2.1) is crucial for our discussion. As was noted in Section 1, the intuitive meaning of Z_j is 'aggregate innovation' (i.e., the 'innovation' of the aggregated process Y_j in (1.2)) and Φ_j as 'aggregate volatility', the latter being completely determined by the current regime (its duration, type) and the previous history up to time S_{j-1}. Representation (2.1) is obviously not unique; a natural choice of Φ_j, at least in the case when Y_j has finite conditional variance w.r.t. $\sigma\{U_j, \zeta_j, \mathscr{G}_{j-1}\}$, is the conditional standard deviation:

$$\Phi_j = \mathrm{Var}^{1/2}[Y_j | U_j, \zeta_j, \mathscr{G}_{j-1}]. \tag{2.2}$$

If the conditional law $[Y_j | U_j, \zeta_j, \mathscr{G}_{j-1}]$ is centered Gaussian, then $Z_j = \Phi_j^{-1} Y_j \sim N(0,1)$, implying that Y_j of (2.1) is a conditionally heteroskedastic series with i.i.d. Gaussian innovations Z_j and (heavy-tailed) volatility Φ_j, the heavy-tailedness being a consequence of Assumption A_2 below. In some cases, Φ_j is a simple function of the current regime variables alone, such as $\Phi_j = U_j$ in the switching mean example of Section 3, or $\Phi_j = \zeta_j$ in the volatility example of Section 4. See also (5.28), (5.31) for simple expressions of Φ_j in the case of slope switching between 0 and some nonrandom $A \geqslant 1$.

Assumption $\mathbf{A_2}$. There exist (nonrandom) constants $0 < \lambda < 2$, $c_0 > 0$, $C > 0$, a (nonrandom) function $h(v) \to 0$ $(v \to 0)$ and a r.v. $\Phi^0 > 0$ such that

$$\mathrm{P}[\Phi^0 > u] \sim c_0 u^{-\lambda} \ (u \to \infty), \qquad \mathrm{P}[\Phi_j > u | \mathscr{G}_{j-1}] \leqslant C u^{-\lambda} \quad (\forall u > 0) \tag{2.3}$$

and such that

$$\left|\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j}|\mathcal{G}_{j-1}] - \mathrm{E}\mathrm{e}^{\mathrm{i}v\Phi^0}\right| \leqslant |v|^{\lambda} h(v). \tag{2.4}$$

The first relation of (2.3) implies that the distribution of Φ^0 belongs to the domain of attraction of a totally skewed to the right λ-stable law (i.e., a stable law with stability parameter λ and skewness parameter equal to 1; see Samorodnitsky and Taqqu (1994, p. 13)). Together with (2.4), it implies a similar tail relation for the *conditional* distribution:

$$\mathrm{P}[\Phi_j > u|\mathcal{G}_{j-1}] \sim c_0 u^{-\lambda} \quad (u \to \infty), \tag{2.5}$$

with the same nonrandom $c_0 > 0$ as in (2.3). While condition (2.5) seems close to (2.3)–(2.4), the latter conditions are technically more convenient for proving limit theorems.

Assumption $\mathbf{A_3}$. There exists a r.v. Z^0 such that for any fixed $K > 0$

$$\sup_{|u| \leqslant K} \left|\mathrm{E}[\mathrm{e}^{\mathrm{i}uZ_j}|\Phi_j, \mathcal{G}_{j-1}] - \mathrm{E}\mathrm{e}^{\mathrm{i}uZ^0}\right| \leqslant \delta_K(\Phi_j), \tag{2.6}$$

where $\delta_K(u)$ is a nonrandom function such that $\delta_K(u) \to 0$ $(u \to \infty)$.

Assumption A_3 is equivalent to weak convergence of the conditional distribution $\mathrm{P}[Z_j \leqslant x|\Phi_j, \mathcal{G}_{j-1}]$ to the distribution $\mathrm{P}[Z^0 \leqslant x]$ as $\Phi_j \to \infty$. Typically, $\Phi_j \to \infty$ implies $U_j \to \infty$ and vice versa, so that (2.6) says that the distribution of $Z_j = \Phi_j^{-1} Y_j$ tends to some distribution Z^0 independent of $\mathcal{G}_{j-1}$ as the interval length U_j increases. Under conditional Gaussianity of the law$[Y_j|U_j, \zeta_j, \mathcal{G}_{j-1}]$ and the choice (2.2) of Φ_j, relation (2.6) is obviously satisfied with $Z^0 \sim N(0,1)$ and $\delta_K(u) \equiv 0$. 'Aggregate innovation' Z_j being often a normalized sum of random variables, A_3 may also entail some form of central limit theorem as the interval length $U_j \to \infty$, in which case Z^0 again may be a normal or stable r.v. However, A_3 also applies to the situations as in Theorem 5.3 (slope switching above the unit root), where Z^0, differently from above, is given by infinite geometric series of noise variables (see (5.30) below). Let us finally note that a similar but stronger version of A_3 is given by the uniform bound

$$\sup_{x \in \mathbb{R}} |\mathrm{P}[Z_j \leqslant x|\Phi_j, \mathcal{G}_{j-1}] - \mathrm{P}[Z^0 \leqslant x]| \leqslant \delta(\Phi_j), \tag{2.7}$$

where $\delta(u)$ is a nonrandom function such that $\delta(u) \to 0$ as $u \to \infty$.

Assumption $\mathbf{A_4}$. There exist $r > \lambda$ and a (nonrandom) constant $C_0 < \infty$ such that

$$\mathrm{E}[|Z_j|^r|\Phi_j, \mathcal{G}_{j-1}] + \mathrm{E}|Z^0|^r \leqslant C_0.$$

Moreover, if $\lambda \geqslant 1$ then $\mathrm{E}[Z_j|\Phi_j, \mathcal{G}_{j-1}] = \mathrm{E}Z^0 = 0$.

Assumption A_4 implies that Z_j have lighter conditional tails as Φ_j. The zero conditional expectation condition is consistent with the 'aggregate innovation' interpretation of Z_j, and implies the martingale difference property of Y_j (2.1) when

$\lambda\geqslant 1$. In particular, if $\lambda\geqslant 1$ and $r\geqslant 2$, A_3–A_4 imply that Z_j, $j\geqslant 1$ are uncorrelated, with zero mean and variance uniformly bounded in j.

Assumption A_5. $\sum_{t=1}^{n} X_t - \sum_{j=1}^{N_n} Y_j = o_p(n^{1/\lambda})$, as $n\to\infty$.

Assumption A_5 is the most intuitive of A_1–A_5. By the definition of Y_j, the difference of the two sums in A_5 equals $\sum_{S_{N_n}<t\leqslant n} X_t$, where the number of summands $n-S_{N_n}=O_p(1)$, see Feller (1971). Therefore this difference is also bounded in probability under weak additional assumptions on the sequence X_t. In particular, A_5 holds if $\sup_{j\geqslant 1}\mathrm{E}|X_j|^{\delta}<\infty$ for some $\delta>0$; see the proof of Corollary 3.1 in Appendix B.

Introduce a Lévy process $W_\lambda(\tau)$, $\tau\geqslant 0$ with independent and stationary increments and the characteristic function

$$\mathrm{E}e^{\mathrm{i}aW_\lambda(\tau)} = \exp\{-\tau\mu^{-1}|a|^{\lambda}\omega(a/|a|)\},\quad a\in\mathbb{R},$$

where $\omega(z)=\frac{\Gamma(2-\lambda)}{\lambda|\lambda-1|}\left((c_+ + c_-)\cos(\frac{\pi\lambda}{2})+\mathrm{i}\,\mathrm{sgn}(z)\sin(\frac{\pi\lambda}{2})(c_+ - c_-)\right)$, $\mathrm{i}=\sqrt{-1}$, and where

$$c_+ = c_0\mathrm{E}|Z^0|^{\lambda}I(Z^0>0),\quad c_- = c_0\mathrm{E}|Z^0|^{\lambda}I(Z^0<0). \tag{2.8}$$

Write $\to_{\mathrm{fdd}}$ for weak convergence of finite dimensional distributions.

Theorem 2.1. *Let X_t be a renewal regime switching process satisfying Assumptions A_1–A_5, $0<\lambda<2$. Then*

$$\left\{n^{-1/\lambda}\sum_{s=1}^{[n\tau]} X_s,\ \tau\geqslant 0\right\}\to_{\mathrm{fdd}}\{W_\lambda(\tau),\ \tau\geqslant 0\}. \tag{2.9}$$

As noted in the Introduction, A_1–A_5 help to reduce the proof of (2.9) to the convergence

$$\left\{n^{-1/\lambda}\sum_{j=1}^{[n\tau/\mu]} \Phi_j^0 Z_j^0,\ \tau\geqslant 0\right\}\to_{\mathrm{fdd}}\{W_\lambda(\tau),\tau\geqslant 0\}, \tag{2.10}$$

where Φ_j^0, $j\geqslant 1$ and Z_j^0, $j\geqslant 1$ are both i.i.d. sequences, also independent of each other, Z_j^0 being a copy of Z^0, and Φ_j^0 a copy of Φ^0. Relation (2.10) follows by the central limit theorem for i.i.d. summands $Y_j^0=\Phi_j^0 Z_j^0$, provided their distribution belongs to the domain of attraction of λ-stable law with characteristic function $e^{-|a|^{\lambda}\omega(a/|a|)}$. The last fact follows from A_2 to A_4 and the classical Breiman's lemma (Lemma A.1 below) about tail behavior of the product of two *independent* random variables. Therefore, Theorem 2.1 can be considered as a generalization of Breiman's lemma for *dependent* random variables.

Remark 2.1. A natural question in the context of Theorem 2.1 concerns functional convergence in the Skorokhod space $D[0,1]$, the limit process $W_\lambda(\tau)$ being a.s. discontinuous on $[0,1]$. It is well-known that the convergence in (2.10) for the approximating i.i.d. sequence $\Phi_j^0 Z_j^0$ extends to the convergence in $D[0,1]$ (Skorokhod, 1964). As noted by a referee, a stronger approximation assumption, viz.

$$\sup_{\tau\in[0,1]} n^{-1/\lambda}\left|\sum_{s=1}^{[n\tau]} X_s - \sum_{j=1}^{[n\tau/\mu]} \Phi_j^0 Z_j^0\right| = \mathrm{o_p}(1), \tag{2.11}$$

together with A_1–A_5, imply the functional convergence in Theorem 2.1, too. On the other hand, assumption (2.11) seems to be quite restrictive and needs further investigation in concrete cases. It is also known that the functional convergence in the commonly used Skorokhod J_1-topology does not hold for some simple mean switching models, see Mikosch et al. (2002, p. 33, 40), Pipiras et al. (2004).

3. Renewal regime switching in levels

Consider the simplest stochastic regime switching model

$$X_t = \mu_t + \varepsilon_t, \tag{3.1}$$

where ε_t is a zero mean stationary process and μ_t is a randomly switching mean. The processes μ_t and ε_t are usually assumed independent. We assume that the mean process μ_t is a stationary renewal reward process, i.e.

$$\mu_t = \zeta_j, \quad S_{j-1} < t \leqslant S_j, \tag{3.2}$$

where S_j is a stationary renewal process with interrenewal distribution U, $\mu = \mathrm{E}U < \infty$, and ζ_j are i.i.d. random variables, independent of the renewal process S_j. By the independence of μ_t and ε_t,

$$\mathrm{Cov}(X_0, X_t) = \mathrm{Var}(\zeta) p_t + \mathrm{Cov}(\varepsilon_0, \varepsilon_t),$$

where $p_t = \mathrm{P}[S_j \notin (0,t)\ \forall j]$ is the probability that the interval $(0,t)$ is void of renewal points. It is well-known that for a stationary renewal process, this probability is given by $p_t = \mu^{-1}\sum_{u=t}^{\infty} \mathrm{P}[U \geqslant u]$. If the tail distribution of U decays as in (3.4) below, with $\lambda > 1$, the probability p_t decays as $t^{1-\lambda}$ and hence the autocovariance functions of μ_t and X_t are nonsummable for $1 < \lambda < 2$. More precisely, assuming that the autocovariance of ε_t decays as $\mathrm{o}(t^{1-\lambda})$, we obtain

$$\mathrm{Cov}(X_0, X_t) \sim \mathrm{Cov}(\mu_0, \mu_t) \sim c_2 t^{1-\lambda}, \quad c_2 = c_1 \mathrm{Var}(\zeta)/\mu. \tag{3.3}$$

This means that both processes μ_t and X_t have covariance long memory. Related results can be found in Liu (2000), Jensen and Liu (2001), Davidson and Sibbertsen (2002).

Asymptotic behavior of partial sums in the renewal mean switching model (3.1) and in some related models was discussed in Taqqu and Levy (1986), Mikosch et al. (2002), Pipiras et al. (2004), Davidson and Sibbertsen (2002) and other papers. The main emphasis of these studies is *aggregation*, or the possibility of obtaining Gaussian long-memory process as the limit of an aggregated sum of independent copies of (3.1). According to the popular idea of Granger (1980), this provides a possible explanation of observed long-memory property in economic time series.

If the process ε_t is covariance stationary and short memory, in the sense that its covariance function is absolutely summable, then $\sum_{t=1}^{n} \varepsilon_t = \mathrm{O_p}(n^{1/2})$. The behavior of partial sums of X_t in (3.1) in the long-memory case is thus determined by the behavior of sums of μ_t, or i.i.d. random variables $Y_j = U_j\zeta_j$. The product form of Y_j suggests using Theorem 2.1 with $\Phi_j = U_j, Z_j = \zeta_j, \Phi^0 = U, Z^0 = \zeta$. A short proof of the following corollary is given in Appendix B.

Corollary 3.1. *Let X_t be a renewal mean switching process defined in* (3.1), (3.2), *with interrenewal distribution U satisfying*

$$\mathrm{P}[U > u] \sim c_1 u^{-\lambda} \quad (u \to \infty) \tag{3.4}$$

for some $c_1 > 0$, $1 < \lambda < 2$, and ζ satisfying $\mathrm{E}\zeta = 0$, $\mathrm{E}|\zeta|^r < \infty$ for some $r > \lambda$. Suppose ε_t is a stationary zero mean process whose autocovariance function is absolutely summable. Then the convergence (2.9) *holds, where constants $c_+ = c_1 \mathrm{E}|\zeta|^\lambda I(\zeta > 0)$, $c_- = c_1 \mathrm{E}|\zeta|^\lambda I(\zeta < 0)$.*

A generalization of model (3.1), (3.2) is discussed in Davidson and Sibbertsen (2002), where (U_j, ζ_j) form a generally dependent stationary sequence. In the case when this sequence is i.i.d. (with ζ_j not necessarily independent of U_j) and the U_j's are heavy tailed, they prove a similar result to our Corollary 3.1.

Intuitively, the fact that the limit process in Corollary 3.1 has jumps, can be explained as follows. Consider the simplest case of mean switching between two values ± 1 with equal probabilities $P[\zeta_j = \pm 1] = 1/2$, and let $\varepsilon_t = 0$. Then, the integrated process $\sum_{t=1}^{k} X_t = \sum_{t=1}^{k} \mu_t$ is a 'broken line' with slope ± 1 on intervals $(S_{j-1}, S_j]$ where $\zeta_j = \pm 1$. The rescaled partial sums process $n^{-1/\lambda} \sum_{t=1}^{[n\tau]} X_t$ is a similar 'broken line' but with slope $\pm n^{1-1/\lambda} \to \pm\infty$ on corresponding random intervals $(S_{j-1}/n, S_j/n]$. Because of (3.4), almost all of these intervals have length $\mathrm{O}(n^{-1})$ but a few 'long' intervals have typical length $\mathrm{O}(n^{1/\lambda - 1})$, see Embrechts et al. (1997, Chapter 8.6), which still tends to zero as $\lambda > 1$. The increment of the partial sums process on such 'long' interval is proportional to $n^{1-1/\lambda} n^{1/\lambda - 1} = 1$; in other words, this increment does not vanish in the limit $n \to \infty$ but instead becomes a jump in the trajectory of the limiting process $W_\lambda(\tau)$.

4. Renewal regime switching in volatility

Let us discuss regime switching in volatility, or

$$X_t = \sigma_t \varepsilon_t, \tag{4.1}$$

where ε_t is a stationary process, and $\sigma_t > 0$ ('volatility') is a regime process independent of ε_t. To simplify our discussion, we shall consider the case when ε_t is i.i.d. noise, with generic distribution ε.

Let σ_t be a stationary renewal reward process similar to (3.2):

$$\sigma_t = \zeta_j, \quad S_{j-1} < t \leqslant S_j, \tag{4.2}$$

where S_j, ζ_j satisfy the same conditions as in (3.2), with the only difference that now $\zeta_j > 0$ a.s.

The model (4.1) was discussed in Liu (2000) in the finite variance case $\mathrm{E}\zeta^2 < \infty, \mathrm{E}\varepsilon^2 < \infty, \mathrm{E}\varepsilon = 0$ and a heavy tailed duration distribution U. Let

$$\mathrm{P}[U > u] \sim c_1 u^{-\beta} \quad (u \to \infty) \tag{4.3}$$

for some $c_1 > 0$, $\beta > 1$. Exactly as in (3.3), in this case one has $\mathrm{Cov}(\sigma_0, \sigma_t) \sim c_2 t^{1-\beta}$, so that for $1 < \beta < 2$ (and $\mathrm{E}\varepsilon^2 < \infty, \mathrm{E}\zeta^2 < \infty$) the stochastic volatility model (4.1) has covariance long memory. A similar conclusion is given in Liu (2000, Theorem 2.1). Under the same conditions, Liu (2000, Theorem 2.2) proved that partial sums process of X_t in (4.1) converges to a Brownian motion, under standard normalization $n^{1/2}$.

It is interesting to compare this result with our Theorem 2.1. If we put $\Phi_j = \zeta_j$, $\Phi^0 = \zeta$, $Z_j = \sum_{S_{j-1} < t \leqslant S_j} \varepsilon_t$, Assumption A$_3$ holds with $Z^0 = \sum_{i=1}^{U} \varepsilon_i, \mathrm{E}(Z^0)^2 = \mathrm{E}U\mathrm{E}\varepsilon^2 < \infty$ but A$_2$ cannot hold with $\lambda < 2$, as $\mathrm{E}\Phi^2 = \mathrm{E}\zeta^2 < \infty$.

The following Corollary 4.1 complements the results in Liu (2000), by considering the case of *infinite variance* stochastic volatility (4.2). Let

$$\mathrm{P}[\zeta > z] \sim c_2 z^{-\lambda} \quad (z \to \infty) \tag{4.4}$$

for some $0 < \lambda < 2$, $c_2 > 0$. By stationarity of the renewal process, the distribution of σ_t coincides with ζ and therefore (4.4) implies $\mathrm{E}\sigma_t^2 = \infty$. In view of (4.1), (4.2), Theorem 2.1 naturally applies with

$$\Phi_j = \zeta_j, \quad Z_j = \sum_{S_{j-1} < t \leqslant S_j} \varepsilon_t. \tag{4.5}$$

Corollary 4.1. *Let X_t be the stochastic volatility model of* (4.1), *where ζ_j satisfy* (4.4), *with some $0 < \lambda < 2$, and where ε_t are i.i.d.*, $\mathrm{E}|\varepsilon|^r < \infty$ *for some $r > \lambda$ and* $\mathrm{E}\varepsilon = 0$ *whenever $\lambda \geqslant 1$. Let U, ζ be independent*, $\mathrm{E}U < \infty$. *Then the convergence* (2.9) *holds.*

See Appendix B for the proof of the above result. Note that it does not require heavy-tailedness of U nor any other condition on U except $\mu = \mathrm{E}U < \infty$. In this sense, Corollary 4.1 is not related to long memory in stochastic volatility. Of course, if we assume long-tailedness of U as in (4.3), the infinite variance volatility model of Corollary 4.1 will display long memory, in the sense that power series $|X_t|^\delta = \sigma_t^\delta |\varepsilon_t|^\delta$ will have autocorrelations $\mathrm{O}(t^{1-\beta})$ decaying as in Liu's model, for any $\delta > 0$ such that $\mathrm{E}\zeta^{2\delta}\mathrm{E}|\varepsilon|^{2\delta} < \infty$. See also Liu (2000, p. 149).

The fact that duration distribution U has no effect (except for the mean $\mathrm{E}U$) on the limit distribution $W_\lambda(\tau)$ is in contrast with the results of Sections 3 and 5. It this sense, models with regime switching in volatility seem to be different from models with switching of mean or slope. The same lack of effect of U on the limit distribution occurs also in the finite variance case studied by Liu (2000). As Liu says on p.149: '... regardless of regime switching and even in a quite peculiar way, we still have Brownian motion as our limit instead of any jump process...'

Intuitively, the absence of jumps in the limit process of the volatility model with finite variance can be explained by a similar reasoning as their presence in the switching mean model of Section 3. Consider the case of volatility switching between two values $0<\sigma_-<\sigma_+$. The process $\sum_{t=1}^{k} X_t = \sum_{t=1}^{k} \sigma_t \varepsilon_t$ on each interval $k \in (S_{j-1}, S_j]$ is a random walk with zero mean and respective variance $\sigma_\pm^2$. Therefore $n^{-1/2}\sum_{t=1}^{[n\tau]} X_t$ behaves as random walk normalized by $n^{-1/2}$ and its increment on random interval $(S_{j-1}/n, S_j/n]$ of length $\mathrm{O}(n^{1/\beta-1})$ is of vanishing magnitude $\mathrm{O}((n^{1/\beta-1})^{1/2}) = \mathrm{o}(1)$, for any $\beta>1$. In other words, even 'long' durations $U_j = \mathrm{O}(n^{1/\beta})$ between consecutive switches of the volatility cannot produce a jump in the limit $n \to \infty$.

Let us note, finally, that the above reasoning does not apply to the infinite variance volatility switching model discussed in Corollary 4.1. In that case, jumps in the limit λ-stable process arise from occasional 'large' values $\mathrm{O}(n^{1/\lambda})$ of ζ_j's, similarly as in the classical central limit theorem for sums of i.i.d. r.v.'s.

5. Renewal regime switching in slope

One of the most interesting cases of stochastic regime switching concerns the slope coefficient a_t in AR(1) model (1.1). General properties of AR(1) equation with random and/or time-dependent coefficient can be found in Vervaat (1979), Tjøstheim (1986), Brandt (1986), Karlsen (1990), among others. Tong (1990) discusses various regime switching time series models including (1.1), with a special emphasis on threshold models. According to the so-called 'threshold principle', regimes are naturally introduced via thresholds, e.g., in the simplest SETAR(1) model, $a_t \equiv a^{(s_t)}$, where $s_t = j$ whenever $X_{t-1} \in S^{(j)}$ with $(S^{(j)})$ constituting some partition of $\mathbb{R}$. In the Markov switching regime model, s_t is an outcome of an (unobserved) finite-state Markov chain independent of ε_t, see, e.g., Hamilton (1994, Chapter 22).

Long-memory properties and asymptotic behavior of partial sums for renewal regime switching in slope was recently studied in Leipus and Surgailis (2003a), Leipus et al. (2004). Below, we extend these results and discuss these questions in the context of Theorem 2.1, in particular, the verification of Assumptions A_1–A_5.

Consider the equation

$$X_t = a_t X_{t-1} + \varepsilon_t, \tag{5.1}$$

where $\varepsilon_t, t \in \mathbb{Z}$ are i.i.d. innovations, and $a_t, t \in \mathbb{Z}$ is a strictly stationary ergodic process, independent of $\varepsilon_t, t \in \mathbb{Z}$. A stationary solution of (5.1) is given by the infinite series

$$\begin{aligned} X_t &= \varepsilon_t + a_t\varepsilon_{t-1} + a_t a_{t-1}\varepsilon_{t-2} + \cdots \\ &= \varepsilon_t + \sum_{s<t} \varepsilon_s \prod_{s<u\leqslant t} a_u. \end{aligned} \tag{5.2}$$

According to Brandt (1986), the series (5.2) converges in probability if conditions

$$\mathrm{E}\log|a_0|<0, \quad \mathrm{E}\log_+|\varepsilon_0|<\infty, \tag{5.3}$$

are satisfied, where $\log_+ x = \log(x \vee 1)$. Note the first condition of (5.3) is satisfied if either $P[a_0 = 0] > 0$, or $P[|a_0| \leqslant 1] = 1$ and $P[|a_0| < 1] > 0$. The value $a_t = a$ of the slope coefficient determines the current regime of the process X_t, i.e., in the terminology of Section 2, $a_t = R_t$ is the regime process. Correspondingly, one can have three types of behavior: (1) stationarity, or I(0) regime $0 \leqslant a < 1$, (2) random walk, or I(1) regime $a = 1$, and (3) exponential growth, or I(∞) regime $a > 1$.

Note that Theorem 2.1 does not directly apply to X_t of (5.2), as the conditional expectation $\mathrm{E}[\mathrm{Y}_j|\Phi_j, \mathscr{G}_{j-1}] = \sum_{s \leqslant S_{j-1}} \varepsilon_s \mathrm{E}\left[\sum_{S_{j-1} < t \leqslant S_j} \mathrm{a}_t \cdots \mathrm{a}_{s+1} | \Phi_j, \mathscr{G}_{j-1}\right] \neq 0$ in general. On the other hand, A_1 and A_4 imply $\mathrm{E}[Y_j|\Phi_j, \mathscr{G}_{j-1}] = \mathrm{E}[\Phi_j^{-1} Z_j|\Phi_j, \mathscr{G}_{j-1}] = \Phi_j^{-1}\mathrm{E}[Z_j|\Phi_j, \mathscr{G}_{j-1}] = 0$, for $\lambda \geqslant 1$. Therefore X_t of (5.2) need to be centered by corresponding conditional expectations. For $S_{j-1} < t \leqslant S_j$, let $X_t^1 = \mathrm{E}[X_t|\mathscr{G}_{j-1}, U_j, \zeta_j]$, $X_t^0 = X_t - \mathrm{E}[X_t|\mathscr{G}_{j-1}, U_j, \zeta_j]$. Clearly, the convergence (2.9) follows from

$$n^{-1/\lambda} \sum_{t=1}^{[n\tau]} X_t^0 \rightarrow_{\mathrm{fdd}} W_\lambda(\tau), \quad \sum_{t=1}^{n} X_t^1 = \mathrm{o_p}(n^{1/\lambda}). \tag{5.4}$$

Theorem 2.1 can be used to prove the first relation in (5.4) while the second one needs additional argument. Consider the representation (2.1) of 'centered aggregates' $Y_j^0 = \sum_{S_{j-1} < t \leqslant S_j} X_t^0$. Note

$$Y_j^0 = \sum_{S_{j-1} < s \leqslant S_j} \varepsilon_s \sum_{s \leqslant t \leqslant S_j} \zeta_j^{t-s} \tag{5.5}$$

is a weighted sum of random number $U_j = S_j - S_{j-1}$ of i.i.d. r.v.'s $\varepsilon_s, S_{j-1} < s \leqslant S_j$ with random weights $\sum_{s \leqslant t \leqslant S_j} \zeta_j^{t-s}$ depending on the current regime $a_t = \zeta_j$. According to our definition of renewal regime process, ζ_j, U_j are independent and therefore the distribution of Y_j^0 is completely determined by generic distributions ε, ζ and U. The choice of the representation $Y_j^0 = \Phi_j Z_j$ depends on tail properties of ε. In the finite variance case $\sigma^2 = \mathrm{E}\varepsilon^2 < \infty$, $\mathrm{E}\varepsilon = 0$, let

$$\Phi_j = \mathrm{Var}^{1/2}[Y_j^0|U_j, \zeta_j, \mathscr{G}_{j-1}], \quad Z_j = \Phi_j^{-1} Y_j^0, \tag{5.6}$$

as in (2.2). Note Y_j^0 and Φ_j are independent of $\mathscr{G}_{j-1}$ and

$$\Phi_j^2 = \sigma^2 \sum_{s=1}^{U_j} (1 + \zeta_j + \cdots + \zeta_j^{U_j - s})^2$$

depends on the distributions ζ and U only. Note that at $\zeta_j = 1$ (the unit root), this conditional variance grows as U_j^3. Indeed, $\Phi_j^2 = \sigma^2 \sum_{s=1}^{U_j} (U_j - s + 1)^2 = \sigma^2(U_j(U_j + 1)(2U_j + 1)/6) \sim (\sigma^2/3)U_j^3$, implying $P[\Phi_j > u] \sim P[\zeta_1 = 1]P[U_j > (3u/\sigma^2)^{2/3}] \sim c_0 u^{-\lambda}$ with $\lambda = 2\beta/3$, in the case of (5.2) switching between I(0) and I(1) regimes and duration distribution $P[U_j > u] \sim c_1 u^{-\beta}$ (see (4.3)). In some other situations and especially for nonlinear models, determining the tail index λ of 'aggregate volatility' Φ_j is not so obvious and may present technical difficulty. Another technical problem is the control of centering conditional expectations X_t^1 for proving the second relation in (5.4). Leipus and Surgailis (2003a) show that X_t^1 actually have short memory, in the sense that $\sum_{t=0}^{\infty} |\mathrm{E}X_0^1 X_t^1| < \infty$, and therefore $\sum_{t=1}^{n} X_t^1 = \mathrm{O_p}(n^{1/2}) = \mathrm{o_p}(n^{1/\lambda})$ as $\lambda < 2$.

The above mentioned technical difficulties do not arise in the simplest situation when a_t switches between two values: value $a = 0$ and some deterministic value $A \geqslant 1$. At the moment when a_t assumes value 0, the process X_t drops into i.i.d. regime and 'forgets' all previous history. We call such switching mechanism *memoryless regime switching*. In this case, sums of X_t between consecutive moments of $a_t = 0$ are conditionally independent random variables, similarly as Y_j^0 in (5.5).

Let us precise what we mean by memoryless regime switching. Let

$$\cdots < S_{j-1} < S_j^0 < S_j < S_{j+1}^0 < S_{j+1} < \cdots \tag{5.7}$$

be an alternating stationary process of successive switching times of the slope coefficient:

$$a_t = \begin{cases} A & S_{j-1} < t \leqslant S_j^0, \\ 0 & S_j^0 < t \leqslant S_j. \end{cases} \tag{5.8}$$

The corresponding regime durations will be denoted by

$$U_j^1 = S_j^0 - S_{j-1}, \quad U_j^0 = S_j - S_j^0. \tag{5.9}$$

Eq. (5.8) implies that X_t stays in i.i.d. regime during time interval $[S_j^0 + 1, S_j]$. Below we assume that (U_j^0, U_j^1), $j \in \mathbb{Z}$ *is a sequence of i.i.d. random vectors*, while the components U_j^0, U_j^1 themselves may be mutually independent or dependent. This alternating regime process fits into Definition 2.1 by putting $U_j = U_j^0 + U_j^1$, $\zeta_j = U_j^0$, in which case regime between successive moments S_j is specified by duration U_j^0, and $U_j = S_j - S_{j-1}$ are i.i.d.

In a realistic model, U_j^1 and U_j^0 could have different distributions and/or probability tails, because they correspond to different economic situations: the first one to a period of high economic activity ('wild fluctuations in stock market'), and the second one to a usual 'stabilization' period. It is quite common that the lengths of these periods are correlated between themselves: after a long period of high activity, one should expect a longer period of stabilization and vice versa. One of the simplest cases of memoryless regime switching in slope is the Blanchard's bubbles' model mentioned in Section 1. In this model, a_t are given by

$$a_t = Ab_t, \tag{5.10}$$

where b_t is (i.i.d.) Bernoulli process taking value 1 with probability π and value 0 with probability $1 - \pi$. In such case, generic durations U^0 and U^1 are mutually independent and geometrically distributed with parameters π and $1 - \pi$, respectively. A generalization of (5.10) given by a stationary Markov chain taking two values 0 and A also yields independent and geometrically distributed U^0 and U^1, see Example 5.1 below.

In the engineering literature, the model a_t (5.8) with independent durations U^0, U^1 and $A = 1$ is commonly referred to as an *on/off process*; see Willinger et al. (1997), Heath et al. (1998) and the references therein. It models the evolution of an idealized source which produces data at a constant rate in *on* state and produces no data in *off* state. It is argued that both *on* and *off* times are well modelled by heavy tailed distributions. In such case, Heath et al. (1998) obtain the precise long-memory decay of the covariance function of on/off process, using advanced renewal theory methods. Similar result is obtained in Jensen and Liu (2001). Our model (5.8) generalizes on/off process by allowing consecutive *on* and *off* durations to be mutually dependent. Asymptotic decay of the autocovariance of (5.8) is discussed in Leipus and Surgailis (2003b). It is easy to see that marginal probabilities of stationary process (5.8) are given by

$$\mathrm{P}[a_t = A] = \mu_1/\mu, \quad \mathrm{P}[a_t = 0] = \mu_0/\mu, \tag{5.11}$$

where $\mu_i = \mathrm{E}U^i$, $i = 0, 1$ and $\mu = \mu_0 + \mu_1$.

Let us turn to the properties of solution of (5.1) with a_t defined by (5.8). In such case, (5.2) becomes

$$X_t = \begin{cases} \varepsilon_t + A\varepsilon_{t-1} + \cdots + A^{t-S_{j-1}}\varepsilon_{S_{j-1}} & S_{j-1} < t \leqslant S_j^0, \\ \varepsilon_t & S_j^0 < t \leqslant S_j. \end{cases} \tag{5.12}$$

Write $\varepsilon \in DA(\alpha)$ $(0<\alpha<2)$ if there exist constants $c_\varepsilon^{\pm} \geqslant 0$, $c_\varepsilon^+ + c_\varepsilon^- > 0$ such that

$$\mathrm{P}[\varepsilon > x] \sim c_\varepsilon^+ x^{-\alpha} \quad (x \to \infty), \qquad \mathrm{P}[\varepsilon < x] \sim c_\varepsilon^- |x|^{-\alpha} \quad (x \to -\infty) \tag{5.13}$$

and, moreover, $\mathrm{E}\varepsilon = 0$ whenever $\alpha > 1$. We also write $\varepsilon \in DA(2)$ if $\mathrm{E}\varepsilon^2 < \infty$, $\mathrm{E}\varepsilon = 0$. Condition $\varepsilon \in DA(\alpha)$ implies that the distribution ε belongs to the domain of normal attraction of α-stable law (Ibragimov and Linnik, 1971), in other words,

$$n^{-1/\alpha} \sum_{t=1}^{n} \varepsilon_t \to_{\mathrm{d}} Z^0, \tag{5.14}$$

where Z^0 is α-stable r.v. $(0<\alpha\leqslant 2)$ and $\to_{\mathrm{d}}$ stands for convergence in distribution.

Clearly, if $\varepsilon \in DA(\alpha)$ $(0<\alpha\leqslant 2)$ then the second condition of (5.3) is satisfied. The first condition of (5.3) is satisfied as $\mathrm{P}[a_0 = 0] > 0$, see (5.11). As a consequence, (5.12) is a (unique) strictly stationary solution of (5.1) with a_t from (5.8) (Brandt, 1986). However, this solution need not have finite variance. A necessary and sufficient condition for (5.12) to be covariance stationary is given in the following theorem. Let $\bar{U}^1$ be a r.v. taking values $v = 1, 2, \ldots$ with probabilities $\mathrm{P}[\bar{U}^1 = v] = \mu_1^{-1}\mathrm{P}[U^1 \geqslant v]$.

Theorem 5.1. *Let* $\varepsilon \in DA(2)$. *Eq.* (5.1) *with* a_t *from* (5.8) *admits a covariance stationary solution* X_t *defined by* (5.12) *if and only if*

$$\sum_{v=1}^{\infty} A^{2v}\mathrm{P}[\bar{U}^1 \geqslant v] < \infty. \tag{5.15}$$

In this case,

$$\mathrm{Cov}(X_0, X_t) = (\sigma^2\mu_1/\mu)A^{-t}\sum_{v=t}^{\infty} A^{2v}\mathrm{P}[\bar{U}^1 \geqslant v], \tag{5.16}$$

where $\sigma^2 = \mathrm{E}\varepsilon^2$.

An immediate consequence of the above theorem is the fact that in the case $A > 1$, the covariance function of the covariance stationary solution X_t from (5.12) decays exponentially as $\mathrm{O}(A^{-t})$. However, in the case $A = 1$ the solution X_t may exhibit covariance stationary long memory. A similar fact was earlier observed in Pourahmadi (1988) (see also Leipus and Surgailis, 2003a).

Corollary 5.1. *Assume* $A = 1$ *and* $\varepsilon \in DA(2)$. *Moreover, assume*

$$\mathrm{P}[U^1 > u] \sim c_2 u^{-\beta} \quad (u \to \infty), \tag{5.17}$$

where $c_2 > 0$ *and* $\beta > 2$. *Then*

$$\mathrm{Cov}(X_0, X_t) \sim c_3 t^{-(\beta-2)}, \tag{5.18}$$

where $c_3 = c_2\sigma^2/(\mu(\beta-1)(\beta-2))$.

Indeed, (5.17) implies

$$\mathrm{P}[\bar{U}^1 \geqslant v] = \mu_1^{-1}\sum_{u=v}^{\infty}\mathrm{P}[U^1 \geqslant u] \sim (c_2/\mu_1)\sum_{u=v}^{\infty} u^{-\beta} \sim (c_2/\mu_1(\beta-1))v^{1-\beta}.$$

From this and (5.16), we get

$$\mathrm{Cov}(X_0, X_t) = (\sigma^2\mu_1/\mu)\sum_{v=t}^{\infty}\mathrm{P}[\bar{U}^1 \geqslant v] \sim \frac{c_2\sigma^2}{\mu(\beta-1)}\sum_{v=t}^{\infty} v^{1-\beta} \sim c_3 t^{2-\beta},$$

in accordance with (5.18).

The following theorems give conditions for the convergence to a stable limit of partial sums of (5.12) in the cases $A = 1$ and $A > 1$, respectively.

Theorem 5.2 (*Memoryless regime switching between I(0) and I(1)*). *Assume* $A = 1$. *Let* $\varepsilon \in DA(\alpha)$ $(0 < \alpha \leqslant 2)$ *and let condition* (5.17) *be satisfied, where*

$$1 < \beta < 1 + \alpha. \tag{5.19}$$

Moreover, assume that there exist $r < \alpha$ *and a constant* $C < \infty$ *such that*

$$\mathrm{E}[U^0 | U^1 = n] \leqslant C n^{r(1+\alpha)/\alpha}, \quad n \geqslant 1. \tag{5.20}$$

Then the convergence (2.9) *to a* λ*-stable Lévy process holds*, *with*

$$\lambda = \frac{\alpha\beta}{1+\alpha} \tag{5.21}$$

and the constant $c_0 = c_2(1+\alpha)^{-\beta/(1+\alpha)}$ in (2.8).

On the intuitive level, the above result can be explained as follows. Let $\mathrm{E}\varepsilon_t^2 < \infty$, or $\alpha = 2$. Note that the increment of the process $\sum_{t=1}^{k} X_t$ on interval $(S_{j-1}, S_j^0]$ follows

I(2) (integrated random walk) and its magnitude is proportional to $(S_j^0 - S_{j-1})^{3/2} = (U_j^1)^{3/2}$. Therefore increments of $n^{-1/\lambda}\sum_{t=1}^{[n\tau]} X_t$ on 'long' intervals $(S_{j-1}/n, S_j^0/n]$ of length $O(n^{1/\beta-1})$ have magnitude $O(n^{-1/\lambda}(n^{1/\beta})^{3/2}) = O(1)$ (where $\lambda = 2\beta/3$, see (5.21)), which does not vanish as $n \to \infty$. Hence occasional 'long' durations of the unit root regime $a_t = 1$ in the model (5.1) give rise to jumps in the limit process $W_\lambda(\tau)$, similarly as in the switching mean example of Section 3.

Theorem 5.3 (*Memoryless regime switching between I(0) and I(∞)*). *Assume* $A > 1$. *Let* $\varepsilon \in DA(\alpha)$ $(0 < \alpha \leqslant 2)$ *and let*

$$P[U^1 > u] \sim c_4 w^u \quad (u \to \infty), \tag{5.22}$$

where $c_4 > 0$ *and*

$$A^{-\alpha} < w < 1. \tag{5.23}$$

Moreover, assume that there exist $r < \alpha$ *and a constant* $C < \infty$ *such that*

$$E[U^0 | U^1 = n] \leqslant CA^{rn}, \quad n \geqslant 1. \tag{5.24}$$

Then the convergence (2.9) *to a* λ*-stable Lévy process holds, with*

$$\lambda = -\frac{\log w}{\log A} \tag{5.25}$$

and the constant $c_0 = c_4$ *in* (2.8).

The proofs of these theorems are given in Appendix B. Here we give some comments on the choice of normalization Φ_j in (2.1). Note that Y_j for $j \geqslant 2$ can be split

$$Y_j = \sum_{S_{j-1} < t \leqslant S_j^0} X_t + \sum_{S_j^0 < t \leqslant S_j} X_t =: Y_j' + Y_j'', \tag{5.26}$$

where the sum $Y_j' = \sum_{S_{j-1} < t \leqslant S_j^0} (\varepsilon_t + A\varepsilon_{t-1} + \cdots + A^{t-S_{j-1}-1}\varepsilon_{S_{j-1}+1})$ for fixed $U_j^1 = n$ has the same distribution as (we remind that $S_0 = 0$)

$$\begin{aligned} T(A,n) &= \sum_{t=1}^{n} (\varepsilon_t + A\varepsilon_{t-1} + \cdots + A^{t-1}\varepsilon_1) \\ &= \varepsilon_n + (1+A)\varepsilon_{n-1} + \cdots + (1 + A + \cdots + A^{n-1})\varepsilon_1. \end{aligned} \tag{5.27}$$

From $\varepsilon \in DA(\alpha)$ and the classical central limit theorem it easily follows that in the case $0 \leqslant A \leqslant 1$ the sum (5.27), normalized by $(1^\alpha + (1+A)^\alpha + \cdots + (1 + A + \cdots + A^{n-1})^\alpha)^{1/\alpha}$, has a limit α-stable distribution. This explains the choice of α-stable Z^0 and

$$\Phi_j = \Phi(U_j^1), \quad \Phi^0 = \Phi(U^1), \tag{5.28}$$

in the proof of Theorem 5.2, where

$$\Phi(n) = (1^\alpha + 2^\alpha + \cdots + n^\alpha)^{1/\alpha}. \tag{5.29}$$

In the case of Theorem 5.3, i.e. if $A>1$, it follows from definition (5.27) that $A^{-n}T(A,n)$ tends in distribution to a r.v. Z^0 given by the convergent series

$$Z^0 = (A-1)^{-1}\sum_{j=0}^{\infty} A^{-j}\varepsilon_j, \tag{5.30}$$

leading to the choice

$$\Phi_j = A^{U_j^1}, \quad \Phi^0 = A^{U^1}. \tag{5.31}$$

Clearly, (5.30) need not have stable distribution or even probability density. It is interesting to note that $Z^0 =_{\mathrm{d}} -A(A-1)^{-1}\tilde{X}_t$, where $\tilde{X}_t = -\sum_{j=0}^{\infty} A^{-j-1}\varepsilon_{t+j}$ is a noncausal stationary solution of $\tilde{X}_t = A\tilde{X}_{t-1} + \varepsilon_t$, $A>1$, and $=_{\mathrm{d}}$ stands for equality of distributions.

Conditions (5.20) and (5.24) are rather weak. They are needed to verify Assumption A_3 (see Appendix B). In the case when U^1 and U^0 are independent, they are automatically satisfied as we assume $\mu_0 = \mathrm{E}U^0 < \infty$. More generally, if U^1 and U^0 are dependent, these conditions roughly say that U^0 cannot grow very fast as $U^1 = n \to \infty$.

Conditions (5.19) and (5.23) (more precisely, the upper bound in (5.19) and the lower bound in (5.23)), combined with the corresponding tail conditions (5.17), (5.22) on the distribution of *on* interval U^1, seem crucial for λ-stable limit behavior of partial sums of X_t ($\lambda<\alpha$). Examples 5.1 and 5.2 below show that these bounds are quite sharp. One may expect that if (5.19), (5.23) are violated, partial sums process in Theorems 5.2 and 5.3 will converge to a α-stable Lévy process (Brownian motion if $\alpha=2$), under usual normalization $n^{1/\alpha}$.

Example 5.1. Let $A>1$, $\varepsilon \in DA(2)$ and let a_t be a stationary Markov chain with two states 0 and A and transition probabilities p_0 and p_A of staying in the same state, $0 \leqslant p_0, p_A < 1$. Then a_t can be represented as (5.8), where durations U_j^0 and U_j^1 are independent and geometrically distributed:

$$\mathrm{P}[U^1 = k] = (1-p_A)p_A^{k-1}, \quad \mathrm{P}[U^0 = k] = (1-p_0)p_0^{k-1}.$$

Then, if $p_A < A^{-2}$, the corresponding Markov regime switching process X_t has finite variance and exponentially decaying autocovariance, see Theorem 5.1, implying $\sum_{t=1}^n X_t = \mathrm{O_p}(n^{1/2})$. On the other hand, if $p_A > A^{-2}$, the process X_t satisfies conditions of Theorem 5.3 and $\sum_{t=1}^n X_t = \mathrm{O_p}(n^{1/\lambda})$ with $\lambda = -\log p_A / \log A < 2$. It is interesting to note that for the Blanchard's model with a_t as in (5.10), Lux and Sornette (2002) obtained the same tail index λ for the stationary solution X_t itself, in the case $0<\lambda<1$ when this solution has infinite expectation $\mathrm{E}|X_t| = \infty$.

Example 5.2. Let $A=1$, $\varepsilon \in DA(2)$ and let a_t be a stationary on/off process with independent *on* and *off* durations U^1 and U^0, where $\mathrm{E}U^0 < \infty$ and U^1 has a discrete Pareto distribution $\mathrm{P}[U^1 = k] = c_0 k^{-\beta-1}$ with parameter $\beta>1$. Then, if $\beta>3$, from Corollary 5.1 we have $\mathrm{Cov}(X_0, X_t) = \mathrm{O}(t^{-(\beta-2)})$ implying $\sum_{t=1}^n X_t = \mathrm{O_p}(n^{1/2})$. On the other hand, if $1<\beta<3$, Theorem 5.2 implies $\sum_{t=1}^n X_t = \mathrm{O_p}(n^{1/\lambda})$ with $\lambda =$

$2\beta/3<2$. Note that in this case Corollary 5.1 yields $\mathrm{Var}\left(\sum_{t=1}^{n} X_t\right) = \mathrm{O}(n^{4-\beta})$ growing faster than $\mathrm{O}(n^{2/\lambda})$.

6. Conclusion

It is well-known that covariance long memory similar to that in I(d) models can arise from structural, or regime changes with a heavy tailed duration distribution. Typical examples of models with long memory caused by regime switches is I(0) series with heavy tailed switching of mean and AR(1) process switching between i.i.d. regime and the unit root. However, unlike their second order properties, distributional properties of regime switching models with long memory seem to be very different from I(d) models: the latter models generally lead to a Gaussian but strongly persistent asymptotic process (fractional Brownian motion), and the former models to a heavy tailed stable process with independent increments.

We introduce a class of general regime switching models whose natural temporal aggregates between regime switching times have a characteristic stochastic volatility representation, with 'aggregate volatilities' largely determined by current regime variables (duration and type) and following a heavy tailed distribution, while 'aggregate innovations' are essentially independent of regime variables and have relatively light distribution tails. It is shown in the paper that the partial sums process of such stationary regime switching models converges to a stable Lévy process with independent increments. The intuitive meaning of the last result is that the covariance long memory of regime switching model does not persist in the distributional limit but instead 'transforms into excess variability'. Our results also apply to regime switching models with infinite variance.

Acknowledgements

The authors are grateful to three anonymous referees for valuable comments and suggestions.

Appendix A. Proof of Theorem 2.1

The following lemma is commonly attributed to Breiman (1965); see e.g. Pipiras et al. (2004, Lemma 1.1).

Lemma A.1. *Let* $\Phi^0 \geqslant 0$ *and* Z^0 *be independent random variables such that* $\mathrm{P}[\Phi^0 > u] \sim c_0 u^{-\lambda}$, $(u \to \infty)$ *for some* c_0, $\lambda > 0$, *and* $\mathrm{E}|Z^0|^r < \infty$ *for some* $r > \lambda$. *Let* $Y^0 = \Phi^0 Z^0$. *Then*

$$P[Y^0 > x] \sim c_+ x^{-\lambda} \quad (x \to \infty), \qquad \mathrm{P}[Y^0 \leqslant x] \sim c_- |x|^{-\lambda} \quad (x \to -\infty),$$

where c_+, c_- *are given by* (2.8).

Proof of Theorem 2.1. We shall prove the convergence of one-dimensional distributions at $\tau = 1$ only. For simplicity of notation, put $\mu = \mathrm{E}U = 1$. With Assumption A_5 in mind, it suffices to show

$$n^{-1/\lambda} \sum_{i=1}^{N_n} Y_i \to_{\mathrm{d}} W_\lambda(1). \tag{A.1}$$

We split the proof of (A.1) into three following steps.

Step 1: Approximation of the sum $\sum_{i=1}^{N_n} Y_i$ of random number N_n of summands by the sum $\sum_{i=1}^{n} Y_i$, in the sense that

$$Q_n := \sum_{i=1}^{N_n} Y_i - \sum_{i=1}^{n} Y_i = \mathrm{o_p}(n^{1/\lambda}). \tag{A.2}$$

Step 2: Approximation of $\sum_{j=1}^{n} Y_j$ by $\sum_{j=1}^{n} Y_j^0$, where the $Y_j^0 = \Phi_j^0 Z_j^0$, $j \geqslant 1$ are i.i.d. as in (2.10), in the sense that for each $v \in \mathbb{R}$

$$\left| \mathrm{E}\mathrm{e}^{\mathrm{i}vn^{-1/\lambda} \sum_{j=1}^{n} Y_j} - \mathrm{E}\mathrm{e}^{\mathrm{i}vn^{-1/\lambda} \sum_{j=1}^{n} Y_j^0} \right| = \mathrm{o}(1). \tag{A.3}$$

Step 3: Application of Lemma A.1 (proof of (2.10)).

We start with (the most difficult) Step 2, or the approximation of a sum of dependent r.v.'s Y_j by a corresponding sum of independent r.v.'s Y_j^0, in distribution. To that end, we need (i) to approximate the conditional distribution of each summand Y_j by (unconditional) distribution of Y_j^0, in the sense which is explained below, and (ii) to extend the approximation from summands to sums. For (i), we use the telescoping argument popular in the probability theory. We recall that the idea of telescoping is to replace summands consecutively one by one, so that each time we need to compare two sums which differ by only one summand. For (ii), Assumptions A_2–A_4 are used to show that the conditional distribution tails of Y_j and $Y_j^0 = \Phi_j^0 Z_j^0$ coincide, more precisely, that

$$\Delta(v) := \left| \mathrm{E}[\mathrm{e}^{\mathrm{i}vY_j} - \mathrm{e}^{\mathrm{i}vY_j^0} | \mathcal{G}_{j-1}] \right| \leqslant |v|^\lambda \tilde{\delta}(v) = \mathrm{o}(|v|^\lambda) \tag{A.4}$$

with a (nonrandom) $\tilde{\delta}(v) \to 0$ as $v \to 0$.

To prove (A.4), write $\Delta(v) \leqslant \Delta_1(v) + \Delta_2(v)$, where

$$\Delta_1(v) := \left| \mathrm{E}\left[\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0} | \Phi_j, \mathcal{G}_{j-1}] | \mathcal{G}_{j-1} \right] \right|, \quad \Delta_2(v) := \left| \mathrm{E}\left[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0} - \mathrm{e}^{\mathrm{i}v\Phi_j^0 Z_j^0} | \mathcal{G}_{j-1} \right] \right|.$$

Choose a large $K > 0$, then $\Delta_1(v) \leqslant \sum_{i=1}^{3} \Delta_{1i}(v, K)$, where

$$\Delta_{11}(v, K) = \left| \mathrm{E}\left[\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0} | \Phi_j, \mathcal{G}_{j-1}] I(K^{-1} < |v|\Phi_j \leqslant K) | \mathcal{G}_{j-1} \right] \right|,$$

$$\Delta_{12}(v, K) = \left| \mathrm{E}\left[\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0} | \Phi_j, \mathcal{G}_{j-1}] I(|v|\Phi_j \leqslant K^{-1}) | \mathcal{G}_{j-1} \right] \right|,$$

$$\Delta_{13}(v, K) = \left| \mathrm{E}\left[\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0} | \Phi_j, \mathcal{G}_{j-1}] I(|v|\Phi_j > K) | \mathcal{G}_{j-1} \right] \right|.$$

Since $|\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0}|\Phi_j, \mathscr{G}_{j-1}]| \leqslant 2$, using (2.3) we obtain

$$\Delta_{13}(v, K) \leqslant 2\mathrm{P}[\Phi_j > |v|^{-1}K|\mathscr{G}_{j-1}] \leqslant C|v|^{\lambda}/K^{\lambda}, \tag{A.5}$$

where $C < \infty$ is a nonrandom constant.

Next, consider $\Delta_{12}(v, K)$. We shall use the following well-known fact: for any $0 < r \leqslant 2$ and any r.v. ξ (for $1 < r \leqslant 2$, assume $\mathrm{E}\xi = 0$ in addition), the following inequality holds: for any real number u

$$|\mathrm{E}e^{iu\xi} - 1| \leqslant 3\min(1, |u|^r \mathrm{E}|\xi|^r). \tag{A.6}$$

Using (A.6), Assumptions A_2, A_4 and integration by parts, we obtain

$$\begin{aligned}
\Delta_{12}(v, K) &\leqslant C|v|^r \mathrm{E}\Big[\Phi_j^r\Big(\mathrm{E}[|Z_j|^r|\Phi_j, \mathscr{G}_{j-1}] + \mathrm{E}|Z_j^0|^r\Big) I(|v|\Phi_j \leqslant K^{-1})\Big|\mathscr{G}_{j-1}\Big] \\
&\leqslant C|v|^r \mathrm{E}\Big[\Phi_j^r I(|v|\Phi_j \leqslant K^{-1})\Big|\mathscr{G}_{j-1}\Big] \\
&= -C|v|^r \int_0^{|v|^{-1}K^{-1}} u^r \,\mathrm{dP}[\Phi_j > u|\mathscr{G}_{j-1}] \\
&= C|v|^r \Bigg(-|v|^{-r}K^{-r}\mathrm{P}[\Phi_j > v^{-1}K^{-1}|\mathscr{G}_{j-1}] \\
&\quad + r\int_0^{|v|^{-1}K^{-1}} \mathrm{P}[\Phi_j > u|\mathscr{G}_{j-1}] u^{r-1}\,\mathrm{d}u \Bigg) \\
&\leqslant C|v|^r \int_0^{|v|^{-1}K^{-1}} u^{r-1-\lambda}\,\mathrm{d}u \\
&\leqslant C|v|^{\lambda}/K^{r-\lambda}.
\end{aligned} \tag{A.7}$$

By (A.5), (A.7), $\sup_{|v|<1} |v|^{-\lambda}(\Delta_{12}(v, K) + \Delta_{13}(v, K))$ can be made arbitrarily small by choosing K large enough. Then (A.4) follows for $\Delta_1(v)$ if we show for any $K < \infty$

$$\lim_{v \to 0} |v|^{-\lambda}\Delta_{11}(v, K) = 0. \tag{A.8}$$

Let $\delta_1 > 0$ be an arbitrary small number and let $\delta = \delta_1 K^{-\lambda}$. Given δ and K, by Assumption A_3 one can find $L = L(K, \delta) > 0$ such that

$$\sup_{|u| \leqslant K} \left|\mathrm{E}[\mathrm{e}^{\mathrm{i}uZ_j}\Big|\Phi_j, \mathscr{G}_{j-1}] - \mathrm{E}\mathrm{e}^{\mathrm{i}uZ_j^0}\right| = \sup_{|u| \leqslant K}\left|\mathrm{E}[\mathrm{e}^{\mathrm{i}uZ_j} - \mathrm{e}^{\mathrm{i}uZ_j^0}|\Phi_j, \mathscr{G}_{j-1}]\right| \leqslant \delta$$

holds on $\{\Phi_j > L\}$. Clearly, on the set $\{K^{-1} < |v|\Phi_j \leqslant K\}$ we have

$$\left|\mathrm{E}[\mathrm{e}^{\mathrm{i}v\Phi_j Z_j} - \mathrm{e}^{\mathrm{i}v\Phi_j Z_j^0}|\Phi_j, \mathscr{G}_{j-1}]\right| \leqslant \sup_{|u| \leqslant K}\left|\mathrm{E}[\mathrm{e}^{\mathrm{i}uZ_j} - \mathrm{e}^{\mathrm{i}uZ_j^0}|\Phi_j, \mathscr{G}_{j-1}]\right| \leqslant \delta \tag{A.9}$$

provided $|v| < 1/(KL)$. Consequently, if $|v| < 1/(KL)$, where $L = L(K, \delta)$ is defined above, by (A.9) and Assumption A_2 it follows that

$$\begin{aligned}
\Delta_{11}(v, K) &\leqslant \delta\mathrm{P}[K^{-1} < |v|\Phi_j|\mathscr{G}_{j-1}] \\
&= \delta\mathrm{P}[\Phi_j > K^{-1}|v|^{-1}|\mathscr{G}_{j-1}] \leqslant C\delta K^{\lambda}|v|^{\lambda} = C\delta_1|v|^{\lambda},
\end{aligned}$$

implying $\limsup_{v\to 0}|v|^{-\lambda}\Delta_{11}(v,K)\leqslant C\delta_1$. By arbitrariness of $\delta_1>0$, this proves (A.8) and hence (A.4) for $\Delta_1(v)$.

Let us prove (A.4) for $\Delta_2(v)$. By A_2,

$$\begin{aligned}\Delta_2(v) &= \left|\int \mathrm{E}(\mathrm{e}^{\mathrm{i}vz\Phi_j}-\mathrm{e}^{\mathrm{i}vz\Phi_j^0}|\mathscr{F}_{S_{j-1}})\mathrm{P}[Z^0\in \mathrm{d}z]\right|\\ &\leqslant \int |vz|^{\lambda}h(vz)\mathrm{P}[Z^0\in\mathrm{d}z]\\ &= |v|^{\lambda}\mathrm{E}|Z^0|^{\lambda}h(vZ^0)=|v|^{\lambda}h^0(v),\end{aligned}$$

where $h^0(v){:=}\mathrm{E}|Z^0|^{\lambda}h(vZ^0)\to 0$ as $v\to 0$ by the dominated convergence theorem. This completes the proof of (A.4).

Let us prove (A.3). Denote $W_n{:=}n^{-1/\lambda}\sum_{j=1}^n Y_j$, $W_n^0{:=}n^{-1/\lambda}\sum_{j=1}^n Y_j^0$,

$$W_{n,k}{:=}n^{-1/\lambda}\left(\sum_{j=1}^{k} Y_j+\sum_{j=k+1}^{n} Y_j^0\right),$$

so that $W_n=W_{n,n}$ and $W_n^0=W_{n,0}$. Then by telescoping identity,

$$\mathrm{E}\mathrm{e}^{\mathrm{i}uW_n}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_n^0}=\left(\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,n}}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,n-1}}\right)+\cdots+\left(\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,1}}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,0}}\right).$$

Using the fact that Y_j, $j\leqslant k-1$ are $\mathscr{G}_{k-1}$ measurable while Y_j^0, $j\leqslant n$ are independent of $\mathscr{G}_n$, one can write

$$\begin{aligned}&\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,k}}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,k-1}}\\ &= \mathrm{E}\mathrm{e}^{\mathrm{i}un^{-1/\lambda}\left(\sum_{j=1}^{k}Y_j+\sum_{j=k+1}^{n}Y_j^0\right)}-\mathrm{E}\mathrm{e}^{\mathrm{i}un^{-1/\lambda}\left(\sum_{j=1}^{k-1}Y_j+\sum_{j=k}^{n}Y_j^0\right)}\\ &= \mathrm{E}\left[\mathrm{e}^{\mathrm{i}un^{-1/\lambda}\sum_{j=1}^{k-1}Y_j}\mathrm{E}[\mathrm{e}^{\mathrm{i}un^{-1/\lambda}Y_k}-\mathrm{e}^{\mathrm{i}un^{-1/\lambda}Y_k^0}|\mathscr{G}_{k-1}]\right]\mathrm{E}\mathrm{e}^{\mathrm{i}un^{-1/\lambda}\sum_{j=k+1}^{n}\Phi_j^0Z_j^0}.\end{aligned}$$

By (A.4), uniformly in k, for any fixed u

$$\begin{aligned}|\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,k}}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_{n,k-1}}|&\leqslant \mathrm{E}\left|\mathrm{E}[\mathrm{e}^{\mathrm{i}un^{-1/\lambda}Y_k}-\mathrm{e}^{\mathrm{i}un^{-1/\lambda}Y_k^0}|\mathscr{G}_{k-1}]\right|\\ &= \mathrm{o}((n^{-1/\lambda})^{\lambda})=\mathrm{o}(n^{-1}).\end{aligned}$$

Consequently, for any u, we obtain $|\mathrm{E}\mathrm{e}^{\mathrm{i}uW_n}-\mathrm{E}\mathrm{e}^{\mathrm{i}uW_n^0}|=n\mathrm{o}(n^{-1})=o(1)$, or (A.3).

To complete the proof of Theorem 2.1, it suffices to show (A.2), or Step 1, as Step 3 follows by Lemma A.1. For any δ_1, $\delta_2>0$ one has

$$\begin{aligned}\mathrm{P}[|Q_n|>\delta_1 n^{1/\lambda}]&\leqslant \mathrm{P}[|Q_n|>\delta_1 n^{1/\lambda},(1-\delta_2)n\leqslant N_n\leqslant(1+\delta_2)n]\\ &\quad+\mathrm{P}[|N_n-n|>\delta_2 n]=:r_1(n)+r_2(n).\end{aligned}$$

By the law of large numbers, for any $\delta>0$, $\delta_2>0$ one can find n_0 such that $r_2(n)\leqslant\delta$, $\forall n>n_0$. It suffices to show that for any given $\delta,\delta_1>0$ one can find $\delta_2=\delta_2(\delta,\delta_1)>0$ small enough so that for all n

$$r_1(n)\leqslant\delta. \tag{A.10}$$

Due to the fact that A_1–A_4 hold uniformly in j, this follows from

$$\kappa(n,\delta_2):=\mathrm{P}\left[\sup_{1\leqslant k\leqslant \delta_2 n}\left|\sum_{j=1}^{k} Y_j\right| > \delta_1 n^{1/\lambda}\right] \leqslant \delta. \tag{A.11}$$

Consider first the case $1\leqslant\lambda<2$. Note $(Y_j,\mathscr{G}_j)_{j>1}$ is a martingale difference by Assumption A_4. Let $\bar{Y}_j:=\bar{\Phi}_j Z_j$, where

$$\bar{\Phi}_j:=\begin{cases}\Phi_j & \text{if } \Phi_j\leqslant n^{1/\lambda},\\ 0 & \text{otherwise.}\end{cases}$$

Then $(\bar{Y}_j,\mathscr{G}_j)_{j>1}$ is again a martingale difference sequence. Then

$$\begin{aligned}\kappa(n,\delta_2) &= \mathrm{P}\left[\sup_{1\leqslant k\leqslant \delta_2 n}\left|\sum_{j=1}^{k} Y_j\right| > \delta_1 n^{1/\lambda};\ \Phi_j=\bar{\Phi}_j \quad \text{for all } 1\leqslant j\leqslant \delta_2 n\right]\\ &\quad + \mathrm{P}\left[\sup_{1\leqslant k\leqslant \delta_2 n}\left|\sum_{j=1}^{k} Y_j\right| > \delta_1 n^{1/\lambda};\ \Phi_j\neq\bar{\Phi}_j \quad \text{for some } 1\leqslant j\leqslant \delta_2 n\right]\\ &\leqslant \mathrm{P}\left[\sup_{1\leqslant k\leqslant \delta_2 n}\left|\sum_{j=1}^{k} \bar{Y}_j\right| > \delta_1 n^{1/\lambda}\right] + \mathrm{P}\left[\Phi_j\neq\bar{\Phi}_j \quad \text{for some } 1\leqslant j\leqslant \delta_2 n\right]\\ &=:\kappa_1(n,\delta_2)+\kappa_2(n,\delta_2).\end{aligned}$$

By Doob's inequality for martingales, for any $\lambda<r\leqslant 2$,

$$\kappa_1(n,\delta_2)\leqslant C\delta_1^{-r}n^{-r/\lambda}\mathrm{E}\left|\sum_{j=1}^{[\delta_2 n]}\bar{Y}_j\right|^r \leqslant C\delta_1^{-r}n^{-r/\lambda}\sum_{j=1}^{[\delta_2 n]}\mathrm{E}\left|\bar{Y}_j\right|^r,$$

where, using A_2 and A_4, similarly as in (A.7),

$$\mathrm{E}|\bar{Y}_j|^r = \mathrm{E}\left[\Phi_j^r\mathrm{E}[|Z_j|^r|\mathscr{G}_{j-1},\Phi_j];\Phi_j\leqslant n^{1/\lambda}\right]\leqslant C\mathrm{E}[\Phi_j^r;\Phi_j\leqslant n^{1/\lambda}]\leqslant Cn^{(r/\lambda)-1}.$$

We thus obtain $\kappa_1(n,\delta_2)\leqslant C\delta_2\delta_1^{-r}$. Similarly, by A_2, $\kappa_2(n,\delta_2)\leqslant\sum_{j=1}^{[\delta_2 n]}\mathrm{P}[\Phi_j>n^{1/\lambda}]\leqslant C\delta_2$. Therefore $\kappa_i(n,\delta_2)$, $i=1,2$ can be made arbitrarily small by an appropriate choice of δ_2, uniformly in n. This proves (A.11) and (A.10) for $\lambda\geqslant 1$. Case $0<\lambda<1$ follows by a similar argument and using the simple triangle inequality in L^r $(r<1)$ instead of Doob's inequality. □

Appendix B. Other proofs

Proof of Corollary 3.1. We apply Theorem 2.1 with $Y_j=U_j\zeta_j$, $\Phi_j=U_j$, $Z_j=\zeta_j$, $\Phi^0=U$, $Z^0=\zeta$. By independence of U_j, ζ_j and relation (3.4), A_1–A_3 are trivially satisfied, with $h(u) = \delta(u) \equiv 0$. Assumption A_4 follows from $E|\zeta|^r < \infty$, $E\zeta = 0$. It remains to verify A_5. More generally, assume $\sup_{t \geqslant 1} E|X_t|^\delta < \infty$ for some $0 < \delta \leqslant 1$ (for X_t in (3.1), this assumption is clearly satisfied with $\delta = 1$, as $E|X_t| \leqslant E|\zeta| + E|\varepsilon| < \infty$). Let V_n denote the difference of the two sums in A_5, i.e. $V_n = \sum_{S_{N_n} < t \leqslant n} X_t$. For any $K, L > 0$ one can write

$$P[|V_n| > K] \leqslant P[n - N_n > L] + P\left[\sum_{n-L<t\leqslant n} |X_t| > K\right].$$

As $n - S_{N_n} = O_p(1)$, see Feller (1971, Chapter 11, Example 10), the first probability on the r.h.s. can be made arbitrary small by taking L large enough. The second probability on the r.h.s. does not exceed $K^{-\delta} E(\sum_{n-L<t\leqslant n} |X_t|)^\delta \leqslant K^{-\delta} \sum_{n-L<t\leqslant n} E|X_t|^\delta$, for arbitrary $0 < \delta \leqslant 1$. Therefore this probability is less than $CLK^{-\delta} \leqslant L^{-1}$, for some constant C independent of n and $K = (CL^2)^{1/\delta}$. Hence $V_n = O_p(1)$, thereby proving A_5. $\square$

Proof of Corollary 4.1. For Φ_j, Z_j defined in (4.5), A_1 follows by definition, A_2 follows by tail condition (4.4), A_3 is immediate with $Z^0 = \sum_{i=1}^U \varepsilon_i$, and A_5 can be proved as in the proof of Corollary 3.1 above, as $E|X_t|^\delta = E\zeta^\delta E|\varepsilon|^\delta < \infty$ for $\delta > 0$ small enough. Finally, A_4 follows easily by the independence of U and ε_t:

$$E|Z_j|^r = E|Z^0|^r = EE\left[\left|\sum_{i=1}^U \varepsilon_i\right|^r \middle| U\right] \leqslant 2E\left[\sum_{i=1}^U E|\varepsilon_i|^r\right] = 2EUE|\varepsilon|^r < \infty. \qquad \square$$

Proof of Theorem 5.1. Let us first give a rigorous construction of a stationary process (5.7) with a given joint distribution (U^0, U^1) (for mutually independent U^0 and U^1, this construction is given in Heath et al. (1998)). Let $U := U^0 + U^1$, $EU =: \mu = \mu_0 + \mu_1$, and let

$$\cdots < S_{-1} < 0 \leqslant S_0 < S_1 < \cdots \tag{B.1}$$

be a (double-sided) stationary renewal process with interarrival distribution U. The well-known construction of (B.1) starts with a joint distribution of the pair (S_{-1}, S_0):

$$P[S_0 = u, S_{-1} = -v] := \mu^{-1} P[U = u + v], \quad u = 0, 1, \ldots, \ v = 1, 2, \ldots \tag{B.2}$$

and an i.i.d. sequence U_j, $j \neq 0$, independent of (S_{-1}, S_0) and distributed according to U. The moments $S_j, j \neq 0, -1$ are then defined by

$$S_j := \begin{cases} S_0 + \sum_{k=1}^j U_j & j = 1, 2, \ldots, \\ S_{-1} - \sum_{k=-1}^{j+1} U_j & j = -2, -3, \ldots \end{cases} \tag{B.3}$$

Let U_j^1, $j \in \mathbb{Z}$ be conditionally independent given S_i, $i \in \mathbb{Z}$ defined as above, and distributed according to the same (conditional) probability

$$\begin{aligned}\mathrm{P}[U_j^1 = k | S_i = s_i, i \in Z] &:= \mathrm{P}[U^1 = k | U = s_j - s_{j-1}] \\ &= \mathrm{P}[U^1 = k | U^0 + U^1 = s_j - s_{j-1}],\end{aligned} \tag{B.4}$$

$k = 1, 2, \ldots, s_j - s_{j-1} - 1$, $j \in \mathbb{Z}$. Put $S_j^0 := S_j + U_j^1$, $j \in \mathbb{Z}$. From stationarity of (B.3) and (B.4) it easily follows stationarity of (5.7) and (5.8).

To prove the theorem, let $\sigma = 1$, for simplicity. Consider the stationary solution X_t as given by (5.2). This solution has finite variance if and only if (see Pourahmadi, 1988) $\sum_{u=0}^{\infty} \mathrm{E}[a_0^2 \ldots a_{-u}^2] < \infty$. Clearly,

$$\begin{aligned}\mathrm{E}[a_0^2 \ldots a_{-u}^2] &= A^{2u+2}\mathrm{P}[a_0 = \cdots = a_{-u} = A] \\ &= A^{2u+2}\mathrm{P}[S_{-1} < -u, S_0^0 \geqslant 0] \\ &= A^{2u+2} \sum_{s_{-1} < -u, s_0 > 0} \mathrm{P}[S_{-1} = s_{-1}, S_0 = s_0, S_0^0 \geqslant 0].\end{aligned}$$

According to (B.2)–(B.4),

$$\begin{aligned}&\mathrm{P}[S_{-1} = s_{-1}, S_0 = s_0, S_0^0 \geqslant 0] \\ &\quad = \mathrm{P}[S_{-1} = s_{-1}, S_0 = s_0]\mathrm{P}[S_0^0 \geqslant 0 | S_{-1} = s_{-1}, S_0 = s_0] \\ &\quad = \mu^{-1}\mathrm{P}[U = s_0 - s_{-1}]\mathrm{P}[U^1 \geqslant -s_{-1} | U = s_0 - s_{-1}] \\ &\quad = \mu^{-1}\mathrm{P}[U = s_0 - s_{-1}, U^1 \geqslant -s_{-1}].\end{aligned}$$

Therefore,

$$\begin{aligned}&\sum_{s_{-1} < -u, s_0 > 0} \mathrm{P}[S_{-1} = s_{-1}, S_0 = s_0, S_0^0 \geqslant 0] \\ &\quad = \mu^{-1} \sum_{s_{-1} < -u, s_0 > 0} \mathrm{P}[U = s_0 - s_{-1}, U^1 \geqslant -s_{-1}] \\ &\quad = \mu^{-1} \sum_{s_{-1} < -u} \mathrm{P}[U^1 \geqslant -s_{-1}] \\ &\quad = (\mu_1/\mu)\mathrm{P}[\bar{U}^1 > u]\end{aligned}$$

and we obtain $\sum_{u=0}^{\infty} \mathrm{E}[a_0^2 \ldots a_{-u}^2] = (\mu_1/\mu)\sum_{v=1}^{\infty} A^{2v}\mathrm{P}[\bar{U}^1 \geqslant v]$, thereby proving the first part of the theorem.

In a similar way,

$$\begin{aligned}\mathrm{E}X_0X_t &= \mathrm{E}[a_t \ldots a_1] + \sum_{u=0}^{\infty} \mathrm{E}[a_t \ldots a_1 a_0^2 \ldots a_{-u}^2] \\ &= A^t\mathrm{P}[a_t = \cdots = a_1 = A] \\ &\qquad + \sum_{u=0}^{\infty} A^{t+2u+2}\mathrm{P}[a_t = \cdots = a_1 = a_0 = \cdots = a_{-u} = A]\end{aligned}$$

$$
\begin{aligned}
&= (\mu_1/\mu)A^t\left(\mathrm{P}[\bar{U}^1 \geqslant t] + \sum_{v=1}^{\infty} A^{2v}\mathrm{P}[\bar{U}^1 \geqslant t+v]\right) \\
&= (\mu_1/\mu)A^{-t}\sum_{u=t}^{\infty} A^{2u}\mathrm{P}[\bar{U}^1 \geqslant u].
\end{aligned}
$$

Theorem 5.1 is proved. $\square$

Proof of Theorem 5.2. Let Φ_j, Φ^0 be defined as in (5.28), (5.29). Note $\Phi(n) \sim n^{(1+\alpha)/\alpha}(1+\alpha)^{-1/\alpha}$, therefore from (5.17)

$$
\mathrm{P}[\Phi^0 > x] \sim \mathrm{P}[U^1 > (1+\alpha)^{1/(1+\alpha)}x^{\alpha/(1+\alpha)}] \sim c_0 x^{-\lambda} \quad (x \to \infty),
$$

where $c_0 = c_2(1+\alpha)^{-\beta/(1+\alpha)}$. Whence, Assumption A_2 follows, with $h(u) \equiv 0$.

Let us verify A_3. To that end, split $Y_j = Y_j' + Y_j''$ as in (5.26). Clearly, distributions of Y_j', Y_j'' depend on U_j^1, U_j^0 and do not depend on $\mathscr{G}_{j-1}$, moreover, $Y_j' = \sum_{S_{j-1} < t \leqslant S_j^0}(t - S_{j-1})\varepsilon_t$ and $Y_j'' = \sum_{S_j^0 < t \leqslant S_j} \varepsilon_t$ are conditionally independent given Φ_j. Using stationarity of the renewal process and the fact that $\Phi_j \to \infty$ is equivalent to $U_j^1 \to \infty$, A_3 follows from

$$
\mathrm{E}[\exp\{\mathrm{i}u\Phi(n)^{-1}Y''(U^0)\}|U^1 = n] \to 1, \tag{B.5}
$$

$$
\mathrm{E}[\exp\{\mathrm{i}u\Phi(n)^{-1}Y'(U^1)\}|U^1 = n] \to \mathrm{E}\mathrm{e}^{\mathrm{i}uZ^0}, \tag{B.6}
$$

where

$$
Y'(n) = \sum_{k=1}^{n} k\varepsilon_k, \quad Y''(n) = \sum_{k=1}^{n} \varepsilon_k.
$$

Relation (B.6) is equivalent to the convergence $\sum_{k=1}^{n} b_{nk}\varepsilon_k \to_{\mathrm{d}} Z^0$ of weighted sum of i.i.d. r.v.'s ε_k, with weights $b_{nk} := \Phi(n)^{-1}k$ satisfying $\sum_{k=1}^{n} b_{nk}^{\alpha} = 1$ (see the definition (5.29) of $\Phi(n)$). Then (B.6) follows from assumption $\varepsilon \in DA(\alpha)$ and standard probabilistic argument (Araujo and Giné, 1980, Theorem 2.3.5).

If U^0 and U^1 are independent, (B.5) is obvious by $Y''(U^0) = \mathrm{O_p}(1)$ and $\Phi(n)^{-1} \to 0$. On the other hand, if $U^0 = G(U^1)$ is some function of U^1, then (5.20) implies $G(n) = \mathrm{O}(n^{r(1+\alpha)/\alpha})$ and $Y''(U^0) = \mathrm{O_p}((U^0)^{1/\alpha}) = \mathrm{O_p}(n^{r(1+\alpha)/\alpha^2}) = \mathrm{o_p}(\Phi(n))$ on the set $\{U^1 = n\}$, due to $\Phi(n) = \mathrm{O}(n^{(1+\alpha)/\alpha})$ and $r < \alpha$. In the general case, in order to prove (B.5) we need to show that for any $\in\, > 0$

$$
I(n, \in) := \mathrm{P}[|\Phi(n)^{-1}Y''(U^0)| > \,\in |U^1 = n] \to 0
$$

as $n \to \infty$. For any $k \geqslant 1$ (which will be chosen later), we can write

$$
\begin{aligned}
I(n, \in) \leqslant &\sum_{j=1}^{k} \mathrm{P}\left[\left|\Phi(n)^{-1}\sum_{s=1}^{j}\varepsilon_s\right| > \,\in, U^0 = j \,\middle|\, U^1 = n\right] \\
&+ \mathrm{P}[U^0 > k|U^1 = n] =: I_1 + I_2.
\end{aligned}
$$

Using the independence of (U^0, U^1) and $\varepsilon_t, 1 \leqslant t \leqslant n$,

$$I_1 = \sum_{j=1}^{k} \mathrm{P}\left[\Phi(n)^{-1} j^{1/\alpha} \left| j^{-1/\alpha} \sum_{s=1}^{j} \varepsilon_s \right| > \epsilon \middle| U^0 = j, U^1 = n\right] \mathrm{P}[U^0 = j | U^1 = n]$$
$$\leqslant \max_{1 \leqslant j \leqslant k} \mathrm{P}\left[\Phi(n)^{-1} k^{1/\alpha} \left| j^{-1/\alpha} \sum_{s=1}^{j} \varepsilon_j \right| > \epsilon \right].$$

Now, since $\Phi(n) \sim n^{(1+\alpha)/\alpha}$ and $j^{-1/\alpha} \sum_{s=1}^{j} \varepsilon_s = \mathrm{O}_\mathrm{p}(1)$, see (5.14), therefore $I_1 \to 0$ if k is chosen so that $\Phi(n)^{-1} k^{1/\alpha} \to 0$. In particular, one can choose $k = n^{1+\alpha-\gamma}$ with some $\gamma > 0$. With this k and using (5.20), we obtain

$$I_2 = \mathrm{P}[U^0 > k | U^1 = n] \leqslant n^{-1-\alpha+\gamma} \mathrm{E}[U^0 | U^1 = n] \leqslant C n^{-1-\alpha+\gamma+r(1+\alpha)/\alpha} \to 0$$

provided $\gamma > 0$ was taken small enough ($\gamma < (1+\alpha)(1-(r/\alpha))$). This proves (B.5) and A_3.

Let us verify A_4. Similarly as in (B.5), (B.6), it suffices to show $J' \leqslant C$, $J'' \leqslant C$, for

$$J' := \Phi(n)^{-r} \mathrm{E}[|Y'(U^1)|^r | U^1 = n], \quad J'' := \Phi(n)^{-r} \mathrm{E}[|Y''(U^0)|^r | U^1 = n] \tag{B.7}$$

and some $\lambda < r < \alpha$. For J'', by applying Marcinkiewicz–Zygmund moment inequality and (5.20), we get

$$J'' \leqslant \Phi(n)^{-r} \mathrm{E}|\varepsilon_1|^r E[U^0 | U^1 = n] \leqslant C n^{-r(1+\alpha)/\alpha + r(1+\alpha)/\alpha} = C.$$

The estimate $J' \leqslant C$ follows from the bound $\mathrm{P}[\Phi(n)^{-1} |Y'(n)| > x] \leqslant C x^{-\alpha}$ of the tail distribution function of weighted sum $\Phi(n)^{-1} Y'(n) = \sum_{k=1}^{n} b_{nk} \varepsilon_k$ of i.i.d. r.v.'s ε_k; see Mikosch and Samorodnitsky (2000, Lemma A.4). This proves Assumption A_4 and Theorem 5.2. □

Proof of Theorem 5.3. As in (5.31), let $\Phi_j = A^{U_j^1}$, $\Phi^0 = A^{U^1}$. Clearly, Φ_j is independent of the past history $\mathscr{G}_{j-1}$ and has the same distribution as Φ^0. Condition (5.22) easily implies

$$\mathrm{P}[\Phi^0 > x] = \mathrm{P}[U^1 > \log x / \log A] \sim c_3 w^{\log x / \log A} = c_0 x^{-\lambda}$$

as $x \to \infty$, where λ is defined in (5.25). Whence A_2 follows, with $h(u) \equiv 0$.

To verify A_3, it suffices to show that for each $u \in \mathbb{R}$

$$\mathrm{E}[\mathrm{e}^{\mathrm{i}uZ_j} | U_j^1 = n] \to \mathrm{E}\mathrm{e}^{\mathrm{i}uZ^0} \quad \text{as} \quad n \to \infty, \tag{B.8}$$

where r.v. Z^0 is defined in (5.30). Similarly as in the proof of Theorem 5.2, this follows from

$$\mathrm{E}[\exp\{\mathrm{i}uA^{-U^1} Y''(U^0)\} | U^1 = n] \to 1, \tag{B.9}$$
$$\mathrm{E}[\exp\{\mathrm{i}uA^{-U^1} Y'(U^1)\} | U^1 = n] \to \mathrm{E}\mathrm{e}^{\mathrm{i}uZ^0}, \tag{B.10}$$

as $n \to \infty$, where $Y'(n) := \varepsilon_n + (1+A)\varepsilon_{n-1} + \cdots + (1 + A + \cdots + A^{n-1})\varepsilon_1$, $Y''(n) := \varepsilon_1 + \cdots + \varepsilon_n$.

To show (B.9), as in proof of Theorem 5.2 we write $I(n,\in) := \mathrm{P}[A^{-n}|Y''(U^0)| > \in |U^1 = n] \leqslant I_1 + I_2$, where

$$I_1 := \sum_{j=1}^{k} \mathrm{P}\left[A^{-n}\left|\sum_{s=1}^{j} \varepsilon_s\right| > \in, U^0 = j \middle| U^1 = n\right], \quad I_2 := \mathrm{P}[U^0 > k | U^1 = n].$$

Similarly as in the proof of Theorem 5.2, it follows easily that $I_1 \to 0$ provided $A^{-n}k^{1/\alpha} \to 0$, and we can choose $k = A^{n\alpha}n^{-1}$. Then, using condition (5.24), $I_2 \leqslant k^{-1}\mathrm{E}[U^0|U^1 = n] \leqslant CA^{-n\alpha}nA^{nr} = CnA^{-n(\alpha-r)} \to 0$, thus proving (B.9).

To prove (B.10), write

$$\begin{aligned} A^{-n}Y'(n) &= A^{-n}(A-1)^{-1}\sum_{i=1}^{n}(A^{n-i+1}-1)\varepsilon_i \\ &=_{\mathrm{d}} (A-1)^{-1}\sum_{i=0}^{n-1} A^{-i}\varepsilon_i - (A-1)^{-1}A^{-n}\sum_{i=0}^{n-1}\varepsilon_i = Z^0 - R_n, \end{aligned}$$

where $R_n := (A-1)^{-1}\sum_{i=n}^{\infty} A^{-i}\varepsilon_i + (A-1)^{-1}A^{-n}\sum_{i=0}^{n-1}\varepsilon_i$. Using $A > 1$ and $\varepsilon \in DA(\alpha)$, $R_n = \mathrm{o_p}(1)$ easily follows, thereby implying (B.10).

It remains to verify Assumption A_4, or $J' \leqslant C,\ J'' \leqslant C$, where

$$J' := A^{-nr}\mathrm{E}[|Y'(n)|^r], \quad J'' := A^{-nr}\mathrm{E}[|Y''(U^0)|^r | U^1 = n]$$

for some $\lambda < r < \alpha$, cf. (B.7). Using the fact that $\varepsilon \in DA(\alpha)$ implies $\mathrm{E}|\varepsilon|^r < \infty$ for any $0 < r < \alpha \leqslant 2$ as well as the inequality $\mathrm{E}|\sum_{i=1}^{n} a_i\varepsilon_i|^r \leqslant 2\sum_{i=1}^{n}|a_i|^r\mathrm{E}|\varepsilon_i|^r$ which is valid for any $0 < r \leqslant 2$, we obtain

$$J' \leqslant C\left(\sum_{i=1}^{n} A^{-ri}\mathrm{E}|\varepsilon_i|^r + A^{-rn}\sum_{i=1}^{n}\mathrm{E}|\varepsilon_i|^r\right) \leqslant C\left(\sum_{i=1}^{n} A^{-ri} + A^{-rn}n\right) \leqslant C.$$

Finally, using (5.24)

$$J'' \leqslant CA^{-nr}\mathrm{E}\left[\sum_{i=1}^{U^0}\mathrm{E}|\varepsilon_i|^r \middle| U^1 = n\right] \leqslant CA^{-nr}\mathrm{E}[U^0|U^1 = n] \leqslant C.$$

This completes the proof of Theorem 5.3. □

References

Araujo, A., Giné, E., 1980. The Central Limit Theorem for Real and Banach Valued Random Variables. Wiley, New York.

Blanchard, O.J., 1979. Speculative bubbles, crashes and rational expectations. Economics Letters 3, 387–389.

Blanchard, O.J., Watson, M.W., 1982. Bubbles, rational expectations and speculative markets. In: Wachtel, P. (Ed.), Crisis in Economic and Financial Structure: Bubbles, Bursts, and Shocks. Lexington Books, Lexington.

Brandt, A., 1986. The stochastic equation $Y_{n+1} = A_nY_n + B_n$ with stationary coefficients. Advances in Applied Probability 18, 211–220.

Breiman, L., 1965. On some limit theorems similar to the arc-sin law. Theory of Probability and Applications 10, 323–331.

Chow, Y.-F., Liu, M., 1999. Long swings with memory and stock market fluctuations. Journal of Financial and Quantitative Analysis 34, 341–367.
Cox, D.R., 1984. Long-range dependence: a review. In: David, H.A., David, H.T. (Eds.), Statistics: An Appraisal. Iowa State University Press, Iowa, pp. 55–74.
Davidson, J., Sibbertsen, Ph., 2002. Generating schemes for long memory processes: regimes, aggregation and linearity. Preprint.
Davydov, Yu.A., 1970. The invariance principle for stationary processes. Theory of Probability and Applications 15, 487–498.
Davydov, Yu.A., 1973. Mixing conditions for Markov chains. Theory of Probability and Applications 18, 312–328.
Dehling, H., Philipp, W., 2002. Empirical process techniques for dependent data. In: Dehling, H., Mikosch, T., Sørensen, M. (Eds.), Empirical Process Techniques for Dependent Data. Birkhäuser, Boston, pp. 3–113.
Diebold, F.X., Inoue, A., 2001. Long memory and regime switching. Journal of Econometrics 105, 131–159.
Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.
Feller, W., 1971. An Introduction to Probability Theory and its Applications, vol. 2. Wiley, New York.
Franses, P.H., van Dijk, D., 2000. Non-linear Time Series Models in Empirical Finance. Cambridge University Press, Cambridge.
Gourieroux, C., Jasiak, J., 2001. Memory and infrequent breaks. Economics Letters 70, 29–41.
Granger, C.W.J., 1980. Long memory relationships and the aggregation of dynamic models. Journal Econometrics 14, 227–238.
Granger, C.W.J., 2000. Current perspectives on long memory processes. Academia Economic Papers 28, 1–16.
Granger, C.W.J., Hyung, N., 2004. Occasional structural breaks and long memory, with application to the S&P 500 absolute stock returns. Journal of Empirical Finance 11, 399–421.
Granger, C.W.J., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford.
Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.
Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton.
Heath, D., Resnick, S., Samorodnitsky, G., 1998. Heavy tails and long range dependence in on/off processes and associated fluid models. Mathematics of Operations Research 23, 145–165.
Hsieh, M., Hurvich, C.M., Soulier, Ph., 2003. Asymptotics for duration-driven long range dependent processes. Preprint.
Ibragimov, I.A., Linnik, Yu.V., 1971. Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen.
Jensen, M.J., Liu, M., 2001. Do long swings in the business cycle lead to strong persistence in output? Preprint.
Karlsen, H.A., 1990. Existence of moments in a stationary stochastic difference equation. Advances in Applied Probability 22, 129–146.
Leipus, R., Surgailis, D., 2003a. Random coefficient autoregression, regime switching and long memory. Advances in Applied Probability 35, 737–754.
Leipus, R., Surgailis, D., 2003b. Long range dependence in on/off process with dependent on and off intervals. Preprint.
Leipus, R., Viano, M.-C., 2003. Long memory and stochastic trend. Statistics and Probability Letters 61, 177–190.
Leipus, R., Paulauskas, V., Surgailis, D., 2004. Random coefficient AR(1) process with heavy-tailed renewal switching coefficient and heavy tailed noise. Preprint.
Liu, M., 2000. Modeling long memory in stock market volatility. Journal of Econometrics 99, 139–171.
Lobato, I.N., Savin, N.E., 1998. Real and spurious long-memory properties of stock-market data (with comments). Journal of Business and Economic Statistics 16, 261–283.

Lux, T., Sornette, D., 2002. On rational bubbles and fat tails. Journal of Money, Credit, and Banking 34, 589–610.

Mikosch, T., Samorodnitsky, G., 2000. The supremum of a negative drift random walk with dependent heavy-tailed steps. Annals of Applied Probability 10, 1025–1064.

Mikosch, T., Resnick, S., Rootzén, H., Stegeman, A., 2002. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? Annals of Applied Probability 12, 23–68.

Parke, W.R., 1999. What is fractional integration? Review of Economics and Statistics 81, 632–638.

Pipiras, V., Taqqu, M.S., Levy, J.B., 2004. Slow, fast and arbitrary growth conditions for renewal-reward processes when both the renewals and the rewards are heavy-tailed. Bernoulli 10, 121–163.

Pourahmadi, M., 1988. Stationarity of the solution of $X_t = A_t X_{t-1} + \varepsilon_t$ and analysis of non-Gaussian dependent variables. Journal of Time Series Analysis 9, 225–239.

Samorodnitsky, G., Taqqu, M.S., 1994. Stable Non-Gaussian Random Processes. Chapman and Hall, New York.

Skorokhod, A.V., 1964. Random Processes With Independent Increments. Nauka, Moscow (in Russian).

Taqqu, M.S., Levy, J.B., 1986. Using renewal processes to generate long-range dependence and high variability. In: Eberlein, E., Taqqu, M.S. (Eds.), Dependence in Probability and Statistics. Birkhäuser, Boston, pp. 73–89.

Tjøstheim, D., 1986. Some doubly stochastic time series models. Journal of Time Series Analysis 7, 51–72.

Tong, H., 1990. Non-linear Time Series: A Dynamical System Approach. Clarendon Press, Oxford.

Vervaat, W., 1979. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. Advances in Applied Probability 11, 750–783.

Willinger, W., Taqqu, M.S., Sherman, R., Wilson, D.V., 1997. Self-similarity through high variability: Statistical analysis of Ethernet LAN traffic at the source level. IEEE/ACM Transactions in Networking 5, 71–86.

Wu, G., Xiao, Z., 2002. Are there speculative bubbles in stock markets? Evidence from an alternative approach. Preprint.