Rate of escape and central limit theorem for the supercritical Lamperti problem

Mikhail V. Menshikov[a], Andrew R. Wade[b],*

[a] *Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK*
[b] *Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, UK*

Abstract

The study of discrete-time stochastic processes on the half-line with mean drift at x given by $\mu_1(x) \to 0$ as $x \to \infty$ is known as *Lamperti's problem*. We give sharp almost-sure bounds for processes of this type in the case where $\mu_1(x)$ is of order $x^{-\beta}$ for some $\beta \in (0, 1)$. The bounds are of order $t^{1/(1+\beta)}$, so the process is super-diffusive but sub-ballistic (has zero speed). We make minimal assumptions on the moments of the increments of the process (finiteness of $(2 + 2\beta + \varepsilon)$-moments for our main results, so fourth moments certainly suffice) and do not assume that the process is time-homogeneous or Markovian. In the case where $x^\beta \mu_1(x)$ has a finite positive limit, our results imply a strong law of large numbers, which strengthens and generalizes earlier results of Lamperti and Voit. We prove an accompanying central limit theorem, which appears to be new even in the case of a nearest-neighbour random walk, although our result is considerably more general. This answers a question of Lamperti. We also prove transience of the process under weaker conditions than those that we have previously seen in the literature. Most of our results also cover the case where $\beta = 0$. We illustrate our results with applications to birth-and-death chains and to multi-dimensional non-homogeneous random walks.

* Corresponding author. Tel.: +44 0141 548 3663; fax: +44 0141 548 3345.
  E-mail addresses: mikhail.menshikov@durham.ac.uk (M.V. Menshikov), andrew.wade@strath.ac.uk (A.R. Wade).

1. Introduction

In a pioneering series of papers [14–16] published in the early 1960s, Lamperti systematically studied how the asymptotic behaviour of a non-negative real-valued discrete-time stochastic process with asymptotically zero drift is governed by the (first two) moment functions of its increments. In the last two decades, there has been renewed interest in Lamperti's problem and in particular in its applications to studying the behaviour of complicated multi-dimensional processes (see e.g. [8,18]). A special case of Lamperti's problem supported on $\mathbb{Z}^+ := \{0, 1, 2, \ldots\}$ is that of asymptotically-zero-drift birth-and-death chains, for which exact calculations are often possible (using for instance Karlin–McGregor theory [4,12,13]); although classically well-studied, there has been recent renewed interest in such birth-and-death chains (see e.g. [5,6]), particularly in the context of modelling random polymers (see e.g. [3,9]). The study of continuous-time analogues of the general Lamperti problem seems to have begun only recently: see e.g. [7].

Let us describe Lamperti's problem informally. Consider a stochastic process $X = (X_t)_{t \in \mathbb{Z}^+}$ on $[0, \infty)$. For now, suppose that X is a time-homogeneous Markov process (that is, a Markov process with stationary transition probabilities) and that its increment moment functions

$$\mu_k(x) = \mathbf{E}[(X_{t+1} - X_t)^k \mid X_t = x] \tag{1.1}$$

are well defined for $k \geq 0$; one way to ensure this is to impose a uniform bound on the increments. (We will relax all of these conditions shortly.) Lamperti's problem is to determine how the asymptotic behaviour of X depends upon μ_1 and μ_2.

Under mild regularity conditions, the behaviour of X is rather standard when, outside some bounded set, $\mu_1(x) \equiv 0$ (the zero-drift case) or $\mu_1(x)$ is uniformly bounded to one side of 0. Roughly speaking, in the zero-drift case X behaves like a simple symmetric random walk and is null-recurrent, in the case of uniformly negative drift X is positive-recurrent with exponentially decaying stationary distribution, and in the case of uniformly positive drift X is transient with positive speed (i.e., ballistic).

This motivates the study of the *asymptotically-zero-drift* regime, in which $\mu_1(x) \to 0$ as $x \to \infty$, to investigate phase transitions. It turns out that there is a rich spectrum of possible behaviours of X, governed by μ_1 and μ_2; we mention heavy-tailed positive-recurrence, transience with sub-linear rate of escape (diffusive and super-diffusive motion both being possible), weak convergence to a Bessel process, and so on.

Results of Lamperti [14,16] imply that from the point of view of the recurrence classification of X, the case where $|\mu_1(x)|$ is of order x^{-1} and $\mu_2(x)$ is of order 1 is critical. In the present paper we are interested in the *supercritical* case where $\mu_1(x)$ is positive and of order $x^{-\beta}$, $\beta \in (0, 1)$. Here, under mild conditions, transience is assured: our primary interest is to quantify this transience by studying the *rate of escape* and accompanying second-order behaviour.

As well as being of interest in their own right, stochastic processes on the half-line with mean drift asymptotically zero are important for the study of multi-dimensional processes by the method of Lyapunov-type functions (see e.g. [8]). In this context, it is particularly desirable to work in some generality without imposing, for instance, assumptions of the Markov property, a countable state-space, or uniformly bounded increments. Thus we work in more generality than the model outlined informally above. To start with, the assumption on uniformly bounded increments can be relaxed, and replaced by an appropriate moments condition. Another important

relaxation (building on the ideas in Lamperti's first paper on the topic [14]) is that we do not need X to be a Markov process. It is invaluable with regard to applications to be able to dispense with the Markov assumption. The prototypical illustration of this latter point is provided by the case where X is given by $X_t = \|Y_t\|$, the norm of some multi-dimensional (perhaps Markov but not necessarily spatially homogeneous) process. If Y_t has mean drift zero, X_t will typically have $\mu_1(x) \to 0$ as $x \to \infty$.

Relaxing the Markov assumption leads to a slight complication in defining the correct analogues of (1.1), but does not complicate our proofs which are based on general martingale arguments. The process X will be taken to be adapted to some filtration $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$. Important families of processes that fit into our framework include non-Markov processes where $\mathcal{F}_t = \sigma(X_0, X_1, \ldots, X_t)$ and the law of X_{t+1} depends on the entire previous history of the process, as well as processes where X_t is not Markov by itself, but $X_t = f(Y_t)$ for some Markov process Y_t on a general space Σ, a measurable function $f : \Sigma \to [0, \infty)$, and $\mathcal{F}_t = \sigma(Y_0, Y_1, \ldots, Y_t)$. The first of these two situations was treated by Lamperti in [14, Section 3], and the second in [14, Section 4] (see also [16, Section 5]); we work somewhat more generally.

In the next section, we will describe more precisely the model that we consider and give our main results. In Section 3, we give two applications of our results to stochastic processes of interest in their own right. The first is the birth-and-death chain case; even in this classical setting, some of our results seem to be new. Our second example is a model inaccessible to many classical methods: a multi-dimensional non-homogeneous random walk. In the latter setting, our results add to the analysis of MacPhee et al. [17].

2. Model, results, and discussion

2.1. The model and main results

We now introduce our notation and assumptions. Let $X = (X_t)_{t \in \mathbb{Z}^+}$ be a discrete-time stochastic process adapted to a filtration $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$ and taking values in an unbounded subset $\mathcal{S}$ of $[0, \infty)$. In applications $\mathcal{S}$ may be countable (e.g. the birth-and-death chain example in Section 3.1) or uncountable (e.g. the non-homogeneous random walk example in Section 3.2, or the application to stochastic billiards in [18]); it is thus desirable to make no further restriction on $\mathcal{S}$.

The central object in all that follows will be the conditional mean increment (the one-step mean drift) $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t]$. Many of the conditions in our theorems will suppose that an inequality holds involving the $\mathcal{F}_t$-measurable random variables $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t]$ and X_t; such inequalities will have to hold a.s. and in an appropriate asymptotic sense (as $X_t \to \infty$). It will be convenient therefore to introduce some notation for upper and lower bounds on the mean increment $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t]$ as functions of X_t.

Shortly, we will briefly define $\underline{\mu}_1 : \mathcal{S} \to \mathbb{R}$ and $\overline{\mu}_1 : \mathcal{S} \to \mathbb{R}$ such that, for all $t \in \mathbb{Z}^+$,

$$\underline{\mu}_1(X_t) \le \mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] \le \overline{\mu}_1(X_t), \quad \text{a.s.} \tag{2.1}$$

If X is a Markov process, $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \mathbf{E}[X_{t+1} - X_t \mid X_t]$, a.s., and we can take

$$\underline{\mu}_1(x) = \inf_{t \in \mathbb{Z}^+} \mathbf{E}[X_{t+1} - X_t \mid X_t = x], \quad \text{and} \quad \overline{\mu}_1(x) = \sup_{t \in \mathbb{Z}^+} \mathbf{E}[X_{t+1} - X_t \mid X_t = x];$$

if additionally X is time-homogeneous then $\underline{\mu}_1(x) \equiv \overline{\mu}_1(x) \equiv \mu_1(x)$, where $\mu_k : \mathcal{S} \to \mathbb{R}$ is given by

$$\mu_k(x) = \mathbf{E}[(X_{t+1} - X_t)^k \mid X_t = x] \quad (t \in \mathbb{Z}^+), \tag{2.2}$$

provided the expectation exists. Loosely speaking, in the general case $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t]$ involves additional randomness in $\mathcal{F}_t$, once X_t has been fixed. Thus $\overline{\mu}_1(x)$ should be the (essential) supremum over this additional randomness given $\{X_t = x\}$. For $\underline{\mu}_1$, the situation is analogous.

Let us now formally define $\underline{\mu}_1$ and $\overline{\mu}_1$. Suppose that $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t]$ exists for all $t \in \mathbb{Z}^+$. By standard theory of conditional expectations (see e.g. [2, Section 9.1]), for each $t \in \mathbb{Z}^+$ there exist a Borel-measurable function $\phi_t : \mathcal{S} \to \mathbb{R}$ and an $\mathcal{F}_t$-measurable random variable ψ_t such that $\mathbf{E}[\psi_t \mid X_t] = 0$ and, a.s.,

$$\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \mathbf{E}[X_{t+1} - X_t \mid X_t] + \psi_t = \phi_t(X_t) + \psi_t. \tag{2.3}$$

Set $\mu_1(t; x) := \phi_t(x) + \psi_t$, an $\mathcal{F}_t$-measurable random variable. Then, for $x \in \mathcal{S}$, define

$$\overline{\mu}_1(x) := \sup_{t \in \mathbb{Z}^+} \operatorname{ess\,sup} \mu_1(t; x), \tag{2.4}$$

$$\underline{\mu}_1(x) := \inf_{t \in \mathbb{Z}^+} \operatorname{ess\,inf} \mu_1(t; x). \tag{2.5}$$

Provided the expectations in question exist, $\overline{\mu}_1(x), \underline{\mu}_1(x)$ are (non-random) $\mathbb{R}$-valued functions of $x \in \mathcal{S}$; clearly $\overline{\mu}_1(x) \geq \underline{\mu}_1(x)$ for all $x \in \mathcal{S}$. Then (2.4) and (2.5) define functions with the property (2.1). We provide some further discussion of the definitions in (2.4) and (2.5), and give some illustrative examples, in Section 2.2 below.

In the time-homogeneous Markov case, the statement of our results is simplified, and $\overline{\mu}_1, \underline{\mu}_1$ can be replaced simply by μ_1 defined by (2.2) everywhere. One such example, which might also be useful for orientation purposes, is the birth-and-death chain example described in Section 3.1 below. As mentioned above, in applications, it can be important to dispense with the Markovian assumption. It often turns out to be the case in applications that, as $x \to \infty$, $\overline{\mu}_1(x) \sim \underline{\mu}_1(x)$; the example in Section 3.2 below demonstrates such a case, and also the importance of not having to assume a Markov property for X.

Returning to the general setting, for our purposes the most interesting case is when $\overline{\mu}_1(x), \underline{\mu}_1(x) \to 0$ as $x \to \infty$. Results of Lamperti [14,16] show that from the point of view of the recurrence classification of X, the case where $\overline{\mu}_1(x), \underline{\mu}_1(x)$ are of order $1/x$ is critical (assuming some natural regularity conditions). Our focus in the present paper is the *supercritical* case where $\overline{\mu}_1(x), \underline{\mu}_1(x)$ are of order $x^{-\beta}$ (in the positive direction) for some $\beta \in (0, 1)$. In this case, Lamperti [14] proved that X is transient (that is, $X_t \to \infty$ a.s.) under certain regularity assumptions; we give a proof of this result under weaker conditions (Theorem 2.1). Our primary interest, however, is the nature of the transience, in particular the *rate of escape*, i.e., the speed at which $X_t \to \infty$. The results of this paper give sharp bounds of order $t^{1/(1+\beta)}$ for X_t (Theorem 2.3), which in the special case where $\overline{\mu}_1(x) \sim \underline{\mu}_1(x) \sim \rho x^{-\beta}$ imply a strong law of large numbers (Theorem 2.4) that improves upon results of Lamperti [15] and Voit [24]. We also study the second-order behaviour, obtaining a central limit theorem (Theorem 2.5) to accompany the law of large numbers. Although not our primary concern, most of our results also cover the case where $\beta = 0$.

Let us state our basic assumption.

(A0) Let $X = (X_t)_{t \in \mathbb{Z}^+}$ be a stochastic process on the unbounded set $\mathcal{S} \subseteq [0, \infty)$ adapted to the filtration $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$. Suppose that, for some $x_0 \in \mathcal{S}$, $\mathbf{P}[X_0 \le x_0 \mid \mathcal{F}_0] = 1$.

We also assume the following condition.

(A1) Suppose that $\limsup_{t \to \infty} X_t = \infty$ a.s.

Condition (A1) is necessary for our questions of interest to be non-trivial, and is usually straightforward to verify in a particular application: for instance, a sufficient condition is that for any $y \in (0, \infty)$ there exist $w : \mathbb{Z}^+ \to \mathbb{Z}^+$ and $\varepsilon > 0$ such that

$$\inf_{t \in \mathbb{Z}^+} \mathbf{P}[X_{t+w(t)} > y \mid \mathcal{F}_t] > \varepsilon, \quad \text{a.s.}$$

Indeed, if X is an irreducible time-homogeneous Markov chain and $\mathcal{S}$ is countable, (A1) holds automatically. For suitable concepts of irreducibility in more general state-spaces, see [20].

We also need to assume some regularity condition on the increments of X. For our purposes, we will need a moment bound of the form

$$\sup_{t \in \mathbb{Z}^+} \mathbf{E}[|X_{t+1} - X_t|^\gamma \mid \mathcal{F}_t] \le B, \quad \text{a.s.,} \tag{2.6}$$

for some $B < \infty$ and $\gamma > 0$. If (2.6) holds with $\gamma \ge 1$, $\overline{\mu}_1$ and $\underline{\mu}_1$ given by (2.4) and (2.5) exist as $\mathbb{R}$-valued functions. Assumption of (2.6) amounts to, in some sense, the choice of a correct scale for the process X.

Our first result yields transience of the supercritical Lamperti problem.

Theorem 2.1. *Suppose that* (A0) *and* (A1) *hold, and that there exists* $\beta \in [0, 1)$ *such that* (2.6) *holds for some* $\gamma > 1 + \beta$ *and*

$$\liminf_{x \to \infty}(x^\beta \underline{\mu}_1(x)) > 0.$$

Then X *is transient, i.e.,* $X_t \to \infty$ *a.s. as* $t \to \infty$.

Theorem 2.1 proves transience under weaker conditions than we have seen previously published; for instance, Lamperti [14, Theorem 3.2] (see also [20, Section 9.5.3]) assumed (2.6) with $\gamma > 2$ and also that $\mathbf{E}[(X_{t+1} - X_t)^2 \mid \mathcal{F}_t] \ge v$ a.s. for $v > 0$; Lamperti [14] was mainly concerned with the critical case ($\beta = 1$), where such stronger conditions are natural, but they are not necessary here, as Theorem 2.1 shows.

Next we move on to our main topic, the quantitative asymptotic behaviour of X. The first natural question is what bounds we can obtain under conditions of comparable strength to those in Theorem 2.1. We have the following upper bound.

Theorem 2.2. *Suppose that* (A0) *holds, there exists* $\beta \in [0, 1)$ *such that*

$$\limsup_{x \to \infty}(x^\beta \overline{\mu}_1(x)) < \infty,$$

and (2.6) *holds for some* $\gamma > 1 + \beta$. *Then, for any* $\varepsilon > 0$, *a.s., for all but finitely many t,*

$$\sup_{0 \le s \le t} X_s \le t^{\frac{1}{1+\beta}} (\log t)^{\frac{1}{1+\beta} + \varepsilon}. \tag{2.7}$$

Next we impose stronger conditions on X in order to obtain a tighter upper bound, as well as a complementary lower bound. Our bounds will involve the constants $\lambda(a, \beta)$ defined for $a \in (0, \infty)$, $\beta \in (0, 1)$ by

$$\lambda(a, \beta) := (a(1 + \beta))^{\frac{1}{1+\beta}}. \tag{2.8}$$

The next result gives sharp almost-sure bounds on X.

Theorem 2.3. *Suppose that* (A0) *and* (A1) *hold, and that, for some* $\beta \in [0, 1)$ *and some* a, $A \in (0, \infty)$ *with* $a \leq A$,

$$a = \liminf_{x \to \infty} (x^\beta \underline{\mu}_1(x)) \leq \limsup_{x \to \infty} (x^\beta \overline{\mu}_1(x)) = A. \tag{2.9}$$

Suppose that (2.6) *holds for some* $\gamma > 2 + 2\beta$. *Then, a.s.,*

$$\lambda(a, \beta) \leq \liminf_{t \to \infty} \frac{X_t}{t^{1/(1+\beta)}} \leq \limsup_{t \to \infty} \frac{X_t}{t^{1/(1+\beta)}} \leq \lambda(A, \beta).$$

Remark 1. The proof of the upper bound on X_t given by Theorem 2.3 only uses the condition on $\overline{\mu}_1$ in (2.9) and not the condition on $\underline{\mu}_1$ there.

Note that, since $\beta < 1$, certainly taking $\gamma = 4$ in (2.6) suffices for Theorem 2.3. Theorem 2.3 implies that in the case $\beta \in (0, 1)$ the transience given in Theorem 2.1 is *super-diffusive* but *sub-ballistic*, since $1/2 < 1/(1 + \beta) < 1$. This should be contrasted with the critically transient case ($\beta = 1$) where the drift is $O(x^{-1})$ and X is transient, in which case there are upper and lower bounds for X_t of order about $t^{1/2}$ known under additional conditions, see [18, Section 4.1], where for instance it is shown in [18, Theorem 4.2] that $X_t \geq t^{1/2}(\log t)^{-D}$ for some $D \in (0, \infty)$ and all but finitely many t (in the critically transient birth-and-death chain case, certain sharp bounds are a by-product of the invariance principle of [6]).

An immediate corollary of Theorem 2.3, obtained on taking $a = A = \rho$, is the following strong law of large numbers.

Theorem 2.4. *Suppose that* (A0) *and* (A1) *hold, and that, for some* $\beta \in [0, 1)$,

$$\lim_{x \to \infty} x^\beta \overline{\mu}_1(x) = \lim_{x \to \infty} x^\beta \underline{\mu}_1(x) = \rho \in (0, \infty). \tag{2.10}$$

Suppose that (2.6) *holds for some* $\gamma > 2 + 2\beta$. *Then, as* $t \to \infty$, *a.s.,*

$$\frac{X_t}{t^{1/(1+\beta)}} \longrightarrow \lambda(\rho, \beta). \tag{2.11}$$

Lamperti [15, Theorem 7.1] obtained a weaker version of Theorem 2.4 under more restrictive conditions. Specifically, [15, Theorem 7.1] assumes that X is a time-homogeneous Markov process with $\lim_{x \to \infty} x^\beta \mu_1(x) = \rho$ and $\sup_x |\mu_k(x)| < \infty$ for all k, where μ_k is given by (2.2). Then [15, Theorem 7.1] says that (2.11) holds *with convergence in probability*. Lamperti [15, p. 768] asks whether his result "can be strengthened to almost sure convergence"; Theorem 2.4 answers this affirmatively, and also shows that the assumptions in [15] can be relaxed to a significant extent. Theorem 2.4 also generalizes a result of Voit [24] in the birth-and-death chain case: see Section 3.1 below.

It is natural to ask whether, under the assumptions of Theorem 2.4, there is a central limit theorem to accompany the law of large numbers. This question was raised by Lamperti [15, p. 768], and seems to have remained open even for the case of a birth-and-death chain. The following result shows that there is a central limit theorem, provided that we impose a somewhat stronger version of (2.10) and an asymptotic stability condition on the second moments of the increments. Here and subsequently ' $\xrightarrow{d}$ ' denotes convergence in distribution. Unlike our preceding results, the case $\beta = 0$ is excluded from the following theorem.

Theorem 2.5. *Suppose that* (A0) *and* (A1) *hold, and that, for some $\beta \in (0, 1)$ and $\rho \in (0, \infty)$, as $x \to \infty$,*

$$\underline{\mu}_1(x) = \rho x^{-\beta} + o(x^{-\beta - \frac{1-\beta}{2}}); \qquad \overline{\mu}_1(x) = \rho x^{-\beta} + o(x^{-\beta - \frac{1-\beta}{2}}). \tag{2.12}$$

Suppose that (2.6) *holds for some $\gamma > 2 + 2\beta$, and that, for some $\sigma^2 \in (0, \infty)$,*

$$\mathbf{E}[(X_{t+1} - X_t)^2 \mid \mathcal{F}_t] \to \sigma^2, \quad \text{a.s., as } t \to \infty. \tag{2.13}$$

Then, as $t \to \infty$,

$$\frac{X_t - \lambda(\rho, \beta)t^{1/(1+\beta)}}{t^{1/2}} \xrightarrow{d} Z\sigma \sqrt{\frac{1+\beta}{1+3\beta}},$$

where Z is a standard normal random variable.

2.2. Further remarks on $\underline{\mu}_1$ and $\overline{\mu}_1$; examples

We now briefly discuss further the definitions in (2.4) and (2.5), and give some examples for particular classes of process X that should help to clarify the nature of the crucial functions $\underline{\mu}_1$ and $\overline{\mu}_1$. Recall that

$$\text{ess sup } \mu_1(t; x) = \inf\{z \in \mathbb{R} : \mathbf{P}[\mu_1(t; x) > z] = 0\},$$

with a similar expression for ess inf. Some intuitive feeling for the quantities $\underline{\mu}_1, \overline{\mu}_1$ is best gained by specializing our general framework to some particular families of processes.

Markov processes. If $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$ and X is Markov, we have that $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \mathbf{E}[X_{t+1} - X_t \mid X_t]$, a.s., so that, with the notation of (2.3),

$$\mu_1(t; x) = \phi_t(x) = \mathbf{E}[X_{t+1} - X_t \mid X_t = x], \quad \text{a.s.,}$$

for any t. When the state-space $\mathcal{S}$ is countable, this last quantity is simply expressed in terms of the one-step transition probabilities $\mathbf{P}[X_{t+1} = y \mid X_t = x]$. In the case of general $\mathcal{S}$, $\mu_1(t; x)$ can be expressed in terms of a corresponding Markov transition kernel. In either case, we then have that $\overline{\mu}_1(x) = \sup_{t \in \mathbb{Z}^+} \mathbf{E}[X_{t+1} - X_t \mid X_t = x]$, with a similar expression for $\underline{\mu}_1(x)$. If X is additionally time-homogeneous, $\mathbf{E}[X_{t+1} - X_t \mid X_t = x]$ does not depend on t, so $\overline{\mu}_1(x) \equiv \underline{\mu}_1(x)$.

History-dependent processes. Suppose, more generally, that $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$ and the law of X_{t+1} depends only upon $(X_0, \ldots, X_t)$. For convenience, take $\mathcal{S}$ to be countable. Then we can write

$$\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \sum_{x_0, \ldots, x_t \in \mathcal{S}} \mathbf{E}[X_{t+1} - X_t \mid X_0 = x_0, \ldots, X_t = x_t]$$

$$\times \mathbf{1}\{X_0 = x_0, \ldots, X_t = x_t\}.$$

This last expression can be written as $\mu_1(t; X_t)$, where $\mu_1(t; x)$ is given by

$$\sum_{x_0,\ldots,x_{t-1}\in\mathcal{S}} \mathbf{E}[X_{t+1} - X_t \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}, X_t = x]$$

$$\times \mathbf{1}\{X_0 = x_0, \ldots, X_{t-1} = x_{t-1}\}.$$

It follows that, in this case,

$$\overline{\mu}_1(x) = \sup_{t\in\mathbb{Z}^+} \sup_{\substack{x_0,\ldots,x_{t-1}\in\mathcal{S}: \\ \mathbf{P}[X_0=x_0,\ldots,X_{t-1}=x_{t-1}]>0}} \mathbf{E}[X_{t+1} - X_t \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}, X_t = x],$$

with an analogous expression for $\underline{\mu}_1$. In the case where $\mathcal{S}$ is uncountable, the expressions are similar but may be understood in terms of regular conditional distributions. This formulation is essentially used by Lamperti [14, p. 322].

Functions of Markov processes. Suppose that $(Y_t)_{t\in\mathbb{Z}^+}$ is a Markov process on some state-space Σ, and that, for some measurable function $f : \Sigma \to [0, \infty)$, $X_t = f(Y_t)$. Set $\mathcal{F}_t = \sigma(Y_0, \ldots, Y_t)$. Then X has state-space $\mathcal{S} = f(\Sigma)$, and X is typically non-Markovian; see e.g. [22] for a discussion on the latter point. Now $\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \mathbf{E}[X_{t+1} - X_t \mid Y_t]$, a.s., and if Σ (hence $\mathcal{S}$) is countable, we may write

$$\mathbf{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \sum_{x\in\mathcal{S}} \sum_{y\in\Sigma: f(y)=x} \mathbf{E}[X_{t+1} - X_t \mid Y_t = y]\mathbf{1}\{Y_t = y, X_t = x\}.$$

Expressing the latter quantity as $\mu_1(t; X_t)$ entails

$$\mu_1(t; x) = \sum_{y\in\Sigma: f(y)=x} \mathbf{E}[X_{t+1} - X_t \mid Y_t = y]\mathbf{1}\{Y_t = y\}.$$

It follows that, in this case,

$$\overline{\mu}_1(x) = \sup_{t\in\mathbb{Z}^+} \sup_{y\in\Sigma: f(y)=x,\, \mathbf{P}[Y_t=y]>0} \mathbf{E}[X_{t+1} - X_t \mid Y_t = y],$$

and similarly for $\underline{\mu}_1$. This situation often arises in applications, where f may be, for instance, a Lyapunov-type function applied to a multi-dimensional process. See the example in Section 3.2 below, as well as [14, Section 4] and [16, Section 5].

2.3. Open problems and paper outline

We finish this section by mentioning some possible directions for future work. A natural question is whether Theorem 2.3 holds under a weaker moments condition. Also of interest is whether any weak limit theory analogous to Theorem 2.5 is available when (2.10) holds but (2.12) does not.

In [19], an analogue of Lamperti's problem was considered for processes with $\mathbf{E}[X_{t+1} - X_t \mid X_t = x] \approx cx^\alpha t^{-\beta}$, loosely speaking. It seems likely that for appropriate α, β one could obtain results similar to ours in that setting.

The outline of the remainder of the paper is as follows. In Section 3, we discuss two applications of our main theorems, specifically to birth-and-death chains (nearest-neighbour random walks on $\mathbb{Z}^+$) in Section 3.1, and to non-homogeneous random walks in $\mathbb{R}^d$ in Section 3.2. Section 4 is devoted to the proofs of our theorems. In Section 4.1 we give a brief overview of our proofs. In Section 4.2, 4.3, 4.4 and 4.5, we prove Theorems 2.2, 2.3, 2.5 and 2.1 respectively;

finally, in Section 4.6, we prove our result (Theorem 3.2) on the non-homogeneous random walk presented in Section 3.2 below.

3. Applications

3.1. Birth-and-death chains

Suppose that X is an irreducible time-homogeneous Markov chain supported on the countable set $\mathcal{S} = \mathbb{Z}^+$ with jumps of size at most 1. Specifically, suppose that there exist sequences a_x, b_x, c_x ($x \in \mathbb{N} := \{1, 2, 3, \ldots\}$) with $a_x > 0$, $b_x \geq 0$, $c_x > 0$ and $a_x + b_x + c_x = 1$ for all $x \in \mathbb{N}$. Define the transition law of X for $t \in \mathbb{Z}^+$ as follows: for $x \in \mathbb{N}$,

$$\mathbf{P}[X_{t+1} = x + 1 \mid X_t = x] = a_x,$$
$$\mathbf{P}[X_{t+1} = x \mid X_t = x] = b_x,$$
$$\mathbf{P}[X_{t+1} = x - 1 \mid X_t = x] = c_x,$$

and with reflection from 0 governed by $\mathbf{P}[X_{t+1} = 1 \mid X_t = 0] = 1$. Of course in this setting X has uniformly bounded increments, so that (2.6) holds for all $\gamma > 0$, and is an irreducible time-homogeneous Markov chain on $\mathbb{Z}^+$, so that (A1) holds as well.

Such an X is known as a *birth-and-death chain* or birth-and-death random walk. Such processes have been extensively studied in various contexts, and are often amenable to explicit computation. Early contributions to the theory of such random walks, particularly to the recurrence/transience classification, are due to Harris [10] and Hodges and Rosenblatt [11]. Orthogonal polynomials provide one fruitful tool for analysis of such processes (see e.g. [4] for a survey); this approach dates back at least to Karlin and McGregor [12,13].

For $x \in \mathbb{N}$, with $\mu_1(x)$ and $\mu_2(x)$ defined by (2.2), we have that

$$\mu_1(x) = a_x - c_x, \qquad \mu_2(x) = 1 - b_x > 0.$$

The asymptotically-zero-drift case is the case where $\lim_{x \to \infty}(a_x - c_x) = 0$. There is an extensive literature concerned with various special cases where $|x\mu_1(x)| = O(1)$. For recent work, we refer to [3,5,6]; the papers of Csáki, Földes and Révész cited include references to some of the older literature. We are in the supercritical case if, for $\beta \in (0, 1)$,

$$\lim_{x \to \infty} x^\beta (a_x - c_x) = \rho \in (0, \infty). \tag{3.1}$$

In this case, the following law of large numbers is due to Voit [24, Theorem 2.11] (in fact Voit works in a more general setting of random walks on polynomial hypergroups, which do not concern us here). Note that there is a misprint in the limiting constant in the statement of Theorem 2.11 of [24] (the proof there does yield the correct constant): the $1/(1 + \alpha)$ power should be applied to the entire limiting expression, not just the μ there; this misprint persists into [5, Theorem D].

Proposition 3.1 (*[24, Theorem 2.11]*). *Suppose that X is a birth-and-death chain specified by a_x, b_x, c_x as described above. Suppose that (3.1) holds for $\beta \in (0, 1)$ and $\rho \in (0, \infty)$. Suppose also that the two limits*

$$\lim_{x \to \infty} a_x \quad \text{and} \quad \lim_{x \to \infty} c_x \text{ exist in } (0, 1) \tag{3.2}$$

(in which case they must take the same value). Then (2.11) holds.

Proposition 3.1 is a special case of our Theorem 2.4, under the additional assumption (3.2). Theorem 2.4 shows that the assumption (3.2) is not necessary for the result: only the mean is important, not the absolute probabilities of going left or right.

Our Theorem 2.5 above has the following immediate (and apparently new) consequence in the birth-and-death chain case. Under the assumption that $\lim_{x \to \infty} b_x = b$, (3.2) holds (with limit $\frac{1-b}{2}$ for a_x and c_x), so this central limit theorem can be seen as the natural companion to Voit's law of large numbers [24, Theorem 2.11].

Theorem 3.1. *Suppose that X is a birth-and-death chain specified by a_x, b_x, c_x as described above. Suppose that, for some $\beta \in (0, 1)$ and $\rho \in (0, \infty)$,*

$$a_x - c_x = \rho x^{-\beta} + o(x^{-\beta - \frac{1-\beta}{2}}); \qquad \lim_{x \to \infty} b_x = b \in [0, 1).$$

Then, as $t \to \infty$, for a standard normal random variable Z,

$$\frac{X_t - \lambda(\rho, \beta) t^{1/(1+\beta)}}{t^{1/2}} \xrightarrow{d} Z \sqrt{\frac{(1-b)(1+\beta)}{1+3\beta}}. \tag{3.3}$$

We make some final remarks on the case, of secondary interest to us here, where $\beta = 0$. Our Theorem 2.4 applies to the case $\beta = 0$, i.e., where $a_x - c_x \to \rho \in (0, 1]$ as $x \to \infty$, in which case our result says that $t^{-1} X_t \to \rho$ a.s. as $t \to \infty$. This particular result has been previously obtained by Pakes [21, Proposition 4], under some more restrictive conditions, including $\rho = 1$ and $b_x \equiv 0$, and also, for general ρ but again under conditions more restrictive than ours, in a result of Voit [23, Corollary 2.6]. When $\beta = 0$, the second-order behaviour of X is somewhat different (our Theorem 3.1 does not apply). See for instance [21, Theorem 7] and [23, Theorems 2.7–2.10].

3.2. Rate of escape for a non-homogeneous random walk on $\mathbb{R}^d$

In this section we illustrate the application of our results to a non-homogeneous random walk model similar to that of [17]. Fix $d \in \{2, 3, \ldots\}$. Let $\Xi = (\xi_t)_{t \in \mathbb{Z}^+}$ be a time-homogeneous Markov process with state-space an unbounded subset Σ of $\mathbb{R}^d$. The law of the increment $\xi_{t+1} - \xi_t$ then depends only on the position of ξ_t; this is formalized in general in terms of Markov transition kernels (see [20, Section 3.4]), so we may use the notation $\mathbf{P}[\, \cdot \mid \xi_t = \mathbf{x}]$ for the conditional distributions and $\mathbf{E}[\, \cdot \mid \xi_t = \mathbf{x}]$ for the corresponding expectations.

Write $\| \cdot \|$ for the Euclidean norm on $\mathbb{R}^d$ and $\mathbf{0}$ for the origin. We use the notation $\mathbf{x}$ for a point of $\mathbb{R}^d$, and, when $\mathbf{x} \neq \mathbf{0}$, $\hat{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$ for the corresponding unit vector. We use '·' to denote the usual scalar product on $\mathbb{R}^d$. We assume that there exist $\rho \in (0, \infty)$ and $\beta \in (0, 1)$ such that, for $\mathbf{x} \in \Sigma$,

$$\mathbf{E}[(\xi_{t+1} - \xi_t) \cdot \hat{\mathbf{x}} \mid \xi_t = \mathbf{x}] = \rho \|\mathbf{x}\|^{-\beta} + o(\|\mathbf{x}\|^{-\beta}), \tag{3.4}$$

as $\|\mathbf{x}\| \to \infty$. We will also assume a moment bound on the size of the jumps:

$$\sup_{\mathbf{x} \in \Sigma} \mathbf{E}[\|\xi_{t+1} - \xi_t\|^\gamma \mid \xi_t = \mathbf{x}] < \infty. \tag{3.5}$$

By an analysis (presented in Section 4.6) of the process X defined by $X_t = \|\xi_t\|$, we will see that the following result is a consequence of our general Theorems 2.1 and 2.4. The condition

$\lim \sup_{t \to \infty} \|\xi_t\| = \infty$ a.s. is ensured by, for instance, a reasonable 'irreducibility' condition, such as (A1) in [17] in the case where $\Sigma = \mathbb{Z}^d$.

Theorem 3.2. *Suppose that, for some $\rho \in (0, \infty)$ and $\beta \in (0, 1)$, (3.4) holds and that $\lim \sup_{t \to \infty} \|\xi_t\| = \infty$ a.s. Then*

(i) *if (3.5) holds for some $\gamma > 1 + \beta$, $\|\xi_t\| \to \infty$ a.s. as $t \to \infty$;*
(ii) *if (3.5) holds for some $\gamma > 2 + 2\beta$,*

$$\frac{\|\xi_t\|}{t^{1/(1+\beta)}} \longrightarrow \lambda(\rho, \beta), \quad \text{a.s.,}$$

as $t \to \infty$, for λ the constant defined at (2.8).

The *qualitative* behaviour of non-homogeneous random walks Ξ satisfying (3.4) and (3.5) of this section was studied under slightly different conditions in [17] (it was assumed there that $\Sigma = \mathbb{Z}^d$ and that the moment bound (3.5) hold for $\gamma = 2$). In particular, if we impose an additional condition on the non-radial drift field such as

$$\sup_{\mathbf{u} \in \mathbb{R}^d : \mathbf{u} \cdot \mathbf{x} = 0, \, \|\mathbf{u}\| = 1} |\mathbf{E}[(\xi_{t+1} - \xi_t) \cdot \mathbf{u} \mid \xi_t = \mathbf{x}]| = O(\|\mathbf{x}\|^{-\beta - \varepsilon}),$$

for some $\varepsilon > 0$, it was shown in [17, Theorem 2.2] that Ξ has a limiting direction. That is, there exists a random unit vector $\mathbf{u}$ such that $\xi_t / \|\xi_t\| \to \mathbf{u}$ a.s., as $t \to \infty$. Combined with our quantitative result, Theorem 3.2(ii), this implies that under the conditions of Theorem 2.2 of [17], with the stronger condition that (3.5) holds for $\gamma > 2 + 2\beta$,

$$\frac{\xi_t}{t^{1/(1+\beta)}} \longrightarrow \mathbf{u}\lambda(\rho, \beta), \quad \text{a.s., as } t \to \infty.$$

4. Proofs of theorems

4.1. Overview of the proofs

In [18, Section 3], general techniques were developed for obtaining almost-sure bounds for stochastic processes using Lyapunov functions. In [18, Section 4], those techniques were applied to the *critical* regime of the Lamperti problem (i.e., drifts of order $1/x$ at x). The results of [18, Section 3] are a useful starting point for us, enabling us to prove Theorem 2.2, but they yield bounds that are considerably less sharp than those that we ultimately require for Theorem 2.3. The sharp bounds in Theorem 2.3, and the second-order behaviour in Theorem 2.5, require a different approach. Throughout Sections 4.2–4.4, we work with the process $X_t^{1+\beta}$, for which we can establish sharp estimates (see Lemma 4.3 below). In particular, Doob's decomposition for $X_t^{1+\beta}$ will be the basis for our proofs of Theorems 2.3 and 2.5.

The proof of Theorem 2.1 (in Section 4.5) is somewhat different in flavour, and uses ideas more closely related to those of Lamperti [14]. The proof of Theorem 3.2 (in Section 4.6) demonstrates the utility of our general results in dealing with multi-dimensional processes.

4.2. Proof of Theorem 2.2

To prove Theorem 2.2, we will need the following result, contained in [18, Theorem 3.2]. Let $(Y_t)_{t \in \mathbb{Z}^+}$ be an $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$-adapted process taking values in an unbounded subset of $[0, \infty)$.

Lemma 4.1. *Let $B \in (0, \infty)$ be such that, for all $t \in \mathbb{Z}^+$,*

$$\mathbf{E}[Y_{t+1} - Y_t \mid \mathcal{F}_t] \leq B \quad \text{a.s.}$$

Then, for any $\varepsilon > 0$, a.s., for all but finitely many $t \in \mathbb{Z}^+$,

$$\sup_{0 \leq s \leq t} Y_s \leq t (\log t)^{1+\varepsilon}.$$

To prove Theorem 2.2, we will apply Lemma 4.1 to the process $Y_t = X_t^{1+\beta}$. To show that this choice of Y_t satisfies the hypothesis of Lemma 4.1, we thus need to show that the expected increment is bounded above. Lemma 4.3 below will take care of this. First we need a technical estimate. For ease of notation, we write $\Delta_t := X_{t+1} - X_t$ throughout the remainder of the paper.

Lemma 4.2. *Suppose that* (A0) *holds and that* (2.6) *holds for $\gamma \in (0, \infty)$. For any $r \in (0, \gamma)$ and $\delta \in (0, 1)$, there exists $C \in (0, \infty)$ for which, for all $t \in \mathbb{Z}^+$ and all $x \geq 0$,*

$$\mathbf{E}\left[|\Delta_t|^r \mathbf{1}\{|\Delta_t| \geq x^{1-\delta}\} \mid \mathcal{F}_t \right] \leq C(1 + x)^{-(\gamma - r)(1 - \delta)}, \quad \text{a.s.}$$

Proof. Suppose that (2.6) holds and fix $\delta \in (0, 1)$. By Markov's inequality,

$$\mathbf{P}[|\Delta_t| \geq x^{1-\delta} \mid \mathcal{F}_t] \leq x^{\gamma(\delta-1)} \mathbf{E}[|\Delta_t|^\gamma \mid \mathcal{F}_t] = O(x^{\gamma(\delta-1)}), \quad \text{a.s.}, \tag{4.1}$$

using (2.6). Now, by Hölder's inequality, for any $r \in (0, \gamma)$,

$$\mathbf{E}\left[|\Delta_t|^r \mathbf{1}\{|\Delta_t| \geq x^{1-\delta}\} \mid \mathcal{F}_t \right] \leq \mathbf{E}[|\Delta_t|^\gamma \mid \mathcal{F}_t]^{\frac{r}{\gamma}} \mathbf{P}[|\Delta_t| \geq x^{1-\delta} \mid \mathcal{F}_t]^{1 - \frac{r}{\gamma}}$$

$$= O(x^{(\gamma - r)(\delta - 1)}), \quad \text{a.s.},$$

using (2.6) and (4.1). $\quad \square$

The next result gives (in part (i)) the desired upper bound for the expected increments of $X_t^{1+\beta}$, and also provides (in part (ii)) a corresponding lower bound. Part (iii) is a technical estimate on the higher moments of the increments that we will need later in our proof of Theorem 2.3.

Lemma 4.3. *Suppose that* (A0) *holds and that, for $\beta \in [0, 1)$, $\gamma > 1 + \beta$,* (2.6) *holds.*

(i) *Suppose that, for some $A \in (0, \infty)$, $\limsup_{x \to \infty}(x^\beta \overline{\mu}_1(x)) \leq A$. Then, for any $\varepsilon > 0$, there exists $K \in (0, \infty)$ such that, for all $t \in \mathbb{Z}^+$, on $\{X_t > K\}$,*

$$\mathbf{E}[X_{t+1}^{1+\beta} - X_t^{1+\beta} \mid \mathcal{F}_t] \leq A(1 + \beta) + \varepsilon, \quad \text{a.s.} \tag{4.2}$$

(ii) *Suppose that, for some $a \in (0, \infty)$, $\liminf_{x \to \infty}(x^\beta \underline{\mu}_1(x)) \geq a$. Then, for any $\varepsilon > 0$, there exists $K \in (0, \infty)$ such that, for all $t \in \mathbb{Z}^+$, on $\{X_t > K\}$,*

$$\mathbf{E}[X_{t+1}^{1+\beta} - X_t^{1+\beta} \mid \mathcal{F}_t] \geq a(1 + \beta) - \varepsilon, \quad \text{a.s.} \tag{4.3}$$

(iii) *Let $r \in [1, \frac{\gamma}{1+\beta})$. Then there exists $C \in (0, \infty)$ such that, for all $t \in \mathbb{Z}^+$,*

$$\mathbf{E}[|X_{t+1}^{1+\beta} - X_t^{1+\beta}|^r \mid \mathcal{F}_t] \leq C X_t^{\beta r}, \quad \text{a.s.} \tag{4.4}$$

Proof. In this proof and all of the proofs that follow, C will denote a constant whose value may change from line to line. Recall that $\Delta_t = X_{t+1} - X_t$. First we prove parts (i) and (ii) of the

lemma. Let $\delta \in (0, 1)$ and define the event $E_t := \{|\Delta_t| < X_t^{1-\delta}\}$; denote the complement of E_t by E_t^c. The basic idea is as follows. We will show that the difference $X_{t+1}^{1+\beta} - X_t^{1+\beta}$ can be written as

$$(1 + \beta)X_t^\beta \left(\Delta_t - \Delta_t \mathbf{1}(E_t^c) + R_1(X_t, \Delta_t)\right) + R_2(X_t, \Delta_t),$$

where $\mathbf{E}[|\Delta_t \mathbf{1}(E_t^c)| \mid \mathcal{F}_t] = o(X_t^{-\beta})$, $\mathbf{E}[|R_1(X_t, \Delta_t)| \mid \mathcal{F}_t] = o(X_t^{-\beta})$ and $\mathbf{E}[|R_2(X_t, \Delta_t)| \mid \mathcal{F}_t] = o(1)$. Then, on taking expectations, we see that the dominant term is $(1+\beta)X_t^\beta \mathbf{E}[\Delta_t \mid \mathcal{F}_t]$, which gives the results in (i) and (ii). In the above display, the term R_1 comes from the error term in the Taylor expansion of $(X_t + \Delta_t)^{1+\beta} - X_t^{1+\beta}$ on the event E_t, while the term R_2 is $X_{t+1}^{1+\beta} - X_t^{1+\beta}$ on the event E_t^c, which is an event of small probability under our assumption of (2.6).

We now give the details of the argument sketched above. Since $X_t + \Delta_t \geq 0$, Taylor's theorem with Lagrange form for the remainder implies that

$$X_{t+1}^{1+\beta} - X_t^{1+\beta} = (X_t + \Delta_t)^{1+\beta} - X_t^{1+\beta} = (1 + \beta)X_t^\beta \Delta_t \left(1 + \eta \frac{\Delta_t}{X_t}\right)^\beta,$$

where $\eta = \eta(X_t, \Delta_t) \in [0, 1]$. Since there exists $C \in (0, \infty)$ such that $|(1 + y)^\beta - 1| \leq C|y|$ for all $y \in [-1, 1]$, we have that

$$\Delta_t \left(1 + \eta \frac{\Delta_t}{X_t}\right)^\beta \mathbf{1}(E_t) = \Delta_t \mathbf{1}(E_t) + R_1(X_t, \Delta_t),$$

where $|R_1(X_t, \Delta_t)| \leq C|\Delta_t|^2 X_t^{-1} \mathbf{1}(E_t) \leq C|\Delta_t|^{1+\beta} X_t^{-1+(1-\delta)(1-\beta)}$. Since (2.6) holds for $\gamma > 1 + \beta$, it follows that

$$\mathbf{E}[|R_1(X_t, \Delta_t)| \mid \mathcal{F}_t] \leq C X_t^{-\beta-\delta(1-\beta)} = o(X_t^{-\beta}), \tag{4.5}$$

as $X_t \to \infty$, since $\beta < 1$. Then we have that

$$X_{t+1}^{1+\beta} - X_t^{1+\beta} = (1 + \beta)X_t^\beta \Delta_t \mathbf{1}(E_t) + (1 + \beta)X_t^\beta R_1(X_t, \Delta_t) + R_2(X_t, \Delta_t), \tag{4.6}$$

where

$$R_2(X_t, \Delta_t) = (1 + \beta)X_t^\beta \Delta_t \left(1 + \eta \frac{\Delta_t}{X_t}\right)^\beta \mathbf{1}(E_t^c). \tag{4.7}$$

Since on E_t^c we have $X_t^\beta \leq |\Delta_t|^{\beta/(1-\delta)} = O(|\Delta_t|^{\beta+\delta})$ for $\beta \in [0, 1)$ and $\delta > 0$ small enough, it follows that, for small enough δ,

$$|R_2(X_t, \Delta_t)| \leq C|\Delta_t|^{1+\beta+\delta} \mathbf{1}(E_t^c), \quad \text{a.s.}, \tag{4.8}$$

for some constant $C \in (0, \infty)$ not depending on t, X_t or Δ_t. Then, taking expectations in (4.8) and using the $r = 1 + \beta + \delta$ case of Lemma 4.2, we have that, for $t \in \mathbb{Z}^+$,

$$\mathbf{E}[|R_2(X_t, \Delta_t)| \mid \mathcal{F}_t] \leq C(1 + X_t)^{-(\gamma-1-\beta-\delta)(1-\delta)} = o(1), \tag{4.9}$$

as $X_t \to \infty$, taking $\delta \in (0, 1)$ small enough so that $1 + \beta + \delta < \gamma$. Also, we have that

$$\mathbf{E}[\Delta_t \mathbf{1}(E_t) \mid \mathcal{F}_t] = \mathbf{E}[\Delta_t \mid \mathcal{F}_t] - \mathbf{E}[\Delta_t \mathbf{1}(E_t^c) \mid \mathcal{F}_t]$$

$$= \mathbf{E}[\Delta_t \mid \mathcal{F}_t] + O(X_t^{-(\gamma-1)(1-\delta)}), \tag{4.10}$$

by the $r = 1$ case of Lemma 4.2. Since $\gamma > 1 + \beta$, we can choose $\delta < \frac{\gamma - 1 - \beta}{\gamma - 1}$ so that this last error term is $o(X_t^{-\beta})$ as $X_t \to \infty$. Then, taking expectations in (4.6) and using (4.5), (4.9) and (4.10), it follows that, a.s., for all $t \in \mathbb{Z}^+$,

$$\mathbf{E}[X_{t+1}^{1+\beta} - X_t^{1+\beta} \mid \mathcal{F}_t] = (1 + \beta)X_t^\beta \mathbf{E}[\Delta_t \mid \mathcal{F}_t] + o(1).$$

Under the conditions of part (i) of the lemma, we have that $\mathbf{E}[\Delta_t \mid \mathcal{F}_t] \leq (A + o(1))X_t^{-\beta}$, as $X_t \to \infty$, a.s. Then, (4.2) follows. Similarly, (4.3) follows under the conditions of part (ii) of the lemma.

It remains to prove part (iii) of the lemma. From (4.6) and (4.8), together with the fact that $|X_t^\beta R_1(X_t, \Delta_t)| \leq C|\Delta_t|^{1+\beta}$, we have that, for $\delta > 0$,

$$|X_{t+1}^{1+\beta} - X_t^{1+\beta}|^r \leq C \left(X_t^\beta |\Delta_t| + |\Delta_t|^{1+\beta+\delta} \right)^r, \quad \text{a.s.}$$

Since $r \geq 1$, Minkowski's inequality implies that

$$\mathbf{E}[|X_{t+1}^{1+\beta} - X_t^{1+\beta}|^r \mid \mathcal{F}_t] \leq C \left(X_t^\beta \mathbf{E}[|\Delta_t|^r \mid \mathcal{F}_t]^{1/r} + \mathbf{E}[|\Delta_t|^{(1+\beta+\delta)r} \mid \mathcal{F}_t]^{1/r} \right)^r, \quad \text{a.s.}$$

Taking δ small enough so that $(1 + \beta + \delta)r \leq \gamma$, which we can do since $r < \gamma/(1 + \beta)$, we have from (2.6) that both of the expectations on the right-hand side of the last display are uniformly bounded above. Thus (4.4) follows. $\square$

We can now give the proof of Theorem 2.2.

Proof of Theorem 2.2. Lemma 4.3(i) shows that, under the conditions of Theorem 2.2, it is legitimate to apply Lemma 4.1 to the process $Y_t = X_t^{1+\beta}$. This yields the result. $\square$

4.3. Proof of Theorem 2.3

Armed with the estimates in Lemma 4.3, we can now work towards a proof of Theorem 2.3. The next result shows that, for large t, $X_t^{1+\beta}$ is, to first order, well approximated by the quantity A_t defined by $A_0 := 0$ and for $t \in \mathbb{N}$ by

$$A_t := \sum_{s=0}^{t-1} \mathbf{E}[X_{s+1}^{1+\beta} - X_s^{1+\beta} \mid \mathcal{F}_s]. \tag{4.11}$$

Under the conditions of Theorem 2.3, A_t will be seen to grow linearly with t.

Lemma 4.4. *Suppose that* (A0) *holds, that, for $\beta \in [0, 1)$,* $\limsup_{x \to \infty}(x^\beta \overline{\mu}_1(x)) < \infty$, *and that* (2.6) *holds for some $\gamma > 2 + 2\beta$. Define A_t by* (4.11). *Then, as $t \to \infty$,*

$$t^{-1} \left| X_t^{1+\beta} - A_t \right| \to 0, \quad \text{a.s.}$$

Proof. Write $Y_t := X_t^{1+\beta}$. By Doob's decomposition (see e.g. [25, p. 120]), taking $D_t := \mathbf{E}[Y_{t+1} - Y_t \mid \mathcal{F}_t]$ and writing $A_t := \sum_{s=0}^{t-1} D_s$, we have that $(M_t)_{t \in \mathbb{Z}^+}$, defined by $M_0 := Y_0$ and for $t \in \mathbb{N}$ by

$$M_t := Y_t - A_t = Y_t - \sum_{s=0}^{t-1} D_s, \tag{4.12}$$

is a martingale adapted to $(\mathcal{F}_t)_{t\in\mathbb{Z}^+}$. Taking expectations in the identity $M_{t+1}^2 - M_t^2 = (M_{t+1} - M_t)^2 + 2M_t(M_{t+1} - M_t)$, we see by the martingale property that

$$\mathbf{E}[M_{t+1}^2 - M_t^2 \mid \mathcal{F}_t] = \mathbf{E}[(M_{t+1} - M_t)^2 \mid \mathcal{F}_t]. \tag{4.13}$$

Moreover, by (4.12),

$$\begin{aligned}
\mathbf{E}[(M_{t+1} - M_t)^2 \mid \mathcal{F}_t] &= \mathbf{E}[(Y_{t+1} - Y_t - D_t)^2 \mid \mathcal{F}_t] \\
&= \mathbf{E}[(Y_{t+1} - Y_t)^2 \mid \mathcal{F}_t] - D_t^2, \tag{4.14}
\end{aligned}$$

where we have expanded the term $(Y_{t+1} - Y_t - D_t)^2$ and used the fact that D_t is $\mathcal{F}_t$-measurable. Now, by (4.13), (4.14) and the $r = 2$ case of (4.4) (which is valid since $\gamma > 2(1 + \beta)$), we have that, for all $t \in \mathbb{N}$,

$$\mathbf{E}[M_{t+1}^2 - M_t^2] = \mathbf{E}[\mathbf{E}[M_{t+1}^2 - M_t^2 \mid \mathcal{F}_t]] \leq C\mathbf{E}[X_t^{2\beta}].$$

Now, since $\beta \in [0, 1)$, Jensen's inequality implies that

$$\mathbf{E}[X_t^{2\beta}] \leq \left(\mathbf{E}[X_t^{1+\beta}]\right)^{\frac{2\beta}{1+\beta}} = O\left(t^{\frac{2\beta}{1+\beta}}\right),$$

by (4.2) and the fact that Y_0 is uniformly bounded (from the final part of (A0)). Thus M_t^2 is a non-negative submartingale with

$$\mathbf{E}[M_t^2] \leq \mathbf{E}[Y_0^2] + \sum_{s=0}^{t-1} \mathbf{E}[M_{s+1}^2 - M_s^2] = O\left(t^{\frac{1+3\beta}{1+\beta}}\right),$$

again using the fact that Y_0 is uniformly bounded. Doob's submartingale inequality (see e.g. [25, p. 137]) then implies that, for any $\varepsilon > 0$,

$$\mathbf{P}\left[\sup_{0 \leq s \leq t} M_s^2 > t^{\frac{1+3\beta}{1+\beta}+\varepsilon}\right] \leq t^{-\frac{1+3\beta}{1+\beta}-\varepsilon}\mathbf{E}[M_t^2] = O(t^{-\varepsilon}).$$

Hence the Borel–Cantelli lemma implies that, for any $\varepsilon > 0$, a.s.,

$$\sup_{0 \leq s \leq 2^m} |M_s| \leq (2^m)^{\frac{1+3\beta}{2+2\beta}+\varepsilon},$$

for all but finitely many $m \in \mathbb{Z}^+$. Since, for any $t \in \mathbb{N}$, we have $2^{m(t)} \leq t < 2^{m(t)+1}$ for some $m(t) \in \mathbb{Z}^+$, we have that, for any $\varepsilon > 0$, a.s., for all but finitely many $t \in \mathbb{Z}^+$,

$$\sup_{0 \leq s \leq t} |M_s| \leq \sup_{0 \leq s \leq 2^{m(t)+1}} |M_s| \leq (2^{m(t)+1})^{\frac{1+3\beta}{2+2\beta}+\varepsilon} \leq Ct^{\frac{1+3\beta}{2+2\beta}+\varepsilon},$$

for some $C \in (0, \infty)$ not depending on t. Since $\beta < 1$, we may take ε small enough so that $\frac{1+3\beta}{2+2\beta} + \varepsilon \leq 1 - \varepsilon$. Then, we have that $|A_t - Y_t| = O(t^{1-\varepsilon})$ as $t \to \infty$, a.s. $\qquad\square$

Remark 2. The decomposition in Lemma 4.4 is central to our proof of Theorem 2.3. Here, the behaviour of the supercritical case ($\beta < 1$) is very different to that of the critical case when $\beta = 1$ (see [18, Section 4]) where, even in the transient case, there is no decomposition available into a dominant 'drift' part (like A_t) and a smaller 'variation' part (like M_t). Thus proving (particularly lower) bounds in the critical case needs a rather different approach: see [18].

Now we can complete the proof of Theorem 2.3.

Proof of Theorem 2.3. Under the conditions of the theorem, we have from Lemma 4.3 that (4.2) and (4.3) hold. Moreover, we know from Theorem 2.1 that $X_t \to \infty$ as $t \to \infty$, a.s., so that the ε terms in (4.2) and (4.3) may be taken to be arbitrarily small for all t large enough. Hence, with A_t defined by (4.11), we have that, for any $\varepsilon > 0$, a.s.,

$$a(1 + \beta) - \varepsilon \leq t^{-1} A_t \leq A(1 + \beta) + \varepsilon,$$

for all but finitely many $t \in \mathbb{Z}^+$. Now, from Lemma 4.4, we have that $X_t^{1+\beta} = A_t + o(t)$, a.s., so that for any $\varepsilon > 0$, a.s.,

$$a(1 + \beta) - \varepsilon \leq t^{-1} X_t^{1+\beta} \leq A(1 + \beta) + \varepsilon,$$

for all but finitely many $t \in \mathbb{Z}^+$. This proves the theorem. $\square$

4.4. Proof of Theorem 2.5

The basic ingredients of the proof of Theorem 2.5 are already in place in the decomposition used in the proof of Lemma 4.4, but we need to revisit some of our earlier estimates and obtain sharper bounds under the conditions of Theorem 2.5. First, we have the following refinement of Lemma 4.3 in this case.

Lemma 4.5. *Suppose that* (A0) *holds and that, for* $\beta \in [0, 1)$, $\rho \in (0, \infty)$, (2.12) *holds. Suppose that* (2.6) *holds for* $\gamma > 2 + 2\beta$. *Then, as* $t \to \infty$,

$$\mathbf{E}[X_{t+1}^{1+\beta} - X_t^{1+\beta} \mid \mathcal{F}_t] = \rho(1 + \beta) + o(t^{-\frac{1-\beta}{2+2\beta}}), \quad \text{a.s.} \tag{4.15}$$

If, in addition, (2.13) *holds for* $\sigma^2 \in (0, \infty)$, *then, as* $t \to \infty$,

$$\mathbf{E}[(X_{t+1}^{1+\beta} - X_t^{1+\beta})^2 \mid \mathcal{F}_t] = \sigma^2(1 + \beta)^2 \lambda(\rho, \beta)^{2\beta} t^{\frac{2\beta}{1+\beta}}(1 + o(1)), \quad \text{a.s.} \tag{4.16}$$

Proof. We follow a similar argument to the proof of Lemma 4.3. We again use the notation $E_t := \{|\Delta_t| < X_t^{1-\delta}\}$ and E_t^c for the complementary event. We need to obtain better estimates for the error terms in (4.6) than we did in the proof of Lemma 4.3. For this reason the $\delta \in (0, 1)$ there will not be arbitrarily small, so we cannot use (4.8). Instead, with $R_2(X_t, \Delta_t)$ as defined at (4.7), since on E_t^c we have $X_t^{\beta} \leq |\Delta_t|^{\beta/(1-\delta)}$ where $\beta < 1$, we have that, a.s.,

$$|R_2(X_t, \Delta_t)| \leq C|\Delta_t|^{1+\frac{\beta}{1-\delta}} \mathbf{1}(E_t^c),$$

so that, from Lemma 4.2, a.s.,

$$\mathbf{E}[|R_2(X_t, \Delta_t)| \mid \mathcal{F}_t] \leq C(1 + X_t)^{(1+\frac{\beta}{1-\delta}-\gamma)(1-\delta)} = C(1 + X_t)^{1-\delta+\beta-(1-\delta)\gamma}.$$

Now, take $\delta = \frac{1-\beta}{2} - \varepsilon$ for $\varepsilon > 0$ small enough so that $\delta > 0$. It follows that, provided $\gamma > 2$, we can take $\varepsilon > 0$ small enough so that $\mathbf{E}[|R_2(X_t, \Delta_t)| \mid \mathcal{F}_t] = o(X_t^{-\frac{1-\beta}{2}})$. Next, recall (see just above (4.5)) that $|X_t^{\beta} R_1(X_t, \Delta_t)| \leq C X_t^{\beta-1}|\Delta_t|^2$. Since (2.6) holds for $\gamma > 2$, we can take expectations to obtain

$$\mathbf{E}[|X_t^{\beta} R_1(X_t, \Delta_t)| \mid \mathcal{F}_t] = O(X_t^{\beta-1}) = o(X_t^{-\frac{1-\beta}{2}}),$$

as $X_t \to \infty$, since $\beta < 1$. Moreover, from (4.10), we have that, again taking ε small enough and using the fact that $\gamma > 2$,

$$\mathbf{E}[\Delta_t \mathbf{1}(E_t) \mid \mathcal{F}_t] = \mathbf{E}[\Delta_t \mid \mathcal{F}_t] + o(X_t^{-\beta - \frac{1-\beta}{2}}) = \rho X_t^{-\beta} + o(X_t^{-\beta - \frac{1-\beta}{2}}), \quad \text{a.s.,}$$

by (2.12). With these sharper bounds, from (4.6) and the present choice of δ, we obtain

$$\mathbf{E}[X_{t+1}^{1+\beta} - X_t^{1+\beta} \mid \mathcal{F}_t] = (1 + \beta)\rho + o(X_t^{-\frac{1-\beta}{2}}), \quad \text{a.s.}$$

Under the conditions of the lemma, Theorem 2.4 applies, so $X_t \sim \lambda(\rho, \beta)t^{\frac{1}{1+\beta}}$. Thus we obtain (4.15). The argument for (4.16) is similar, starting by squaring (4.6) and then taking $\delta > 0$ small enough, so we omit the details. $\square$

Lemma 4.6. *Suppose that* (A0) *holds and that, for* $\beta \in (0, 1)$, $\rho \in (0, \infty)$, (2.12) *holds. Suppose that* (2.6) *holds for* $\gamma > 2 + 2\beta$ *and that* (2.13) *holds for* $\sigma^2 \in (0, \infty)$. *Then, with* M_t *as defined at* (4.12), *we have that, as* $t \to \infty$,

$$t^{-\frac{1+3\beta}{2+2\beta}} M_t \xrightarrow{d} Z\sigma\lambda(\rho, \beta)^\beta \sqrt{\frac{(1 + \beta)^3}{(1 + 3\beta)}},$$

where Z *is a standard normal random variable.*

Proof. We will apply a standard martingale central limit theorem. Set $M_{t,s} = t^{-\frac{1+3\beta}{2+2\beta}}(M_s - M_{s-1})$, for $1 \leq s \leq t$. For fixed t, $(M_{t,s})_s$ is a martingale difference sequence with $\mathbf{E}[M_{t,s} \mid \mathcal{F}_{s-1}] = 0$. Moreover,

$$\sum_{s=1}^{t} \mathbf{E}[M_{t,s}^2 \mid \mathcal{F}_{s-1}] = t^{-\frac{1+3\beta}{1+\beta}} \sum_{s=1}^{t} \mathbf{E}[(M_s - M_{s-1})^2 \mid \mathcal{F}_{s-1}]$$

$$= t^{-\frac{1+3\beta}{1+\beta}} \sum_{s=1}^{t} \left(\mathbf{E}[(X_s^{1+\beta} - X_{s-1}^{1+\beta})^2 \mid \mathcal{F}_{s-1}] + O(1) \right), \quad \text{a.s.,}$$

by (4.14), using the fact that $|D_{s-1}| = O(1)$ by (4.15). Now, applying (4.16), we obtain

$$\sum_{s=1}^{t} \mathbf{E}[M_{t,s}^2 \mid \mathcal{F}_{s-1}] = t^{-\frac{1+3\beta}{1+\beta}} \sigma^2 (1 + \beta)^2 \lambda(\rho, \beta)^{2\beta} (1 + o(1)) \sum_{s=1}^{t} s^{\frac{2\beta}{1+\beta}}$$

$$= \sigma^2 \frac{(1 + \beta)^3}{1 + 3\beta} \lambda(\rho, \beta)^{2\beta} + o(1), \quad \text{a.s.,} \tag{4.17}$$

where we have used the fact that $\beta > 0$ to obtain the $o(1)$ bound in the first equality. We also need to verify a form of the conditional Lindeberg condition. We claim that, for any $\varepsilon > 0$,

$$\sum_{s=1}^{t} \mathbf{E}[M_{t,s}^2 \mathbf{1}\{|M_{t,s}| > \varepsilon\} \mid \mathcal{F}_{s-1}] = o(1), \quad \text{a.s.,} \tag{4.18}$$

as $t \to \infty$. To see this, take $p \in (2, \frac{\gamma}{1+\beta})$. By the elementary inequality $|M_{t,s}|^2 \mathbf{1}\{|M_{t,s}| > \varepsilon\} \leq \varepsilon^{2-p}|M_{t,s}|^p$, we have that, for any $\varepsilon > 0$,

$$\mathbf{E}[M_{t,s}^2 \mathbf{1}\{|M_{t,s}| > \varepsilon\} \mid \mathcal{F}_{s-1}] = O(\mathbf{E}[|M_{t,s}|^p \mid \mathcal{F}_{s-1}]).$$

Then, we have that

$$\mathbf{E}[|M_{t,s}|^p \mid \mathcal{F}_{s-1}] = t^{-\frac{(1+3\beta)p}{2+2\beta}} \mathbf{E}[(M_s - M_{s-1})^p \mid \mathcal{F}_{s-1}]$$
$$= t^{-\frac{(1+3\beta)p}{2+2\beta}} \mathbf{E}[(X_s^{1+\beta} - X_{s-1}^{1+\beta} - D_{s-1})^p \mid \mathcal{F}_{s-1}],$$

where, by (4.15), $|D_{s-1}| = O(1)$, and by the $r = p$ case of (4.4),

$$\mathbf{E}[(X_s^{1+\beta} - X_{s-1}^{1+\beta})^p \mid \mathcal{F}_{s-1}] \leq C X_{s-1}^{\beta p} \leq C s^{\frac{\beta p}{1+\beta}}, \quad \text{a.s.},$$

by Theorem 2.4. Hence, by Minkowski's inequality,

$$\mathbf{E}[|M_{t,s}|^p \mid \mathcal{F}_{s-1}] \leq C t^{-\frac{(1+3\beta)p}{2+2\beta}} s^{\frac{\beta p}{1+\beta}}, \quad \text{a.s.},$$

for some $C \in (0, \infty)$. Thus, we obtain

$$\sum_{s=1}^{t} \mathbf{E}[M_{t,s}^2 \mathbf{1}\{|M_{t,s}| > \varepsilon\} \mid \mathcal{F}_{s-1}] \leq C t^{1 + \frac{\beta p}{1+\beta} - \frac{(1+3\beta)p}{2+2\beta}}, \quad \text{a.s.}$$

From this last bound, we verify (4.18), since $p > 2$. Given (4.17) and (4.18), we can apply a standard central limit theorem for martingale differences (e.g. [1, Theorem 35.12, p. 476]) to complete the proof. □

Proof of Theorem 2.5. Recall the decomposition at (4.12). Under the conditions of the theorem, Theorem 2.4 and the proof of Lemma 4.4 imply that $M_t = o(A_t)$, a.s., so that

$$X_t = (A_t + M_t)^{\frac{1}{1+\beta}} = A_t^{\frac{1}{1+\beta}} + \frac{1}{1+\beta} M_t A_t^{-\frac{\beta}{1+\beta}} (1 + o(1)).$$

Here, we have from (4.15) that, a.s., $A_t = \rho(1 + \beta)t + o(t^{\frac{1+3\beta}{2+2\beta}})$. It follows that, a.s.,

$$X_t = \lambda(\rho, \beta)t^{\frac{1}{1+\beta}} + o(t^{1/2}) + \frac{1}{1+\beta}\lambda(\rho, \beta)^{-\beta} t^{-\frac{\beta}{1+\beta}} M_t(1 + o(1)).$$

Rearranging, we obtain, a.s.,

$$\frac{X_t - \lambda(\rho, \beta)t^{\frac{1}{1+\beta}}}{t^{1/2}} = \frac{1}{1+\beta}\lambda(\rho, \beta)^{-\beta} t^{-\frac{1+3\beta}{2+2\beta}} M_t(1 + o(1)) + o(1).$$

Now, on letting $t \to \infty$, Lemma 4.6 completes the proof. □

4.5. Proof of Theorem 2.1

Our proof of Theorem 2.1 under the minimal moments conditions stated in that theorem requires some delicate analysis in a similar vein to the estimates in Section 4.3. The key is the following lemma.

Lemma 4.7. *Suppose that* (A0) *and* (A1) *hold, and that there exists* $\beta \in [0, 1)$ *such that* (2.6) *holds for* $\gamma > 1 + \beta$ *and*

$$\liminf_{x \to \infty}(x^\beta \underline{\mu}_1(x)) > 0.$$

Then there exist $\nu > 0$ *and* $M_0 \in (0, \infty)$ *such that, for all* $t \in \mathbb{Z}^+$, *on* $\{X_t > M_0\}$,

$$\mathbf{E}[(1 + X_{t+1})^{-\nu} - (1 + X_t)^{-\nu} \mid \mathcal{F}_t] \leq 0, \quad \text{a.s.}$$

Proof. For ease of notation, write $W_t := (1 + X_t)^{-\nu}$ and, as before, $\Delta_t = X_{t+1} - X_t$. Note that our assumption on $\underline{\mu}_1$ implies that, for some $c > 0$, $\mathbf{E}[\Delta_t \mid \mathcal{F}_t] \geq cX_t^{-\beta}$ a.s., for all sufficiently large X_t and all t. Let $\delta \in (0, 1)$. First, since $(1 + x)^{-\nu}$ is a decreasing function of $x \geq 0$, for any $x \geq 0$, we have

$$W_{t+1} - W_t \leq (W_{t+1} - W_t)\mathbf{1}\{|\Delta_t| < x^{1-\delta}\} + (W_{t+1} - W_t)\mathbf{1}\{\Delta_t \leq -x^{1-\delta}\}. \tag{4.19}$$

Moreover,

$$W_{t+1} - W_t = (1 + X_t)^{-\nu} \left[\left(1 + \frac{\Delta_t}{1 + X_t} \right)^{-\nu} - 1 \right].$$

Hence, by Taylor's theorem with Lagrange remainder,

$$(W_{t+1} - W_t)\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\}$$
$$= (1 + X_t)^{-\nu} \left[-\nu \frac{\Delta_t}{1 + X_t} \left(1 + \frac{\eta \Delta_t}{1 + X_t} \right)^{-1-\nu} \right] \mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\},$$

where $\eta = \eta(X_t, \Delta_t) \in [0, 1]$. Here, we have that, since $|(1 + y)^{-1-\nu} - 1| \leq C|y|$ for some $C \in (0, \infty)$ and any $y \in (-1, 1)$,

$$\Delta_t \left(1 + \frac{\eta \Delta_t}{1 + X_t} \right)^{-1-\nu} \mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\} = \Delta_t \mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\}$$
$$+ O(|\Delta_t|^2 X_t^{-1} \mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\}),$$

as $X_t \to \infty$. Hence, as $X_t \to \infty$, a.s.,

$$(W_{t+1} - W_t)\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\} = -(\nu + o(1))X_t^{-1-\nu}\Delta_t\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\} + S(X_t, \Delta_t), \tag{4.20}$$

where $|S(X_t, \Delta_t)| \leq C|\Delta_t|^2 X_t^{-2-\nu}\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\}$. We have that

$$\mathbf{E}[|S(X_t, \Delta_t)| \mid \mathcal{F}_t] \leq CX_t^{-2-\nu}X_t^{(1-\delta)(1-\beta)}\mathbf{E}[|\Delta_t|^{1+\beta} \mid \mathcal{F}_t] = O(X_t^{-1-\beta-\nu-\delta(1-\beta)}), \tag{4.21}$$

since (2.6) holds for $\gamma > 1 + \beta$. Moreover, since $\gamma > 1 + \beta$, (4.10) implies that we can take $\delta > 0$ small enough so that, as $X_t \to \infty$,

$$\mathbf{E}[\Delta_t\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\} \mid \mathcal{F}_t] = \mathbf{E}[\Delta_t \mid \mathcal{F}_t] + o(X_t^{-\beta}), \quad \text{a.s.} \tag{4.22}$$

Hence, taking expectations in (4.20), and using (4.21) and (4.22) together with the assumption that $\mathbf{E}[\Delta_t \mid \mathcal{F}_t] \geq (c + o(1))X_t^{-\beta}$, we have that, as $X_t \to \infty$,

$$\mathbf{E}\left[(W_{t+1} - W_t)\mathbf{1}\{|\Delta_t| < X_t^{1-\delta}\} \mid \mathcal{F}_t \right] \leq -(c\nu + o(1))X_t^{-1-\beta-\nu}, \quad \text{a.s.} \tag{4.23}$$

On the other hand, since $W_t \in [0, 1]$ a.s., we have that

$$\mathbf{E}\left[(W_{t+1} - W_t)\mathbf{1}\{\Delta_t \leq -X_t^{1-\delta}\} \mid \mathcal{F}_t \right] \leq \mathbf{P}[|\Delta_t| \geq X_t^{1-\delta} \mid \mathcal{F}_t] = O(X_t^{\gamma(\delta-1)}), \tag{4.24}$$

by (4.1). This last bound is $O(X_t^{-1-\beta-\delta})$, provided $\delta \leq (\gamma - 1 - \beta)/(1 + \gamma)$. From (4.19) with (4.23) and (4.24), we therefore conclude that, a.s., as $X_t \to \infty$,

$$\mathbf{E}[W_{t+1} - W_t \mid \mathcal{F}_t] \leq -(c\nu + o(1))X_t^{-1-\beta-\nu} + O(X_t^{-1-\beta-\delta}).$$

Now, taking $\nu \in (0, \delta)$ completes the proof. $\quad\square$

Proof of Theorem 2.1. To complete the proof, we use a well-known martingale idea (see e.g. [8, Theorem 2.2.2] in the countable Markov chain case). With M_0 the constant in Lemma 4.7, let $M > M_0$. For $s \in \mathbb{Z}^+$, set $T_s := \min\{t > s : X_t \le M\}$, an $(\mathcal{F}_t)_{t\in\mathbb{Z}^+}$-stopping time. We proceed to show that, for some $c > 0$, on $\{X_s > 2M\}$,

$$\mathbf{P}[T_s = \infty \mid \mathcal{F}_s] > c, \quad \text{a.s.}, \tag{4.25}$$

for all $s \in \mathbb{Z}^+$. By Lemma 4.7, we have that $(1 + X_{t\wedge T_s})^{-\nu}$ is a non-negative supermartingale adapted to $(\mathcal{F}_t)_{t\in\mathbb{Z}^+}$ for $t \ge s$. Hence, for any given s, $(1 + X_{t\wedge T_s})^{-\nu}$ converges a.s. as $t \to \infty$ to some limit, say L. Then, if $X_s > 2M$, we have

$$(1 + 2M)^{-\nu} \ge (1 + X_s)^{-\nu} \ge \mathbf{E}[L \mid \mathcal{F}_s],$$

by the supermartingale property. Moreover, on $\{T_s < \infty\}$, we have that $(1 + X_{t\wedge T_s})^{-\nu}$ converges to $(1 + X_{T_s})^{-\nu}$, so that

$$\mathbf{E}[L \mid \mathcal{F}_s] \ge \mathbf{E}[(1 + X_{T_s})^{-\nu}\mathbf{1}\{T_s < \infty\} \mid \mathcal{F}_s] \ge (1 + M)^{-\nu}\mathbf{P}[T_s < \infty \mid \mathcal{F}_s],$$

since $X_{T_s} \le M$ a.s. Thus, we obtain

$$\mathbf{P}[T_s < \infty \mid \mathcal{F}_s] \le \left(\frac{1 + 2M}{1 + M}\right)^{-\nu} < 1 - c,$$

for some $c > 0$, and so we obtain (4.25), as required. From the assumption that $\limsup_{t\to\infty} X_t = \infty$ a.s., we have that, a.s., there exist infinitely many $\mathbb{Z}^+$-valued stopping times $s_1 < s_2 < \cdots$ such that $X_{s_i} > 2M$. By standard arguments (such as Lévy's extension of the Borel–Cantelli lemmas) we can then conclude from (4.25) that $X_t > M$ for all but finitely many $t \in \mathbb{Z}^+$, a.s. This argument holds for any $M > M_0$, and so we have that $\lim_{t\to\infty} X_t = \infty$ a.s., completing the proof of transience. $\square$

4.6. Proof of Theorem 3.2

Let Ξ be as defined in Section 3.2, and take $\mathcal{F}_t = \sigma(\xi_0, \xi_1, \ldots, \xi_t)$. We will consider the process X defined by $X_t = \|\xi_t\|$. Thus we are in the final case described in Section 2.2, where X_t is a function of a Markov process. We will show that, under the conditions of Theorem 3.2, X so defined is an instance of the supercritical Lamperti problem and satisfies the conditions of Theorem 2.1 or 2.4 as appropriate. Write $\mathcal{S} = \cup_{\mathbf{x}\in\Sigma}\{\|\mathbf{x}\|\}$ for the state-space of X. The next lemma will allow us to apply our general theorems with $X_t = \|\xi_t\|$.

Lemma 4.8. *Suppose that, for some $\beta \in (0, 1)$, $\rho \in (0, \infty)$ and $\gamma > 1 + \beta$, Ξ satisfies (3.4) and (3.5). Then, X defined by $X_t = \|\xi_t\|$ is a stochastic process on $\mathcal{S}$ satisfying*

$$\sup_{t\in\mathbb{Z}^+}\sup_{\mathbf{x}\in\Sigma} \mathbf{E}[|X_{t+1} - X_t|^\gamma \mid \xi_t = \mathbf{x}] < \infty, \tag{4.26}$$

$$\lim_{x\to\infty}(x^\beta \underline{\mu}_1(x)) = \lim_{x\to\infty}(x^\beta \overline{\mu}_1(x)) = \rho. \tag{4.27}$$

Proof. For ease of notation, write $D_t = \xi_{t+1} - \xi_t$. By the triangle inequality,

$$|X_{t+1} - X_t| = |\|\xi_t + D_t\| - \|\xi_t\|| \le \|D_t\|. \tag{4.28}$$

Thus, with (4.28), (3.5) implies (4.26). Thus, it remains to prove (4.27). In this case, it suffices to show that, as $\|\mathbf{x}\| \to \infty$,

$$\|\mathbf{x}\|^\beta \mathbf{E}[X_{t+1} - X_t \mid \xi_t = \mathbf{x}] \to \rho.$$

Suppose $\xi_t = \mathbf{x} \in \Sigma$, and take $\delta \in (0, 1)$. Then, by (4.28),

$$\mathbf{E}[(X_{t+1} - X_t)\mathbf{1}\{\|D_t\| > \|\mathbf{x}\|^{1-\delta}\}] \leq \mathbf{E}[\|D_t\|\mathbf{1}\{\|D_t\| > \|\mathbf{x}\|^{1-\delta}\}] = O(\|\mathbf{x}\|^{-(\gamma-1)(1-\delta)}),$$

by an argument similar to the proof of Lemma 4.2. Since $\gamma > 1 + \beta$, we can take $\delta > 0$ small enough so that this last bound is $o(\|\mathbf{x}\|^{-\beta})$. On the other hand, applying Taylor's theorem on $\mathbb{R}^d$, we have that, when $\xi_t = \mathbf{x}$,

$$
\begin{aligned}
(X_{t+1} - X_t)\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} &= (\|\mathbf{x} + D_t\| - \|\mathbf{x}\|)\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} \\
&= \|\mathbf{x} + \eta D_t\|^{-1}(\eta\|D_t\|^2 + D_t \cdot \mathbf{x})\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\},
\end{aligned}
$$

where $\eta = \eta(\mathbf{x}, D_t) \in [0, 1]$. Hence,

$$
\begin{aligned}
\mathbf{E}&[(X_{t+1} - X_t)\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}] \\
&= (1 + o(1))\mathbf{E}[(D_t \cdot \hat{\mathbf{x}})\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}] \\
&\quad + O(\|\mathbf{x}\|^{-1})\mathbf{E}[\|D_t\|^2\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}],
\end{aligned}
\tag{4.29}
$$

as $\|\mathbf{x}\| \to \infty$. Here, we have that

$$
\begin{aligned}
\mathbf{E}[\|D_t\|^2\mathbf{1}\{\|D_t\| \leq \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}] &\leq \mathbf{E}[\|D_t\|^{1+\beta} \mid \xi_t = \mathbf{x}]\|\mathbf{x}\|^{(1-\beta)(1-\delta)} \\
&= O(\|\mathbf{x}\|^{(1-\beta)(1-\delta)}),
\end{aligned}
$$

since (3.5) holds for $\gamma > 1 + \beta$. Thus the second term on the right-hand side of (4.29) is $O(\|\mathbf{x}\|^{-\beta-\delta(1-\beta)}) = o(\|\mathbf{x}\|^{-\beta})$, since $\beta < 1$ and $\delta > 0$. Moreover, for the first term on the right-hand side of (4.29), we have that

$$\left|\mathbf{E}[(D_t \cdot \hat{\mathbf{x}})\mathbf{1}\{\|D_t\| > \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}]\right| \leq \mathbf{E}[\|D_t\|\mathbf{1}\{\|D_t\| > \|\mathbf{x}\|^{1-\delta}\} \mid \xi_t = \mathbf{x}],$$

which is $o(\|\mathbf{x}\|^{-\beta})$, as we saw above. Combining our calculations, we have shown that

$$\mathbf{E}[X_{t+1} - X_t \mid \xi_t = \mathbf{x}] = (1 + o(1))\mathbf{E}[(\xi_{t+1} - \xi_t) \cdot \hat{\mathbf{x}} \mid \xi_t = \mathbf{x}] + o(\|\mathbf{x}\|^{-\beta}).$$

Hence, from (3.4), we obtain (4.27). $\square$

Proof of Theorem 3.2. Lemma 4.8 shows that, under the conditions of Theorem 3.2, X defined by $X_t = \|\xi_t\|$ satisfies all the conditions of Theorem 2.1, which yields part (i) of the theorem. Lemma 4.8 also shows that, provided (3.5) holds for $\gamma > 2 + 2\beta$, all the conditions of Theorem 2.4 are satisfied, which yields part (ii). $\square$

Acknowledgements

References

[1] P. Billingsley, Probability and Measure, 3rd ed., Wiley, New York, 1995.

[2] K.L. Chung, A Course in Probability Theory, 2nd ed., Academic Press, San Diego, 1974.

[3] J. De Coninck, F. Dunlop, T. Huillet, Random walk weakly attracted to a wall, J. Stat. Phys. 133 (2008) 271–280.

[4] P. Coolen-Schrijner, E.A. van Doorn, Analysis of random walks using orthogonal polynomials, J. Comput. Appl. Math. 99 (1998) 387–399.

[5] E. Csáki, A. Földes, P. Révész, Transient nearest neighbour random walk on the line, J. Theoret. Probab. 22 (2009) 100–122.

[6] E. Csáki, A. Földes, P. Révész, Transient nearest neighbour random walk and Bessel process, J. Theoret. Probab. 22 (2009) 992–1009.

[7] D. DeBlassie, R. Smits, The influence of a power law drift on the exit time of Brownian motion from a half-line, Stochastic Process. Appl. 117 (2007) 629–654.

[8] G. Fayolle, V.A. Malyshev, M.V. Menshikov, Topics in the Constructive Theory of Countable Markov Chains, Cambridge University Press, 1995.

[9] G. Giacomin, Random Polymer Models, Imperial College Press, London, 2007.

[10] T.E. Harris, First passage and recurrence distributions, Trans. Amer. Math. Soc. 73 (1952) 471–486.

[11] J.L. Hodges Jr., M. Rosenblatt, Recurrence-time moments in random walks, Pacific J. Math. 3 (1953) 127–136.

[12] S. Karlin, J. McGregor, Representation of a class of stochastic processes, Proc. Natl. Acad. Sci. USA 41 (1955) 387–391.

[13] S. Karlin, J. McGregor, Random walks, Illinois J. Math. 3 (1959) 66–81.

[14] J. Lamperti, Criteria for the recurrence or transience of stochastic processes I, J. Math. Anal. Appl. 1 (1960) 314–330.

[15] J. Lamperti, A new class of probability limit theorems, J. Math. Mech. 11 (1962) 749–772.

[16] J. Lamperti, Criteria for stochastic processes II: passage-time moments, J. Math. Anal. Appl. 7 (1963) 127–145.

[17] I.M. MacPhee, M.V. Menshikov, A.R. Wade, Angular asymptotics for multi-dimensional non-homogeneous random walks with asymptotically zero drifts. Preprint arXiv:0910.1772 (2009).

[18] M.V. Menshikov, M. Vachkovskaia, A.R. Wade, Asymptotic behaviour of randomly reflecting billiards in unbounded tubular domains, J. Stat. Phys. 132 (2008) 1097–1133.

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

[20] S. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, 2nd ed., Cambridge University Press, 2009.

[21] A.G. Pakes, Some remarks on a one-dimensional skip-free process with repulsion, J. Aust. Math. Soc. 30 (1980) 107–128.

[22] M. Rosenblatt, Functions of a Markov process that are Markovian, J. Math. Mech. 8 (1959) 585–596.

[23] M. Voit, Central limit theorems for random walks on $\mathbb{N}_0$ that are associated with orthogonal polynomials, J. Multivariate Anal. 34 (1990) 290–322.

[24] M. Voit, Strong laws of large numbers for random walks associated with a class of one-dimensional convolution structures, Monatsh. Math. 113 (1992) 59–74.

[25] D. Williams, Probability with Martingales, Cambridge University Press, 1991.