Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 130 (2020) 1853–1878

stochastic processes and their applications

www.elsevier.com/locate/spa

Stable limits for Markov chains via the Principle of Conditioning

Mohamed El Machkouri[a], Adam Jakubowski[b,*], Dalibor Volný[a]

[a] *Université de Rouen Normandie, LMRS, Avenue de l'Université, BP 12 76801 Saint-Étienne-du-Rouvray cedex, France*
[b] *Nicolaus Copernicus University, Faculty of Mathematics and Computer Science, ul. Chopina 12/18, 87-100 Toruń, Poland*

Received 25 June 2018; received in revised form 2 June 2019; accepted 4 June 2019
Available online 12 June 2019

Abstract

We study limit theorems for partial sums of instantaneous functions of a homogeneous Markov chain on a general state space. The summands are heavy-tailed and the limits are stable distributions.

We show that if the transition operator of the chain is operator uniformly integrable and the chain is ρ-mixing, then the limit is the same as if the summands were independent.

We provide an example of a Markov chain that is operator uniformly integrable (and admits a spectral gap) while it is not hyperbounded.

What makes our assumptions working is a new, efficient version of the Principle of Conditioning.

MSC: 60F05; 60F17; 60E07; 60J05; 60J35

Keywords: Stable laws; Markov chains; Spectral gap; Operator uniform integrability; Principle of conditioning; Hyperbouneduess

1. Introduction

Our motivation comes from the paper by Jara, Komorowski and Olla [36], where a fractional diffusion was obtained as a scaled limit of functionals of Markov chains forming a probabilistic solution to a linear Boltzmann equation. The main tool used in [36] was a functional limit theorem on convergence to stable Lévy processes due to Durret and Resnick [20] and the assumptions that made this functional limit theorem working were L^2-spectral gap and

* Corresponding author.
E-mail addresses: mohamed.elmachkouri@univ-rouen.fr (M. El Machkouri), adjakubo@mat.umk.pl (A. Jakubowski), dalibor.volny@univ-rouen.fr (D. Volný).

https://doi.org/10.1016/j.spa.2019.06.002
0304-4149/

strong contractivity properties of the Markov transition operator. In the particular example considered in [36] the *ultraboundedness* of the transition operator was used, but in the general considerations (Theorem 2.4, *ibid.*) properties related to a weaker notion of *hyperboundedness* were assumed (We refer to Section 2 for formal definitions and discussion of all these notions).

Later Cattiaux and Manou-Abi [12] reexamined the limit theorems from [36] in the context of the general theory of convergence to stable laws for sums of stationary sequences. They considered standard mixing conditions (ϕ-, ρ-, α- mixing) and anti-clustering condition D', introduced in [14] and discussed in [17]. While the discussion in [12] was quite extensive, it did not address the question whether the strong assumption of hyperboundedness of the transition operator can be essentially weakened.

In the present paper we suggest replacing the hyperboundedness with the *uniform integrability in* L^2 (2-U.I. in short) of the transition operator, a notion introduced in [52]. We believe that this is the proper minimal form for operator contractivity whenever limit theorems for Markov chains with stable limits are considered. Our main results are formulated in Section 3. In Theorem 3.1 we obtain limit theorems assuming the 2-U.I. condition and ρ-mixing. In Proposition 3.3 we study the relations between hyperboundedness and ρ-mixing and in Corollary 3.5 we provide the corresponding limit theorems. The proofs of both main results are deferred to Section 5. In Section 2 we gather all necessary information, notation and comments related to the models considered in the paper.

What allows considerable weakening of the assumptions and removing technicalities is a new efficient version of the Principle of Conditioning that operates with conditional characteristic functions rather than predictable characteristics and therefore keeps integrability requirements at the minimal possible level. Recall that the Principle of Conditioning is a heuristic rule that transforms limit theorems for independent random variables into limit theorems for dependent random variables. The mentioned above functional limit theorem by Durret and Resnick [20] is a particular manifestation of this rule. We state our new result (Theorem A.3) and give more comments and references on the Principle of Conditioning in Appendix.

In Section 4 we give five examples, each of a different nature.

First we provide an example of a Markov chain with the transition operator that is uniformly integrable in L^2 (and admits an L^2-spectral gap) while it is not hyperbounded. The constructed Markov chain is integer-valued but can be easily modified to obtain the absolutely continuous stationary distribution on $\mathbb{R}^+$. This shows that our theory substantially extends [36, Theorem 2.4] and [12].

Then, using the standard example of a stationary AR(1) sequence with Gaussian innovations, which is hyperbounded and ρ-mixing, we demonstrate that our theory does not imply ϕ-mixing (or uniform ergodicity). This proves that the hyperboundedness is not as demanding as it looks like.

On the other hand, using our Corollary 3.5, we can show that some popular Markov chains, like ARCH or GARCH processes, *are not* hyperbounded.

Finally we contribute to the problem of m-skeletons, providing an example of a Markov chain such that the 3-skeleton $\{X_{3n}\}$ is not i.i.d. but still satisfies the assumptions of our Theorem 3.1, while the partial sums of the whole sequence remain bounded in probability.

2. Preliminaries

2.1. Transition operator

Let $\{X_n\}_{n\geq 0}$ be a Markov chain with state space $(\mathbb{S}, \mathcal{S})$ and the transition probability $P(x, dy)$ on $\mathbb{S} \times \mathcal{S}$. We will always assume that $P(x, dy)$ admits a stationary distribution π on $(\mathbb{S}, \mathcal{S})$, i.e.

$$\pi(A) = \int_{\mathbb{S}} \pi(dx) P(x, A), \quad A \in \mathcal{S}. \tag{1}$$

The transition probability defines the *transition operator* that acts by the formula

$$(Pf)(x) = \int_{\mathbb{S}} P(x, dy) f(y) \tag{2}$$

and is a positive contraction on every space $L^p(\pi) = L^p(\mathbb{S}, \mathcal{S}, \pi)$, $p \in [1, +\infty]$.

2.2. 2-U.I. condition

Following [52] we will say that the transition operator P is:

uniformly integrable in L^2 (or 2-U.I.) if

$$\{|Pf|^2 ; \|f\|_{L^2(\pi)} \leq 1\} \quad \text{is uniformly } \pi\text{-integrable.} \tag{3}$$

hyperbounded if there exists $q > 2$ such that $P : L^2(\pi) \to L^q(\pi)$ is a bounded linear operator, i.e.

$$\sup\{\int |Pf|^q \, d\pi \, ; \|f\|_{L^2(\pi)} \leq 1\} < +\infty. \tag{4}$$

ultrabounded if

$$\sup\{\|Pf\|_\infty ; \|f\|_{L^1(\pi)} \leq 1\} < +\infty. \tag{5}$$

There are simple relations between these notions. First, we have

$$\int |Pf|^q \, d\pi = \| \, |Pf|^2 \, \|^{q/2}_{L^{q/2}(\pi)},$$

so (4) implies boundedness of $\{|Pf|^2 ; \|f\|_{L^2(\pi)} \leq 1\}$ in $L^{q/2}(\pi)$, hence uniform integrability, if $q > 2$. Therefore the hyperboundedness implies the uniform integrability in L^2. Second, applying the Jensen inequality, we get for any $p > 1$ that (5) implies

$$\sup\{\|Pf\|_\infty ; \|f\|_p \leq 1\} < +\infty.$$

Thus the ultraboundedness implies the hyperboundedness.

Notice that the hyperboundedness of the transition operator is, in a sense, independent of the particular choice of $q > 2$. Indeed, by the Riesz–Thorin theorem, if P is a bounded linear operator from $L^p(\pi)$ to $L^q(\pi)$, with $1 < p < q < +\infty$, then for any other $1 < p' < +\infty$ there is $q' > p'$, $q' < +\infty$, such that P is a bounded linear operator from $L^{p'}$ to $L^{q'}$.

Conditions like (3)–(5) are usually considered in the context of hypercontractivity of Markov semigroups and all examples mentioned in [52] (as well as most of examples in [12]) are related to the continuous time Markov processes analysis.

In the present paper we deal with discrete time Markov chains and show that also in this more elementary setting there are natural examples of Markov chains with contracting properties of the transition operator describable by relations (3)–(5).

For example, suppose that P is given by a density $p(x, y)$ with respect to π, i.e.

$$Pf(x) = \int_{\mathbb{S}} \pi(dy)p(x, y)f(y).$$

Then P is ultrabounded if $p(x, y)$ is a bounded function in (x, y) (as in the main model in [36]), and it is hyperbounded if $p(x, y) \in L^q(\pi \times \pi)$ for some $q > 2$ (see [12, p. 480]). In Section 4.1 we shall provide an example of a countable-space Markov chain with P that is 2-U.I. but not hyperbounded.

Remark 2.1. By the linearity of P, if any of conditions (3)–(5) holds for *real-valued* functions f, then it is satisfied also for *complex-valued* functions f.

2.3. Geometric ergodicity and strong mixing

The chain $\{X_n\}_{n\geq 0}$ is *geometrically ergodic*, if there is a number $0 \leq \eta < 1$ and a function $C : \mathbb{S} \to \mathbb{R}^+$ such that

$$\|P^n(x, \cdot) - \pi\|_{TV} \leq C(x)\eta^n, \quad \text{for } \pi\text{-a.e. } x \in \mathbb{S},\ n \in \mathbb{N},$$

where $\|\cdot\|_{TV}$ is the total variance distance (see [48, Theorem 2.1]). It is well known that the geometric ergodicity of a Markov chain is equivalent (under natural conditions) to the exponential absolute regularity (see e.g. [8, Theorem 21.19, p. 325]), hence implies also *the strong mixing at geometric rate*. The latter property means (in the world of Markov chains) that there is a constant $0 < C < +\infty$ such that for every *bounded* measurable function $h : (\mathbb{S}, \mathcal{S}) \to (\mathbb{R}^1, \mathcal{B}^1)$ and every n

$$\int_{\mathbb{S}} \pi(dx)\Big|P^n h(x) - \int_{\mathbb{S}} h\, d\pi\Big| \leq C\eta^n \|h\|_\infty. \tag{6}$$

See [9, Chapter 4] or [19] for definitions and properties of this and other mixing conditions.

2.4. L^2-spectral gap and ρ-mixing

The transition operator P is said to have an L^2-*spectral gap* if there is a number $a < 1$ such that

$$\sup\{\|Pf\|_{L^2(\pi)}\,;\ \int_{\mathbb{S}} f(x)\, d\pi(x) = 0,\ \|f\|_{L^2(\pi)} \leq 1\} \leq a.$$

If some power P^m admits an L^2-spectral gap of size at least $1-a$, then by iteration we obtain that there is a constant $1 \leq D < +\infty$ such that for $f \in L^2_0(\pi) = \{f \in L^2(\pi)\,;\ \pi(f) = \int_{\mathbb{S}} f(x)\pi(dx) = 0\}$

$$\|P^n f\|_{L^2(\pi)} \leq Da^n \|f\|_{L^2(\pi)}, \quad n = 1, 2, \ldots. \tag{7}$$

This means that $\{X_n\}$ satisfies "an L^2 norm condition" of [49] and by Theorem 2, p. 217, *ibid.*, a central limit theorem with the standard normalization $\sqrt{n}$ holds for the stationary sequence $\Psi(X_0), \Psi(X_1), \ldots$ whenever $\int \Psi(x)\pi(dx) = 0$ and $\int \Psi^2(x)\pi(dx) < +\infty$. (A proof of this limit theorem that is preferred nowadays can be found e.g. in [23]).

In the contemporary language we say that the chain is ρ-mixing (for Markov chains necessarily at the geometric rate). A particular consequence of this property that is crucial for our paper is that for any function $g \in L^2(\pi)$ we have

$$\mathbb{V}\mathrm{ar}\Big(\sum_{j=0}^{n-1} g(X_j)\Big) \le n\Big(1 + 2\frac{D}{(1-a)}\Big)\mathbb{V}\mathrm{ar}\big(g(X_0)\big). \tag{8}$$

For reversible, ψ-irreducible and aperiodic Markov chains the spectral gap property is known to be equivalent to the geometric ergodicity.

If $\{X_n\}$ is not reversible, then the spectral gap property implies the geometric ergodicity (see [38, Theorem 1.3]), but there are Markov chains that are geometrically ergodic and do not have an L^2 spectral gap (see [38, Theorem 1.4]). It is remarkable that the central limit theorem need not hold for such Markov chains (see [7,24,25]).

Notice that if one is interested in a central limit theorem to hold for particular instantaneous function of the underlying Markov chain, then sufficient conditions weaker than the L^2 spectral gap are known (see e.g. [41]).

2.5. *Stable limits*

In the present paper the limiting distribution μ will be *stable* with exponent $\alpha \in (0, 2)$. It is well-known (see e.g. [51] or [27]) that its characteristic function admits the Lévy–Khintchine representation

$$\hat{\mu}(\theta) = \exp\Big(i\theta a^h + \int \big(e^{i\theta x} - 1 - i\theta x \mathbb{1}_{\{|x|\le h\}}\big)\, \nu_{\alpha,c_+,c_-}(dx)\Big), \tag{9}$$

where $c_+, c_- \ge 0$, $c_+ + c_- > 0$ and $a^h \in \mathbb{R}^1$, the Lévy measure ν_{α,c_+,c_-} has the density

$$p_{\alpha,c_+,c_-}(x) = \alpha\Big(c_+ x^{-(\alpha+1)}\mathbb{1}_{\{x>0\}} + c_-|x|^{-(\alpha+1)}\mathbb{1}_{\{x<0\}}\Big),$$

and $h > 0$ is a fixed level of truncation. We will denote the stable distribution with characteristic function (9) by $\delta_{a^h} *$-Poiss$(\alpha, c_+, c_-)c_h$.

In the main results of the paper we shall consider somewhat less general limits μ_α with characteristic function of the form

$$\hat{\mu}_\alpha(\theta) = \begin{cases} \exp\big(\int \big(e^{i\theta x} - 1\big)\nu_{\alpha,c_+,c_-}(dx)\big), & \alpha \in (0,1); \\ \exp\big(\int \big(e^{i\theta x} - 1\big)\nu_{1,c,c}(dx)\big), & \alpha = 1; \\ \exp\big(\int \big(e^{i\theta x} - 1 - i\theta x\big)\nu_{\alpha,c_+,c_-}(dx)\big), & \alpha \in (1,2). \end{cases} \tag{10}$$

A reader familiar with the terminology would observe that completing the above list with probability laws of the form $\delta_a * \mu_1$, $a \ne 0$, we obtain all *strictly stable* laws on $\mathbb{R}^1$

Notice that the integrals under the exponents in (9) or (10) can be evaluated, but obtained this way formulas are usually meaningless within the limit theory.

3. Results

Let $\{X_n\}$ be a Markov chain on the space $(\mathbb{S}, \mathcal{S})$ with a stationary distribution π. Define $\mathcal{F}_n = \sigma\{X_j\,;\ j \le n\}$.

We will study distributional limits for suitably normalized and centered partial sums of the form

$$S_n = \sum_{j=1}^{n} \Psi(X_j),$$

where $\Psi : (\mathbb{S}, \mathcal{S}) \to (\mathbb{R}^1, \mathcal{B}^1)$ is a measurable function.

We will assume that the probability law $\pi \circ \Psi^{-1}$ belongs to the domain of attraction of μ_α, $0 < \alpha < 2$. This means (see e.g. [22, Theorem 1a, p. 313]) that

$$\pi\big(x\,;\ |\Psi(x)| > t\big) = t^{-\alpha}\ell(t), \tag{11}$$

where $\ell(t)$ is a slowly varying function as $t \to \infty$, and there exist the limits

$$\lim_{t\to\infty} \frac{\pi\big(x\,;\ \Psi(x) > t\big)}{\pi\big(x\,;\ |\Psi(x)| > t\big)} = \frac{c_+}{c_+ + c_-}, \quad \lim_{t\to\infty} \frac{\pi\big(x\,;\ \Psi(x) < -t\big)}{\pi\big(x\,;\ |\Psi(x)| > t\big)} = \frac{c_-}{c_+ + c_-}. \tag{12}$$

Theorem 3.1. *Let $\{X_n\}$ be a Markov chain on the space $(\mathbb{S}, \mathcal{S})$, with the transition operator P and a stationary distribution π. We assume that the chain is ρ-mixing and that P satisfies the 2-U.I. condition.*

Let $\Psi : (\mathbb{S}, \mathcal{S}) \to (\mathbb{R}^1, \mathcal{B}^1)$ be such that $\pi \circ \Psi^{-1}$ belongs to the domain of attraction of the stable distribution μ_α, $\alpha \in (0, 2)$ (i.e. both (11) and *(12) are fulfilled). Let $B_n \to \infty$ satisfies*

$$\frac{n}{B_n^\alpha}\ell(B_n) \to c_+ + c_-. \tag{13}$$

(i) *If $\alpha \in (0, 1)$ or $\alpha = 1$ and $c_+ = c_- = c$ then*

$$\frac{\Psi(X_1) + \Psi(X_2) + \cdots + \Psi(X_n)}{B_n} \longrightarrow_{\mathcal{D}} \mu_\alpha. \tag{14}$$

(ii) *If $\alpha \in (1, 2)$, then*

$$\frac{\sum_{j=1}^{n}\big(\Psi(X_j) - \mathbb{E}\,\Psi(X_j)\big)}{B_n} \longrightarrow_{\mathcal{D}} \mu_\alpha, \tag{15}$$

and

$$\frac{\sum_{j=1}^{n}\big(\mathbb{E}\big(\Psi(X_j)|\mathcal{F}_{j-1}\big) - \mathbb{E}\,\Psi(X_j)\big)}{B_n} \underset{\mathcal{P}}{\longrightarrow} 0. \tag{16}$$

Remark 3.2. In the recent paper [44] Mikosch and Wintenberger consider a similar setting of (multidimensional) instantaneous functions of Markov chains. They operate with two types of assumptions. While their drift Condition $\mathbf{DC_p}$ implies geometric ergodicity and is close in spirit to "dynamical" properties of the transition operator P, their other assumption $\mathbf{RV}_\alpha$ (regular variation of finite dimensional distributions, see [2]) is probabilistic in nature and usually requires knowledge of the structure of the Markov chain, going beyond the properties of non-iterated P. See the example in Section 4.3 for the flavor of the results obtained in [44].

As Example 4.1 shows, the hyperboundedness of the transition operator P is strictly stronger than the 2-UI property. On the other hand the hyperboundedness is the most natural sufficient condition for the 2-UI property. It is therefore interesting that the hyperboundedness is close to the L^2-spectral gap property.

Proposition 3.3. *Suppose that the transition operator P is hyperbounded.*

(i) *If the chain is reversible and ergodic, then P has an L^2-spectral gap.*

(ii) *If the chain is strongly mixing at geometric rate (geometrically ergodic), then P^m has an L^2-spectral gap for some $m \in \mathbb{N}$.*

In both cases the chain is ρ-mixing.

Remark 3.4. We owe the statement (ii) and its proof given in Section 5.4 to the anonymous reviewer.

Corollary 3.5. *Let $\{X_n\}$ be a Markov chain on the space $(\mathbb{S}, \mathcal{S})$, with the transition operator P and a stationary distribution π.*

Let $\Psi : (\mathbb{S}, \mathcal{S}) \to (\mathbb{R}^1, \mathcal{B}^1)$ satisfies both (11) and (12) *and $B_n \to \infty$ is defined by* (13).

If P is hyperbounded and the chain is either reversible and ergodic or strongly mixing at geometric rate then

$$\frac{\Psi(X_1) + \Psi(X_2) + \cdots + \Psi(X_n)}{B_n} \longrightarrow_{\mathcal{D}} \mu_\alpha,$$

provided $\alpha \in (0, 1)$ or $\alpha = 1$ and $c_+ = c_- = c$ or $\alpha \in (1, 2)$ and $\int \Psi(x)\pi(dx) = 0$.

Remark 3.6. The fact that hyperboundedness, ergodicity and reversibility imply an L^2-spectral gap is a good piece of mathematics that was recently obtained by Miclo [43].

In the context of our results one can pose another question: is it true that the 2-U.I. property implies an L^2-spectral gap for reversible and ergodic Markov chains?

Remark 3.7. In assumptions of Corollary 3.5 we can strengthen the convergence in probability in (16) to the a.s. convergence.

Indeed, denote by $\|\cdot\|_{s\to t}$ the operator norm of a linear map between $L^s(\pi)$ and $L^t(\pi)$ and suppose that $\|P\|_{2\to q} < +\infty$, for some $q > 2$. Let $\alpha \in (1, 2)$ and take $0 < r < \alpha - 1$. We know that $\mathbb{E}\big|\Psi(X_0)\big|^{\alpha-r} < +\infty$. Since P is also a bounded map from $L^1(\pi)$ to $L^1(\pi)$, we can apply the Riesz–Thorin interpolation theorem (see e.g. [3, Theorem 1.1.1]) to get that

$$\|P\|_{(\alpha-r)\to\beta} \leq \|P\|_{2\to q}^{2(\alpha-r-1)/(\alpha-r)} < +\infty,$$

where

$$\beta = \frac{q(\alpha - r)}{2(q-1) - (q-2)(\alpha - r)} > \alpha.$$

This means that $\mathbb{E}\big|\mathbb{E}\big(\Psi(X_1)\big|\mathcal{F}_0\big)\big|^\beta = \int \pi(dx)|P\Psi(x)|^\beta < +\infty$ for some $\beta > \alpha$. By the remark at the end of Annex C in [47, p. 185] and Corollary 3.2 (i), p. 55 *ibid.*, we have

$$\frac{\sum_{j=1}^n \big(\mathbb{E}\big(\Psi(X_j)|\mathcal{F}_{j-1}\big) - \mathbb{E}\Psi(X_j)\big)}{n^{1/(\beta-\varepsilon)}} \to 0, \quad \text{a.s.},$$

for every $\varepsilon > 0$, $\beta - \varepsilon > \alpha$. But in such a case also $n^{1/(\beta-\varepsilon)}/B_n \to 0$ and our claim follows.

Remark 3.8. It is worth stressing that in our results for $\alpha = 1$ we need only that *the limit* is symmetric and not $\pi \circ \Psi^{-1}$ itself.

4. Examples

4.1. Example related to the 2-U.I. condition

We are going to construct a discrete in time and space example of the transition operator that exhibits the L^2-spectral gap property, satisfies the 2-U.I. condition but is not hyperbounded. This will show that our theory essentially extends Theorem 2.4 of [36] and the results of [12]. Notice also that all the examples of operators provided in [52] and related to the 2-U.I. condition are taken from the stochastic analysis.

Example 4.1. The example is a variant of Rosenblatt's family of examples [49, pp. 213–214], but it occurs also in many other places, e.g. in [42, p.54], in the context of the backward recurrence time chain.

Let $T : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{N} = \{0, 1, 2, \dots\}$ be an integer valued non-negative random variable such that

$$\mathbb{E}T < +\infty, \; \mathbb{P}(T \geq j) > 0, \; j \in \mathbb{N}.$$

(Other requirements imposed on the distribution of T will be specified later). Let the transition probabilities $p_{j,k}$ be given by the formula

$$p_{j,k} = \begin{cases} \dfrac{\mathbb{P}(T = j)}{\mathbb{P}(T \geq j)}, & \text{if } k = 0; \\ \dfrac{\mathbb{P}(T \geq j+1)}{\mathbb{P}(T \geq j)}, & \text{if } k = j+1; \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\pi(j) = \frac{\mathbb{P}(T \geq j)}{1 + \mathbb{E}T}, \; j = 0, 1, 2, \dots,$$

is the unique stationary distribution for $P = [p_{j,k}]$ and the transition operator reads

$$Pf(j) = \frac{\mathbb{P}(T = j)}{\mathbb{P}(T \geq j)} f(0) + \frac{\mathbb{P}(T \geq j+1)}{\mathbb{P}(T \geq j)} f(j+1).$$

Let $\{X_n\}$ be a Markov chain on $\mathbb{S} = \mathbb{N}$ with the transition probabilities $[p_{j,k}]$.

Lemma 4.2. *If* $3\,\mathbb{E}T < \mathbb{P}(T = 0)$ *and*

$$\mathbb{P}(T \geq 1) \geq \sup_{k \geq 1} \frac{\mathbb{P}(T \geq k+1)}{\mathbb{P}(T \geq k)}, \tag{17}$$

then the Markov chain $\{X_n\}$ *has the* L^2*-spectral gap property.*

Proof. Let $f \in L^2_0(\pi)$ and $\|f\|_{L^2(\pi)} = 1$. These relations imply that

$$\big|f(0)\big| = \Big| - \sum_{j=1}^{\infty} f(j)\mathbb{P}(T \geq j)\Big| \leq \sum_{j=1}^{\infty} \big(|f(j)|\sqrt{\mathbb{P}(T \geq j)}\big)\sqrt{\mathbb{P}(T \geq j)}$$

$$\leq \sqrt{\sum_{j=1}^{\infty} f^2(j)\mathbb{P}(T \geq j)}\sqrt{\sum_{j=1}^{\infty} \mathbb{P}(T \geq j)} = \sqrt{\big(1 + \mathbb{E}T - f^2(0)\big)\mathbb{E}T}.$$

Hence

$$|f(0)| \le \sqrt{\mathbb{E}T}. \tag{18}$$

In a similar way we obtain

$$\begin{aligned}\Big|\sum_{j=1}^{\infty} f(j)\frac{\mathbb{P}(T=j-1)}{\mathbb{P}(T\ge j-1)}\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T}\Big| &\le \sum_{j=1}^{\infty}|f(j)|\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T}\\ &\le \sqrt{\sum_{j=1}^{\infty} f^2(j)\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T}}\sqrt{\sum_{j=1}^{\infty}\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T}} \le \sqrt{\frac{\mathbb{E}T}{1+\mathbb{E}T}}.\end{aligned} \tag{19}$$

We are ready for estimates of $\mathbb{E}_\pi\left(|Pf|^2\right) = \left(1/(1+\mathbb{E}T)\right)\sum_{j=0}^{\infty}\left|Pf(j)\right|^2\mathbb{P}(T\ge j)$.

$$\begin{aligned}\sum_{j=0}^{\infty}\left|Pf(j)\right|^2\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T} &= \sum_{j=0}^{\infty}\left|\frac{\mathbb{P}(T=j)}{\mathbb{P}(T\ge j)}f(0)+\frac{\mathbb{P}(T\ge j+1)}{\mathbb{P}(T\ge j)}f(j+1)\right|^2\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T}\\ &= \frac{f^2(0)}{1+\mathbb{E}T}\sum_{j=0}^{\infty}\frac{\mathbb{P}^2(T=j)}{\mathbb{P}(T\ge j)} + \frac{2f(0)}{1+\mathbb{E}T}\sum_{j=0}^{\infty}f(j+1)\frac{\mathbb{P}(T=j)}{\mathbb{P}(T\ge j)}\mathbb{P}(T\ge j+1)\\ &\quad + \frac{1}{1+\mathbb{E}T}\sum_{j=0}^{\infty}f^2(j+1)\frac{\mathbb{P}^2(T\ge j+1)}{\mathbb{P}(T\ge j)} = J_1+J_2+J_3.\end{aligned}$$

We have by (18)

$$J_1 \le \frac{\mathbb{E}T}{1+\mathbb{E}T}\sum_{j=0}^{\infty}\frac{\mathbb{P}^2(T=j)}{\mathbb{P}(T\ge j)} \le \mathbb{E}T,$$

while by (18) and (19)

$$J_2 \le 2\sqrt{\mathbb{E}T}\sqrt{\frac{\mathbb{E}T}{1+\mathbb{E}T}} \le 2\mathbb{E}T.$$

Finally, by (17),

$$J_3 \le \mathbb{P}(T\ge 1)\sum_{j=1}^{\infty}f^2(j)\frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T} \le \mathbb{P}(T\ge 1).$$

Therefore

$$\mathbb{E}_\pi\left(|Pf|^2\right) \le 3\,\mathbb{E}T + \mathbb{P}(T\ge 1) = 1-\left(\mathbb{P}(T=0)-3\mathbb{E}T\right) = a < 1.$$

The proof of Lemma 4.2 is complete.

It remains to show that for some specific distribution of T the 2-U.I. condition holds, but there is no hyperboundedness. Choose $\gamma\in(0,1)$ and set

$$\mathbb{P}(T\ge 1)=\gamma, \mathbb{P}(T\ge 2)=\gamma^3,\ldots,\mathbb{P}(T\ge j)=\gamma^{1+2+\cdots+j}=\gamma^{j(j+1)/2},\ldots.$$

Clearly, $\mathbb{P}(T\ge j+1)/\mathbb{P}(T\ge j)=\gamma^{j+1},\ j=0,1,2,\ldots$ and for $\gamma<1/5$

$$\mathbb{E}T < \frac{\gamma}{1-\gamma} < (1/3)(1-\gamma) = (1/3)\mathbb{P}(T=0),$$

so that the assumptions of Lemma 4.2 are satisfied and the corresponding Markov chain $\{X_n\}$ has the L^2-spectral gap property.

In order to prove that the 2-U.I. condition holds, it is enough to show that

$$\lim_{k\to\infty} \sup_{\|f\|_{L^2(\pi)}\le 1} \sum_{j=k}^{\infty} \big|Pf(j)\big|^2 \frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T} = 0.$$

Notice that $\|f\|_{L^2(\pi)} \le 1$ implies that $f^2(j) \le (1+\mathbb{E}T)/\mathbb{P}(T\ge j)$, $j = 0,1,2,\ldots$ Keeping this in mind we can proceed as follows.

$$\begin{aligned}
&\sum_{j=k}^{\infty} \big|Pf(j)\big|^2 \frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T} \\
&\qquad\le \frac{2f^2(0)}{1+\mathbb{E}T}\sum_{j=k}^{\infty}\frac{\mathbb{P}^2(T=j)}{\mathbb{P}(T\ge j)} + \frac{2}{1+\mathbb{E}T}\sum_{j=k}^{\infty} f^2(j+1)\frac{\mathbb{P}^2(T\ge j+1)}{\mathbb{P}(T\ge j)} \\
&\qquad\le 2\mathbb{P}(T\ge k) + 2\sum_{j=k}^{\infty}\frac{\mathbb{P}(T\ge j+1)}{\mathbb{P}(T\ge j)} = 2\mathbb{P}(T\ge k) + 2\sum_{j=k}^{\infty}\gamma^{j+1} \to 0.
\end{aligned}$$

Next consider a sequence $\{f_k\}$ of functions in $L^2(\pi)$ given by

$$f_k(j) = \begin{cases} \sqrt{\dfrac{1+\mathbb{E}T}{\mathbb{P}(T\ge k)}}, & \text{if } j=k; \\ 0, & \text{otherwise.}\end{cases}$$

Take any $q > 2$. We have, if $k\to\infty$,

$$\begin{aligned}
\|Pf_k\|^q_{L^q(\pi)} &= \sum_{j=0}^{\infty} \big|Pf_k(j)\big|^q \frac{\mathbb{P}(T\ge j)}{1+\mathbb{E}T} \\
&= \left(\frac{1+\mathbb{E}T}{\mathbb{P}(T\ge k)}\right)^{q/2}\left(\frac{\mathbb{P}(T\ge k)}{\mathbb{P}(T\ge k-1)}\right)^{q}\frac{\mathbb{P}(T\ge k-1)}{1+\mathbb{E}T} \\
&= \big(1+\mathbb{E}T\big)^{q/2-1}\frac{\big(\mathbb{P}(T\ge k)\big)^{q/2}}{\big(\mathbb{P}(T\ge k-1)\big)^{q-1}} = \big(1+\mathbb{E}T\big)^{q/2-1}\gamma^{w(k)} \to +\infty,
\end{aligned}$$

for $w(k) = qk(k+1)/4 - (q-1)k(k-1)/2 = (1/4)\big(k^2(2-q)+k(3q-2)\big) \to -\infty$. It follows that the transition operator P cannot be a bounded linear map from $L^2(\pi)$ to $L^q(\pi)$.

Example 4.3. On the probability space $(\mathbb{N},\pi)$ considered in Example 4.1 one cannot define Ψ with the distributional properties (11) and (12). This drawback disappears, however, after easy modification.

Let $\{X_k\}_{k\ge0}$ be the Markov chain from the previous example and let $\{\epsilon_k\}_{k\ge0}$ be an i.i.d. sequence of random variables distributed uniformly on $(0,1)$ that is independent of $\{X_k\}$. Set

$$Y_k = X_k + \epsilon_k,\ \ k = 0,1,2,\ldots.$$

Then it is not difficult to check that $\{Y_k\}$ is a Markov chain on $(\mathbb{R}^+,\mathcal{B}^+)$ with the transition probabilities ($Leb(B)$ means the Lebesgue measure of B and $\lfloor x\rfloor$ is the integer part of x)

$$\begin{aligned}
\mathbb{P}\big(Y_1\in A\big|Y_0=x\big) &= \mathbb{P}\big(Y_1\in A\big|X_0=\lfloor x\rfloor\big) \\
&= \mathbb{P}\big(X_1=0\big|X_0=\lfloor x\rfloor\big)Leb\big(A\cap[0,1)\big) \\
&\quad+ \mathbb{P}\big(X_1=\lfloor x+1\rfloor\big|X_0=\lfloor x\rfloor\big)Leb\big(A\cap[\lfloor x+1\rfloor,\lfloor x+2\rfloor)\big)
\end{aligned}$$

and the stationary distribution

$$\widetilde{\pi}(A) = \sum_{j=0}^{\infty} \mathbb{P}\big(X_0 = j\big) Leb\big(A \cap [j, j+1)\big).$$

The transition operator $\widetilde{P}$ acts similarly to P from the previous example.

$$(\widetilde{P} f)(x) = \Big(\int_0^1 f(u) d u\Big) \mathbb{P}\big(X_1 = 0 \big| X_0 = \lfloor x \rfloor\big) + \Big(\int_{\lfloor x+1 \rfloor}^{\lfloor x+2 \rfloor} f(u) d u\Big) \mathbb{P}\big(X_1 = \lfloor x+1 \rfloor \big| X_0 = \lfloor x \rfloor\big).$$

Therefore the 2-U.I. and the L^2-spectral gap properties, as well as the lack of hyperboundedness can be verified by obvious adaptation of the corresponding computations performed for P. And the stationary distribution $\widetilde{\pi}$ is absolutely continuous so that our probability space $(\mathbb{S}, \mathcal{S}, \widetilde{\pi})$ supports a function Ψ with the desired distributional properties.

4.2. *Gaussian hyperboundedness*

Let us examine a standard example — the stationary $AR(1)$ process with Gaussian innovations, already considered by Doob [18, p. 218]. For $0 < |\rho| < 1$ set

$$P(x, dy) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}} \, dy, \quad \pi(dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$

It is well-known that the Markov chain $\{X_n\}$ corresponding to $P(x, dy)$ has the L^2-spectral gap property. Using [12, p. 480], we shall show that $\{X_n\}$ is also hyperbounded. Indeed, $P(x, dy) = p(x, y)\pi(dy)$, where

$$p(x, y) = \frac{1}{\sqrt{1-\rho^2}} \exp\Big(-\frac{\rho^2}{2(1-\rho^2)} x^2 + \frac{\rho x y}{1-\rho^2} - \frac{\rho^2}{2(1-\rho^2)} y^2\Big).$$

And we have $\int \pi(dx)\pi(dy) p(x, y)^q < +\infty$, whenever $2 < q < \frac{1+|\rho|}{|\rho|}$. Applying Theorem 3.1 (or Corollary 3.5) we obtain

$$\frac{\Psi(X_1) + \Psi(X_2) + \cdots + \Psi(X_n)}{B_n} \longrightarrow_{\mathcal{D}} \mu_\alpha, \tag{20}$$

for suitably chosen Ψ and B_n. Notice that this simple example is not covered by Theorem 2.4 in [36], because relation (2.9) of Condition 2.3 *ibid.* is not satisfied.

We evoke this classic example for two reasons. It was Doob [18, p. 218] who pointed out that this Markov chain does not satisfy Doeblin's condition (D). And since the work of Davydov [16] we know that Doeblin's condition means essentially ϕ-mixing of a Markov chain. It follows that the limit theory developed in our paper is much broader than results depending on uniform ergodicity of Markov chains, as presented e.g. in [13].

The other reason is that $\{\Psi(X_n)\}$ is a particular case of a more general example considered by Davis in [14]. Dealing with stationary sequences and using analogs of the well-known in the Extreme Value Limit Theory conditions D and D', Davis proved in Theorem 2, p. 265 *ibid.* a limit theorem for $\alpha \in (0, 1)$ and adding a technical condition D'' he proved in Theorem 3, p.266, *ibid.* a limit theorem for $\alpha \in [1, 2)$. As was observed in [33, Theorem 4.2] condition D'' is satisfied by ρ-mixing sequences with sufficiently fast decay.

It is possible that our 2-U.I. condition implies condition D'. But we are not able to prove this fact.

The main example considered by Davis and described on p. 267 *ibid.* is a sequence of instantaneous functions of a stationary Gaussian sequence (possibly non-Markovian) with the covariances $r_n = \mathbb{E}X_0X_n$ satisfying either $r_n \log n \to 0$ or $\sum_{n=1}^{\infty} r_n^2 < +\infty$ (clearly, our $\{X_n\}$ fulfills these requirements). For detailed discussion of various aspects of results based on condition D' we refer to [17,39] and [12].

4.3. ARCH processes with heavy tails are not hyperbounded

An ARCH(1) process is a Markov chain given by the recurrence formula

$$X_{j+1} = \sqrt{\beta + \lambda X_j^2} Z_{j+1}, \ \ j \geq 0, \tag{21}$$

where $\beta, \lambda > 0$ and $\{Z_n\}_{n\in\mathbb{N}}$ is an i.i.d. sequence, independent of X_0. In order to comply with references we shall assume that $Z_n \sim \mathcal{N}(0, 1)$.

For basic information on ARCH processes and the properties used below we refer both to the classic book [21] and to the recent source [11].

In the range of parameters $\beta > 0$ and $\lambda \in (0, 2e^{\gamma})$ (where γ is the Euler constant) the process $\{X_j\}$ admits a stationary distribution given by

$$X_0 \sim r_0 \sqrt{\beta \sum_{m=1}^{\infty} Z_m^2 \prod_{j=1}^{m-1} (\lambda Z_j^2)},$$

where r_0 is a Rademacher random variable ($P(r_0 = \pm 1) = 1/2$), independent of $\{Z_n\}$. This stationary distribution exhibits power decay of the tails. Namely, if $\kappa > 0$ is the unique positive solution of the equation $\mathbb{E}(\lambda Z_1^2)^u = 1$, then, as $x \to \infty$,

$$\mathbb{P}\big(X_0 > x\big) = \mathbb{P}\big(X_0 < -x\big) \sim \frac{C_{\beta,\lambda}}{2} \, x^{-2\kappa}, \tag{22}$$

where

$$C_{\beta,\lambda} = \frac{\mathbb{E}\Big[\big(\beta + \lambda X_0^2\big)^{\kappa} - \big(\lambda X_0^2\big)^{\kappa}\Big]}{\kappa \lambda^{2\kappa} E\Big[\big(Z_1^{2\kappa} \ln(\lambda Z_1^2)\big)\Big]} \in (0, +\infty).$$

It follows that $\lambda > 1$ implies "really" heavy tails and it is likely that the partial sums of $\{X_j\}$ properly normalized converge to stable laws. Indeed, Davis and Mikosch [15] showed that the partial sums under the natural normalization converge to *some* stable limit and Bartkiewicz et al. [1] identified the parameters of the limit.

For purposes of the present example, let us denote by $\mu_{\alpha,\tau}$ the symmetric α-stable distribution given for $\alpha \in (0, 2)$ and $\tau > 0$ by

$$\widehat{\mu_{\alpha,\tau}}(\theta) = \exp\Big(\tau\alpha \int_{\mathbb{R}} \big(e^{i\theta u} - 1\big)|u|^{-(\alpha+1)} \, du\Big).$$

If our Theorem 3.1 or Corollary 3.5 were applicable to $\{X_j\}_{j\geq 0}$, then we would have

$$\frac{X_1 + X_2 + \cdots + X_n}{(nC_{\beta,\lambda})^{\frac{1}{2\kappa}}} \longrightarrow_{\mathcal{D}} \mu_{2\kappa,1}.$$

It is, however, proved in [1] that

$$\frac{X_1 + X_2 + \cdots + X_n}{(nC_{\beta,\lambda})^{\frac{1}{2\kappa}}} \longrightarrow_{\mathcal{D}} \mu_{2\kappa,\tau},$$

where $0 < \tau = E\big[|1 + S_\infty|^{2\kappa} - |S_\infty|^{2\kappa}\big] < 1$ and the series

$$S_\infty = \sum_{j=1}^{\infty} \lambda^{j/2} \Big[\prod_{k=1}^{j-1} |Z_k|\Big] Z_j$$

converges a.s.

Moreover, ARCH(1) processes are strongly mixing at geometric rate, as is shown in [15, p. 2077] (see also [2, Theorem 2.8]).

Therefore the transition operator of an ARCH(1) process *is not hyperbounded*, if $\lambda > 1$.

Remark 4.4. If $2\kappa \in (1, 2)$, then the corresponding ARCH(1) process $\{X_j\}$ forms a stationary sequence of martingale differences, partial sums of which normalized by $n^{1/2\kappa}$ are weakly convergent, but to a different limit than in the independent case. This is in striking contrast to the properties of martingale difference sequences with finite variance!

Remark 4.5. If $\lambda \in (0, 1)$, then by taking $\Psi(x) = \text{const} \cdot |x|^{\upsilon} \text{sign}(x)$, with υ sufficiently large, we obtain a Markov chain $\{\Psi(X_j)\}$ satisfying the distributional relations (11) and (12), related to some ARCH process $\{X'_j\}$ with $\lambda' > 1$. There seems to be, however, no simple correspondence between the asymptotics of partial sums of $\{\Psi(X_j)\}$ and $\{X'_j\}$ and therefore we are not able to reduce the case $\lambda \in (0, 1)$ to the case $\lambda > 1$.

4.4. m-skeletons

It is well known that iterating the transition operator improves its properties from many viewpoints. So it may happen that some power P^m is hyperbounded, for instance, while P itself not. Such situation implies that for $\{\Psi(X_{k\cdot m})\}_{k=0,1,2,\ldots}$ (the m-skeleton) some α-stable limit theorem holds and one may hope to extend this property to the whole sequence. This is impossible in general, as the simple counterexample provided already by Rosenblatt [49, p. 195] shows. Indeed, take an i.i.d. sequence $\{Y_n\}$ of strictly stable random variables and consider a Markov chain on $\mathbb{S} = \mathbb{R}^2$ given by the formula $X_n = (Y_n, Y_{n-1})$. Take $\Psi(x, y) = x - y$. Then $\sum_{j=0}^{n-1} \Psi(X_n)$ remains stochastically bounded while the 2-skeleton consists of independent random variables and therefore satisfies the corresponding limit theorem.

Rosenblatt's example is of probabilistic provenience. Some people may prefer another example given below that is closer to thinking in terms of dynamical systems.

Example 4.6. Set $\mathbb{S} = [0, 3)$ and let *Leb* be the Lebesgue measure restricted to $\mathbb{S}$. For $x \in [0, 1)$ and $B \in \mathcal{B}_{[0,1)} \cup \mathcal{B}_{[2,3)}$ define

$$P(x, \{x + 1\}) = P(x + 1, \{x + 2\}) = 1, \;\; P(x + 2, B) = \frac{Leb(B)}{2}.$$

The invariant measure π is given by the density $p(x) = \frac{1}{4}\mathbb{1}_{[0,2)}(x) + \frac{1}{2}\mathbb{1}_{[2,3)}(x)$. Elementary calculations show that for $f \in L^2_0(\pi)$ we have $\mathbb{E}_\pi\Big(\big(P^3 f\big)^2\Big) \le \frac{27}{32}\mathbb{E}_\pi\big(f^2\big)$, i.e. the 3-skeleton has the spectral gap property. Another elementary calculation shows that also $\|P^3 f\|_\infty \leq 3\|f\|_1$, i.e. the 3-skeleton is ultrabounded.

Now take $\psi(\cdot) : [0,1) \to \mathbb{R}^1$ with a symmetric α-stable distribution μ and define

$$\Psi(x) = \begin{cases} \psi(x), & \text{if } x \in [0,1); \\ -\psi(x-1), & \text{if } x \in [1,2); \\ 0, & \text{if } x \in [2,3). \end{cases}$$

One verifies directly that

$$\pi\big(\Psi > r\big) = \frac{1}{2} Leb\big(\psi > r\big), \quad \pi\big(\Psi < -r\big) = \frac{1}{2} Leb\big(\psi < -r\big).$$

Therefore the 3-skeleton $\{\Psi(X_{3n})\}$ satisfies all assumptions of our Theorem 3.1, while the partial sums of the whole sequence $\{\Psi(X_k)\}$ are bounded in probability.

5. Proofs

5.1. Some auxiliary results

We begin with establishing an important property of conditional distributions $P(x, dy) \circ \Psi^{-1}$ that is a consequence of solely (11)–(12).

Proposition 5.1. *Suppose that (11) and (12) hold. Let B_n be defined by (13). Then*

$$n\Big|1 - \mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2 \underset{\mathcal{P}}{\longrightarrow} 0, \quad \theta \in \mathbb{R}^1. \tag{23}$$

Proof. Recall that if B_n is defined by (13) then $B_n = n^{1/\alpha}\tilde{\ell}(n)$, where $\tilde{\ell}(t)$ is a slowly varying function. Let $h > 0$ be fixed. Using the inequality $|1 + ix - e^{ix}| \leq \frac{1}{2}|x|^2$, we have

$$\begin{aligned} & n\Big|1 - \mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2 \\ & \leq 2n\Big|1 + i\theta\mathbb{E}\big(\frac{\Psi(X_1)}{B_n}\mathbb{1}_{\{|\Psi(X_1)|\leq hB_n\}}\big|\mathcal{F}_0\big) - \mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2 \\ & \qquad + 2nB_n^{-2}\theta^2\Big|\mathbb{E}\big(\Psi(X_1)\mathbb{1}_{\{|\Psi(X_1)|\leq hB_n\}}\big|\mathcal{F}_0\big)\Big|^2 \\ & \leq nB_n^{-4}\theta^4\Big(\mathbb{E}\big(\Psi(X_1)^2\mathbb{1}_{\{|\Psi(X_1)|\leq hB_n\}}\big|\mathcal{F}_0\big)\Big)^2 + 16n\Big(\mathbb{P}\big(|\Psi(X_1)| > hB_n\big|\mathcal{F}_0\big)\Big)^2 \\ & \qquad + 2nB_n^{-2}\theta^2\Big|\mathbb{E}\big(\Psi(X_1)\mathbb{1}_{\{|\Psi(X_1)|\leq hB_n\}}\big|\mathcal{F}_0\big)\Big|^2 = \theta^4 I_{n,1}^h + 16 I_{n,2}^h + 2\theta^2 I_{n,3}^h. \end{aligned} \tag{24}$$

At first we shall examine the convergence of $I_{n,3}^h$. If $\alpha \in (1,2)$ then

$$\Big|\mathbb{E}\big(\Psi(X_1)\mathbb{1}_{\{|\Psi(X_1)|\leq hB_n\}}\big|\mathcal{F}_0\big)\Big|^2 \to \Big|\mathbb{E}\big(\Psi(X_1)\big|\mathcal{F}_0\big)\Big|^2 \text{ a.s.},$$

while $nB_n^{-2} = n^{1-2/\alpha}(\tilde{\ell}(n))^{-2} \to 0$. Consequently, $I_{n,3}^h \to 0$ a.s.

Now suppose that $\alpha \in (0,1]$. Take $0 < r < \alpha/2$. We have

$$\mathbb{E}\Big(\mathbb{E}\big(|\Psi(X_1)|^{\alpha-r}\big|\mathcal{F}_0\big)\Big) = \mathbb{E}|\Psi(X_1)|^{\alpha-r} < +\infty,$$

and so

$$(\alpha - r)\int_0^\infty t^{\alpha-r-1} P\big(X_0, |\Psi|^{-1}(t, +\infty)\big)\,dt = \mathbb{E}\big(|\Psi(X_1)|^{\alpha-r}\big|\mathcal{F}_0\big) < +\infty \text{ a.s.}$$

It follows that

$$\begin{aligned}
I^h_{n,3} &= nB_n^{-2}\Big|\mathbb{E}\big(\Psi(X_1)\mathbb{1}_{\{|\Psi(X_1)|\le hB_n\}}\big|\mathcal{F}_0\big)\Big|^2\\
&\le nB_n^{-2}\Big|\mathbb{E}\big(\big|\Psi(X_1)\big|\mathbb{1}_{\{|\Psi(X_1)|\le hB_n\}}\big|\mathcal{F}_0\big)\Big|^2\\
&\le nB_n^{-2}\Big|\int_0^{hB_n} P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big|^2\\
&= nB_n^{-2}\Big|\int_0^{hB_n} t^{1-\alpha+r}t^{\alpha-r-1}P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big|^2\\
&\le nB_n^{-2}h^{2(1-\alpha+r)}B_n^{2(1-\alpha+r)}\Big|\int_0^{\infty} t^{\alpha-r-1}P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big|^2\\
&= n^{-1+2r/\alpha}\big(\tilde{\ell}(n)\big)^{-2(\alpha-r)}h^{2(1-\alpha+r)}\Big(\frac{1}{\alpha-r}\mathbb{E}\big(|\Psi(X_1)|^{\alpha-r}\big|\mathcal{F}_0\big)\Big)^2 \to 0, \text{ a.s.}
\end{aligned}$$

Similarly, if $\alpha \in (0,2)$ and $0 < r < \alpha/2$, then we have

$$\begin{aligned}
I^h_{n,1} &= nB_n^{-4}\Big|\mathbb{E}\big(\Psi(X_1)^2\mathbb{1}_{\{|\Psi(X_1)|\le hB_n\}}\big|\mathcal{F}_0\big)\Big|^2\\
&\le 4nB_n^{-4}\Big(\int_0^{hB_n} t\,P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big|^2\Big)\\
&= 4nB_n^{-4}\Big(\int_0^{hB_n} t^{2-\alpha+r}t^{\alpha-r-1}P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big)^2\\
&\le 4nB_n^{-4}h^{2(2-\alpha+r)}B_n^{2(2-\alpha+r)}\Big(\int_0^{\infty} t^{\alpha-r-1}P\big(X_0, |\Psi|^{-1}(t,+\infty)\big)\,dt\Big)^2\\
&= 4n^{-1+2r/\alpha}\big(\tilde{\ell}(n)\big)^{-2(\alpha-r)}h^{2(2-\alpha+r)}\Big(\frac{1}{\alpha-r}\mathbb{E}\big(|\Psi(X_1)|^{\alpha-r}\big|\mathcal{F}_0\big)\Big)^2 \to 0, \text{ a.s.}
\end{aligned}$$

It remains to show that $I^h_{n,2} \underset{\mathcal{P}}{\longrightarrow} 0$. This condition is not related to truncated moments and therefore requires a different type argument. Notice that the convergence in probability is metrizable and so it is enough to show that in every subsequence n' one can find a further subsequence n'' along which $I^h_{n'',2} \underset{\mathcal{P}}{\longrightarrow} 0$. So choose n' and consider random variables $Y_{n'}$ defined on $(\mathbb{S}, \mathcal{S}, \pi)$ by the formula

$$Y_{n'}(x) = n'P\big(x, |\Psi|^{-1}(hB_{n'}, +\infty)\big).$$

We know from (11), (12), (13) and the continuity of the stable Lévy measure that

$$\int_{\mathbb{S}} \pi(dx)Y_{n'}(x) = n'\mathbb{P}\big(|\Psi(X_1)| > hB_{n'}\big) \to (c_+ + c_-)h^{-\alpha},$$

hence, in particular, random variables $\{Y_{n'}\}$ are uniformly tight. Let $\{n''\}$ be a subsequence such that $Y_{n''} \longrightarrow_{\mathcal{D}} Y_\infty$. By the Skorokhod representation theorem one can construct random variables $\widetilde{Y}_{n''}$ and $\widetilde{Y}_\infty$, defined on the standard probability space $\big([0,1], \mathcal{B}_{[0,1]}, Leb\big)$ and such that

$$\widetilde{Y}_{n''} \sim Y_{n''}, \quad \widetilde{Y}_\infty \sim Y_\infty,$$

and

$$\widetilde{Y}_{n''}(\omega) \to \widetilde{Y}_\infty(\omega), \quad \text{for almost all } \omega \in [0,1].$$

This implies that

$$\frac{1}{n''}\widetilde{Y}^2_{n''}(\omega) \to 0, \quad \text{for almost all } \omega \in [0,1].$$

But under the initial distribution π we have

$$n''\Big(P\big(x, |\Psi|^{-1}(hB_{n''}, +\infty)\big)\Big)^2 = \frac{1}{n''}Y^2_{n''} \sim \frac{1}{n''}\widetilde{Y}^2_{n''}. \tag{25}$$

It follows that

$$n''\Big(P\big(x, |\Psi|^{-1}(hB_{n''}, +\infty)\big)\Big)^2 \underset{\mathcal{P}}{\longrightarrow} 0.$$

Remark 5.2. It is clear that the convergences $I^h_{n,1} \underset{\mathcal{P}}{\longrightarrow} 0$ and $I^h_{n,3} \underset{\mathcal{P}}{\longrightarrow} 0$ can be obtained also by the last method. But the proofs given above lead to the a.s. convergence and provide some idea about the rate of convergence.

Remark 5.3. It is also clear that relation (25) can be extended to

$$(n'')^{1+\delta}\Big(P\big(x, |\Psi|^{-1}(hB_{n''}, +\infty)\big)\Big)^2 = \frac{1}{(n'')^{1-\delta}}Y^2_{n''} \sim \frac{1}{(n'')^{1-\delta}}\widetilde{Y}^2_{n''},$$

hence, in fact, we have

$$n^\delta I^h_{n,2} \underset{\mathcal{P}}{\longrightarrow} 0,$$

for every $\delta \in [0, 1)$. Gathering information on $I^h_{n,1}$, $I^h_{n,2}$ and $I^h_{n,3}$ we obtain existence of some $\delta > 0$ such that

$$n^{1+\delta}\Big|1 - \mathbb{E}\big(e^{i\theta\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2 \underset{\mathcal{P}}{\longrightarrow} 0, \quad \theta \in \mathbb{R}^1.$$

5.2. An application of the principle of conditioning

Now we are ready to prove a universal (i.e. independent of $\alpha \in (0, 2)$) limit theorem.

Proposition 5.4. *Let $\{X_n\}$ be a Markov chain on the space $(\mathbb{S}, \mathcal{S})$, with the transition operator P and a stationary distribution π. We assume that P satisfies the 2-U.I. condition and the chain is ρ-mixing.*

Let $\alpha \in (0, 2)$ and $h > 0$. Let $\Psi : (\mathbb{S}, \mathcal{S}) \to (\mathbb{R}^1, \mathcal{B}^1)$ be such that $\pi \circ \Psi^{-1}$ belongs to the domain of attraction of the stable distribution μ_α, $\alpha \in (0, 2)$ (i.e. both (11) and (12) are fulfilled). Let $B_n \to \infty$ be defined by (13).

Set $S^h_n = \sum_{j=1}^n \Psi(X_j) - \mathbb{E}\big(\Psi(X_j)\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\big|\mathcal{F}_{j-1}\big)$. Then

$$\frac{S^h_n}{B_n} \longrightarrow_{\mathcal{D}} c_h\text{-}Poiss(\alpha, c_+, c_-). \tag{26}$$

Proof. Choose $\theta \in \mathbb{R}^1$ and notice that by Proposition 5.1 relation (23) holds. We will show that this relation can be strengthened to

$$n\mathbb{E}\Big|1 - \mathbb{E}\big(e^{i\theta\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2 \to 0. \tag{27}$$

It is enough to show that $n\big|1-\mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\big|^2$ is a uniformly integrable sequence. By the 2-U.I. condition we have to prove that the sequence $\{Z_n=\sqrt{n}\big(1-e^{i\theta\,\Psi(X_1)/B_n}\big)\}$ is bounded in L^2.

$$\begin{aligned}
&\mathbb{E}\Big|\sqrt{n}\big(1-e^{i\theta\,\Psi(X_1)/B_n}\big)\Big|^2\\
&\quad= n\,\mathbb{E}\Big(\big(1-\cos(\theta\,\Psi(X_1)/B_n)\big)^2+\big(\sin(\theta\,\Psi(X_1)/B_n)\big)^2\Big) \qquad (28)\\
&\quad\le \theta^2(1+\theta^2/4)\frac{n}{B_n^2}\mathbb{E}\,\Psi(X_1)^2\mathbb{1}_{\{|\,\Psi(X_1)|\le B_n\}}+5n\mathbb{P}\big(|\,\Psi(X_1)|>B_n\big)\\
&\quad\le \theta^2(1+\theta^2/4)\frac{n}{B_n^2}2\int_0^{B_n} t\mathbb{P}\big(|\,\Psi(X_1)|>t\big)\,dt+5n\mathbb{P}\big(|\,\Psi(X_1)|>B_n\big).
\end{aligned}$$

The last expression converges to $\big(2\theta^2(1+\theta^2/4)+5\big)(c_++c_-)$ by (13) and the direct half of the Karamata theorem (see [6, Theorem 1.5.11, p. 28]).

Given (27) we obtain the crucial relation (40)

$$\mathbb{E}\Big(\sum_{j=1}^n\Big|1-\mathbb{E}\left(e^{i\theta\,\Psi(X_j)/B_n}|\mathcal{F}_{j-1}\right)\Big|^2\Big)=n\mathbb{E}\Big|1-\mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big)\Big|^2\to 0,\quad \theta\in\mathbb{R}^1.$$

By Theorem A.3 it is enough to prove (41), i.e.

$$\begin{aligned}
\Phi_n^h(\theta) &:= \sum_{j=1}^n \mathbb{E}\big(e^{i\theta\,\Psi(X_j)/B_n}|\mathcal{F}_{j-1}\big)-1-i\theta B_n^{-1}\mathbb{E}\big(\Psi(X_j)\mathbb{1}_{\{|\,\Psi(X_j)|\le hB_n\}}\big|\mathcal{F}_{j-1}\big) \qquad (29)\\
&\xrightarrow[\mathcal{P}]{} \int\big(e^{i\theta x}-1-i\theta x\mathbb{1}_{\{|x|\le h\}}\big)\,\nu_{\alpha,c_+,c_-}(dx)=:\Phi^h(\theta).
\end{aligned}$$

Let us notice that by (11) and (12) we have $\mathbb{E}\,\Phi_n^h(\theta)\to\Phi^h(\theta)$. Taking all these facts together we obtain the final condition to be verified:

$$\sum_{j=1}^n \chi_{n,\theta}^h(X_{j-1})\xrightarrow[\mathcal{P}]{}0, \qquad (30)$$

where

$$\begin{aligned}
\chi_{n,\theta}^h(x) &= \int\big(\exp\big(i\theta\,\Psi(y)/B_n\big)-1-i\theta\big(\Psi(y)/B_n\big)\mathbb{1}_{\{|\,\Psi(y)|\le hB_n\}}\big)P(x,dy)\\
&\quad-\Big(\int\big(\exp\big(i\theta\,\Psi(y)/B_n\big)-1-i\theta\big(\Psi(y)/B_n\big)\mathbb{1}_{\{|\,\Psi(y)|\le hB_n\}}\big)\pi(dy)\Big).
\end{aligned}$$

By (8), for some finite constant D' we have

$$\mathbb{V}\mathrm{ar}\Big(\sum_{j=1}^n\chi_{n,\theta}^h(X_{j-1})\Big)\le nD'\mathbb{V}\mathrm{ar}\big(\chi_{n,\theta}^h(X_0)\big)=n\mathbb{V}\mathrm{ar}\big(W_{n,\theta}^h\big)\le n\mathbb{E}\big|W_{n,\theta}^h\big|^2,$$

where

$$W_{n,\theta}^h=1+i\theta\mathbb{E}\Big(\frac{\Psi(X_1)}{B_n}\mathbb{1}_{\{|\,\Psi(X_1)|\le hB_n\}}\Big|\mathcal{F}_0\Big)-\mathbb{E}\big(e^{i\theta\,\Psi(X_1)/B_n}\big|\mathcal{F}_0\big).$$

Therefore it is enough to prove that $n\mathbb{E}\big|W_{n,\theta}^h\big|^2\to 0$. By inspection of (24) we see that

$$n\big|W_{n,\theta}^h\big|^2\le\frac{1}{2}\theta^4I_{n,1}^h+8I_{n,2}^h\xrightarrow[\mathcal{P}]{}0.$$

As before, this convergence can be strengthened to the convergence in L^1 by applying the 2-U.I. condition. Indeed, both sequences

$$\{Z_n' = \sqrt{n}\big(\frac{\Psi(X_1)^2}{B_n^2}\mathbb{1}_{\{|\Psi(X_1)|\le hB_n\}}\big)\} \text{ and } \{Z_n'' = \sqrt{n}\mathbb{1}_{\{|\Psi(X_1)|>hB_n\}}\}$$

are bounded in L^2 by arguments essentially identical to those used in (28) in the proof of L^2-boundedness of $\{Z_n = \sqrt{n}\big(1 - e^{i\theta\Psi(X_1)/B_n}\big)\}$. Our L^1- claim follows from the observation that

$$n|W_{n,\theta}^h|^2 \le \frac{1}{2}\theta^4\Big(\mathbb{E}\big(Z_n'|\mathcal{F}_0\big)\Big)^2 + 8\Big(\mathbb{E}\big(Z_n''|\mathcal{F}_0\big)\Big)^2.$$

We have thus completed the proof of Proposition 5.4.

5.3. *Proof of Theorem 3.1*

Given (26), i.e.

$$\frac{\sum_{j=1}^n \Psi(X_j) - \mathbb{E}\big(\Psi(X_j)\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\big|\mathcal{F}_{j-1}\big)}{B_n} \longrightarrow_{\mathcal{D}} c_h\text{-Poiss}(\alpha, c_+, c_-),$$

we shall apply classic Theorem 4.2 from [5] in a way suitable for each case $\alpha \in (0, 1)$, $\alpha = 1$ or $\alpha \in (1, 2)$.

The reasoning for $\alpha \in (0, 1)$ is based on the direct half of Karamata's theorem [6, Theorem 1.5.11, p. 28]. We shall show that

$$\lim_{h\to 0}\limsup_{n\to\infty}\mathbb{E}\Big|\sum_{j=1}^n \mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big| = 0, \tag{31}$$

and that

$$c_h\text{-Poiss}(\alpha, c_+, c_-) \Rightarrow \mu_\alpha, \quad \text{as } h \to 0.$$

The latter relation holds because $\int |x|\mathbb{1}_{\{|x|\le 1\}}\,\nu_{\alpha,c+,c_-}(dx) < +\infty$ for $\alpha \in (0, 1)$. In order to prove (31) we proceed in the standard way.

$$\begin{aligned}\mathbb{E}\Big|\sum_{j=1}^n\mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big| &\le n\mathbb{E}\Big(\Big|\frac{\Psi(X_1)}{B_n}\Big|\mathbb{1}_{\{|\Psi(X_1)|\le hB_n\}}\Big)\\ &\le \frac{n}{B_n}\int_0^{hB_n}\mathbb{P}\big(|\Psi(X_1)| > t\big)\,dt \to_{n\to\infty} (1-\alpha)^{-1}h^{1-\alpha}(c_+ + c_-) \to_{h\to 0} 0.\end{aligned}$$

The proof for $\alpha = 1$ is somewhat different. Let us notice first that due to the symmetry of $\nu_{1,c,c}$ we have the equality

$$c_h\text{-Poiss}(\alpha, c, c) = \mu_1, \quad h \in \mathbb{R}^1,$$

hence

$$c_h\text{-Poiss}(\alpha, c, c) \Rightarrow_{h\to 0} \mu_1.$$

Let $h > h' > 0$. By (29) we have also

$$\Phi_n^h(\theta) - \Phi_n^{h'}(\theta) = -i\theta\sum_{j=1}^n B_n^{-1}\mathbb{E}\big(\Psi(X_j)\mathbb{1}_{\{h'B_n<|\Psi(X_j)|\le hB_n\}}\big|\mathcal{F}_{j-1}\big) \underset{\mathcal{P}}{\longrightarrow} 0, \quad \theta \in \mathbb{R}^1.$$

Therefore

$$\lim_{h'\to 0}\limsup_{n\to\infty}\mathbb{P}\Big(\Big|\sum_{j=1}^n \mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{h'B_n<|\Psi(X_j)|\le hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big|>\varepsilon\Big)=0,\quad \varepsilon>0.$$

Now let $\alpha\in(1,2)$. Find $\zeta_n^h\in L_0^2(\pi)$ such that

$$\mathbb{E}\big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\big|\mathcal{F}_{j-1}\big)-\mathbb{E}\big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\big)=\zeta_n^h(X_{j-1}),\ \ j\in\mathbb{N}.$$

We shall prove that for every $h>0$

$$\sum_{j=1}^n \zeta_n^h(X_{j-1}) \underset{\mathcal{P}}{\longrightarrow} 0. \tag{32}$$

Similarly as before, by ρ-mixing $\mathbb{V}\text{ar}\Big(\sum_{j=1}^n \zeta_n^h(X_{j-1})\Big) \le nD'\mathbb{V}\text{ar}\big(\zeta_n^h(X_0)\big)$, and our task consists in proving that the last expression converges to 0.

By (11) and (12) and the general limit theorem for triangular arrays of row-wise independent random variables (see e.g. Theorem 2.35 in [27]) we have

$$n\mathbb{V}\text{ar}\Big(\frac{\Psi(X_1)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\Big)\longrightarrow_{n\to\infty}\int_{-h}^{h}x^2\nu_{\alpha,c_+,c_-}(dx)<+\infty.$$

Therefore the sequence $\{n\,\big|\zeta_n^h(X_0)\big|^2\}$ is uniformly integrable by the 2-U.I. property. On the other hand

$$\begin{aligned} n\,\big|\zeta_n^h(X_0)\big|^2 &\le 2nB_n^{-2}\Big|\mathbb{E}\Big(\Psi(X_j)\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big|^2 \\ &\quad + 2nB_n^{-2}\Big|\mathbb{E}\Big(\Psi(X_j)\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\Big)\Big|^2 \to 0, \text{ a.s,}\end{aligned}$$

by the same argument as in the verification of $I_{n,3}^h\to 0$ a.s. in the proof of Proposition 5.1. Hence $n\,\mathbb{E}\big|\zeta_n^h(X_0)\big|^2\to 0$, and so (32) holds.

It follows that in (26) we can replace conditional expectations with expectations, i.e. for every $h>0$

$$\frac{\sum_{j=1}^n\big(\Psi(X_j)-\mathbb{E}\,\Psi(X_j)\mathbb{1}_{\{|\Psi(X_j)|\le hB_n\}}\big)}{B_n}\longrightarrow_{\mathcal{D}} c_h\text{-Poiss}(\alpha,c_+,c_-). \tag{33}$$

Applying the direct part of the Karamata Theorem we obtain

$$\begin{aligned} n\Big|\mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|>hB_n\}}\Big)\Big| &\le n\mathbb{E}\Big(\Big|\frac{\Psi(X_1)}{B_n}\Big|\mathbb{1}_{\{|\Psi(X_1)|>hB_n\}}\Big) \\ &\le \frac{n}{B_n}\Big(\int_{hB_n}^{\infty}\mathbb{P}\big(|\Psi(X_1)|>t\big)\,dt + hB_n\mathbb{P}\big(|\Psi(X_1)|>hB_n\big)\Big) \\ &\to_{n\to\infty}\frac{\alpha}{\alpha-1}h^{1-\alpha}(c_++c_-)\to_{h\to\infty}0.\end{aligned} \tag{34}$$

We have also

$$c_h\text{-Poiss}(\alpha,c_+,c_-)\Rightarrow\mu_\alpha,\quad \text{as } h\to\infty, \tag{35}$$

due to the fact that $\int|x|\mathbb{1}_{\{|x|\ge 1\}}\,\nu_{\alpha,c+,c-}(dx)<+\infty$, if $\alpha\in(1,2)$. Relations (33)–(35) prove (15).

In order to prove (16) we use (32), (34) and

$$\lim_{h\to\infty}\limsup_{n\to\infty}\mathbb{E}\Big|\sum_{j=1}^{n}\mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\geq hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big| = 0.$$

The above statement holds because

$$\mathbb{E}\Big|\sum_{j=1}^{n}\mathbb{E}\Big(\frac{\Psi(X_j)}{B_n}\mathbb{1}_{\{|\Psi(X_j)|\geq hB_n\}}\Big|\mathcal{F}_{j-1}\Big)\Big| \leq n\mathbb{E}\Big(\Big|\frac{\Psi(X_1)}{B_n}\Big|\mathbb{1}_{\{|\Psi(X_1)|\geq hB_n\}}\Big)$$
$$\to_{n\to\infty}\frac{\alpha}{\alpha-1}h^{1-\alpha}(c_+ + c_-)\to_{h\to\infty} 0.$$

5.4. *Proof of* Proposition 3.3

(i) If P is hyperbounded, self-adjoint and ergodic then it admits an L^2-spectral gap by [43, Theorem 1].

(ii) Suppose that $\|P\|_{2\to q} = H < +\infty$. Keeping in mind that $\|P\|_{1\to 1} = 1$ we obtain by the Riesz–Thorin interpolation theorem that for $\theta_0 = (q-2)/(3q-2)$

$$\|P\|_{2/(1+\theta_0)\to 2/(1-\theta_0)} \leq H^{\theta_0}$$

But $\|P\|_{r\to r} = 1$ for every $r \geq 1$ hence also

$$\|P^m\|_{2/(1+\theta_0)\to 2/(1-\theta_0)} \leq H^{\theta_0},\ m\in\mathbb{N}.$$

Let us consider an operator Q_m given by

$$\big(Q_m f\big)(x) = P^m\big(f - \int_{\mathbb{S}} f\,d\pi\big)(x).$$

We have $\|Q_m(f)\|_{2/(1-\theta_0)} \leq H^{\theta_0}\|f - \int_{\mathbb{S}} f d\pi\|_{2/(1+\theta_0)} \leq 2H^{\theta_0}\|f\|_{2/(1+\theta_0)}$, hence

$$\|Q_m\|_{2/(1+\theta_0)\to 2/(1-\theta_0)} \leq 2H^{\theta_0}. \tag{36}$$

On the other hand, due to the strong mixing at geometric rate we have

$$\int_{\mathbb{S}}\pi(dx)\big|(Q_m h)(x)\big| = \int_{\mathbb{S}}\pi(dx)\big|(P^m h)(x) - \int_{\mathbb{S}} h\,d\pi\big| \leq C\eta^m\|h\|_\infty,$$

for some $C > 0$ and $0 < h < 1$ (see (6)), what gives

$$\|Q_m\|_{\infty\to 1} \leq C\eta^m. \tag{37}$$

Given (36) and (37) we again apply the Riesz–Thorin theorem with $\theta_1 = \theta_0/(1+\theta_0)$ and obtain

$$\|Q_m\|_{2\to 2} \leq \big(2H^{\theta_0}\big)^{\theta_1}\big(C\eta^m\big)^{1-\theta_1}.$$

It follows that there exists $m_0 \in \mathbb{N}$ such that $\|Q_{m_0}\|_{2\to 2} \leq 1/2$. This means that for every function $g \in L^2_0(\pi)$

$$\|P^{m_0} g\|_{L^2(\pi)} = \|Q_{m_0} g\|_{L^2(\pi)} \leq \|g\|_{L^2(\pi)}/2.$$

Acknowledgments

We would like to express our gratitude to the anonymous reviewers for their insightful and inspiring opinions on the paper which helped us substantially improve the results. We also thank Zbigniew Szewczak for pointing us the article by Wu [52].

A. Jakubowski gratefully acknowledges the support of the University of Rouen Normandie.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Complements on the principle of conditioning

As mentioned in Introduction, the Principle of Conditioning (PoC) is a heuristic rule that allows producing limit theorems for dependent random variables given limit theorems for independent random variables. For example, applying the PoC one obtains the following theorem on convergence to stable laws.

Theorem A.1. *Let* $\{X_{n,j}\,;\ j\in\mathbb{N}, n\in\mathbb{N}\}$ *be an array of random variables, which are row-wise adapted to a sequence of filtrations* $\{\{\mathcal{F}_{n,j}\,;\ j=0,1,\ldots\}\,;\ n\in\mathbb{N}\}$. *Let* $h>0$ *and let* $k_n\to\infty$ *be a sequence of numbers.*

The following conditions

$$\max_{1\le j\le k_n}\mathbb{P}\big(|X_{n,j}|>\varepsilon\big|\mathcal{F}_{n,j-1}\big)\underset{\mathcal{P}}{\longrightarrow}0,\quad \varepsilon>0;$$

$$\sum_{j=1}^{k_n}\mathbb{P}\big(X_{n,j}>x|\mathcal{F}_{n,j-1}\big)\underset{\mathcal{P}}{\longrightarrow}c_+x^{-\alpha},\quad x>0;$$

$$\sum_{j=1}^{k_n}\mathbb{P}\big(X_{n,j}<x\big|\mathcal{F}_{n,j-1}\big)\underset{\mathcal{P}}{\longrightarrow}c_-|x|^{-\alpha},\quad x<0;$$

$$\sum_{j=1}^{k_n}\mathbb{E}\big(X_{n,j}\mathbb{1}_{\{|X_{n,j}|\le h\}}\big|\mathcal{F}_{n,j-1}\big)\underset{\mathcal{P}}{\longrightarrow}a^h;$$

$$\sum_{j=1}^{k_n}\mathbb{V}ar\big(X_{n,j}\mathbb{1}_{\{|X_{n,j}|\le h\}}\big|\mathcal{F}_{n,j-1}\big)\underset{\mathcal{P}}{\longrightarrow}\int_{\{|x|\le h\}}x^2\,\nu_{\alpha,c_+,c_-}(dx);$$

imply that

$$\sum_{j=1}^{k_n}X_{n,j}\longrightarrow_{\mathcal{D}}\delta_{a^h}*c_h\text{-}Poiss(\alpha,c_+,c_-),\tag{38}$$

where $\delta_{a^h}*c_h\text{-}Poiss(\alpha,c_+,c_-)$ *is the stable distribution with the characteristic function (9).*

In other words the PoC says that if we replace in a limit theorem for row-wise independent summands:

- the expectations by conditional expectations with respect to the past,
- the convergence of numbers by convergence in probability of random variables appearing in the conditions,

then still the conclusion (in our case: (38)) will hold. In fact, one can also replace the summation to constants by summation to stopping times.

We refer to [29] for exposition of results related to various versions of the PoC, beginning with the Brown–Eagleson martingale CLT [10], through multidimensional [4,37] and functional [20,26] limit theorems, up to the PoC in infinite dimensional Hilbert [28,30] and Banach [50] spaces. The ideas standing behind the PoC motivated further research devoted to so called decoupling inequalities, described in detail in the well-known books by Kwapień and Woyczyński [40] and de la Peña and Giné [45]. It might be interesting to realize that the tools developed to cope with the PoC find unexpected applications even today [34,46].

Behind the verbal form of the PoC there is a result on convergence of conditional characteristic functions (see [28]).

Theorem A.2. *Let the system $\{X_{n,j}, \mathcal{F}_{n,j}\}$ be as in Theorem A.1. If for some $z \in \mathbb{C}, z \neq 0$, we have*

$$\phi_n(\theta) = \prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) \underset{\mathcal{P}}{\longrightarrow} z,$$

then also

$$\mathbb{E}\exp(i\theta \sum_{j=1}^{k_n} X_{n,j}) \longrightarrow z.$$

In particular, if for some probability measure μ on $\mathbb{R}^1$ we have

$$\phi_n(\theta) \underset{\mathcal{P}}{\longrightarrow} \hat{\mu}(\theta), \quad \theta \in \mathbb{R}^1, \tag{39}$$

then

$$\sum_{j=1}^{k_n} X_{n,j} \longrightarrow_{\mathcal{D}} \mu.$$

Mimicking the case of independent random variables one can prove that conditions obtained by the PoC imply (39). But in many cases this is not the most efficient way of applying the PoC. It was observed in [31] that for highly structured models we can often check (39) directly and that going this way we can keep integrability requirements at the minimal possible level.

We extend the results of [28,32] and [31] in the following theorem that provides a convenient tool in many cases of interest.

Theorem A.3. *Let $\{X_{n,j}\,;\, j \in \mathbb{N}, n \in \mathbb{N}\}$ be an array of random variables, which are row-wise adapted to a sequence of filtrations $\{\{\mathcal{F}_{n,j}\,;\, j = 0, 1, \ldots\}\,;\, n \in \mathbb{N}\}$.*

Suppose that the following condition holds.

$$\sum_{j=1}^{k_n} \left|1 - \mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right)\right|^2 \underset{\mathcal{P}}{\longrightarrow} 0, \quad \theta \in \mathbb{R}^1. \tag{40}$$

Let A_n be arbitrary random variables and $\Phi(\theta) \in \mathbb{C}$ be a constant for each $\theta \in \mathbb{R}^1$. The following conditions are equivalent:

$$\Big(\sum_{j=1}^{k_n} \left(\mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) - 1\right)\Big) - i\theta A_n \underset{\mathcal{P}}{\longrightarrow} \Phi(\theta). \tag{41}$$

$$\Big(\prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right)\Big) e^{-i\theta A_n} \underset{\mathcal{P}}{\longrightarrow} e^{\Phi(\theta)}. \tag{42}$$

In either case we have also

$$\mathbb{E}\exp(i\theta(\sum_{j=1}^{k_n} X_{n,j} - A_n)) \longrightarrow e^{\Phi(\theta)}. \tag{43}$$

In particular, if $e^{\Phi(\theta)} = \hat{\mu}(\theta)$, $\theta \in \mathbb{R}^1$, *for some probability measure* μ, *then either of conditions* (41) *or* (42) *implies*

$$\sum_{j=1}^{k_n} X_{n,j} - A_n \longrightarrow_{\mathcal{D}} \mu.$$

Proof. Set

$$\phi_n(\theta) = \prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right); \quad \Phi_n(\theta) = \sum_{j=1}^{k_n} \big(\mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) - 1\big).$$

If $z \in \mathbb{C}$ satisfies $|z| \leq 1$, then $|z - e^{z-1}| \leq 5|z-1|^2$. Hence we have

$$\begin{aligned} |\phi_n(\theta)e^{-i\theta A_n} - \exp\big(\Phi_n(\theta) - i\theta A_n\big)| &= |\phi_n(\theta) - \exp\big(\Phi_n(\theta)\big)| \\ &\leq \sum_{j=1}^{k_n} |\mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) - \exp\big(\mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) - 1\big)| \\ &\leq 5\sum_{j=1}^{k_n} \left|\mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right) - 1\right|^2 \underset{\mathcal{P}}{\longrightarrow} 0, \quad \text{by (40)}. \end{aligned}$$

We have thus established the equivalence of (41) and (42). To prove that (42) implies (43) we need a suitable version of Lemma 2 in [35].

Lemma A.4. *For every* $\varepsilon > 0$

$$\begin{aligned} &\left|\mathbb{E}\exp(i\theta(\sum_{j=1}^{k_n} X_{n,j} - A_n)) - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\right| \\ &\quad \leq 2(1 + \frac{1}{\varepsilon})\mathbb{P}\big(|\phi_n(\theta)| < \varepsilon\big) + \frac{1}{\varepsilon}\mathbb{E}\left|\phi_n(\theta)e^{-i\theta A_n} - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\right|. \end{aligned} \tag{44}$$

Proof. We follow the idea of the proof of Theorem A in [28]. Define

$$\phi_{n,k}(\theta) = \prod_{j=1}^{k} \mathbb{E}\left(e^{i\theta X_{n,j}} | \mathcal{F}_{n,j-1}\right).$$

Fix $\theta \in \mathbb{R}^1$ and $\varepsilon > 0$ and consider random variables

$$X^*_{n,k} = X_{n,k} \mathbb{1}_{\{|\phi_{n,k}(\theta)| \geq \varepsilon\}}.$$

Then we have both

$$\left|\mathbb{E}\exp(i\theta(\sum_{j=1}^{k_n} X_{n,j} - A_n)) - \mathbb{E}\exp(i\theta(\sum_{j=1}^{k_n} X^*_{n,j} - A_n))\right| \leq 2\mathbb{P}\big(|\phi_n(\theta)| < \varepsilon\big),$$

and, if we set $\phi_n^*(\theta) = \prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}^*} | \mathcal{F}_{n,j-1}\right)$,

$$\mathbb{E}\big|e^{-i\theta A_n}\phi_n(\theta) - e^{-i\theta A_n}\phi_n^*(\theta)\big| \le 2\mathbb{P}\big(|\phi_n(\theta)| < \varepsilon\big).$$

The advantage of random variables $\{X_{n,j}^*\}$ consists in the fact that

$$|\phi_n^*(\theta)| = \Big|\prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}^*} | \mathcal{F}_{n,j-1}\right)\Big| \ge \varepsilon,$$

and so, by the backward induction (or the martingale property)

$$\mathbb{E}\frac{\exp(i\theta\big(\sum_{j=1}^{k_n} X_{n,j}^* - A_n\big))}{e^{-i\theta A_n}\phi_n^*(\theta)} = \mathbb{E}\frac{\exp(i\theta\big(\sum_{j=1}^{k_n} X_{n,j}^*\big))}{\prod_{j=1}^{k_n} \mathbb{E}\left(e^{i\theta X_{n,j}^*} | \mathcal{F}_{n,j-1}\right)} = 1.$$

Therefore,

$$\begin{aligned}
&\big|\mathbb{E}\exp(i\theta\big(\sum_{j=1}^{k_n} X_{n,j}^* - A_n\big)) - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\big| \\
&\quad = \Big|\mathbb{E}\frac{\exp(i\theta\big(\sum_{j=1}^{k_n} X_{n,j}^*\big))}{\phi_n^*(\theta)}\phi_n^*(\theta)e^{-i\theta A_n} \\
&\qquad\qquad - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\mathbb{E}\frac{\exp(i\theta\big(\sum_{j=1}^{k_n} X_{n,j}^*\big))}{\phi_n^*(\theta)}\Big| \\
&\quad \le \frac{1}{\varepsilon}\mathbb{E}|\phi_n^*(\theta)e^{-i\theta A_n} - \mathbb{E}\phi_n(\theta)e^{-i\theta A_n}| \\
&\quad \le \frac{1}{\varepsilon}\Big(2\mathbb{P}\big(|\phi_n(\theta)| < \varepsilon\big) + \mathbb{E}\big|\phi_n(\theta)e^{-i\theta A_n} - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\big|\Big).
\end{aligned}$$

Proof of Theorem A.3(continued). Now assume that (42) holds. Let $\varepsilon = 1/2|e^{\Phi(\theta)}|$. Then $\mathbb{P}\big(|\phi_n(\theta)| < \varepsilon\big) = \mathbb{P}\big(|\phi_n(\theta)e^{-i\theta A_n}| < \varepsilon\big) \to 0$ and by the dominated convergence $\mathbb{E}\big|\phi_n(\theta)e^{-i\theta A_n} - \mathbb{E}\big(\phi_n(\theta)e^{-i\theta A_n}\big)\big| \to 0$.

References

[1] K. Bartkiewicz, A. Jakubowski, T. Mikosch, O. Wintenberger, Stable limits for sums of dependent infinite variance random variables, Probab. Theory Related Fields 150 (2011) 337–372.

[2] B. Basrak, R.A. Davis, T. Mikosch, Regular variation of GARCH processes, Stochastic Process. Appl. 99 (2002) 95–115.

[3] J. Bergh, J. Löfström, Interpolation Spaces. An Introduction, in: Grundlehren Math. Wiss, vol. 223, Springer, Heidelberg, 1976.

[4] M. Beśka, A. Kłopotowski, L. Słomiński, Limit theorems for random sums of dependent d-dimensional random vectors, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 61 (1982) 43–57.

[5] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[6] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, in: Encyclopedia Math. Appl., vol. 27, Cambridge Univ. Press, Cambridge, 1987.

[7] R.C. Bradley, Information regularity and the central limit question, Rocky Mt. J. Math. 13 (1983) 77–97.

[8] R.C. Bradley, Introduction to Strong Mixing Conditions, Vol. II, Kendrick Press, Heber City, 2007.

[9] R.C. Bradley, Introduction to Strong Mixing Conditions, Vol. I, Kendrick Press, Heber City, 2007.

[10] B.M. Brown, G.K. Eagleson, Martingale convergence to infinitely divisible laws with finite variances, Trans. Amer. Math. Soc. 162 (1971) 449–453.

[11] D. Buraczewski, E. Damek, T. Mikosch, Stochastic Models with Power-Law Tails. the Equation $X = AX + B$, Springer, 2016.

[12] P. Cattiaux, M. Manou-Abi, Limit theorems for some functionals with heavy tails of a discrete time markov chain, ESAIM Probab. Stat. 18 (2014) 468–486.

[13] R. Cogburn, The central limit theorem for Markov processes, in: L.M. Le Cam, J. Neyman, E.L. Scott (Eds.), Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory, University of California Press, 1972, pp. 485–512.

[14] R.A. Davis, Stable limits for partial sums of dependent random variables, Ann. Probab. 11 (1983) 262–269.

[15] R.A. Davis, T. Mikosch, Limit theory for the sample acf of stationary process with heavy tails with applications to ARCH, Ann. Probab. 26 (1998) 2049–2080.

[16] Yu. A. Davydov, Mixing conditions for Markov chains, Teor. Veroyatnost. Primienen. 18 (1973) 321–338.

[17] M. Denker, A. Jakubowski, Stable limit distributions for strongly mixing sequences, Statist. Probab. Lett. 8 (1989) 477–483.

[18] J.L. Doob, Stochastic Processes, Wiley, New York, 1953.

[19] P. Doukhan, Mixing: Properties and Examples, in: Lecture Notes in Statist, vol. 85, Springer, New York, 1994.

[20] R. Durret, S. Resnick, Functional limit theorems for dependent random variables, Ann. Probab. 6 (1978) 829–846.

[21] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, Springer, Berlin, 1997.

[22] W. Feller, An Introduction to Probability Theory and its Applications. Volume II, second ed., Wiley, New York, 1970.

[23] M.I. Gordin, B.A. Lifshitz, A central limit theorem for Markov processes, Sov. Math. Dokl. 19 (1978) 392–394.

[24] O. Häggström, On the central limit theorem for geometrically ergodic Markov chains, Probab. Theory Related Fields 132 (2005) 74–82.

[25] O. Häggström, Acknowledgment of priority concerning on the central limit theorem for geometrically ergodic Markov chains, Probab. Theory Related Fields 135 (2006) 470.

[26] J. Jacod, A. Kłopotowski, J. Mémin, Théorème de la limite centrale et convergence fonctionelle vers un processus à accroissements indépendants: la méthode des martingales, Ann. Inst. H. Poincaré, Sect. B 18 (1982) 1–45.

[27] J. Jacod, A. Shiryayev, Limit Theorems for Stochastic Processes, second ed., in: Grundlehren Math. Wiss., vol. 288, Springer, Heidelberg, 2003.

[28] A. Jakubowski, On limit theorems for sums of dependent Hilbert space valued random variables, Lect. Notes Statist. 2 (1980) 178–187.

[29] A. Jakubowski, Principle of conditioning in limit theorems for sums of random variables, Ann. Probab. 14 (1986) 902–915.

[30] A. Jakubowski, Tightness criteria for random measures with application to the principle of conditioning in Hilbert spaces, Probab. Math. Statist. 9 (1988) 95–114.

[31] A. Jakubowski, Principle of conditioning revisited, Demonstratio Math. XLV (2012) 325–336.

[32] A. Jakubowski, A. Kłopotowski, Quelques remarques sur les démonstrations des théorèmes limite pour des vecteurs d-dimensionells aléatoires non indépendants, Publ. Sém. Probab. Rennes exp. no 3 (1980) 1–16.

[33] A. Jakubowski, M. Kobus, α-Stable limit theorems for sums of dependent random vectors, J. Multivariate Anal. 29 (1989) 219–251.

[34] A. Jakubowski, M. Riedle, Stochastic integration with respect to cylindrical Lévy processes, Ann. Probab. 45 (2017) 4273–4306.

[35] A. Jakubowski, L. Słomiński, Extended convergence to continuous in probability processes with independent increments, Probab. Theory Related Fields 72 (1986) 55–82.

[36] M. Jara, T. Komorowski, S. Olla, Limit theorems for additive functionals of a Markov chain, Ann. Appl. Probab. 19 (2009) 2270–2300.

[37] A. Kłopotowski, Limit theorems for sums of dependent random vectors in $\mathbb{R}^d$, Diss. Math. 151 (1977) 1–62.

[38] I. Kontoyiannis, S. Meyn, Geometric ergodicity and the spectral gap of non-reversible Markov chains, Probab. Theory Related Fields 154 (2012) 327–339.

[39] D. Krizmanic, Functional Limit Theorems for Weakly Dependent Regularly Varying Time Series Ph.D. thesis, University of Zagreb, 2010.

[40] S. Kwapień, W.A. Woyczyński, RandOm Series and Stochastic Integrals: Single and Multiple, Birkhäuser, Basel, 1992.

[41] M. Maxwell, M. Woodroofe, Central limit theorems for additive functionals of Markov chains, Ann. Probab. 28 (2000) 713–724.
[42] S. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, second ed., Cambridge, Cambridge, 2009.
[43] L. Miclo, On hyperboundedness and spectrum of Markov operators, Invent. math. 200 (2015) 311–343.
[44] T. Mikosch, O. Wintenberger, The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains, Probab. Theory Related Fields 159 (2014) 157–196.
[45] V. de la Peña, E. Giné, Decoupling: From Dependence to Independence, Springer, New York, 1999.
[46] G. Peccati, M.S. Taqqu, Stable convergence of generalized l^2 stochastic integrals and the principle of conditioning, Electron. J. Probab. 12 (2007) 447–480.
[47] E. Rio, Asymptotic Theory of Weakly Dependent Random Processes, Springer, Heidelberg, 2017.
[48] G.O. Roberts, J.S. Rosenthal, Geometric ergodicity and hybrid Markov chains, Electron. Commun. Probab. 2 (1997) 13–25.
[49] M. Rosenblatt, Markov Processes, in: Structure and Asymptotic Behavior, Springer Verlag, New York, 1971.
[50] J. Rosiński, Central limit theorems for dependent random vectors in banach spaces, in: J.-A. Chao, W.A. Woyczyński (Eds.), Martingale Theory in Harmonic Analysis and Banach Spaces, in: Lecture Notes in Math., vol. 939, 1982, pp. 157–180.
[51] G. Samorodnitsky, M.S. Taqqu, Stable Non-Gaussian RandOm Processes: Stochastic Models with Infinite Variance, Chapman & Hall/CRC, Boca Raton, 1994.
[52] L. Wu, Uniformly integrable operators and large deviations for Markov processes, J. Funct. Anal. 172 (2000) 301–376.