Annales Univ. Sci. Budapest., Sect. Comp. 54 (2023) 171–198

ELEPHANT RANDOM WALKS; A REVIEW

Allan Gut (Uppsala, Sweden)
Ulrich Stadtmüller (Ulm, Germany)

Communicated by Imre Kátai
(Received April 12, 2023; accepted June 16, 2023)

Abstract. In the simple random walk the steps are independent, whereas in the elephant random walk (ERW), which was introduced by Schütz and Trimper in 2004 [33], the next step always depends on the whole path so far. In a series of earlier papers we have investigated some variations and extensions, in particular cases when the elephant has a restricted memory. In the present paper we summarize and extend some results on elephant random walks, closing with some remarks and open questions.

1. Introduction

In the classical *simple* random walk the steps are equal to plus or minus one and independent: $P(X = 1) = 1 - P(X = -1)$; the walker has no memory. The walk is, in particular, Markovian. The other extreme is when the walker has a complete memory, that is, when the next step depends on the whole process so far. This is the case for the so called elephant random walk (ERW) that was introduced by Schütz and Trimper [33] in 2004. The name refers to the fact that elephants have a very long memory. The first mathematically more rigorous papers in the area are, to the best of our knowledge, due to Bercu [3] and Coletti et al. [13]. A main point is that the ERW is subject to a phase transition, which results in a diffusive regime, a critical regime, and a superdiffusive regime, with somewhat different asymptotics. It turns, för example, out that usual strong laws hold in the former cases but not in the superdiffusive one.

Key words and phrases: Elephant random walks, number of zeros, law of large numbers, central limit theorem, Markov chain.
2010 Mathematics Subject Classification: 60F05, 60G50, 60F15, 60J10.

Our interest in the topic centers around cases when the walker has some (restricted) memory. In a series of earlier papers we have investigated some variations and extensions in this direction, such as models in which the elephant remembers only some distant past, only a recent past or a mixture of both. In those models there is, contrary to the classical case, no phase transition. Moreover, in some instances the behavior is dramatically different.

Some of our work was motivated by simulations [34, 16, 31]. There are also connections to the concept of memory lapse; [18].

After describing some models and tools we review the classical case, after which we present a collection of our findings for the various generalizations of the ERW and replenish them with related results from the literature. We add some new results on first passage times. Proofs are scarce; for details we refer to the original papers. We close with some remarks for future thoughts and an appendix containing some mathematical tools.

2. Models

Following [3], the elephant random walk is defined as a random walk in which the first step X_1 equals 1 with probability $s \in [0, 1]$ and -1 with probability $1 - s$. After n steps, at position $S_n = \sum_{k=1}^{n} X_k$, the next step is defined as

$$X_{n+1} = \begin{cases} +X_K, & \text{with probability} \quad p \in [0,1], \\ -X_K, & \text{with probability} \quad 1-p, \end{cases} \tag{2.1}$$

where $K = K(n)$ has a uniform distribution on the integers $1, 2, \ldots, n$. For convenience, we assume that $s = p$.

Via martingale methods Bercu, in [3], establishes asymptotics for the ERW. It turns out that the results depend crucially on the parameter p. More precisely, there is a diffusive regime ($0 \leq p < 3/4$), a critical regime ($p = 3/4$) and a superdiffusive one ($3/4 < p \leq 1$).

In [22] we investigated analogous questions under the assumption that the elephant has a limited memory, such as remembering only the first step, the last step or both. We describe the memory via the index set $\mathfrak{M}_n \subset \{1, \ldots, n\}$. Since there is no martingale around we had to rely on alternative methods.

In [24] we additionally assumed that the elephant may stay put at each step. The rule for a next step is then governed by

$$X_{n+1} = \begin{cases} +X_K, & \text{with probability} \quad p \in [0,1], \\ -X_K, & \text{with probability} \quad q \in [0,1], \\ \quad 0, & \text{with probability} \quad r \in [0,1], \end{cases} \tag{2.2}$$

where $p+q+r=1$, and where $K=K(n)$ has a uniform distribution over $\mathfrak{M}_n$. Let us mention in passing that, in the special case when the memory consists of the most recent step only, the ERW stops as soon as a zero appears.

When one of us (A.G.) introduced Svante Janson to the world of elephant random walks he remarked "det blir en väldig massa nollor" (there will be a lot of zeros). This triggered us to make a quantitative investigation of this fact (cf. [23]).

Notation

We use the standard δ_a to denote the distribution function with a jump of height one at a and $\mathcal{N}_{\mu,\sigma^2}$ for the normal distribution with mean μ and variance σ^2. The arrows $\stackrel{a.s.}{\to}$, $\stackrel{p}{\to}$ and $\stackrel{d}{\to}$ denote convergence almost surely, in probability and in distribution, respectively, $\stackrel{d}{=}$ denotes equality in distribution, and $\mathbb{1}\{\cdot\}$ denotes the indicator function of the set in braces. As for distributions, W and $S(1/2,1/2)$ denote the Wiener process and the stable subordinator with parameters $1/2$, respectively. Constants c and C are numerical constants that may change between appearances.

3. Conditioning

In the classical case the behavior of the next step is governed by the relation

$$E(X_{n+1}\mid\mathcal{G}_n)=(2p-1)\cdot\frac{S_n}{n},\tag{3.1}$$

where $\mathcal{G}_n=\sigma\{X_1,X_2,\ldots,X_n\}$, so that, noticing that $X_k^2=1$ for all k,

$$E(S_{n+1}\mid\mathcal{G}_n) = \big(1+\frac{2p-1}{n}\big)\cdot S_n;\tag{3.2}$$

$$E((S_{n+1})^2\mid\mathcal{G}_n) = \big(1+\frac{2(2p-1)}{n}\big)\cdot(S_n)^2+1.\tag{3.3}$$

In the delayed case the relations become somewhat more complicated. With $V_n=\sum_{k=1}^n X_k^2$,

$$\begin{aligned}E(S_{n+1}\mid\mathcal{G}_n) &= \big(1+\frac{p-q}{n}\big)\cdot S_n;\\ E((S_{n+1})^2\mid\mathcal{G}_n) &= \big(1+\frac{2(p-q)}{n}\big)\cdot(S_n)^2+(1-r)\frac{V_n}{n};\\ E((V_{n+1})^2\mid\mathcal{G}_n) &= \Big(1+\frac{1-r}{n}\Big)V_n.\end{aligned}$$

Taking expectations on both sides provides difference equations that can be taken care of (via Proposition A.2).

Next we need analogs for ERWs with restricted memories. Toward that end, let $\mathcal{F}_n = \sigma\{X_k,\, k \in \mathfrak{M}_n\}$ for $n \geq 1$. The analog of (3.1) then turns out as

$$E(X_{n+1} \mid \mathcal{F}_n) = (2p-1) \cdot \frac{\sum_{i \in \mathfrak{M}_n} X_i}{|\mathfrak{M}_n|}. \tag{3.4}$$

Conditioning on steps that are not contained in the memory means conditioning on steps that the elephant does not remember. Thus, if $A \subset \{1, 2, \ldots, n\}$ is an arbitrary set of indices, then

$$E(X_{n+1} \mid \sigma\{A \cup \mathfrak{M}_n\}) = E(X_{n+1} \mid \mathcal{F}_n) = (2p-1) \cdot \frac{\sum_{i \in \mathfrak{M}_n} X_i}{|\mathfrak{M}_n|}, \tag{3.5}$$

where, throughout, $2p-1$ is replaced by $p-q$ in the delayed case.

4. The classical case

In order to analyze the behavior of the classical ERW we have to find proper representations. Using the notation of [3] we set

$$\gamma_n = 1 + \frac{2p-1}{n} \qquad \text{and} \qquad a_n = \prod_{k=1}^{n-1} \gamma_n^{-1} = \frac{\Gamma(n)\Gamma(2p)}{\Gamma(n+2p-1)}.$$

It follows, in view of Lemma A.1, that $\{M_n = a_n S_n,\, n \geq 1\}$ is a martingale.

Exploiting the martingale property Bercu [3] establishes (*i.a.*) the following result, from which we learn that the process has different regimes depending on the parameter p.

Theorem 4.1. (a) *If* $0 < p < 3/4$ *then*

$$\frac{S_n}{n} \overset{a.s.}{\to} 0 \quad \textit{and} \quad \frac{S_n}{\sqrt{n}} \overset{d}{\to} \mathcal{N}_{0,1/(3-4p)} \qquad \textit{as} \quad n \to \infty.$$

(b) *If* $p = 3/4$ *then*

$$\frac{S_n}{\sqrt{n}\log n} \overset{a.s.}{\to} 0 \quad \textit{and} \quad \frac{S_n}{\sqrt{n\,\log n}} \overset{d}{\to} \mathcal{N}_{0,1} \qquad \textit{as} \quad n \to \infty.$$

(c) *If* $p > 3/4$ *then*

$$\frac{S_n}{n^{2p-1}} \overset{a.s.}{\to} L \qquad \textit{as} \quad n \to \infty, \ \textit{for some nondegenerate rv } L,$$

defined in Theorem 3.7 *of* [3, 5], *having moments*

$$E(L) = \frac{1-2p}{\Gamma(2p)} \quad \textit{and} \quad E(L^2) = \frac{1}{(4p-3)\Gamma(2(2p-1))}.$$

Remark 4.1. (Cf. [3]) For $0 < p < 3/4$, there is also a LIL,

$$\limsup_{n\to\infty}(\liminf_{n\to\infty}) \frac{S_n}{\sqrt{2n\log\log n}} = \frac{1}{\sqrt{3-4p}} \quad \Big(-\frac{1}{\sqrt{3-4p}}\Big),$$

and a quadratic strong law,

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{S_k^2}{k^2} = \frac{1}{3-4p} \qquad \text{a.s.}$$

Further representations can be obtained by imbedding $\{M_n, n \geq 1\}$ into a Wiener process $\{W(t), t \geq 0\}$, with stopping times T_n, $n \geq 1$, so that $M_n = W(T_n)$ for all n, or asymptotically by a weak invariance principle; cf. Proposition A.3.

4.1. General steps

It is a relatively easy task to extend the results from [3] to a model in which arbitrary steps are allowed. Namely, let $\{\widetilde{S}_n, n \geq 1\}$ be an ERW, and suppose that R is a random variable with distribution function F_R that is independent of the walk. If $\widetilde{S}_n/c_n \overset{a.s.}{\to} Z$ as $n \to \infty$ for some normalizing positive sequence $c_n \to \infty$ as $n \to \infty$ and some random variable Z, it follows from Proposition A.1 that $R\widetilde{S}_n/c_n \overset{a.s.}{\to} RZ$ as $n \to \infty$. We can therefore extend Theorems 3.1, 3.4 and (the first half of) Theorem 3.7 of [3]. Namely, consider the ERW for which $\widetilde{X}_1 \equiv 1$, and let the random variables $\widetilde{X}_n$, $n \geq 2$. Furthermore, let R be a random variable, independent of $\{\widetilde{X}_n, n \geq 1\}$, and consider $X_n = R \cdot \widetilde{X}_n$, $n \geq 1$, and, hence, $S_n = R \cdot \widetilde{S}_n$.

Theorem 4.2. (a) *For* $0 < p < 3/4$, $\quad \dfrac{S_n}{n} \overset{a.s.}{\to} 0 \quad$ *as* $\quad n \to \infty$;

(b) *For* $p = 3/4$, $\quad \dfrac{S_n}{\sqrt{n}\log n} \overset{a.s.}{\to} 0 \quad$ *as* $\quad n \to \infty$;

(c) *For* $3/4 < p < 1$, $\quad \dfrac{S_n}{n^{2p-1}} \overset{a.s.}{\to} RL \quad$ *as* $\quad n \to \infty$,

where L *is the same as in Theorem* 4.1.

As for convergence in distribution, we have to distinguish more carefully between the three cases.

Theorem 4.3. *For $0 < p < 3/4$,*

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} \int\limits_{\mathbb{R}\setminus\{0\}} \mathcal{N}_{0,\frac{1}{3-4p}}(\cdot/|t|)\,dF_R(t) + P(R=0)\cdot\delta_{[0,\infty)}(\cdot) \quad \textit{as} \quad n\to\infty.$$

Moreover, if $E(R^2) < \infty$, then

$$E(S_n/\sqrt{n}) \to 0 \quad \textit{and} \quad E((S_n/\sqrt{n})^2) \to E(R^2)/(3-4p) \quad \textit{as} \quad n\to\infty.$$

For the critical case, $p = 3/4$, one similarly obtains, using [3], *Theorem* 3.6, *that*

$$\frac{S_n}{\sqrt{n\log n}} \xrightarrow{d} \int\limits_{\mathbb{R}\setminus\{0\}} \mathcal{N}_{0,1}(\cdot/|t|)\,dF_R(t) + P(R=0)\cdot\delta_{[0,\infty)}(\cdot) \quad \textit{as} \quad n\to\infty.$$

Remark 4.2. The supercritical case, $3/4 < p < 1$, has a different evolution and no analogous result exists.

4.2. Restricted memories

In this subsection we present some results, together with hints on proofs in order to give a flavor of our procedure. For additional details we refer to our original paper [22].

The easiest case is $\mathfrak{M}_n = \{1\}$, because then every step equals X_1 with probability p and $-X_1$ with probability $1-p$. This means that S_n operates like an ordinary simple random walk from step 2 on. Considering the two initial values this yields

$$\frac{S_n}{n} \xrightarrow{d} \begin{cases} 2p-1, & \text{with probability} \quad p, \\ -(2p-1), & \text{with probability} \quad 1-p, \end{cases} \quad \text{as} \quad n\to\infty,$$

and

$$E\Big(\frac{S_n}{n}\Big) \to (2p-1)^2 \quad \text{and} \quad \operatorname{Var}(S_n/n) \to 4p(1-p)(2p-1)^2 \quad \text{as} \quad n\to\infty.$$

If $\mathfrak{M}_n = \{1,2\}$ there are three possibilities for the first two summands; both equal to $+1$, both equal to -1 and unequal. Analogous arguments show that

$$\frac{S_n}{n} \xrightarrow{d} \begin{cases} 2p-1, & \text{with probability} \quad p^2, \\ 0, & \text{with probability} \quad 1-p, \\ -(2p-1), & \text{with probability} \quad p(1-p), \end{cases} \quad \text{as} \quad n\to\infty,$$

and that

$$E\Big(\frac{S_n}{n}\Big) \to p(2p-1)^2 \;\text{ and }\; \operatorname{Var}(S_n/n) \to p(1-p)(2p-1)^2(4p^2+1) \text{ as } n\to\infty.$$

Theorem 4.4. *For* $\mathfrak{M}_n = \{1, \ldots, m\}$ *with* $q_k = P(S_m = m - 2k)$ *and* $p_k = \frac{m-2k}{m}(2p-1)$ *for* $0 \le k \le m$,

$$\frac{S_n}{n} \xrightarrow{d} \sum_{k=0}^{m} q_k \delta_{p_k} \quad \text{as } n \to \infty.$$

Remark 4.3. The quantities q_k are in general not easily computed. However, $q_0 = p^m$ and $q_m = (1-p)p^{m-1}$.

The other extreme is when only the most recent step is remembered, viz., $\mathfrak{M}_n = \{n\}$. This model is also called a correlated random walk (CRW), see e.g. [32].

We first assume $X_1 = 1$, and denote the partial sums by T_n, $n \ge 1$. Then,

$$E(E(X_{n+1} \mid \mathcal{F}_n)) = E\big(E(X_{n+1} \mid X_n)\big) = (2p-1) \cdot E(X_n) \quad \text{for all} \quad n.$$

By iterating it follows, for $n \ge 0$, that

$$E(X_{n+1}) = (2p-1)^n E(X_1) = (2p-1)^n,$$

and that

$$E(T_{n+1}) = \frac{1-(2p-1)^{n+1}}{2(1-p)}, \qquad \text{hence,} \quad E\Big(\frac{T_n}{n}\Big) \to 0 \quad \text{as} \quad n \to \infty.$$

For the second moments we obtain, via conditioning,

$$E(T_{n+1}^2) = E(T_n^2) + 2(2p-1)E(T_n X_n) + 1\,.$$

For the middle term we similarly obtain

$$E(T_n X_n) = E(X_n^2) + E(T_{n-1}E(X_n \mid \mathcal{G}_{n-1})) = 1 + (2p-1)E(T_{n-1}X_{n-1})\,,$$

which, after iteration, yields

$$E(T_n X_n) = 1 + \sum_{k=1}^{n-1}(2p-1)^k = \frac{1-(2p-1)^n}{2(1-p)}\,.$$

Combining the two relations and telescoping finally yields, as $n \to \infty$,

$$E(T_{n+1}^2) = \frac{np}{1-p} + \mathcal{O}(1), \qquad \text{hence,} \quad \operatorname{Var}(T_{n+1}) = \frac{np}{1-p} + \mathcal{O}(1)\,.$$

Noticing that $S_n = X_1 T_n$ and that $X_1 = \pm 1$ establishes the following result:

Theorem 4.5. $\frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_{0,\frac{p}{1-p}}$ *as* $n \to \infty$.

The Markov property provides, in addition, a strong law:

$$\frac{S_n}{n} \xrightarrow{a.s.} 0 \quad \textit{as} \quad n \to \infty.$$

Remark 4.4. If the elephant remembers a fixed, but finite, number, k say, of the most recent steps, the sequence of steps forms a Markov chain of order k, and we obtain, by (basically) the same arguments as above that $S_n/\sqrt{n}$ will be asymptotically normal (a Markov chain of order k can be considered as a k-dimensional Markov chain).

We close this subsection with a combination of early and recent steps, notably the simpliest case.

Theorem 4.6. *For* $\mathfrak{M}_n = \{1, n\}$,

$$\frac{S_n}{n} \xrightarrow{d} S = \begin{cases} \dfrac{2p-1}{3-2p}, & \textit{with probability} \quad p, \\ -\dfrac{2p-1}{3-2p}, & \textit{with probability} \quad 1-p, \end{cases} \qquad \textit{as} \quad n \to \infty.$$

Moreover, $E(S_n/n)^r \to E(S^r)$ *for all* $r > 0$*, since* $|S_n/n| \le 1$ *for all* n.

Here is a quick sketch of the proof. For $X_1 = 1$ we have $E(X_2) = 2p-1$, and, via conditioning, that

$$E(X_{n+1}) = (2p-1)E\Big(\frac{X_1 + X_n}{2}\Big) = \frac{2p-1}{2} \cdot (1 + E(X_n)).$$

Exploiting Proposition A.2(i) we obtain, for $n \ge 1$,

$$E(X_n) = \frac{2p-1}{3-2p} + \Big(\frac{2p-1}{2}\Big)^{n-1} \cdot \frac{4(1-p)}{3-2p},$$

and, hence, after summation, that

$$E(T_n) = n \cdot \frac{2p-1}{3-2p} + \frac{8(1-p)}{(3-2p)^2} + o(1) \quad \text{as} \quad n \to \infty.$$

The relations for the second and mixed moments are

$$E(T_{n+1}^2) = E(T_n^2) + (2p-1)E(T_n) + (2p-1)E(T_n X_n) + 1,$$

and

$$E(T_{n+1}X_{n+1}) = \frac{2p-1}{2} \cdot E(T_n X_n) + \frac{2p-1}{2} E(T_n) + 1,$$

respectively.

Some tedious computations and appeals to Proposition A.2 complete the proof of the theorem.

4.3. Increasing memory

What happens if we have a slowly increasing memory; $\mathfrak{M}_n = \{m_n, \dots, n\}$ with $m_n \nearrow \infty$ but $m_n/n \to 0$? Here is a main result. For details and further results we refer to [25].

Theorem 4.7. (a) *If* $0 < p < 3/4$, *then*

$$\frac{S_n\sqrt{m_n}}{n} \xrightarrow{d} \mathcal{N}_{0,(2p-1)^2/(3-4p)} \quad \text{as} \quad n \to \infty\,.$$

(b) *If* $p = 3/4$, *then*

$$\frac{S_n\sqrt{m_n/\log m_n}}{n} \xrightarrow{d} \mathcal{N}_{0,1/4} \quad \text{as} \quad n \to \infty\,.$$

(c) *If* $3/4 < p < 1$, *then*

$$\frac{S_n m_n^{2(1-p)}}{n} \xrightarrow{d} (2p-1)\,L \quad \text{as} \quad n \to \infty\,,$$

where the random variable L *is the same as in Theorem* 4.1.

Remark 4.5. If $p = 1/2$, we recall that the ERW reduces to coin-tossing, hence, $S_n/\sqrt{n} \xrightarrow{d} \mathcal{N}_{0,1}$ as $n \to \infty$. Note also that Theorem 4.7 (a) reduces to $S_n\sqrt{m_n}/n \xrightarrow{p} 0$ as $n \to \infty$ in that case.

Remark 4.6. Another approach is given in [29] by Laulin. She replaces the uniform distribution of the index variable K in the classical model with probabilities $P(K = k) = \frac{(\beta+1)\Gamma(\beta+k)\Gamma(n)}{\Gamma(k)\Gamma(\beta+n+1)}$ for $1 \le k \le n$ and $\beta \ge 0$. The case $\beta = 0$ corresponds to the classical case, and a large β corresponds to an elephant losing its early memory. In Laulin's case there is a phase transition with change point $p = \frac{4\beta+3}{4(\beta+1)}$. Hence, it is the early memory which causes the phase transition.

5. ERWs with delays — ERWD [24]

Next we introduce the possibility of delays in that the elephant, in addition, has a choice of staying put in every step. The steps are then governed by

$$X_{n+1} = \begin{cases} +X_K, & \text{with probability} \quad p \in [0,1], \\ -X_K, & \text{with probability} \quad q \in [0,1], \\ \;\;0, & \text{with probability} \quad r \in [0,1], \end{cases} \tag{5.1}$$

where $p + q + r = 1$, and where K has a uniform distribution on $\mathfrak{M}_n$. Here $X_1 = 1, -1, 0$ with probabilities p, q, and r respectively. Everything reduces, of course, to the classical case if $r = 0$. Since it turns out that a key task is to keep track of the number of zeros, it is convenient to begin by considering the following *one-zero* ERW.

5.1. The Bernoulli-ERW — BERW

We naturally define the BERW via

$$X_{n+1} = \begin{cases} X_K, & \text{with probability} \quad \alpha \in [0,1], \\ 0, & \text{with probability} \quad 1-\alpha. \end{cases} \tag{5.2}$$

Just as the ERW is a special case of an ERWD with $r = 0$, the BERW is a special case of an ERWD with $q = 0$.

For the following result we need to define the Mittag–Leffler distribution:

Definition 5.1. Let $0 < \alpha < 1$. The Mittag–Leffler distribution with parameter α, $M\!L_\alpha$, is a distribution with moment generating function $M_\alpha(t) = \sum_{n=0}^{\infty} t^n/\Gamma(1+n\,\alpha)$.

Theorem 5.1. ([26]) *Let $0 < \alpha < 1$. Then,*

$$P\Big(\frac{S_n}{n^\alpha} \le x\Big) \longrightarrow \alpha \cdot M\!L_\alpha(x) + (1-\alpha)\cdot \delta_0(x) \qquad \text{as} \quad n \to \infty\,,$$

where $M\!L_\alpha$ denotes the Mittag–Leffler distribution with parameter α.

Moreover, for the limit variable Y_α,

$$E(Y_\alpha) = \frac{1}{\Gamma(\alpha)} \qquad \text{and} \qquad \operatorname{Var}(Y_\alpha) = \frac{2\alpha}{\Gamma(1+2\alpha)} - \frac{1}{\Gamma^2(\alpha)}\,.$$

The proof amounts to the usual computations, however, of all moments, after which one identifies them with the Mittag–Leffler distribution.

5.1.1. Restricted memories

We confine ourselves to stating results for our two most elementary cases from [26] and refer to our paper for more.

Theorem 5.2. *If $\mathfrak{M}_n = \{1\}$, then* $\displaystyle\frac{S_n}{n} \overset{a.s.}{\to} p\,\mathbb{1}\{X_1 = 1\}$ *as* $n \to \infty$.

Theorem 5.3. *If $\mathfrak{M}_n = \{n\}$, then* $S_n \overset{a.s.}{\to} S$ *as* $n \to \infty$,

where S is a geometric random variable with mean $E(S) = p/(1-p)$.

We now return to the ERWD and its asymptotic behavior.

5.2. The ERWD with full memory

Recall (5.1) and set $\tilde{p}_r = 2p - 1 + r$.

Proposition 5.4. *As $n \to \infty$,*

$$E(S_n) \sim \frac{n^{\tilde{p}_r}}{\Gamma(\tilde{p}_r)};$$

$$\operatorname{Var}(S_n) \sim \begin{cases} \dfrac{(1-r)\, n^{1-r}}{\Gamma(1-r)(1-r-2\tilde{p}_r)}, \\ \dfrac{1-r}{\Gamma(1-r)} n^{2\tilde{p}_r} \log n = \dfrac{1-r}{\Gamma(1-r)}\, n^{1-r} \log n, \\ n^{2\tilde{p}_r} \Big(\dfrac{1-r}{(2\tilde{p}_r - (1-r))\Gamma(2\tilde{p}_r)} - \dfrac{1}{\Gamma^2(\tilde{p}_r)} \Big), \end{cases}$$

for $0 < p < \frac{3}{4}(1-r)$, $p = \frac{3}{4}(1-r)$, and $\frac{3}{4}(1-r) < p < 1$, respectively.

In this setting there are two martingales involved, namely,

$$\Big\{ \frac{\Gamma(n)\Gamma(\tilde{p}_r+1)}{\Gamma(n+\tilde{p}_r)}\, S_n,\ n \geq 1 \Big\} \quad \text{and} \quad \Big\{ \frac{\Gamma(n)\Gamma(1-r+1)}{\Gamma(n+1-r)}\, V_n,\ n \geq 1 \Big\},$$

where $V_n = \sum_{k=1}^{n} X_k^2$ for all n. In the diffusive case with $|X_1| = 1$, we use Theorem 1 in [28] and obtain $S_n/\sqrt{V_n}$ is asymptotically normal and V_n is a BERW with parameter $\alpha = 1-r$. This leads (cf. [28]) to F_{r,σ^2} being the distribution function of a random variable $Z = \sqrt{Y} \cdot W$, where $Y \in ML_{1-r}$ is independent $W \in \mathcal{N}_{0,\sigma^2}$. Therefore,

$$F_{r,\sigma^2}(x) = \int_0^\infty \int_{-\infty}^{x/v} e^{-w^2/(2\sigma^2)} \frac{dw}{\sqrt{2\pi\sigma^2}} ML_{1-r}(dv),$$

which is a symmetric distribution with $E(Z) = 0$ and $E(Z^2) = \frac{\sigma^2}{\Gamma(1-r)}$.

The following result is obtained by adapting the proofs from [4].

Theorem 5.5. ([24]) (a) *For $p < \frac{3}{4}(1-r)$ and $\sigma^2 = 1/(3(1-r)-4p)$,*

$$P\Big(\frac{S_n}{\sqrt{n^{1-r}}} \leq x \Big) \xrightarrow{d} (1-r)F_{r,\sigma^2}(x) + r\,\delta_0(x) \quad as \quad n \to \infty.$$

Furthermore

$$\frac{S_n}{n^{1-r}} \xrightarrow{a.s.} 0 \quad as \quad n \to \infty.$$

(b) *For $p = \frac{3}{4}(1-r)$,*

$$P\Big(\frac{S_n}{\sqrt{n^{1-r}\log n}} \leq x \Big) \xrightarrow{d} (1-r)F_{r,1} + r\,\delta_0(x) \quad as \quad n \to \infty.$$

And, again,

$$\frac{S_n}{n^{1-r}} \stackrel{a.s.}{\to} 0 \quad \text{as} \quad n \to \infty.$$

(c) *For* $\frac{3}{4}(1-r) < p < 1$,

$$\frac{S_n}{n^{2p-1+r}} \stackrel{a.s.}{\to} L \cdot \mathbb{1}\{X_1 = 1\} - L \cdot \mathbb{1}\{X_1 = -1\} \quad \text{as} \quad n \to \infty,$$

where L *is a non-degenerate random variable with moments*

$$E(L) = \frac{1}{\Gamma(2p+r)} \quad \text{and} \quad E(L^2) = \frac{1-r}{(2\tilde{p}_r - (1-r))\,\Gamma(2\tilde{p}_r)},$$

with $\tilde{p}_r = 2p - 1 + r$.

5.3. Restricted memories [24]

If the walker only remembers the first step, the game is over if that step equals zero.

Theorem 5.6. *If* $\mathfrak{M}_n = \{1\}$, *then*

$$\frac{S_n}{n} \stackrel{d}{\to} \begin{cases} p-q, & \text{with probability} \quad p, \\ 0, & \text{with probability} \quad r, \\ -(p-q), & \text{with probability} \quad q, \end{cases} \quad \text{as} \quad n \to \infty.$$

Moreover,

$$E(S_n/n) \to (p-q)^2 \text{ and } \operatorname{Var}(S_n/n) \to (p-q)^2\big(p+q-(p-q)^2\big) \quad \text{as} \quad n \to \infty.$$

The case $\mathfrak{M}_n = \{1, 2\}$ is similar with six cases to take care of.

A very special model is when the elephant only remembers the most recent step, because in that case the walk terminates as soon as a zero appears. Letting $\tau = \min\{n : X_n = 0\}$ we observe that τ is geometric with mean $1/r$. The $\tau - 1$ zero-truncated summands $\tilde{X}_k$ are thus coin-tossing random variables with mean $(p-q)/(1-r) = (p-q)/(p+q)$. They are, however, not independent of the number of them, and we can therefore not apply e.g. Theorem 2.15.1 in [20].

Theorem 5.7. *Suppose that* $\mathfrak{M}_n = \{n\}$, *let* $\tau = \min\{n : X_n = 0\}$. *Then*

(a) τ *is geometric with mean* $\dfrac{1}{1-p-q}$;

(b) $S_n \stackrel{a.s.}{\to} S_\tau \quad \text{as} \quad n \to \infty$;

(c) $E(S_\tau) = \dfrac{1-p}{1-p+q} \quad$ *and* $\quad E(S_\tau^2) = 1 + \dfrac{(p-q)(1-2p+(p+q)^2)}{qr(1-p+q)}$.

We close with the case when the memory consists of the first and most recent steps.

Theorem 5.8. *For* $\mathfrak{M}_n = \{1, n\}$,

$$\frac{S_n}{n} \xrightarrow{d} S = \begin{cases} \dfrac{p-q}{2+q-p}, & \text{with probability} \quad p, \\ 0, & \text{with probability} \quad r, \\ -\dfrac{p-q}{2+q-p}, & \text{with probability} \quad q, \end{cases} \quad \text{as} \quad n \to \infty.$$

Moreover, $E(S_n/n)^r \to E(S^r)$ *for all* $r > 0$*, since* $|S_n/n| \le 1$ *for all* n*.*

6. More on the number of zeros in ERWDs [23, 26]

Since it is mathematically more convenient to investigate the number of non-zeros, set, for $n \in \mathbb{N}$

$$N_n = \sum_{k=1}^{n} I_k, \text{ where } I_n = \mathbb{1}_{\{X_n=0\}} \text{ and } N_n^* = \sum_{k=1}^{n} I_k^*, \text{ where } I_n^* = \mathbb{1}_{\{X_n\neq 0\}},$$

and note that any starred result can easily be transferred to a non-starred one, since $N_n + N_n^* = n$.

6.1. The classical case

The starred process forms a BERW. Exploiting the fact that

$$\Big\{\Big(M_n^* = \frac{\Gamma(n)\Gamma(2-r)}{\Gamma(n+1-r)} N_n^*, \mathcal{F}_n\Big)\; n \ge 1\Big\} \quad \text{is nonnegative a martingale,}$$

with $\{\mathcal{F}_n, n \ge 1\}$ being the natural filtration, one can prove the following result.

Theorem 6.1. *There exists a random variable* Y*, such that* $M_n^* \xrightarrow{a.s.} Y$ *as* $n \to \infty$*, and, hence, that*

$$\frac{N_n^*}{n^{1-r}} = \frac{n - N_n}{n^{1-r}} \xrightarrow{a.s.} \frac{Y}{\Gamma(2-r)} \quad \text{as} \quad n \to \infty.$$

Moreover, convergence holds in L^1*, in particular,*

$$E\Big(\frac{N_n^*}{n^{1-r}}\Big) = E\Big(\frac{n - N_n}{n^{1-r}}\Big) \to \frac{1}{\Gamma(1-r)} \quad \text{as} \quad n \to \infty.$$

Proof. We first note that $E(I_1^*) = 1 - r$ and that $E(I_2^*) = (1-r)^2$. Furthermore,

$$E(I_{n+1}^* \mid \mathcal{F}_n) = (1-r) \cdot \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{I_k^* \neq 0\}} = \frac{1-r}{n} N_n^*,$$

and thus

$$E(N_{n+1}^* \mid \mathcal{F}_n) = \frac{1-r}{n} N_n^* + N_n^*,$$

so that

$$\begin{aligned} E(N_{n+1}^*) &= \prod_{k=1}^{n} \frac{k+1-r}{k} E(N_1^*) = \frac{1-r}{\Gamma(2-r)} \cdot \frac{\Gamma(n+1+(1-r))}{\Gamma(n+1)} = \\ &= \frac{\Gamma(n+2-r)}{\Gamma(1-r)\Gamma(n+1)}. \end{aligned}$$

Lemma A.2 now tells us that, as $n \to \infty$,

$$E(N_{n+1}^*) = \frac{n^{1-r}}{\Gamma(1-r)} + n^{-r} \frac{(1-r)(2-r)}{2\Gamma(1-r)} + \mathcal{O}(n^{-r-1}). \tag{6.1}$$

A similar, but more complicated, calculation of second moments leads, via Lemma A.1, to asymptotics of the quadratic variation $\{< M_n^* >, n \geq 1\}$. Almost sure convergence for the starred process now follows via a standard convergence result, see, e.g., p. 510 in [20]. Mean convergence is due to the boundedness of the variances. The limit of the expected values is immediate from (6.1). ■

Remark 6.1. It follows from the convergence of the expected values that $E(Y) = 1 - r$. Comparing with Theorem 5.1 shows, in addition, that $Y \stackrel{d}{=} \stackrel{d}{=} \Gamma(2-r)\big((1-r)M\!L_{(1-r)} + r\delta_0\big)$.

6.2. Restricted memories

When $\mathfrak{M}_n = \{1\}$ it follows (again) that the indicators are independent and identically distributed. Moreover, $E(I_1^*) = 1 - r$ and $E(I_{n+1}^* \mid X_1) = (1-r)\,|X_1|$ for all n, which imples that, given $I_1^* = 1$, $\{N_n^*, n \geq 1\}$, is a binomial random walk, and is identically zero if $I_1^* = 0$. Note that Theorems 6.2 and 6.4 are slight corrections of the corresponding results in [23].

Theorem 6.2. *For* $\mathfrak{M}_n = \{1\}$,

$$\frac{N_n^*}{n} \stackrel{a.s.}{\to} (1-r)\, I_1^* \quad \text{and} \quad \frac{N_n}{n} \stackrel{a.s.}{\to} 1 - (1-r)\, I_1^* \quad \text{as} \quad n \to \infty.$$

For ERWDs with $\mathfrak{M}_n = \{n\}$ there will be almost surely a finite number of non-zero steps. Recall from Theorem 5.7 that $\tau = \min\{n : X_n = 0\}$ has a geometric distribution with mean $1/r$. Moreover, $P(N_n^* = k) = (1-r)^k r$, for $k = 0, \dots, n-1$, that $P(N_n^* = n) = (1-r)^n$. The generating function therefore turns out as

$$g_{N_n^*}(t) = \frac{r\,(1-((1-r)t)^n}{1-(1-r)t} + (1-r)^n t^n \to \frac{r}{1-(1-r)t} = g(t) \quad \text{as} \quad n \to \infty$$

(provided $t < 1/(1-r)$). The limiting generating function is that of a geometric random variable with mean $(1-r)/r$; in particular, the generating function of $\tau - 1$.

Theorem 6.3. *For $\mathfrak{M}_n = \{n\}$,*

$$N_n^* = n - N_n \overset{a.s.}{\to} Z \quad \text{as} \quad n \to \infty,$$

where $Z = \tau - 1$ is a geometric random variable with mean $\frac{1-r}{r}$. Moreover, all moments converge.

Proof. Almost sure convergence holds since N_n^* is monotone increasing and bounded almost surely. That $Z = \tau - 1$ is, again, immediate. Convergence of the generating functions does the rest. ■

Remark 6.2. Note that $\{I_n^*, n \geq 1\}$ is a two state Markov chain where one state is absorbing.

Theorem 6.4. *Suppose that $\mathfrak{M}_n = \{1, n\}$. Then*

$$\frac{N_n^*}{n} \overset{a.s.}{\to} \frac{1-r}{1+r} I_1^* \quad \text{and} \quad \frac{N_n}{n} \overset{a.s.}{\to} 1 - \frac{1-r}{1+r} I_1^* \quad \text{as} \quad n \to \infty.$$

Proof. If $X_1 = 0$ the random walk stays put at zero. We therefore suppose in the following that $X_1 \neq 0$, and, hence, that $I_1^* = 1$. Then, $E(I_2^* \mid I_1^*) = 1 - r = E(I_2^*)$, and, generally, for $n \geq 1$,

$$E(I_{n+1}^*) = E\big(E(I_{n+1}^* \mid \mathcal{F}_n)\big) = \frac{1-r}{2} E(I_1^*) + \frac{1-r}{2} E(I_n^*) = \frac{1-r}{2} + \frac{1-r}{2} E(I_n^*).$$

Adding the extreme members yields, for $n \geq 1$,

$$E(N_{n+1}^*) - 1 = \frac{n(1-r)}{2} + \frac{1-r}{2} E(N_n^*),$$

which, after an application of Proposition A.2, tells us that

$$E(N_n^*) = \frac{n(1-r)}{1+r} + \frac{4r}{(1+r)^2} + o(1) \quad \text{as} \quad n \to \infty.$$

Being in the branch with $I_1^* = 1$ we are faced with a stationary ergodic Markov chain, which asserts the validity of the first strong law, after which the second one follows in the usual manner. ■

7. Recurrence, renewal theory and related questions

Apart from Theorem 7.1 the results in this section are new.

Recurrence and transience are connected with return times to zero.

7.1. Full memory

In [14] it was shown that the classical elephant random walk is recurrent for $0 < p \le 3/4$ and transient for $p > 3/4$. After each visit to zero the walk starts afresh, although the situation is different each time, since the walker remembers the whole past. For the diffusive case a more detailed analysis has been provided by Bertoin in [9]:

Theorem 7.1. *Let $0 < p < 3/4$.*

(a) *Let $Z(n) = Card\{1 \le j \le n : S_j = 0\}$ be the number of zeros up to step n. Then there exists a nondegenerate random variable V_p, such that*

$$\frac{Z(n)}{\sqrt{n}} \xrightarrow{d} V_p \quad \text{as} \quad n \to \infty.$$

(b) *The return time to 0, $\tau_0 = \min\{n : S_n = 0\}$ has a heavy-tailed distribution with*

$$P(\tau_0 > n) \sim \frac{\sqrt{6-8p}}{\sqrt{\pi}\,\Gamma(2p)}\, n^{2p-3/2} \qquad \text{as} \quad n \to \infty.$$

Hence $E(\tau_0) < \infty$ iff $p < 1/4$. The result remains true for any further visit to zero.

The basic tools are the embedding of a relevant martingale process into a Wiener process with suitable stopping times and excursion theory for Wiener processes.

Consider possible asymptotics for the first passage times

$$\tau_n^{(+)} = \min\{k \in \mathbb{N} : S_k \ge n\} \; = \min\{k : S_k = n\}, \quad n \ge 1.$$

These are stopping times which are closely related to renewal theory.

A first observation is that the stopping times are not proper random variables in the superdiffusive case.

Proposition 7.2. *$\tau_n^{(+)}$ is not a proper random variable for $3/4 < p < 1$, at least for $n \in \mathbb{N}$ large enough.*

Proof. If $\tau_n^{(+)}$ is finite a.s. for a subsequence $n_k \nearrow \infty$, then, as $\tau_{n_k}^{(+)} \nearrow \infty$,

$$\frac{n_k}{\tau_{n_k}^{(+)}} \xrightarrow{a.s.} L \quad \text{as} \quad k \to \infty.$$

However, since $E(S_n/n^{2p-1} \mid X_1 = -1) \sim -\frac{2p}{\Gamma(2p)}$ as $n \to \infty$, L cannot be a positive random variable, thus yielding a contradiction. Hence, $\tau_{n_k}^{(+)} = \infty$ for all $k \geq k_0$ on a set of positive measure, which implies that $\tau_n^{(+)} = \infty$ for all $n \geq k_0$. ■

In the diffusive case we obtain strong laws and almost sure bounds.

Theorem 7.3. (i) *For* $0 < p < 3/4$, $\displaystyle\liminf_{n\to\infty} \frac{\tau_n^{(+)} \log\log n}{n^2} \geq \frac{3-4p}{2}$ *a.s.*

(ii) *For* $p = 3/4$, $\displaystyle\liminf_{n\to\infty} \frac{\tau_n^{(+)} \log n \log\log\log n}{n^2} \geq \frac{1}{2}$ *a.s.*

Proof. (i): It follows from the LIL, $\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = \frac{1}{\sqrt{3-4p}}$ a.s. (cf. [3], Theorem 3.2), that

$$\limsup_{n\to\infty} \frac{S_{\tau_n^{(+)}}}{\sqrt{2\tau_n^{(+)} \log\log \tau_n^{(+)}}} \leq \frac{1}{\sqrt{3-4p}} \text{ a.s.}$$

along the sequence $\{\tau_n^{(+)}, n \geq 1\}$, which, since $S_{\tau_n^{(+)}} = n$, is the same as

$$\limsup_{n\to\infty} \frac{n}{\sqrt{2\tau_n^{(+)} \log\log \tau_n^{(+)}}} \leq \frac{1}{\sqrt{3-4p}} \text{ a.s.}$$

(ii): The proof follows the same pattern, departing from [3], Theorem 3.5. ■

Remark 7.1. (i) For $0 < p \leq 3/4$ it follows, in particular, that $\tau_n^{(+)}/n \overset{a.s.}{\to} \infty$ and, hence, that $E(\tau_n^{(+)})/n \to \infty$ as $n \to \infty$. This might be interpreted as a negative renewal theorem, in that the almost sure limit equals $1/\mu < \infty$ for random walks with positive mean μ ([19], Theorem 3.6.1).

(ii) For $3/4 < p < 1$ the LLN for S_n does not imply an LLN for $\tau_n^{(+)}$.

Next in turn is a result on distributional convergence.

Theorem 7.4. (i) *For* $0 < p < 3/4$,

$$P\Big(\frac{\tau_n^{(+)}}{n^2} \leq x\Big) \overset{d}{\to} G_p(x) \quad \text{as} \quad n \to \infty,$$

where

$$\begin{aligned} G_p(x) &= P\Big(\frac{1}{\sqrt{3-4p}} \max_{0<t\leq x} \{t^{2p-1} W(t^{3-4p})\} \geq 1\Big) = \\ &= P\Big(\max_{0<t\leq x^{3-4p}} \Big\{t^{(2p-1)/(3-4p)} W(t)\Big\} \geq \sqrt{3-4p}\Big). \end{aligned}$$

(ii) *For* $p = 3/4$, $\quad P\Big(\frac{\tau_n^{(+)} \log n}{n^2} \le x\Big) \xrightarrow{d} 2\Big(1 - \mathcal{N}_{0,1}(1/\sqrt{x})\Big) \quad$ *as* $\quad n \to \infty.$

Proof. (i) Using the weak invariance from Proposition A.3 we find that

$$
\begin{aligned}
P(\tau_n^{(+)} > un^2) &= P(\max_{n \le k \le u\, n^2} S_k < n) = P\Big(\max_{\frac{1}{n} \le \frac{k}{n^2} \le u} \Big\{\frac{S_{\frac{k}{n^2} n^2}}{n}\Big\} < 1\Big) = \\
&= P\Big(\frac{1}{\sqrt{3-4p}} \max_{0 \le t \le u} \{t^{2p-1} W(t^{3-4p})\} < 1\Big)(1 + o(1)).
\end{aligned}
$$

(ii) This part follows the same pattern. For $v(k,n) = \exp\{\frac{k \log n}{n^2}\, n^2/\log n\}$ we have

$$
\begin{aligned}
P\Big(\frac{\tau_n^{(+)} \log n}{n^2} > u\Big) &= P\Big(\max_{n \le k \le u\, n^2/\log n} S_k < n\Big) = \\
&= P\Big(\max_{\log n/n \le k \log n/n^2 \le u} \Big\{\frac{S_{v(k,n)}}{\sqrt{(n^2/\log n)\log n}}\Big\} < 1\Big) = \\
&= \Big(1 + o(1)\Big) P\Big(\max_{0 \le t \le u} W(t) < 1\Big) = \\
&= \Big(1 + o(1)\Big)\Big(1 - 2\,P\Big(W(u) \ge 1\Big)\Big) = \\
&= \Big(1 + o(1)\Big)\Big(1 - 2\,P\Big(W(1) \ge 1/\sqrt{u}\Big)\Big),
\end{aligned}
$$

where we used the change of variable $t = \frac{k \log n}{n^2}$ and Proposition A.3(c). ∎

Remark 7.2. With suitable constants, $c_1, c_2,\, d_1, d_2$, it follows from the proof that, for x large,

$$
\begin{Bmatrix} \sqrt{c_1/x^{3-4p}} \\ \sqrt{d_1/x} \end{Bmatrix} \le 1 - G_p(x) \le \begin{cases} \sqrt{c_2/x}, & \text{for} \quad 0 < p \le 1/2, \\ \sqrt{d_2/x^{3-4p}}, & \text{for} \quad 1/2 \le p < 3/4, \end{cases}
$$

by using that $t^{2p-1} W(t^{3-4p}) \xrightarrow{a.s.} 0$ as $t \to 0+$, and that $x^{2p-1} \le t^{2p-1} \le \delta^{2p-1}$ for $\delta \le t \le x$, e.g., for $0 < p \le 1/2$.

7.1.1. Moments of first passage times

We begin with some inequalities. For $n \geq 2$,

$$\begin{aligned}\sum_{k\geq n} k\,P(\tau_n^{(+)}=k) \quad &\geq \quad \sum_{k\geq n} k\,P\big(\tau_{n-1}^{(+)}=k-1\big)\times\\ &\times\frac{(k-1-(n-1))(1-p)/2+(n-1+k-1)p/2}{k-1}\geq\\ &\geq \sum_{k\geq n}(k-1)\,P\big(\tau_{n-1}^{(+)}=k-1\big)\Big(\frac{(k-1-(n-1))}{2(k-1)}+\frac{n-1}{k-1}p\Big)\geq\\ &\geq \sum_{k\geq n}P\big(\tau_{n-1}^{(+)}=k-1\big)\Big(\frac{1}{2}(k-1)+(n-1)(p-1/2)\Big),\end{aligned}$$

and

$$\sum_{k\geq 0}(2k+1)\,P\big(\tau_1^{(+)}=2k+1\big)\geq p+\sum_{k\geq 1}(2k+1)\,P\big(\tau_0=2k\big)\,\frac{k(1-p)+kp}{2k},$$

which, together yield the following

Lemma 7.1. *For $0<p<1$,*

$$\begin{aligned}E(\tau_1^{(+)}) \quad &\geq \quad p+\frac{1}{2}E(\tau_0) \qquad \textit{and}\\ E(\tau_n^{(+)}) \quad &\geq \quad \frac{1}{2}E(\tau_{n-1}^{(+)})+(n-1)(p-\frac{1}{2})P(\tau_{n-1}^{(+)}<\infty) \qquad \textit{for} \quad n\geq 2.\end{aligned}$$

From the previous subsection we know that $E(\tau_0)=\infty$ for $1/4<p<3/4$, which immediately tells us that

Theorem 7.5. *For $1/4\leq p<3/4$ we have $E(\tau_n^{(+)})=\infty$ for all $n\geq 1$.*

Remark 7.3. We guess that $E(\tau_n^{(+)})<\infty$ for $0<p<1/4$.

7.2. Restricted memories

Theorem 7.6. *Suppose that $\mathfrak{M}_n=\{1\}$.*
(i) *As $n\to\infty$,*

$$\frac{\tau_n^{(+)}}{n}\overset{a.s.}{\to}\begin{cases}\frac{1}{2p-1}, & \textit{for} \quad p>1/2 \textit{ and } X_1=1,\\ \infty, & \textit{for} \quad p>1/2 \textit{ and } X_1=-1,\\ \frac{1}{1-2p}, & \textit{for} \quad p<1/2 \textit{ and } X_1=-1,\\ \infty, & \textit{for} \quad p<1/2 \textit{ and } X_1=1.\end{cases}$$

(ii) *If* $p > 1/2$ *and* $X_1 = 1$, *then*

$$\frac{\tau_n^{(+)} - n/(2p-1)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_{0,\frac{p(1-p)}{(2p-1)^3}} \qquad \text{as} \quad n \to \infty.$$

If $p < 1/2$ *and* $X_1 = -1$, *then*

$$\frac{\tau_n^{(+)} - n/(1-2p)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_{0,\frac{p(1-p)}{(1-2p)^3}} \qquad \text{as} \quad n \to \infty.$$

(iii) *If* $p = 1/2$, *then*

$$\frac{\tau_n^{(+)}}{n^2} \xrightarrow{d} S(1/2, 1/2) \qquad \text{as} \quad n \to \infty.$$

Proof. (i) and (ii): If $p > 1/2$ and $X_1 = 1$, then $S_n = 1 + T_{n-1}$, where T_n is a sum of i.i.d. random variables $Y_k = \pm 1$ with probabilities p and $1-p$, respectively. If $p > 1/2$ and $X_1 = -1$, then $S_n = -1 + \tilde{T}_{n-1}$, and the trend is negative. The conclusions follow from results on stopped random walks, see e.g. [19], Chapter 3.

The other cases follow via analogous arguments.

(iii): For the case $p = 1/2$ we are faced with a simple symmetric random walk with well known results. ∎

Theorem 7.7. *Suppose that* $\mathfrak{M}_n = \{1, 2\}$, *and that* $p > 1/2$. *Then,*
(i)

$$\frac{\tau_n^{(+)}}{n} \xrightarrow{a.s.} \begin{cases} \frac{1}{2p-1}, & \text{for} \quad X_1 = X_2 = 1, \\ \infty, & \text{for} \quad X_1 = X_2 = -1. \end{cases} \qquad \text{as} \quad n \to \infty.$$

(ii) *If* $X_1 = X_2 = 1$, *then*

$$\frac{\tau_n^{(+)} - n/(2p-1)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_{0,\frac{p(1-p)}{(2p-1)^3}} \qquad \text{as} \quad n \to \infty.$$

If $X_1 \cdot X_2 = -1$, *then*

$$\frac{\tau_n^{(+)}}{n^2} \xrightarrow{d} S(1/2, 1/2) \qquad \text{as} \quad n \to \infty.$$

The proof is similar to that of Theorem 7.6.

Theorem 7.8. *If* $\mathfrak{M}_n = \{n\}$, *then* $\dfrac{p\,\tau_n^{(+)}}{(1-p)n^2} \xrightarrow{d} S(1/2, 1/2) \qquad \text{as} \quad n \to \infty.$

Proof. As in the proof of Theorem 7.4 we use the fact that a stationary recurrent, hence ergodic, Markov chain follows a weak invariance principle with a limiting Wiener process. ■

Theorem 7.9. *For $0 < p < 1$, τ_0 is a proper random variable with infintely many returns to zero.*

Proof. Apply the LIL for ergodic Markov chains with a finite state space. ■

Remark 7.4. We conjecture that $E(\tau_0) = \infty$, and, hence, that the ERW is nullrecurrent.

Theorem 7.10. *Let $\mathfrak{M} = \{1, n\}$ and set $\sigma_T^2 = 1 + \frac{2p-1}{(3-2p)^3}(4p^2 - 24p + 19)$ (cf. [22]).*

(i) *As $n \to \infty$,*

$$\frac{\tau_n^{(+)}}{n} \overset{a.s.}{\to} \begin{cases} \frac{3-2p}{2p-1,} & for \quad p > 1/2 \text{ and } X_1 = 1, \\ \infty, & for \quad p > 1/2 \text{ and } X_1 = -1, \\ \frac{3-2p}{1-2p}, & for \quad p < 1/2 \text{ and } X_1 = -1, \\ \infty, & for \quad p < 1/2 \text{ and } X_1 = 1. \end{cases}$$

(ii) *For $p > 1/2$ and $X_1 = 1$,*

$$\frac{\tau_n^{(+)} - n(3-2p)/(2p-1)}{\sigma_T\sqrt{n((3-2p)/(2p-1))^3}} \overset{d}{\to} \mathcal{N}_{0,1} \qquad as \quad n \to \infty.$$

(iii) *For $p < 1/2$ and $X_1 = -1$,*

$$\frac{\tau_n^{(+)} - n(3-2p)/(1-2p)}{\sigma_T\sqrt{n((3-2p)/(1-2p))^3}} \overset{d}{\to} \mathcal{N}_{0,1} \qquad as \quad n \to \infty.$$

(iv) *If $p = 1/2$, then*

$$\frac{\tau_n^{(+)}}{n^2} \overset{d}{\to} S(1/2, 1/2) \qquad as \quad n \to \infty.$$

Proof. (i): If $p > 1/2$ and $X_1 = 1$ then $\{X_n, n \geq 1\}$ is a stationary recurrent Markov chain with $S_n/n \overset{a.s.}{\to} (2p-1)/(3-2p)$ as $n \to \infty$ (see e.g. [22]) which yields the first line. If $p < 1/2$ and $X_1 = -1$ then $E(S_n) \sim n(1-2p)/(3-2p)$ and we obtain the third line. In the other two cases the result follows from the LIL and the negative trend.

(ii): Using the Skorohod representation of the CLT (see e.g. [20], Section 5.13) $\big(S_n - n(2p-1)/(3-2p)\big)/\big(\sigma_T\sqrt{n}\big) \overset{d}{\to} \mathcal{N}_{0,1}$ as $n \to \infty$, we proceed as before and obtain

$$\big(S_{\tau_n^{(+)}} - \tau_n^{(+)}(2p-1)/(3-2p)\big)/\big(\sigma_T\sqrt{\tau_n^{(+)}}\big) \overset{d}{\to} \mathcal{N}_{0,1} \qquad \text{as} \quad n \to \infty,$$

which, together with part (i), and the symmetry of the normal distribution, provides the result.

(iii): As in (ii).

(iv): Again we are in the situation of a classical symmetric random walk. ∎

8. Final remarks and suggestions

(i) We have described the classical elephant random walk as well as walks with various restricted memories, in particular when $\mathfrak{M}_n = \{1, 2, \ldots, m\}$ for some fixed m. Details are provided for $m = 1$ and $m = 2$. For larger sizes, see Ben-Ari et al., [2]. Other variations are $\mathfrak{M}_n = \{m, m+1, \ldots, n\}$ or a combination of them. Details are here given for the cases $\mathfrak{M}_n = \{n\}$ and $\mathfrak{M}_n = \{1, n\}$. We have also discussed the situation if the elephant remembers an increasing number of early steps. Related to this is the work of Laulin [29], in which the elephant has a decreasing memory in the distant past, i.e., has a kind of amnesia. It would be interesting to know what happens if the elephant remembers an increasing number of late steps, as well as in both ends. This would *i.a.* be of interest with respect to phase changes (which we have only encountered in the classical case).

(ii) The different initial steps (+1, −1, or 0) cause multiple branches. Asymptotic normality therefore cannot hold. However, with a random centering the following result from [22] holds for $\mathfrak{M}_n = \{1, 2\}$:

$$\frac{S_n - n(2p-1)\,(X_1 + X_2)/2}{\sqrt{n}} \xrightarrow{d} p \cdot \mathcal{N}_{0,4p(1-p)} + (1-p) \cdot \mathcal{N}_{0,1} \quad \text{as} \quad n \to \infty.$$

(iii) The classical elephant random walk has a connection to Pólya urn models. Namely, let $U(n) = \big(B(n), R(n)\big)^T$, where $B(n), R(n)$ denote the number of black and red balls at time n, respectively, and T denotes transpose. Now, suppose that

$$\mathbf{U}(n+1) = \mathbf{U}(n) + \mathbf{A}(n)\,(1, 0)^T \quad \text{with a random matrix} \quad \mathbf{A}(n),$$

where

$$\mathbf{A}(n) = \begin{cases} \mathbf{I}_2, & \text{with probability } p, \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & \text{with probability } 1-p. \end{cases}$$

Then $S_n = \mathbf{U}(n)\,(1, -1)^T$ and one can transfer results. For more see [1].

(iv) Multivariate elephant random walks have also been discussed in the literature, see e.g., [6, 8, 11].

(v) Another topic that might be of interest is the gambler's ruin problem for elephants, with as well as without delays. (For the classical problem with delays, cf. [21]).

(vi) Another angle is the statistical one, e.g., how to estimate the unknown parameter p from the history of the process, cf., Bercu and Laulin [7].

(vii) When preparing for the next step the elephant has throughout chosen uniformly among then previous ones. It would be interesting to find out what happens if different steps have different weights, for example if more emphasis is given on recent times (or past times).

(viii) There exist other models related to ours. Engländer and Volkov, [17], is devoted to a variation in that the next step is not generated by flipping a coin, rather by turning it over or not. They have a somewhat different focus, in particular, they allow for different p-values in each step.

In addition, there is a large literature dealing with so-called correlated random walks, although with different aims. Chen and Renshaw [12] investigate a walk in dimension d and the probability of returns. Menshikov and Volkov, [30], consider continuous time processes generalizing the ERW and questions of transience and recurrence, and Comets et al. [15] study a kind of self-avoiding walk in $\mathbb{R}^d$.

A. Appendix

Here we collect some auxiliary material from probability and analysis.

A.1. A martingale lemma

Lemma A.1. *Let* $\{U_n, n \geq 1\}$ *be a sequence of random variables adapted to* $\mathcal{F}_n$, $n \geq 1$, *with*

$$E(U_{n+1} \mid \mathcal{F}_n) = a_n U_n + b_n \quad \text{for} \quad n \geq 1,$$

with two squences $\{a_n\}$ *and* $\{b_n\}$, $n \geq 1$, *where* $a_n \neq 0$ *for all* n. *Then*

$$\{(M_n = \alpha_n U_n + \beta_n, \mathcal{F}_n), n \geq 1\} \quad \text{is a martingale,}$$

where $\alpha_1 = 1$, $\beta_1 = 0$, *and*

$$\alpha_n = \prod_{k=1}^{n-1} \frac{1}{a_k} \quad \text{and} \quad \beta_n = -\sum_{k=1}^{n-1} \alpha_{k+1} b_k \quad \text{for} \quad n \geq 2\,.$$

The proof amounts to checking that the martingale condition is satisfied.

A.2. Asymptotics for the Gamma-function ([23], Lemma 2.2)

Lemma A.2. *For $x \in \mathbb{R}$,*

$$\frac{\Gamma(n+1+x)}{\Gamma(n+1)} = n^x\Big(1+\frac{x(1+x)}{2n}+\mathcal{O}(n^{-2})\Big) \quad \text{as} \quad n\to\infty.$$

A.3. Disturbed limit distributions

The following result is a special case of the Cramér–Slutsky theorem.

Proposition A.1. *Let $\{U_n,\, n \geq 1\}$ be a sequence of random variables, and suppose that V is independent of all of them. If $U_n \stackrel{d}{\to} U$ as $n\to\infty$, then $U_nV \stackrel{d}{\to} UV$ as $n\to\infty$.*

Proof. Using characteristic functions and bounded convergence we have, as $n\to\infty$,

$$\begin{aligned}\varphi_{U_nV}(t) &= E\exp\{itU_nV\} = E\Big(E(\exp\{itU_nV\}\mid V)\Big) = E\varphi_{U_n}(tV)\\ &\to E\varphi_U(tV) = E\Big(E(\exp\{itUV\}\mid V)\Big) = E\exp\{itUV\}) = \varphi_{UV}(t).\end{aligned}$$

An application of the continuity theorem for characteristic functions finishes the proof. ∎

A.4. Difference equations

In the proofs we use several difference equations. For easy reference we summarize some well-known facts about linear difference equations (consult e.g. [27]).

Proposition A.2. (i) *Consider the first order equation*

$$x_{n+1} = a\,x_n + b_n, \quad \text{for} \quad n\geq 1, \quad \text{with} \quad x_1 = x_1^* \quad \text{as initial value.}$$

Then

$$x_n = a^{n-1}x_1^* + \sum_{j=0}^{n-2} a^j b_{n-1-j}.$$

If, in addition, $|a|<1$ and $b_n = bn^\gamma$ with $\gamma > -1$, then

$$x_n = \frac{b_{n-1}}{1-a} - \frac{\gamma a b_{n-1}}{n(1-a)^2}\big(1+o(1)\big) \quad \text{as} \quad n\to\infty.$$

(ii) *If, in particular, $|a|<1$ and $x_{n+1} = ax_n + b$, then*

$$x_n = \frac{b}{1-a} + a^{n-1}\Big(x_1^* - \frac{b}{1-a}\Big) = \frac{b}{1-a}\big(1+o(1)\big) \quad \text{as} \quad n\to\infty.$$

(iii) *Consider the homogeneous second order equation*

$$x_{n+1} = a\,x_n + b\,x_{n-1}, \quad \textit{for} \quad n \geq 2, \quad \textit{with} \quad x_1^*,\, x_2^* \quad \textit{given.}$$

Then, with $\lambda_{1/2} = (a \pm \sqrt{a^2+4b})/2$, *provided* $a^2+4b \neq 0$,

$$x_n^h = c_1\lambda_1^n + c_2\lambda_2^n \quad \textit{with} \quad c_1,\, c_2 \quad \textit{chosen such that} \quad x_i^h = x_i^* \quad \textit{for} \quad i = 1, 2.$$

(iv) *As for the inhomogeneous second order equation*

$$x_{n+1} = a\,x_n + b\,x_{n-1} + d_n, \quad \textit{for} \quad n \geq 2, \quad \textit{with} \quad x_1^*,\, x_2^* \quad \textit{given,}$$

we have $x_n = x_n^h + y_n$, *where* y_n *is some solution of the inhomogeneous equation, where the constants* c_1, c_2 *in* x_n^h *are chosen properly. If* $d_n \equiv d$ *and* $a+b \neq 1$ *we may choose* $y_n = d/(1-a-b)$.

A.5. Asymptotics

Proposition A.3. *Suppose that* S_n, $n \geq 1$, *is a classical* ERW, *and let* L *be as described in, e.g.,* [3].

(a) *For* $0 < p < 3/4$ *we have* $S_n/\sqrt{n} \xrightarrow{d} \mathcal{N}_{0,1/(3-4p)}$, *hence,*

$$\frac{S_n^2(3-4p)}{n} \xrightarrow{d} \chi_1^2 \quad \textit{as} \quad n \to \infty,$$

whereas, for $p = 3/4$,

$$\frac{S_n^2}{n\log n} \xrightarrow{d} \chi_1^2 \quad \textit{as} \quad n \to \infty,$$

and, for $3/4 < p < 1$,

$$\frac{S_n}{n^{2p-1}} \xrightarrow{a.s.} L \quad \textit{as} \quad n \to \infty.$$

(b) *If* $Z \in \chi_1^2$, *then* $1/Z \in S(1/2, 1/2)$, *and if* $U_n \xrightarrow{d} \chi_1^2$ *as* $n \to \infty$ *for a sequence of positive random variables* $\{U_n,\, n \geq 1\}$, *then* $1/U_n \xrightarrow{d} S(1/2,1/2)$ *as* $n \to \infty$.

(c) *Moreover, the following invariance principles, due to* [1], *hold on the Sko-*

rohod space $(D[0,\infty), \mathcal{D}[0,\infty))$ (see e.g. [10], Section 16) *as* $n \to \infty$*:*

$$\begin{aligned}
\frac{S_{[nt]}}{\sqrt{n}} &\overset{d}{\to} \frac{t^{2p-1}}{\sqrt{3-4p}} W(t^{3-4p}), \qquad \textit{for} \quad 0 < p < 3/4, \\
&\qquad \textit{i.e., the process oscillates strongly,} \\
&\qquad \textit{although with slower speed as } p \textit{ increases;} \\
\frac{S_{[n^t]}}{\sqrt{n^t \log n}} &\overset{d}{\to} W(t), \qquad \textit{for} \quad p = 3/4; \\
\Big\{\frac{S_{[nt]}}{\sqrt{n}},\, t \geq 0\Big\} &\overset{a.s.}{\to} \big\{t^{2p-1} L,\, t \geq 0\big\}, \qquad \textit{for} \quad 3/4 < p < 1.
\end{aligned}$$

Acknowledgement. We wish to thank a referee for a most careful reading of our manuscript.

References

[1] **Baur, E. and J. Bertoin,** Elephant random walks and their connection to Pólya-type urns, *Phys. Rev. E*, **94** (2016), 052134.

[2] **Ben-Ari, I., J. Green, T. Meredith, H. Panzo and X. Tan,** Finite-memory elephant random walk and the central limit theorem for additive functionals, *Brazil. J. Probab. Statist.*, **35** (2021), 242–262.

[3] **Bercu, B.,** A martingale approach for the elephant random walk, *J. Phys. A: Math. Theor.*, **51** (2018), 015201.

[4] **Bercu, B.,** On the elephant random walk with stops playing hide and seek with the Mittag–Leffler distribution, *J. Statistical Physics*, **189** (2022), 1–27.

[5] **Bercu, B., M.-L. Chabanol and J.-J. Ruch,** Hypergeometric identities arising from the elephant random walk, *JMAA*, **480** (2019), 123360.

[6] **Bercu, B. and L. Laulin,** On the multi-dimensional elephant random walk, *Journal of Statistical Physic*, **175** (2019), 1146–1163.

[7] **Bercu, B. and L. Laulin,** How to estimate the memory of the elephant random walk, *Communications in Statistics, Theory and Methods*, (2022), 1–22.

[8] **Bertenghi, M.,** Functional limit theorems for the multi-dimensional elephant random walk, *Stochastic Models*, **38** (2022), 37–50.

[9] **Bertoin, J.,** Counting the zeros of an elephant random walk, *Trans Amer. Math. Soc.*, **375** (2022), 5539–5560.

[10] **Billingsley, P.,** *Convergence of Probability Measures*, 2nd edn, Wiley, New York, 1999.

[11] **Chen, J. and L. Laulin,** Analysis of smoothly amnesia-reinforced multidimensional elephant random walk, arXiv preprint (2023) https://arxiv.org/abs/2301.08644

[12] **Chen, A. and E. Renshaw,** The Gillis-Domb-Fisher correlated random walk, *J. Appl. Probab.*, **29** (1992), 792–813.

[13] **Coletti, C.F., R. Gava and G.M. Schütz,** Central limit theorem and related results for the elephant random walk, *J. Math. Phys.*, **58(5)** (2017), 053303, 8.

[14] **Coletti, C.F. and I. Papageorgiou,** Asymptotic analysis of the elephant random walk, *J. Statist. Mechanics, Theory Methods*, (2021), 013205.

[15] **Comets, F., M.V. Menshikov and A.R. Wade,** Random walks avoiding their convex hull with a finite memory, *Indigationes Mathematicae*, **31** (2020), 117–140.

[16] **Cressoni, J.C., M.A.A. da Silva and G.M. Viswanathan,** Amnestically induced persistence in random walks, *J. Phys. A.: Math. Theor.*, **46** (2007), 505002.

[17] **Engländer, J. and S. Volkov,** Turning a coin over instead of tossing it, *J. Theor. Probab.* **31** (2018), 1097–1118.

[18] **González-Navarette, M. and R. Lambert,** Non-Markovian random walks with memory lapses, *J. Math. Phys.*, **59** (2018), 113301.

[19] **Gut, A.,** *Stopped Random Walks*, 2nd edn, Springer-Verlag, New York, 2009.

[20] **Gut, A.,** *Probability: A Graduate Course*, 2nd edn, Springer-Verlag, New York, 2013.

[21] **Gut, A.,** The gambler's ruin problem with delays, *Stat. & Prob. Letters*, **83** (2013), 2549–2552.

[22] **Gut, A. and U. Stadtmüller,** Variations of the elephant random walk, *J. Appl. Probab.*, **58** (2021), 805–829.

[23] **Gut, A. and U. Stadtmüller,** The number of zeros in Elephant random walks with delays, *Stat. & Prob. Letters* **174** (2021), 109112.

[24] **Gut, A. and U. Stadtmüller,** Elephant random walks with delays, *Revue Roumaine de Mathématiques Pures et Appliquées*, Tome LXVII, No. 1-2 (2022).

[25] **Gut, A. and U. Stadtmüller,** Elephant random walks with increasing memory, *Stat. & Prob. Letters*, **189** (2022), 109598.

[26] **Gut, A. and U. Stadtmüller,** Bernoulli elephant random walks, Preprint, (2022).

[27] **Kelley, W.-G. and A.C. Peterson,** *Difference Equations: An Introduction with Applications*, Academic Press, 2001.

[28] **Heyde, C.C.,** On the Central Limit Theorem and Iterated Logarithm Supplements to the Martingale Convergence Theorem, *J. Appl. Probab.*, **14** (1977), 758–775.

[29] **Laulin, L.,** Introducing smooth amnesia to the memory of the elephant random walk, *Electron. Commun. Prob.*, **27** (2022), 1–12.

[30] **Menshikov, M. and S. Volkov,** Urn-related random walk with drift $\rho x^{\alpha}/t^{\beta}$, *Electron. J. Probab.*, **13** (2008), 944–960.

[31] **Moura, Th.R.S., G.M. Viswanathan, M.A.A. daSilva, J.C. Cressoni and L.R. daSilva,** Transient superdiffusion in random walks with a q-exponentially decaying memory profile, *Physica A.: Statist. Mech. Appl.*, **453** (2016), 259–263.

[32] **Renshaw, E. and R. Henderson,** The Correlated Random Walk, *J. Appl. Probab.*, **18**, 403–414.

[33] **Schütz, G.M. and S. Trimper,** Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk, *Phys. Rev. E*, **70** (2004), 045101.

[34] **da Silva, M.A.A., J.C. Cressoni and G.M. Viswanathan,** Discrete-time non-Markovian random walks: The effect of memory limitations on scaling, *Physica A*, **364** (2006), 70–78.

A. Gut
Uppsala University
Department of Mathematics
Uppsala
Sweden
allan.gut@math.uu.se

U. Stadtmüller
Ulm University
Department of Mathematics
Ulm
Germany
ulrich.stadtmueller@uni-ulm.de