



---

Year: 2020

---

## Scaling exponents of step-reinforced random walks

Bertoin, Jean

Abstract: Let  $X_1, X_2, \dots$  be i.i.d. copies of some real random variable  $X$ . For any deterministic  $2, 3, \dots$  in  $(0, 1)$ , a basic algorithm introduced by H.A. Simon yields a reinforced sequence  $X^1, X^2, \dots$  as follows. If  $n = 0$ , then  $X^n$  is a uniform random sample from  $X^1, \dots, X^{n-1}$ ; otherwise  $X^n$  is a new independent copy of  $X$ . The purpose of this work is to study the scaling exponents of the sequence  $X^1 + \dots + X^n$  with that of its step-reinforced version  $S^{(n)} = X^1 + \dots + X^n$ . Depending on the tail of  $X$  and on asymptotic behavior of the so-called elephant random walk and its variations.

DOI: <https://doi.org/10.1007/s00440-020-01008-2>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-193745>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Bertoin, Jean (2020). Scaling exponents of step-reinforced random walks. Probability Theory and Related Fields:online.

DOI: <https://doi.org/10.1007/s00440-020-01008-2>



# Scaling exponents of step-reinforced random walks

Jean Bertoin<sup>1</sup>

Received: 17 February 2020 / Revised: 16 September 2020 / Accepted: 25 September 2020  
© The Author(s) 2020

## Abstract

Let  $X_1, X_2, \dots$  be i.i.d. copies of some real random variable  $X$ . For any deterministic  $\varepsilon_2, \varepsilon_3, \dots$  in  $\{0, 1\}$ , a basic algorithm introduced by H.A. Simon yields a reinforced sequence  $\hat{X}_1, \hat{X}_2, \dots$  as follows. If  $\varepsilon_n = 0$ , then  $\hat{X}_n$  is a uniform random sample from  $\hat{X}_1, \dots, \hat{X}_{n-1}$ ; otherwise  $\hat{X}_n$  is a new independent copy of  $X$ . The purpose of this work is to compare the scaling exponent of the usual random walk  $S(n) = X_1 + \dots + X_n$  with that of its step reinforced version  $\hat{S}(n) = \hat{X}_1 + \dots + \hat{X}_n$ . Depending on the tail of  $X$  and on asymptotic behavior of the sequence  $(\varepsilon_n)$ , we show that step reinforcement may speed up the walk, or at the contrary slow it down, or also does not affect the scaling exponent at all. Our motivation partly stems from the study of random walks with memory, notably the so-called elephant random walk and its variations.

**Keywords** Reinforcement · Random walk · Scaling exponent · Heavy tail distribution

**Mathematics Subject Classification** 60G50 · 60G51 · 60K35

## 1 Introduction

In 1955, Herbert A. Simon [24] introduced a simple reinforcement algorithm that runs as follows. Consider a deterministic sequence  $(\varepsilon_n)$  in  $\{0, 1\}$  with  $\varepsilon_1 = 1$ . The  $n$ -th step of the algorithm corresponds to an innovation if  $\varepsilon_n = 1$ , and to a repetition if  $\varepsilon_n = 0$ . Specifically, denote the number of innovations after  $n$  steps by

$$\sigma(n) = \sum_{i=1}^n \varepsilon_i \quad \text{for } n \geq 1,$$

and let also  $X_1, X_2, \dots$  denote a sequence of different items (in [24], these items are words). One constructs recursively a random sequence of items  $\hat{X}_1, \hat{X}_2, \dots$  by

---

✉ Jean Bertoin  
jean.bertoin@math.uzh.ch

<sup>1</sup> Institute of Mathematics, University of Zurich, Zurich, Switzerland

deciding that  $\hat{X}_n = X_{\sigma(n)}$  if  $\varepsilon_n = 1$ , and that  $\hat{X}_n = \hat{X}_{U(n)}$  if  $\varepsilon_n = 0$ , where  $U(n)$  is random with the uniform distribution on  $[n - 1] = \{1, \dots, n - 1\}$  and  $U(2), U(3), \dots$  are independent.

Simon was especially interested in regimes where either  $\varepsilon_n$  converges in Césaro’s mean to some limit  $q \in (0, 1)$ , which we will refer to as steady innovation with rate  $q$ , or  $\sigma(n)$  grows like<sup>1</sup>  $n^\rho$  for some  $\rho \in (0, 1)$  as  $n \rightarrow \infty$ , which we will call slow innovation with exponent  $\rho$ . By analyzing the frequencies of words with some fixed number of occurrences, he pointed out that these regimes yield a remarkable one-parameter family of power tail distributions, that are known nowadays as the Yule–Simon laws and arise in a variety of empirical data. This is also closely related to preferential attachment dynamics, see e.g. [10] for an application to the World Wide Web. Clearly, repetitions in Simon’s algorithm should be viewed as a linear reinforcement, the probability that a given item is repeated being proportional to the number of its previous occurrences.

In the present work, the items  $X_1, X_2, \dots$  are i.i.d. copies of some real random variable  $X$ , which we further assume to be independent of the uniform variables  $U(2), U(3), \dots$ . Note that although the variable  $\hat{X}_n$  has the same distribution as  $X$  for every  $n \geq 0$ , the reinforced sequence  $(\hat{X}_n)$  is not stationary; it can also be seen that its tail sigma-field is not even independent of  $\hat{X}_1$ . Picking up on a key question in the general area of reinforced processes (see notably the survey [21] by Pemantle, and also some more recent works [2, 14, 15, 18, 22] and references therein), our purpose is to analyze how reinforcement affects the growth of partial sums. Specifically, we write

$$S(n) = X_1 + \dots + X_n$$

for the usual random walk with step distribution  $X$ , and

$$\hat{S}(n) = \hat{X}_1 + \dots + \hat{X}_n$$

for its reinforced version, and we would like to compare  $\hat{S}(n)$  and  $S(n)$  when  $n \gg 1$ . The main situation of interest is when  $S$  has a scaling exponent  $\alpha \in (0, 2]$ , in the sense that

$$\lim_{n \rightarrow \infty} n^{-1/\alpha} S(n) = Y \quad \text{in law,} \tag{1}$$

where  $Y$  denotes an  $\alpha$ -stable variable. Recall that this holds if and only if the typical step  $X$  belongs to the domain of normal attraction (without centering) of a stable distribution, in the terminology of Gnedenko and Kolmogorov [16]. We shall refer to (1) as an instance of  $\alpha$ -diffusive asymptotic behavior, the usual diffusive situation corresponding to  $\alpha = 2$ .

The asymptotic behavior of the step-reinforced random walk  $\hat{S}$  has been considered previously in the literature when  $\varepsilon_2, \varepsilon_3, \dots$  are random and given by i.i.d. samples of the Bernoulli law with parameter  $q \in (0, 1)$ . This is of course a most important case

<sup>1</sup> The precise definition of these regimes will be given later on.

of a steady regime with innovation rate  $q$  a.s. It has been shown recently in [8] that when  $X \in L^2(\mathbb{P})$ , the asymptotic growth of  $\hat{S}$  exhibits a phase transition at  $q_c = 1/2$ . Specifically, assuming for simplicity that  $X$  is centered, then on the one hand for  $q < 1/2$ , there is some non-degenerate random variable  $V$  such that

$$\lim_{n \rightarrow \infty} n^{-1+q} \hat{S}(n) = V \quad \text{a.s.} \quad (2)$$

In other words, (2) shows that for  $q < 1/2$ ,  $\hat{S}$  has scaling exponent  $\hat{\alpha} = 1/(1 - q)$ , or equivalently, grows with exponent  $1/\hat{\alpha}$ , and in particular is super-diffusive since  $1/\hat{\alpha} > 1/2$ . On the other hand, the step-reinforced random walk remains diffusive for  $q > 1/2$ , in the sense that  $n^{-1/2} \hat{S}(n)$  converges in law to some Gaussian variable. This phase transition was first established when  $X$  is a Rademacher variable, i.e.  $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ . Indeed, Kürsten [20] observed that  $\hat{S}$  is then a version of the so-called *elephant random walk*, a nearest neighbor process with memory which was introduced by Schütz and Trimper [23] and has then raised much interest. The description of the asymptotic behavior of the elephant random walk has motivated many works, see notably [3,5,6,12,13,19].

Further, when the typical step  $X$  has a symmetric stable distribution with index  $\alpha \in (0, 2]$ ,  $\hat{S}$  is the so-called *shark random swim*, which has been studied in depth by Businger [11]. Its large time asymptotic behavior exhibits a similar phase transition for  $\alpha > 1$ , now for the critical parameter  $q_c = 1 - 1/\alpha$ . When  $\alpha \leq 1$ , there is no such phase transition and  $\hat{S}$  has the same scaling exponent  $\alpha$  as  $S$ . See also [7] for related results in the setting of Lévy processes.

The results that we just recalled suggest that, more generally, for any steady innovation regime and any typical step  $X$  belonging to the domain of normal attraction of an  $\alpha$ -stable distribution (i.e. such that (1) is fulfilled), then the following should hold. First, for  $\alpha \in (0, 1)$ , the random walk  $S$  and its step-reinforced version  $\hat{S}$  should have the same scaling exponent  $\hat{\alpha} = \alpha$ , independently of the innovation rate. Second, for  $\alpha \in (1, 2]$ , if the innovation rate  $q$  is larger than  $q_c = 1 - 1/\alpha$ , then again the scaling exponent of  $S$  and  $\hat{S}$  should coincide, whereas if  $q < 1 - 1/\alpha$ , then the super- $\alpha$ -diffusive behavior (2) should hold and  $\hat{S}$  should thus have scaling exponent  $\hat{\alpha} = 1/(1 - q) < \alpha$ . We shall see in Theorems 2 and 4 that this guess is indeed correct. In particular, the weaker the innovation (or equivalently the stronger the reinforcement), the faster the step-reinforced random walk  $\hat{S}$  grows.

An informal explanation for this phase transition is as follows. When  $\alpha \in (1, 2]$ , the  $\alpha$ -diffusive behavior (1) of  $S$  relies on some kind of balance between its positive and negative steps (recall that  $X$  must be centered, i.e.  $\mathbb{E}(X) = 0$ ). The reinforcement effect of Simon's algorithm for sufficiently small innovation rates  $q$  yields certain steps to be repeated much more often than the others, up to the point that this balance is disrupted. More precisely, we shall see that in a steady regime with innovation rate  $q$ , the maximal number of repetitions of a same item up to the  $n$ -th step of Simon's algorithm grows with exponent  $1 - q$ . For  $q > q_c = 1 - 1/\alpha$ , this is smaller than the growth exponent  $1/\alpha$  of  $S$  and repetitions have only a rather limited impact on the asymptotic behavior of  $\hat{S}$ . At the opposite, for  $q < q_c$ , some increments have been

repeated much more often and the growth of  $\hat{S}$  is then rather governed by the latter, yielding (3).

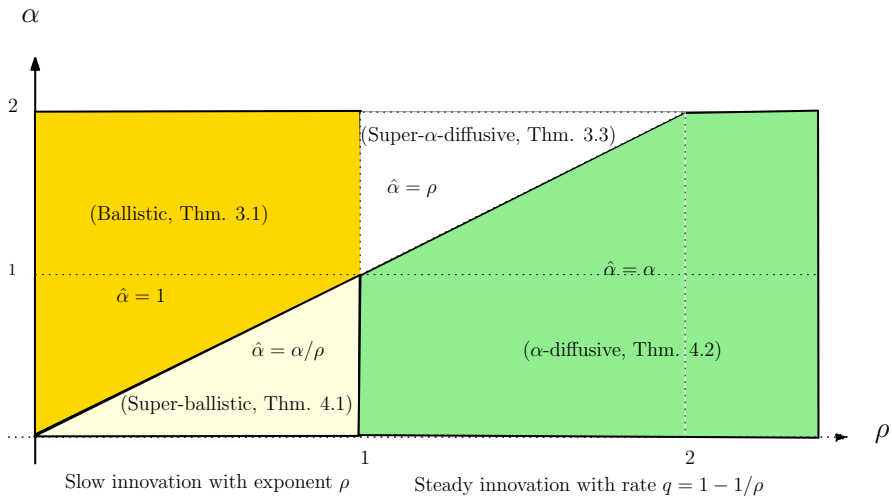
We now turn our attention to regimes with slow innovation. Extrapolating from the steady regime, we might expect that reducing the innovation should again speed up the step-reinforced random walk. This intuition turns out to be wrong, and we will see that at the opposite, in slow regimes, diminishing the innovation actually slows down the walk. More precisely, there is another phase transition when  $\alpha \in (0, 1)$ , occurring now for the critical innovation exponent  $\rho_c = \alpha$ . Specifically, if  $\rho < \alpha$ , then we shall see in Theorem 1 that  $\hat{S}$  has always scaling exponent  $\hat{\alpha} = 1$  (i.e. a ballistic asymptotic behavior which contrasts with the growth with exponent  $1/\alpha > 1$  for  $S$ ), whereas for  $\rho > \alpha$ , we will see in Theorem 3 that  $\hat{S}$  has rather scaling exponent  $\hat{\alpha} = \alpha/\rho > \alpha$ , that it grows now with exponent  $1/\hat{\alpha} = \rho/\alpha > 1$ , but nonetheless still significantly slower than  $S$ . On the other hand, when  $\alpha \geq 1$ , there is no phase transition for slow innovation regimes and  $\hat{S}$  has always scaling exponent  $\hat{\alpha} = 1$ .

This apparently surprising feature can be explained informally as follows. As it was argued above, for  $\alpha \in (1, 2]$ ,  $\mathbb{E}(X) = 0$  and the super-diffusive regime (2) results from the disruption of the balance between positive and negative steps when certain steps are repeated much more than others. At the opposite, for  $\alpha \in (0, 1)$ , the typical step  $X$  has a heavy tail distribution with  $\mathbb{E}(|X|) = \infty$ . In this situation, it is well-known that for  $n \gg 1$ ,  $|S(n)|$  has roughly the same size as its largest step up to time  $n$ ,  $\max\{|X_i| : 1 \leq i \leq n\}$ . Regimes with slow innovation delay the occurrence of rare events at which steps are exceptionally large. Therefore they induce a slow down effect for the step-reinforced random walk, up to the point that when the innovation exponent drops below a critical value,  $\hat{S}$  has merely a ballistic growth. This aspect will be further discussed quantitatively in Sect. 5.

A somewhat simpler version of the main results of our work are summarized in Fig. 1. It expresses the scaling exponent  $\hat{\alpha}$  of  $\hat{S}$  in terms of the scaling exponent  $\alpha \in (0, 2]$  of  $S$  and the innovation parameter  $\rho > 0$ . The slow regime corresponds to  $\rho \in (0, 1)$  and  $\rho$  is then the innovation exponent as usual. The steady regime corresponds to  $\rho > 1$ , and then the rate of innovation is given by  $q = 1 - 1/\rho$ . This new parametrization for steady regimes of innovation may seem artificial; nonetheless we stress that the same is actually used for the definition of the one-parameter family of Yule–Simon distributions; see Lemma 3.

The cornerstone of our approach is provided by Lemma 2, where we observe that the process that counts the number of occurrences of a given item in Simon's algorithm can be turned into a square integrable martingale. The latter is a close relative to another martingale that occurs naturally in the setting of the elephant random walk; see [6, 12, 13, 19], among others. The upshot of Lemma 2 is that this yield useful estimates for these numbers of occurrences and their asymptotic behaviors, which hold uniformly for all items.

The plan for the rest of this article is as follows. Section 2 is devoted to preliminaries on the stable central limit theorem, on martingales induced by occurrence counting processes in Simon's algorithm, and on the Yule–Simon distributions. We state and prove our main results in Sects. 3 and 4. Finally, several comments are given in Sect. 5.



**Fig. 1** Scaling exponent  $\hat{\alpha}$  of a step-reinforced random walk in terms of the innovation parameter  $\rho$  and the scaling exponent  $\alpha$  of the original random walk. Results above the diagonal  $\rho = \alpha$  are strong (convergence in probability), those below the diagonal are weak (convergence in distribution)

## 2 Preliminaries

Given two sequences  $a(n)$  and  $b(n)$  of positive real numbers, it will be convenient to use the following notation throughout this work:

$$\begin{aligned}
 a(n) \sim b(n) &\iff \lim_{n \rightarrow \infty} a(n)/b(n) = 1, \\
 a(n) \approx b(n) &\iff \lim_{n \rightarrow \infty} a(n)/b(n) \text{ exists in } (0, \infty), \\
 a(n) \asymp b(n) &\iff 0 < \inf_{n \geq 1} a(n)/b(n) \leq \sup_{n \geq 1} a(n)/b(n) < \infty.
 \end{aligned}$$

### 2.1 Background on the stable central limit theorem

We assume in this section that the step distribution belongs to the domain of normal attraction (without centering) of some stable distribution, i.e. that (1) holds for some  $\alpha \in (0, 2]$ . The Cauchy case  $\alpha = 1$  has some peculiarities and for the sake of simplicity, it will be ruled out from time to time. We present some classical results in this framework that will be useful later on.

We start by recalling that for  $\alpha = 2$ , (1) holds if and only if  $X$  is centered with finite variance; see Theorem 4 on p. 181 in [16]. For  $\alpha \in (0, 1)$ , (1) is equivalent to

$$\lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X > x) = c_+ \quad \text{and} \quad \lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X < -x) = c_-$$

for some nonnegative constants  $c_+$  and  $c_-$  with  $c_+ + c_- > 0$ . Finally, for  $\alpha \in (1, 2)$ , (1) holds if and only if the same as above is fulfilled and furthermore  $X$  is centered. See Theorem 5 on p. 181-2 in [16].

We denote the characteristic function of  $X$  by

$$\Phi(\theta) = \mathbb{E}(\exp(i\theta X)) \quad \text{for } \theta \in \mathbb{R},$$

and the characteristic exponent of the stable variable  $Y$  by  $\varphi_\alpha$ , that is  $\varphi_\alpha : \mathbb{R} \rightarrow \mathbb{C}$  is the unique continuous function with  $\varphi_\alpha(0) = 0$  such that

$$\mathbb{E}(\exp(i\theta Y)) = \exp(-\varphi_\alpha(\theta)) \quad \text{for } \theta \in \mathbb{R}.$$

In particular,  $\varphi_\alpha$  is homogeneous with degree  $\alpha$  in the sense that

$$\varphi_\alpha(c\theta) = c^\alpha \varphi_\alpha(\theta) \quad \text{for all } c > 0 \text{ and } \theta \in \mathbb{R}.$$

In this setting, (1) can be expressed classically as

$$\lim_{n \rightarrow \infty} \Phi(\theta n^{-1/\alpha})^n = \exp(-\varphi_\alpha(\theta)), \quad \text{for all } \theta \in \mathbb{R}, \tag{3}$$

but we shall rather use a logarithmic version of (3).

Pick  $r > 0$  sufficiently small so that  $|1 - \Phi(\theta)| < 1$  whenever  $|\theta| \leq r$ , and then define  $\varphi : [-r, r] \rightarrow \mathbb{C}$  as the continuous determination of the logarithm of  $\Phi$  on  $[-r, r]$ , i.e. the unique continuous function with  $\varphi(0) = 0$  and such that  $\Phi(\theta) = \exp(-\varphi(\theta))$  for all  $\theta \in [-r, r]$ . Theorem 2.6.5 in Ibragimov and Linnik [17] entails that (3) can be rewritten in the form

$$\lim_{t \rightarrow \infty} t\varphi(\theta t^{-1/\alpha}) = \varphi_\alpha(\theta), \quad \text{for all } \theta \in \mathbb{R}. \tag{4}$$

We stress that the parameter  $t$  in (4) is real, whereas  $n$  in (3) is an integer, and as a consequence, we have also that

$$\varphi(\theta) = O(|\theta|^\alpha) \quad \text{as } \theta \rightarrow 0.$$

### 2.2 Martingales in Simon’s algorithm

Recall Simon’s algorithm from the Introduction, and in particular that  $\sigma(n)$  stands for the number of innovations up to the  $n$ -th step. In this work, we will be mostly concerned with the cases where either the sequence  $\sigma(\cdot)$  is regularly varying with exponent  $\rho \in (0, 1)$ , that is

$$\lim_{n \rightarrow \infty} \frac{\sigma(\lfloor cn \rfloor)}{\sigma(n)} = c^\rho \quad \text{for all } c > 0, \tag{5}$$

or

$$\sum_{n=1}^{\infty} n^{-2} |\sigma(n) - qn| < \infty \quad \text{for some } q \in (0, 1). \tag{6}$$

It is easily checked that (6) implies  $\sigma(n) \sim qn$ , and conversely, (6) holds whenever  $\sigma(n)/n = q + O(\log^{-\beta} n)$  for some  $\beta > 1$ . We refer to (5) as the *slow regime with innovation exponent*  $\rho \in (0, 1)$ , and to (6) as the *steady regime with innovation rate*  $q \in (0, 1)$ . Often, it is convenient to set  $\rho = 1/(1 - q)$  for  $q \in (0, 1)$  and then view  $\rho \in (0, 1) \cup (1, \infty)$  as a parameter for the innovation, with  $\rho > 1$  corresponding to steady regimes.

Several of our results however rely on much weaker assumptions; in any case we shall always assume at least that the total number of innovations is infinite and that the number of repetitions is not sub-linear, i.e.

$$\sigma(\infty) = \infty \quad \text{and} \quad \limsup_{n \rightarrow \infty} n^{-1} \sigma(n) < 1. \tag{7}$$

Simon’s algorithms induces a natural partition of the set of indices  $\mathbb{N} = \{1, 2, \dots\}$  into a sequence of blocks  $B_1, B_2, \dots$ , where

$$B_j = \{k \in \mathbb{N} : \hat{X}_k = X_j\}.$$

In words,  $B_j$  is the set of steps of Simon’s algorithm at which the  $j$ -th item  $X_j$  is repeated. We consider for every  $n \in \mathbb{N}$  the restriction of the preceding partition to  $[n] = \{1, \dots, n\}$  and write

$$B_j(n) = B_j \cap [n] = \{k \in [n] : \hat{X}_k = X_j\};$$

plainly  $B_j(n)$  is nonempty if and only if  $j \leq \sigma(n)$ . Last, we set

$$|B_j(n)| = \text{Card } B_j(n)$$

for the number of elements of  $B_j(n)$ , and arrive at the following basic expression for the step-reinforced random walk:

$$\hat{S}(n) = \sum_{j=1}^{\infty} |B_j(n)| X_j = \sum_{j=1}^{\sigma(n)} |B_j(n)| X_j, \tag{8}$$

where the  $X_j$  are i.i.d. copies of  $X$ , and further independent of the random coefficients  $|B_j(n)|$ .



The identity (8) incites us to investigate the asymptotic behavior of the coefficients  $|B_j(n)|$ . In this direction, we introduce the quantities

$$\pi(n) = \prod_{j=2}^n \left(1 + \frac{1 - \varepsilon_j}{j - 1}\right), \quad n \in \mathbb{N}, \tag{9}$$

and the times of innovation

$$\tau(j) = \inf\{n \in \mathbb{N} : \sigma(n) = j\} = \min B_j = \inf\{n \in \mathbb{N} : |B_j(n)| = 1\}.$$

We stress that these quantities are deterministic, since the sequence  $(\varepsilon_n)$  is deterministic.

We start with a simple lemma:

**Lemma 1** *The following assertions hold:*

- (i) Assume that  $\sigma(n) = O(n^\rho)$  for some  $\rho < 1$ . Then  $\pi(n) \approx n$ .
- (ii) Assume (6); then  $\pi(n) \approx n^{1-q}$ .
- (iii) Assume (7); then the series  $\sum_{n=1}^\infty 1/(n\pi(n))$  converges.

**Proof** We have from the definition of  $\pi(n)$  that

$$\pi(n) = \exp \left( \sum_{j=2}^n \log \left( 1 + \frac{1 - \varepsilon_j}{j - 1} \right) \right) \approx \exp \left( \sum_{j=2}^n \frac{1 - \varepsilon_j}{j - 1} \right).$$

Next we observe by summation by parts that

$$\sum_{j=2}^n \frac{1 - \varepsilon_j}{j - 1} = \frac{n - \sigma(n)}{n - 1} + \sum_{j=2}^{n-1} \frac{j - \sigma(j)}{j(j - 1)}.$$

Assume first  $\sigma(n) = O(n^\rho)$  for some  $\rho < 1$ . Then  $\sum_{j=2}^\infty \sigma(j)j^{-2} < \infty$ , which yields

$$\lim_{n \rightarrow \infty} \left( \sum_{j=2}^n \frac{1 - \varepsilon_j}{j - 1} - \log n \right) \text{ exists in } \mathbb{R},$$

and (i) follows.

Next, when (6) holds, we write

$$\sum_{j=2}^{n-1} \frac{j - \sigma(j)}{j(j - 1)} = (1 - q) \sum_{j=2}^{n-1} \frac{1}{j - 1} - \sum_{j=2}^{n-1} \frac{\sigma(j) - qj}{j(j - 1)}.$$

The second series in the right-hand side converges absolutely; as a consequence,

$$\lim_{n \rightarrow \infty} \left( \sum_{j=2}^n \frac{1 - \varepsilon_j}{j - 1} - (1 - q) \log n \right) \text{ exists in } \mathbb{R},$$

and (ii) follows.

Finally, assume (7). There is  $a < 1$  such that  $\sigma(k) \leq ak$  for all  $k$  sufficiently large. It follows that there is some  $b > 0$  such that for all  $n$ ,

$$\sum_{j=2}^n \frac{1 - \varepsilon_j}{j - 1} \geq (1 - a) \log n - b.$$

We conclude that  $1/(n\pi(n)) = O(n^{a-2})$ , which entails the last claim. □

The next result determines the asymptotic behavior of the sequences  $|B_j(\cdot)|$  for all  $j \in \mathbb{N}$ , and will play therefore a key role in our analysis.

**Lemma 2** *Assume (7). For every  $j \in \mathbb{N}$ , the process started at time  $\tau(j)$ ,*

$$\pi(n)^{-1} |B_j(n)|, \quad n \geq \tau(j),$$

*is a square integrable martingale. We denote its terminal value by*

$$\Gamma_j = \lim_{n \rightarrow \infty} \pi(n)^{-1} |B_j(n)|,$$

*and have*

$$\mathbb{E}(\Gamma_j) = \frac{1}{\pi(\tau(j))} \text{ and } \text{Var}(\Gamma_j) \leq \frac{2}{\pi(\tau(j))} \sum_{n=\tau(j)}^{\infty} \frac{1}{n\pi(n)}.$$

**Proof** The martingale property is immediate from Simon’s algorithm. More precisely, for any  $n \geq \tau(j)$ , we have  $\pi(n+1) = \pi(n)$  and  $|B_j(n+1)| = |B_j(n)|$  when  $\varepsilon_{n+1} = 1$  (by innovation), whereas when  $\varepsilon_{n+1} = 0$ , we have  $\pi(n+1) = \pi(n)(1 + 1/n)$  and further (by reinforcement)

$$\mathbb{P}(|B_j(n+1)| = |B_j(n)| + 1 \mid \mathcal{F}_n) = |B_j(n)|/n$$

and

$$\mathbb{P}(|B_j(n+1)| = |B_j(n)| \mid \mathcal{F}_n) = 1 - |B_j(n)|/n$$

where  $(\mathcal{F}_n)_{n \geq 1}$  denotes the natural filtration of Simon’s algorithm. The claimed martingale property follows, and as a consequence, there is the identity

$$\mathbb{E}(|B_j(n)|) = \pi(n)/\pi(\tau(j)) \quad \text{for all } n \geq \tau(j). \tag{10}$$

We next have to check that the mean of the quadratic variation of the martingale  $|B_j(\cdot)|/\pi(\cdot)$  satisfies

$$\sum_{n=\tau(j)}^{\infty} \mathbb{E} \left( \left| \frac{|B_j(n+1)|}{\pi(n+1)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \right) \leq \frac{2}{\pi(\tau(j))} \sum_{n=\tau(j)}^{\infty} \frac{1}{n\pi(n)};$$

thanks to Lemma 1, the remaining assertions are then immediate.

In this direction, we first note that the terms in the sum on the left-hand side above that correspond to an innovation (i.e.  $\varepsilon_{n+1} = 1$ ) are zero and can thus be discarded. Let  $\varepsilon_{n+1} = 0$ , so that  $\pi(n+1) = \pi(n)(1 + 1/n)$ . We then have

$$\begin{aligned} & \mathbb{E} \left( \left| \frac{|B_j(n+1)|}{\pi(n+1)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \right) \\ & \leq \mathbb{E} \left( \left| \frac{|B_j(n)| + 1}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \frac{|B_j(n)|}{n} \right) + \mathbb{E} \left( \left| \frac{|B_j(n)|}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \right). \end{aligned}$$

On the one hand, since

$$\frac{|B_j(n)| + 1}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} = \frac{1 - |B_j(n)|/n}{\pi(n)(1 + 1/n)} \in \left[ 0, \frac{1}{\pi(n)} \right],$$

we deduce from (10) the bound

$$\mathbb{E} \left( \left| \frac{|B_j(n)| + 1}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \frac{|B_j(n)|}{n} \right) \leq \frac{1}{n\pi(n)\pi(\tau(j))}.$$

On the other hand, since

$$\left| \frac{|B_j(n)|}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \leq \frac{|B_j(n)|^2}{\pi(n)^2 n^2} \leq \frac{|B_j(n)|}{\pi(n)^2 n},$$

using again (10), we get

$$\mathbb{E} \left( \left| \frac{|B_j(n)|}{\pi(n)(1 + 1/n)} - \frac{|B_j(n)|}{\pi(n)} \right|^2 \right) \leq \frac{1}{n\pi(n)\pi(\tau(j))}.$$

The proof of the statement is now complete. □

As an immediate consequence, we point at the following handier estimate for the second moment of  $\Gamma_j$ .

**Corollary 1** *Assume (7) and further that  $\pi(n) \asymp n^a$  for some  $a > 0$ . Then*

$$\mathbb{E}(\Gamma_j^2) \asymp \frac{1}{\tau(j)^{2a}}.$$

**Proof** On the one hand, there is the lower bound  $\mathbb{E}(\Gamma_j^2) \geq \mathbb{E}(\Gamma_j)^2$ . On the other hand, our assumption also entails that for some  $b, b' > 0$ , we have

$$\sum_{n \geq \ell} 1/(n\pi(n)) \leq b \sum_{n \geq \ell} n^{-1-a} \leq b'\ell^{-a},$$

and we conclude with Lemma 2. □

### 2.3 Yule–Simon distributions

Recall that the slow and the steady regimes have been defined by (5) and (6), respectively. Simon [24] observed that in each regime, the empirical measure of the sizes of the blocks  $|B_j(n)|$  converges to a deterministic distribution.

**Lemma 3** (Simon [24]) *Let  $\rho > 0$ . For  $0 < \rho < 1$ , consider the regime (5) of slow innovation with exponent  $\rho$ , whereas for  $\rho > 1$ , set  $q = 1 - 1/\rho \in (0, 1)$  and consider the regime (6) of steady innovation with rate  $q$ . In both regimes, for every  $k \in \mathbb{N}$ , we have*

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma(n)} \text{Card}\{j \leq \sigma(n) : |B_j(n)| = k\} = \rho B(k, \rho + 1),$$

where  $B$  is the Beta function and the convergence holds in  $L^p$  for any  $p \geq 1$ .

The limiting distribution in the statement is called the Yule–Simon distribution with parameter  $\rho$ . Strictly speaking, Simon only established the stated convergence in expectation. A classical argument of propagation of chaos yields the stronger convergence in probability; see e.g. Section 5 in [4], and since the random variables in the statement are obviously bounded by 1, convergence in  $L^p$  also holds for any  $p \geq 1$ .

The next lemma will be needed to check some uniform integrability properties.

**Lemma 4** *Let  $0 < \beta \leq \rho$  and assume either (i) or (ii) is fulfilled, where:*

- (i)  $\rho \in (0, 1)$  and the slow regime (5) holds with exponent  $\rho$ ,
- (ii)  $\rho > 1$  and the steady regime (6) holds with innovation rate  $q = 1 - 1/\rho$ .

Then

$$\sup_{n \geq 1} \frac{1}{\sigma(n)} \sum_{j=1}^{\sigma(n)} \mathbb{E}(|B_j(n)|^\beta) < \infty.$$

**Remark 1** Since  $B(k, \rho + 1) \sim \Gamma(\rho + 1)k^{-(\rho+1)}$  as  $k \rightarrow \infty$ , we have that

$$\sum_{k=1}^{\infty} k^\beta \rho B(k, \rho + 1) < \infty$$

for any  $\beta < \rho$ , in agreement with Fatou’s lemma and Lemmas 3 and 4.

**Proof** (i) Recall from Lemma 1 that in the slow regime, there are the bounds  $n/c \leq \pi(n) \leq cn$  for all  $n \in \mathbb{N}$ , where  $c > 1$  is some constant. Since, from Lemma 2,

$$\mathbb{E}(|B_j(n)|) = \pi(n)/\pi(\tau(j)) \leq c^2 n/\tau(j),$$

we get by Jensen’s inequality that

$$\sum_{j=1}^{\sigma(n)} \mathbb{E}(|B_j(n)|^\beta) \leq c^{2\beta} n^\beta \sum_{j=1}^{\sigma(n)} \tau(j)^{-\beta} = O(n^\beta \sigma(n)(\tau(\sigma(n)))^{-\beta}),$$

where for the  $O$  upperbound, we used the fact that the inverse function  $\tau$  of  $\sigma$  is regularly varying with exponent  $1/\rho$  (Theorem 1.5.12 in [9]), and Proposition 1.5.8 in [9] since  $-\beta/\rho > -1$ . On the other hand, since  $\tau$  is the right-inverse of  $\sigma$ , we have  $\tau(\sigma(n)) \leq n \leq \tau(\sigma(n) + 1)$ , so again by regular variation,  $\tau(\sigma(n)) \sim n$ . Finally

$$\sum_{j=1}^{\sigma(n)} \mathbb{E}(|B_j(n)|^\beta) = O(\sigma(n)),$$

as we wanted to verify.

(ii) The proof is similar to (i), using now that there exists  $c > 0$  such that

$$\mathbb{E}(|B_j(n)|^2) \leq c(n/j)^{2-2q} \quad \text{for all } j \in \mathbb{N} \text{ and } n \geq \tau(j),$$

as it is readily seen from Corollary 1. □

### 3 Strong limit theorems

In this section, we will establish two strong limit theorems for step-reinforced random walks, the first concerns slow innovation regimes, and the second steady ones.

#### 3.1 Ballistic behavior

**Theorem 1** *Suppose that*

$$\sigma(n) = O(n^\rho) \quad \text{as } n \rightarrow \infty,$$

*for some  $\rho \in (0, 1)$ , and that*

$$\mathbb{P}(|X| > x) = O(x^{-\beta}) \quad \text{as } x \rightarrow \infty,$$

*for some  $\beta > \rho$ . Then*

$$\lim_{n \rightarrow \infty} n^{-1} \hat{S}(n) = V' \quad \text{a.s.}$$

where  $V'$  is some non-degenerate random variable.

We will deduce Theorem 1 by specializing the following more general result.

**Lemma 5** Assume (7) and set

$$\Gamma_j^* = \sup_{n \geq \tau(j)} |B_j(n)|/\pi(n), \quad j \in \mathbb{N}.$$

Provided that

$$\sum_{j=1}^{\infty} \Gamma_j^* |X_j| < \infty \quad a.s., \tag{11}$$

we have

$$\lim_{n \rightarrow \infty} \hat{S}(n)/\pi(n) = V \quad a.s.,$$

with

$$V = \sum_{j=1}^{\infty} \Gamma_j X_j.$$

**Proof** Thanks to (11), the claim follows from (8) and Lemma 2 by dominated convergence. □

**Proof of Theorem 1** Recall from Lemma 1(i) that  $\pi(n) \approx n$ . From Lemma 5, it thus suffices to check that

$$\sum_{j=1}^{\infty} \mathbb{E} \left( (\Gamma_j^* |X_j|) \wedge 1 \right) < \infty, \tag{12}$$

since then, the condition (11) follows.

Without loss of generality, we may assume that  $\beta < 1$ . Pick  $a > 0$  sufficiently large so that

$$\sigma(n) \leq an^\rho \quad \text{for all } n \geq 1$$

and

$$\mathbb{P}(|X| > x) \leq ax^{-\beta} \quad \text{for all } x > 0.$$

Since  $X_j$  is a copy of  $X$  which is independent of  $\Gamma_j^*$ , we have

$$\mathbb{E} \left( (\Gamma_j^* |X_j|) \wedge 1 \right) = \int_0^1 \mathbb{P}(\Gamma_j^* |X_j| > x) dx \leq a \mathbb{E}((\Gamma_j^*)^\beta) \int_0^1 x^{-\beta} dx = \frac{a \mathbb{E}((\Gamma_j^*)^\beta)}{1 - \beta}.$$

Recall from Lemma 2 that  $|B_j(\cdot)|/\pi(\cdot)$  is a closed martingale with terminal value  $\Gamma_j$ . Then by Doob’s maximal inequality, there is some numerical constant  $c_\beta > 0$  such that  $\mathbb{E}((\Gamma_j^*)^\beta) \leq c_\beta \mathbb{E}(\Gamma_j)^\beta$ , and hence again from Lemma 2,

$$\mathbb{E}\left((\Gamma_j^*|X_j|) \wedge 1\right) = O(\tau(j)^{-\beta}).$$

Finally, since  $\tau(j) \geq (j/a)^{1/\rho}$ , we conclude that

$$\mathbb{E}\left((\Gamma_j^*|X_j|) \wedge 1\right) = O(j^{-\beta/\rho}) \quad \text{as } j \rightarrow \infty,$$

which ensures (12) since  $\beta > \rho$ . □

### 3.2 Super- $\alpha$ -diffusive behavior

We next turn our attention to the steady regime.

**Theorem 2** *Suppose (6) holds with  $q < 1/2$  and that*

$$\mathbb{E}(|X|^\beta) < \infty \quad \text{and} \quad \mathbb{E}(X) = 0,$$

for some  $\beta > 1/(1 - q)$ . Then

$$\lim_{n \rightarrow \infty} n^{q-1} \hat{S}(n) = V' \quad \text{in } L^\beta(\mathbb{P}) \text{ and a.s.}$$

where  $V'$  is some non-degenerate random variable.

The proof of Theorem 2 relies on the following martingale convergence result.

**Lemma 6** *Assume (7) and let  $\beta \in (1, 2]$ . Suppose that  $X \in L^\beta(\mathbb{P})$  with  $\mathbb{E}(X) = 0$ , and further that*

$$\sum_{j=1}^\infty \mathbb{E}(\Gamma_j^\beta) < \infty. \tag{13}$$

The process

$$V_n = \sum_{j=1}^n \Gamma_j X_j, \quad n \in \mathbb{N}$$

is then a martingale bounded in  $L^\beta(\mathbb{P})$ ; we write  $V_\infty$  for its terminal value. We have

$$\lim_{n \rightarrow \infty} \hat{S}(n)/\pi(n) = V_\infty \quad \text{in } L^\beta(\mathbb{P}) \text{ and a.s.}$$

**Proof** The assertion that the process  $V_n$  is a martingale is straightforward since the variables  $X_j$  are i.i.d., centered, and independent of the  $\Gamma_j$ . The assertion of boundedness in  $L^\beta(\mathbb{P})$  then follows from the assumption (13), the Burkholder–Davis–Gundy inequality, and the fact that, for any sequence  $(y_j)_{j \in \mathbb{N}}$  of nonnegative real numbers, since  $\beta \leq 2$ ,

$$\left( \sum_{j=1}^{\infty} y_j^2 \right)^{\beta/2} \leq \sum_{j=1}^{\infty} y_j^\beta.$$

The convergence of  $\hat{S}(n)/\pi(n)$  in  $L^\beta(\mathbb{P})$  is proven similarly. Specifically, we observe from (8) that

$$V_{\sigma(n)} - \hat{S}(n)/\pi(n) = \sum_{j=1}^{\sigma(n)} (\Gamma_j - |B_j(n)|/\pi(n)) X_j,$$

and recall that the variables  $X_j$  are independent of those appearing in Simon’s algorithm. By the Burkholder–Davis–Gundy inequality, there exists a constant  $c_\beta \in (0, \infty)$  such that

$$\mathbb{E} \left( \left| V_{\sigma(n)} - \hat{S}(n)/\pi(n) \right|^\beta \right) \leq c_\beta \mathbb{E}(|X|^\beta) \sum_{j=1}^{\sigma(n)} \mathbb{E}(|\Gamma_j - |B_j(n)|/\pi(n)|^\beta).$$

We know from Lemma 2 that for each  $j \geq 1$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E}(|\Gamma_j - |B_j(n)|/\pi(n)|^\beta) = 0,$$

and further by Jensen’s inequality, that

$$\mathbb{E}(|\Gamma_j - |B_j(n)|/\pi(n)|^\beta) \leq 2^\beta \mathbb{E}(\Gamma_j^\beta).$$

The assumption (13) enables us to complete the proof of convergence of the sequence  $(\hat{S}(n)/\pi(n))$  in  $L^\beta(\mathbb{P})$  by dominated convergence.

The almost sure convergence then follows from the observation that the process  $\hat{S}(n)/\pi(n)$  is a martingale (in the setting of the elephant random walk, a similar property has been pointed at in [6,12,13,19]). Indeed, we see from Simon’s algorithm and the assumption  $\mathbb{E}(X) = 0$  that

$$\mathbb{E}(\hat{X}_{n+1} \mid \hat{X}_1, \dots, \hat{X}_n) = \begin{cases} 0 & \text{if } \varepsilon_{n+1} = 1, \\ \hat{S}(n)/n & \text{if } \varepsilon_{n+1} = 0. \end{cases}$$

This immediately entails our assertion. □



**Proof of Theorem 2** Recall that we assume that  $\mathbb{E}(|X|^\beta) < \infty$  for some  $\beta > 1/(1 - q)$ . Since  $q < 1/2$ , we can further suppose without loss of generality that  $\beta \leq 2$ . Then, by Jensen’s inequality, we have

$$\sum_{j=1}^\infty \mathbb{E}(\Gamma_j^\beta) \leq \sum_{j=1}^\infty \mathbb{E}(\Gamma_j^2)^{\beta/2},$$

and we just need to check that the right-hand side is finite, as then an appeal to Lemma 6 completes the proof.

It follows from (6) and Lemma 1(ii) that

$$\tau(n) \sim n/q \quad \text{and} \quad \pi(n) \asymp n^{1-q}, \tag{14}$$

and then from Corollary 1 that  $\mathbb{E}(\Gamma_j^2) \asymp j^{-2+2q}$ . Since  $\beta - q\beta > 1$ , the series  $\sum_{j \geq 1} j^{-\beta+q\beta}$  converges, and the proof is finished.  $\square$

### 4 Weak limit theorems

In this section, we will establish two weak limit theorems for step-reinforced random walks, depending on the innovation regimes.

#### 4.1 Super-ballistic behavior

**Theorem 3** *Suppose that  $X$  belongs to the domain of normal attraction of a stable law (i.e. (1) holds) with index  $\alpha \in (0, 1)$ , and that (5) holds for some  $\rho \in (\alpha, 1)$ . Then*

$$\lim_{n \rightarrow \infty} \sigma(n)^{-1/\alpha} \hat{S}(n) = Y' \quad \text{in law}$$

where  $Y'$  is an  $\alpha$ -stable random variable.

Under the assumptions of Theorem 3, the step-reinforced random walk grows roughly like  $n^{\rho/\alpha}$ , and since  $1 < \rho/\alpha < 1/\alpha$ , its asymptotic behavior is both super-ballistic and sub- $\alpha$ -diffusive.

**Proof** Note first that, since  $\rho > \alpha$ ,  $n\sigma(n)^{-1/\alpha}$  goes to 0 as  $n \rightarrow \infty$ , and *a fortiori* so does  $|B_j(n)|\sigma(n)^{-1/\alpha}$  uniformly for all  $j \in \mathbb{N}$ . We fix  $\theta \in \mathbb{R}$  and get from (8) that for  $n$  sufficiently large

$$\mathbb{E}(\exp(i\theta\sigma(n)^{-1/\alpha} \hat{S}(n))) = \mathbb{E} \left( \exp \left( - \sum_{j=1}^{\sigma(n)} \varphi(\theta\sigma(n)^{-1/\alpha} |B_j(n)|) \right) \right)$$

We focus on the sum in the right-hand side, and first consider the terms with  $|B_j(n)| \leq k$  for some fixed  $k \in \mathbb{N}$ . Write

$$\sum_{j:|B_j(n)| \leq k} \varphi(\theta \sigma(n)^{-1/\alpha} |B_j(n)|) = \frac{1}{\sigma(n)} \sum_{\ell=1}^k \varphi(\theta \sigma(n)^{-1/\alpha} \ell) \sigma(n) N_\ell(n),$$

where  $N_\ell(n) = \text{Card}\{j \leq \sigma(n) : |B_j(n)| = \ell\}$ . Next, recall from (4) that as  $n \rightarrow \infty$ ,

$$\varphi(\theta \sigma(n)^{-1/\alpha} \ell) \sigma(n) \sim \varphi_\alpha(\theta \ell) = \varphi_\alpha(\theta) \ell^\alpha.$$

We now deduce from Lemma 4 that for any fixed  $k \in \mathbb{N}$ , there is the convergence

$$\lim_{n \rightarrow \infty} \sum_{j:|B_j(n)| \leq k} \varphi(\theta \sigma(n)^{-1/\alpha} |B_j(n)|) = \varphi_\alpha(\theta) \sum_{\ell=1}^k \ell^\alpha \rho \mathbf{B}(\ell, \rho + 1) \quad \text{in } L^p(\mathbb{P})$$

for every  $p \geq 1$ .

We can next complete the proof by an argument of uniform integrability. Recall that  $\varphi(\lambda) = O(|\lambda|^\alpha)$  as  $\lambda \rightarrow 0$  and pick  $\beta \in (\alpha, \rho)$ . There exists  $a > 0$  such that for all  $n$  sufficiently large and all  $k \geq 1$ , there is the upper bound

$$\sum_{j:|B_j(n)| > k} \varphi(\theta \sigma(n)^{-1/\alpha} |B_j(n)|) \leq a \frac{k^{\alpha-\beta}}{\sigma(n)} \sum_{j=1}^\infty |B_j(n)|^\beta,$$

and the same inequality holds with  $\varphi_\alpha$  replacing  $\varphi$ . We can then deduce from the preceding paragraph in combination with Lemma 4 that actually

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{\sigma(n)} \varphi(\theta \sigma(n)^{-1/\alpha} |B_j(n)|) = \varphi_\alpha(\theta) \sum_{\ell=1}^\infty \ell^\alpha \rho \mathbf{B}(\ell, \rho + 1) \quad \text{in probability.}$$

It now suffices to recall that  $\Re \varphi \geq 0$ , so by dominated convergence,

$$\lim_{n \rightarrow \infty} \mathbb{E}(\exp(i\theta \sigma(n)^{-1/\alpha} \hat{S}(n))) = \exp\left(-\varphi_\alpha(\theta) \sum_{\ell=1}^\infty \ell^\alpha \rho \mathbf{B}(\ell, \rho + 1)\right),$$

which completes the proof. □

### 4.2 $\alpha$ -Diffusive behavior

**Theorem 4** *Suppose that  $X$  belongs to the domain of normal attraction without centering of a stable law (i.e. (1) holds) with index  $\alpha \in (0, 2]$ , and that (6) holds for some*

$q \in (0, 1)$ . Suppose further that  $q > 1 - 1/\alpha$  when  $\alpha > 1$ . Then

$$\lim_{n \rightarrow \infty} n^{-1/\alpha} \hat{S}(n) = Y' \quad \text{in law}$$

where  $Y'$  is an  $\alpha$ -stable random variable.

The proof of Theorem 4 requires the following uniform bounds

**Lemma 7** Suppose (6) holds for some  $q \in (0, 1)$  and take any  $\beta \in (0, 1/(1 - q))$ . Then

$$\lim_{n \rightarrow \infty} \sup_{j \geq 1} |B_j(n)| n^{-1/\beta} = 0 \quad \text{in probability.}$$

**Proof** The claim is obvious when  $\beta < 1$ , so we focus on the case  $\beta \geq 1$ . In this direction, recall from Lemma 2 that  $|B_j(n)|/\pi(n)$  is a square integrable martingale with terminal value  $\Gamma_j$ . Recall also from Lemma 1(ii) and Corollary 1, that in the regime (6),  $\pi(n) \approx n^{1-q}$  and  $\mathbb{E}(\Gamma_j^2) \asymp j^{2q-2}$ . There is thus some constant  $a > 0$ , such that for any  $\eta > 0$  arbitrarily small, we have

$$\mathbb{P}(|B_j(n)| > \eta n^{1/\beta}) \leq a \eta^{-2} n^{2-2q-2/\beta} j^{2q-2}. \tag{15}$$

Suppose first that  $q < 1/2$ , so  $\sum_{j \geq 1} j^{2q-2} < \infty$  and therefore

$$\sum_{j=1}^{\infty} \mathbb{P}(|B_j(n)| > \eta n^{1/\beta}) = O(n^{2-2q-2/\beta}).$$

Since  $1 - q < 1/\beta$ , our claim follows.

Then suppose that  $q = 1/2$ ; using  $\sum_{j \leq n} j^{-1} \sim \log n$  and  $|B_j(n)| = 0$  for  $j > n$ , we get

$$\sum_{j=1}^{\infty} \mathbb{P}(|B_j(n)| > \eta n^{1/\beta}) = O(n^{1-2/\beta} \log n).$$

Since  $1/\beta > 1/2$ , our assertion is verified.

Finally, suppose that  $q > 1/2$ ; using  $\sum_{j \leq n} j^{2q-2} \approx n^{2q-1}$  and  $|B_j(n)| = 0$  for  $j > n$ , we get

$$\sum_{j=1}^{\infty} \mathbb{P}(|B_j(n)| > \eta n^{1/\beta}) = O(n^{1-2/\beta}).$$

Since again  $1/\beta > 1/2$ , the proof is complete. □

Lemma 7 enables us to duplicate the argument for the proof of Theorem 3, as the reader will readily check.

## 5 Miscellaneous remarks

- Technically, the fact that the indices of the steps at which innovations occur are deterministic eases our approach by pointing right from the start at the relevant quantities. Although our statements are only given for deterministic sequences  $(\varepsilon_n)$ , they also apply to random sequences  $(\varepsilon_n)$  independent of  $(X_n)$ , provided of course that we can check that the requirements hold a.s. A basic example, which has been chiefly dealt with in the literature, is when the  $\varepsilon_j$  are i.i.d. samples of the Bernoulli distribution with parameter  $q \in (0, 1)$ , as then (6) obviously holds a.s. Plainly independence of the  $\varepsilon_j$  is not a necessary assumption, and much less restrictive correlation structures suffice. For instance, if we merely suppose that each  $\varepsilon_j$  has the Bernoulli law with parameter  $q_j$  such that  $\sum_{n \geq 2} n^{-2} |\sum_{j=2}^n (q_j - q)| < \infty$ , and that  $|\text{Cov}(\varepsilon_j, \varepsilon_\ell)| \leq |j - \ell|^{-a}$  for some  $a > 0$ , then one readily verifies that (6) is fulfilled a.s. Similar examples can be developed to get slow innovation regimes, for instance assuming that each variable  $\varepsilon_j$  has a Bernoulli law with  $q(j) \approx j^{\rho-1}$  and again a mild condition on the correlation.
- Dwelling on an informal comment made in the Introduction, it may be interesting to compare the step-reinforced random walk  $\hat{S}(n)$  with its maximal step  $\hat{X}_n^* = \max_{1 \leq j \leq n} |\hat{X}_j|$ . Assume  $\alpha \in (0, 2)$ , and that  $\mathbb{P}(|X| > x) \approx x^{-\alpha}$  (recall Sect. 2.1 about characterization of stable domains of normal attraction). Plainly, there is the identity  $\hat{X}_n^* = X_{\sigma(n)}^*$ , where  $X_n^* = \max_{1 \leq j \leq n} |X_j|$ , from which we deduce that  $\sigma(n)^{-1/\alpha} \hat{X}_n^*$  converges in distribution as  $n \rightarrow \infty$  to some Fréchet variable. Comparing with the results in Sects. 3 and 4, we now see that in the slow regime with innovation exponent  $\rho \in (0, 1)$ ,  $\hat{S}$  grows with the same exponent as  $\hat{X}^*$  when  $\alpha > \rho$ , and with a strictly larger exponent if  $\alpha < \rho$ . Similarly, in the steady regime with innovation rate  $q \in (0, 1)$ ,  $\hat{S}$  grows with the same exponent as  $\hat{X}^*$  when  $\alpha > \rho = 1/(1 - q)$  and with a strictly larger exponent if  $\alpha < \rho$ . In other words, the maximal step  $\hat{X}^*$  has a sensible impact in the strong limit theorems of Sect. 3, but its role is negligible for the weak limit theorems of Sect. 4.
- We have worked in the real setting for the sake of simplicity only; the arguments work as well for random walks in  $\mathbb{R}^d$  with  $d \geq 2$ . In this direction, one notably needs a multidimensional version of (4), which can be found in Section 2 of Aaronson and Denker [1]. The same sake of simplicity (possibly combined with the author's laziness) motivated our choice of working with domains of normal attraction rather than with domains of attraction. Most likely, dealing with this more general setting would only require very minor modifications of the present arguments and results.
- It would be interesting to complete the strong limit results (Theorems 1 and 2) and investigate the fluctuations  $n^{-1/\hat{\alpha}} \hat{S}(n) - V'$  as  $n \rightarrow \infty$ . In the setting of the elephant random walk, Kubota and Takei [19] have recently established that these fluctuations are Gaussian.
- The case where the generic step  $X$  has the standard Cauchy distribution is remarkable, due to the feature that for any  $a, b > 0$ ,  $aX_1 + bX_2$  has the same distribution as  $(a + b)X$ , where  $X_1$  and  $X_2$  are two independent copies of  $X$ . It follows that

$n^{-1}\hat{S}(n)$  has the standard Cauchy distribution for all  $n$ , independently of the choice of the sequence  $(\varepsilon_n)$ . This agrees of course with Theorems 1 and 4.

**Funding** Open access funding provided by University of Zurich.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Aaronson, J., Denker, M.: Characteristic functions of random variables attracted to 1-stable laws. *Ann. Probab.* **26**(1), 399–415 (1998)
2. Angel, O., Crawford, N., Kozma, G.: Localization for linearly edge reinforced random walks. *Duke Math. J.* **163**(5), 889–921 (2014)
3. Baur, E.: On a class of random walks with reinforced memory. *J. Stat. Phys.* (2020). <https://doi.org/10.1007/s10955-020-02602-3>
4. Baur, E., Bertoin, J.: On a two-parameter Yule–Simon distribution. [arXiv:2001.01486](https://arxiv.org/abs/2001.01486), to appear in *Progress in Probability: Festschrift R.A. Doney*; Birkhäuser
5. Baur, E., Bertoin, J.: Elephant random walks and their connection to Pólya-type urns. *Phys. Rev. E* **94**, 052134 (2016)
6. Bercu, B.: A martingale approach for the elephant random walk. *J. Phys. A* **51**(1), 015201 (2018)
7. Bertoin, J.: Noise reinforcement for Lévy processes. *Ann. Inst. H. Poincaré Probab. Stat.* **56**(3), 2236–2252 (2020)
8. Bertoin, J.: Universality of noise reinforced Brownian motions. In: *In and Out of Equilibrium 3, celebrating Vidas Sidoravicius* (to appear), *Progress in Probability*, Birkhäuser. [arXiv:2002.09166](https://arxiv.org/abs/2002.09166)
9. Bingham, N.H., Goldie, C.M., Teugels, J.L.: *Regular Variation*. *Encyclopedia of Mathematics and Its Applications*, vol. 27. Cambridge University Press, Cambridge (1987)
10. Bornholdt, S., Ebel, H.: World Wide Web scaling exponent from Simon's 1955 model. *Phys. Rev. E* **64**, 035104 (2001)
11. Businger, S.: The shark random swim (Lévy flight with memory). *J. Stat. Phys.* **172**(3), 701–717 (2018)
12. Coletti, C.F., Gava, R., Schütz, G.M.: Central limit theorem and related results for the elephant random walk. *J. Math. Phys.* **58**(5), 053303 (2017)
13. Coletti, C.F., Gava, R., Schütz, G.M.: A strong invariance principle for the elephant random walk. *J. Stat. Mech. Theory Exp.* **12**, 123207 (2017)
14. Cotar, C., Thacker, D.: Edge- and vertex-reinforced random walks with super-linear reinforcement on infinite graphs. *Ann. Probab.* **45**(4), 2655–2706 (2017)
15. Disertori, M., Sabot, C., Tarrès, P.: Transience of edge-reinforced random walk. *Commun. Math. Phys.* **339**(1), 121–148 (2015)
16. Gnedenko, B.V., Kolmogorov, A.N.: *Limit distributions for sums of independent random variables*. Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills., Ont. (1968)
17. Ibragimov, I.A., Linnik, Y.V.: *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, Groningen (1971). With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman
18. Kious, D., Sidoravicius, V.: Phase transition for the once-reinforced random walk on  $\mathbb{Z}^d$ -like trees. *Ann. Probab.* **46**(4), 2121–2133 (2018)
19. Kubota, N., Takei, M.: Gaussian fluctuation for superdiffusive elephant random walks. *J. Stat. Phys.* **177**(6), 1157–1171 (2019)

20. Kürsten, R.: Random recursive trees and the elephant random walk. *Phys. Rev. E* **93**(3), 032111 (2016)
21. Pemantle, R.: A survey of random processes with reinforcement. *Probab. Surv.* **4**, 1–79 (2007)
22. Sabot, C., Tarrès, P.: Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. *JEMS* **17**(9), 2353–2378 (2015)
23. Schütz, G.M., Trimper, S.: Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk. *Phys. Rev. E* **70**, 045101 (2004)
24. Simon, H.A.: On a class of skew distribution functions. *Biometrika* **42**(3/4), 425–440 (1955)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.